CRANFIELD UNIVERSITY

CARLOS VRANCKEN

DEVELOPMENT OF A METHOD TO CLASSIFY AND ANALYSE THE COMPOSITION OF MIXED WASTE MATERIALS IN REAL-TIME

SCHOOL OF WATER, ENERGY AND ENVIRONMENT

PhD

Academic Year: 2015 - 2022

Supervisor: Stuart Wagland
Associate Supervisor: Phil Longhurst
July 2022

CRANFIELD UNIVERSITY

School of Water, Energy and Environment (SWEE)

PhD

Academic Year 2015 - 2022

CARLOS VRANCKEN

Development of a method to classify and analyse the composition of mixed waste materials in real-time

Supervisor: Stuart Wagland
Associate Supervisor: Phil Longhurst
July 2022

© Cranfield University 2022. All rights reserved. No part of this publication may be reproduced without the written permission of the copyright owner.

ABSTRACT

There is a need for innovative technologies to classify and monitor the composition of solid waste in real-time. This research project has highlighted which information is required to improve current process designs. It also identified visible spectrum cameras as the solution that can better inform waste composition and quality without requiring complementing technologies. The experiments applied deep learning methods to classify the materials based on their images, and a method to analyse the composition of mixed waste was developed.

There is a high variability in the appearance of waste materials in the context of a material recovery facility. An image capture setup using multiple cameras and light sources was implemented and tested to acquire a representative set of images. The hardware captures images from different angles, with enhanced shadow details, and different exposure levels. Image processing software further augmented the data by rotating and changing the images resolutions. The images were converted to greyscale to increase the method robustness without affecting classification performance.

Deep convolutional neural networks were trained on the augmented datasets. The trained networks obtained state-of-the-art performance when tested and validated for the task of waste material classification. Based on this, a composition analysis methodology was developed and tested with mixed material samples. The methodology provides results as accurate as current manual solutions, but it can analyse a waste stream on a conveyor belt in real-time. The findings and observations from the experimental results contribute to knowledge in three main areas: data capture, data processing, and deep learning.

This thesis presents the progressive development of the methodology and discusses different applications for waste management. The composition analysis can provide real-time waste data to improve the overall efficiency of the waste treatment industry. This information can be also used by stakeholders for better decision-making in the future.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisors, Dr. Stuart Wagland and Prof. Phil Longhurst. They understood the importance of this research and pushed it through (some might consider a fire to be a bad omen, but not them!), and they even trusted me to deliver it. Stuart and Phil provided me with guidance and motivation through each stage of the process. They allowed me to participate in the decisions regarding the path of my research, inspired my interest in the development of innovative methods and solutions for this area, and encouraged me to learn and apply new technical skills.

Thanks to all my family and friends, following my progress and encouraging me from a distance. To Chiara and Blessing, for being my buddies and sharing coffee or a meal whilst talking about research or about our lives: you made my days on-campus much, much easier.

To my wife Gaby, who always find the right words and never doubt me (even and specially when I was not confident myself): thank you for your love, support and extra efforts (super tag-team parenting!).

To Danilo and Mirko, who were both born during this and cannot yet read this: watching you grow and enjoying your smiles, hugs, laughs and games were the best part of my PhD (and who cares about sleep, right!?)

Dedicated to my father, who always believed in me. And to Maia, who inspired me to get started on this path and is the first thought in my mind as I take this and every step of my journey.

TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS iii
LIST OF FIGURES ix
LIST OF TABLES xii
LIST OF EQUATIONS xvii
LIST OF ABBREVIATIONS xix
1 Introduction 23
1.1 Waste management systems 23
1.2 Monitoring waste quality 25
1.3 Image classification 25
1.3.1 Texture and material recognition 26
1.3.2 Waste analysis 28
1.4 Aim and objectives 29
1.4.1 Overall aim and research question 29
1.4.2 Objectives 30
1.5 Thesis structure 31
1.6 References 33
2 Critical review of real-time methods for solid waste characterisation: informing material recovery and fuel production 41
2.1 Introduction 41
2.1.1 Manual sorting 44
2.1.2 Biological processes 45
2.1.3 Mechanical processes 45
2.2 Advanced sensors 48
2.2.1 Fluorescence 48
2.2.2 Hydrophobicity 48
2.2.3 Laser induced breakdown spectroscopy 49
2.2.4 Microwaves 49
2.2.5 Resistivity and capacity 50
2.2.6 Stiffness 50
2.2.7 Ultrasound imaging 50
2.2.8 Optical imaging using visible wavelength sensors 51
2.2.9 Optical imaging using infrared wavelength sensors 52
2.2.10 X-ray imaging 54
2.3 Process analytical technology 55
2.3.1 Critical quality attributes 56
2.3.2 Process monitoring tools 81
2.4 Conclusions and perspectives 86
2.5 References 87

3 Deep learning in material recovery: development of method to create training database 107
3.1 Introduction 107
3.2 Materials and methods 112
3.2.1 Material samples 112
3.2.2 Image capture 113
3.2.3 Image processing 116
3.2.4 Deep neural network 118
3.2.5 Accuracy testing 119
3.2.6 Data analysis 120
3.2.7 Experimental design 121
3.3 Results and discussion 122
3.3.1 Experiment 1: Software rotation 122
3.3.2 Experiment 2: Patch overlap 124
3.3.3 Experiment 3: Number of lights 125
3.3.4 Experiment 4: Number of cameras 126
3.3.5 Experiment 5: Combined software techniques 128
3.3.6 Experiment 6: Combined hardware techniques 128
3.3.7 Experiment 7: Combined all (software and hardware) techniques 129
3.3.8 Time requirements 130
3.3.9 Accuracy vs configuration 131
3.4 Conclusions 132
3.5 References 134
4 Deep learning for waste management: effects of colour changes and identification of paper grades 141
4.1 Introduction 141
4.2 Materials and methods 144
4.2.1 Material samples 144
4.2.2 Image capture 144
4.2.3 Image processing 145
4.2.4 Deep neural network 146
4.2.5 Accuracy testing 146
4.2.6 Data analysis 149
4.2.7 Experimental design 149
4.3 Results 152
4.3.1 Experiment 1: Effects of the number of training classes 152
4.3.2 Experiment 2: Effects of colour content 155
4.3.3 Experiment 3: Effects of changes in the hue channel 159
4.3.5 Experiment 4: Effects of changes in the saturation channel 161
4.3.7 Experiment 5: Effects of changes in the brightness channel 162
4.3.8 Experiment 6: Effects of the number of training patches 162
4.3.9 Experiment 7: Paper grade identification 166

4.4 Discussion 168
4.5 Conclusions 172
4.6 References 174
5 Multi-scale training of deep networks for the classification of waste material 181
5.1 Introduction 181
5.2 Materials and methods 185
5.2.1 Material samples 185
5.2.2 Image capture 186
5.2.3 Image processing 187
5.2.4 Deep learning 187
5.2.5 Accuracy testing 188
5.2.6 Interpretation of the deep models 189
5.2.7 Experimental design 190
5.3 Results 192
5.4 Discussion 200
5.5 Conclusions 204
5.6 References 205
6 Determination of mixed waste composition using deep neural networks 213
6.1 Introduction 213
6.2 Materials and methods 217
6.2.1 Material samples 217
6.2.2 Image capture 220
6.2.3 Image processing 221
6.2.4 Deep learning 222
6.2.5 Classification performance and composition analysis 223
6.2.6 Experimental design 225
6.3 Results 227
6.3.1 Experiment 1 227
6.3.2 Experiment 2 228
6.3.3 Experiment 3 231
6.4 Discussion 238
6.5 Conclusions 243
6.6 References 244
7 Discussion 253
7.1 Contribution to knowledge 253
7.1.1 Data capture 254
7.1.2 Data processing 257
7.1.3 Deep learning methods 260
7.2 Applications in the wider context 263
7.3 Limitations 264
7.4 Further work 265

7.5 References 267
8 Conclusions 271
8.1 References 277
APPENDICES 281
Appendix A 281

LIST OF FIGURES

Figure 1-1. General methodology and research areas covered in each chapter 32

Figure 2-1. Flow sheet diagram of a typical MRF using manual and MBT processes to produce four outputs: biogas and SRF for energy conversion, materials for recycling, and rejects going into landfill. Adapted from DEFRA (2013). 44

Figure 2-2. Current waste management strategy (left branch), receiving waste of variable content cannot control the product. The implementation of PAT (right branch) could allow the waste processing plants to produce a consistent product. 56

Figure 3-1- Schematics showing the image capture hardware configuration which includes 5 cameras (C1 – C5) and 5 lights (L1 – L5). The lights (not shown) are located next to the cameras, also pointing towards the target area... 114

Figure 3-2. Steps taken to create the training and testing databases..... 115

Figure 3-3 - Effects of using software rotations to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions 123

Figure 3-4 - Effects of using different levels of overlapping between patches to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions 124

Figure 3-5 - Effects of using multiple illuminations to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions 126

Figure 3-6 - Effects of using multiple cameras to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions 127

Figure 3-7 - Comparison of best accuracies achieved in each experiment. The vertical bars denote standard deviation from 3 experiment repetitions.... 132

Figure 4-1 - Diagram describing the process followed for Experiments 3 to 5, to study the effects of colour changes 150

Figure 4-2. Comparison of the effects of changing the hue channel when testing coloured or converted-to-greyscale images. Standard deviations (not shown due to scale) ranged from 0.1% to 2.5% and from 0.1% to 0.7% for colour and greyscale images, respectively. 160

Figure 4-3. Comparison of the effects of changing the saturation channel when testing coloured or converted-to-greyscale images. Standard deviations (not shown due to scale) ranged from 1.1% to 1.6%, and from 0.2% to 0.3%, for colour and greyscale images, respectively. 161

Figure 4-4. Comparison of the effects of changing the brightness channel when testing coloured or converted-to-greyscale images. Standard deviations (not shown due to scale) ranged from 0.2% to 0.7% and from 0.1% to 1.2% for colour and greyscale images, respectively. 162

Figure 4-5. Effects of different number of training patches, on a log-10 scale. Vertical bars indicate standard deviation (note: some vertical bars are not visible due to scale). 163

Figure 4-6. Image of a printed white paper sample incorrectly classified as newspaper. Patches classified as white paper are green and newspaper patches are orange. 167

Figure 5-1 – Average accuracy per class for images of different resolutions classified with 2 different networks, one trained with 55 and the other with 289 PPI images. Vertical bars show the standard error for each tested resolution. 193

Figure 5-2 - Average accuracy per class for test images of different resolutions in the range from 50 to 289 PPI, classified by networks trained with images of multiple resolutions, with and without images of 200 PPI resolution. Vertical bars show the standard errors. Note: the Y-axis range is reduced to the range from 80 to 100%. 195

Figure 5-3 - Correct classification examples, using the trained network from Experiment 8, with test images of the same material sample using different resolutions. (a) red brick 55 PPI, (b) red brick 289 PPI, (c) metal pipes 55 PPI, (d) metal pipes 289 PPI, (e) newspaper 55 PPI, (f) newspaper 289 PPI. Left image is the input, right image is the heatmap generated towards its correct classification. Red colour indicates positive evidence, blue colour negative evidence, no colour is neutral evidence. The dotted-lines square regions on the 55 PPI input image denote the areas analysed by the 289 PPI input patch. 196

Figure 5-4 - Classification performance for 4 different material classes using images of different resolutions. Vertical bars show the standard error of the mean. Note: the Y-axis range is reduced to 50-100%. 198

Figure 5-5 - Wrong classification examples, using the trained network from Experiment 8 for images of different resolutions. (a) glossy magazine 55 PPI, (b) glossy magazine 289 PPI, (c) plastic carrier bag 55 PPI, (d) plastic carrier bag 289 PPI. At the top is the input image, middle shows the heatmap generated towards the correct class, and bottom image is the relevance heatmap for the wrongly predicted class. Red colour indicates positive evidence, blue colour negative evidence, no colour is neutral evidence. 199

Figure 5-6 - Average standard error of the average accuracy per class for the range of spatial resolution in the test images. Vertical bars show the full (min, max) range of standard errors for the 20 subclasses considered. 200

Figure 6-1. Results from experiments 1 and 2 to recognize the percentage of wood in mixture 1, using classification method A 230

Figure 6-2. Results from experiment 2 to recognize cardboard material classes in mixture 1, using classification method A..... 231

Figure 6-3 – True area (%) composition and the estimates from the analysis of Mixture 1, using the network from Experiment 3 and method A. The vertical bars indicate the standard deviation. 233

Figure 6-4. True weight % composition and the estimates from the analysis of mixture 1, using the network from experiment 3 and method A. The vertical bars indicate the standard deviation. (*) Note that the wood and brick compositions are shown on a different scale indicated by the secondary axis due to their higher values. 234

Figure 6-5 – Visualisation of partial results from experiments 3 to recognize the percentage of materials in mixture 1, using classification method A..... 235

Figure 6-6 – Visualisation of partial results from experiments 2 and 3 to recognize the black plastic food trays and white polystyrene foam in an image of all fractions from mixture 2, using classification method A 236

Figure 6-7 – Visualisation of partial results from experiments 2 and 3 to recognize the black plastic food trays and white polystyrene foam in an image of all fractions from mixture 3, using classification method A 237

Figure 7-1. Contributions to knowledge..... 254

Figure A-1 - Images of paper samples used for training, as captured with camera C5 and illumination L5 286

Figure A-2 - Images of cardboard samples used for training, as captured with camera C5 and illumination L5 290

Figure A-3 - Images of paper samples used for testing, as captured with camera C5 and illumination L5 292

Figure A-4 - Images of cardboard samples used for training, as captured with camera C5 and illumination L5 293

LIST OF TABLES

Table 2-1. Mechanical and biological processes, their effects on solid waste streams and operational parameters affecting their efficiency. 46

Table 2-2. Performance rating scale using the different commonly published performance parameters (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006), where R^2 is the coefficient of determination, RPD is the ratio of the standard deviation of the original data to standard error of prediction, RER is the ratio of the range of the original data to standard error of prediction, % error is the percent error between the predicted and the original data, and % rate is the success rate (recognition, classification, separation, etc.) of the method. 57

Table 2-3. Technologies and methods applied to problematic objects recognition. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 59

Table 2-4. Sensor technologies and methods to estimate particle size distribution. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 60

Table 2-5. Sensor technologies for the determination of calorific value. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 62

Table 2-6. Sensor technologies for the determination of ash contents. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 66

Table 2-7. Sensor technologies for the determination of moisture contents. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 69

Table 2-8. Sensor technologies for composition analysis of mixed materials. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 73

Table 2-9. Sensor technologies for the determination of biogenic carbon contents and determination of BMP. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent). 77

Table 2-10. Sensor technologies for the identification of specific contaminants in mixed materials. Performance rating (Malley et al., 2002, 2004;

Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), *** (good), **** (very good), ***** (excellent)..... 79

Table 2-11. Best performing technologies for each indicator when applied to waste materials..... 82

Table 2-12. Critical quality indicators and the available monitoring tools for the different waste process units. 84

Table 3-1 – Raw images captured from a sample of kitchen towel paper (tissue_whitetowel_1.jpg, Figure A-1) with a combination of 5 cameras (c1 to c5) and 6 light sources (i1 to i6). Each image shows an area of 43 by 29 cm. 116

Table 3-2 – Patches extracted from the raw images of kitchen towel paper (tissue_whitetowel_1.jpg, Figure A-1) . The size of each patch is 256 by 256 pixels, equivalent to an area of 2.25 by 2.25 cm. The patches in this table show the same region from the material captured with combination of 5 cameras (c1 to c5) and 6 light sources (i1 to i6) (Table 3-1)..... 117

Table 3-3 - Confusion matrix structure used to show the performance of the models from each of the experiments..... 120

Table 3-4 - Summary of experimental setups and the different configurations tested..... 122

Table 3-5 - Confusion matrix with the average prediction results using 12 rotations (30 degrees). Correct predictions in bold..... 123

Table 3-6 - Confusion matrix with the average prediction results using 25% overlapping. Correct predictions in bold..... 125

Table 3-7 - Confusion matrix with the average prediction results using 6 illuminations Correct predictions in bold. 126

Table 3-8 - Confusion matrix with the average prediction results using 5 cameras Correct predictions in bold. 127

Table 3-9 - Confusion matrix with the average prediction results using a combination of software rotation and patch overlapping. Correct predictions in bold. 128

Table 3-10 - Confusion matrix with the average prediction results using combinations of 5 cameras and 6 illuminations. Correct predictions in bold. 129

Table 3-11 - Confusion matrix with the average prediction results using combinations of 5 cameras and 6 illuminations. Correct predictions in bold. 130

Table 3-12. Best configuration and accuracy obtained in each experiment ... 131

Table 4-1 - Example confusion matrix to illustrate the calculations of the different performance indicators 148

Table 4-2 – Normalised version of the example confusion matrix to enable comparison between different tests. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. 148

Table 4-3. Summary of experimental design and specific goals..... 151

Table 4-4 - Confusion matrix, training with 2 classes, testing on 2 classes. Correct predictions in bold..... 152

Table 4-5 - Confusion matrix, model trained with 2 classes, tested on 2 classes, and evaluated for each subclass. Correct predictions in bold..... 153

Table 4-6 - Confusion matrix, model trained with 10 subclasses, tested on 2 classes. Correct predictions in bold. 153

Table 4-7 - Confusion matrix, model trained with 10 subclasses, tested on 2 classes, and evaluated for each subclass. Correct predictions in bold. ... 154

Table 4-8 - Comparison of accuracies of networks trained to differentiate either 2 or 10 class types, to classify the materials into 2 types, either paper or cardboard. Bold indicates best result..... 155

Table 4-9 - Number of patches per class in the training dataset 156

Table 4-10 - Confusion matrix when using the *Coloured* version of the patches. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results showing the numbers of patches can be seen in Table A-1. 157

Table 4-11 - Confusion matrix when using the greyscale version of the patches. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results showing the numbers of patches can be seen in Table A-2. 158

Table 4-12. Comparison of accuracies per class, obtained when training the network using 10 classes of materials, either with colour or greyscale versions of the images, to classify them into the same 10 classes..... 159

Table 4-13. Confusion matrix presenting the tests results after training the network with 150,000 patches per class..... 164

Table 4-14 - Confusion matrix when using a model trained with a balanced training dataset, with 5,000 patches per class, to classify 10 classes. The

percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results showing the numbers of patches can be seen in Table A-3. 165

Table 4-15. Accuracies obtained when training the network using balanced and unbalanced databases of 10 classes of materials, with greyscale versions of the images, to classify them into the same 10 classes. 166

Table 4-16. Confusion matrix presenting the validation results, using the network trained with 5,000 patches per class to identify 3 classes of material samples. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results in parenthesis indicate numbers of patches. 167

Table 5-1. List of classes used for training and testing purposes. 186

Table 5-2. Summary of experimental design parameters 191

Table 6-1 - Area composition of each of the material groups used for training and testing – the material classes indicate whether they belong to the light (L) or heavy (H) fractions. The shredded fractions (not shown) share the composition of the light fractions. 218

Table 6-2 – Relative area composition of each of the material groups used for training and testing expressed as a percentage to the total area of each mixture – the material classes indicate whether they belong to the light (L) or heavy (H) fractions. The shredded fractions (not shown) share the composition of the light fractions 219

Table 6-3 - Area density and relative weight composition of each material class expressed as a percentage to the total weight of each mixture. The material classes indicate whether they belong to the light (L) or heavy (H) fractions. The shredded fractions (not shown) share the composition of the light fractions 220

Table 6-4. Overall scaled performance indicators obtained with the trained network from Experiment 1, using classification methods A and B, considering the top-1 and all probabilities, respectively, to determine the content of each test patch..... 228

Table 6-5. Overall scaled performance indicators obtained with a network trained with multiple exposure levels in Experiment 2, using classification methods A and B, considering the top-1 and all probabilities, respectively, to determine the content of each test patch..... 229

Table 6-6. Overall scaled performance indicators obtained with the trained network from Experiment 3, using classification methods A and B,

considering the top-1 and all probabilities, respectively, to determine the content of each test patch..... 232

Table A-1 - Confusion matrix when using the Coloured version of the patches. Bold indicates correct predictions. Colour shades indicate the proportion of the number of a predicted class relative to the number of samples for that row..... 294

Table A-2 - Confusion matrix when using the greyscale version of the patches. Bold indicates correct predictions. Colour shades indicate the proportion of the number of a predicted class relative to the number of samples for that row..... 295

Table A-3 - Confusion matrix when using a model trained with a balanced training dataset, with 5,000 patches per class, to classify 10 classes. Bold indicates correct predictions. Colour shades indicate the proportion of the number of a predicted class relative to the number of samples for that row 296

LIST OF EQUATIONS

(3-1)..... 120
(3-2)..... 120
(3-3)..... 120
(4-1)..... 145
(4-2)..... 147
(4-3)..... 147
(4-4)..... 147
(4-5)..... 148
(4-6)..... 149
(4-7)..... 149
(5-1)..... 189
(5-2)..... 189
(5-3)..... 189
(6-1)..... 224
(6-2)..... 224
(6-3)..... 225
(6-4)..... 225
(6-5)..... 225

LIST OF ABBREVIATIONS

ANOVA	Analysis Of Variance
BMP	Biochemical Methane Potential
C&D	Construction and Demolition
C&I	Commercial and Industrial
CCD	Charge-Coupled Device
CEN/TC	European Committee for Standardization Technical Committee
CEPI	Confederation of European Paper Industries
CMOS	Complementary Metal-Oxide-Semiconductor
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CV	Calorific Value
DECC	Department of Energy and Climate Change
DEFRA	Department for Environment, Food and Rural Affairs
EPR	Extended Producer Responsibility
EU	European Union
FDA	Food and Drug Organization
GB	Gigabyte
GCV	Gross Calorific Value
GPU	Graphical Processing Unit
HSI	Hyperspectral Imaging
HSI	Hue, Saturation and Intensity
HSV	Hue, Saturation and Value
ISO	International Organisation for Standardisation
ISWA	International Solid Waste Association
JPEG	Joint Photographic Experts Group
LBP	Local Binary Patterns
LIBS	Laser-Induced Breakdown Spectroscopy
LIR	Long Wavelength Infrared
LRP	Layer-Wise Relevance Propagation
LWIR	Long Wavelength Infrared
MAE	Mean Absolute Error
MB	Megabyte

MBT	Mechanical And Biological Treatment
ME	Mean Error
MINC	Materials In Context Database
MIR	Medium Wavelength Infrared
MP	Megapixel
MRF	Material Recovery Facility
MSW	Municipal Solid Waste
NCV	Net Calorific Value
NIR	Near Infrared
NWIR	Medium Wavelength Infrared
PAT	Process Analytical Technology
PET	Polyethylene Terephthalate
PLS	Partial Least Squares
PLS-DA	Partial Least Squares Discriminant Analysis
PPI	Pixels Per Inch
PSD	Particle Size Distribution
PVC	Polyvinyl Chloride
RDF	Refuse Derived Fuel
RER	Range Error Ratio
RGB	Red, Green and Blue
RMSE	Root Mean Square Error
RPD	Ratio of Predictive Deviation
RWM	Recycling and Waste Management
SD	Standard Deviation
SE	Standard Error
SRF	Solid Refuse Fuel
SWIR	Short Wavelength Infrared
VIS	Visible Wavelength
WDF	Waste Derived Fuel
WEEE	Waste Electrical and Electronic Equipment
WISH	Waste Industry Safety and Health
WRAP	Waste and Resources Action Programme
XRF	Fluorescence X-Ray
XRT	Transmission X-Ray

CHAPTER 1

INTRODUCTION

1 Introduction

This section provides background information, summarises existing research, formulates the research question and objectives, and positions the methodological approach.

1.1 Waste management systems

Sustainable development aims to continuously improve the quality of life through efficient management of resources to reduce environmental issues (EU Council, 2001). In the EU, the landfill directive (EU Council, 1999) was implemented to reduce the reliance on landfills and minimise their environmental impact. Certain types of hazardous waste were banned and pre-processing of waste materials was required before they could be disposed of. The Waste Framework Directive (EU Council, 2008) defined the hierarchy of waste management options and their order of priority, including prevention as the preferred action, followed by reuse, recycling, recovery, and safe disposal as the least preferable option.

Prevention and reuse of materials are incentivised through policies such as extended producer responsibility (EPR), placing the responsibility on manufacturers for the materials used in their products in their entire life cycle, including their monitoring, treatment, and disposal (DEFRA, 2019). Recent legislation and regulations (EU Commission, 2018; EU Parliament, 2018, 2017) were set out to increase the levels of waste material recycling and recovery (DEFRA, 2018).

The Renewable Energy Directive (EU Council, 2009) and the UK energy policy (DECC, 2009) recognise and promote the use of energy from waste as an alternative for sustainable waste management. Waste derived fuels such as refuse derived fuel (RDF) and solid recovered fuel (SRF) can be produced from residual waste, which cannot be reused or recycled but still contains energy that could displace fossil-derived fuels (Velis et al., 2010). The RDF market has grown rapidly as a result of the implementation of the landfill directive and the opening of the export market (UK Environment Agency, 2015). There are no standard quality requirements for RDF, but each end-user defines its own specifications

based on chlorine, ash contents and calorific value. The production of SRF, however, is regulated by EU and UK legislation (EU CEN, 2012; WRAP, 2012) to prevent adverse effects on the environment and on industrial processes.

Material recovery facilities (MRF) are the most adopted strategy to manage waste (Velis et al., 2010). These facilities use mechanical and biological treatment processes to treat solid materials, reducing the volume of waste going into landfill by extracting target materials which can be reused, recycled, or converted into fuel. The input waste materials and the targeted outputs determine the design of an MRF, including which processes to use. The correct implementation of these treatment processes along with constant quality monitoring are important to reduce the levels of contaminants in the different outputs (Lorber et al., 2012; Sarc et al., 2014; Sarc and Lorber, 2013). In the UK, the operation of MRF is regulated and requires constant monitoring through material sampling, testing and recording of the quality of input and output materials (WRAP, 2014).

Whilst the type of incoming waste and the purpose of an MRF determine its configuration, the methods used to sort the materials are affected by the economy and available markets (David McKinnon et al., 2017). In developed countries with stronger economies and quality demanding markets, advanced sorting technologies are used to complement, supplement and partially or totally replace manual sorting in MRF. This promotes technology innovation to advance material identification and separation methods (ISWA, 2016; Sedlak, 2017). On the other hand, the configurations of MRF in developing countries with emerging economics are built with low capital investments and maintenance costs, relying on manual sorting techniques or affordable solutions (Guerrero et al., 2013).

Manual sorting stations in MRF are located next to conveyor belts moving at speeds up to 0.5 m/s (Tchobanoglous and Kreith, 2002). The belts at these stations are up to 107 cm wide when the workers are located on only one side. Multiple sorters typically work together, and each one is responsible for sorting out a specific material class. The manual recovery rate of materials ranges from 60% to 95%, and it is affected by the conveyor belt speed, the type and presentation of materials, worker experience, and fatigue levels. These workers

are exposed to health and safety risks, due to mechanical and operational issues of the moving belts, as well as hazards created by the materials, such as sharp metals or broken glass, and errors caused by human fatigue.

1.2 Monitoring waste quality

Currently, MRF are built based on fixed designs, with flows and processes configured to treat the specified types and flow of input waste and the target outputs. This fixed nature of the MRF means that the quality of the targeted outputs is determined by the quality of the input waste and the performance of the individual treatment processes. However, the composition of input solid waste is highly variable which represents a challenge in waste to energy MRF to meet the required output quality targets with consistency (Cimpan et al., 2015; Flamme and Geiping, 2012).

Other industries have improved the design of their different material and product processes through process analytical technology (PAT) (FDA, 2004). This mechanism requires an understanding of the processes involved and the attributes of the materials or products that have a significant impact on the quality of the output. Real-time data is acquired and used to control process parameters to ensure the intermediate and final outputs. In the context of an MRF, a real-time composition analysis could provide information about the quality of the input and output streams, for example to detect hazardous objects or for quality control purposes. However, it is not clear if or how the composition information could improve the efficiency of an MRF, nor whether other waste information would be required.

1.3 Image classification

The traditional approach to automatic image classification requires custom processing and filters specifically designed to extract useful information or features, using computer vision algorithms such as threshold segmentation and edge detection. This is usually a process of trial and error, requiring expertise and careful analysis of many parameters to adjust each hand-crafted algorithm (Nixon and Aguado, 2012). Each target object class is defined by a number of these

extracted features or descriptors from known example images. To classify a test image, the same hand-crafted filters are applied, and the extracted features compared to the definitions of the different target classes to determine the best match.

Traditional approaches to texture recognition include grey-level co-occurrence matrices (Haralick et al., 1973) and local binary patterns (Ojala et al., 1996). A common implementation of these methods divides high resolution images into patches and use the filters to extract features from each patch. These patches can have different sizes or scales (Cavalin et al., 2013; Cimpoi et al., 2016) and can overlap each other or not (Cavalin et al., 2013; Hafemann et al., 2014). The classification of the entire image is then based on an analysis of the combination of features extracted from each patch.

The use of deep neural networks trained with supervised learning for the task of automatic image recognition has been the focus of many recent studies since the high classification rates obtained by Krizhevsky et al. (2012). A deep neural network is an artificial neural network containing multiple layers of interconnected neurons that engage to solve a task (LeCun et al., 2015). Deep learning is the set of techniques developed to train deep neural networks for a specific task, such as image classification. One common technique requires supervised learning, which indicates the training and testing of the neural networks are achieved with labelled datasets containing pairs of image and descriptive information. A typical deep learning process requires a large training dataset containing many representative images for each target class. Data augmentation can be used to increase the size of a training dataset through image processing to reduce generalisation errors. A trained network is said to generalize when it can perform well the classification task on test data different to the training dataset.

1.3.1 Texture and material recognition

Deep learning has been applied for the task of texture classification using convolutional neural networks (CNN) (Sharan et al., 2013), which are a type of neural network adapted for the specific task of image classification (LeCun et al.,

2015). The classification of textures with CNN are comparable to those obtained with traditional methods using state-of-the-art but hand-crafted descriptors such as LBP, with CNN obtaining better results than LBP when classifying textures with high-intra class variability (Bello-Cerezo et al., 2017). Object or scenes classification rates obtained with CNN can be higher than human perception levels (Russakovsky et al., 2015).

One large study created a database of images to train a deep network for the task of material recognition using contextual clues included the pictures to assist the classification (Bell and Bala, 2015). Another study used deep convolutional neural networks to analyse materials in realistic images containing multiple materials (Cimpoi et al., 2016), using traditional object detection methods to segment the areas which are then classified pixel by pixel.

The work by Sharan et al. (2013) combined a series of hand-crafted algorithms to extract features and classify surfaces into one of 10 common material categories. The extracted features were based on shape, colour, texture, and context information beyond the material in question. This recognises the issues with ambiguity in material recognition, where not a single property is enough to accurately distinguish one material category from another. Their proposed solution used a combination of features to achieve a recognition rate 57.1%. Colour content was highlighted as an important feature, as its removal led to a relative drop in performance of 8.6% (Sharan et al., 2013).

The work by Bell et al. (2015) used a combination of multiple image examples and context information to deal with ambiguity. Three million images of materials were manually labelled into 23 categories, ensuring that different variations of the same materials were capture and represented when training the deep learning models. Furthermore, images were resampled to different scales to ensure that contextual information is considered in the classification task, so that an item, whether red or blue, had more chances to being a cup when the image includes a coffee table.

1.3.2 Waste analysis

In the field of waste management, the classification of images presents a unique challenge compared to object and scenes. For example, multiple waste materials can be visible at the same time, they can be overlapping and be placed at any position in the field of view, and they can be presented at any rotation angle. Several recent studies have used deep learning to classify images of waste objects into three or four different categories (Adedeji and Wang, 2019; Aral et al., 2018; Bircanoğlu et al., 2018; Ozkaya and Seyfi, 2019; Ruiz et al., 2019; Sakr et al., 2016; Yang and Thung, 2016). A related study by Awe et al. (2017) overlapped images of single objects to simulate images of mixed waste but limited the classification to only three categories. However, all these studies only assume and identify one waste object per image, and rely on shape and edge information to classify the item types.

The lack of context of waste materials in a MRF is a key operational difference to classifying those same items in the wild or in their natural placement. Waste materials flow on top of conveyor belts in MRF, and images of these items will have a random background, with either the conveyor belt in the case of a sparse stream, or other types of materials in the case of a very dense stream. In the case of an output stream, after the materials have already been sorted in the MRF, there is a higher probability to see the same material type. For example, most of the items in a stream that positively sorts clear PET materials will be clear PET bottles and trays. However, a conveyor belt upstream, before separation processes, could contain a mix of fibre, textile, plastics, glass, and metal materials all in the same picture. Therefore, any developed technology should not depend on contextual information to inform the classification decision.

A methodology to manually determine the composition of mixed waste, based on visual inspection of images, was previously developed by Wagland et al. (2012). This same technique was applied to determine the energy contents of waste, combining the visual analysis of images with values acquired from a moisture sensor (Wagland et al., 2013). Finally, the method was used to determine the energy contents of SRF mixtures (Peddireddy et al., 2015), analysing images of

shredded materials. This visual analysis was used to estimate the composition of every image of waste into 11 categories like paper, plastics, and metals. The composition was combined with average calorific content for each of the material classes to indirectly estimate the energy contents of the mixtures. The technique takes between 3 and 4 minutes because it requires a person to analyse and categorise 121 points from a single image of waste.

1.4 Aim and objectives

The use of deep neural networks has been thoroughly investigated for the visual classification of objects and materials, and it has been demonstrated that the composition of mixed waste materials can be determined through visual analysis of their images. However, there remains an outstanding question as to whether deep neural networks are a suitable methodology for the automatic classification of materials in the waste industry. This is the first work to investigate the specific challenges of waste recognition through deep learning, including the random presentation of materials, changes to colours and illumination, variable scales and sizes, cluttered appearance of mixed materials, the different common types of solid materials, and their high intra-class variabilities (Adamcová and Vaverková, 2014; Canopoli et al., 2018; Peddireddy et al., 2015; Wagland et al., 2013, 2012). The research aims and objectives were defined as follows.

1.4.1 Overall aim and research question

The aim of this project is to develop, define and evaluate a methodology for the classification of images of waste to analyse the composition of solid waste streams automatically and accurately in material recovery facilities.

The formulated research question is:

Given the current state of available technologies, is it possible to develop and define a methodology for the classification of images of waste to analyse the composition of solid waste automatically and accurately in material recovery facilities?

1.4.2 Objectives

The research objectives set to achieve the research aim are:

1. To identify the state of the art in classification of materials using advanced sensor technologies and assess the potential impact and limitations of using visible sensor cameras in the context of waste management (Chapter 2).

2. To design, develop and test an image acquisition method to capture images of waste samples with representative variability to create a high-quality dataset for training a deep learning model for the classification of images of waste (Chapter 3)

3. To investigate the importance of coloured content in the images for the recognition of waste and critically evaluate the recognition performance when using greyscale images to create a solution to typical operational challenges (Chapter 4)

4. To understand the effects of different spatial scales and apply scale jittering as a data augmentation technique to generalise the solution to different operational conditions in waste management (Chapter 5)

5. To formulate, test and define an end-to-end pipeline of image acquisition, image processing and data augmentation techniques to increase the diversity of images and improve the generalisation of deep learning models for the classification of waste in the field (Chapters 3, 4, and 5)

6. To synthesise the findings from objectives 1 to 5 into a methodology to automatically estimate the composition of mixed waste using image-based classification, and to appraise its practical uses in the waste industry (Chapter 6)

The work discussed in the following chapters evaluates the potential and limitations of deep learning for the identification of images of solid waste materials. A critical literature review analyses and compares different advanced sensor technologies to inform processes and stakeholders of waste material recovery plants. An image capture technique and data processing strategies are implemented and tested, and a methodology is progressively developed for training deep neural networks for waste material classification. The networks are validated through applications in relevant waste classification tasks and their performances are measured and compared to state-of-the-art solutions. Finally, a mixed material composition analysis method is implemented using trained deep neural networks and is validated in the context of a material recovery facility.

1.5 Thesis structure

The thesis is presented as a series of chapters formatted as papers for publication in the *Waste Management* journal. Chapter 1 provided the context and motivation of the work, highlighted the research gap and stated the aim and objectives of the research. Each subsequent chapter builds upon the previous one, starting with a literature review and followed by the experimental work.

Chapter 2 is a literature review of methodologies for waste sorting and characterisation focused on fuel production. It makes the comparison to other industries and approaches the problem through a process analytical technology framework. Based on the findings from this chapter, optical imaging using visible sensor cameras were identified and selected as the data capture technology for the experimental work presented in the subsequent chapters. This chapter was published as: Vrancken, C., Longhurst, P., Wagland, S., 2017. *Critical review of real-time methods for solid waste characterisation: Informing material recovery and fuel production*. Waste Management 61, 40–57.

Chapter 3 developed and tested a methodology to capture images using multiple cameras and illumination sources combined with data augmentation techniques to improve the classification of waste using deep neural networks. The basic methodology established from these findings was applied and extended in the

experimental works presented in the subsequent chapters, each focused on different research areas (Figure 1-1).

Chapter 4 tested the methodology to determine the effects of intra-class variability in the training datasets and recommended a data processing strategy to reduce negative effects of colour changes of waste samples.

Chapter 5 provided a qualitative analysis of the neural network decision process and extended the methodology to improve the classification of images captured with different spatial resolutions.

Chapter 6 improved the image capture to deal with non-uniform illumination of material samples and developed a methodology to determine the composition of mixed waste materials based on classification of patches.

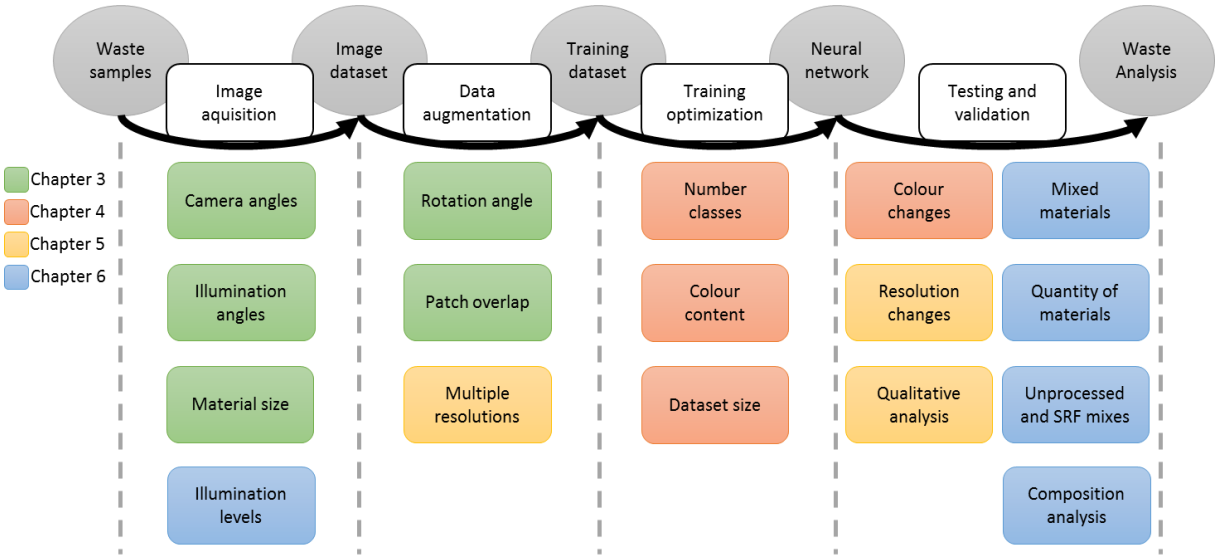


Figure 1-1. General methodology and research areas covered in each chapter

Chapter 7 is a discussion with a critical interpretation of the findings from the experimental works, identifying the contributions to knowledge to academia and industry. The limitations of the proposed methods and strategies are acknowledged, the impact for waste management is explored and further research is recommended.

Chapter 8 contains the conclusions, indicating how the aim and objectives of the project were fulfilled.

1.6 References

Adamcová, D., Vaverková, M., 2014. Degradation of Biodegradable/Degradable Plastics in Municipal Solid-Waste Landfill. Polish Journal of Environmental Studies 23, 1071–1078.

Adedeji, O., Wang, Z., 2019. Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network. Procedia Manufacturing, The 2nd International Conference on Sustainable Materials Processing and Manufacturing, SMPM 2019, 8-10 March 2019, Sun City, South Africa 35, 607–612. https://doi.org/10.1016/j.promfg.2019.05.086

Aral, R.A., Keskin, Ş.R., Kaya, M., Hacıömeroğlu, M., 2018. Classification of TrashNet Dataset Based on Deep Learning Models, in: 2018 IEEE International Conference on Big Data (Big Data). Presented at the 2018 IEEE International Conference on Big Data (Big Data), pp. 2058–2062. https://doi.org/10.1109/BigData.2018.8622212

Awe, O., Mengistu, R., Sreedhar, V., 2017. Final Report: Smart Trash Net: Waste Localization and Classification. Stanford University.

Bell, S., Bala, K., 2015. Learning visual similarity for product design with convolutional neural networks. ACM Transactions on Graphics 34, 98:1-98:10. https://doi.org/10.1145/2766959

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Bello-Cerezo, R., Bianconi, F., Cascianelli, S., Fravolini, M.L., Maria, F. di, Smeraldi, F., 2017. Hand-Designed Local Image Descriptors vs. Off-the-Shelf CNN-Based Features for Texture Classification: An Experimental Comparison, in: Intelligent Interactive Multimedia Systems and Services 2017, Smart Innovation, Systems and Technologies. Presented at the International

Conference on Intelligent Interactive Multimedia Systems and Services, Springer, Cham, pp. 1–10. https://doi.org/10.1007/978-3-319-59480-4_1

Bircanoğlu, C., Atay, M., Beşer, F., Genç, Ö., Kızrak, M.A., 2018. RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks, in: 2018 Innovations in Intelligent Systems and Applications (INISTA). Presented at the 2018 Innovations in Intelligent Systems and Applications (INISTA), pp. 1–7. https://doi.org/10.1109/INISTA.2018.8466276

Canopoli, L., Fidalgo, B., Coulon, F., Wagland, S.T., 2018. Physico-chemical properties of excavated plastic from landfill mining and current recycling routes. Waste Management 76, 55–67. https://doi.org/10.1016/j.wasman.2018.03.043

Cavalin, P.R., Kapp, M.N., Martins, J., Oliveira, L.E., 2013. A multiple feature vector framework for forest species recognition, in: Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM, pp. 16–20.

Cimpan, C., Maul, A., Jansen, M., Pretz, T., Wenzel, H., 2015. Central sorting and recovery of MSW recyclable materials: A review of technological state-of-the-art, cases, practice and implications for materials recycling. Journal of Environmental Management 156, 181–199. https://doi.org/10.1016/j.jenvman.2015.03.025

Cimpoi, M., Maji, S., Kokkinos, I., Vedaldi, A., 2016. Deep Filter Banks for Texture Recognition, Description, and Segmentation. International Journal of Computer Vision 1–30. https://doi.org/10.1007/s11263-015-0872-3

David McKinnon, Jak Fazakerley, Rolph Hultermans, 2017. Waste Sorting Plants: Extracting Value from Waste. ISWA.

DECC, 2009. National Renewable Energy Action Plan for the United Kingdom [WWW Document]. URL https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/47871/25-nat-ren-energy-action-plan.pdf (accessed 10.19.16).

DEFRA, 2019. Consultation on reforming the UK packaging producer responsibility system [WWW Document]. URL

https://consult.defra.gov.uk/environmental-quality/consultation-on-reforming-the-uk-packaging-produce/ (accessed 11.9.19).

DEFRA, 2018. 25 Year Environment Plan [WWW Document]. URL https://www.gov.uk/government/publications/25-year-environment-plan (accessed 10.3.19).

EU CEN, 2012. Solid Recovered Fuels [WWW Document]. URL https://standards.cen.eu/dyn/www/f?p=204:32:0::::FSP_ORG_ID,FSP_LANG_ID:407430,25&cs=1E692895E0FA13AE68B9FA01D5A630ED7 (accessed 5.4.16).

EU Commission, 2018. EUR-Lex - 52018DC0029 - EN - EUR-Lex [WWW Document]. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A29%3AFIN (accessed 12.18.19).

EU Council, 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:FULL:EN:PDF (accessed 10.19.16).

EU Council, 2008. Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste and repealing certain Directives [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:312:0003:0030:en:PDF (accessed 5.21.16).

EU Council, 2001. EU Sustainable Development Strategy [WWW Document]. URL https://ec.europa.eu/environment/sustainable-development/strategy/index_en.htm (accessed 11.14.19).

EU Council, 1999. Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. Official Journal of the European Communities L 182/1, 1–19.

EU Parliament, 2018. Circular economy package: Four legislative proposals on waste [WWW Document]. URL

http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2018)614766 (accessed 4.24.18).

EU Parliament, 2017. The circular economy package: new EU targets for recycling [WWW Document]. URL https://www.europarl.europa.eu/news/en/headlines/society/20170120STO59356/the-circular-economy-package-new-eu-targets-for-recycling (accessed 12.18.19).

Flamme, S., Geiping, J., 2012. Quality standards and requirements for solid recovered fuels: a review. Waste Management & Research 30, 335–353. https://doi.org/10.1177/0734242X12440481

Guerrero, L.A., Maas, G., Hogland, W., 2013. Solid waste management challenges for cities in developing countries. Waste Management 33, 220–232. https://doi.org/10.1016/j.wasman.2012.09.008

Hafemann, L.G., Oliveira, L.S., Cavalin, P., 2014. Forest species recognition using deep convolutional neural networks, in: 2014 22nd International Conference on Pattern Recognition. IEEE, pp. 1103–1107.

Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics 610–621.

ISWA, 2016. Circular Economy: Energy and Fuels [WWW Document]. URL http://www.iswa.org/fileadmin/galleries/Task_Forces/Task_Force_Report_5.pdf (accessed 4.15.16).

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Lorber, K.E., Sarc, R., Aldrian, A., 2012. Design and quality assurance for solid recovered fuel. Waste Management & Research 30, 370–380. https://doi.org/10.1177/0734242X12440484

Nixon, M., Aguado, A.S., 2012. Feature Extraction & Image Processing for Computer Vision, Third Edition, 3rd ed. Academic Press, Inc., USA.

Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition 29, 51–59. https://doi.org/10.1016/0031-3203(95)00067-4

Ozkaya, U., Seyfi, L., 2019. Fine-Tuning Models Comparisons on Garbage Classification for Recyclability. ArXiv abs/1908.04393.

Peddireddy, S., Longhurst, P.J., Wagland, S.T., 2015. Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste Management 40, 9–13. https://doi.org/10.1016/j.wasman.2015.03.015

Ruiz, V., Sánchez, Á., Vélez, J.F., Raducanu, B., 2019. Automatic Image-Based Waste Classification, in: Ferrández Vicente, J.M., Álvarez-Sánchez, J.R., de la Paz López, F., Toledo Moreo, J., Adeli, H. (Eds.), From Bioinspired Systems and Biomedical Applications to Machine Learning, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 422–431. https://doi.org/10.1007/978-3-030-19651-6_41

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y

Sakr, G.E., Mokbel, M., Darwich, A., Khneisser, M.N., Hadi, A., 2016. Comparing deep learning and support vector machines for autonomous waste sorting, in: Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International. IEEE, pp. 207–212.

Sarc, R., Lorber, K., Pomberger, R., Rogetzer, M., Sipple, E., 2014. Design, quality, and quality assurance of solid recovered fuels for the substitution of

fossil feedstock in the cement industry. Waste management & research 32, 565–585. https://doi.org/10.1177/0734242X14536462

Sarc, R., Lorber, K.E., 2013. Production, quality and quality assurance of Refuse Derived Fuels (RDFs). Waste Management 33, 1825–1834. https://doi.org/10.1016/j.wasman.2013.05.004

Sedlak, D., 2017. From the Ontonagon Boulder to the Circular Economy. Environ. Sci. Technol. 51, 1941–1942. https://doi.org/10.1021/acs.est.7b00430

Sharan, L., Liu, C., Rosenholtz, R., Adelson, E.H., 2013. Recognizing Materials Using Perceptually Inspired Features. Int J Comput Vis 103, 348–371. https://doi.org/10.1007/s11263-013-0609-0

Sharan, L., Rosenholtz, R., Adelson, E.H., 2014. Accuracy and speed of material categorization in real-world images. Journal of vision 14, 1–24. https://doi.org/10.1167/14.9.12

Tchobanoglous, G., Kreith, F., 2002. Handbook of solid waste management. New York : McGraw-Hill, 2002.

UK Environment Agency, 2015. Reasons for trends in English refuse derived fuel exports since 2010 [WWW Document]. URL https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/438906/Reasons_for_trends_in_English_refuse_derived_fuel_exports_since_2010_report.pdf (accessed 10.14.16).

Velis, C.A., Longhurst, P.J., Drew, G.H., Smith, R., Pollard, S.J.T., 2010. Production and quality assurance of solid recovered fuels using Mechanical-Biological Treatment (MBT) of waste: a comprehensive assessment. Critical Reviews in Environmental Science and Technology 40, 979–1105. https://doi.org/10.1080/10643380802586980

Wagland, S.T., Dudley, R., Naftaly, M., Longhurst, P.J., 2013. Determination of renewable energy yield from mixed waste material from the use of novel image analysis methods. Waste Management 33, 2449–2456. https://doi.org/10.1016/j.wasman.2013.06.021

Wagland, S.T., Veltre, F., Longhurst, P.J., 2012. Development of an image-based analysis method to determine the physical composition of a mixed waste material. Waste Management 32, 245–248. https://doi.org/10.1016/j.wasman.2011.09.019

WRAP, 2012. A Classification Scheme to Define the Quality of Waste Derived Fuels [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/WDF_Classification_6P%20pdf.pdf (accessed 4.12.16).

WRAP, 2014. Sampling and Testing Guidance for Material Facilities [WWW Document]. URL https://wrap.org.uk/sites/default/files/2020-10/WRAP-MF%20Sampling%20Guidance%20April%202014.pdf (accessed 11.8.19).

Yang, M., Thung, G., 2016. Classification of Trash for Recyclability Status (Report). Stanford University, http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf (accessed 11.08.19).

2 Critical review of real-time methods for solid waste characterisation: informing material recovery and fuel production

Abstract

Waste management processes generally represent a significant loss of material, energy, and economic resources, so legislation and financial incentives are being implemented to improve the recovery of these valuable resources whilst reducing contamination levels. Material recovery and waste derived fuels are potentially valuable options being pursued by industry, using mechanical and biological processes incorporating sensor and sorting technologies developed and optimised for recycling plants. In its current state, waste management presents similarities to other industries that could improve their efficiencies using process analytical technology tools. Existing sensor technologies could be used to measure critical waste characteristics, providing data required by existing legislation, potentially aiding waste treatment processes and assisting stakeholders in decision making. Optical technologies offer the most flexible solution to gather real-time information applicable to each of the waste mechanical and biological treatment processes used by industry. In particular, combinations of optical sensors in the visible and the near infrared range from 800 nm to 2500 nm of the spectrum, and different mathematical techniques, are able to provide material information and fuel properties with typical performance levels between 80% and 90%. These sensors not only could be used to aid waste processes, but to provide most waste quality indicators required by existing legislation, whilst offering better tools to the stakeholders.

Keywords: Solid waste; sensors; waste derived fuel; energy from waste; analysis

2.1 Introduction

Changes in legislation and the introduction of economic incentives are promoting the recovery of valuable resources and the reduction of landfill waste disposal (Council of the European Union, 1999). Sustainable options for waste management are being investigated with a focus on optimising the recovery of

recyclable materials. Residual waste, which is not viable for being reused or recycled, contains calorific value which could partially displace fossil-derived fuels (European Parliament., 2009), and the recovery of this energy is a sustainable option, and certainly more manageable, when compared to landfill biogas (Jeswani and Azapagic, 2016). Therefore, waste is considered a valuable resource with potential of being a major energy contributor (Slade et al., 2011), and many innovations are being introduced to retrieve this energy efficiently (Baytekin et al., 2013; Lombardi et al., 2015). In Europe, energy recovery from waste is one of the recommended operations by the existing legal framework (Council of the European Union., 2008) to reduce landfill use.

Waste (or refuse)- derived fuel (WDF or RDF) is defined as '*a heterogeneous group of non-hazardous wastes that do not cease to be such by virtue of their being used to generate energy without a greater negative environmental impact than landfill disposal*'(WRAP, 2012a). Refuse-derived fuels are prepared by processing waste materials from municipal, industrial and commercial sources, and they are typically used in co-incineration plants in different industries, such as cement kilns and power stations (Lorber et al., 2012; Sarc and Lorber, 2013). In contrast, solid recovered fuel (SRF) is a term used by waste industry to indicate a composition of waste derived fuel which meets strict quality specifications. Existing SRF standards at European and UK levels allow authorities, waste management service providers and equipment manufacturers to work together in a regulated environment.

The European Committee for Standardization (CEN) and Technical Committee (TC) established the standard specifications and requirements for SRF (European Committee for Standardization, 2012). A total of 33 standards and reports were published to cover all the different aspects involved in SRF quality assurance, such as determination of calorific value and moisture contents, and methods for sampling (British Standards Institution, 2011a, 2011b, 2011c). The main indicators used by the CEN/TC 343 standards to characterise SRF are the net calorific value (NCV), the chlorine (Cl) content, and the mercury (Hg) content,

which point to the economic, technical, and environmental quality of the fuel, respectively (British Standards Institution, 2011d).

The most adopted waste management strategy is to implement material recovery facilities (MRFs) (Velis et al., 2010) to process municipal solid waste (MSW) and refuse from commercial and industrial (C&I) or from construction and demolition (C&D) sources. An important goal of an MRF is to reduce the waste volume and mass before sending it to a landfill. Outputs from an MRF can include high-calorific material such as RDF and SRF, recyclable materials such as metals, and stabilised organic waste which could be used as fertiliser. Implementation of the correct waste treatment processes together with quality assurance are key to obtain the highest possible waste-derived fuel quality, and these are affected by the input waste and desired output properties (Lorber et al., 2012; Sarc et al., 2014; Sarc and Lorber, 2013). Data informing on waste processes, including flows, treatment efficiencies and timings are required by stakeholders to create new policies and improve waste treatment (RWM Ambassadors, 2016).

The composition of mixed solid waste is highly variable, including different materials and contaminant levels, which represents a challenge to MRF to meet the required RDF and SRF qualities with consistency (Cimpan et al., 2015; Flamme and Geiping, 2012). Currently, the waste received is processed through a combination of manual sorting, and mechanical and biological technologies (MBT) which are selected according to the type of input waste and desired output characteristics (DEFRA, 2013; Velis et al., 2010). The plants are designed to balance manual sorting processes, with high efficiency but increased labour costs, and automated sorting processes, which require more investment capital (Tchobanoglous and Kreith, 2002; WRAP, 2006).

Input waste is received into a tipping hall, where it is stored protected from the weather for up to two days (WRAP, 2007). The transport of waste from one process unit to another is typically done with conveyor belts, whilst input waste is loaded with grabbing cranes and output material is moved with forklifts and wheel loaders (Tchobanoglous and Kreith, 2002). An illustration of the typical waste flow

between manual, mechanical and biological treatment processes is depicted in Figure *2-1*.

Due to the difficulties and limitations of these processes to separate materials and sort them based on grades, to obtain high quality waste-derived fuel (Cimpan et al., 2015; Rotter et al., 2004), advanced sensors are being developed to improve and complement these methods. These new technologies are required to adapt to short- and long-term changes in waste composition, introduction of new materials and increasing legal requirements (Cimpan et al., 2015). Stakeholders, particularly in developing countries, prefer waste management systems with high efficiency and lower costs (Guerrero et al., 2013). The integration of multiple sensors, improved energy efficiency, and ruggedness to withstand the extreme environment conditions are some of the technical challenges to be faced by future research and development in this subject (Gundupalli et al., 2016).

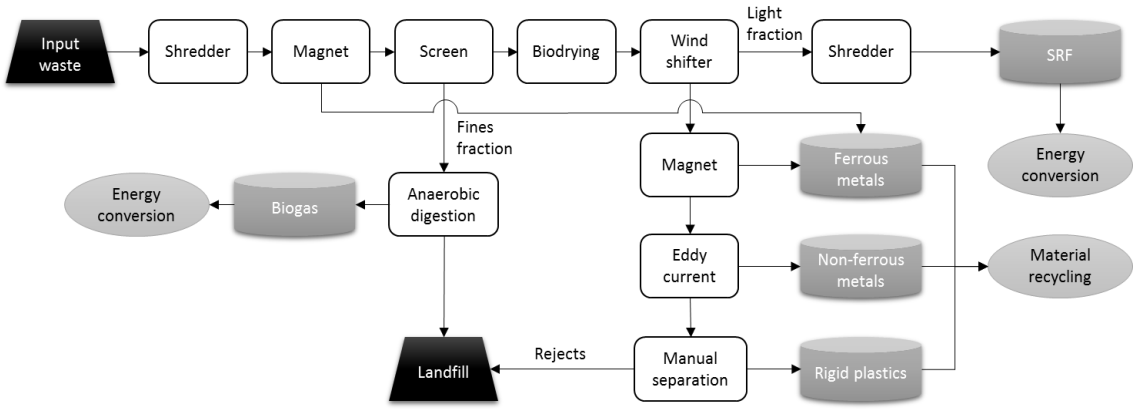


Figure 2-1. Flow sheet diagram of a typical MRF using manual and MBT processes to produce four outputs: biogas and SRF for energy conversion, materials for recycling, and rejects going into landfill. Adapted from DEFRA (2013).

2.1.1 Manual sorting

Hand picking or manual sorting is a physical operation performed by plant personnel working in special facilities installed at different waste processing stages. The goal is to either sort the waste items based on material type, or to remove those items that could reduce the performance of other processes downstream (Tchobanoglous and Kreith, 2002). Humans have a unique ability to

recognise materials in a very short time (Sharan et al., 2014), with observed waste recovery efficiencies up to 95% (Tchobanoglous and Kreith, 2002). Many factors affect this efficiency, such as fatigue, circulation speed, particle size distribution and volumetric flow of the waste material, and environmental aspects such as working temperature and lighting. The overall manual sorting efficiency is proportional to the number of people performing these tasks, but also requiring larger facilities.

2.1.2 Biological processes

Biological processes are applied in an MRF when the input waste stream contains biodegradable materials, which is typically contained in municipal solid waste (MSW) (Nasrullah et al., 2015). Waste from commercial and industrial, or from construction and demolition sources do not usually have a biodegradable fraction (Nasrullah et al., 2014a, 2014b), therefore biological processes are not required (Velis et al., 2010). There are two main types of biological processes used in MBT systems, biodrying and anaerobic digestion. With these processes it is possible to reduce the volume of waste going into landfill, in particular the biodegradable fraction. They can also be used to improve the quality of the waste derived fuel being produced, to produce stabilised bio-waste to be used as soil conditioner, or to generate biogas for energy conversion (DEFRA, 2013).

2.1.3 Mechanical processes

Mechanical processes for the preparation of waste material include size reduction (comminution), and separation based on size, shape, density, and magnetic properties. The performance or efficiency of the separation processes are defined by their recovery rate and purity grade. Recovery rate is defined as the ratio of the mass of the particular component appearing in an output fraction and the total mass of that component in the input stream. The purity grade expresses the level of cleanliness of a particular component in an output fraction, as the ratio of the mass of the particular component present in that output fraction and the total mass of that output fraction. Several factors affect the performance of the different process units, including the moisture level, particle shapes and size distribution, and the presence of large or hard objects (Table 2-1).

Table 2-1. Mechanical and biological processes, their effects on solid waste streams and operational parameters affecting their efficiency.

Process	Main objective	Effects on waste	Factors affecting performance	Main operational variables
Bag splitting	Increase waste stream throughput [1,2]	Enable mechanical separation; increase energy recovery	Presence of large or hard items in the input stream [1,3]	Distance between cutting shears; shears rotational speed[3]
Shredding	Reduce particle size and homogenise waste stream [1,2]	Increase biochemical reactivity; reduce bulk volume; enable mechanical separation; increase energy recovery	Presence of large or hard items in the input stream [1,3]	Distance between cutting shears; shears rotational speed[3]
Screening	Separate a particular shape, and larger from smaller fractions [1,2]	Remove ash and moisture content; increase WDF calorific value [4–7]	Presence of highly ductile material which end up in fines fraction; particle size distribution; throughput[1]	Filter size and shape; residence time[1]
Air separation	Separate lighter from heavier fractions [1,2]	Remove non-combustible material; increase energy recovery [4–7]	Absorbed moisture in light materials which end up in heavy fraction; particle size distribution[8,9]	Distance between separation chambers; air flow rate[9]
Ballistic separation	Separate lighter from heavier, and smaller from larger fractions [1,2]	Remove organics and non-combustible material; increase energy recovery [2,4–6]	Absorbed moisture in light materials which end up in heavy fraction; particle size distribution [2]	Filter size and shape; vibration frequency; inclination angle [2]

Process	Main objective	Effects on waste	Factors affecting performance	Main operational variables
Magnetic separation	Separate ferrous metals from waste stream[1,2,10]	Remove non-combustible material; increase energy recovery [1,2,4–6]	Particle size distribution; moisture level[1,10]	Magnetic field strength[1,10]
Eddy current separation	Separate non-ferrous metals from waste stream[1,2,11]	Separate non-combustible material; increase energy recovery[2,12,13]	Particle size distribution; moisture level; presence of aggregated materials; throughput; humidity [1,11,14]	Electric field strength; distance between separation chambers[1,11]
Biodrying	Reduce moisture and waste stream volume[1,2,15]	Increase energy recovery[4–6,15,16]	Moisture level; composition and presence of organic material[16]	Aeration flow; temperature; residence time[16]
Anaerobic digestion	Recover energy and converts biodegradable contents to gas[1,2]	Reduce overall solid mass[1,2]	Particle size distribution; composition; organic material; biochemical methane potential[17–19]	Feed pump flow; buffer substance concentration; heat; pH[17–19]

[1]Tchobanoglous and Kreith (2002), [2]Velis et al.(2010), [3]Fitzgerald and Themelis (2009), [4]Nasrullah et al.(2015), [5]Nasrullah et al.(2014a), [6]Nasrullah et al. (2014b), [7]Cimpan and Wenzel (2013), [8]Rotter et al.(2004), [9]Shapiro and Galperin (2005), [10]Svoboda and Fujita (2003), [11]Rem et al. (1998), [12]Ionescu et al. (2013), [13]Mesina et al.(2007), [14]Jujun et al.(2014), [15]de Araújo Morais et al.(2008), [16]Velis et al. (2009), [17]Zhang and Banks(2013), [18]John and Singh(2011), [19]Sponza and Ağdağ (2005)

2.2 Advanced sensors

These new types of automated processes use combinations of advanced, non-invasive sensors and mechanical processes to provide fast separation and sorting of materials. Usually, the sensor is used to detect a particular characteristic of a material, classifying it and deciding its destination fraction using a computer system, and then a mechanical process separates it from the mixed stream accordingly (Velis et al., 2010). Advanced online processes are fundamental to reliably sort materials, increasing the waste processing plant efficiency by producing high purity recycled and SRF materials (WRAP, 2010).

2.2.1 Fluorescence

Fluorescence sensors determine the presence of a certain material by measuring the light emission generated after the material is excited with a light. Of particular interest is the fluorescence property of lignin to emit light at near 650nm when excited with visible light. For example, some proposed methods (Ammineni, 2001; Mallapragada, 2004; Ramasubramanian et al., 2005) can separate newspaper from mixed papers in real-time but the results are affected by coloured paper, with red ink reflecting light close to the 650nm wavelength which the sensor can confuse with lignin fluorescence, whilst green and blue inks reduce the lignin fluorescence effect.

2.2.2 Hydrophobicity

By using electrochemical properties of materials, such as their surface hydrophobicity, it is possible to sort plastics by flotation mechanisms (Fraunholcz, 2004; Wang et al., 2015). For example, the method investigated by Saisinchai (2014) separates polyvinyl chloride (PVC) from polyethylene terephthalate (PET) plastics using a froth flotation method with calcium lignosulfonate as the wetting reagent. A review of flotation mechanisms to separate plastics, mainly PVC or PET from polystyrene or polyethylene was conducted by Wang et al. (2015).

2.2.3 Laser induced breakdown spectroscopy

Laser induced breakdown spectroscopy (LIBS) involves focussing a high energy laser onto a material which generates plasma emissions with a particular light spectrum. These emissions are detected with a spectrograph and analysed to determine the material chemical composition, such as contaminants. For example, Barefield et al. (1995) used this technique to detect the presence of Strontium in mixed waste. Similarly, the presence of contaminated waste wood can be detected with LIBS, with measurement times of 3-5 seconds per sample (Rasem Hasan et al., 2011). A comprehensive list of detection limits obtained for different chemicals has been published by Zorov et al. (2015). Huber et al. (2014) investigated the use of LIBS to detect chlorine in industrial waste in real-time, without being affected by the visible colour of the plastics.

The use of LIBS was also explored to identify plastic materials by analysing the emitted plasma to detect the presence of carbon and hydrogen (Gondal and Siddiqui, 2007). The combination of different multivariate correlation methods and LIBS were investigated to sort different types of polymers (Anzano et al., 2011; Banaee and Tavassoli, 2012). Similarly, LIBS was studied by Aguirre et al. (2013) who characterised polymers and other materials used in electronic equipment. In the context of metal scrap sorting, LIBS was used to classify aluminium alloys by Werheit (2011), whilst another project obtained real-time classification of 10 different classes of metal alloys (Merk et al., 2015).

2.2.4 Microwaves

These sensors measure the interaction of electromagnetic fields at microwave frequencies, i.e. electromagnetic waves, and dielectric materials. For example, resonant sensors can be used to measure changes in resonant frequencies due to the absorbed water in a moisture sensitive ceramic. This technique was successfully tested to determine the contents of moisture in construction and demolition materials (Quinn and Kelly, 2010; Sokoll and Jacob, 2007). Wagland et al. (2013) investigated the use of microwaves to measure moisture level in mixed solid waste, using a microwave setup with a transmitter and an array of

sensors to detect signal absorption and transmission in waste materials (wood, cardboard, textile) to estimate their moisture content.

2.2.5 Resistivity and capacity

This group of sensors can detect changes in electrical fields due to variations of pressure, humidity, and conductivity. They range from very simple to complex circuitry designs that allow for increased functionality, such as self-calibration and error compensation. In general, they are low cost, small sized and with low power consumption (Du, 2014). Resistive sensors measure changes in resistivity between two electrodes when in the presence of conductor media such as water. A proposed method (Gawande et al., 2003) calibrated a resistive sensor using mixed solid waste. Capacitive sensors measure the dielectric permittivity of materials between two plates (Udo and Christof, 2010). The method investigated by Fuchs et al. (2008) measures moisture content in municipal solid waste falling off a conveyor belt with a capacitive sensor.

2.2.6 Stiffness

It is possible to sort paper by using its elastic properties, as discussed by Baum (1987). For example, one researched method measures the deflection in path trajectory caused by a jet of air to classify grades of paper based on their different stiffness properties (Katuri, 2006; Ramasubramanian et al., 2006, 2007). This technique is able to sort paper based on their grades, such as newspaper, office paper and cardboard, in real-time on a conveyor belt at a speed of 1.5 m/s (Ramasubramanian et al., 2012).

2.2.7 Ultrasound imaging

This method sends ultrasound pulses to a material, which reflects or scatters back the ultrasound waves as they travel though the material, with varying intensity according to the material density. An ultrasound sensor receives and processes these reflections (ultrasound echoes) to generate an ultrasound image. A proposed method (Faibish et al., 1997) uses a combination of optical and ultrasonic sensors to detect and remove paper from municipal waste in real-time. Ultrasound techniques have also been investigated to recognize and

quantify plastic materials in waste in real-time using a wet medium (Sanaee and Bakker, 2009), showing a marked difference in signal amplitude between polypropylene and polyethylene materials.

2.2.8 Optical imaging using visible wavelength sensors

Visible optical sensors are based on the surface properties of materials to reflect varying intensities of light in wavelengths visible to the human eye (400-700 nm). Visible image sensors, such as complementary metal-oxide semiconductor (CMOS) and charged couple devices (CCD) cameras, produce images with pixel intensities proportional to the visible light received (Janesick, 2007). Unique features in the materials and the generated images, such as shape, colour and surface textures can be used to recognize objects, for example bottles (Nawrocky et al., 2010; Torres-García et al., 2015; Wahab et al., 2006), cans (Torres-García et al., 2015; Yani and Budiman, 2015), minerals (Anding et al., 2011, 2013), and electronic boards (Kleber et al., 2015; Torres-García et al., 2015). Optical imaging has also been studied to determine particle size distributions of different mixed materials (Bianconi et al., 2015; Nakamura et al., 2005; Salehizadeh and Sadeghi, 2010; Urbanski et al., 2011). Typical issues affecting these visible optical imaging methods are changes in illumination, or the presence of different materials or objects presenting similar visible features such as colour.

The method investigated by Huang et al. (2010) uses a 3D camera to extract both colour and shape information to classify waste objects. This setup had a high classification rate for non-ferrous metals, plastics bottles and coins. Identification of polypropylene and acrylonitrile butadiene styrene had the highest rate, but the detection of polymers was affected by their surface colours, as only 20% of black plastics were identified correctly. Rahman et al. (2009a, 2009b, 2010, 2011, 2012b, 2012a, 2012c, 2014, 2015) developed and tested a series of mathematical algorithms to process images acquired with a colour webcam, separating different grades of paper. Visible image processing was also investigated to sort wood according to type, colour and quality grades (Bombardier and Schmitt, 2010; Faria et al., 2008; Hrćka, 2008; Kurdthongmee, 2008; Lu et al., 1997).

The imaging method developed by Wagland et al. (2012, 2013) captured images of mixed solid waste to manually identify materials and estimate the waste composition and biogenic energy potential of waste. This methodology was further researched to estimate the composition of solid recovered fuels (Peddireddy et al., 2015). These methods used waste composition as a surrogate for estimating the fuel properties of waste, either at the collection point or at the end of the MBT process, requiring a database of properties for the individual waste materials.

2.2.9 Optical imaging using infrared wavelength sensors

Infrared sensors are able to detect variations in absorption, transmittance and scattering of light in infrared wavelengths produced by different materials. The infrared spectrum is typically divided into several regions: near infrared (NIR, wavelengths: 700-1400 nm), short-wave infrared (SWIR, wavelengths: 1400-3000 nm), medium-wave infrared (MWIR or MIR, wavelengths: 3000-8000 nm), and long-wave infrared (LWIR or LIR, wavelengths: 8000-15000 nm). Both MIR and LIR methods require longer acquisition times than NIR and SWIR, but offer more defined spectral information. On the other hand, acquisition of NIR and SWIR spectra is quicker, but they require advanced mathematical techniques, such as multivariate statistical analysis to extract information about material components (Reich, 2005).

Typical infrared image sensors include CCD cameras, which produce a digital image with pixel intensities proportional to the infrared light reflected or scattered back by a material in that particular position. Multispectral imaging or hyperspectral imaging (HSI) methods capture the spectral response for each pixel in the digital image, which is affected by the material composition for that particular position. An early multispectral method was developed by Rantanen et al. (1998), using NIR sensors tuned to 3 different SWIR wavelengths to estimate moisture content in pharmaceutical pellets in real-time. More recently, the studies by Achata et al. (2013) and by Wünsch and Jenkins (2013) used NIR hyperspectral imaging to determine moisture content in food and mixed solid waste samples, respectively.

Some methodologies are developed to use more than one region of the light spectrum. For example, McWhirt et al. (2012) used a visible and near infrared (VIS-NIR) system to determine the contents of organic matter in compost sourced from different types of feedstock in real-time. Techniques using NIR-MIR hyperspectral imaging to determine biochemical methane potential from municipal solid waste were tested with positive results using statistical analysis (Lesteur et al., 2011, 2013), whilst similar results were obtained by other researchers in the review article by Ward (2016). The utilisation of NIR-SWIR imaging technology for the recycling of textiles were explored with government and industry support (European Commission, 2012; Ishfaq, 2015; Luiken and Bos, 2010).

Other uses of optical infrared methods are to identify specific substances or materials grades, such as the investigation of Mauruschat et al. (2015) to use NIR hyperspectral imaging to detect chemical preservatives and plastics contaminants in wood. Classification of glass contaminants in mixed glass materials using NIR and MIR spectral imaging was investigated by Farcomeni et al. (2008). Tatzer et al. (2005) used hyperspectral NIR imaging for the recycling of 4 different grades of cardboard and paper, and similar methods were investigated to sort materials from construction and demolition (C&D) sources (Anding et al., 2015; Kuritcyn et al., 2015; Palmieri et al., 2014).

Extensive research has been completed on sorting and recycling plastic materials, with early experiments with NIR hyperspectral imaging to quickly identify plastics from a distance (Huth-Fehre et al., 1995; van den Broek et al., 1996). Wienke et al. (1996) extended this method and used NIR-SWIR multispectral analysis to identify plastics from domestic waste. Waste electrical and electronic equipment (WEEE) contains recyclable plastics which could be sorted by NIR devices (WRAP, 2009). Beigbeder et al. (2013) used NIR-SWIR hyperspectral imaging to sort plastics material from a WEEE stream with similar results to those presented in an industry report (WRAP, 2009). More recently, techniques using NIR-SWIR hyperspectral imaging have been investigated for

separating fossil-based plastics from the more modern and environmentally friendly bio-plastics (Hollstein and Wohllebe, 2015).

Black or very dark plastics present a problem for optical methods as they absorb most of the light in VIS, NIR, and SWIR wavelengths and therefore, the reflected spectra do not contain enough information about the material composition (Huth-Fehre et al., 1995; WRAP, 2011). A potential solution to the black plastics problem was investigated by Kassouf et al. (2014), who used MIR-LIR hyperspectral imaging and statistical analysis methods to successfully identify 5 different types of polymers regardless of their colours.

Applications of NIR hyperspectral technology to waste management have been described by Bonifazi and Serranti (2014), suggesting its feasibility as a quality control tool. The use of NIR spectral information to directly determine the composition of waste was investigated by Smidt et al. (2008), using statistical analysis to classify waste into 3 groups: landfill, composts, and MSW. Pieber et al. (2012) investigated the use of commercial NIR sorting devices for the production of SRF, and their effect on removing chlorine, cadmium and lead contaminants. Krämer and Flamme (2015) presented the results of an industry sponsored project using NIR hyperspectral imaging and statistical analysis to determine SRF composition and fuel properties. This method used a database of material properties to indirectly estimate chlorine, calorific value, and ash contents, as originally described by Wagland et al. (2013) and Peddireddy et al. (2015).

2.2.10 X-ray imaging

X-ray is a form of high energy radiation which can penetrate materials, where it is absorbed or scattered depending on the material molecular density. The production of X-ray energy is done by applying a very high voltage in a vacuum tube, and the detection is typically done by a charge collection device behind the material, which is connected to an electronic system to estimate the transmitted energy. Therefore, transmission X-ray (XRT) measures the variation in transmitted radiation through different materials and can be applied to detecting contaminated plastics or minerals (Firsching et al., 2013), and presence of metals

(Knapp et al., 2014). Overlapping materials can be dealt with dual energy techniques (Li et al., 2016), but the less dense, smaller or thinner pieces of materials do not affect the x-ray radiation enough to produce a change in transmission. Fluorescence X-ray (XRF) is a technique to measure the fluorescent X-ray emitted after a material is excited with X-ray or gamma-ray radiation, and has been extensively investigated to detect preservative inorganic chemicals in waste wood from construction and demolition sources (Blassino et al., 2002; Block et al., 2007; Jacobi et al., 2007; Solo-Gabriele et al., 2004; Yasuda et al., 2006). The XRF methods were able to identify contaminated wood with a wide range of measurement time required per sample (0.25-6 seconds) (Rasem Hasan et al., 2011).

2.3 Process analytical technology

Process analytical technology (PAT) is a framework designed to encourage innovation in manufacturing and quality assurance processes, originally introduced in 2004 by the FDA to the pharmaceutical industry in USA (FDA, 2004). The PAT guidance recommends the use of tools for data acquisition, data and process analysis, and process control. Implementing PAT leads to the understanding of production processes and being able to assess quality in real-time. Ever since the introduction of PAT, the use of advanced technology in the pharmaceutical industry has increased (Dickens, 2010; Simon et al., 2015) with multiple examples of successful implementations (Chen et al., 2011; Mantanus et al., 2011; Reich, 2005; Rosas et al., 2012; Wu et al., 2011).

Process analysis concepts and tools are transferable (Simon et al., 2015), and activities other than pharmaceutical have adopted the PAT approach, including the production of biofuels (Chadwick et al., 2014), biotechnology (Gomes et al., 2015), food production (Tajammal Munir et al., 2015; Troy et al., 2016; van den Berg et al., 2013), and manufacturing of catalysts (Wu et al., 2016). In all these examples, the production plants are managed using PAT tools to collect key process information to feed-back or feed-forward a central control system, improving the quality of the product in real-time.

The existing standards only regulate SRF quality but make no recommendations regarding waste processes or how to achieve the required quality (EU CEN, 2012). Emerging technologies applicable to solid waste management systems (Gundupalli et al., 2016; Hannan et al., 2015; Lu et al., 2013) could make it possible to apply PAT for the production of SRF. Similar to other industries, the implementation of PAT could increase the understanding of the waste processes involved, allowing for better control and quality monitoring and providing a consistent product, as illustrated in Figure 2-2.

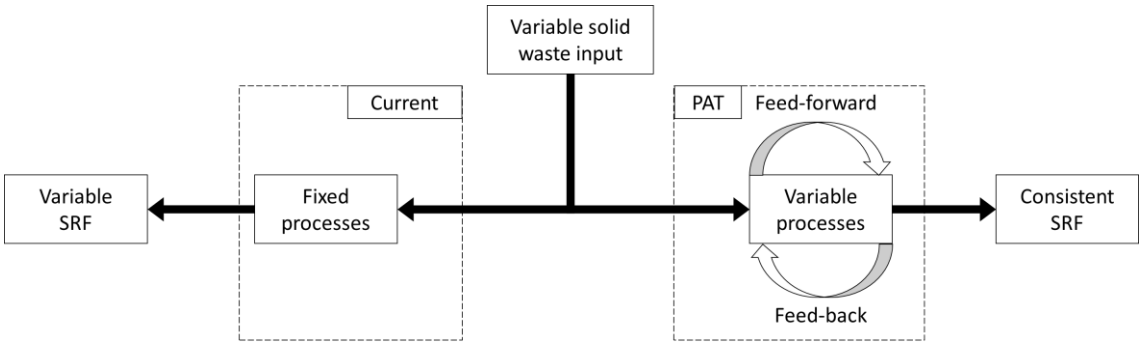


Figure 2-2. Current waste management strategy (left branch), receiving waste of variable content cannot control the product. The implementation of PAT (right branch) could allow the waste processing plants to produce a consistent product.

2.3.1 Critical quality attributes

In a PAT framework, critical quality attributes are measurable characteristics that determine and indicate the quality of the product (FDA, 2009). In the case of solid waste management, these indicators would be the factors affecting the performance of the MBT processes identified in Table 2-1. Additionally, for the case of an MRF producing SRF, the main quality indicators required by EU standards (British Standards Institution, 2011d) would also be measured. Finally, adverse conditions affecting advanced sensors should also be considered, such as dirtiness, deterioration or degradation of material surfaces, or presence of dust and vapour in the air, some of which originate prior to the waste material arriving at the processing plants, whilst others are the result of the MBT operations during the treatment process. These properties are currently difficult to measure and quantify. Effects from the MBT processes could potentially be estimated, thus

they currently represent a limitation and thus a challenge to the full implementation of PAT in the waste management industry. Therefore, further studies regarding their effects are needed.

The following section reviews the available advanced sensor technologies that could be used to measure each of the identified critical quality indicators. The performances of these technologies and methodologies are reported through (recognition and separation) rate values (%), coefficient of determination (R^2) and determination errors (%). However, most studies using discriminant analysis report the Root Mean Square Error (RMSE), the range error ratio (RER) and the ratio of predictive deviation (RPD), which are dimensionless statistical values to compare different models and their performances. Therefore, a qualitative performance rating scale of poor, fair, good, very good, and excellent was adopted (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006), enabling the comparison of published performance results of different technologies and methodologies, as shown in Table 2-2.

Table 2-2. Performance rating scale using the different commonly published performance parameters (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006), where R^2 is the coefficient of determination, RPD is the ratio of the standard deviation of the original data to standard error of prediction, RER is the ratio of the range of the original data to standard error of prediction, % error is the percent error between the predicted and the original data, and % rate is the success rate (recognition, classification, separation, etc.) of the method.

Qualitative rating	**Performance parameters**				
	R2	RPD	RER	% error	% rate
poor	<0.70	<2.5	<9	>30	<70
fair	0.70 to 0.80	2.5 to 3.0	9 to 12	20 to 30	70 to 80
good	0.80 to 0.90	3.0 to 3.5	12 to 15	10 to 20	80 to 90
very good	0.90 to 0.95	3.5 to 4.0	15 to 18	5 to 10	90 to 95
excellent	>0.95	>4.0	>18	<5	>95

2.3.1.1 Problematic objects

The presence of objects and materials containing hazardous substances such as batteries, compressed canisters or gas cylinders poses a health and safety risk to waste treatment plants and increases environmental contamination (European Commission, 2015). Furthermore, certain types of materials can affect the performance of some MBT processes, as a high percentage of hard materials or large pieces can damage bag splitting and shredding equipment, whilst ductile materials can reduce the performance of screening devices (Fitzgerald and Themelis, 2009; Tchobanoglous and Kreith, 2002).

Optical technologies have been used to identify and detect the presence of certain objects and materials, based on visible features such as shapes and colours, resulting in good, very good and excellent qualitative ratings (Table 2-3). These algorithms were developed to find specific objects in real time for industrial applications and no universal methodology has been developed. One identified problem is the detection of black or very dark objects, in particular plastics, which absorb the light and reduce their classification rates by 75%.

Table 2-3. Technologies and methods applied to problematic objects recognition. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Optical VIS-nir-swir	Mixed cans and bottles	Identifies aluminium cans and plastic bottles using a webcam and imaging algorithms using visible shape features.	**** rate=91%	Yani and Budiman (2015)
Optical VIS-nir-swir	Mixed papers	Classifies paper based on quality grades using a webcam and different imaging algorithms using visible surface features.	**** rate>90%	Rahman et al. (2015, 2014, 2012b, 2012a, 2012c, 2011, 2010, 2009a, 2009b)
Optical VIS-nir-swir	Mixed plastic	Machine vision system using visible images to separate PET from non-PET plastic bottles based on shape and colour.	*** rate=80-95%	Wahab et al. (2006)
Optical VIS-nir-swir	Mixed plastic containers	Classifies Polycoat containers based on visible 3D geometrical characteristics.	***** error<4%	Mattone et al. (2000)
Optical VIS-nir-swir	MSW (inorganic)	Classifies non-ferrous metals, plastic bottles and coins using visible 3D features.	**** rate=91-99% (dark objects=20%)	Huang et al. (2010)
Optical VIS-nir-swir	MSW (inorganic)	Uses a webcam and image processing to classify plastic cutlery, plastic bottles, and aluminium cans based on shape.	***** rate=98%	Torres-García et al. (2015)

2.3.1.2 Particle size distribution

The size of the materials following shredding affects the performance of screening devices, ballistic separators, eddy current separators, and anaerobic digestion (Table 2-1). Particle size distribution (PSD) is also an important quality indicator of SRF, and the use of image analysis for large materials and objects is recommended by existing standards (British Standards Institution, 2012). Optical technologies using visible images and computer vision algorithms have been investigated for the determination of PSD in C&D waste, in all cases with fair to very good qualitative ratings based on their accuracy rates (Table 2-4).

Table 2-4. Sensor technologies and methods to estimate particle size distribution. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Optical VIS-nir-swir	C&D waste	Imaging algorithms and regression analysis to estimate PSD from sands and rocks.	** - **** rate=72-93%	Bianconi et al. (2015)
Optical VIS-nir-swir	C&D waste	Imaging algorithms to estimate PSD from recognizable features and a lookup table.	*** - **** rate=82-95%	Di Maria et al. (2016)
Optical VIS-nir-swir	C&D waste	Imaging algorithms to estimate PSD from sand samples.	*** - **** rate=86-91%	Urbanski et al. (2011)
Optical VIS-nir-swir	C&D waste	Imaging algorithms and regression analysis to estimate PSD from sands and rocks.	** - *** rate=79-85%	Salehizadeh and Sadeghi (2010)
Optical VIS-nir-swir	MSW, ash	Image analysis to estimate PSD.	n/a	Nakamura et al. (2005)
Optical VIS-nir-swir	RDF	Machine vision to estimate PSD.	n/a	Immonen (2015)
Optical VIS-nir-swir	SRF	Image analysis to estimate PSD.	n/a	Dunnu et al. (2006)

2.3.1.3 Calorific value

The calorific value an important quality attribute of the produced SRF as it indicates the amount of recoverable energy from waste. Optical technologies have been applied to the determination of gross calorific value (GCV) with fair to excellent results. The indirect determination of GCV based on composition analysis using visible images reached excellent qualitative ratings, matching the best results obtained by more complex systems using visible and infrared hyperspectral analysis (Table 2-5). Spectral analysis reveals that a visible to SWIR spectrum (400-2500 nm) achieves good qualitative ratings, but limiting the bandwidth to NIR and SWIR regions (880-2500 nm) improves the estimates with excellent qualitative ratings. However, those studies which ignored the visible spectrum and were also part of the SWIR regions, obtained only fair or good qualitative ratings, which would indicate that the 1720-2500 nm region of the SWIR spectrum can reveal uniquely important information about GCV.

Table 2-5. Sensor technologies for the determination of calorific value. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Optical VIS-nir-swir	MSW	Estimates GCV from visible image analysis of composition. Requires material database and human intervention.	***** R2=0.99	Wagland et al. (2013)
Optical VIS-NIR-swir (450-950 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	***** R2=0.97,RER=16.7	Everard et al. (2012b)
Optical VIS-NIR-SWIR (400-2498 nm)	biomass	Hyperspectral imaging with PLS regression analysis on straw samples.	**** R2=0.89,RER=15.2	Huang et al. (2008a)
Optical VIS-NIR-SWIR	biomass	Hyperspectral imaging with PLS regression analysis on rice straw samples.	*** R2=0.85,RER=13.4	Huang et al. (2008b)

Technology	Target	Methodology	Performance	Author
(400-2500 nm)			RPD=2.56	
Optical vis-NIR-SWIR (880-1680 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	***** R2=0.97,RER=16.7	Everard et al. (2012b)
Optical vis-NIR-SWIR (1100-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	***** R2=0.99,RER=39.9 RPD=70.8	Fagan et al. (2011)
Optical vis-NIR-SWIR (880-1720 nm)	biomass	Hyperspectral imaging with PLS regression analysis on energy crops pellet samples.	** R2=0.73,RER=3.1 RPD=1.0	Gillespie et al. (2015)
Optical vis-NIR-SWIR (780-2498 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	**** RER=16.5	Lestander and Rhén (2005)

Technology	Target	Methodology	Performance	Author
Optical vis-NIR-SWIR	SRF	Industry sponsored project. Indirect determination from composition analysis.	***** error<5%	Krämer and Flamme (2015)
Optical vis-nir-SWIR (1400-2400 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	*** RER=14.0	Lestander and Rhén (2005)

2.3.1.4 Ash content

The content of ashes is an important quality attribute of the produced SRF as it reduces the amount of recoverable energy from waste and places a burden on the management of residues. Optical technologies using hyperspectral image analysis have been applied to the determination of ash contents in biomass with poor to excellent qualitative ratings. Studies obtained excellent qualitative ratings using the complete NIR and SWIR regions, compared to methods limited to visible and NIR regions producing poor qualitative ratings (Table 2-6). This may indicate that visible or NIR regions do not contain enough information regarding ash contents and a broader spectral analysis is required. Moreover, two studies using both NIR-SWIR region produced only fair qualitative ratings, which may be due to ignoring relevant parts of the NIR region (700-1000 nm) and SWIR region (1720-2500 nm). However, one case study was limited to only part of NIR and SWIR (1400-2400 nm) and still obtained excellent qualitative ratings when analysing woody samples, which implies that different materials may require specific spectrums.

Table 2-6. Sensor technologies for the determination of ash contents. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Optical VIS-NIR-swir (450-950 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	* R2 =0.52,RER=6.7 RPD=1.4	Everard et al. (2012)
Optical VIS-NIR-SWIR (400-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on rice straw samples.	**** R2=0.93,RER=20.1 RPD=5.0	Huang et al. (2008)
Optical VIS-NIR-SWIR (350-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	***** R2 =0.99,RER=10.5	Labbé et al. (2008)
Optical vis-NIR-swir (750-1100 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	* R2=0.58,RER=7.7 RPD=3.6	Fagan et al. (2011)
Optical vis-NIR-SWIR (780-2498 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	***** RER=25.9	Lestander and Rhén (2005)
Optical vis-NIR-SWIR (880-1720 nm)	biomass	Hyperspectral imaging with PLS regression analysis on energy crops pellet samples.	** R2=0.75,RER=6.6 RPD=1.8	Gillespie et al. (2015)

Technology	Target	Methodology	Performance	Author
Optical vis-NIR-SWIR (1000-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on rice straw samples.	** R2 =0.86,RER=1.2	Jin and Chen (2007)
Optical vis-NIR-SWIR (1400-2400 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	***** RER=20.9	Lestander and Rhén (2005)
Optical vis-NIR-SWIR	SRF	Industry sponsored project. Indirect determination from composition analysis.	**** error<10%	Krämer and Flamme (2015)

2.3.1.5 Moisture

The content of moisture can affect the performance of air separation, ballistic separation, magnetic separation, eddy current, and biodrying processes. It is also an important fuel quality indicator of SRF, as it determines the net calorific value of the material (British Standards Institution, 2011c). Different real-time technologies have been investigated to determine moisture in different materials, including capacitive and resistive, microwave and optical sensors. Capacitive sensors produced a range of poor to excellent qualitative ratings, with the best cases obtained when tested on MSW samples (Table 2-7). Microwave sensors obtained fair to excellent qualitative ratings, with frequencies ranging from 9 to 26 GHz when testing on either biomass or MSW samples. The use of a resistive sensor was investigated but found the method limited to cases with 35% moisture or more.

Optical technologies were also investigated to estimate moisture in biomass and waste samples, with qualitative ratings ranging from poor to excellent. Hyperspectral analysis resulted in good to excellent qualitative ratings, and the frequency band seems to be affected by the studied material (Table 2-7). Investigations on moisture in biomass, wood and bioenergy crops such as Miscanthus, using different spectrum regions (visible, NIR, SWIR) all resulted in excellent qualitative ratings. One case limited the method to multispectral analysis of 3 bands in the SWIR region (1740-2145 nm) and obtained excellent qualitative ratings with pharmaceutical pellet samples. This also implies that the choice of frequencies is influenced by the sampled material. Hyperspectral studies performed on rice straw samples obtained only fair qualitative ratings when limited to NIR and SWIR (1000-2500 nm), but very good qualitative ratings when including visible wavelengths (400-2500 nm).

Table 2-7. Sensor technologies for the determination of moisture contents. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Capacitive	biomass	Compares 4 sensors on 98 types of wood and grass samples.	* - **** R2 =0.31-0.94	Jensen et al. (Jensen et al., 2006)
Capacitive	MSW	Unaffected by flow rate, presence of metals or saline contamination.	***** R2 =0.96 error<3.5%	Fuchs et al.(2008)
Microwave	biomass	Uses 9GHz and 12GHz sensors and linear regression analysis on wood sawdust samples.	** - **** RMSE=9.0-22.8%	Zhang and Ogura (Zhang and Ogura, 2010)
Microwave	biomass	Uses a 10GHz sensor and linear regression analysis on peanut hull pellets samples	***** R2 =0.97,RER=19.3	Trabelsi et al. (Trabelsi et al., 2013)
Microwave	concrete and masonry	Sensor embedded in solid walls, using 2.2 to 2.7 GHz signals, not tested on loose materials.	***** error<1%	Sokoll and Jacob (2007)
Microwave	MSW	Sensor tested at 5, 16, and 26 GHz. Cannot detect thin water layers (<0.05mm).	**** Error=3-17%	Wagland et al.(2013)
Optical VIS-NIR-SWIR (400-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on rice straw samples.	**** R2=0.93,RER=19.6 RPD=3.37	Huang et al. (2008)
Optical VIS-NIR-swir (450-950 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	***** R2 =0.96,RER=16.0 RPD=4.9	Everard et al. (2012)

Technology	Target	Methodology	Performance	Author
Optical vis-NIR-SWIR (780-2498 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	***** RER=57.1	Lestander and Rhén (2005)
Optical vis-NIR-SWIR (880-1720 nm)	biomass	Hyperspectral imaging with PLS regression analysis on energy crops pellet samples.	*** R2=0.88,RER=18.1 RPD=4.26	Gillespie et al. (2015)
Optical vis-NIR-SWIR (950-1664 nm)	food	Hyperspectral imaging with PLS regression analysis on coffee, soybeans and wafers samples.	**** - ****** R2=0.95-0.99 RMSEP=0.1-0.3	Achata et al.(2013)
Optical vis-NIR-SWIR (1000-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on rice straw samples.	*** R2 =0.89,RER=5.5	Jin and Chen (2007)
Optical vis-NIR-SWIR (1100-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	***** R2=0.99,RER=39.3 RPD=13.5	Fagan et al. (2011)
Optical vis-NIR-SWIR (1400-2400 nm)	biomass	Hyperspectral imaging with PLS regression analysis on wood samples.	***** RER=51.4	Lestander and Rhén (2005)
Optical vis-NIR-SWIR	SRF	Industry sponsored project.	***** error<5%	Krämer and Flamme (2015)
Optical vis-NIR-SWIR	MSW	Industry sponsored project.	**** error<6%	Wünsch and Jenkins (2013)
Optical vis-nir-SWIR (1740-2145 nm)	pharma. pellets	Multispectral imaging and linear regression analysis using 3 bands.	***** R2 =0.98	Rantanen et al.(1998)
Resistive	MSW	Limited to solid medium with moisture levels > 35%.	n/a	Gawande et al. (2003)

2.3.1.6 Composition

Knowledge of the variable composition of a mixed waste stream could affect the performance of manual separation, biodrying and anaerobic digestion processes, whilst being an important consideration towards the fuel quality of SRF. Different technologies have been investigated to identify materials in specific mixed waste streams. Hydrophobicity methods mechanisms for mixed plastics achieved good qualitative ratings, but they involve submerging the materials in a wet medium, the same with one case study using an ultrasound sensor (Table 2-8). Fluorescence, ultrasound and stiffness sensors can be utilised to separate mixed papers, with excellent qualitative ratings in the case of fluorescence but only when limited to papers printed without coloured inks. Similarly, identifying gold ore from mineral waste with X-ray technology resulted in good qualitative ratings. Good to excellent qualitative ratings were obtained using LIBS and PLS-DA analysis for quick classification of mixed metal alloys.

Optical technologies were also applied to determine composition of mixed waste material, including C&D and MSW. The results are affected by the studied material and the spectral region and image processing algorithms applied. For example, one case obtained good qualitative ratings classifying metals but another one provided excellent qualitative ratings of the full composition of C&D waste materials, and both used only visible imaging (Table 2-8). Hyperspectral analysis of WEEE using NIR and SWIR spectrum (1100-2100 nm) produced good qualitative ratings in identifying plastic contaminants, but another case obtained excellent qualitative ratings when extending the spectrum to include visible and NIR wavelengths (348-1008 nm) to identify the metal composition.

Good to very good qualitative ratings were achieved when estimating the composition of mixed waste using visible images of the waste materials, which indicates that the information in visible wavelength is enough to determine material composition. Multispectral analysis of NIR and SWIR images (1100-2500 nm) to identify plastics in MSW produced good qualitative ratings, which were improved upon when limiting the spectrum to SWIR (1548-2550 nm), obtaining excellent qualitative ratings. This implies that the NIR region may

obscure the information in the SWIR region about the presence of plastics. Experiments with MIR and LIR regions of the spectrum (2500-25000 nm) resulted in good qualitative ratings based on the identification rates of types of wastes, and excellent qualitative ratings when identifying types of plastics without being affected by their colours.

Table 2-8. Sensor technologies for composition analysis of mixed materials. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Fluorescence	Mixed papers	Measures fluorescence of lignin to separate 5 grades of paper. It is affected by colour inks.	***** rate=100%	Ramasubramanian et al. (2005)
Hydrophobicity	Mixed plastics	Review of flotation mechanisms to separate PVC or PET.	**** rates=60-100%, purities=70-100%	Wang et al. (2015)
LIBS VIS-nir-swir (242-548nm)	Mixed metal alloys	Classification of 10 types of metals at a 25 measurements per second rate.	*** Average rate>88%	Merk (2015)
Optical VIS-nir-swir	MSW	Full composition through image analysis. Requires material database and human intervention.	**** R2 =0.91	Wagland et al. (2012)
Optical VIS-nir-swir	SRF	Full composition with image analysis. Requires material database and human intervention.	*** R2 =0.89	Peddireddy et al. (2015)
Optical VIS-nir-swir	Mixed metals	Identifies stainless steel, brass and copper using visible imaging based on colour.	*** rate=80%	Kutila et al. (2005)
Optical VIS-nir-swir	Mixed plastics	Uses imaging and histogram-based algorithms to classify PET and Polycoat containers.	***** rate=94-98%	Nawrocky et al. (2010)

Technology	Target	Methodology	Performance	Author
Optical VIS-NIR-swir	Mixed textiles	Review of commercial devices to identify textiles.	**** rate>90%	Ishfaq (2015)
Optical VIS-NIR-swir (384-1008 nm)	WEEE	Hyperspectral image analysis to identify 6 non-ferrous metals.	***** rate=98%	Picon et al. (2009)
Optical vis-NIR-SWIR (900-2500 nm)	Mixed plastics	Hyperspectral imaging analysis to separates PET from bio-plastics.	n/a	Hollstein and Wohllebe (2015)
Optical vis-NIR-SWIR (1100-2100 nm)	WEEE	Hyperspectral image analysis to identifies plastics.	*** rate=80%, purity=70-80%	Beigbeder et al. (2013)
Optical vis-NIR-SWIR (1100-2550 nm)	MSW	Multispectral imaging analysis (6 bands) to identify plastics.	*** rate>80%	Wienke et al. (1996)
Optical vis-NIR-SWIR	SRF	Industry sponsored project. Full composition analysis, requires material database.	n/a	Krämer and Flamme (2015)
Optical vis-nir-SWIR (1548-2550 nm)	MSW	Multispectral imaging analysis (6 bands) to identify plastics.	***** rate>95% error<1%	van den Broek et al. (1996)
Optical vis-nir-swir-MIR-LIR (2500-16667 nm)	MSW, compost, landfill	Hyperspectral imaging analysis to identify 6 types of plastics, including black ones.	***** rate=100% (blacks=100%)	Kassouf et al. (2014)

Technology	Target	Methodology	Performance	Author
Optical vis-nir-swir-MIR-LIR (2500-25000 nm)	MSW, compost, landfill	Hyperspectral imaging analysis to classify waste into 3 groups.	** rate=71%	Smidt et al. (2008)
Stiffness	Mixed papers	Separates 3 grades of paper. It can confuse a single sheet of paper with cardboard.	n/a	Ramasubramanian et al. (2007, 2006)
Ultrasound	MSW	Identifies and removes paper materials. Combined with an optical sensor.	**** rate=90%, purity=100%	Faibish et al. (1997)
Ultrasound	MSW	Identifies polymers in a wet medium.	n/a	Sanaee and Bakker (2009)
X-ray	Minerals	Uses XRT to detect gold ore.	*** rate=84%	Knapp et al. (2014)

2.3.1.7 Biogenic carbon and biochemical methane potential

The contents of biogenic carbon are an important fuel qualifier as it determines the potential fraction of renewable energy in SRF, and under relevant schemes could result in financial incentives (WRAP, 2012b). Biochemical methane potential (BMP) is a related measure which estimates the maximum amount of biogas to be produced by anaerobic digestion processes. Optical methods have been investigated to measure both of these parameters, with poor to excellent qualitative ratings (Table 2-9). The analysis of biomass samples using hyperspectral imaging in the NIR-SWIR range (1000-2500 nm) produced good qualitative ratings, but another experiment on rice straw samples which included the visible range (400-2500 nm) produced poor qualitative ratings. This might indicate that specific ranges are required for different samples, or that the visible range has to be ignored. Another case study with biomass samples but more limited spectral range (880-1720 nm) also produced poor qualitative ratings, which may indicate that the end of the SWIR range is required. However, a simple method using visible images and human interaction to determine composition and indirectly estimate biogenic fraction produced the best results with very good qualitative ratings.

Table 2-9. Sensor technologies for the determination of biogenic carbon contents and determination of BMP. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
Optical VIS-nir-swir	MSW	Estimates biogenic fraction from visible image analysis of composition. Requires material database and human intervention.	**** error=5%	Wagland et al. (2013)
Optical VIS-NIR-SWIR (400-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on rice straw samples.	* R2=0.67, RER=13.9 RPD=1.7	Huang et al. (2008)
Optical vis-NIR-SWIR	MSW, others	Review of infrared methods to determine biochemical methane potential.	* - *** R2 =0.26-0.88 RPD=1.49-2.88	Ward (2016)
Optical vis-NIR-SWIR (880-1720 nm)	biomass	Hyperspectral imaging with PLS regression analysis on energy crops pellet samples.	* R2=0.69, RER=6.6 RPD=1.8	Gillespie et al. (2015)
Optical vis-NIR-SWIR (1000-2500 nm)	MSW and others	Hyperspectral imaging with PLS regression analysis.	*** R2=0.85 RPD=1.8	Doublet et al. (2013)
Optical vis-NIR-SWIR (1100-2500 nm)	biomass	Hyperspectral imaging with PLS regression analysis on Miscanthus and short rotational coppice willow samples.	*** R2=0.88, RER=10.4 RPD=4.6	Fagan et al. (2011)

2.3.1.9 Contaminants

The detection of contaminants in the form of chemicals or unexpected materials is an important indicator which determines the suitability of the waste streams and the fuel quality of the produced SRF. Optical, X-ray and LIBS methodologies have been used to determine the presence of particular contaminants. For example, when determining chemical preservatives in wood samples, X-ray produced very good to excellent qualitative ratings, and optical hyperspectral analysis in the NIR and SWIR range (1162-2197 nm) produced excellent qualitative ratings (Table 2-10). With the sample application, LIBS produced better results when analysing plasma in a narrow band in the visible spectrum (330-510 nm) than when including at both visible and NIR (200-800 nm). The presence of particular chemicals in samples, such as chlorine and strontium, were determined with LIBS and analysis of visible spectra, both with excellent qualitative ratings.

Table 2-10. Sensor technologies for the identification of specific contaminants in mixed materials. Performance rating (Malley et al., 2002, 2004; Nduwamungu et al., 2009; Williams, 2001; Williams et al., 2006): * (poor), ** (fair), * (good), **** (very good), ***** (excellent).**

Technology	Target	Methodology	Performance	Author
LIBS VIS-nir-swir (400-461nm)	Mixed glass, soil	Detects presence of Strontium.	***** detection limit = 16-30ppm	Barefield et al. (1995)
LIBS VIS-nir-swir (330-510 nm)	Mixed wood	Detects presence of treated wood by identifying Chromium-based contaminants.	**** - ***** rate=92-100%	Moskal and Hahn (2002)
LIBS VIS-nir-swir (330-510 nm)	Mixed wood	Detects presence of treated wood by detecting Chromium-based contaminants.	***** rate=98-100%	Solo-Gabriele et al. (2004)
LIBS VIS-NIR-swir (200-800 nm)	Mixed wood	Detects chemical-based preservatives.	*** - **** R2 =0.77-0.94	Martin et al. (2005)
LIBS vis-NIR-swir (837.6 nm)	C&I waste	Detects presence of chlorine in polymers.	***** detection limit = 18% (w/w)	Huber et al. (2014)
Optical vis-nir-SWIR-MIR (2280-4480 nm)	Mixed glass	Hyperspectral imaging and regression analysis to identifies ceramic glass contaminants.	**** rate=92-98%	Serranti et al. (2006)

Technology	Target	Methodology	Performance	Author
Optical vis-NIR-SWIR (1162-2197 nm).	Mixed wood	Identifies chemical preservatives and 4 types of wood plastic composites.	***** rate=97-99%	Mauruschat et al. (2015)
Optical vis-NIR-SWIR	SRF	Industry sponsored project. Estimates chlorine contents based on composition analysis. It requires material database.	*** error<20%	Krämer and Flamme (2015)
X-ray	Mixed plastics	Uses dual energy XRT to detect presence of brominated flame retardants in polymers.	n/a	Firsching et al. (2013)
X-ray	Mixed wood	Detects presence of treated wood by using XRF to detect inorganic preservatives.	**** rate=85-100%	Rasem Hasan et al. (2011)
X-ray	Mixed wood	Detects presence of treated wood by using XRF to detect Arsenic-based contaminants.	***** rate=100%	Solo-Gabriele et al. (2004)

2.3.2 Process monitoring tools

Eight quality indicators have been identified in the scope of an MRF, which would determine the fuel quality of the produced SRF. Selecting the best tools for each of these indicators shows that optical technologies can be applied to measuring all of them (Table 2-11). Optical methods using visible light cameras combined with computer vision algorithms could measure five of the eight indicators, and the rest would require infrared spectral imaging and mathematical analysis. The visible image processing algorithms would need to be tested waste processing on industrial scale, in particular with dark or black objects, which present a problem to spectral imaging analysis. Alternatively, capacitive and microwaves sensors could be used to determine moisture contents, and LIBS could detect and quantify some contaminants. The presence of chlorine contaminants in waste materials has been investigated but further research is required to monitor presence of mercury and heavy metals in mixed solid waste.

When applied to an MRF, every mechanical and biological process could make use of critical data measured by optical technology using visible sensors (Table 2-12), with some of them requiring extra information which could be provided by either infrared spectroscopy, capacitive or resistive sensors. Problematic objects recognition technologies could offer information regarding presence of hazardous materials, which could trigger an automatic stop of the plant processes until the dangerous item is removed. Multiple object recognition sensors could be deployed in one single plant, in particular before each shredder, looking for specific items, sizes and shapes, according to their position in the process. The identified contaminants or unsafe objects could be sent through alternate treatment routes to be processed as required, reducing risks and potential fines.

Table 2-11. Best performing technologies for each indicator when applied to waste materials.

Indicator	Technologies	Notes
Problematic objects	Optical (visible)	Different algorithms needed depending on location in the process. More tests needed on mixed waste.
Particle size distribution	Optical (visible)	Different algorithms needed depending on location in the process. More tests needed on mixed waste.
Calorific value	Optical (visible)	Needs algorithm development of automatic material recognition on mixed waste. Indirect determination needs database of material properties.
	Optical hyperspectral (800-2500 nm)	Cannot measure black/dark objects. Indirect determination needs database of material properties. Direct determination method needs testing on MSW.
Ash contents	Optical hyperspectral (800-2500 nm)	Cannot measure black/dark objects.
Moisture	Capacitive	Need testing with MSW in industrial scale.
	Microwave	Need testing with MSW in industrial scale.
	Optical hyperspectral (800-2500 nm)	Cannot measure black/dark objects.
Composition	Optical (visible)	Automated method on subset of C&D waste. MSW method needs algorithm development of automatic material recognition. Indirect determination needs database of material properties.
	Optical hyperspectral (800-2500 nm)	Cannot measure black/dark objects.
Biogenic carbon and BMP	Optical (visible)	Needs algorithm development of automatic material recognition on mixed waste. Indirect determination needs database of material properties.
Contaminants	LIBS	Identifies PVC, but quantification of Chlorine contents not in real-time. Can detect preservatives in wood samples.
	Optical hyperspectral (800-2500 nm)	Cannot measure black/dark objects. Indirect estimate of Chlorine based on PVC fraction requires database of material properties.

Particle size distribution sensors could be deployed together with moisture and composition analysis sensors before any stage using separation based on physical properties, such as screening, air and ballistic separators. Gaining real-time information of moisture, composition and PSD could improve the separation processes which would adjust their operational parameters accordingly. Particle size distribution data is required by European standards (EU CEN, 2012) to inform stakeholders about the quality of solid waste streams and waste derived fuels.

Calorific value, ash contents and moisture sensors could be installed at different stages of MRF, allowing the processes to route and possibly mix waste streams of different composition to obtain the desired WDF quality. Mixing wastes with different measured properties, for example a stream with low CV with a stream of high CV could allow an operator to produce the required WDF with the highest possible efficiency by controlling the mix fractions in real-time. Early detection of contaminants such as chlorine could be used to prevent them from going into WDF, recycling or landfill, enabling better material recovery and reducing environmental pollution.

Significantly, current techniques are designed and implemented for, typically, a single application (i.e., detection of specific plastic types or removal of metals). Each technology works in isolation, with little system-level data sharing, validation, and reconciliation. Therefore, no single technology or system provides the operator with all the parameters required, thus further opportunities exist in developing technologies capable of informing and impacting material recovery and fuel production in real time.

Table 2-12. Critical quality indicators and the available monitoring tools for the different waste process units.

Process	Critical quality indicators	Available monitoring tools
Bag splitting	Problematic objects - Presence of large or hard items in the input stream	Optical (visible)
Shredding	Problematic objects - Presence of large or hard items in the input stream	Optical (visible)
Screening	Problematic objects - Presence of highly ductile material which end up in fines fraction	Optical (visible)
	Particle size distribution (PSD)	Optical (visible)
Air separation	Moisture	Optical (infrared hyperspectral) Capacitive Microwave
	Particle size distribution (PSD)	Optical (visible)
Ballistic separation	Moisture	Optical (infrared hyperspectral) Capacitive Microwave
	Particle size distribution (PSD)	Optical (visible)
Magnetic separation	Particle size distribution (PSD)	Optical (visible)
	Moisture	Optical (infrared hyperspectral) Capacitive Microwave
	Particle size distribution (PSD)	Optical (visible)

Process	Critical quality indicators	Available monitoring tools
Eddy current separation	Moisture	Optical (infrared hyperspectral) Capacitive Microwave
	Composition - presence of aggregated materials	Optical (visible) Optical (infrared hyperspectral)
Biodrying	Moisture	Optical (infrared hyperspectral) Capacitive Microwave
	Composition - presence of aggregated materials and organic material	Optical (visible) Optical (infrared hyperspectral)
Anaerobic digestion	Particle size distribution (PSD)	Optical (visible)
	Composition - organic material	Optical (visible) Optical (infrared hyperspectral)
	Biochemical methane potential (BMP)	Optical (visible)

2.4 Conclusions and perspectives

Waste management services currently present a situation similar to that of the pharmaceutical industry over a decade ago, whereby quality was assessed by collecting samples and then assessed using laboratory testing. Existing technology could be used to improve waste derived fuel production processes, by developing better real-time analysis and control tools for material recovery facilities.

This article answers two key questions:

a) which information, or critical quality attributes, would need to be collected for the implementation of PAT in the waste management industry; and
b) which sensor technologies, or process monitoring tools, could provide this information in real-time to improve waste material treatment processes.

Each of the waste processes used in MRFs could benefit from implementing PAT and gaining information about the waste stream. Using real-time quality measurements to adjust the operational parameters of MBT processes could lead to a consistently improved quality of waste derived fuel. This in turn reduces uncertainty due to fluctuations from input waste streams, increases the energy and material recovered and reduces landfill waste disposal.

Optical technologies, in particular those using visible sensors, offer the most flexibility as they can be used to measure most of the critical quality indicators without requiring complementing technologies. Some attributes such as the presence of mercury and other heavy metals, are still not developed for mixed solid materials and could represent a challenge. However, indirect determination of fuel quality properties based on composition analysis and a database of material properties database offer a viable alternative, as they produce similar or better results than direct determination analysis. Near-infrared spectrographic analysis technologies are limited by working distances, particle sizes, and the presence of undetectable materials such as dark objects, metals, and glass, which do not necessarily affect visible-sensor methods. Another potential challenge for both NIR spectroscopy and visible camera methods are the effects

of dirtiness and the degradation of material surfaces, which could cause false-positive results.

The integration of existing applicable technologies, used so far in isolation or for very specific reasons, in the framework of process analytical technologies to provide real-time waste information could improve the overall efficiency of the waste treatment industry. The generated waste data could also provide needed information to stakeholders, such as governments, local authorities, regulators and waste industry operators, for better decision-making in the future.

2.5 References

Achata, E., Esquerre, C., O'Donnell, C., Gowen, A., 2013. Use of near infrared spectral imaging and chemometrics for prediction of moisture content and water activity in different low moisture food systems., in: Bellon-Maurel V., Williams P., Downey G. (Eds.), NIR 2013 Proceedings of 16th International Conference on Near Infrared Spectroscopy: Picking up Good Vibrations, ST5 – Hyperspectral Imaging. La Grande - Motte, France, pp. 713–717.

Ammineni, C., 2001. Design of Lignin Sensor for Identification of Paper Grades for an Automatic Waste Paper Sorting System (MSc thesis). North Carolina State University.

Anding, K., Garten, D., Linß, E., 2013. Application of intelligent image processing in the construction material industry. Acta IMEKO 2, 61–73.

Anding, K., Kuritcyn, P., Linß, E., Latyev, S., 2015. Significant Characteristics in VIS- and IR-Spectrum of CDW for High-Precision Supervised Classification, in: Beyerer, J., Puente Leon, F., Langle, T. (Eds.), OCM 2015 - Optical Characterization of Materials - Conference Proceedings. KIT Scientific Publishing, Karlsruhe, Germany, pp. 161–168.

Anding, K., Linß, E., Träger, H., Rückwardt, M., Göpfert, A., 2011. Optical identification of construction and demolition waste by using image processing and machine learning methods, in: 14th Joint International IMEKO TC1, TC7,

TC13 Symposium on Intelligent Quality Measurements. Jena, Germany, pp. 126–132.

Ángel Aguirre, M., Hidalgo, M., Canals, A., Nóbrega, J.A., Pereira-Filho, E.R., 2013. Analysis of waste electrical and electronic equipment (WEEE) using laser induced breakdown spectroscopy (LIBS) and multivariate analysis. Talanta 117, 419–424. doi:10.1016/j.talanta.2013.09.046

Anzano, J., Bonilla, B., Montull-Ibor, B., Casas-González, J., 2011. Plastic identification and comparison by multivariate techniques with laser-induced breakdown spectroscopy. J. Appl. Polym. Sci. 121, 2710–2716. doi:10.1002/app.33801

Banaee, M., Tavassoli, S.H., 2012. Discrimination of polymers by laser induced breakdown spectroscopy together with the DFA method. Polym. Test. 31, 759–764. doi:10.1016/j.polymertesting.2012.04.010

Barefield, J.E., Koskelo, A.C., Multari, R.A., Cremers, D.A., Gamble, T.K., Han, C.Y., 1995. Laser-induced breakdown spectroscopy for the real-time analysis of mixed waste samples containing Sr. Los Alamos National Lab., NM (United States).

Baum, G.A., 1987. Elastic properties, paper quality, and process control. Appita 40, 289–294.

Baytekin, B., Baytekin, H.T., Grzybowski, B.A., 2013. Retrieving and converting energy from polymers: deployable technologies and emerging concepts. Energy Environ. Sci. 6, 3467–3482. doi:10.1039/C3EE41360H

Beigbeder, J., Perrin, D., Mascaro, J.-F., Lopez-Cuesta, J.-M., 2013. Study of the physico-chemical properties of recycled polymers from waste electrical and electronic equipment (WEEE) sorted by high resolution near infrared devices. Resour. Conserv. Recycl. 78, 105–114. doi:10.1016/j.resconrec.2013.07.006

Bianconi, F., Maria, F.D., Micale, C., Fernández, A., Harvey, R.W., 2015. Grain-size assessment of fine and coarse aggregates through bipolar area morphology. Mach. Vis. Appl. 26, 775–789. doi:10.1007/s00138-015-0692-z

Blassino, M., Solo-Gabriele, H., Townsend, T., 2002. Pilot scale evaluation of sorting technologies for CCA treated wood waste. Waste Manag. Res. 20, 290–301.

Block, C.N., Shibata, T., Solo-Gabriele, H.M., Townsend, T.G., 2007. Use of handheld X-ray fluorescence spectrometry units for identification of arsenic in treated wood. Environ. Pollut. 148, 627–633. doi:10.1016/j.envpol.2006.11.013

Bombardier, V., Schmitt, E., 2010. Fuzzy rule classifier: Capability for generalization in wood color recognition. Eng. Appl. Artif. Intell. 23, 978–988. doi:10.1016/j.engappai.2010.05.001

Bonifazi, G., Serranti, S., 2014. Quality control by Hyper Spectral Imaging (HSI) in solid waste recycling : logics , algorithms and procedures, in: Niel, K.S., Bingham, P.R. (Eds.), Proceedings - SPIE The International Society for Optical Engineering, Image Processing: Machine Vision Applications VII. Bellingham, SPIE, pp. 1–15. doi:10.1117/12.2038374

British Standards Institution, 2012. BS EN 15415-3:2012 Solid recovered fuels — Determination of particle size distribution. Method by image analysis for large dimension particles. BSI, London.

British Standards Institution, 2011a. BS EN 15442:2011 Solid recovered fuels — Methods for sampling. London, UK.

British Standards Institution, 2011b. BS EN 15403:2011 Solid recovered fuels — Methods for the determination of Calorific Value. London, UK.

British Standards Institution, 2011c. BS EN 15414-3:2011 Solid recovered fuels — Determination of moisture content using the oven dry method. Part 3: Moisture in general analysis sample. London, UK.

British Standards Institution, 2011d. BS EN 15359:2011 Solid recovered fuels — specification and classes. BSI, London.

Chadwick, D.T., McDonnell, K.P., Brennan, L.P., Fagan, C.C., Everard, C.D., 2014. Evaluation of infrared techniques for the assessment of biomass and

biofuel quality parameters and conversion technology processes: A review. Renew. Sustain. Energy Rev. 30, 672–681. doi:10.1016/j.rser.2013.11.006

Chen, Z., Lovett, D., Morris, J., 2011. Process analytical technologies and real time process control a review of some spectroscopic issues and challenges. J. Process Control 21, 1467–1482. doi:10.1016/j.jprocont.2011.06.024

Cimpan, C., Maul, A., Jansen, M., Pretz, T., Wenzel, H., 2015. Central sorting and recovery of MSW recyclable materials: A review of technological state-of-the-art, cases, practice and implications for materials recycling. J. Environ. Manage. 156, 181–199. doi:10.1016/j.jenvman.2015.03.025

Council of the European Union., 2008. Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste and repealing certain Directives. [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:312:0003:0030:en:PDF (accessed 5.21.16)

Council of the European Union, 1999. Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. Official J. Eur. Communities L 182/1, 1–19.

DEFRA, 2013. Mechanical biological treatment of municipal solid waste. [WWW Document]. URL https://www.gov.uk/government/publications/mechanical-biological-treatment-of-municipal-solid-waste (accessed 4.24.16).

Dickens, J.E., 2010. Overview of Process Analysis and PAT, in: Bakeev, K.A. (Ed.), Process Analytical Technology. John Wiley & Sons, Ltd, pp. 1–15.

Du, W., 2014. Resistive, Capacitive, Inductive, and Magnetic Sensor Technologies, Series in Sensors. CRC Press.

European Commission, 2015. Draft guidance document on the definition and classification of hazardous waste. [WWW Document]. URL http://ec.europa.eu/environment/waste/hazardous_index.htm (accessed 5.26.16)

European Commission, 2012. Eco-innovation: Textiles 4 textiles Project. [WWW Document]. URL http://ec.europa.eu/environment/eco-innovation/projects/en/projects/t4t (accessed 2.3.16)

European Committee for Standardization, 2012. Solid Recovered Fuels (Published Standards). [WWW Document]. URL https://standards.cen.eu/dyn/www/f?p=204:32:0::::FSP_ORG_ID,FSP_LANG_ID:407430,25&cs=1E692895E0FA13AE68B9FA01D5A630ED7 (accessed 5.4.16)

European Parliament., 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009. [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:FULL:EN:PDF (accessed 10.19.16)

Faibish, S., Bacakoglu, H., Goldenberg, A.A., 1997. An Eye-Hand System for Automated Paper Recycling+. Proc. 1997 IEEE Int. Conf. Robot. Autom. 9–14. doi:10.1109/ROBOT.1997.620008

Farcomeni, A., Serranti, S., Bonifazi, G., 2008. Non-parametric analysis of infrared spectra for recognition of glass and glass ceramic fragments in recycling plants. Waste Manag. 28, 557–564. doi:10.1016/j.wasman.2007.01.019

Faria, J., Martins, T., Ferreira, M., Santos, C., 2008. A computer vision system for color grading wood boards using Fuzzy Logic, in: 2008 IEEE International Symposium on Industrial Electronics. Presented at the 2008 IEEE International Symposium on Industrial Electronics, pp. 1082–1087. doi:10.1109/ISIE.2008.4677036

FDA, 2009. Q8(R2) Pharmaceutical Development, Guidance for industry. Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, USA. [WWW Document]. URL http://www.fda.gov/downloads/Drugs/.../Guidances/ucm073507.pdf (accessed 5.26.16)

FDA, 2004. PAT — a framework for innovative pharmaceutical development, manufacturing, and quality assurance, Guidance for industry. Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, USA. [WWW Document]. URL http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf (accessed 4.15.16)

Firsching, M., Muhlbauer, J., Maurer, A., Nachtrab, F., Uhlmann, N., 2013. Quantitative sorting using dual energy X-ray transmission imaging, in: OCM 2013 - Optical Characterization of Materials - Conference Proceedings. pp. 259–264. doi:10.5445/KSP/1000032143

Fitzgerald, G.C., Themelis, N.J., 2009. Technical and economic impacts of pre-shredding the MSW feed to moving grate WTE boilers, in: 17th Annual North American Waste-to-Energy Conference. American Society of Mechanical Engineers, pp. 237–243.

Flamme, S., Geiping, J., 2012. Quality standards and requirements for solid recovered fuels: a review. Waste Manag. Res. 30, 335–353. doi:10.1177/0734242X12440481

Fraunholcz, N., 2004. Separation of waste plastics by froth flotation—a review, part I. Miner. Eng., Processing and Disposal of Minerals Industry Waste ’03 17, 261–268. doi:10.1016/j.mineng.2003.10.028

Fuchs, A., Zangl, H., Holler, G., Brasseur, G., 2008. Design and analysis of a capacitive moisture sensor for municipal solid waste. Meas. Sci. Technol. 19, 025201–9. doi:10.1088/0957-0233/19/2/025201

Gawande, N.A., Reinhart, D.R., Thomas, P.A., McCreanor, P.T., Townsend, T.G., 2003. Municipal solid waste in situ moisture content measurement using an electrical resistance sensor. Waste Manag. 23, 667–674. doi:10.1016/S0956-053X(03)00100-4

Gomes, J., Chopda, V.R., Rathore, A.S., 2015. Integrating systems analysis and control for implementing process analytical technology in bioprocess

development. J. Chem. Technol. Biotechnol. 90, 583–589. doi:10.1002/jctb.4591

Gondal, M.A., Siddiqui, M.N., 2007. Identification of different kinds of plastics using laser-induced breakdown spectroscopy for waste management. J. Environ. Sci. Health Part A 42, 1989–1997. doi:10.1080/10934520701628973

Guerrero, L.A., Maas, G., Hogland, W., 2013. Solid waste management challenges for cities in developing countries. Waste Manag. 33, 220–232. doi:10.1016/j.wasman.2012.09.008

Gundupalli, S.P., Hait, S., Thakur, A., 2016. A review on automated sorting of source-separated municipal solid waste for recycling. Waste Manag. doi:10.1016/j.wasman.2016.09.015

Hannan, M.A., Al Mamun, M.A., Hussain, A., Basri, H., Begum, R.A., 2015. A review on technologies and their usage in solid waste monitoring and management systems: Issues and challenges. Waste Manag. 43, 509–523. doi:10.1016/j.wasman.2015.05.033

Hollstein, F., Wohllebe, M., Manjon, D., 2015. Identification of bio-plastics by NIR-SWIR-Hyperspectral-Imaging, in: Beyerer, J., Puente Leon, F., Langle, T. (Eds.), OCM 2015 - Optical Characterization of Materials - Conference Proceedings. KIT Scientific Publishing, Karlsruhe, Germany, pp. 71–87. https://doi.org/10.5445/KSP/1000044906

Hrćka, R., 2008. Colour modelling as a tool for wood grading, in: Proceedings of COST E53 Quality Control for Wood and Wood Products. Presented at the COST E53 Quality control for wood and wood products, Delft, The Netherlands, pp. 165–170.

Huang, J., Pretz, T., Bian, Z., 2010. Intelligent solid waste processing using optical sensor based sorting technology. 2010 3rd Int. Congr. Image Signal Process. 4, 1657–1661. doi:10.1109/CISP.2010.5647729

Huber, N., Eschlböck-Fuchs, S., Scherndl, H., Freimund, A., Heitz, J., Pedarnig, J.D., 2014. In-line measurements of chlorine containing polymers in an

industrial waste sorting plant by laser-induced breakdown spectroscopy. Appl. Surf. Sci. 302, 280–285. doi:10.1016/j.apsusc.2013.10.070

Huth-Fehre, T., Feldhoff, R., Kantimm, T., Quick, L., Winter, F., Cammann, K., van den Broek, W., Wienke, D., Melssen, W., Buydens, 1995. NIR-Remote sensing and artificial neural networks for rapid identification of post consumer plastics. J. Mol. Struct. 143–146.

Ishfaq, M., 2015. Infrared technology and its applications in textile recycling technology : improving sustainability in clothing industry (PhD Thesis). Lahti University of Applied Sciences.

Jacobi, G., Solo-Gabriele, H., Townsend, T., Dubey, B., 2007. Evaluation of methods for sorting CCA-treated wood. Waste Manag. 27, 1617–1625. doi:10.1016/j.wasman.2006.09.014

Janesick, J.R., 2007. Photon Transfer. SPIE, 1000 20th Street, Bellingham, WA 98227-0010 USA.

Jeswani, H.K., Azapagic, A., 2016. Assessing the environmental sustainability of energy recovery from municipal solid waste in the UK. Waste Manag. 50, 346–363. doi:10.1016/j.wasman.2016.02.010

Kassouf, A., Maalouly, J., Rutledge, D.N., Chebib, H., Ducruet, V., 2014. Rapid discrimination of plastic packaging materials using MIR spectroscopy coupled with independent components analysis (ICA). Waste Manag. 34, 2131–2138. doi:10.1016/j.wasman.2014.06.015

Katuri, K.C., 2006. Development of online stiffness sensor for high speed sorting of recovered paper (MSc Thesis). North Carolina State University.

Kleber, F., Pramerdorfer, C., Wetzinger, E., Kampel, M., 2015. Optical Sensor Evaluation for Vision Based Recognition of Electronics Waste. J. Environ. Sci. Dev. IJESD 6. doi:10.7763/IJESD.2015.V6.724

Knapp, H., Neubert, K., Schropp, C., Wotruba, H., 2014. Viable Applications of Sensor-Based Sorting for the Processing of Mineral Resources. ChemBioEng Rev. 1, 86–95. doi:10.1002/cben.201400011

Krämer, P., Flamme, S., 2015. Real-time Analysis of Solid Recovered Fuels using Sensor Technology, in: Thomé-Kozmiensky, K.J., Thiel, S. (Eds.), Proceedings of IRRC Waste To Energy. TK Verlag Karl Thomé-Kozmiensky - Neuruppin, pp. 339–348.

Kurdthongmee, W., 2008. Colour classification of rubberwood boards for fingerjoint manufacturing using a SOM neural network and image processing. Comput. Electron. Agric. 64, 85–92. doi:10.1016/j.compag.2008.04.002

Kuritcyn, P., Anding, K., Linß, E., Latyev, S.M., 2015. Increasing the Safety in Recycling of Construction and Demolition Waste by Using Supervised Machine Learning. J. Phys. Conf. Ser. 588. doi:10.1088/1742-6596/588/1/012035

Lesteur, M., Latrille, E., Maurel, V.B., Roger, J.M., Gonzalez, C., Junqua, G., Steyer, J.P., 2011. First step towards a fast analytical method for the determination of Biochemical Methane Potential of solid wastes by near infrared spectroscopy. Bioresour. Technol. 102, 2280–2288. doi:10.1016/j.biortech.2010.10.044

Lesteur, M., Preys, S., Boulanger, A., Latrille, E., Treguer, R., 2013. Fast Prediction of Biochemical Methane Potential (BMP) of Organic Waste by Near Infrared Spectroscopy. Pick. Good Vib. 453.

Li, L., Li, R., Zhang, S., Zhao, T., Chen, Z., 2016. A Dynamic Material Discrimination Algorithm for Dual MV Energy X-ray Digital Radiography☆. Appl. Radiat. Isot. doi:10.1016/j.apradiso.2016.05.018

Lombardi, L., Carnevale, E., Corti, A., 2015. A review of technologies and performances of thermal treatment systems for energy recovery from waste. Waste Manag., Special Thematic Issue: Waste-to-Energy Processes and Technologies 37, 26–44. doi:10.1016/j.wasman.2014.11.010

Lorber, K.E., Sarc, R., Aldrian, A., 2012. Design and quality assurance for solid recovered fuel. Waste Manag. Res. 30, 370–380. doi:10.1177/0734242X12440484

Lu, J.-W., Chang, N.-B., Liao, L., 2013. Environmental Informatics for Solid and Hazardous Waste Management: Advances, Challenges, and Perspectives. Crit. Rev. Environ. Sci. Technol. 43, 1557–1656. doi:10.1080/10643389.2012.671097

Lu, Q., Conners, R.W., Kline, D.E., Araman, P., 1997. A real-time algorithm for color sorting edge-glued panel parts, in: Image Processing, 1997. Proceedings., International Conference on. IEEE, pp. 822–825.

Luiken, A., Bos, P., 2010. Automated identification and sorting of post-consumer textile waste. Melliand Int. 16, 50.

Mallapragada, V., 2004. Online lignin sensor for high-speed sorting of newsprint from mixed waste (MSc thesis). North Carolina State University.

Malley, D.F., Martin, P.D., Ben-Dor, E., 2004. Application in Analysis of Soils, in: Near-Infrared Spectroscopy in Agriculture, Agronomy Monograph. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI, pp. 729–784.

Malley, D.F., McClure, C., Martin, P.D., Firlotte, N., Goldsborough, G., Sheppard, M., 2002. Evaluation of near-infrared spectroscopy as a rapid method for estimating the carbon stored per unit area in a wetland. Final Rep. Manit. Clim. Change Action Fund Proj. 15.

Mantanus, J., Rozet, E., Van Butsele, K., De Bleye, C., Ceccato, A., Evrard, B., Hubert, P., Zi??mons, E., 2011. Near infrared and Raman spectroscopy as Process Analytical Technology tools for the manufacturing of silicone-based drug reservoirs. Anal. Chim. Acta 699, 96–106. doi:10.1016/j.aca.2011.05.006

Mauruschat, D., Plinke, B., Aderhold, J., Gunschera, J., Meinlschmidt, P., Salthammer, T., 2015. Application of near-infrared spectroscopy for the fast

detection and sorting of wood–plastic composites and waste wood treated with wood preservatives. Wood Sci. Technol. doi:10.1007/s00226-015-0785-x

McWhirt, A.L., Weindorf, D.C., Chakraborty, S., Li, B., 2012. Visible near infrared diffuse reflectance spectroscopy (VisNIR DRS) for rapid measurement of organic matter in compost. Waste Manag. Res. 30, 1049–1058. doi:10.1177/0734242X12450601

Merk, S., Scholz, C., Florek, S., Mory, D., 2015. Increased identification rate of scrap metal using Laser Induced Breakdown Spectroscopy Echelle spectra. Spectrochim. Acta Part B At. Spectrosc. 112, 10–15. doi:10.1016/j.sab.2015.07.009

Nakamura, M., Castaldi, M.J., Themelis, N.J., 2005. Measurement of Particle Size and Shape of New York City Municipal Solid Waste and Combustion Residues Using Image Analysis,", in: Proc. 16th Japan Society of Waste Management Experts (JSMWE) Fall Conference. pp. 1–3.

Nasrullah, M., Vainikka, P., Hannula, J., Hurme, M., Kärki, J., 2015. Mass, energy and material balances of SRF production process. Part 3: Solid recovered fuel produced from municipal solid waste. Waste Manag. Res. J. Int. Solid Wastes Public Clean. Assoc. ISWA 33, 146–156.

Nasrullah, M., Vainikka, P., Hannula, J., Hurme, M., Karki, J., 2014a. Mass, energy and material balances of SRF production process. Part 1: SRF produced from commercial and industrial waste. Waste Manag. 34, 1398–1407. doi:10.1016/j.wasman.2014.03.011

Nasrullah, M., Vainikka, P., Hannula, J., Hurme, M., Karki, J., 2014b. Mass, energy and material balances of SRF production process. Part 2: SRF produced from construction and demolition waste. Waste Manag. 34, 2163–2170. doi:10.1016/j.wasman.2014.06.009

Nawrocky, M., Schuurman, D.C., Fortuna, J., 2010. Visual sorting of recyclable goods using a support vector machine. Can. Conf. Electr. Comput. Eng. 4–7. doi:10.1109/CCECE.2010.5575231

Nduwamungu, C., Ziadi, N., Tremblay, G.F., Parent, L.-É., 2009. Near-Infrared Reflectance Spectroscopy Prediction of Soil Properties: Effects of Sample Cups and Preparation. Soil Sci. Soc. Am. J. 73, 1896. doi:10.2136/sssaj2008.0213

Palmieri, R., Bonifazi, G., Serranti, S., 2014. Automatic detection and classification of EOL-concrete and resulting recovered products by hyperspectral imaging, in: Vo-Dinh, T., Lieberman, R.A., Gauglitz, G.G. (Eds.), Proc. SPIE 9106, Advanced Environmental, Chemical, and Biological Sensing Technologies XI. Baltimore, p. 91060D. doi:10.1117/12.2049399

Peddireddy, S., Longhurst, P.J., Wagland, S.T., 2015. Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste Manag. 40, 9–13. doi:10.1016/j.wasman.2015.03.015

Pieber, S., Ragossnig, A., Pomberger, R., Curtis, A., 2012. Biogenic carbon-enriched and pollutant depleted SRF from commercial and pretreated heterogeneous waste generated by NIR sensor-based sorting. Waste Manag. Res. 30, 381–391. doi:10.1177/0734242X12437567

Quinn, B., Kelly, G., 2010. Feasibility of embedded wireless sensors for monitoring of concrete curing and structural health. pp. 764704-764704–12. doi:10.1117/12.847529

Rahman, M.O., Hannan, M.A., Scavino, E., Hussain, A., Basri, H., 2009a. An efficient paper grade identification method for automatic recyclable waste paper sorting. Eur. J. Sci. Res. 25, 96–103.

Rahman, M.O., Hussain, A., Basri, H., 2014. A critical review on waste paper sorting techniques. Int. J. Environ. Sci. Technol. 11, 551–564. doi:10.1007/s13762-013-0222-3

Rahman, M.O., Hussain, A., Basri, N.E.A., Basri, H., Scavino, E., 2012a. Chromaticity based waste paper grade identification. Int Arab J Inf Technol 9, 411–417.

Rahman, M.O., Hussain, A., Basri, N.E.A., Scavino, E., Basri, H., Hannan, M.A., 2010. Waste paper grade identification system using window features. J Comput Inf Syst 6, 2077–2091.

Rahman, M.O., Hussain, A., Scavino, E., Basri, H., Hannan, M.A., 2011. Intelligent computer vision system for segregating recyclable waste papers. Expert Syst. Appl. 38, 10398–10407. doi:10.1016/j.eswa.2011.02.112

Rahman, M.O., Hussain, A., Scavino, E., Hannan, M.A., Basri, H., 2015. DNA computer based algorithm for recyclable waste paper segregation. Appl. Soft Comput. J. 31, 223–240. doi:10.1016/j.asoc.2015.02.042

Rahman, M.O., Hussain, A., Scavino, E., Hannan, M.A., Basri, H., 2012b. Object identification using DNA computing algorithm, in: Evolutionary Computation (CEC), 2012 IEEE Congress on. IEEE, pp. 1–7.

Rahman, M.O., Hussain, A., Scavino, E., Hannan, M.A., Basri, H., 2012c. Real-time waste paper grading using CBR approach. Int J Innov Compuy Inf Control 8, 471–488.

Rahman, M.O., Hussain, A., Scavino, E., Hannan, M.A., Basri, H., 2009b. Segregating recyclable waste papers using co-occurrence features, in: The 9th WSEAS International Conference on Applied Computer Science (ACS'09). pp. 187–191.

Ramasubramanian, M.K., Venditti, R., Katuri, K., 2007. Behavior of paper on a high speed conveyor subjected to air jet impingement: a method for estimating bending stiffness. J. Mech. Mater. Struct. 2, 201–219.

Ramasubramanian, M.K., Venditti, R.A., Ammineni, C., Mallapragada, V., 2005. Optical sensor for noncontact measurement of lignin content in high-speed moving paper surfaces. IEEE Sens. J. 5, 1132–1139. doi:10.1109/JSEN.2005.851007

Ramasubramanian, M.K., Venditti, R.A., Gillella, P.K., 2012. Sensor systems for high speed intelligent sorting of waste paper in recycling. Tappi J. 11, 33–39.

Ramasubramanian, M.K., Venditti, R.A., Katuri, K.C., 2006. Development of bending stiffness sensor for automated sorting of recovered paper. Presented at the 2005 TAPPI Engineering, Pulping, Environmental Conference - Conference Proceedings.

Rantanen, J., Lehtola, S., Rämet, P., Mannermaa, J.P., Yliruusi, J., 1998. On-line monitoring of moisture content in an instrumented fluidized bed granulator with a multi-channel NIR moisture sensor. Powder Technol. 99, 163–170. doi:10.1016/S0032-5910(98)00100-4

Rasem Hasan, A., Schindler, J., Solo-Gabriele, H.M., Townsend, T.G., 2011. Online sorting of recovered wood waste by automated XRF-technology. Part I: Detection of preservative-treated wood waste. Waste Manag. 31, 688–694. doi:10.1016/j.wasman.2010.11.010

Reich, G., 2005. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. Adv. Drug Deliv. Rev. 57, 1109–1143. doi:10.1016/j.addr.2005.01.020

Rosas, J.G., Blanco, M., Gonzalez, J.M., Alcala, M., 2012. Real-time determination of critical quality attributes using near-infrared spectroscopy: A contribution for Process Analytical Technology (PAT). Talanta 97, 163–170. doi:10.1016/j.talanta.2012.04.012

Rotter, V.S., Kost, T., Winkler, J., Bilitewski, B., 2004. Material flow analysis of RDF-production processes. Waste Manag. 24, 1005–1021. doi:10.1016/j.wasman.2004.07.015

RWM Ambassadors, 2016. Waste data in the UK. [WWW Document]. URL https://www.rwmexhibition.com/__media/Ambassador-reports/Waste-Data-Report-RWM-Ambassadors.pdf (accessed 9.7.16)

Saisinchai, S., 2014. Separation of PVC from PET/PVC Mixtures Using Flotation by Calcium Lignosulfonate Depressant. Eng. J. 18, 45–54. doi:10.4186/ej.2014.18.1.45

Salehizadeh, M., Sadeghi, M.T., 2010. Size Distribution Estimation of Stone Fragments via Digital Image Processing, in: Proceedings of the 6th International Conference on Advances in Visual Computing - Volume Part III, ISVC'10. Springer-Verlag, Berlin, Heidelberg, pp. 329–338.

Sanaee, S.A., Bakker, M., 2009. Ultrasound For Monitoring And Quality Inspection In MDS Plastics Recycling, in: ISWA World Congress. Lisbon.

Sarc, R., Lorber, K., Pomberger, R., Rogetzer, M., Sipple, E., 2014. Design, quality, and quality assurance of solid recovered fuels for the substitution of fossil feedstock in the cement industry. Waste Manag. Res. 32, 565–585. doi:10.1177/0734242X14536462

Sarc, R., Lorber, K.E., 2013. Production, quality and quality assurance of Refuse Derived Fuels (RDFs). Waste Manag. 33, 1825–1834. doi:10.1016/j.wasman.2013.05.004

Sharan, L., Rosenholtz, R., Adelson, E.H., 2014. Accuracy and speed of material categorization in real-world images. J. Vis. 14, 1–24. doi:10.1167/14.9.12

Simon, L.L., Pataki, H., Marosi, G., Meemken, F., Hungerbühler, K., Baiker, A., Tummala, S., Glennon, B., Kuentz, M., Steele, G., Kramer, H.J.M., Rydzak, J.W., Chen, Z., Morris, J., Kjell, F., Singh, R., Gani, R., Gernaey, K.V., Louhi-Kultanen, M., O'Reilly, J., Sandler, N., Antikainen, O., Yliruusi, J., Frohberg, P., Ulrich, J., Braatz, R.D., Leyssens, T., von Stosch, M., Oliveira, R., Tan, R.B.H., Wu, H., Khan, M., O'Grady, D., Pandey, A., Westra, R., Delle-Case, E., Pape, D., Angelosante, D., Maret, Y., Steiger, O., Lenner, M., Abbou-Oucherif, K., Nagy, Z.K., Litster, J.D., Kamaraju, V.K., Chiu, M.-S., 2015. Assessment of Recent Process Analytical Technology (PAT) Trends: A Multiauthor Review. Org. Process Res. Dev. 19, 3–62. doi:10.1021/op500261y

Slade, R., Gross, R., Bauen, A., 2011. Estimating bio-energy resource potentials to 2050: learning from experience. Energy Environ. Sci. 4, 2645–2657. doi:10.1039/C0EE00732C

Smidt, E., Meissl, K., Schwanninger, M., Lechner, P., 2008. Classification of waste materials using Fourier transform infrared spectroscopy and soft independent modeling of class analogy. Waste Manag. 28, 1699–1710. doi:10.1016/j.wasman.2007.08.003

Sokoll, T., Jacob, A.F., 2007. In-situ moisture detection system with a vector network analyser. Meas. Sci. Technol. 18, 1088. doi:10.1088/0957-0233/18/4/017

Solo-Gabriele, H.M., Townsend, T.G., Hahn, D.W., Moskal, T.M., Hosein, N., Jambeck, J., Jacobi, G., 2004. Evaluation of XRF and LIBS technologies for on-line sorting of CCA-treated wood waste. Waste Manag. 24, 413–424. doi:10.1016/j.wasman.2003.09.006

Tajammal Munir, M., Yu, W., Young, B.R., Wilson, D.I., 2015. The current status of process analytical technologies in the dairy industry. Trends Food Sci. Technol. 43, 205–218. doi:10.1016/j.tifs.2015.02.010

Tatzer, P., Wolf, M., Panner, T., 2005. Industrial application for inline material sorting using hyperspectral imaging in the NIR range. Real-Time Imaging 11, 99–107. doi:10.1016/j.rti.2005.04.003

Tchobanoglous, G., Kreith, F., 2002. Handbook of solid waste management. New York : McGraw-Hill, 2002.

Torres-García, A., Rodea-Aragón, O., Longoria-Gandara, O., Sánchez-García, F., González-Jiménez, L.E., 2015. Intelligent waste separator. Comput. Sist. 19, 487–500. doi:10.13053/CyS-19-3-2254

Troy, D.J., Ojha, K.S., Kerry, J.P., Tiwari, B.K., 2016. Sustainable and consumer-friendly emerging technologies for application within the meat industry: An overview. Meat Sci. doi:10.1016/j.meatsci.2016.04.002

Udo, K., Christof, H., 2010. Electromagnetic techniques for moisture content determination of materials. Meas. Sci. Technol. 21, 82001. doi:10.1088/0957-0233/21/8/082001

Urbanski, J., Wochna, A., Herman, A., 2011. Automated granulometric analysis and grain-shape estimation of beach sediments using object-based image analysis. J. Coast. Res. 1745.

van den Berg, F., Lyndgaard, C.B., Sørensen, K.M., Engelsen, S.B., 2013. Process Analytical Technology in the food industry. Trends Food Sci. Technol. 31, 27–35. doi:10.1016/j.tifs.2012.04.007

van den Broek, W.H.A.M., Derks, E.P.P.A., van de Ven, E.W., Wienke, D., Geladi, P., Buydens, L.M.C., 1996. Plastic identification by remote sensing spectroscopic NIR imaging using kernel partial least squares (KPLS). Chemom. Intell. Lab. Syst. 35, 187–197. doi:10.1016/S0169-7439(96)00056-1

Velis, C.A., Longhurst, P.J., Drew, G.H., Smith, R., Pollard, S.J.T., 2010. Production and quality assurance of solid recovered fuels using Mechanical-Biological Treatment (MBT) of waste: a comprehensive assessment. Crit. Rev. Environ. Sci. Technol. 40, 979–1105. doi:10.1080/10643380802586980

Wagland, S.T., Dudley, R., Naftaly, M., Longhurst, P.J., 2013. Determination of renewable energy yield from mixed waste material from the use of novel image analysis methods. Waste Manag. 33, 2449–2456. doi:10.1016/j.wasman.2013.06.021

Wagland, S.T., Veltre, F., Longhurst, P.J., 2012. Development of an image-based analysis method to determine the physical composition of a mixed waste material. Waste Manag. 32, 245–248. doi:10.1016/j.wasman.2011.09.019

Wahab, D.A., Hussain, A., Scavino, E., Mustafa, M.M., Basri, H., 2006. Development of a Prototype Automated Sorting System for Plastic Recycling. Am. J. Appl. Sci. 3, 1924–1928. doi:10.3844/ajassp.2006.1924.1928

Wang, C., Wang, H., Fu, J., Liu, Y., 2015. Flotation separation of waste plastics for recycling—A review. Waste Manag. 41, 28–38. doi:10.1016/j.wasman.2015.03.027

Ward, A.J., 2016. Near-Infrared Spectroscopy for Determination of the Biochemical Methane Potential: State of the Art. Chem. Eng. Technol. 39, 611–619. doi:10.1002/ceat.201500315

Werheit, P., Fricke-Begemann, C., Gesing, M., Noll, R., 2011. Fast single piece identification with a 3D scanning LIBS for aluminium cast and wrought alloys recycling. ResearchGate 26, 2166–2174. doi:10.1039/C1JA10096C

Wienke, D., van den Broek, W., Melssen, W., Buydens, L., Feldhoff, R., Huth-Fehre, T., Kantimm, T., Winter, F., Cammann, K., 1996. Near-infrared imaging spectroscopy (NIRIS) and image rank analysis for remote identification of plastics in mixed waste. Fresenius J. Anal. Chem. 354, 823–828.

Williams, P., 2001. Implementation of near-infrared technology, in: Williams, P., Norris, K. (Eds.), Near-Infrared Technology in the Agricultural and Food Industries. American association of cereal chemists, St. Paul (Minnesota), pp. 145–169.

Williams, P., Saranwong, S., Kawano, S., Isaksson, T., Segtnan, V.H., 2006. Applications to Agricultural and Marine Products, in: Near-Infrared Spectroscopy in Food Science and Technology. John Wiley & Sons, Inc., pp. 163–277.

WRAP, 2012a. A Classification Scheme to Define the Quality of Waste Derived Fuels [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/WDF_Classification_6P%20pdf.pdf (accessed 4.12.16).

WRAP, 2012b. EfW Development Guidance [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/O_And_EFW_Guidance_FULL.pdf (accessed 9.6.16).

WRAP, 2011. Commercial scale mixed plastics recycling [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/Commercial Scale Mixed Plastics Recycling 19 6 FINAL FINAL VERSION.pdf

WRAP, 2010. Near Infrared sorting of household plastic packaging - Guidance for new users [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/NIR Good practice guidance for existing NIR users Final.pdf

WRAP, 2009. Separation of mixed WEEE plastics final report [WWW Document].

WRAP, 2007. Recovering value from MRFs [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/RecoveringValueMRFs.3644.pdf

WRAP, 2006. Materials Recovery Facilities [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/MRF_v6_19Dec06_LC.pdf (accessed 8.24.16).

Wu, H., White, M., Khan, M.A., 2011. Quality-by-Design (QbD): An integrated process analytical technology (PAT) approach for a dynamic pharmaceutical co-precipitation process characterization and process design space development. Int. J. Pharm. 405, 63–78. doi:10.1016/j.ijpharm.2010.11.045

Wu, S., Panikar, S.S., Singh, R., Zhang, J., Glasser, B., Ramachandran, R., 2016. A systematic framework to monitor mulling processes using Near Infrared spectroscopy. Adv. Powder Technol. doi:10.1016/j.apt.2016.03.022

Wünsch, C., Jenkins, P., 2013. The preparation of calibration curves for a near infrared sensor for the continuous determination of the water content for various waste fractions, in: Collection of Papers of the International Solid Waste Association (ISWA) Specialised Conference “MSW: Management Systems and Technical Solutions.” Moscow.

Yani, I., Budiman, I., 2015. Development of Identification System of cans And Bottle. J. Phys. Conf. Ser. 622, 012053. doi:10.1088/1742-6596/622/1/012053

Yasuda, K., Tanaka, M., Deguchi, Y., 2006. Basic evaluation of sorting technologies for CCA treated wood waste. WIT Press, pp. 319–327. doi:10.2495/ETOX060321

Zorov, N.B., Popov, A.M., Zaytsev, S.M., Labutin, T.A., 2015. Qualitative and quantitative analysis of environmental samples by laser-induced breakdown spectrometry. Russ. Chem. Rev. 84, 1021–1050. doi:10.1070/RCR4538

3 Deep learning in material recovery: development of method to create training database

Abstract

Increasing the rate of material identification, separation and recovery is a priority in resource management and recovery, and rapid, low-cost imaging and interpretation is key. This study uses different combinations of cameras, illuminations, and data augmentation techniques to create databases of images from 128 material samples to train deep neural networks for the recognition of fibre materials. The best classification accuracy obtained averaged 83.7%, indicating it is possible to achieve state of the art performance and meet industrial expectations. The method overcomes problems such as limited available materials by using a setup with 5 cameras and 5 lights and applying simple software image manipulation techniques. The same method can be used to create training databases for deep neural network to recognise in real-time a wider range of materials typically found in solid waste streams. Furthermore, it offers flexibility as the classification cameras could be deployed at different stages within solid waste processing plants, providing feedback for process control, with the potential of increasing plant efficiency and reducing costs.

Keywords: material recognition; deep neural network; machine learning; waste management; material recovery

3.1 Introduction

Priorities to achieve reductions in waste, retaining materials in use within the economy for longer are evident within many government policies and initiatives such as the circular economy (EU Council, 2009, 1999; EU Parliament, 2018, 2017). Targets for collecting separate materials and recycling municipal wastes to reduce landfill reflect this global awareness. Consequently, innovation is needed to advance material identification and separation technologies within current solid waste treatment processes (ISWA, 2016; Sedlak, 2017).

Currently, in waste material recovery facilities (MRF), the composition of incoming and outgoing waste is analysed manually (WRAP, 2014). Typically, samples of 60 kg are taken every 125 tonnes of incoming waste, and of 10 to 50kg from every 15 to 60 tonnes of output materials, depending on the waste stream. Trained people sort these

samples into fractions, which are weighed and reported to different stakeholders. The typical recognition performance of this process is high, as the sorters can analyse the waste items in detail and asking for expert help if needed. The time required to manually sort waste samples depend on the analysed stream, with more complex ones such as infeed mixed materials, containing more fractions to sort out. The level of experience can also affect the manual sorting process, and even during the same day and with the same person, tiredness can become an important factor (Tchobanoglous and Kreith, 2002) affecting material recovery rates. Furthermore, this process is not scalable to moving belts, where the entire waste flow would need to be continuously analysed. As the needs for more and better waste data are required, for example for the implementation of schemes like Extended Producer Responsibility, there is an increased pressure and expectations on the development of automated waste visual detection, with expected accuracy levels of 85% or higher (DEFRA, 2021).

In recent years, many sector studies have used material flow analysis to understand better the volume and composition of wastes (Moriguchi and Hashimoto, 2016; Velis et al., 2013) thus assisting decision making (Turner et al., 2016) and improving material recovery and treatment (Al Sabbagh et al., 2012; Allesch and Brunner, 2017, 2015; Arena and Di Gregorio, 2014; Habib et al., 2014; Pivnenko et al., 2016; Stanisavljevic and Brunner, 2014; Tonini et al., 2014). Knowing the composition of solid waste within process streams in real-time reduces uncertainty (Laner et al., 2016; Rechberger et al., 2014), eliminating the input-output lag during flow analysis (Pivnenko et al., 2016; Zoboli et al., 2016). Integrating material composition sensors at different stages of waste management processes can provide feedback to stakeholders and improve the safety and efficiency of material recovery plants (Vrancken et al., 2017). This study uses off-the-shelf cameras with visible spectrum sensors, identified previously as the most suitable method amongst existing technologies to measure critical quality indicators for waste treatment. This research advances previous research that demonstrated that visual inspection of images of waste can provide good indications of material composition and fuel properties (Peddireddy et al., 2015; Wagland et al., 2013, 2012).

Applying deep learning techniques (LeCun et al., 2015) to train convolutional neural networks (CNN) (Krizhevsky et al., 2012; Simonyan and Zisserman, 2014) to

recognize objects and materials in images is shown to achieve classification rates close to, or even surpassing, human performance (Russakovsky et al., 2015). Typically, CNN are trained with a database of labelled images, containing many examples of each classification category. Trained CNN are then used to classify materials with data from a separate set of test images.

In general, the accuracy of CNN increases with the number of training images (Bell et al., 2015), thus data collection is an important step. The solid waste domain is infinite, as new product types continuously appear on the markets which are then discarded as waste. It is virtually impossible to capture images of every type of item or products in the market, and any training dataset would constitute only a small percentage of the waste that can be found. The challenge is to train the deep learning models with this reduced subset of examples so that it generalises to the large extent of solid waste.

Dataset augmentation is regularly used to improve classification accuracy of smaller datasets, applying transformations such as warping, scaling and colour changing. However, exploiting more images in the training dataset has the disadvantage of increasing the time and storage required, resulting in industry concerns (He and Sun, 2015). To address this, techniques such parallel computing can help reduce the computational burden, while transfer learning can improve accuracy (Rawat and Wang, 2017).

Transfer learning is a technique to improve model performance, which refers to using an existing model, pre-trained on data from a different domain or for a different task. The approach uses the knowledge that the model learnt from previous data and tasks to learn to perform a new task. Different layers of the model architecture can be frozen, so that they are not modified during the training, thus reducing the number of weights.

Training large model architectures with limited number of datapoints can result in overfitting (Caruana et al., 2001), with the network learning to recognise the training images but failing to classify unseen images. The effects can be seen during the training phase, as the error on the test dataset stops reducing and starts to increase after a certain number of training passes, while the error on the training dataset keeps reducing. Stopping the training as soon as the error on the test dataset stops decreasing will limit overfitting.

Another approach to increase generalisation, thus reducing overfitting, is to use dropout (Srivastava et al., 2014). This method applies a random dropout factor, so that on each training pass, a selection of neurons and its connections are hidden, thereby reducing the effective size of the model. The selection of hidden neurons is random, so in every training pass a different subset of the neurons is learning, which effectively means that different, smaller networks are trained. These smaller networks may overfit, but in different ways each time. The resulting trained network combines these learnings which makes it more robust and reduces overfitting (Srivastava et al., 2014).

Whilst others are investigating and developing network architectures to tackle intra-class variability and geometric variance issues (Jaderberg et al., 2015; Laptev et al., 2016), here we focus on generating the best possible training databases given a reduced set of images, using controlled image capture and optimized data augmentation techniques. Classification rates obtained by CNN are negatively affected by large geometric changes in the images, such as rotation (Ciresan et al., 2011; Gong et al., 2014; Razavian et al., 2014) which would be a common occurrence when capturing material images of waste streams on a conveyor belt. Comparing the effects of dataset augmentation with different degrees of image rotation and oversampling (Buda et al., 2017), allows us to improve recognition performance when using images captured in different conditions such as viewing angles and illuminations.

This paper considers only a single CNN architecture and focuses instead on data capture and augmentation techniques to tackle the variability in the waste image data. The CNN architecture used in this study is AlexNet (Krizhevsky et al., 2012), a well-known deep neural network that has been proved to achieve the required operational performance on similar tasks (Bell et al., 2015). The experiments on this work were designed to understand purely the performance effects of different data capture and manipulation methods, and it is assumed that these effects would translate to other network architectures or alternative supervised learning solutions, that are being created and investigated elsewhere. The choice of using AlexNet also allows for benchmark comparisons of the results from this study with previous experiments for similar tasks (Bell et al., 2015), whilst focusing on the effects of data capture, processing, and augmentation.

Newer and more complex CNN model architectures, such as GoogLeNet (Szegedy et al., 2014) and VGG-16 (Simonyan and Zisserman, 2014) can produce better image classification results than AlexNet, but they are 3 to 15 times slower to run, respectively (Canziani et al., 2017). Other alternative model architectures could provide semantic segmentation of the images (Garcia-Garcia et al., 2017), with pixel-wise classification of the images. However, given the cluttered and confusing nature of images of waste streams, an approach using stereo cameras to generate height-maps might get better results. Furthermore, the pixel level of details, identifying the precise shape of waste items as provided by semantic segmentation, is not needed in order to determine the composition of mixed waste, as demonstrated by studies using visual inspection (Peddireddy et al., 2015; Wagland et al., 2013, 2012). Therefore, no attempt was made to use any segmentation solutions and the focus was kept on data generation, processing and augmentation using AlexNet as a benchmark for the classification task.

This study reports how to improve classification rates whilst minimizing the number of training images compared to that of standard datasets (Russakovsky et al., 2015). A series of tests is performed using different methods to acquire and process images then used for training neural networks. Understanding the features that constitute optimal approaches has the potential to lead to robust and efficient learning that recognises materials in recovery facilities, using existing, fast performing deep neural network architectures whilst reducing recognition time and storage requirements.

Besides new types of waste being constantly added to the market, the waste items can be randomly deformed and positioned at any angle on top of conveyor belts. Whilst the positive effects of training with more images, obtained either through image capture or data augmentation, have been discussed before, this paper sets up a systematic study of image data capture and data augmentation. The experiments in this study have been designed to clearly indicate and quantify the effects of using different viewing angles and illuminations, whilst recommending the degrees of image rotation and overlapping to achieve the best possible accuracies.

Finally, whilst the proposed framework was designed to identify solid waste material, the experimental design and results can be transferred to other purposes, using the optimized datasets with different machine learning methodologies and for other

applications. Even inside waste management industry, there are different sectors with very different types of materials, such as construction and demolition, which could benefit from the proposed automated classification solution. Furthermore, the automatic identification of waste items could be integrated with sorting mechanisms, such as robots or air valves, to act on the information provided by the classifiers and not only provide composition analysis but automated sorting of waste.

3.2 Materials and methods

In this section, the methodology is described in detail, including the required materials, the steps, and procedures required by the experimental design, and the analysis applied to the data.

3.2.1 Material samples

The study limited the number of material samples, grouping these into 2 main classes: paper and cardboard, with each class containing multiple different items to provide intra-class variance information. The materials used for these experiments are household waste items previously discarded in waste bins. Typically, these materials are collected and transported in lorries to a waste material recovery facility where they are dumped onto a tipping floor area, before proceeding to being sorted out by the different mechanical processes. While this series of steps will physically mix and compress the waste items, changing their overall shapes, it does not change the local appearance and features of their materials. In the case of cardboards, the boxes are flattened; in the case of papers, they can become crumpled. Both conditions were considered and included in the testing dataset, so that the results are relevant to the expected operational conditions.

Paper samples were represented by white, printed, crumpled, and shredded office paper; brown packaging filler paper; red, white towel, and hygienic tissue paper; glossy magazine paper; folded and crumpled newspaper. Similarly, cardboard samples were represented by plain and printed brown corrugated cardboard, and a selection of household food boxes. A total of 123 material samples, including both paper (66 samples, Figure A-1) and cardboard (29 samples, Figure A-2) were used for training, and a separate set of 28 material samples were used for testing purposes, consisting

of mixed paper (18 samples, Figure A-3) and mixed cardboard (10 samples, Figure A-4).

3.2.2 Image capture

The experimental setup used to collect the library of raw images, consisted of 5 Sony DSC-WX500 cameras, labelled C1 to C5, and 5 Gantom GT21 light sources, labelled L1 to L5 (Figure 3-1). In spherical coordinates, cameras C1, C2, C3, and C4 were placed at a zenith angle of 30 degrees and azimuth angles of 0, 90, 180 and 270 degrees, respectively, with camera C5 placed at a zenith angle of 0 degrees. All the cameras were mounted pointing towards the centre, with material samples placed, at a focal distance of 1,095 mm. The captured images had a resolution of 4,896 by 3,264 pixels, representing the 43 by 29 cm field of view.

The entire field of view of the cameras, the target area, was covered with photographic green screen material, on top of which the individual samples were placed before being photographed (Figure 3-1). The green background material was cleaned manually after each material sample was photographed to prevent cross-contamination. During the experiments, each camera was controlled individually to capture images with different viewing angles, using either camera C1, C2, C3, C4 or C5.

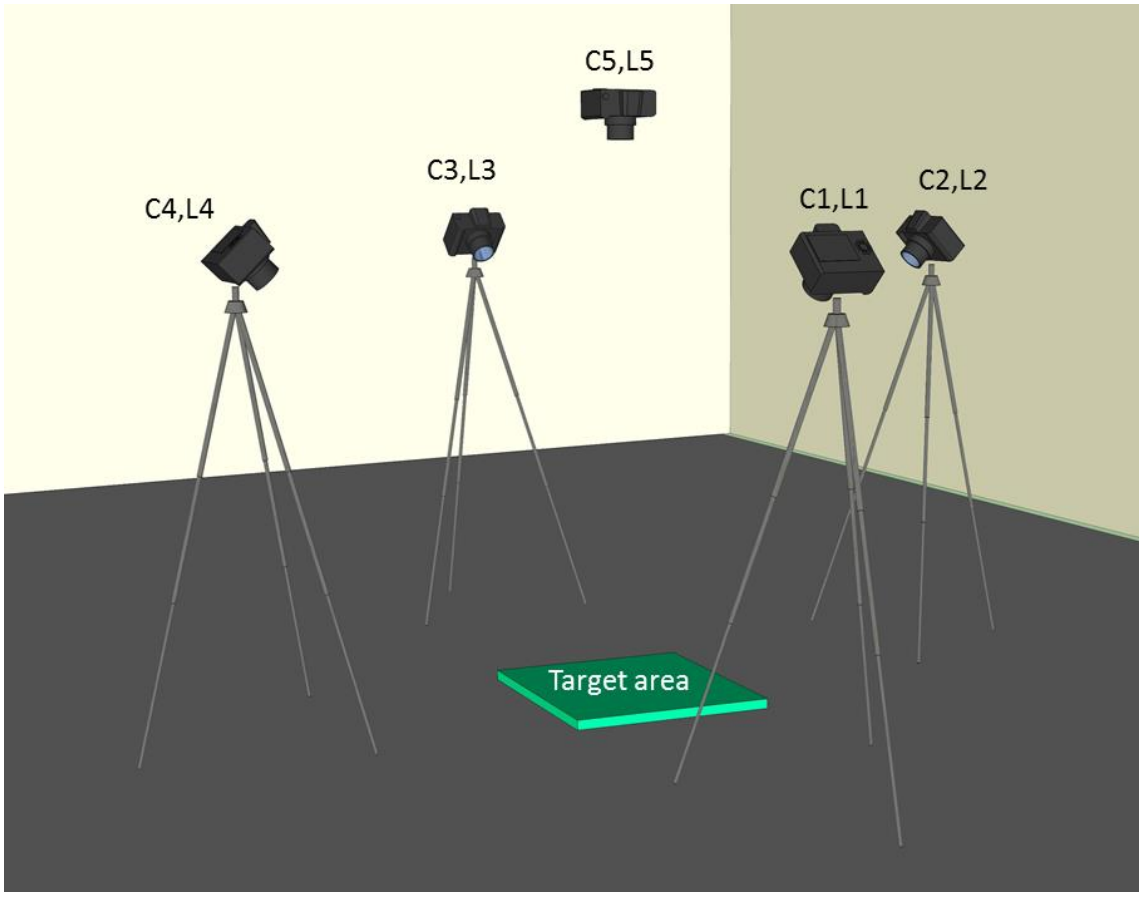


Figure 3-1- Schematics showing the image capture hardware configuration which includes 5 cameras (C1 – C5) and 5 lights (L1 – L5). The lights (not shown) are located next to the cameras, also pointing towards the target area.

The five light sources L1 to L5 were placed next to each of the respective cameras, all set to maximum brightness with a colour temperature of 5,600 K. Each light source illuminated the target area with a uniform beam diameter of 560 mm. During the experiments, each light was controlled individually to capture images with 6 different illuminations using L1, L2, L3, L4, L5, or a configuration where all the lights were powered, labelled L6.

In the first step of the methodology, shown in Figure 3-2, a collection of 30 raw images per sample was captured using all possible combinations of 5 cameras and 6 illuminations. The second step in the method used only the pictures taken by cameras and lights defined by the experimental design. For example, when the experiment considered only 1 camera and 1 light, a single image was selected per material sample, i.e., the one captured using camera C5 and light C5.

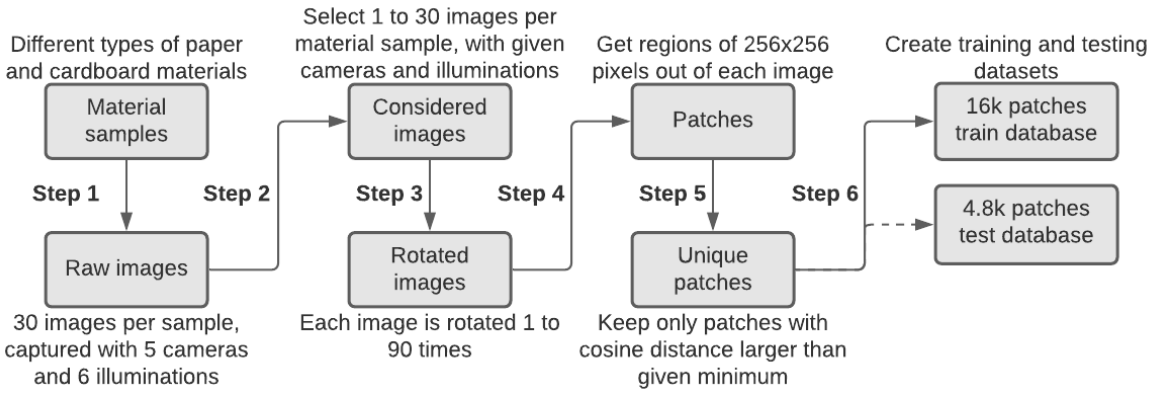
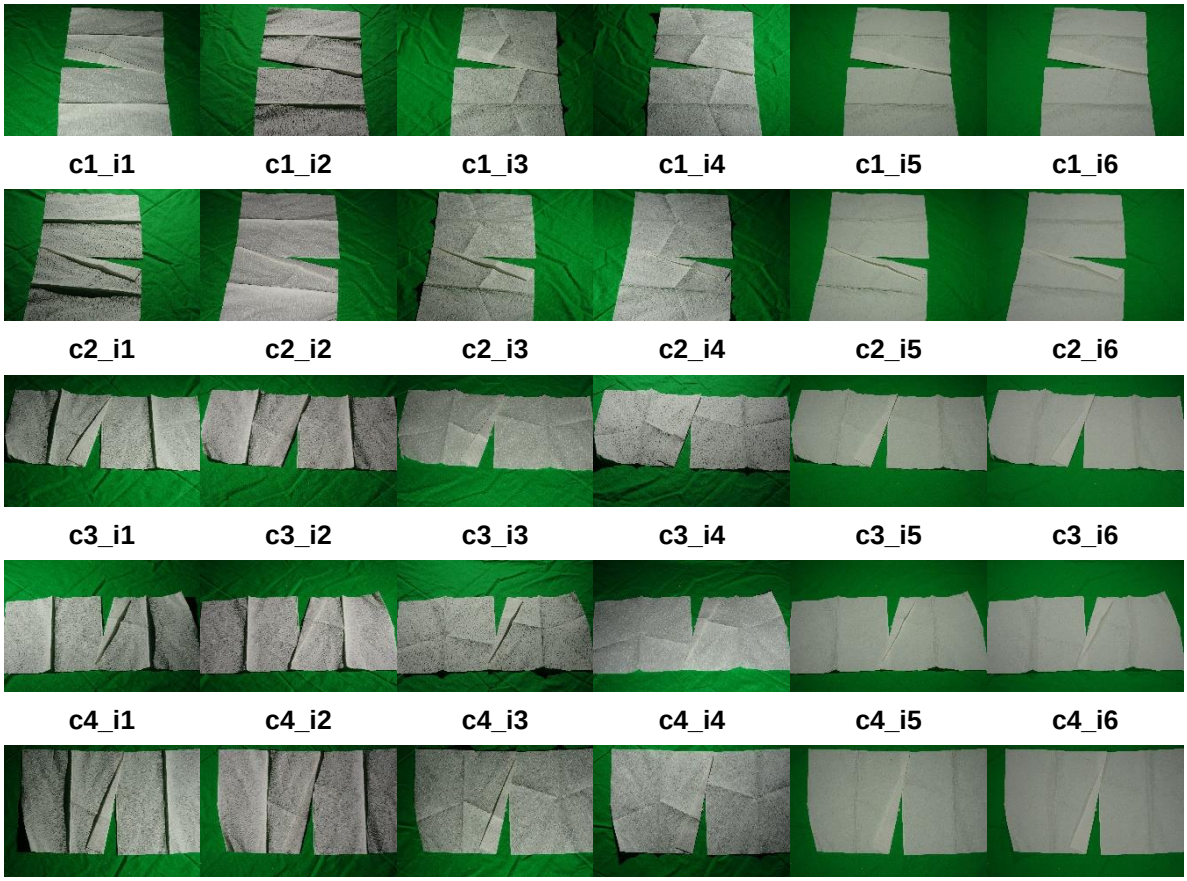


Figure 3-2. Steps taken to create the training and testing databases

When the experiment considered all 5 cameras and 6 lights configurations, 30 raw images per material sample were selected. Examples of images captured using the 30 different combinations of cameras and light sources are shown in Table 3-1.

Table 3-1 – Raw images captured from a sample of kitchen towel paper (tissue_whitetowel_1.jpg, Figure A-1) with a combination of 5 cameras (c1 to c5) and 6 light sources (i1 to i6). Each image shows an area of 43 by 29 cm.



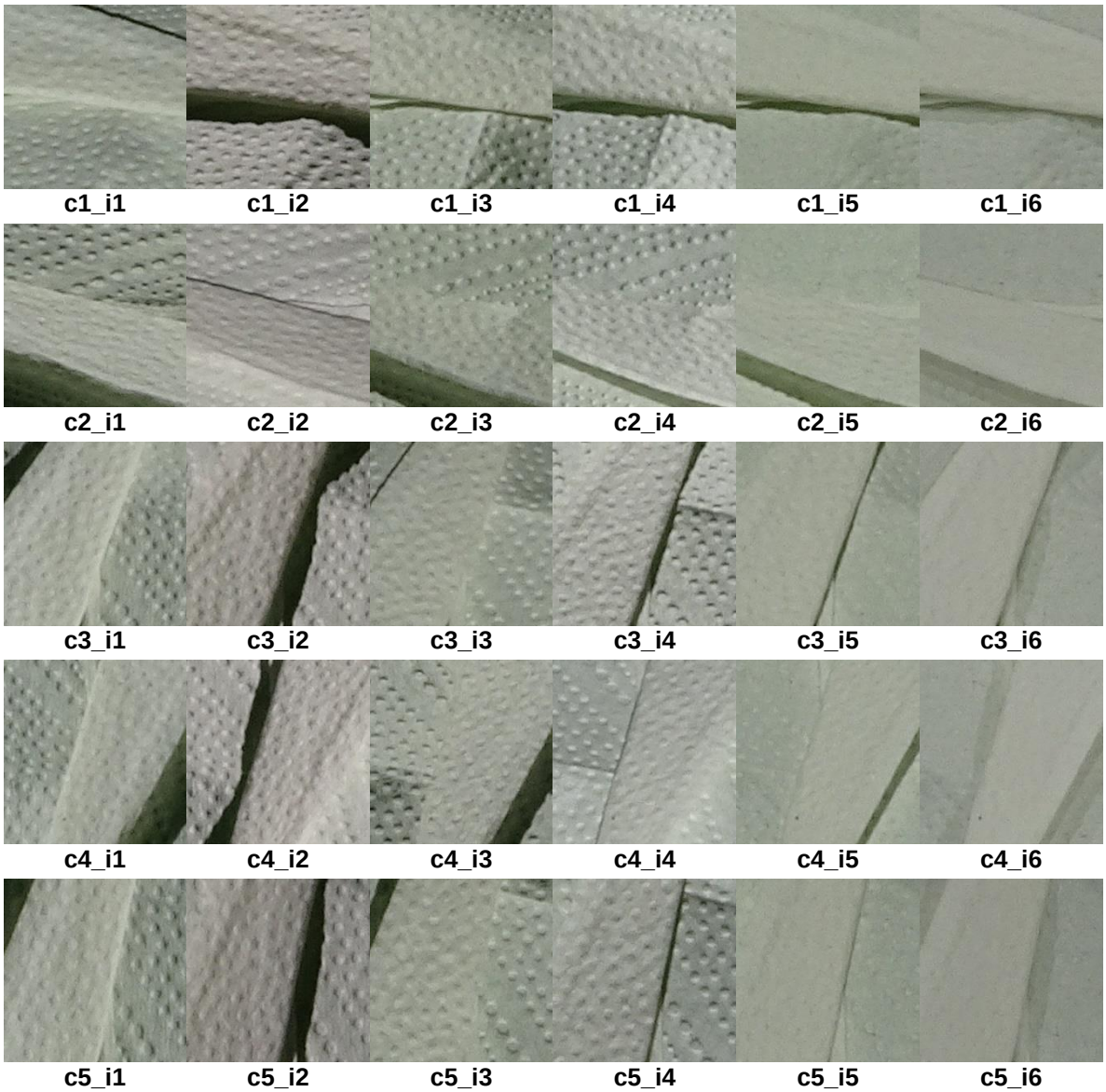
3.2.3 Image processing

Every image was subjected to a series of transformation, implemented in C++ using OpenCV (“OpenCV library 3.1.0,” 2015). The images were rotated in Step 3 using Lanczos resampling over an 8 by 8 pixels neighbourhood. This allowed the rotation of images with respect to the centre at different angles.

In step 4, each rotated image was then split into patches of 256 by 256 pixels (Table 3-2), with varying ranges of overlap between patches. A hue, saturation, and value (HSV) colour-based detection routine was applied to each patch to ensure that only the patches without the green background were kept. All these processes were implemented to run on a single CPU except for the patch extraction step which was

multi-threaded and ran on 4 parallel processes. The time taken in each step was measured and recorded for comparison purposes.

Table 3-2 – Patches extracted from the raw images of kitchen towel paper (tissue_whitetowel_1.jpg, Figure A-1) . The size of each patch is 256 by 256 pixels, equivalent to an area of 2.25 by 2.25 cm. The patches in this table show the same region from the material captured with combination of 5 cameras (c1 to c5) and 6 light sources (i1 to i6) (Table 3-1)



3.2.4 Deep neural network

AlexNet (Krizhevsky et al., 2012) was the convolutional neural network architecture used. Transfer learning was used for the new task, fine-tuning a model trained originally on ImageNet and fine-tuned with the MINC database (Bell et al., 2015), with the output layer fc8 modified to predict 2 types of materials. No layers were frozen, so that all the weights were modified during training. The framework used to implement, train, and test the neural networks was Caffe version 1.00 (Jia et al., 2014) running in parallel on four nVidia K80 GPUs.

Step five of the method step tested each patch by applying the previously trained MINC neural network. The considered output was the fc7 network layer, a 4096-dimensional vector, which was used to calculate the cosine distance to each of the other patches (Bell and Bala, 2015). The cosine distance between two similar images is smaller than the calculated between two unique images. This step ensures the datasets contain only unique patches by requiring a minimum cosine distance value between them. Additionally, each time a patch was fed into the network for training, it was cropped to 227 by 227 pixels from a random region. This reduces overfitting as it avoids duplicate images which would force the network to memorise data instead of learning features.

In the final step, training databases were created by randomly selecting unique patches. In the case of training data sets, 16,000 patches were included for each experiment. The random selection generated unbalanced training datasets, where the number of training patches for paper are different to the training patches for cardboard. The number of paper and cardboard patches on the training dataset for each experiment was noted and reported.

The training step was performed for a minimum of 20 epochs, which meant that each patch in the training database was used at least 20 times. The learning rate was adjusted dynamically using the Adam optimisation algorithm with parameters $\alpha = 10^{-5}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ (Kingma and Ba, 2014). To reduce overfitting, where the network would perform well on the training data but fail on the test patches, the dropout factor in AlexNet was set to 0.5. Furthermore, the trained networks were tested every 4 epochs and the training step was considered complete when the accuracy on the testing dataset stopped improving.

3.2.5 Accuracy testing

A single testing database was generated from 28 samples of paper and cardboard materials that had not been included in the training database. The paper class, see Figure A-3, included white paper (5 samples containing a mix of blank, printed, crumpled, and shredded material), brown paper, tissue paper (8 samples containing a mix of white and red materials), magazine paper (7 samples of different pages), and newspaper (2 samples of different pages). The cardboard class (Figure A-4) included corrugated cardboard (4 samples), and household food cardboard (6 samples).

The test images were captured using 1 camera and 1 light (top view, using C5, L5). The test patches were obtained from random 227 by 227 pixels regions within the raw images, equivalent to 2 by 2 cm, with no overlapping. Only those patches with a cosine distance of 0.2 or more to any other patches in the database were kept. Out of these unique patches, 4,800 were randomly selected to generate the test dataset (paper patches: 3,146; cardboard patches: 1,654).

The performance of the models trained in each experiment were tested on the same test dataset. The final fully connected layer in the network (fc8) contains a 2-way softmax activation function (Bridle, 1990; Krizhevsky et al., 2012), that generates an output with probability distribution over the 2 classes. Each probability output is in the interval [0,1], and the sum of both probabilities is 1. The class with the highest probability was interpreted as the predicted material for each patch in the testing database.

A binary confusion matrix (Table 3-3) was generated with the predictions from the models from each experiment. Each patch is classified into either paper (positive prediction) or cardboard (negative prediction). Misclassifications are also accounted for, either paper classified as cardboard (false negative prediction), or cardboard classified as paper (false positive prediction).

Table 3-3 - Confusion matrix structure used to show the performance of the models from each of the experiments

Data class	Classified as *pos* (paper)	Classified as *neg* (cardboard)
*Positive (*paper*)*	True positive (*tp*)	False negative (*fn*)
*Negative (*cardboard*)*	False positive (*fp*)	True negative (*tn*)

The accuracy metric (Sokolova and Lapalme, 2009) was then calculated to evaluate how effective the model is to classify the patches, using the Equation (3-1). To understand how the model performs on one class or the other, the per-class accuracy is calculated with the recall formula for that class. The per-class accuracy represents the fraction of correct predictions for that subset of test images, and is reported for both data classes (Equations (3-2) and (3-3))

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions} \quad \textbf{(3-1)}$$

$$recall_{pos} = \frac{tp}{tp + fn} = \frac{number\ of\ correct\ positive\ (paper)\ predictions}{total\ number\ of\ positive\ (paper)\ patches} \quad \textbf{(3-2)}$$

$$recall_{neg} = \frac{tn}{tn + fp} = \frac{number\ of\ correct\ negative\ (cardboard)\ predictions}{total\ number\ of\ negative\ (cardboard)\ patches} \quad \textbf{(3-3)}$$

Where *tp=true positive rate*; *tn=true negative rate*; *fp=false positive rate*; fn=*false negative rate*

3.2.6 Data analysis

Each experiment was performed three times, generating an average confusion matrix with rounded up numbers, and calculating the mean accuracy and the standard deviation of the mean accuracy. Significant differences between results were

determined with 1-way ANOVA followed by Dunnett's tests, and 2-way ANOVA followed by Tukey's tests, using 95% confidence intervals.

3.2.7 Experimental design

A series of experiments to evaluate the effects of the illumination and viewing angles configurations, the size of the material samples, software image rotation, patch overlap, and the individuality and the number of training patches were designed. These are shown in Table A1 as a summary of experimental setups and the different configurations tested. The experiments consisted of creating a training database of patches using their specific configuration, training a deep neural network with the created databases, and finally applying the trained networks to the testing database to obtain the mean accuracies. The experiments were performed three times to estimate the uncertainty of the results.

Experiment 1 uses software techniques to increase the number of images, rotating the images by 5, 10, 15, 30, and 90 degrees. Experiment 2 overlaps the extracted patches by 0, 12.5, 25, 37.5, 50, 62.5, 75, and 87.5 percent. Experiments 3 use different hardware configurations to use 1 to 6 illuminations, and experiment 4 uses 1 to 5 cameras.

Finally, experiments 5, 6 and 7 use combinations of techniques to increase the number of training images. Experiment 5 combines the software techniques, rotation and overlapping, using the best results obtained in experiments 1 and 2. Experiment 6 combines hardware techniques, using multiple cameras and illuminations, using the results obtained in experiments 3 and 4. Finally, experiment 7 uses a combination of both software and hardware techniques, including rotation, overlapping, multiple cameras and multiple illumination configurations.

Table 3-4 - Summary of experimental setups and the different configurations tested

Experiment	Cosine distance	Number of rotations (including original)	Overlapping (%)	Lights	Cameras
1	0.2	1, 4, 12, 24, 36, 72	0	L5	C5
2	0.2	1	0, 12.5, 25, 37.5, 50, 62.5, 75, 87.5	L5	C5
3	0.2	1	0	L5, L5+L1, L5+L1+L2, L5+L1+L2+L3, L5+L1+L2+L3+L4, L5+L1+L2+L3+L4+L6	C5
4	0.2	1	0	L5	C5, C5+C1, C5+C1+C2, C5+C1+C2+C3, C5+C1+C2+C3+C4
5	0.2	12*1	25*2	L5	C5
6	0.2	1	0	L1, L2, L3, L4, L5, L6*3	C1, C2, C3, C4, C5*4
7	0.2	12*1	25*2	L1, L2, L3, L4, L5, L6*3	C1, C2, C3, C4, C5*4

Note: *n indicates value(s) based on the results from experiment n.

3.3 Results and discussion

This section objectively reports the results from each of the experiments, and it offers an interpretation, by putting them in context and explaining why they matter.

3.3.1 Experiment 1: Software rotation

Using software rotation to increase the number of unique patches resulted in significant differences to using only patches from the original captured images. The best results were achieved when using a total of 12 rotations of 30 degrees, using a training dataset with 16,000 patches (paper: 10,263; cardboard: 5,737). Training with this model resulted in an accuracy of 81.4% (SD=1.5%). This represents a relative improvement of 14% when compared to a baseline result using only 1 rotation and no augmentation (Figure 3-3).

There were no further benefits when rotating an image more than 12 times, with the results showing that accuracy drops. This could mean that the extra images generated are not providing new information to train the model and may be resulting in overfitting. The results of using software rotations clearly indicate that this augmentation technique is greatly beneficial. The trade-off is the extra time and storage required, which is analysed in depth in Section 3.3.8.

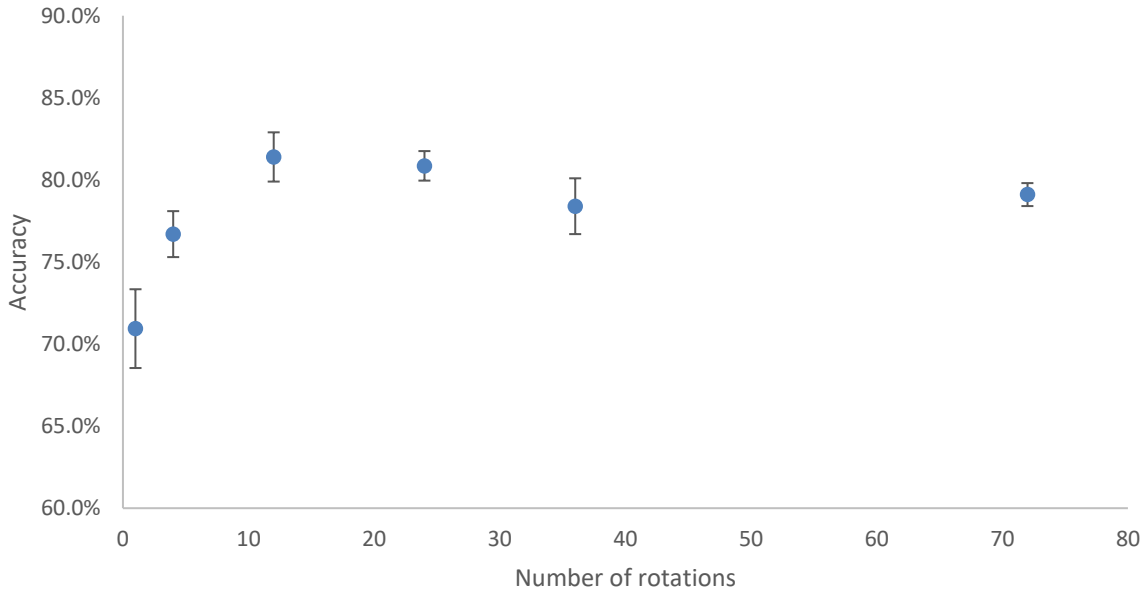


Figure 3-3 - Effects of using software rotations to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions

The results from each of the 3 repetitions when using 12 rotations were averaged to generate a single confusion matrix (Table 3-5). The analysis shows that the model correctly classified 82.7% of the paper patches and 79.0% of the cardboard patches. This could indicate that the cardboard class, due to its higher intraclass variability, is harder to classify than the paper class. The overall accuracy achieved was 81.4%.

Table 3-5 - Confusion matrix with the average prediction results using 12 rotations (30 degrees). Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**82.7% (2,601)**	17.3% (545)
cardboard	21.0% (348)	**79.0% (1,306)**

3.3.2 Experiment 2: Patch overlap

The results show significant differences between the baseline, using 0% overlap, and using 25% and 62.5%. The best accuracy of 76.6% (SD=1.6%) was obtained using a 25% overlap, using a training dataset with 16,000 patches (paper: 10,452; cardboard: 5,548). This result represents a relative gain of 8% to not using any patch overlapping. There are no further benefits by overlapping patches more than 25%, with the results showing no significant difference when further increasing the overlapping (Figure 3-4).

This means that the extra patches generated by overlapping the training images contain no extra information that the network can use for training. This is the same effect observed in experiment 1, where rotating more than 12 times was not beneficial. The results of using patch overlapping clearly indicate that this augmentation technique is greatly favourable to training the model, though it does not increase the accuracy as much as rotating the images.

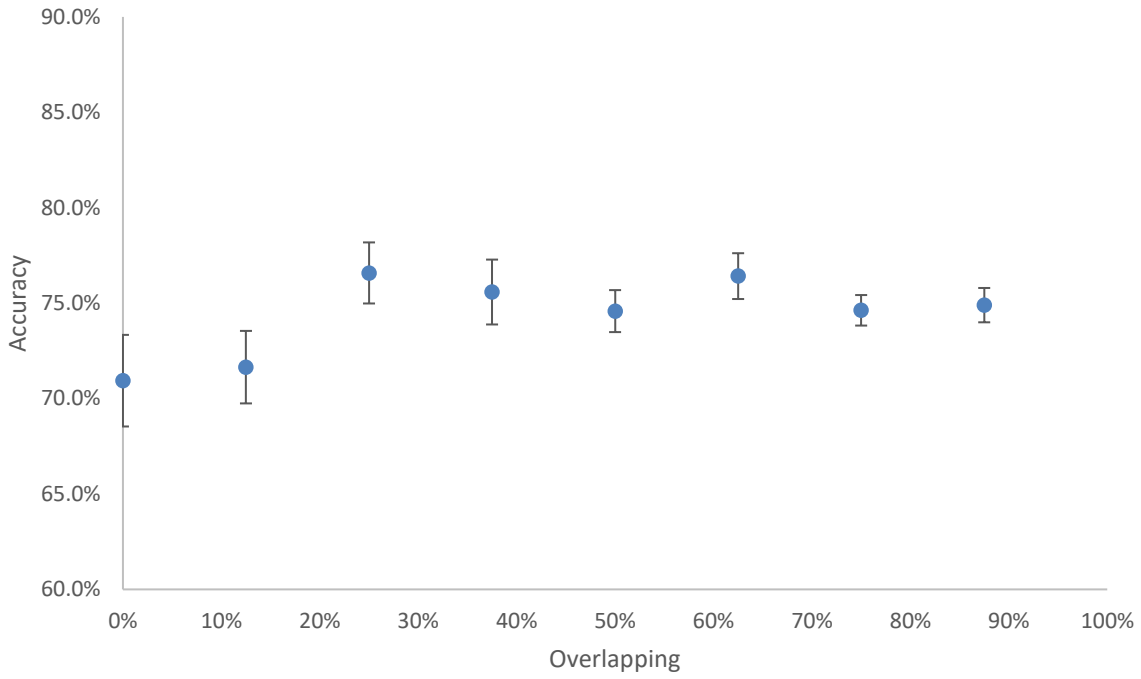


Figure 3-4 - Effects of using different levels of overlapping between patches to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions

The results from each of the 3 repetitions when using 25% overlapping were averaged to generate a single confusion matrix (Table 3-6). The analysis shows that the model

correctly classified 80.5% of the paper patches and 69.2% of the cardboard patches. This could indicate that the cardboard class, due to its higher intraclass variability, is harder to classify than the paper class. In contrast to software rotation, patch overlap cannot add representative data for the cardboard class. The overall accuracy achieved was 76.6%.

Table 3-6 - Confusion matrix with the average prediction results using 25% overlapping. Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**80.5% (2,532)**	19.5% (614)
cardboard	30.8% (510)	**69.2% (1,144)**

3.3.3 Experiment 3: Number of lights

Increasing the number of illuminations resulted in improved accuracies, but the only significant difference was achieved when increasing from 1 to 6 illuminations. Using 6 different illuminations, using a training dataset with 16,000 patches (paper: 9,464; cardboard: 6,536), achieves the maximum accuracy of 75.3% (SD=0.7%). This represents a relative improvement of 6.1% from the baseline using just 1 illumination, which resulted in 70.9% accuracy (Figure 3-5).

The results indicate that each new added illumination adds extra information that the network can use for training. The results of using different hardware illumination clearly indicate that this augmentation technique is beneficial to training the model, though it does not increase the accuracy as much as rotating the images or overlapping the patches.

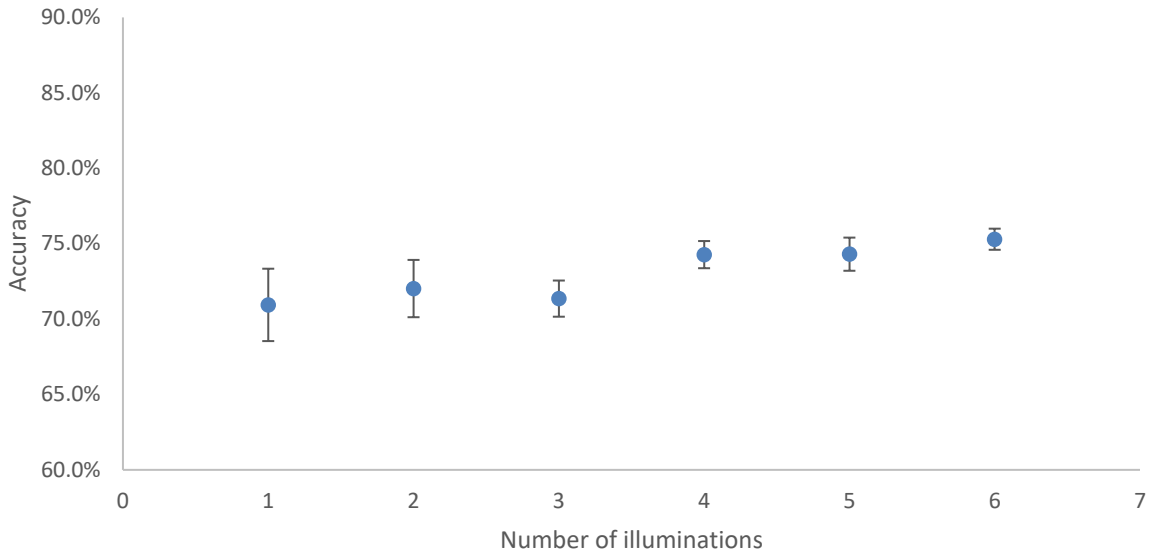


Figure 3-5 - Effects of using multiple illuminations to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions

The results from each of the 3 repetitions when using 6 illuminations were averaged to generate a single confusion matrix (Table 3-7). The analysis shows that the model correctly classified 76.1% of the paper patches and 73.7% of the cardboard patches. This could be explained by the higher intraclass variability in the cardboard class. The overall accuracy achieved was 75.3%.

Table 3-7 - Confusion matrix with the average prediction results using 6 illuminations Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**76.1% (2,396)**	23.9% (751)
cardboard	26.3% (435)	**73.7% (1,218)**

3.3.4 Experiment 4: Number of cameras

Adding images obtained from increasing numbers of cameras resulted in improved accuracies. However, the only significant difference was observed when increasing the number of cameras from 1 to 5. Using 5 cameras results in a training dataset with 16,000 patches (paper: 9,952; cardboard: 6,048) that achieved a maximum accuracy of 75.8% (SD=1.1%) (Figure 3-6).

This result improves the baseline results when using 1 camera by a relative 6.9%. The results indicate that each new camera adds extra information that the network can use for training. This clearly indicate that this augmentation technique is beneficial to training the model, though it does not increase the accuracy as much as rotating the images or overlapping the patches.

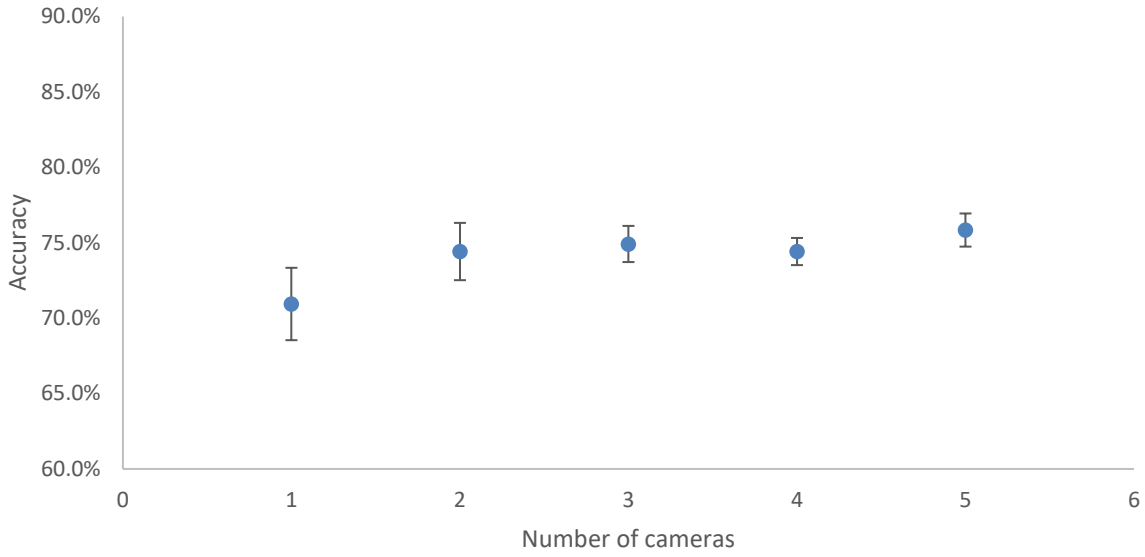


Figure 3-6 - Effects of using multiple cameras to augment the training dataset. The vertical bars denote standard deviation from 3 experiment repetitions

The results from each of the 3 repetitions when using 5 cameras were averaged to generate a single confusion matrix (Table 3-8). The analysis shows that the model correctly classified 76.9% of the paper patches and 73.7% of the cardboard patches. This could indicate that the cardboard class, due to its higher intraclass variability, is harder to classify than the paper class. The overall accuracy achieved was 75.8%.

Table 3-8 - Confusion matrix with the average prediction results using 5 cameras Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**76.9% (2,421)**	23.1% (726)
cardboard	26.3% (435)	**73.7% (1,218)**

3.3.5 Experiment 5: Combined software techniques

This experiment combines the best result from Section 3.3.2 of increasing the overlapping factor to 25%, and the best result from Section 3.3.1, rotating the images 12 times by 30 degrees. This model, trained with a training dataset with 16,000 patches (paper: 9,986; cardboard: 6,014) generated with rotated images and overlapped patches, resulted in an accuracy of 81.8% (SD=1.6%).

This result is significantly better result than the baseline accuracy of 70.9% obtained using only 1 rotation and no overlapping. It is also significantly better than the 76.6% accuracy obtained using just 25% overlapping and no rotations. Compared to just rotating the images, the results are better by a relative 0.5%, and are not significantly different.

The results from each of the 3 repetitions when using 5 cameras were averaged to generate a single confusion matrix (Table 3-9). The analysis shows that the model correctly classified 83.0% of the paper patches and 79.6% of the cardboard patches. As with the individual techniques, the results indicate that the model cannot generalise for the cardboard class as well as for the paper class, due higher intra-class variability in the former. The overall accuracy achieved was 81.8%.

Table 3-9 - Confusion matrix with the average prediction results using a combination of software rotation and patch overlapping. Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**83.0% (2,611)**	17.0% (535)
cardboard	20.4% (338)	**79.6% (1,316)**

3.3.6 Experiment 6: Combined hardware techniques

To understand the extent to which combining hardware techniques can address the effects of a limited data, 30 images were captured per material sample, combining all 5 cameras and 6 different illuminations. The training dataset contained 16,000 patches (paper: 7,893; cardboard: 8,107). The model trained with this combination of different cameras and illuminations resulted in an accuracy of 79.3% (SD=0.9%), which is significantly better than the baseline results of 70.9%, using 1 camera and 1 illumination. Furthermore, it significantly improves the results obtained in Section 3,

using only 6 illuminations, and those from Section 4, using only 5 cameras, by a relative 5.3% and 4.5%, respectively. This means that combining hardware techniques further overcomes the constraints of having a limited data set for training a CNN, adding more variety of information that the model can use to classify these materials.

The results from each of the 3 repetitions when using 5 cameras were averaged to generate a single confusion matrix (Table 3-10). The analysis indicates that the model correctly classified 81.7% of the paper patches and 74.7% of the cardboard patches. The accuracy achieved was 79.3%.

Table 3-10 - Confusion matrix with the average prediction results using combinations of 5 cameras and 6 illuminations. Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**81.7% (2,572)**	18.3% (575)
cardboard	25.3% (419)	**74.7% (1,234)**

3.3.7 Experiment 7: Combined all (software and hardware) techniques

This experiment uses a combination of 5 cameras, 6 illuminations, 12 rotations, and a 25% patch overlap to generate a training dataset with 16,000 patches (paper: 8,077; cardboard: 7,923). Testing the trained model results in an accuracy of 83.7% (SD=0.9%). This result is significantly better than using the combined hardware techniques of 5 cameras and 6 illuminations, improving it by a relative 5.5%. The difference to using combined software techniques is not significantly different, but it is better by a relative 2.5%.

The resulting accuracy of 83.7% is significantly different to the baseline accuracy of 70.9%, improving it by a relative 17.9%. This indicates the full potential of the proposed methodology, as the combination of hardware and software techniques compensates for the limited availability of material samples, enabling the model to achieve state of the art performance (Bell et al., 2015).

The results from each of the 3 repetitions when using 5 cameras were averaged to generate a single confusion matrix (Table 3-11). The analysis shows that the model correctly classified 91.6% of the paper patches and 68.6% of the cardboard patches. The difference in performance per class between paper and cardboard is the highest

of all the conducted experiments. The cardboard class has higher intra-class variability, and the combination of augmentation techniques cannot add enough information for the model to extract. This is because, even though each technique was applied, the total number of training images is the same as in previous experiments. This limits the added information, and the resulting dataset was not rich enough to represent the cardboard class. The accuracy achieved was 83.7%.

Table 3-11 - Confusion matrix with the average prediction results using combinations of 5 cameras and 6 illuminations. Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**91.6% (2,882)**	8.4% (265)
cardboard	31.4% (519)	**68.6% (1,134)**

3.3.8 Time requirements

The configuration used in each experiment affected the time required to generate the training datasets and the classification accuracy. The training dataset was always the same size, as 16,000 patches were always selected. Therefore, the time taken to complete the training was also the same, as it was setup to complete after a maximum of 20 epochs.

The only trade-offs were the added time to capture the needed pictures of each sample. All the possible combinations of 5 cameras and 6 illuminations took an average of 7.5 minutes. This translates to an average of 15 seconds per capture, including the time to configure the cameras, the lights, and the wireless transfer for further computer processing. Considering that the experiments used a training dataset from 95 material samples, this step had the largest impact on the required time to completion.

The rotation and patch overlapping techniques had a relatively small impact on the total time, adding a maximum of 4 minutes when working with images captured with all the cameras and illuminations, at an average of 0.147 seconds per image. The average time to create a training database file was 6 minutes and 40 seconds; this was not affected by the technique as each model was trained with the same number of patches. Finally, the training step for each experiment took an average of 28 minutes, as they all were completed after the same number of epochs.

The best result using images obtained with the combination of software and hardware techniques (experiment 7), took 12 hours and 17 minutes, compared to 1 hour when using no augmentation techniques (baseline). Even though this is a significant level of effort, the resulting model improved the accuracy by a relative 18%, increasing from the baseline of 70.9% to 83.7%.

3.3.9 Accuracy vs configuration

Comparing the individual effects in classification accuracy for each of the 4 main variables studied in the experiments, the single most beneficial change was to apply the image rotations, resulting in a relative gain of 14.8% (Table 3-12, experiment 1). Increasing the number of illuminations was the single technique that resulted in the lowest improvement of a relative 6.1% against the baseline results.

The combination of techniques in pairs resulted in significantly better results than the baseline. The software augmentation techniques, combining rotation and overlapping, proved more beneficial than the pair of hardware techniques of more cameras and illuminations, improving the results by 15.3% and 11.8%, respectively.

Table 3-12. Best configuration and accuracy obtained in each experiment

Parameter	**Value**							
Experiment	Baseline	1	2	3	4	5	6	7
Number of cameras	1	1	1	1	5	1	5	5
Number of illuminations	1	1	1	6	1	1	6	6
Number of rotations*	1	12	1	1	1	12	1	12
Patch overlap (%)	0	0	25	0	0	25	0	25
Accuracy	70.9%	81.4%	76.6%	75.3%	75.8%	81.8%	79.3%	83.7%

*: The number of rotations includes the original image.

The combination of both hardware and software techniques resulted in the best performance (Figure 3-7). The achieved classification accuracy of 83.7% surpassed both the 81.4% performance obtained with CNN (Bell et al., 2015) and the observed 83% human-level accuracy (Sharan et al., 2009), in both cases for similar tasks of

material classification. Furthermore, the achieved performance is very close already to the 85% accuracy expected by the industry by 2025 (DEFRA, 2021). The key difference to prior work is the data requirements, as this method uses a limited set of 151 images, compared to 436,749 images (Bell et al., 2015). The proposed methodology can be implemented in a semi-automatic way, needing only the material sample to create a number of patches representing that material. This contrasts with prior work which requires extensive human classification and annotation of images (Bell et al., 2015).

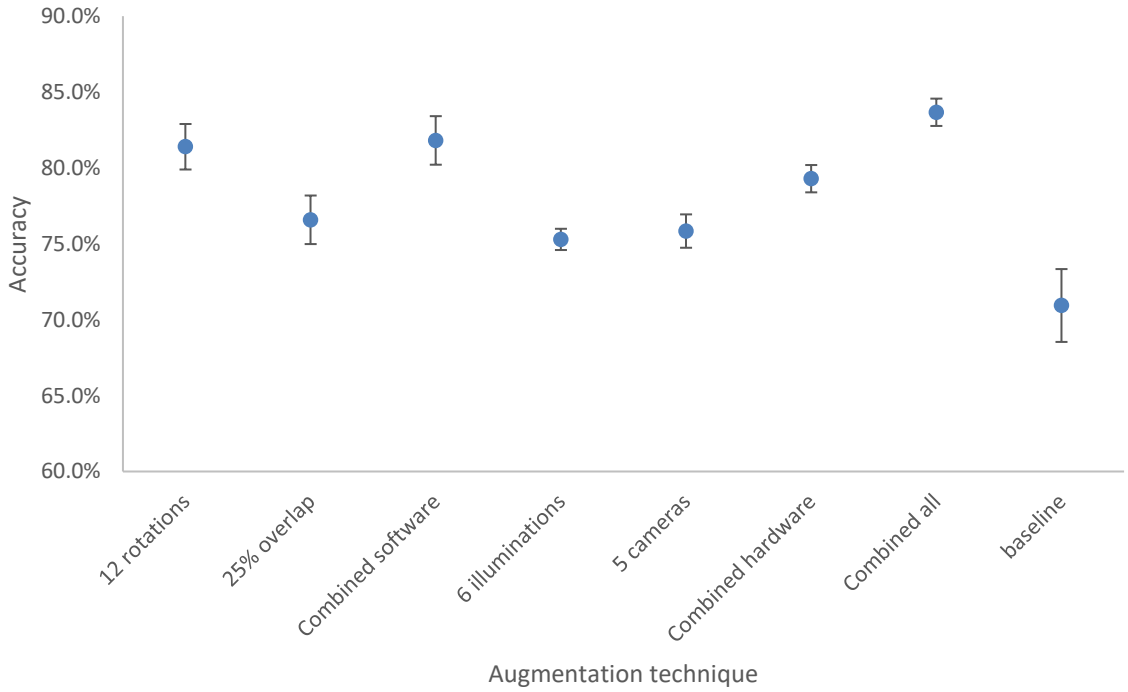


Figure 3-7 - Comparison of best accuracies achieved in each experiment. The vertical bars denote standard deviation from 3 experiment repetitions

3.4 Conclusions

This study takes an innovative approach to solve a problem in waste industry, applying existing AI and computer vision techniques for the automatic classification of solid waste material. There is currently no similar automated solution for benchmarking the results, so the expectations were set out to meet manual analysis performance, which are based on visual classification of materials. The proposed methodology successfully implemented a well-known CNN architecture, which was previously

demonstrated to achieve human-level performance, on this new domain of waste material classification.

Based on the results obtained in this work, and building upon previous research (Peddireddy et al., 2015; Wagland et al., 2013, 2012), this technique could be developed into an automated waste monitoring solution. According to the current industry practices, the composition data is measured manually on spot samples representing less than 1% of the waste streams. In contrast, the proposed AI-based solution, using a camera and a computer running inference to analyse the waste stream flowing on conveyor belts, could estimate its composition continuously and in real-time. This data could then be used to understand the current operational conditions, to adjust and control MRF processes and optimise their efficiency, and to inform future capital investments.

The achieved level of performance of 83.7% compares favourably to previous results in similar tasks, is comparable to current manual implementations in the industry, and is very close to the industry expectations for future requirements such as EPR in the UK This study shows that even if the available number of material samples are limited, the proposed methodology extracts information for training a neural network and achieve state of the art performance for material classification. As opposed to other applications where context data, such as the presence of other objects or materials, orientation, and size can be used to identify a material or an object, in this study the deep neural network was successfully trained only with the characteristics of waste material surfaces as captured with visible cameras enhanced with multiple illuminations.

Whereas this study used two categories with large intra-class variability, the classes could be split so that, for example, white paper, old newspaper, tissue, and magazine would be classified as different types of materials, providing useful information to fibre recycling industries. Furthermore, the proposed image capture and data augmentation methods could be applied to generate training databases for the recognition of more classes of material, such as plastics, textiles, glass, metal, and minerals, or even to objects, for example for the recognition of hazardous items such as batteries and gas canisters.

3.5 References

Al Sabbagh, M.K., Velis, C.A., Wilson, D.C., Cheeseman, C.R., 2012. Resource management performance in Bahrain: a systematic analysis of municipal waste management, secondary material flows and organizational aspects. Waste Manag Res 30, 813–824. https://doi.org/10.1177/0734242X12441962

Allesch, A., Brunner, P., 2017. Material Flow Analysis as a Tool to improve Waste Management Systems: The Case of Austria. Environ. Sci. Technol. 51, 540–551. https://doi.org/10.1021/acs.est.6b04204

Allesch, A., Brunner, P., 2015. Material Flow Analysis as a Decision Support Tool for Waste Management: A Literature Review. Journal of Industrial Ecology 19, 753–764. https://doi.org/10.1111/jiec.12354

Arena, U., Di Gregorio, F., 2014. A waste management planning based on substance flow analysis. Resources, Conservation and Recycling, SI:Packaging Waste Recycling 85, 54–66. https://doi.org/10.1016/j.resconrec.2013.05.008

Bell, S., Bala, K., 2015. Learning visual similarity for product design with convolutional neural networks. ACM Transactions on Graphics 34, 98:1-98:10. https://doi.org/10.1145/2766959

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Bridle, J.S., 1990. Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition, in: Soulié, F.F., Hérault, J. (Eds.), Neurocomputing, NATO ASI Series. Springer, Berlin, Heidelberg, pp. 227–236. https://doi.org/10.1007/978-3-642-76153-9_28

Buda, M., Maki, A., Mazurowski, M.A., 2017. A systematic study of the class imbalance problem in convolutional neural networks. arXiv:1710.05381 [cs, stat].

Canziani, A., Paszke, A., Culurciello, E., 2017. An Analysis of Deep Neural Network Models for Practical Applications. arXiv:1605.07678 [cs].

Caruana, R., Lawrence, S., Giles, C., 2001. Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping, in: Advances in Neural Information Processing Systems. MIT Press.

Ciresan, D.C., Meier, U., Masci, J., Gambardella, L.M., Schmidhuber, J., 2011. Flexible, High Performance Convolutional Neural Networks for Image Classification, in: Proceedings of the 22nd International Joint Conference on Artificial Intelligence. pp. 1237–1242. https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-210

DEFRA, 2021. Extended Producer Responsibility for Packaging [WWW Document]. URL https://consult.defra.gov.uk/extended-producer-responsibility/extended-producer-responsibility-for-packaging/ (accessed 11.8.21).

EU Council, 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:FULL:EN:PDF (accessed 10.19.16).

EU Council, 1999. Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. Official Journal of the European Communities L 182/1, 1–19.

EU Parliament, 2018. Circular economy package: Four legislative proposals on waste [WWW Document]. URL http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2018)614766 (accessed 4.24.18).

EU Parliament, 2017. The circular economy package: new EU targets for recycling [WWW Document]. URL https://www.europarl.europa.eu/news/en/headlines/society/20170120STO59356/the-circular-economy-package-new-eu-targets-for-recycling (accessed 12.18.19).

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Garcia-Rodriguez, J., 2017. A Review on Deep Learning Techniques Applied to Semantic Segmentation. arXiv:1704.06857 [cs].

Gong, Y., Wang, L., Guo, R., Lazebnik, S., 2014. Multi-scale Orderless Pooling of Deep Convolutional Activation Features, in: Computer Vision – ECCV 2014, Lecture

Notes in Computer Science. Presented at the European Conference on Computer Vision, Springer, Cham, pp. 392–407. https://doi.org/10.1007/978-3-319-10584-0_26

Habib, K., Schibye, P.K., Vestbø, A.P., Dall, O., Wenzel, H., 2014. Material Flow Analysis of NdFeB Magnets for Denmark: A Comprehensive Waste Flow Sampling and Analysis Approach. Environ. Sci. Technol. 48, 12229–12237. https://doi.org/10.1021/es501975y

He, K., Sun, J., 2015. Convolutional neural networks at constrained time cost, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5353–5360. https://doi.org/10.1109/CVPR.2015.7299173

ISWA, 2016. Circular Economy: Resources and opportunities [WWW Document]. URL http://www.iswa.org/fileadmin/galleries/Task_Forces/Final_Task_Force_Report.pdf (accessed 4.16.16).

Jaderberg, M., Simonyan, K., Zisserman, A., kavukcuoglu, koray, 2015. Spatial Transformer Networks, in: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 28. Curran Associates, Inc., pp. 2017–2025.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

Laner, D., Rechberger, H., Feketitsch, J., Fellner, J., 2016. A Novel Approach to Characterize Data Uncertainty in Material Flow Analysis and its Application to Plastics Flows in Austria. Journal of Industrial Ecology 20, 1050–1063. https://doi.org/10.1111/jiec.12326

Laptev, D., Savinov, N., Buhmann, J.M., Pollefeys, M., 2016. TI-POOLING: Transformation-Invariant Pooling for Feature Learning in Convolutional Neural Networks, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 289–297. https://doi.org/10.1109/CVPR.2016.38

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Moriguchi, Y., Hashimoto, S., 2016. Material Flow Analysis and Waste Management, in: Taking Stock of Industrial Ecology. Springer, Cham, pp. 247–262. https://doi.org/10.1007/978-3-319-20571-7_12

OpenCV library 3.1.0 [WWW Document], 2015. URL https://docs.opencv.org/3.1.0/ (accessed 5.2.18).

Peddireddy, S., Longhurst, P.J., Wagland, S.T., 2015. Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste Management 40, 9–13. https://doi.org/10.1016/j.wasman.2015.03.015

Pivnenko, K., Laner, D., Astrup, T.F., 2016. Material Cycles and Chemicals: Dynamic Material Flow Analysis of Contaminants in Paper Recycling. Environ. Sci. Technol. 50, 12302–12311. https://doi.org/10.1021/acs.est.6b01791

Rawat, W., Wang, Z., 2017. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. Neural Computation 29, 2352–2449. https://doi.org/10.1162/neco_a_00990

Razavian, A.S., Azizpour, H., Sullivan, J., Carlsson, S., 2014. CNN Features off-the-shelf: an Astounding Baseline for Recognition. arXiv:1403.6382 [cs].

Rechberger, H., Cencic, O., Frühwirth, R., 2014. Uncertainty in Material Flow Analysis. Journal of Industrial Ecology 18, 159–160. https://doi.org/10.1111/jiec.12087

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y

Sedlak, D., 2017. From the Ontonagon Boulder to the Circular Economy. Environ. Sci. Technol. 51, 1941–1942. https://doi.org/10.1021/acs.est.7b00430

Sharan, L., Rosenholtz, R., Adelson, E., 2009. Material perception: What can you see in a brief glance? Journal of Vision 9, 784. https://doi.org/10.1167/9.8.784

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs].

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Information Processing & Management 45, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research 15, 1929–1958.

Stanisavljevic, N., Brunner, P.H., 2014. Combination of material flow analysis and substance flow analysis: A powerful approach for decision support in waste management. Waste Manag Res 32, 733–744. https://doi.org/10.1177/0734242X14543552

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2014. Going Deeper with Convolutions. arXiv:1409.4842 [cs].

Tonini, D., Dorini, G., Astrup, T.F., 2014. Bioenergy, material, and nutrients recovery from household waste: Advanced material, substance, energy, and cost flow analysis of a waste refinery process. Applied Energy 121, 64–78. https://doi.org/10.1016/j.apenergy.2014.01.058

Tchobanoglous, G., Kreith, F., 2002. Handbook of solid waste management. New York : McGraw-Hill, 2002.

Turner, D.A., Williams, I.D., Kemp, S., 2016. Combined material flow analysis and life cycle assessment as a support tool for solid waste management decision making. Journal of Cleaner Production 129, 234–248. https://doi.org/10.1016/j.jclepro.2016.04.077

Velis, C.A., Wagland, S., Longhurst, P., Robson, B., Sinfield, K., Wise, S., Pollard, S., 2013. Solid recovered fuel: materials flow analysis and fuel property development during the mechanical processing of biodried waste. Environmental science & technology 47, 2957–65. https://doi.org/10.1021/es3021815

Vrancken, C., Longhurst, P., Wagland, S., 2017. Critical review of real-time methods for solid waste characterisation: Informing material recovery and fuel production. Waste Management 61, 40–57. https://doi.org/10.1016/j.wasman.2017.01.019

Wagland, S.T., Dudley, R., Naftaly, M., Longhurst, P.J., 2013. Determination of renewable energy yield from mixed waste material from the use of novel image analysis methods. Waste Management 33, 2449–2456. https://doi.org/10.1016/j.wasman.2013.06.021

Wagland, S.T., Veltre, F., Longhurst, P.J., 2012. Development of an image-based analysis method to determine the physical composition of a mixed waste material. Waste Management 32, 245–248. https://doi.org/10.1016/j.wasman.2011.09.019

Wiebel, C.B., Valsecchi, M., Gegenfurtner, K.R., 2013. The speed and accuracy of material recognition in natural images. Atten Percept Psychophys 75, 954–966. https://doi.org/10.3758/s13414-013-0436-y

WRAP, 2014. Sampling and Testing Guidance for Material Facilities [WWW Document]. URL https://wrap.org.uk/sites/default/files/2020-10/WRAP-MF%20Sampling%20Guidance%20April%202014.pdf (accessed 11.8.19).

Zoboli, O., Laner, D., Zessner, M., Rechberger, H., 2016. Added Values of Time Series in Material Flow Analysis: The Austrian Phosphorus Budget from 1990 to 2011. Journal of Industrial Ecology 20, 1334–1348. https://doi.org/10.1111/jiec.12381

4 Deep learning for waste management: effects of colour changes and identification of paper grades

Abstract

Material recognition using visible sensor cameras could determine critical waste characteristics to improve treatment and recovery processes whilst providing key information for stakeholders. This study extends a previously developed method to train and use deep learning for the classification of materials. The results show a relative accuracy improvement of 14% obtained by dividing the materials into more categories with less intra-class variability. Converting the images to greyscale version is shown to increase accuracy rates by 13%, with a higher robustness to changes in colour characteristics of the test materials. The study recommends a minimum of 5,000 training images per class and compares the trade-offs in computing and storage requirements. The method is successfully applied to automatic paper grade identification, achieving rates higher than 98%, which compares favourably to state-of-the-art industrial solutions.

Keywords: material classification, optical sensor, machine learning, convolutional neural network

4.1 Introduction

The development of automatic classification of materials using optical cameras has been highlighted as a priority for solid waste management (Vrancken et al., 2017). The recognition of materials has been investigated for different applications and using many imaging methods, from using custom-designed imaging algorithms (Hu et al., 2011; Liu et al., 2010) to deep learning techniques (Bell et al., 2015; Bello-Cerezo et al., 2017; Bian et al., 2018; Cimpoi et al., 2016). A comprehensive review of methodologies can be found elsewhere (Cavalin and Oliveira, 2017).

The way that solid waste materials are processed in the industry present unique challenges that can affect the performance of general material recognition methods. Images of waste are difficult to recognize because they offer no context information (Schwartz and Nishino, 2015), the samples can be rotated to any angle, and their shapes can be distorted. In a typical waste processing plant, materials flow from one treatment unit to another using material feeders and conveyor belts (Velis et al., 2013).

This offers the opportunity to capture images of materials at different stages of treatment, from material reception to output, providing key information for process control (Vrancken et al., 2017).

An image capture and processing methodology was developed to train a deep neural network for the recognition of solid waste materials with an accuracy rate of 83.7% (Vrancken et al., 2019). The suggested methodology uses a combination of multiple cameras and light sources followed by data augmentation with simple image processing techniques to recognize material in arbitrary rotations. This study is based on that methodology, exploring its limitations, and extending it to increase the classification performance, by focusing on the following specific issues waste material management:

A. A given class of material can include samples that are visually very different to each other, which translates into high intra-class variability. For example, a sample of newspaper looks very different to a sample of red tissue, although they can both be classified as paper (Chapter 3)

B. Solid waste materials can degrade over time when exposed to the sun, moisture, and to other materials (Adamcová and Vaverková, 2014; Canopoli et al., 2018)

The learnings from objectives A) and B) are combined to develop:

C. An automated methodology to solve a specific problem in waste management to classify cardboard and paper materials (INGEDE, 2009)

It has been previously reported that deep neural networks trained using datasets with high intra-class variability produce worst results than using traditional local image descriptors to recognize materials (Bello-Cerezo et al., 2017). However, in the scope of waste material recognition, the materials would be represented by classes with a range of intra-class variability. One way to reduce intra-class variability is to subdivide each material class into subclasses, for example red tissue and newspaper, for training the model, and combine the predictions into their parent class (paper). This study quantifies intra-class variability, and its effects are examined, comparing the performance of deep neural networks trained with different number of classes to represent the materials.

Previous reports of training deep neural networks indicate that balanced datasets, with the same number of images per class, result in better classification accuracies than using unbalanced classes (Bell et al., 2015; Buda et al., 2017). This is validated in this study, whilst also determining the minimum number of images per class to reach acceptable levels of classification accuracy. The computational and time requirements for the training steps are also measured to indicate the trade-offs to assist in the decision-making process.

Material degradation over time when exposed to the elements would affect their colour properties. Therefore, the importance of colour content for the recognition of materials is investigated. The deep networks in this study are trained with colour images and then using the greyscale version of the images, comparing the resulting classification accuracies. Furthermore, tests are performed to study the effects of changing colours in images after a network is already trained.

This is achieved by converting the test images from the red, green, and blue (RGB) to the hue, saturation, and intensity (HSI) colour model (Gonzalez and Woods, 2001, p. 295). The HSI model decouples the colour information, which is described by the hue and saturation channels, from the intensity. The hue channel represents the perceived colour in a palette range from 0 to 360 degrees, where primary colours are separated by 120 degrees, and secondary colours are at 60 degrees from the primaries. The saturation describes the relative purity of the colour, or the percentage of the colour of the given hue mixed up with pure white. The intensity channel is achromatic and indicates brightness level with a range from 0 to 100.

Finally, the proposed methodology is applied for the specific case of fibre material classification in waste management. The estimated global production of recovered paper in 2018 was of 229 million tonnes, an increase of 60% compared to the year 2000, with a market value of 9.8 billion US Dollars (Food and Agriculture Organization of the United Nations, 2018). Advanced material recognition could be used to automatically sort papers of different quality grades, such as white office paper, newspaper, and cardboard (EU Council, 2013).

In the paper industry, there are international standards (British Standards Institution, 2014) specifying the which materials constitute high quality or contamination of waste fibre streams. Furthermore, there are standard methods using spot sampling and

manual inspection to qualify mixed paper waste streams, with accepted measurement errors of 20% (INGEDE, 2009). In this study, a deep neural network will be trained to identify paper grades, comparing its success rates with published industrial and research results.

4.2 Materials and methods

In this section, the methodology is described in detail, including the required materials, the steps, and procedures required by the experimental design, and the applied data analysis.

4.2.1 Material samples

The study used 26 material samples of paper and cardboard classes for training purposes, grouped into 10 different subclasses. Paper samples were represented by 8 sub-classes, including white paper, crumpled or shredded white paper, brown packaging paper, red tissue paper, white or hygienic tissue paper, and glossy magazine paper. The samples representing the cardboard main class included old corrugated cardboard and household drink or food boxes subclasses. A separate set of 23 material samples were used for testing purposes, consisting of 14 paper samples and 9 cardboard samples matching the same 10 subclasses used for training. Finally, a separate validation dataset was created with materials samples not included in either the training or the testing samples. The validation dataset consisted of 12 material samples: 3 sheets of blank white paper, 3 sheets of printed white paper, 3 sheets of newspaper, and 3 pieces of corrugated cardboard.

4.2.2 Image capture

The experimental setup, described in detail before (Vrancken et al., 2019), used 5 visible sensor cameras and 5 light sources, pointing at the target at a focal distance of 1,095 cm. The full sensor is used to image the 43 by 29 cm field of view with an object resolution of 289 pixels per inch (PPI). The camera focal length is set to 16 mm and the lens aperture set to f/8, which results in a depth of field of 394 mm. The camera ISO sensitivity is set to the minimum possible value of 80, and the exposure time is automatically controlled by the camera.

Each of the cameras and lights were controlled individually to capture 30 images per material sample, using the combination of 5 different viewing angles, one per camera,

and 6 different illuminations, one per light source plus a flooded illumination using all lights together.

4.2.3 Image processing

Every image was subjected to a series of transformation, implemented in C++ using OpenCV ("OpenCV library 3.1.0," 2015). The images were rotated 11 times with respect to the centre at different angles, by 30 degrees to each other, using Lanczos resampling over an 8 by 8 pixels neighbourhood (Kylberg and Sintorn, 2016). Each rotated image was then split into patches of 256 by 256 pixels, allowing 25% horizontal and 25% vertical overlap between patches. All these processes were implemented to run on a single CPU except for the patch extraction step which was multi-threaded and ran on 4 parallel processes.

A hue, saturation, and intensity (HSI) colour detection ensured that the patches containing any green background colour were discarded. To get greyscale images, the colour images were converted pixel by pixel using the formula in Equation (4-1).

$$grayscale = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \quad \textbf{(4-1)}$$

Where *R, G,* and *B* are the pixel intensity values for the red, green, and blue channels in the original colour image, respectively.

Finally, a 3-channel greyscale image version was created, by repeating the greyscale image into RGB channels, thus creating a greyscale image inside a colour image structure, to be able to use the same neural network architecture for both colour and greyscale images.

For some experiments, the test images were manipulated by shifting the hue value of each pixel between -180 and 180 degrees, in steps of 10 degrees. The hue value after manipulation is wrapped around the range from 0 to 360 degrees. On other experiments, the saturation values were offset by a range of values from -50 to 0%, in steps of 10%. Similarly, the intensity (brightness) values were offset by a range of values from -50 to 50%, in steps of 10%. For both the saturation and the brightness, their values after manipulation were limited to the range 0 to 100%.

4.2.4 Deep neural network

The framework used to implement, train, and test the neural networks was Caffe version 1.00 (Jia et al., 2014) running in parallel on 4 nVidia K80 GPUs. The convolutional neural network architecture used is AlexNet (Krizhevsky et al., 2012), previously trained ImageNet dataset and finetuned with the MINC database (Bell et al., 2015). The learning rate was adjusted dynamically using the Adam optimisation algorithm with parameters $\alpha = 10^{-5}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ (Kingma and Ba, 2014). To reduce overfitting, where the network would perform well on the training data but fail on the test patches, the dropout factor in AlexNet was set to 0.5.

Each patch in the training database was tested by applying the previously trained MINC neural network. The considered output was the fc7 network layer, a 4096-dimensional vector, which was used to calculate the cosine distance to each of the other patches in the database (Bell and Bala, 2015). The cosine distance between two similar images is smaller than the calculated between two different images. This step ensures the datasets contain only unique patches by requiring a minimum cosine distance value of 0.2 for experiments 1 through 5, and cosine distances between 0.005 and 0.1 for experiment 6.

A different training database was generated for each experiment with the remaining unique patches. Each training experiment was performed for a minimum of 20 epochs, so that each patch in the training database was used at least 20 times. The training was interrupted when the accuracy results stopped improving. Each time a patch of 256 by 256 pixels was fed into the network for training, where it was cropped to 227 by 227 pixels from a random region to prevent replicate input data. These steps were taken to minimize the overfitting effects during training and obtain a generalized trained network, instead of memorizing training data.

4.2.5 Accuracy testing

Testing databases were generated from the 23 material samples that had not been included in the training database. The test patches were obtained from 227 by 227 pixels regions within the raw images, equivalent to 2 by 2 cm, with no overlapping. All the patches were stored in the testing databases, regardless of the cosine distance to the other test patches. Accuracy tests were performed on these testing databases, and every test patch was assigned a probability of being one of the classes. The class

with the highest probability was interpreted as the predicted material for each patch in the testing database, which was then used to calculate the classification accuracy of the trained network.

A multi-class confusion matrix was generated with the predictions from the models in each experiment. The accuracy was calculated to indicate the fraction of correct predictions (Grandini et al., 2020) using the Equation (3-1). The accuracy metric could be misleading when testing unbalanced datasets, as some classes are not represented as much as others and contribute less to the accuracy calculation (Grandini et al., 2020; Sokolova and Lapalme, 2009). Therefore, the accuracy on each subclass was analysed individually, using the recall formula in Equation (3-2) to illustrate the performance on each class, as implemented in similar works (Bell et al., 2015). Finally, the average per-class accuracy (Bell et al., 2015; Grandini et al., 2020; Sokolova and Lapalme, 2009) was calculated with Equation (3-3) as an indicator of the performance of the model across the different classes on the test dataset.

$$accuracy = \frac{number\ of\ correct\ predictions}{total\ number\ of\ patches} \quad \textbf{(4-2)}$$

$$per\ class\ accuracy = accuracy_i = recall_i = \frac{tp_i}{tp_i + fn_i} = \frac{number\ of\ correct\ class\ i\ predictions}{total\ number\ of\ class\ i\ patches} \quad \textbf{(4-3)}$$

Where *tp* are the correct predictions (true positive), *fn* are incorrect predictions (false negatives), for the subset of examples for class *i*

$$average\ per\ class\ accuracy = \frac{\sum_{i=1}^{n} accuracy_i}{n} \quad \textbf{(4-4)}$$

The following is a simplified example, showing a confusion matrix with 3 classes (Table 4-1) and the calculation of the different indicators with Equations (4-5), (4-6), and (4-7):

Table 4-1 - Example confusion matrix to illustrate the calculations of the different performance indicators

		predictions				
		paper	cardboard	metal	patches per class	correct predictions (tp) per class
actual	paper	20	3	2	25	20
	plastic	3	15	3	21	15
	metal	3	12	15	30	15

Table 4-2 – Normalised version of the example confusion matrix to enable comparison between different tests. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type.

		predictions		
		paper	cardboard	metal
actual	paper	**80.0%**	12.0%	8.0%
	plastic	14.3%	**71.4%**	14.3%
	metal	10.0%	40.0%	**50.0%**

$$accuracy = \frac{20 + 15 + 15}{25 + 21 + 30} = 67.1\%$$

$$paper\ class\ accuracy = \frac{20}{25} = 80.0\%$$
$$plastic\ class\ accuracy = \frac{15}{21} = 71.4\%$$
$$metal\ class\ accuracy = \frac{15}{30} = 50.0\%$$

$$average\ per\ class\ accuracy = \frac{80\% + 71\% + 50\%}{3} = 65.8\%$$

$$accuracy = \frac{20 + 15 + 15}{25 + 21 + 30} = 67.1\% \quad \textbf{(4-5)}$$

$$paper\ class\ accuracy = \frac{20}{25} = 80.0\% \quad \textbf{(4-6)}$$
$$plastic\ class\ accuracy = \frac{15}{21} = 71.4\%$$
$$metal\ class\ accuracy = \frac{15}{30} = 50.0\%$$

$$average\ per\ class\ accuracy = \frac{80\% + 71\% + 50\%}{3} = 65.8\% \quad \textbf{(4-7)}$$

4.2.6 Data analysis

The predicted material for every test patch was compared to the actual material class, calculating average accuracy which is then stated as a percental value. The results of repeated tests were used to calculate the standard deviation (SD) of the mean. Significant differences between the results were assessed with parametric t-tests with Welch's corrections using 95% confidence intervals.

4.2.7 Experimental design

A series of six experiments, summarised in Table 4-3, was designed to evaluate the effects of intra-class variability, colour properties and the number of images of materials. A final experiment applies the method to identify paper grades. All the experiments consisted of creating a training database of patches using a specific configuration, training a deep neural network with the created databased, and finally applying the trained network to the relevant testing database to obtain the mean accuracies. Each experiment was performed three times to estimate the uncertainty of the test results

Experiment 1 compared the classification rates obtained when training the neural networks using the 2 main classes, paper and cardboard, and using the 10 subclasses, to sort the test materials into either paper or cardboard. Experiment 2 trained the neural networks using 10 classes of materials and compared the classification rates into those same 10 classes, when using the original colour images and using the greyscale version of the same training and testing images. Experiments 3 through 5 evaluated the effects of changing the hue, the saturation, and the

brightness of the colour test images, respectively, when classifying them using the neural network trained with either colour or greyscale images (Figure 4-1).

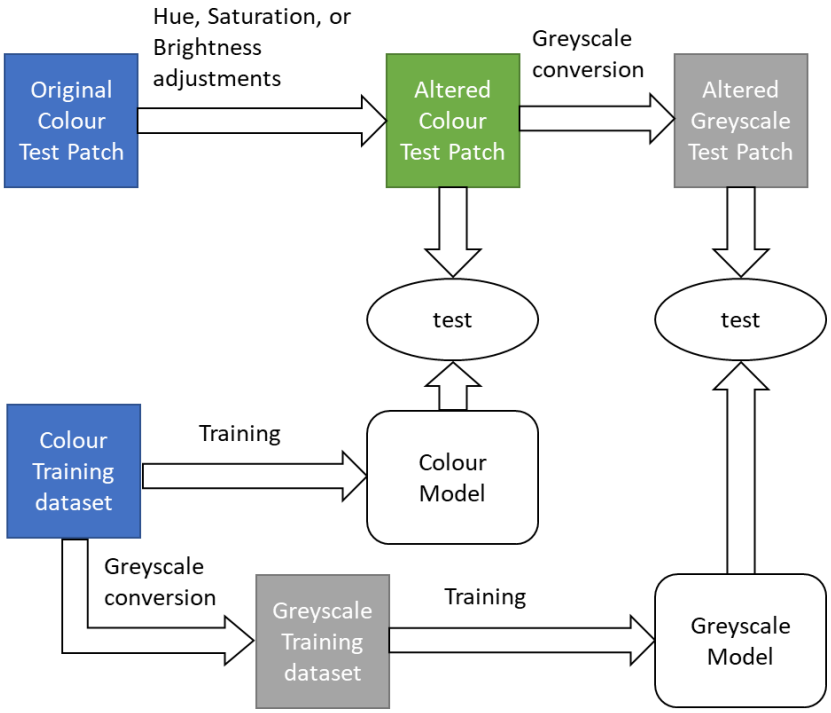


Figure 4-1 - Diagram describing the process followed for Experiments 3 to 5, to study the effects of colour changes

In experiment 6, the neural networks were trained and tested using greyscale images of 3 classes of material: white paper, old newspaper, and corrugated cardboard. To study the effects of different number of training samples, 8 balanced databases were produced with increasing numbers of patches per class. The number of patches per class is based on the results of experiments 1-5.

Table 4-3. Summary of experimental design and specific goals.

Experiment #	To study the effects of:	Images used	# Training classes	# Testing classes
1	Number of classes	Colour	2 and 10	2
2	Colour content	Colour and greyscale	10	2 and 10
3	Changes in the hue channel of the test images	Colour and greyscale	10	10
4	Changes in the saturation channel of the test images	Colour and greyscale	10	10
5	Changes in the brightness channel of the test images	Colour and greyscale	10	10
6	Number of training images per class	Grayscale	3	3

Finally, in experiment 7 the networks trained by experiment 6 were used to classify the validation dataset, which included 90 images of blank white paper, 90 images of printed white paper, 90 images of newspaper, and 90 images of corrugated cardboard. These images were different to the ones used for either training or testing the networks. Each image was divided into patches sized 227 by 227 pixels, which were then individually tested. A majority voting system determined the predicted material in each image, counting the predicted class of each patch as one vote. The predicted class of a sample was the one with the highest number of votes.

4.3 Results

This section objectively reports the key results from each of the experiments, highlighting the model performance achieved with the tested conditions.

4.3.1 Experiment 1: Effects of the number of training classes

The images obtained from the set of 26 training material samples were augmented with rotations and split into partially overlapping patches to create the training dataset. The resulting training dataset contained a total of 569,167 patches: 311,580 from the paper class and 257,587 from cardboard. This database was used to train the network for the task of classifying patches into 2 categories, either paper or cardboard. The test dataset consisted of 143,104 patches (paper: 93,544; cardboard: 49,560), extracted from unseen images of the separate set of 23 material samples.

The results show an accuracy of 76.8% (SD=0.9%). The accuracies for individual classes, which represent the fraction of test images from the subclass that was correctly classified, were 71.5% (SD=4.3%) for paper test samples and 86.9% (SD=2.8%) for the cardboard test samples. The mean cosine similarity of the test dataset before training was 0.44 (SD=0.16) and 0.59 (SD=0.21) after training.

Table 4-4 - Confusion matrix, training with 2 classes, testing on 2 classes. Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**71.5% (66,840)**	28.5% (26,704)
cardboard	13.1% (6,484)	**86.9% (43,076)**

A deeper analysis of the accuracy for each of the 10 subclasses (Table 4-5) shows the worst accuracy was obtained for red tissue paper with 51.9% (SD=8.8%) whilst the best accuracy was achieved for household cardboard with 92.5% (SD=0.5%).

Table 4-5 - Confusion matrix, model trained with 2 classes, tested on 2 classes, and evaluated for each subclass. Correct predictions in bold.

2 CLASSES	10 SUBCLASSES	PAPER	CARDBOARD
Paper	White	**88.5% (5,728)**	11.5% (746)
	White (printed)	**69.4% (4,465)**	30.6% (1,969)
	White (crumpled, shredded)	**85.1% (9,634)**	14.9% (1,682)
	Brown filler	**63.2% (7,818)**	36.8% (4,548)
	Tissue (red)	**51.9% (3,315)**	48.1% (3,071)
	Tissue (white, hygienic)	**83.4% (10,486)**	16.6% (2,089)
	Magazine	**57.6% (7,100)**	42.4% (5,227)
	Newspaper	**71.3% (18,294)**	28.7% (7,372)
Cardboard	Corrugated	21.9% (4,233)	**78.1% (15,116)**
	Household	7.5% (2,251)	**92.5% (27,960)**

Training with 10 classes but classifying each test material into 2 classes, either paper or cardboard, resulted in an accuracy of 87.3% (SD=0.5%) (Table 4-6). The accuracy for paper test samples was 83.6% (SD=1.7%) whilst a 94.3% (SD=1.5%) accuracy was obtained for the cardboard test samples. The cosine similarity of the test dataset before training was 0.60 (SD=0.14) and 0.74 (SD=0.14) after training.

Table 4-6 - Confusion matrix, model trained with 10 subclasses, tested on 2 classes. Correct predictions in bold.

True class	Classified as paper	Classified as cardboard
paper	**83.6% (78,211)**	16.4% (15,335)
cardboard	5.7% (2,806)	**94.3% (46,753)**

Analysing each of the 10 subcategories, the accuracy rate ranged from 57.4% (SD=1.4%) for magazine paper, to 99.1% (SD=0.4%) for tissue paper. The accuracy

rate of 87.3% obtained when training with 10 classes is significantly better than the 76.8% obtained when training with only 2 classes (Table 4-7).

Table 4-7 - Confusion matrix, model trained with 10 subclasses, tested on 2 classes, and evaluated for each subclass. Correct predictions in bold.

2 CLASSES	10 SUBCLASSES	PAPER	CARDBOARD
Paper	White	**94.1% (6,094)**	5.9% (380)
	White (printed)	**92.7% (5,962)**	7.3% (471)
	White (crumpled, shredded)	**87.1% (9,856)**	12/9% (1,461)
	Brown filler	**76.2% (9,417)**	23.8% (2,949)
	Tissue (red)	**93.8% (5,993)**	6.2% (394)
	Tissue (white, hygienic)	**99.1% (12,460)**	0.9% (115)
	Magazine	**57.4% (7,080)**	42.6% (5,247)
	Newspaper	**83.2% (21,348)**	16.8% (4,319)
Cardboard	Corrugated	6.5% (1,256)	**93.5% (18,092)**
	Household	5.1% (1,550)	**94.9% (28,660)**

The results of classifying patches into 2 classes, obtained when training with 2 and with 10 classes, were compared in terms of accuracies for each subclass (Table 4-8). Every subclass performed better when tested with the network trained with 10 subclasses. The highest relative increase in accuracy of 80.9% was obtained for the *tissue (red)* subclass, which went from 51.9%, when the model trained with 2 classes, to 93.9% when the model was trained with 10 classes. The smallest effect was observed on the *magazine* subclass, which showed a relative increase of 1.1%, achieving 58.2% accuracy when trained with 10 classes.

Table 4-8 - Comparison of accuracies of networks trained to differentiate either 2 or 10 class types, to classify the materials into 2 types, either paper or cardboard. Bold indicates best result.

2 CLASSES	10 SUBCLASSES	ACCURACY PER CLASS WHEN TRAINING WITH 2 CLASSES	ACCURACY PER CLASS WHEN TRAINING WITH 10 SUBCLASSES
Paper	White	88.5%	**94.1%**
	White (printed)	69.4%	**92.7%**
	White (crumpled, shredded)	85.1%	**87.1%**
	Brown filler	63.2%	**76.2%**
	Tissue (red)	51.9%	**93.9%**
	Tissue (white, hygienic)	83.4%	**99.1%**
	Magazine	57.6%	**58.2%**
	Newspaper	71.3%	**83.2%**
Cardboard	Corrugated	78.1%	**93.5%**
	Household	92.5%	**94.9%**

4.3.2 Experiment 2: Effects of colour content

The same image capture and augmentation methodology was used to generate a training dataset of 523,945 images split into 10 different classes, instead of only paper and cardboard. The dataset was highly unbalanced, from the 230 patches for the smaller class (*paper white*), to 222,564 patches in the largest class (*cardboard household*) (Table 4-9).

Table 4-9 - Number of patches per class in the training dataset

Class	Patches	Class	Patches	Class	Patches
Cardboard Household	222,564	Paper Magazine	29,364	Paper Tissue (red)	2,339
Paper Newspaper	154,737	Paper Brown filler	24,756	Paper white	2,30
Paper White (crumpled, shredded)	48,789	Paper Tissue (white, hygienic)	7,754		
Cardboard Corrugated	29,694	Paper White (printed)	3,718		

This experiment resulted in an accuracy of 72.6% (SD=1.1%). But the test dataset is highly imbalanced, so the accuracy is biased towards the class with the largest number of test patches, in this case the *cardboard household* class. The individual class accuracies ranged from 30.5% (SD=6.3%) for white paper to 94.8% (SD=0.7%) for household cardboard boxes. The average per-class accuracy was 66.1% (SD=0.7%) for the classification of 10 classes using colour images (Table 4-10).

The same experiment, but training and testing the network using greyscale images, resulted in an accuracy of 77.4% (SD=0.4%) (Table 4-12), which is significantly better. The average per-class accuracy was a significantly better 74.6% (SD=0.6%), which is a relative improvement of 12.9% over the performance of the colour classification model.

Table 4-10 - Confusion matrix when using the *Coloured* version of the patches. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results showing the numbers of patches can be seen in Table A-1.

	Paper white	Paper white printed	Paper white (crumpled, shredded)	Paper brown filler	Paper tissue red	Paper tissue (white, hygienic)	Paper magazine	Paper newspaper	Cardboard corrugated	Cardboard household
Paper white	**30.5%**	1.3%	0.1%	0.0%	0.0%	0.0%	62.2%	0.0%	0.0%	5.9%
Paper white printed	3.8%	**46.6%**	0.1%	0.0%	0.0%	0.3%	40.5%	1.4%	0.0%	7.3%
Paper white (crumpled, shredded)	0.9%	8.0%	**61.7%**	0.1%	0.0%	0.4%	7.4%	8.7%	0.0%	12.9%
Paper brown filler	0.0%	0.0%	0.1%	**33.5%**	0.0%	13.0%	0.0%	29.6%	3.2%	20.6%
Paper tissue red	0.0%	0.0%	0.0%	0.0%	**93.8%**	0.0%	0.0%	0.0%	0.1%	6.1%
Paper tissue (white, hygienic)	0.6%	0.2%	1.8%	1.0%	0.0%	**91.7%**	1.5%	2.3%	0.0%	0.9%
Paper magazine	0.3%	2.4%	0.4%	0.0%	0.0%	0.4%	**44.7%**	9.2%	0.0%	42.6%
Paper newspaper	0.1%	2.3%	0.6%	0.1%	0.0%	1.9%	6.0%	**72.1%**	0.5%	16.3%
Cardboard corrugated	0.0%	0.0%	0.0%	0.3%	0.0%	3.0%	0.7%	2.5%	**91.1%**	2.4%
Cardboard household	0.0%	0.1%	0.1%	0.0%	0.1%	0.1%	1.5%	3.1%	0.1%	**94.8%**

When working with greyscale images, the worst class accuracy was obtained with the magazine paper test samples, reaching 43.4% (SD=2.3%), with 38.8% (SD=4.8%) of them being incorrectly classified as household cardboard, and 10.7% (SD=1.5%) as newspaper (Table 4-11). On the other hand, the corrugated cardboard test samples were correctly identified in 95.3% (SD=1.2%) of the cases, with 1.5% (SD=0.4%) being wrongly assigned the category of household cardboard and 1.4% (SD=0.5%) as newspaper.

Table 4-11 - Confusion matrix when using the greyscale version of the patches. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results showing the numbers of patches can be seen in Table A-2.

	Paper white	Paper white printed	Paper white (crumpled, shredded)	Paper brown filler	Paper tissue red	Paper tissue (white, hygienic)	Paper magazine	Paper newspaper	Cardboard corrugated	Cardboard household
Paper white	**75.1%**	5.4%	0.8%	0.0%	0.0%	1.2%	8.8%	0.0%	0.0%	8.6%
Paper white printed	7.5%	**64.7%**	0.1%	0.0%	0.0%	0.4%	16.1%	7.2%	0.3%	3.7%
Paper white (crumpled, shredded)	0.9%	5.5%	**63.3%**	0.3%	0.0%	2.7%	4.1%	11.9%	0.1%	11.1%
Paper brown filler	0.0%	0.0%	0.1%	**83.9%**	0.2%	0.2%	0.0%	4.7%	8.3%	2.7%
Paper tissue red	0.0%	0.0%	0.3%	0.9%	**82.4%**	0.6%	0.2%	5.3%	6.5%	3.7%
Paper tissue (white, hygienic)	1.0%	0.6%	4.0%	3.6%	0.5%	**76.4%**	0.8%	4.2%	5.6%	3.3%
Paper magazine	0.8%	4.6%	1.3%	0.1%	0.0%	0.2%	**43.4%**	10.7%	0.2%	38.8%
Paper newspaper	0.1%	2.3%	1.9%	0.9%	0.6%	1.1%	1.6%	**74.4%**	1.5%	15.6%
Cardboard corrugated	0.0%	0.0%	0.0%	0.3%	1.2%	0.2%	0.0%	1.4%	**95.3%**	1.5%
Cardboard household	0.2%	0.3%	0.6%	0.6%	0.2%	0.4%	1.4%	7.6%	1.1%	**87.5%**

The results obtained when training with coloured and greyscale patches were directly compared, by looking at the accuracy for each subclass (Table 4-12). Some classes performed better with colour patches, with paper tissue (white, hygienic) being the most notable example, achieving 20.1% higher accuracy than with greyscale patches. When converting to greyscale, the biggest change in accuracy was observed on the paper brown filler subclass, which increased by a relative 150.2% when using greyscale instead of coloured patches.

Table 4-12. Comparison of accuracies per class, obtained when training the network using 10 classes of materials, either with colour or greyscale versions of the images, to classify them into the same 10 classes.

2 CLASSES	10 SUBCLASSES	ACCURACY WHEN TRAINING WITH COLOUR	ACCURACY WHEN TRAINING WITH GREYSCALE
Paper	White	30.5%	**75.2%**
	White (printed)	46.6%	**64.7%**
	White (crumpled, shredded)	61.7%	**63.3%**
	Brown filler	33.5%	**83.9%**
	Tissue (red)	**93.8%**	82.4%
	Tissue (white, hygienic)	**91.7%**	76.4%
	Magazine	**44.7%**	43.4%
	Newspaper	72.1%	**74.4%**
Cardboard	Corrugated	91.1%	**95.3%**
	Household	**94.8%**	87.5%

When classifying the greyscale test images into only 2 main class types, paper or cardboard, the accuracy obtained was 87.1% (SD=0.7%). This result is not significantly different to the 87.3% rate obtained in Experiment 1 using 10 training classes (Table 4-8). The accuracy rates when classifying into 2 classes were significantly higher, for both colour and greyscale images, than the classification rate with 10 subclasses.

4.3.3 Experiment 3: Effects of changes in the hue channel

This section presents the results of using the trained networks from Experiment 2, using 10 material types, to classify the test samples after their hue channel was altered. The resulting accuracy ranged from 36.0% (SD=1.0%) when the hue was shifted by -130 degrees, to 75.0% (SD=1.3%) when no hue shifting was applied, with an average of 47.0% (SD=1.0%). The accuracy was reduced between 25.5% to 30% with hue changes from -30 degree to 30 degrees, which was significantly lower than the results obtained with no hue changes (Figure 4-2).

When classifying the greyscale test images, the accuracy ranged from 66.7%% (SD=0.7%), when the hue was shifted by -170 degrees, to 76.4% (SD=0.4%) when the hue was shifted by 10 degrees, with an average of 72.9% (SD=0.2%). The accuracy results obtained with hue changes from -30 degree to 30 degrees changes were not significantly different to the results obtained with no hue changes.

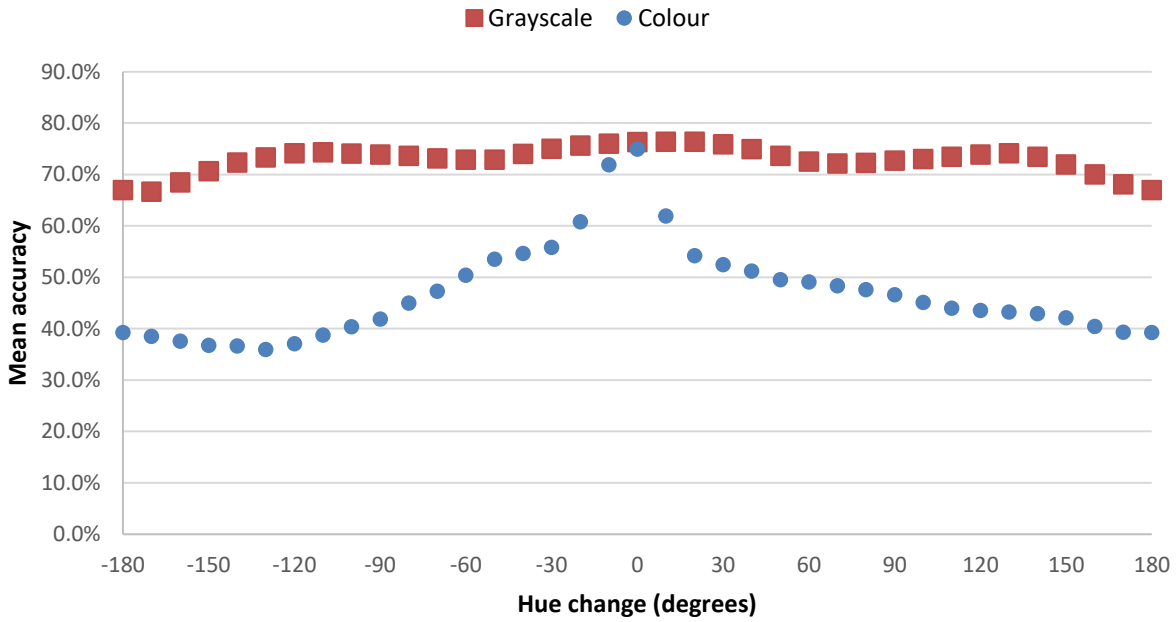


Figure 4-2. Comparison of the effects of changing the hue channel when testing coloured or converted-to-greyscale images. Standard deviations (not shown due to scale) ranged from 0.1% to 2.5% and from 0.1% to 0.7% for colour and greyscale images, respectively.

4.3.5 Experiment 4: Effects of changes in the saturation channel

This section presents the results of using the trained networks from Experiment 2, using 10 material types, to classify the test samples after their saturation channel was altered. The average saturation value for all the test images was 15%. After classifying the coloured test images, the accuracy ranged from 66.1% (SD=1.6%) when the saturation level was reduced by 50%, to 75.3% (SD=1.1%) when saturation was decreased by 10% (Figure 4-3), with an average of 71.3% (SD=1.4%). Reducing the saturation by 20% or more significantly reduced the accuracy of the model when working with colour images (Figure 4-3).

When classifying the greyscale test images, the accuracy rate ranged from 75.5% (SD=0.2%), when the saturation level was reduced by 50%, to 76.6% (SD=0.2%) when saturation was decreased by 10%, with an average of 76.1% (SD=0.3%). Reducing the saturation by 50% presented the only significant reduction in accuracy, as the other results show no significant differences.

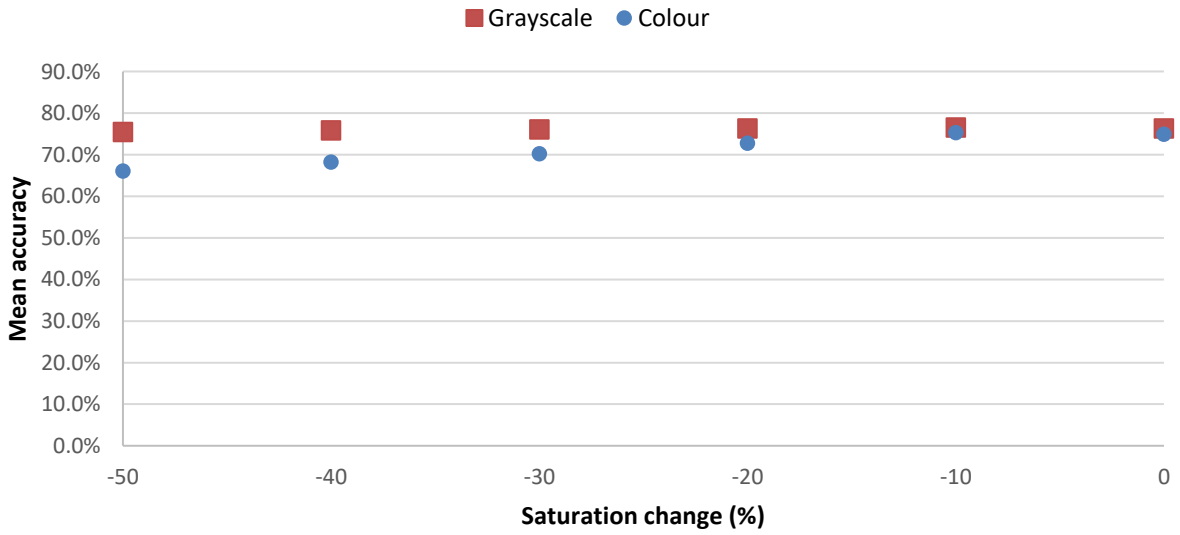


Figure 4-3. Comparison of the effects of changing the saturation channel when testing coloured or converted-to-greyscale images. Standard deviations (not shown due to scale) ranged from 1.1% to 1.6%, and from 0.2% to 0.3%, for colour and greyscale images, respectively.

4.3.7 Experiment 5: Effects of changes in the brightness channel

This section presents the results of using the trained networks from Experiment 2, using 10 material types, to classify the test samples after their brightness channel was altered. The average brightness value of the test images was 46%. After classifying the coloured test images, the accuracy rate ranged from 53.1% (SD=1.5%), when the brightness level was increased by 50%, to 75.8% (SD=1.8%) when brightness was decreased by 10%, with an average of 66.8% (SD=1.9%) (Figure 4-4).

When classifying the greyscale test images, the accuracy rate ranged from 50.3% (SD=1.0%), when the brightness level was decreased by 50%, to 76.4% (SD=0.5%) when brightness was increased by 10%, with an average of 66.8% (SD=0.1%).

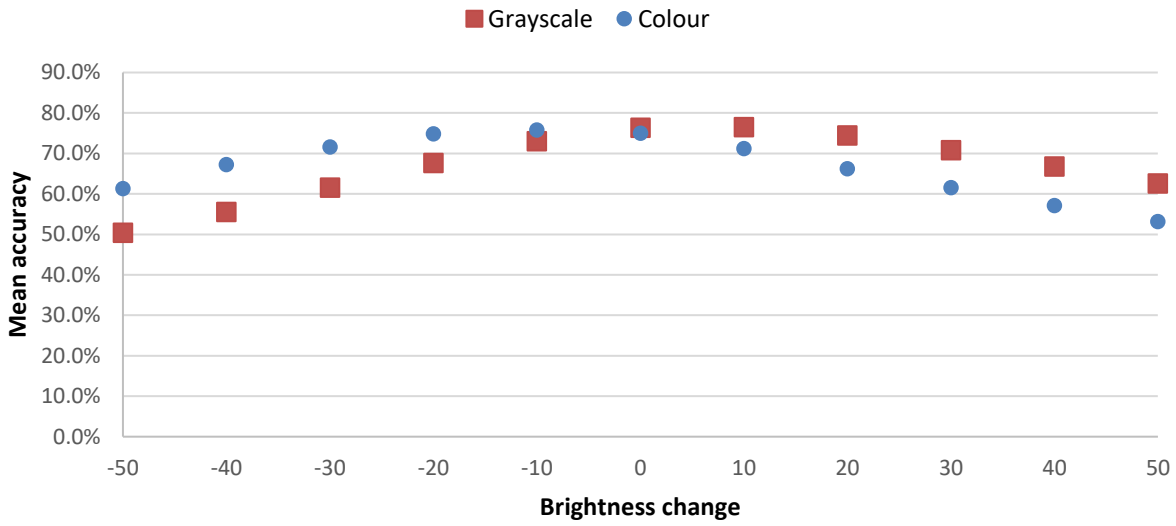


Figure 4-4. Comparison of the effects of changing the brightness channel when testing coloured or converted-to-greyscale images. Standard deviations (not shown due to scale) ranged from 0.2% to 0.7% and from 0.1% to 1.2% for colour and greyscale images, respectively.

4.3.8 Experiment 6: Effects of the number of training patches

This first part of this experiment shows the results of several tests, each one using a different number of training patches in each of 3 material classes: white paper, newspaper, and corrugated cardboard. A main database was created using different minimum cosine distances between the patches in each category, set after trial and

error to generate at least 150,000 samples in each class. The cosine distances used were 0.005 for white paper, 0.05 for newspaper, and 0.1 for corrugated cardboard, which created 218,440, 480,761 and 156,088 patches, respectively.

The main database was then used to generate 8 different, balanced training databases using different numbers of patches per class. The minimum number of patches per class tested was 230, which is the average minimum number of training samples in any of the classes created in experiments 1-5 using a cosine distance of 0.2. The other seven databases were generated with increasing numbers of patches per class: 1,000, 3,000, 4,000, 5,000, 10,000, 50,000 and finally 150,000, which is the rounded-up maximum number of patches that could be obtained from every class in the main database.

Training with 230 patches in each class resulted in a classification rate of 26.8% (SD=9.6%). Increasing the number of training patches per class to 4,000 resulted in a classification rate of 34.1% (SD=1.2%), which represented an improvement of 50.0%. However, increasing by a further 25%, from 4,000 to 5,000 patches per class resulted in an accuracy rate of 93.2% (SD=1.1%), equivalent to a relative increase of 173.1% (Figure 4-5).

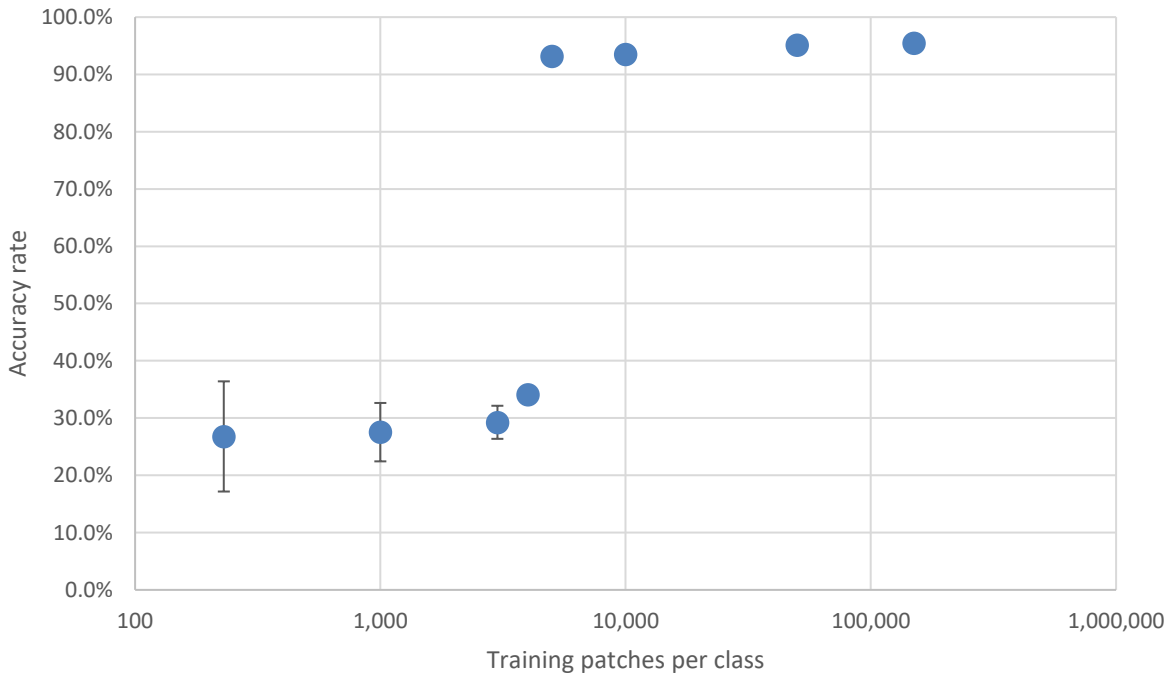


Figure 4-5. Effects of different number of training patches, on a log-10 scale. Vertical bars indicate standard deviation (note: some vertical bars are not visible due to scale).

Further increments of the number of training patches per samples produced relatively smaller increases of mean accuracies. To achieve an improvement of at least 1% in the classification rate, the number of patches per class had to be increased from 5,000 to 50,000, obtaining an accuracy rate of 95.1% (SD=0.4%). The best result was obtained using 150,000 training patches per sample, with an accuracy rate of 95.4% (SD=0.3%). The accuracies per subclass, when using this best result using 150,000 training patches, ranged from 91.09% (SD=1.95%) for newspaper to 98.23% (SD=0.94%) for white paper (Table 4-13).

Table 4-13. Confusion matrix presenting the tests results after training the network with 150,000 patches per class.

		Predicted material		
		White paper	Newspaper	Cardboard
True material	White paper	**98.2% (12,677)**	1.7% (223)	0.0% (6)
	Newspaper	6.1% (1,577)	**91.1% (23,379)**	2.8% (710)
	Cardboard	0.5% (94)	2.0% (394)	**97.5% (18,860)**

To validate this result, the test in experiment 2 was repeated, this time using a balanced database with 5,000 patches in each of the 10 classes using greyscale images. The total number of patches for the balanced dataset was 50,000, compared to 524,035 patches in the unbalanced database. The accuracies per class, when trained with a balanced dataset, ranged from 49.9% (*Paper Magazine*) to 99.8% (*paper white*) (Table 4-14).

The average per-class accuracy was 74.8% (SD=0.4%), which is better but not significantly different than the 74.6% obtained in Experiment 2 using the imbalanced dataset. A class-by-class comparison shows most classes performing better when using the balanced training dataset (Table 4-15).

Table 4-14 - Confusion matrix when using a model trained with a balanced training dataset, with 5,000 patches per class, to classify 10 classes. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results showing the numbers of patches can be seen in Table A-3.

	Paper white	Paper white printed	Paper white (crumpled, shredded)	Paper brown filler	Paper tissue red	Paper tissue (white, hygienic)	Paper magazine	Paper newspaper	Cardboard corrugated	Cardboard household
Paper white	**99.8%**	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Paper white printed	15.0%	**73.9%**	0.0%	0.0%	0.0%	0.2%	10.3%	0.3%	0.1%	0.3%
Paper white (crumpled, shredded)	4.7%	9.4%	**59.9%**	0.8%	0.2%	6.2%	6.5%	6.1%	0.1%	6.1%
Paper brown filler	0.0%	0.0%	0.2%	**90.7%**	1.2%	1.4%	0.1%	1.2%	4.1%	1.1%
Paper tissue red	0.0%	0.0%	0.0%	1.8%	**92.7%**	1.1%	0.1%	0.8%	2.9%	0.5%
Paper tissue (white, hygienic)	6.2%	0.3%	2.5%	2.8%	2.4%	**79.9%**	0.9%	1.4%	2.8%	1.0%
Paper magazine	2.4%	15.2%	2.3%	0.4%	0.1%	0.3%	**49.8%**	8.8%	0.2%	20.6%
Paper newspaper	0.9%	14.1%	3.3%	2.4%	1.5%	3.5%	4.5%	**54.4%**	1.2%	14.2%
Cardboard corrugated	0.4%	0.0%	0.0%	2.0%	8.4%	3.7%	0.5%	0.7%	**83.7%**	0.6%
Cardboard household	1.9%	2.4%	2.6%	3.0%	1.3%	2.9%	8.2%	12.8%	1.9%	**62.9%**

However, creating the balanced database with 5,000 patches per class and training the network took an average time of 17hr 22min, compared to 33hr 47min required in Experiment 2 for the unbalanced dataset. Similarly, the storage requirements for the database were reduced from the 98GB used in Experiment 2, to 9.35GB when using 5,000 patches per class in this experiment.

Table 4-15. Accuracies obtained when training the network using balanced and unbalanced databases of 10 classes of materials, with greyscale versions of the images, to classify them into the same 10 classes.

CLASS	SUBCLASS	UNBALANCED DATABASE		BALANCED DATABASE	
		Training samples	Accuracy per class	Training samples	Accuracy per class
Paper	White	230	75.2%	5,000	99.8%
	White (printed)	3,718	64.7%	5,000	73.9%
	White (crumpled, shredded)	48,879	63.3%	5,000	59.9%
	Brown filler	24,756	83.9%	5,000	90.7%
	Tissue (red)	2,339	82.4%	5,000	92.7%
	Tissue (white, hygienic)	7,754	76.4%	5,000	79.9%
	Magazine	29,364	43.4%	5,000	49.9%
	Newspaper	154,737	74.4%	5,000	54.4%
Cardboard	Corrugated	29,694	95.3%	5,000	83.7%
	Household	222,564	87.5%	5,000	62.9%

4.3.9 Experiment 7: Paper grade identification

The networks trained in the first part of Experiment 6, using 5,000 patches per class to classify 3 classes of materials, were used to identify a set of 360 validation images. This dataset included images of plain and printed white papers, newspapers and corrugated cardboards that were not used before for neither training nor testing. Each image is assumed to contain only 1 type of material which is determined by a majority vote from each patch predicted class. The test resulted in an overall success rate of 98.3%. The confusion matrix indicates that all the cardboard and newspaper images were correctly identified, representing a 100% success rate in both cases (Table 4-16).

Table 4-16. Confusion matrix presenting the validation results, using the network trained with 5,000 patches per class to identify 3 classes of material samples. The percentages and colour shades indicate the proportion of patches predicted as a given material (column) relative to the total number of analysed patches for that row. The percentages in the diagonal (bold) indicate the correct predictions for each material type. The results in parenthesis indicate numbers of patches.

		Predicted material		
		White paper	Newspaper	Cardboard
True material	White paper	**96.7% (174)**	3.3% (6)	0.0% (0)
	Newspaper	0.0% (0)	**100.0% (90)**	0.0% (0)
	Cardboard	0.0% (0)	0.0% (0)	**100.0% (0)**

The recognition rate for images of white paper was 96.7% (SD=0.5%). A deeper analysis show that the 90 images of plain white paper were 100% correctly identified whilst the classification rate for the 90 printed white paper was 93.3% (SD=0.5%). The 6 classification errors consisted of images of printed white paper being classified as newspaper, as the one shown in (Figure 4-6). In this example, 87 patches were analysed and only 29 were classified as white paper, whilst the rest were erroneously classified as newspaper.

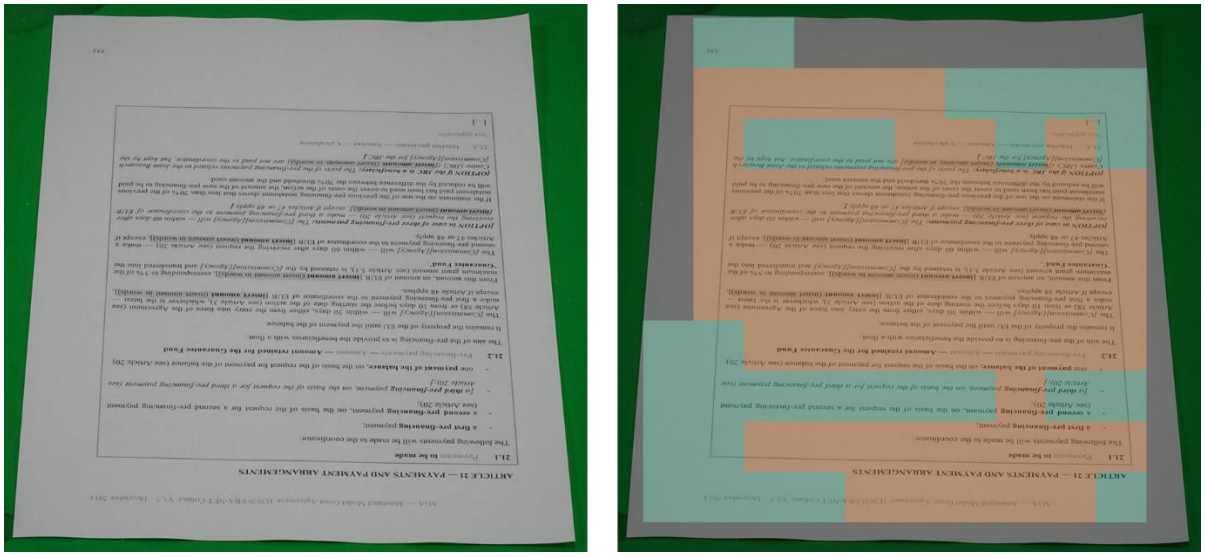


Figure 4-6. Image of a printed white paper sample incorrectly classified as newspaper. Patches classified as white paper are green and newspaper patches are orange.

4.4 Discussion

The results from Experiment 1 when training and classifying with 2 categories, paper and cardboard, show an accuracy of 76.8% which is significantly lower than 83.7% achieved in previous results (Chapter 3). This can be explained with the smaller number of training material samples used for this experiment (23 vs 123), which generates a training dataset with less representation of each class. As a result, and even though the number of training patches was higher (569,167 vs 16,000) the model in this experiment fails to generalise as well as the model trained with the higher number of material samples. On the other hand, the results compare favourably to the application of AlexNet to other tasks, which achieved an accuracy of 62.5% (top-1 error=37.5%) when classifying images into 1,000 classes with a training dataset with 1.2 million images (Krizhevsky et al., 2012).

The classification performance improved significantly when subdividing those 2 categories into 10 different subclasses, achieving an accuracy of 87.3%, significantly higher than the 83.7% accuracy obtained previously (Chapter 3). These results, which were consistent for each of the 10 subclasses, indicate that the proposed methodology benefits from using more classes, with each class containing more closely similar images, thus reducing the intra-class variability. This is reflected in the achieved cosine similarity of 0.74, significantly higher than 0.60 for the 2-class model. It is possible that other network architectures, such as GoogleNet and VGG-16, could perform better with high-intraclass variability, so the relative benefits could be reduced. However, those models are more complex which translates into more computations and longer prediction times.

Using greyscale images, the accuracy for 10 classes of 77.4% was lower than the 81.4% achieved with AlexNet for a material classification task (Bell et al., 2015). This can be explained by the difference in the training datasets. First, the number of training patches was 6 times smaller in this experiment (total of 523,945, with an average of 52,394 per class, vs a total of 2,996,674, with an average of 130,290 per class). Second, the experiments in this study used a more unbalanced training dataset (ratio of largest to the smallest class of 968, while the other study used a more balanced training dataset with a ratio of 38). Both factors have been demonstrated to affect performance (Bell et al., 2015), and it is expected that an increase in accuracy could

be achieved by increasing the number of patches per class in a balanced dataset. This was demonstrated in this study in Experiment 6, using a balanced dataset with 150,000 patches per class, to obtain an accuracy of 95.4% when classifying 3 material classes.

The results obtained in Experiment 2, when training and testing greyscale images indicate that the deep convolutional network can recognize materials without any colour content. This implies that the information in the images regarding surface features, such as shadows and differences in contrast levels, are enough to recognize materials. In general, using colour information did not achieve a significant improvement, which means that colour does not add any useful information. This contrasts with the results obtained using traditional, hand-crafted feature extractors to classify surfaces, which showed a relative increase of 8.6% in performance when colour information is included in the classification (Sharan et al., 2013). This would implicate that the CNN is able to extract more information from the material surfaces than a colour-based hand-crafted feature classifier. The use of CNN reduces the reliance on colour as a distinctive feature, which makes the solution more robust to material degradations affecting colour properties.

However, even though the accuracy rates were significantly better with greyscale than using colour, the benefit did not apply to all the 10 different classes. Whilst the recognition rate of plain white paper was increased more than twofold, the identification of magazine paper remained at a rate lower than 50%, being confused mostly with household cardboard material. Based on the previous results, this low classification rate could be caused by the high intra-class variation in the magazine paper and household cardboard classes. Household cardboard material, such as cereal boxes and biscuit packaging, usually include an external fibre layer with graphics and text, which looks very similar to some magazine contents. Increasing the training dataset, in particular with more examples of household cardboard boxes, would reduce this confusion by increasing the representation of this class which has high variability.

As expected, changes in the colour (hue, saturation, and brightness) of the test images negatively affected the performance of the trained networks, but the effect is lessened when converting them image to greyscale. Changing the hue In Experiment 3, which

shifted the actual colour of the images before being tested, had the largest negative effect on the colour-based method. This again indicates that colour is one important feature being learned by the networks. One potential way to overcome this issue would be to augment the training dataset with variations of the original images using different shifts of hue. However, this would increase the computation and storage requirements.

The presented alternative of removing the colour information, however, made the trained network more resilient to hue changes of the test images. The worst effect, after changing completely the colour of the images as in a worst-case scenario, was a relative decrease in accuracy of 12%. Minor changes of the hue, on the other hand, resulted in minimum degradation of the performance, achieving accuracies not significantly different than without hue changes. This compares favourably to working with colour images, where even minor hue changes resulted in significant reductions of accuracy of up to 30%. Furthermore, removing the colour information did not impact the accuracy on the original images, achieving better accuracy with greyscale images than with colour images, even if the hue was not changed.

Similarly, in Experiment 4, using greyscale images produced better classification results than using colour images when testing images with lower saturation levels. The results when working with greyscale levels showed no significant effects after 40% reductions of the saturation level, whilst saturation changes over 10% can significantly reduce the accuracy when working with colour images.

The changes in brightness levels studied in Experiment 5 negatively affected both the colour and greyscale versions. The colour method was more resilient to reductions in brightness levels, but the opposite was observed with greyscale images. This can be explained by the colour to greyscale conversion formula used, which generated greyscale images with 20% lower brightness levels than the original colour version. If the charts are shifted by 20%, so that a greyscale change of +30% matches a change of +10% of the colour images, the results would not show any significant difference between colour and greyscale predictions.

This leaves more overhead in the greyscale images for increasing the brightness without clipping the intensity level and preserving the details. Therefore, the greyscale images are more resilient to increases in brightness than the colour version. On the other hand, the colour images, with a higher brightness level, are more resilient to

reductions in brightness than the greyscale version, as more details are preserved in the darker regions of the image.

In both cases, colour and greyscale, the results could be partially improved by ensuring that the test images are captured with the same brightness level as the training images. This could be ensured by using enough illumination combined with the camera automatic exposure level detection. Alternatively, augmenting the training image dataset with different brightness levels would reduce the sensitivity to these changes. However, a factor to consider when doing this is the dynamic range of the images, which is the ratio of the brightest to the darkest areas of an image. Due to the nature of the materials, this range may exceed several orders of magnitude.

Experiment 6 showed the important effect of the number of training patches per class on the resulting performance. Even though increasing the number of training images is expected to increase the performance of the model, the results obtained in this experiment showed that using a balanced training dataset with 5,000 patches in each class resulted in similar accuracy rates to using 150,000 patches. However, using a balanced dataset with 4,000 patches per class showed a pronounced reduction in performance. This suggests that, when using a balanced training dataset, the network needs a minimum number of 5,000 patches per class to improve its training and therefore, its resulting classification rate. The reasons for this significant increase of performance between 4,000 patches and 5,000 patches per class remain unclear and has been identified as an area of future research.

This was validated by training a network using 10 different classes of materials with 5,000 patches per class to obtain a similar performance to a non-balanced, larger training database. This suggests that the minimum number of patches is not directly affected by the number of classes. However, the classification performance for some materials was negatively affected, such as newspaper and household cardboard. These materials have a high intra-class variability, with samples visibly different to each other, thus they might require more training samples per class to achieve better classification performance. As expected, using a balanced dataset with less patches per class required less computation and storage for training purposes.

In Experiment 7, the balanced training dataset with 5,000 patches from greyscale images for the classification of 3 materials from Experiment 6 was applied to the

identification of paper material of different grades. The obtained identification rates were higher than those reported by industry (Remade Scotland, 2005) and by previous research studies (Rahman et al., 2014; Tatzer et al., 2005). The only item to be incorrectly identified was a sheet of printed white paper, that was confused with newspaper when the sample had large sections of printed areas. The performance could be improved by handling the edges of the images differently, for example by mirroring or extending the edges instead of discarding the edge patches. As the edges of printed white paper are typically blank, they would be correctly recognized, and the chances of a correct identification of the sample would therefore be increased. This is a very important results because this methodology could be deployed to assist or replace current manual quality assurance methods (INGEDE, 2009). A solution like the proposed could be deployed, for example, to mobile phones, so that a quality assurance officer could take a picture of a mixed waste load to be automatically analysed, instead of manually counting the items per class. This could speed up the quality assurance process and produce better, non-subjective composition analysis, while at the same time keeping the evidence data in the form of pictures for added transparency.

4.5 Conclusions

The accuracy rates when sorting the materials into paper and cardboard improved significantly by subdividing those categories into more, closely related classes. Similar classification rates can be obtained by using a balanced training dataset with a reduced number of images per class than a highly unbalanced dataset with a larger total number of samples. This study showed how this also translates into using less computation time and storage resources for network training. However, further studies are required to adjust the number, and size, of the material classes based on their respective intra-class variability to achieve optimal, balanced training datasets.

This study shows that solid waste material recognition using deep neural networks perform better after removing colour information from the images. When converted to greyscale format, the images can be used for training a deep network to successfully recognize materials based on the surface features captured with multiple cameras and multiple light sources. This would reduce the cost of the hardware, as greyscale sensors are typically cheaper than RGB cameras.

The method becomes more robust when using greyscale images, reducing the negative impact from changes in the colour properties of the materials. This mean that the solution would be more reliable when deployed in the field. For example, it would not be affected as much by changes in the hue that could be expected from different packaging designs for the same products. Similarly, the trained model could generalise to the introduction of new products using similar packaging than existing branded product. The model would be resilient to colour ink transfer, affecting packaging colours, during waste storage and transport. Similarly, it would not be affected by the reduction in colour saturation that happens over time when waste is stored outdoors and exposed to the sun.

This study also showed the negative effects that changing brightness can have on performance. This is of particular importance when deploying a solution in the field, as the illumination could vary during a day and depending on the annual season. This would change the brightness of images of waste materials captured next to a window in the facility, with direct or indirect sunlight hitting the field of view of the camera. In this case, the recommendation would be to use automatic exposure in the camera, which would reduce the camera gain and the exposure time so that the waste is capture with the correct exposure level.

The results obtained in this study expand the original methodology, to generalise the neural network training not only regarding the free-angle rotation of the waste material, but to become less dependent of their colour. The multiple illuminations not only generate multiple different training images from a single material but ensure the images have enough surface features to be recognized in greyscale format. Remaining challenges and issues were acknowledged to cause false recognition of materials, and potential solutions involving simple image processing techniques were proposed.

Despite the identified limitations, this study showed that this methodology can currently achieve top identification rates for different grades of paper material. The technique could be extended to classify more types of paper materials, and effectively implemented in the paper recycling industry for the determination of paper grades according to the standard guidelines.

4.6 References

Adamcová, D., Vaverková, M., 2014. Degradation of Biodegradable/Degradable Plastics in Municipal Solid-Waste Landfill. Polish Journal of Environmental Studies 23, 1071–1078.

Bell, S., Bala, K., 2015. Learning visual similarity for product design with convolutional neural networks. ACM Transactions on Graphics 34, 98:1-98:10. https://doi.org/10.1145/2766959

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Bello-Cerezo, R., Bianconi, F., Cascianelli, S., Fravolini, M.L., Maria, F. di, Smeraldi, F., 2017. Hand-Designed Local Image Descriptors vs. Off-the-Shelf CNN-Based Features for Texture Classification: An Experimental Comparison, in: Intelligent Interactive Multimedia Systems and Services 2017, Smart Innovation, Systems and Technologies. Presented at the International Conference on Intelligent Interactive Multimedia Systems and Services, Springer, Cham, pp. 1–10. https://doi.org/10.1007/978-3-319-59480-4_1

Bian, P., Li, W., Jin, Y., Zhi, R., 2018. Ensemble feature learning for material recognition with convolutional neural networks. EURASIP Journal on Image and Video Processing 2018, 64. https://doi.org/10.1186/s13640-018-0300-z

British Standards Institution, 2014. BS EN 643:2014 Paper and board. European list of standard grades of paper and board for recycling. BSI, London, UK.

Buda, M., Maki, A., Mazurowski, M.A., 2017. A systematic study of the class imbalance problem in convolutional neural networks. arXiv:1710.05381 [cs, stat].

Canopoli, L., Fidalgo, B., Coulon, F., Wagland, S.T., 2018. Physico-chemical properties of excavated plastic from landfill mining and current recycling routes. Waste Management 76, 55–67. https://doi.org/10.1016/j.wasman.2018.03.043

Cavalin, P., Oliveira, L.S., 2017. A Review of Texture Classification Methods and Databases, in: 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images

Tutorials (SIBGRAPI-T). Presented at the 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), pp. 1–8. https://doi.org/10.1109/SIBGRAPI-T.2017.10

CEPI, 2013. European List of Standard Grades of Paper and Board for Recycling [WWW Document]. URL http://www.cepi.org/system/files/public/documents/publications/recycling/2013/CEPI_EN%20643_brochure_FINAL.pdf (accessed 4.15.16).

Cimpoi, M., Maji, S., Kokkinos, I., Vedaldi, A., 2016. Deep Filter Banks for Texture Recognition, Description, and Segmentation. International Journal of Computer Vision 1–30. https://doi.org/10.1007/s11263-015-0872-3

EU Council, 2013. COST Action E48 – The Future of Paper Recycling in Europe: Opportunities and Limitations [WWW Document]. URL https://ec.europa.eu/growth/tools-databases/eip-raw-materials/en/community/document/cost-action-48-final-report-future-paper-recycling-opportunities-and-limitations (accessed 8.7.18).

Food and Agriculture Organization of the United Nations, 2018. FAOSTAT Statistical Database [WWW Document]. URL http://www.fao.org/faostat/en/#data/FO (accessed 10.10.19).

Gonzalez, R.C., Woods, R.E., 2001. Digital Image Processing: United States Edition, 2nd edition. ed. Pearson, Upper Saddle River, N.J.

Grandini, M., Bagli, E., Visani, G., 2020. Metrics for Multi-Class Classification: an Overview. arXiv:2008.05756 [cs, stat].

Hu, D., Bo, L., Ren, X., 2011. Toward robust material recognition for everyday objects, in: Hoey, J., McKenna, S., Trucco, E. (Eds.), Proceedings of the British Machine Vision Conference 2011. British Machine Vision Association, BMVA, Dundee, pp. 1–11. https://doi.org/10.5244/C.25

INGEDE, 2009. INGEDE Method 7 - Visual Inspection for Recovered Paper for Deinking, Unbaled Delivery [WWW Document]. URL http://pub.ingede.com/en/methods/ (accessed 11.8.21).

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv:1408.5093 [cs].

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

Kylberg, G., Sintorn, I.-M., 2016. On the influence of interpolation method on rotation invariance in texture recognition. EURASIP Journal on Image and Video Processing 2016. https://doi.org/10.1186/s13640-016-0117-6

Liu, C., Sharan, L., Adelson, E.H., Rosenholtz, R., 2010. Exploring features in a Bayesian framework for material recognition, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Presented at the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 239–246. https://doi.org/10.1109/CVPR.2010.5540207

OpenCV library 3.1.0 [WWW Document], 2015. URL https://docs.opencv.org/3.1.0/ (accessed 5.2.18).

Rahman, M.O., Hussain, A., Basri, H., 2014. A critical review on waste paper sorting techniques. International Journal of Environmental Science and Technology 11, 551–564. https://doi.org/10.1007/s13762-013-0222-3

Remade Scotland, 2005. Business Case for Automated Paper Sorting [WWW Document]. URL http://www.remade.org.uk/reports/materials/paper/business-case-for-automated-paper-sorting.aspx (accessed 8.8.18).

Schwartz, G., Nishino, K., 2015. Automatically discovering local visual material attributes. Conference on Computer Vision and Pattern Recognition (CVPR) 07-12-June, 3565–3573. https://doi.org/10.1109/CVPR.2015.7298979

Sharan, L., Liu, C., Rosenholtz, R., Adelson, E.H., 2013. Recognizing Materials Using Perceptually Inspired Features. Int J Comput Vis 103, 348–371. https://doi.org/10.1007/s11263-013-0609-0

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Information Processing & Management 45, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Tatzer, P., Wolf, M., Panner, T., 2005. Industrial application for inline material sorting using hyperspectral imaging in the NIR range. Real-Time Imaging 11, 99–107. https://doi.org/10.1016/j.rti.2005.04.003

Velis, C.A., Wagland, S., Longhurst, P., Robson, B., Sinfield, K., Wise, S., Pollard, S., 2013. Solid recovered fuel: materials flow analysis and fuel property development during the mechanical processing of biodried waste. Environmental science & technology 47, 2957–65. https://doi.org/10.1021/es3021815

Vrancken, C., Longhurst, P., Wagland, S., 2017. Critical review of real-time methods for solid waste characterisation: Informing material recovery and fuel production. Waste Management 61, 40–57. https://doi.org/10.1016/j.wasman.2017.01.019

CHAPTER 5

MULTIPLE SPATIAL SCALES

This chapter provided a qualitative analysis of the neural network decision process and extended the methodology to improve the classification of images captured with different spatial resolutions

5 Multi-scale training of deep networks for the classification of waste material

Abstract

There is a need for advanced tools to monitor solid waste material to improve their classification and sorting rates. This study extends a previously developed methodology for the classification of images of waste materials captured with visible sensor cameras and deep learning. The deployment of this method in material recovery facilities could result in images captured at different spatial resolutions. Classifying images with resolutions different to the images in the training datasets were shown to reduce the classification rates by up to 80%. A set of known techniques are applied here to combine images of different spatial resolutions to augment the training datasets. The resulting neural networks achieved an average accuracy per class of 88.5% over the range of resolutions that could be expected from an installation in a material recovery facility. The results were validated using a qualitative analysis to visualise and understand how the model reaches its classification decisions.

5.1 Introduction

There is a trend in government initiatives and policies to improve waste management (EU Commission, 2018; EU Council, 2009, 1999; EU Parliament, 2018, 2017). These are set out to minimise waste material going into landfills thus reducing levels of pollution, with high environmental and economic impact, in particular in developing countries (McKinsey & Company, 2016). In the UK, the plan is to incentivise producers to improve their products designs and to promote the use of recovered materials in secondary markets (DEFRA, 2018a).

High quality data to provide key information to stakeholders is needed for monitoring purposes and it is highlighted as an indicator of progress (DEFRA, 2018b), consistent with discussions and insight gained from recent studies (Laner et al., 2016; Rechberger et al., 2014; Vrancken et al., 2017). Access to real-time

information regarding waste material quantities and qualities, to monitor and evaluate the different processes and available resources, can provide useful control mechanisms whilst assisting future policy making (Vrancken et al., 2017). Funding and other research and development mechanisms are being provided to create innovative solutions and to improve existing methods (DEFRA, 2018b).

Waste materials from household, commercial and industrial sources are typically transported to material recovery facilities (MRF) (WRAP, 2006), where they are sorted and processed through a series of mechanical and biological treatments. Inside an MRF, the materials flow from one process to another on top of a system of conveyor belts (Tchobanoglous and Kreith, 2002). The equipment in a MRF is used to feed, condition, sort and refine the materials, either manually or automatically, through physical or visual properties of the different items. Advanced optical sorting units using near-infrared (NIR) sensors can be used to identify, quantify and separate certain materials, but these machines can be costly to install and run, in particular for MRF with low volumes (Cimpan et al., 2016).

Previous research suggested using visible cameras combined with deep learning as an alternative method to identify and quantify the waste material in the different streams (Chapters 3 and 4). Waste materials present specific challenges compared to typical material classification because they offer no global context information, such as object shapes, or regular orientations. As a result, the materials must be identified purely on their local visible characteristics, which can be a disadvantage and lower their rates of classification (Schwartz and Nishino, 2016). Nonetheless, recent results indicate that a deep network trained only with images of materials can result in average accuracy rates of 74.8% or higher (Chapters 3 and 4), whilst being resilient to different rotations and changes in colours.

The deployed cameras could be positioned at different places inside an MRF, to visualize waste material at different stages with varying composition. For example, a camera might be looking at all the waste at the tipping station to get overall information of the incoming waste. Similarly, a camera could look at the

stream of waste materials going into the first feeder, detecting large fractions of contaminants, whilst other cameras might be looking at the light and heavy fractions after the stream is separated by an air sifter, or a camera might be analysing the output of the shredder before waste is processed by a baler. The waste composition information at different stages could be used to optimise the processes and improve the overall performance of the MRF (Vrancken et al., 2017).

The diverse positioning and configuration options of the camera could influence the actual visible area or field of view. Given the camera imaging sensor resolution, the spatial resolution is measured by the number of pixels per inch of visible area. Given a camera sensor, the field of view and the resulting spatial resolution are dictated by the existing facility operating setup, depending mostly on the conveyor belt width. The width of conveyor belts varies throughout a facility, from wide 2.4 metre belts at the input of the facility to narrow 600mm belts at the output stages. In other cases, the field of view, determined by the configuration of the camera sensor and lens and their respective positioning towards the target could be arbitrarily chosen. For example, a given camera and focal length could be used to look at a conveyor belt and then moved higher up to look at a large waste tipping hall. Similarly, different cameras with higher pixel resolutions can be used to capture images of the same field of view.

The spatial resolution might also change over time, for example if the camera is mounted to look at accumulated waste streams in tipping halls before they are fed into the processing plant. These heaps of waste will change in height, which would result in the object being closer or farther to a fixed top-mounted camera. Another example would be if a person is manually operating the camera to take pictures of the materials at different stages. When using low spatial resolution to visualize a material, its smaller features might not be represented. On the other hand, when using high spatial resolution to represent the same material, large scale features may be visualized. The material features visualised with different resolutions are therefore expected to affect the deep network classification performance. In these scenarios, a trained network would have to recognize

materials with a range of spatial resolutions. Furthermore, neural network that can deal with multiple resolutions would make deployment of such technology less complicated, as it would not have to be re-trained to target any specific spatial resolution.

Previous work using deep learning resulted in good classification performances by training with multiple resolutions of images to train separate networks which are then combined to provide a single prediction. Previous research, focused on object recognition in scenes, using limited image size (512 by 512 pixels) and 2 scale variations, obtained 75% accuracy rate (Wu et al., 2015). Another methodology, using multiple scale images to train multiple networks, obtained 80% accuracy rate for the recognition of the styles of 210 artists in artwork images (van Noord and Postma, 2017).

Another approach is scale jittering, which involves training a single deep network with multiple spatial resolutions images (Girshick, 2015; Simonyan and Zisserman, 2014; Szegedy et al., 2015). Trained this way, a deep network learns to recognize features for the same object at different scales whilst improving the performance, obtaining 75.2% accuracy rate (Simonyan and Zisserman, 2014). The result would be a scale-invariant classifier, though it has been found that standard deep networks could fail to develop this property (Ngiam et al., 2010), and would have to be specifically designed to deal with multiple scaled versions of the same feature (Xu et al., 2014).

A trained deep neural network can classify an image with good performance, but it provides no indications of how its internal filters work to reach its decisions. There are several methods that visualise the effects that every pixel has for the classification of the image (Bach et al., 2015; Simonyan et al., 2013; Zeiler and Fergus, 2014). This type of assessment could lead to improvements to the data capture and image processing methodology for better classification performance (Lapuschkin et al., 2017, 2016b; Seibold et al., 2018). The Layer-wise Relevance Propagation (LRP) method (Bach et al., 2015) is the most suitable to explain the network decisions for the classification of images (Samek et al., 2017). The LRP algorithm computes relevance values which are represented in an heatmap

image, where each pixel value represents whether it contributes, opposes, or has no effects towards a certain classification.

The experiments in this study explore scale jittering for deep learning, augmenting the training data with multiple different resolutions of the same material samples. The range of practical spatial or spatial resolutions for different waste management applications are determined and used to set up the different experiments. The accuracy rates of the networks trained with single and multiple scale resolutions are compared and the findings are used to extend a previously developed methodology (Chapters 3 and 4) to train deep networks for the classification of images of waste material captured with different spatial resolutions. The resulting methodology is verified and validated through a qualitative analysis using the LRP method to understand and interpret the model by visualising which areas of the test images are responsible for their classification.

5.2 Materials and methods

In this section, the methodology is described in detail, including the required materials, the steps, and procedures required by the experimental design, and the analysis applied to the data.

5.2.1 Material samples

This study used 65 different samples for training, taken from 21 common waste items, or subclasses, grouped into 7 main classes of materials (Table 5-1). These samples represent solid materials typically found in waste streams (Séverin et al., 2010). A separate set of 30 material samples were used for testing purposes matching the same 21 subclasses used for training. The numbers of samples in each subclass were selected to better represent the materials, accounting for their different intra-class variabilities. A larger number of training and testing samples were used from material types with higher intra-class variability, such as household food and drink cardboard boxes, than those with lower intra-class variability, such as stainless-steel pipes. The classification groups included a

special subclass, number 21, which represented the green textile material used as background during the image capture.

Table 5-1. List of classes used for training and testing purposes.

ID	Main class	Subclass	Training samples	Testing samples
1	Paper	Plain, crumpled and printed office sheets	3	3
2	Paper	Shredded office sheets	1	1
3	Paper	Packaging void fill	2	1
4	Paper	Red tissue	1	1
5	Paper	Hygienic and kitchen towel	2	2
6	Paper	Glossy magazine	3	1
7	Paper	Newspaper	6	2
8	Cardboard	Corrugated	6	2
9	Cardboard	Household food and drink boxes	8	4
10	Plastic	Balls, straws and wall plugs	4	2
11	Plastic	Assorted toys	1	1
12	Plastic	Carrier bags	5	1
13	Plastic	Black food trays	3	1
14	Plastic	Black paint trays	2	1
15	Plastic	White polystyrene foam	3	1
16	Wood	Raw and processed	7	1
17	Textile	Synthetic cloths and towels	2	1
18	Inert	Red bricks	2	1
19	Inert	Stones	2	1
20	Metal	Stainless steel pipes	1	1
21	Background	Green cotton fabric	1	1

5.2.2 Image capture

The images used for training and testing purposes were captured using a setup previously developed (Chapter 3) with 5 visible sensor cameras and 5 light sources. All the cameras and lights are pointing to the target area which is at a focal distance of 1,095 cm. The camera focal length is set to 16 mm and the lens

aperture set to f/8, which results in a field of view of 43 by 29 cm with a depth of field of 394 mm. The camera image resolution was set to 4,896 by 3,264 pixels, resulting in a spatial resolution of 289 pixels per inch (PPI). Each material sample was photographed 30 times, using the combination of 6 illuminations, one per light source plus a flooded illumination using all lights together, and 5 viewing angles, one per camera. The sensitivity of each camera was set to ISO 80 to minimise sensor noise, and the exposure time was determined automatically by the camera algorithm for each capture.

5.2.3 Image processing

Every image was subjected to a series of transformations implemented in C++ (“OpenCV library 3.1.0,” 2015) following strategies developed previously (Chapters 3 and 4). When required by the experiment, the images were down sampled to different sizes, representing different spatial resolutions, using Lanczos resampling over an 8 by 8 pixels neighbourhood (Kylberg and Sintorn, 2016). The training images were rotated around their centre with Lanczos interpolation using an 8 by 8 pixels kernel.

Each rotated image was then divided into patches of 256 by 256 pixels, partially overlapping in each dimension with the neighbouring patches in the case of training, and no overlapping for the testing patches. A hue, saturation, and intensity (HSI) colour-detection method was used to ensure that the patches from classes 1 to 20 containing any pixel with green colour matching the subclass 21 were discarded. The HSI method was also used to extract patches with only green colour matching the background to create the subclass 21. Therefore, the material classes 1 to 20 did not contain any background material, and class 21 contained only background material. The patches were converted into a 3-channel grayscale version with each channel containing the same values.

5.2.4 Deep learning

The deep neural networks were trained and tested using Caffe framework version 1.00 (Jia et al., 2014). The convolutional network architecture was AlexNet (Krizhevsky et al., 2012), with the final layer fc8 modified to identify 21 classes.

Transfer learning was implemented in each training exercise, reusing the weight values obtained with the MINC database (Bell et al., 2015). The Adam optimisation algorithm was used to adjust the learning rate dynamically, with parameters $\alpha = 10^{-5}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ (Kingma and Ba, 2014). The dropout factor was set to 0.5 to reduce overfitting and to allow the network to better generalise its predictions.

A vector of 4,096 features to represent each patch in the training database was calculated by applying an AlexNet network previously trained with the MINC database. The features vector was obtained from the fc7 network layer. These feature vectors were used to calculate the cosine distance between each of the patches in the training database. Patches that were considered too similar were removed from the databases, and the remaining unique patches were used to generate training databases.

Each experiment used a different training dataset of unique patches. Each training dataset contained 5,000 patches per class. Training was conducted for 20 epochs so that every patch in the database was passed through the neural network 20 times. The patches were cropped to 227 by 227 pixels from a random region to prevent replicate input data and reduce overfitting. Every 4 epochs, the accuracy of the trained networks was evaluated with the test dataset, and the training was deemed finished when the test accuracy stopped improving.

5.2.5 Accuracy testing

The performance of the different trained networks was determined through the classification of the patches in the test databases. These databases were generated using patches obtained from images of 30 material samples not included in the training database. The trained network classified each of the test patches and the output for each patch is interpreted as the probability that the patch belongs to each of the 21 classes. The class with the highest probability was considered the predicted material for that patch. These classifications were used to calculate the accuracy rate for each class using equations (5-1) and (5-2). The average accuracy per class was calculated using equation (5-3) to indicate the accuracy of the model on the dataset. The tests were repeated 3 times for

each experiment to calculate the standard error to indicate the precision of the model.

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions} \quad \textbf{(5-1)}$$

$$accuracy_i = recall_i = \frac{tp_i}{tp_i + fn_i} = \frac{number\ of\ correct\ class\ i\ predictions}{total\ number\ of\ class\ i\ examples} \quad \textbf{(5-2)}$$

$$average\ accuracy\ per\ class = \frac{\sum_{i=1}^{n} accuracy_i}{n} \quad \textbf{(5-3)}$$

Where *tp=true positive rate*; *tn=true negative rate*; *fp=false positive rate*; fn=*false negative rate; n=number of classes (20); i=material class (i<=20)*

5.2.6 Interpretation of the deep models

The LRP Toolbox for Artificial Neural Networks (Lapuschkin et al., 2016a) was used to verify and understand the predictions from the trained neural networks. This provided a mathematical interpretation of the neural networks using the Layer-wise Relevance Propagation (LRP) methodology. The LRP method takes an input image and runs it forward through the trained neural network to estimate the contribution of each pixel of the input image to the prediction. The outputs resemble heatmap images with pixel value representing whether it contributes positively or negatively towards a certain classification. In the experiments, heatmaps were generated both for the true class and for the class assigned by the trained network. These heatmap images were visually inspected and analysed to gain understanding of deep learning in the context of solid waste material classification.

5.2.7 Experimental design

The experiments were designed to evaluate the effects of changes in image scale for training and testing purposes. The image resolutions were selected based on use case scenarios assuming cameras sensors with different pixel resolutions. Each scenario assumes the cameras will be set on top of, and looking down to, a typical 1-metre wide conveyor belt transporting the waste material, making sure the entire width is covered by the different field of views.

The same raw training images were downscaled to different resolutions to generate the training datasets used for the different experiments. Using the recommendations from previous results (Chapter 4), balanced datasets with 5,000 patches per class were generated for each experiment. When using higher resolutions, an increasing number of unique patches were extracted from each material sample. In all cases, 5,000 patches were randomly selected from the pool of unique patches to end up with equally balanced databases for all experiments.

Similarly, the raw test images were downscaled to different spatial resolution to generate the test datasets to measure the classification performance of the trained networks. Each test was performed using test patches with spatial resolutions in the range from 25 to 289 PPI. The average accuracy per class was calculated for each of test resolution in every experiment using Equation (5-2).

The first experiment used 16-megapixel (MP) images for training, which would require 3 cameras to cover the 1-metre wide conveyor belt with 289 PPI resolution. This is equivalent to using a single 86 MP resolution sensor (“Teledyne DALSA Falcon4,” 2018) to capture a field of view of 100 cm x 67 cm, covering 0.67 m2 area of the conveyor belt (Table 5-2). Experiments 2 and 3 used the equivalents of a single 16 MP camera sensor, with the longest side of the sensor perpendicular and parallel to the conveyor belt width, respectively. Experiments 4 and 5 changed the scale of the images to train the network using the equivalent of images taken with an 8 MP sensor, typically found in a 4K camera, and using different orientations.

Experiment 6 trained the neural network with the combination of training images of used in experiments 1 through 5. Experiment 7 shows the test accuracies when training the network with images equivalent to a high-resolution 41 MP sensor camera ("The Imaging Source DFK AFU420-L62," 2017). Experiment 8 evaluated the network accuracy when training with the images combined from experiment 6 and 7.

Table 5-2. Summary of experimental design parameters

Experiment #	Train resolution (PPI)	Equivalent field of view (cm)	Equivalent image size (pixels)	Description
1	289	43 x 29 (*)	4,896 x 3,264	Original, 16-megapixel (MP) resolution images, 0.12 m2 area
		100 x 67	11,378 x 7,585	Equivalent to 86 MP camera covering 0.67 m2 area
2	124	100 x 67	4,896 x 3,264	Equivalent to 16 MP camera covering 0.67 m2 area
3	83	100 x 150	3,264 x 4,896	Equivalent to 16 MP camera, covering 1.50 m2 area
4	98	100 x 56	3,840 x 2160	Equivalent to 8 MP (4K) camera, covering 0.56 m2 area
5	55	100 x 178	2,160 x 3,840	Equivalent to 8 MP (4K) camera, covering 1.78 m2 area
6	All the above	All the above	All the above	All the above
7	200	100 x 67	7,874 x 5,249	Equivalent to 41 MP camera, covering 0.67 m2 area
8	All the above	All the above	All the above	All the above

(*) a minimum of 3 16 MP cameras pointing to side-to-side areas would be required to capture the width of a 1-metre conveyor belt, covering an effective 129 cm x 29 cm area (0.37 m^2).

5.3 Results

Each experiment resulted in a trained network that was used to classify test images of different resolutions, and the average accuracy per class was calculated using Equation (5-2) for each of these tests. The highest average accuracy per class in Experiment 1, using a network trained with 289 PPI images, was 86.7% (SE=1.8%) when testing images of 280 PPI (Figure 5-1). When testing images with lower resolutions on that same trained network, the average accuracy per class decreased, with significant differences observed between the average accuracy per class at 289 PPI and the obtained with any spatial resolutions lower than 190 PPI. The worst case was obtained when testing 25 PPI images, with an average accuracy per class of 19.2% (SE=7.1%).

Experiments 2,3, and 4, classifying images with networks trained with 83, 98 and 124 PPI, respectively resulted in maximum average accuracies per class of 90.6% (SE=2.0%), 89.6% (SE=2.0%), and 87.6% (SE=2.8%). These peaks were all observed around the resolution of the images used for training. Furthermore, the classification performance decreased as the resolution of the test images diverged from the training resolution to minimum average accuracies per class between 16.5% and 33%.

During experiment 5, when classifying with a network trained using 55 PPI images, the highest average accuracy per class of 92.3% (SE=2.2%) was obtained for test images of 50 PPI resolution. Significant differences were observed between classifying a 55 PPI image and test images of resolutions higher than 70 PPI (Figure 5-1). The worst performance was obtained when testing images of 289 PPI, with an average accuracy per class of 13.2% (SE=3.5%).

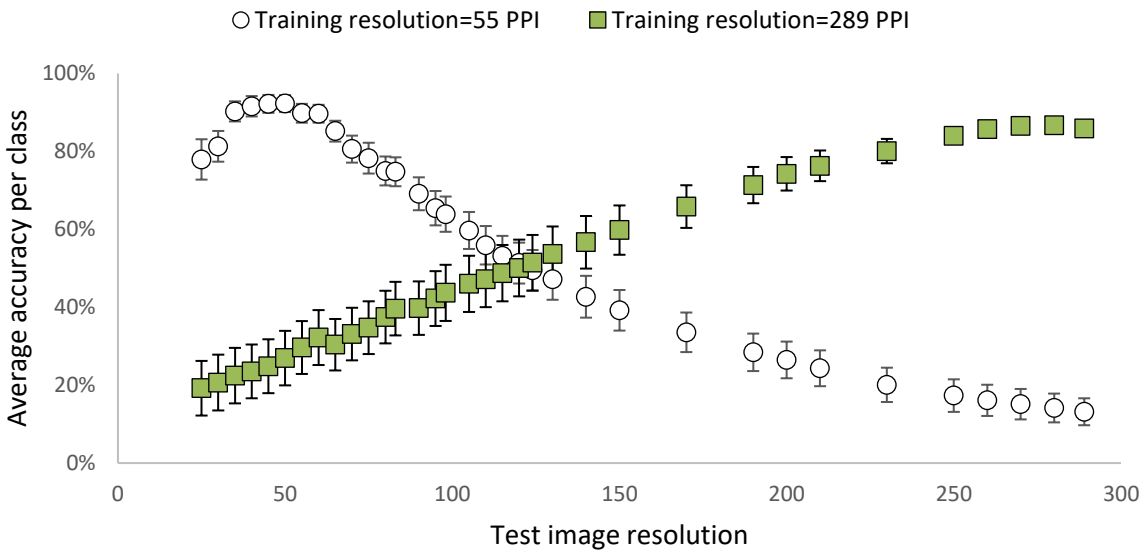


Figure 5-1 – Average accuracy per class for images of different resolutions classified with 2 different networks, one trained with 55 and the other with 289 PPI images. Vertical bars show the standard error for each tested resolution.

The neural network in Experiment 6 was trained with a combination of images of 5 different spatial resolutions (55, 83, 98, 124, and 289 PPI). The highest average accuracy per class was not significantly different to any the above individual experiment results, with 91.5% (SE=1.8%) observed on test images with a 60 PPI spatial resolution. However, the lowest average accuracy per class in experiment 6 was significantly better to any the obtained when training with only one resolution in experiments 1 to 5, with an average accuracy per class of 70.2% (SE=6.7%) at 25 PPI resolution. Furthermore, considering only the test resolutions in the range between 50 and 289 PPI, the lowest average accuracy per class was 84.2% (SE=2.3%) obtained when classifying images with a 190 PPI resolution (Figure 5-2). The resulting average accuracy per class for test resolutions between 50 and 289 PPI was 87.9% with a standard error of 2.3%.

The lowest average accuracy per class from experiment 6 was observed at 190 PPI, which is close to the middle point in the largest gap in training resolutions, between 124 and 289 PPI. The average accuracy per class at 190 PPI represents relative drops in performance of 3.4% and 1.1%, compared to the average accuracy per class obtained at 124 and 289 PPI, respectively. In the other training

resolutions gaps, the effects to the classification performance were relatively smaller, with average relative drops of 0.8% between 55 and 83 PPI, 0.5% between 83 and 98 PPI, and 1.2% between 98 and 124 PPI.

Experiment 7 was set up as experiments 1 to 5 but using images with a single resolution of 200 PPI to train the network. This resulted in similar performance to the observed for experiments 1 to 5, with a peak average accuracy per class of 87% (SE=2.5%) obtained when testing images with 210 PPI resolution. The worst average accuracy per class was 28.1 % (SE=8.4%), obtained when classifying images with 25 PPI resolution

Experiment 8 combined the training images from experiments 6 and 7 to train the network with 6 different spatial resolutions (55, 83, 98, 124, 200, and 289 PPI). The observed peak performance is lower than the results from experiment 6 but not significantly, with an average accuracy per class of 90.6% (SE=2.0%) when classifying test images of 60 PPI resolution. Considering only the test resolutions in the range between 50 and 300 PPI, the lowest average accuracy per class of 85.6% obtained when testing images of 289 PPI was not significantly different to the results from Experiment 6. In the range from 50 to 300 PPI, the average accuracy per class was 88.5% with a standard error of 1.2% (Figure 5-2), which is better but not significantly different to the results from experiment 6. The largest difference between experiment 6 and 8 was found when classifying test images of 190 PPI resolution, with an absolute improvement of 3.56%.

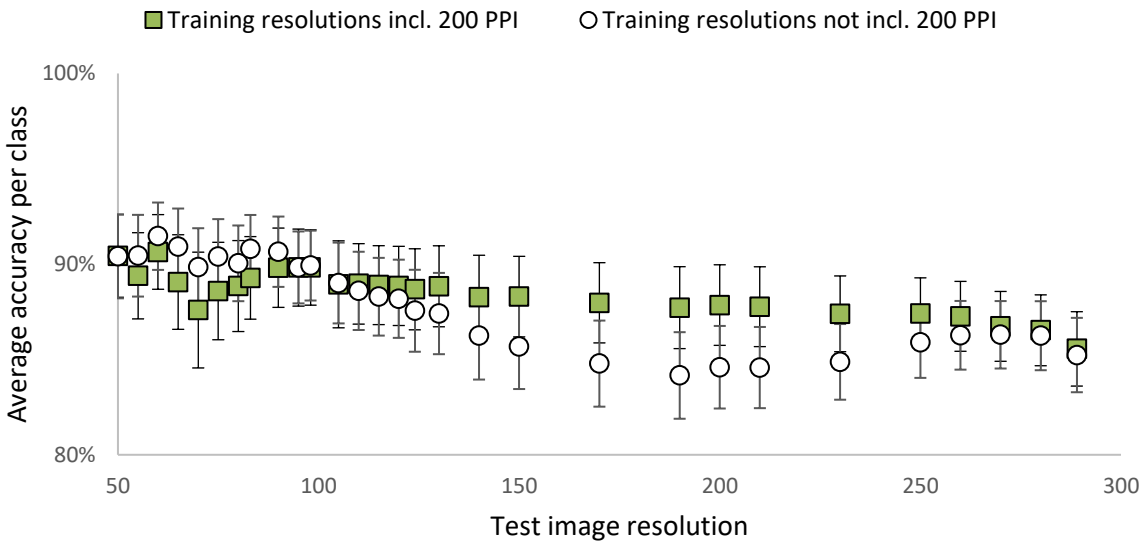


Figure 5-2 - Average accuracy per class for test images of different resolutions in the range from 50 to 289 PPI, classified by networks trained with images of multiple resolutions, with and without images of 200 PPI resolution. Vertical bars show the standard errors. Note: the Y-axis range is reduced to the range from 80 to 100%.

The classification mechanism used by the deep network trained in experiment 8 was analysed using the LRP method for different material classes. The results show how the model can identify the same features from the materials when identifying images with 55 PPI resolution or with higher 289 PPI resolution (Figure 5-3). In the case of red bricks (a, b), the model considers the porous surface as relevant features for classification, displayed with red colour, and ignores holes in the material sample. The light reflections off the metal tubes (c, d) are the most significant features, enough to classify the patches correctly even though the model also found areas with negative evidence, displayed in blue colour. In the case of newspaper material (e, f), the printed edge details are highlighted by the model as the most important characteristic from the images for both resolutions.

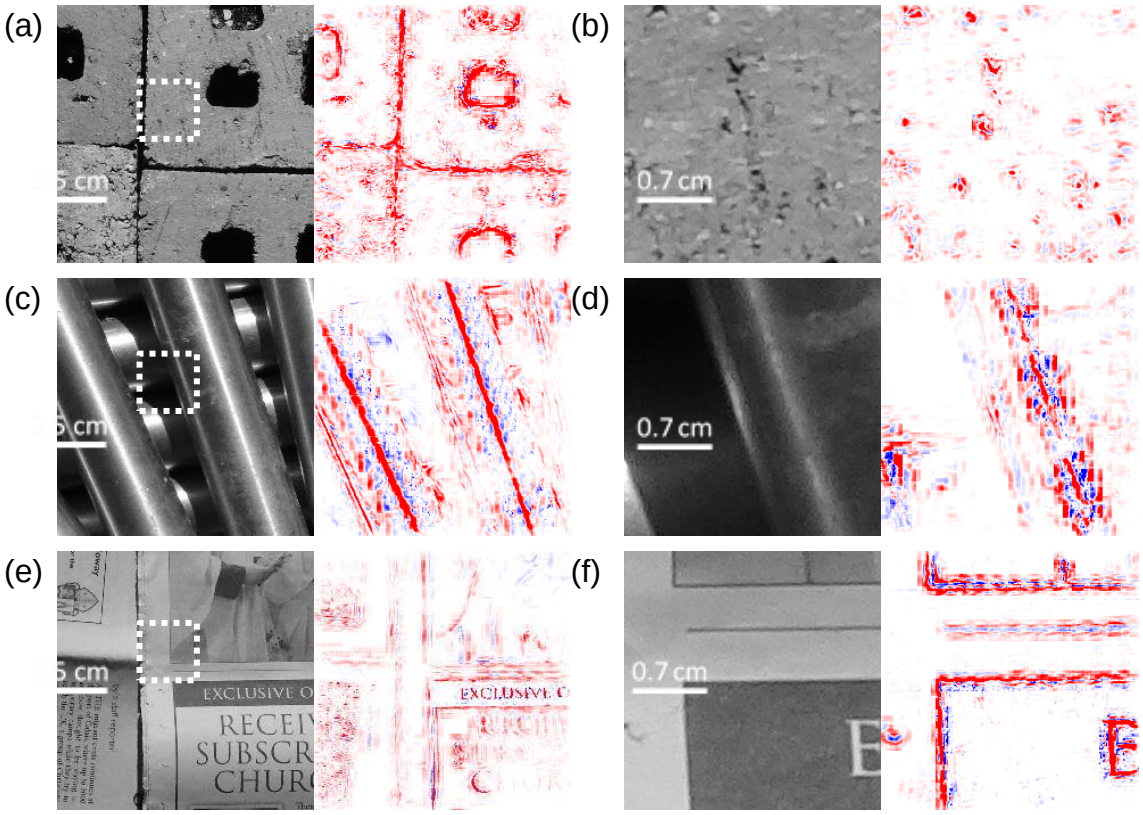


Figure 5-3 - Correct classification examples, using the trained network from Experiment 8, with test images of the same material sample using different resolutions. (a) red brick 55 PPI, (b) red brick 289 PPI, (c) metal pipes 55 PPI, (d) metal pipes 289 PPI, (e) newspaper 55 PPI, (f) newspaper 289 PPI. Left image is the input, right image is the heatmap generated towards its correct classification. Red colour indicates positive evidence, blue colour negative evidence, no colour is neutral evidence. The dotted-lines square regions on the 55 PPI input image denote the areas analysed by the 289 PPI input patch.

The test results from experiment 8 indicate that some materials, such as newspaper and household cardboard boxes, are better classified when testing 55 PPI images, with accuracies of 90.7% (SE=1.6%) and 95.8% (SE=1.2%), respectively (Figure 5-4). On the other hand, a lower accuracy of 68.9% (SE=12.9%) for patches of glossy magazines was obtained with 55 PPI spatial resolution, with 18.7% of the patches misclassified as newspaper material and 12.5% as food and drinks cardboard boxes. The highest accuracy for patches of glossy paper was 84.6% (SE=0.7%) when testing the trained network with 210 PPI images. The accuracy for glossy paper dropped to 78.3% (SE=0.4%) at 289

PPI, with 6% of the patches wrongly classified as office paper, 5% as newspaper, and 9% as household cardboard boxes.

The classification accuracy of other materials, such as red tissue paper, peaked at the highest resolution test images. For spatial resolutions up to 150 PPI, the trained network wrongly classified an average of 8% of the red tissue paper patches as textile materials and 6% as corrugated cardboard. But the class accuracy increases with higher resolutions to obtain the highest accuracy rate of 93.7% at 289 PPI. The background material presented similar results to red tissue paper, as the accuracy for this class was less than 70% with spatial resolutions lower than 98 PPI. The maximum accuracy for the class of background material in experiment 8 was 92.8% obtained with 289 PPI images. However, the lowest accuracies for plastic carrier bags with 64.9% (SE=1.7%), and newspaper with 72.6% (SE=1.5%) were obtained at 289 PPI. Carrier bags samples were mainly misidentified as plastic toys (7.2%), office paper (6.3%) and newspaper (5.8%). In the case of newspaper samples, they were mostly misclassified as office paper (13.6%) and household cardboard boxes (8.5%).

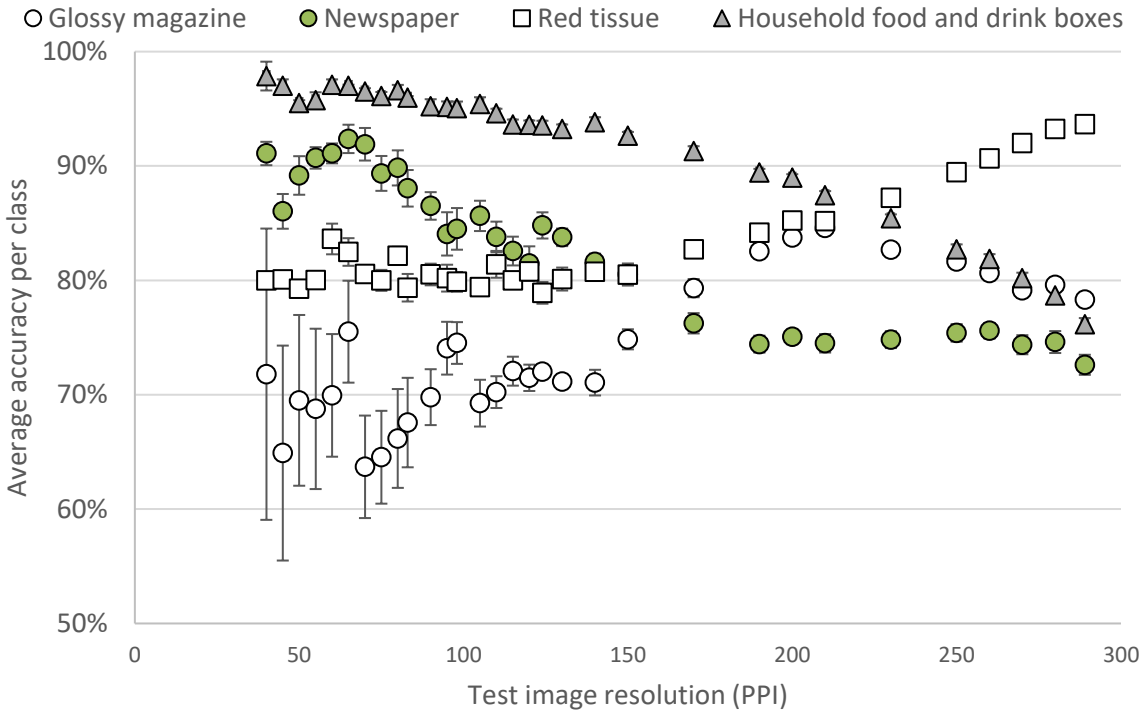
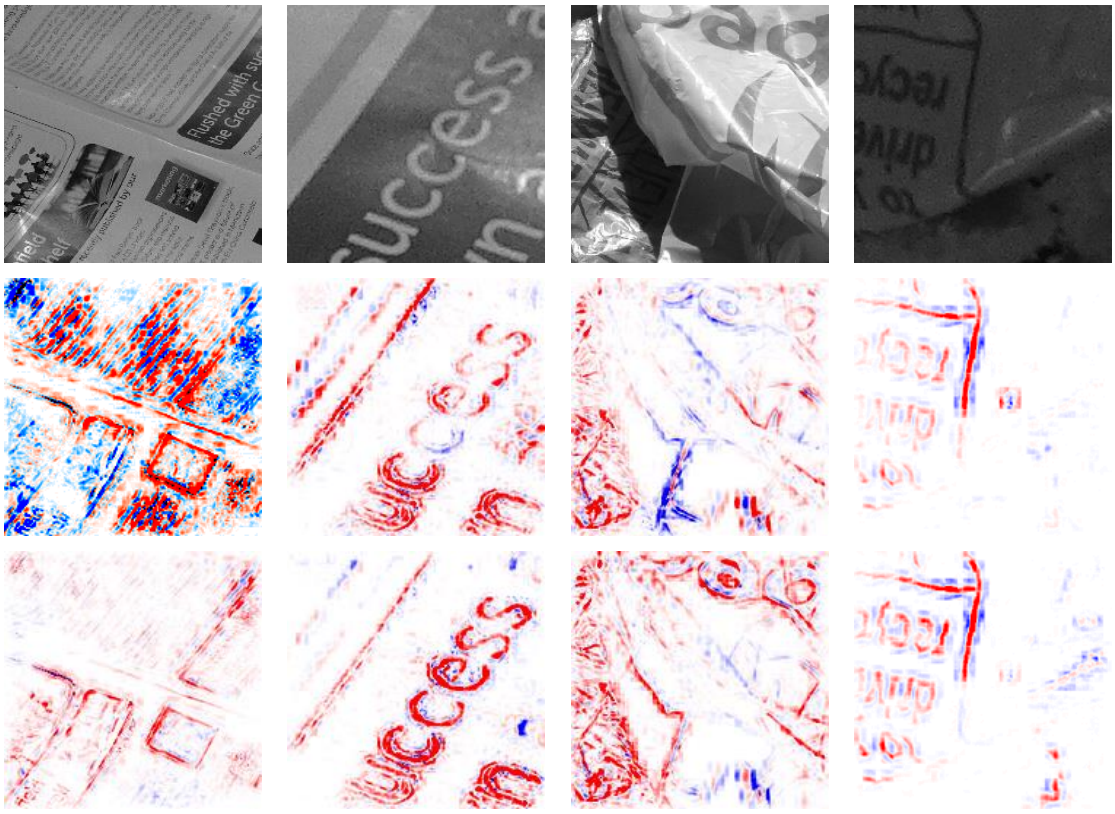


Figure 5-4 - Classification performance for 4 different material classes using images of different resolutions. Vertical bars show the standard error of the mean. Note: the Y-axis range is reduced to 50-100%.

The LRP method was applied to a test patch of glossy magazine material at 55 PPI resolution which was wrongly classified as newspaper. The resulting heatmap shows some areas identified as positive evidence for its correct classification (Figure 5-5), highlighting graphic art edges and some writing in red colour. But this same heatmap also shows a higher number of areas with negative evidence, shown in blue colours, as the model did not expect all the writing. This is consistent with the results of the LRP analysis towards its classification as newspaper, which instead shows all the writing in red colour, as the model expects a newspaper to contain lots of writing. Similar insights can be derived from the 289 PPI test sample of glossy magazine.

The heatmaps obtained from a carrier bags patch with 55 PPI resolution, which was incorrectly classified as newspaper material, show similar areas with positive evidence shown in red colour for both carrier bags and for newspaper. However, there are areas with negative evidence towards its correct classification,

highlighting the straight edges and writing in the central vertical region of the patch. Instead, these same areas were mostly considered as positive evidence towards its classification as newspaper material. In the case of the 289 PPI carrier bag sample, the heatmaps shows some extra positive evidence towards its classification as newspaper where the network found a corner edge, towards the lower right region of the image.



(a) (b) (c) (d)

Figure 5-5 - Wrong classification examples, using the trained network from Experiment 8 for images of different resolutions. (a) glossy magazine 55 PPI, (b) glossy magazine 289 PPI, (c) plastic carrier bag 55 PPI, (d) plastic carrier bag 289 PPI. At the top is the input image, middle shows the heatmap generated towards the correct class, and bottom image is the relevance heatmap for the wrongly predicted class. Red colour indicates positive evidence, blue colour negative evidence, no colour is neutral evidence.

The standard error of the mean, calculated from the repeated test results, was higher for some materials when captured at the lowest resolutions. For example, the standard error for glossy magazine material was reduced from 12.1% at 55

PPI to 0.44% as the resolution of the test images increased to 289 PPI (Figure 5-4). However, the standard error for other material classes, such as stones, is 1.2% at 289 PPI, higher than 0.2% obtained at 55 PPI. The differences are not statistically significant because of the range of standard errors, but the average standard error for the 20 material classes decreased from 1.57% at 55 PPI to a minimum of 0.82% as the resolution increased to 210 PPI (Figure 5-6). The average range of the standard error for each class at 210 is 0.55%, between 0.82% and 1.37%. The lowest average range of standard error was 0.2%, between 0.9% and 1.1%, obtained at 289 PPI.

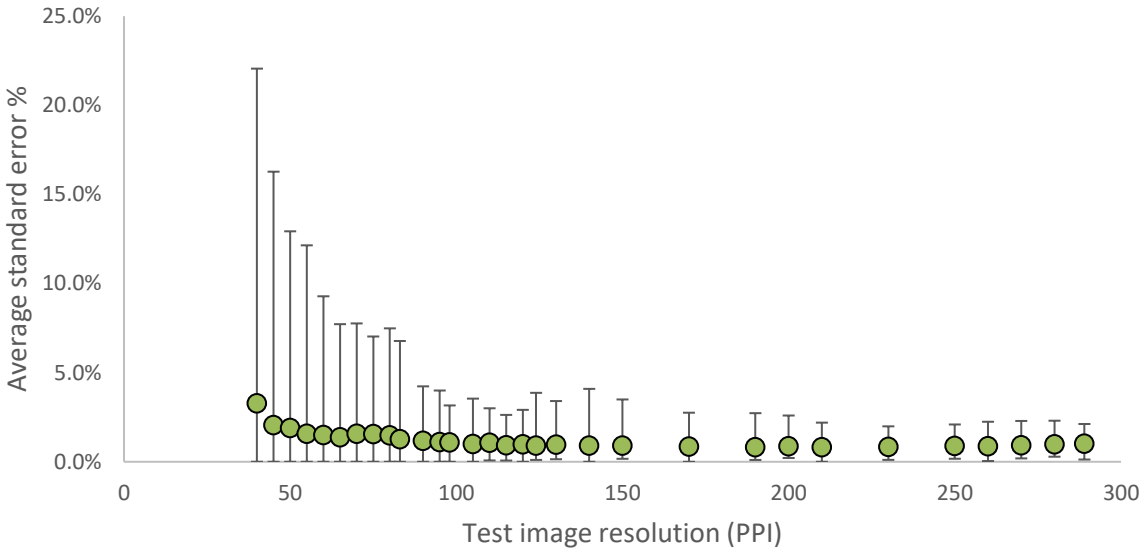


Figure 5-6 - Average standard error of the average accuracy per class for the range of spatial resolution in the test images. Vertical bars show the full (min, max) range of standard errors for the 20 subclasses considered.

5.4 Discussion

The average accuracy per class of 85.9% obtained in Experiment 1, training and testing the network at 289 PPI, showed an improvement compared to previous results of 83.7% and 74.8% (Chapters 3 and 4). This could be due to the increased number of material classes used for training, which reduces the intra-class variabilities, making it easier to the neural network to learn the differentiating characteristic of each material. This in turn leads to increased classification performance.

The performance of the following experiments, 2 to 5 and 7, are consistent with experiment 1. They also show that the average classification rates improve when lowering the resolution of the training and testing images, reaching a maximum of 89.8% with 55 PPI. Even though the differences are not significant, this would indicate that a properly trained network can classify the materials based on larger features better than with smaller and finer details, as picked up by higher resolution images.

The results of experiments 1 to 5 and 7 indicate that the highest classification performance is obtained when testing images of the same spatial resolution as the images used for training. As expected, the performance of the neural networks to predict material classes is negatively affected by changes in the resolution of the test images. These performance drops occurred when the resolution of the test images is either increased or decreased, which means that the entire range of training resolutions are required to improve performance.

Training the network with a combination of images of different spatial resolutions significantly improved the classification rates, as indicated by the results from Experiment 6. Even though the highest average accuracy per class of 91.5% is not significantly different to those obtained using single training resolutions, the lowest average accuracy per class was significantly increased to 70%. Furthermore, the results indicate that the level of performance can be maintained in the range of resolutions used for training, in this case 50 to 300 PPI, with an average accuracy per class of 87.9% and a standard error of 2.3%.

Additionally, when comparing the average accuracy per class obtained at test resolutions matching the training resolutions from experiments 1 to 5 (289, 124, 98, 83, 55 PPI, respectively), no significant differences were found to the highest average accuracy per class obtained in those experiments. This implies that combining databases of images of different resolution not only improves the average accuracy per class over a range of resolutions, but it has no negative effect compared to the targeted experiments. Even if the goal is to classify images with a known resolution, a deep network trained with multiple resolutions can yield the same or even higher average accuracy per class than a targeted training

using a single resolution. The non-target resolutions add useful information to the training dataset.

The results from Experiment 6 illustrate the reduction of classification performance when testing resolutions not represented in the training image dataset. Accordingly, the reduction of those gaps in training resolutions resulted in improved average accuracy per class. This was demonstrated by Experiment 8, which added training images with 200 PPI resolution to the training database of experiment 6, thus reducing the large gap between 124 and 289 PPI.

The performance obtained in the range of training resolution gaps from experiment 8 show a lowest classification of 87.6% obtained at 70 PPI, in the gap between the 55 and 83 PPI training resolutions. This represented an average drop of performance of 0.2% compared to the results at 55 and 83 PPI, less than the drops of performance obtained on experiment 6. The average accuracy per class of 87.6% at 70 PPI was not significantly different, but 3.4% better than the lowest classification rate from experiment 6 at 190 PPI.

When looking at the entire testing resolution range, the lowest average accuracy per class from experiment 8 was 85.6% (SE=1.9%), obtained at 289 PPI, matching the results from experiment 1 at that test resolution. Furthermore, comparing the results at 55, 83, 98, and 124 PPI, there are no significant differences to the classification performances and standard errors obtained with the targeted experiments 1 to 5. This would indicate that there are no negative effects to the performance from using multiple training resolutions. This also implies that the chosen deep network has enough layers to learn the differentiating features of the 21 material classes with 6 different resolutions.

The findings from Experiment 8 are consistent with the results from experiments 1 to 5 and 7, as the highest average accuracy per class was obtained at lower resolutions. This would indicate an overall preference for the deep network to classify the materials based on their larger visible features. However, this does not apply equally to all the different considered classes, as some samples were better recognized at higher resolutions, such as red tissue paper, background materials, and glossy magazines.

The minimum accuracy rates of 85.6% improved the results of 75.2% obtained by previous studies using a training dataset with multiple resolutions (Simonyan and Zisserman, 2014). The main difference is that the experiments in this study used multiple fixed resolutions compared to randomly reducing the resolution during training. Using fixed resolutions ensures the creation of more unique patches, which help the trained model to behave as scale-invariant during testing. The results are also better than the 80% accuracy obtained by combining multiple networks, where each model is with different resolutions (van Noord and Postma, 2017).

Red tissue paper and the background textile materials do not contain any writing or other distinguishing printed features, so the network learnt to recognize them through their fine texture details. For example, with spatial resolutions up to 150 PPI, the trained network wrongly classified some of the red tissue paper patches as textile materials and corrugated cardboard, which are other materials with not writings or printed features. When the images of red tissue paper were captured with a spatial resolution higher than 150 PPI, the texture details became more visible, reducing the number of classification errors to obtain the highest accuracy rate of 93.7% at 289 PPI.

In the case of glossy magazines, the highest accuracy was obtained with a spatial resolution of 210 PPI, and the accuracy decreased with either lower or higher resolutions. Based on the qualitative analysis of glossy magazine test patches, the trained network based its classification on the graphic art details such as text boxes and figures, and some of the captions and writings on this material towards its classification. At resolutions lower than 210 PPI, a patch of glossy magazine could include large areas of writings that would be wrongly classified by the network as newspaper material. A similar observation can be drawn from the patches with higher spatial resolutions, as the visualised area is smaller so it cannot include large features such as graphic art, and only the smaller features such as writings could be present. As a result, the trained networks would also classify these high-resolution patches as newspapers. The highest accuracy

obtained at a spatial resolution of 210 could indicate that this is the optimal balance between large and fine details for this material classification.

The variability of the Experiment 8 results obtained after repeating the methodology 3 times, illustrates the difference in accuracy rates obtained for each of the material classes in the range of considered resolutions. The variability was higher for lower resolutions, as indicated by the higher mean and range of standard errors observed at 55 PPI than 289 PPI. This would mean that, even though the average accuracy per class is higher at 55 PPI, the methodology could fail to train the networks to reliably classify the materials classes with the expected average accuracy per class. The lowest average standard error, indicating the lowest variability when training the deep network, is obtained with 210 PPI resolution images, which would result in an average accuracy per class of 87.8%.

5.5 Conclusions

The findings from this study were used to develop, test, and validate a technique to perform material classification on images with different resolutions. Using a training dataset with multiple images of multiple spatial resolutions resulted in state-of-the-art results, correctly classifying an average of 88.5% of patches captured with a range of spatial resolutions between 50 and 289 PPI. Furthermore, adding multiple lower resolutions to the training dataset improved the classification performance of the network on the high-resolution images against a model trained with only high-resolution images.

The results from this study mean that a scale-invariant solution could be effectively deployed in the field and achieve the required level of operational performance in any common situation. The trained network would be able to analyse waste using different sensor resolutions, as shown by the results, ranging from 8 to 86 megapixels. Similarly, a 16-megapixel sensor could be used to correctly analyse waste flows on belts of different widths, from 43cm wide (using 289 DPI) to 2.5m wide (using 50 PPI). This range covers most of the conveyor belts used throughout the industry in waste material recovery facilities. The

proposed methodology offers robustness and flexibility in the installation of such a solution in the field.

Analysing the mechanisms used by the trained networks to recognize and classify the different materials clearly showed the importance of illumination, in agreement to previous research. These differentiating features, such as edges and illumination highlights, can be seen in the entire the range of resolutions studied. These features, though scale variant, share similarities at different resolutions, which could explain why the network filters were able to capture their features, in contrast to previous findings. A spatial resolution of 210 PPI offers a balance of high and low frequency content for the classification of some classes of solid waste materials and increases the precision of the methodology.

In this work, the standard deep learning network architecture managed to learn and recognize materials at different scales with state-of-the-art performance without overfitting. Further qualitative studies could lead to improved methodologies as the strengths and weaknesses of the trained networks classification methods are further recognized. The impact of this method in waste management could extend beyond material recovery facilities, as its simplicity could enable the technology being implemented at multiple stages, from helping segregation of waste at homes to international transport and controls.

5.6 References

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., Samek, W., 2015. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. PLOS ONE 10, e0130140. https://doi.org/10.1371/journal.pone.0130140

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Cimpan, C., Maul, A., Wenzel, H., Pretz, T., 2016. Techno-economic assessment of central sorting at material recovery facilities – the case of

lightweight packaging waste. Journal of Cleaner Production 112, Part 5, 4387–4397. https://doi.org/10.1016/j.jclepro.2015.09.011

Conveyor Belting [WWW Document], 2019. . Conveyor Belting. URL https://www.conveyorbelting.net/ (accessed 9.30.19).

DEFRA, 2018a. 25 Year Environment Plan [WWW Document]. URL https://www.gov.uk/government/publications/25-year-environment-plan (accessed 10.3.19).

DEFRA, 2018b. Resources and waste strategy for England [WWW Document]. URL https://www.gov.uk/government/publications/resources-and-waste-strategy-for-england (accessed 10.3.19).

EU Commission, 2018. EUR-Lex - 52018DC0029 - EN - EUR-Lex [WWW Document]. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A29%3AFIN (accessed 12.18.19).

EU Council, 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:FULL:EN:PDF (accessed 10.19.16).

EU Council, 1999. Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. Official Journal of the European Communities L 182/1, 1–19.

EU Parliament, 2018. Circular economy package: Four legislative proposals on waste [WWW Document]. URL http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2018)614766 (accessed 4.24.18).

EU Parliament, 2017. The circular economy package: new EU targets for recycling [WWW Document]. URL https://www.europarl.europa.eu/news/en/headlines/society/20170120STO59356/the-circular-economy-package-new-eu-targets-for-recycling (accessed 12.18.19).

Girshick, R., 2015. Fast r-cnn, in: Proceedings of the IEEE International Conference on Computer Vision. pp. 1440–1448.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv:1408.5093 [cs].

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

Kylberg, G., Sintorn, I.-M., 2016. On the influence of interpolation method on rotation invariance in texture recognition. EURASIP Journal on Image and Video Processing 2016. https://doi.org/10.1186/s13640-016-0117-6

Laner, D., Rechberger, H., Feketitsch, J., Fellner, J., 2016. A Novel Approach to Characterize Data Uncertainty in Material Flow Analysis and its Application to Plastics Flows in Austria. Journal of Industrial Ecology 20, 1050–1063. https://doi.org/10.1111/jiec.12326

Lapuschkin, S., Binder, A., Montavon, G., Müller, K., Samek, W., 2016a. The LRP Toolbox for Artificial Neural Networks. Journal of Machine Learning Research 17, 1–5.

Lapuschkin, S., Binder, A., Montavon, G., Müller, K.-R., Samek, W., 2016b. The LRP toolbox for artificial neural networks. Journal of Machine Learning Research 17, 1–5.

Lapuschkin, S., Binder, A., Müller, K., Samek, W., 2017. Understanding and Comparing Deep Neural Networks for Age and Gender Classification. arXiv:1708.07689 [cs, stat].

McKinsey & Company, 2016. Managing waste in emerging markets [WWW Document]. URL https://www.mckinsey.com/business-functions/sustainability/our-insights/managing-waste-in-emerging-markets (accessed 10.3.19).

Ngiam, J., Chen, Z., Chia, D., Koh, P.W., Le, Q.V., Ng, A.Y., 2010. Tiled convolutional neural networks, in: Lafferty, J.D., Williams, C.K.I., Shawe-Taylor, J., Zemel, R.S., Culotta, A. (Eds.), Advances in Neural Information Processing Systems 23. Curran Associates, Inc., pp. 1279–1287.

OpenCV library 3.1.0 [WWW Document], 2015. URL https://docs.opencv.org/3.1.0/ (accessed 5.2.18).

Rechberger, H., Cencic, O., Frühwirth, R., 2014. Uncertainty in Material Flow Analysis. Journal of Industrial Ecology 18, 159–160. https://doi.org/10.1111/jiec.12087

Samek, W., Binder, A., Montavon, G., Lapuschkin, S., Müller, K., 2017. Evaluating the Visualization of What a Deep Neural Network Has Learned. IEEE Transactions on Neural Networks and Learning Systems 28, 2660–2673. https://doi.org/10.1109/TNNLS.2016.2599820

Schwartz, G., Nishino, K., 2016. Material Recognition from Local Appearance in Global Context. arXiv:1611.09394 [cs].

Seibold, C., Samek, W., Hilsmann, A., Eisert, P., 2018. Accurate and Robust Neural Networks for Security Related Applications Exampled by Face Morphing Attacks. arXiv:1806.04265 [cs].

Séverin, M., Velis, C.A., Longhurst, P.J., Pollard, S.J.T., 2010. The biogenic content of process streams from mechanical–biological treatment plants producing solid recovered fuel. Do the manual sorting and selective dissolution determination methods correlate? Waste Management, IWWG Task Group on “Thermal Treatments” 30, 1171–1182. https://doi.org/10.1016/j.wasman.2010.01.012

Simonyan, K., Vedaldi, A., Zisserman, A., 2013. Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034.

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs].

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–9.

Tchobanoglous, G., Kreith, F., 2002. Handbook of solid waste management. New York : McGraw-Hill, 2002.

Teledyne DALSA Falcon4 [WWW Document], 2018. . STEMMER IMAGING. URL https://www.stemmer-imaging.com/en-gb/products/series/teledyne-dalsa-falcon4/ (accessed 10.12.19).

The Imaging Source DFK AFU420-L62 [WWW Document], 2017. URL https://www.theimagingsource.com/products/zoom-autofocus-cameras/usb-3.0-color-autofocus-cameras/dfkafu420l62/ (accessed 10.12.19).

van Noord, N., Postma, E., 2017. Learning scale-variant and scale-invariant features for deep image classification. Pattern Recognition 61, 583–592. https://doi.org/10.1016/j.patcog.2016.06.005

Vrancken, C., Longhurst, P., Wagland, S., 2017. Critical review of real-time methods for solid waste characterisation: Informing material recovery and fuel production. Waste Management 61, 40–57. https://doi.org/10.1016/j.wasman.2017.01.019

Vrancken, C., Wagland, S., Longhurst, P., 2019b. Balanced training raw data. 10.17862/cranfield.rd.9968051.

WRAP, 2006. Materials Recovery Facilities [WWW Document]. URL http://www.wrap.org.uk/sites/files/wrap/MRF_v6_19Dec06_LC.pdf (accessed 8.24.16).

Wu, R., Yan, S., Shan, Y., Dang, Q., Sun, G., 2015. Deep image: Scaling up image recognition. arXiv preprint arXiv:1501.02876 7.

Xu, Y., Xiao, T., Zhang, J., Yang, K., Zhang, Z., 2014. Scale-Invariant Convolutional Neural Networks. arXiv:1411.6369 [cs].

Zeiler, M.D., Fergus, R., 2014. Visualizing and Understanding Convolutional Networks, in: Computer Vision – ECCV 2014, Lecture Notes in Computer Science. Springer, Cham, pp. 818–833. https://doi.org/10.1007/978-3-319-10590-1_53

CHAPTER 6

MIXED WASTE COMPOSITION ANALYSIS

This chapter improved the image capture to deal with non-uniform illumination of material samples and developed a methodology to determine the composition of mixed waste materials based on classification of patches

6 Determination of mixed waste composition using deep neural networks

Abstract

New policies and regulations require new technologies to improve the performance of waste management and material recovery systems. A methodology based on imaging and deep neural network image classification was applied and extended to determine the composition of mixed waste materials, assessing its limitations in typical case scenarios. A database of images with different level of exposure levels was used to finetune a deep neural network to improve the classification of pictures of waste captured with non-uniform illumination. The accuracy and precision of the methodology was assessed for the classification of mixed waste of different properties and appearance. The performance obtained was better than a previous study using human analysis of images of waste, and the proposed methodology can provide automatic composition analysis of waste streams in material facilities in real-time.

6.1 Introduction

In recent years, new legislation and regulations were introduced to reduce the amount of waste going into landfills (EU Commission, 2018; EU Council, 2009, 1999; EU Parliament, 2018, 2017). To achieve this, the current levels of waste material recovery need to increase; governments are pushing towards making manufacturers responsible for the materials used in their products throughout their entire life cycle (DEFRA, 2019). This has served as motivation for research and development of advanced waste treatment technologies (DEFRA, 2018). There is an expectation for automated waste analysis solutions, using computer vision and artificial intelligence, to be able to analyse waste streams with accuracy levels of 85% (DEFRA, 2021).

Advanced techniques for mixed solid waste classification could provide key performance data for the optimization of waste treatment processes. This could lead to improved waste management systems, with a reduction in landfill waste disposal, and the information could also be used to inform stakeholders

(Vrancken et al., 2017). Current sampling guidance and regulation in the UK stipulates sending input and output waste samples for laboratory analysis and regular reporting to the government (DEFRA, 2014). An automated methodology could improve and replace current hand-sorting on-site, by analysing the quality of waste in less time and potentially reducing costs and waste storage requirements.

In the context of a material facility, mixed waste materials flow between one mechanical treatment process to the next on top of conveyor belts (Velis et al., 2013). The quantity and the composition of the material streams changes over time and between processes, and the physical properties of the materials can be transformed. For example, a shredder reduces the size of the samples to produce a homogeneous solid refused fuel (SRF) output (Nasrullah et al., 2014; Wagland et al., 2011). Similarly, the visual appearance of the materials can be affected, for example by a vibratory feeder which ensures an even distribution of non-overlapping materials on the conveyor belts.

The variable composition of the input waste stream affects the quality of the material outputs (McKinnon et al., 2017), and a real-time composition analysis at different stages would fit in the context of Industry 4.0 (Fasting, 2019; Lasi et al., 2014). This could facilitate a shift from the current quantitative and fixed designs of waste processing plants, based on estimated throughputs and compositions, to a qualitative design using feedback loops to control the plant based on performance data. Data about the composition and flow of waste between different stages of waste systems could affect the operation of the different processes to maximise the material recovery and meeting regulatory standards without delays (Vrancken et al., 2017).

The different stages of waste management where this analysis could be implemented goes beyond the material facilities. Automated waste composition analysis could also be deployed to improve waste separation at home, to monitor waste transport to local depots, to assess the infeed quality of waste-derived fuels such as refused derived fuel (RDF), and to check and control international shipments of waste materials. Furthermore, waste flow analysis could also be

implemented as part of an extended producer responsibility (EPR) policy, in agreement with the recent plans to incentivise better product designs (DEFRA, 2018).

Visual inspection of images of mixed waste was investigated in a previous study which showed a high correlation between estimated and true composition values (Wagland et al., 2012). Further research combined this image analysis to a moisture detection system to estimate the energy in waste materials (Wagland et al., 2013). A more recent study correlated the image analysis to the energy potential of SRF mixtures (Peddireddy et al., 2015). The technique developed in this series requires a person to manually assess and classify the material in 121 dots for each of the images of waste. The total time required to analyse a single image ranged between 3 and 4 minutes, and the repetitive process could lead to increased human error levels.

Deep learning techniques using convolutional neural networks (CNN) have been developed for the task of automatic image recognition (Krizhevsky et al., 2012; LeCun et al., 2015). In recent years, the performance of deep neural networks has surpassed human perception for the classification of objects and materials (Russakovsky et al., 2015). The classification of materials has also been investigated through the use of deep learning, achieving similar or better performance than humans (Sharan et al., 2014). The study by Bell et al. (2015) achieved 85% accuracy to classify the materials from images containing clues, or contextual information to help their identification.

In waste management systems, there is not context information in the images, and the material samples can be located anywhere within the field of view at arbitrary rotation angles (Chapter 3). The work by Mittal et al. (2016) trained a neural network to recognize the presence or absence of waste in pictures of real-life scenes. The classification was reduced to a binary classification task, detecting waste or no waste, and used context from the pictures, such as landscapes and streets, to reach a classification accuracy of 88%.

The study by Sakr et al. (2016) obtained 83% accuracy using convolutional neural networks to classify single waste objects into 3 classes: plastic, metal, and paper

classes. A series of projects used a database of single waste objects (Yang and Thung, 2016) to perform waste sorting into glass, paper, plastic, and metal material classes with accuracies between 87% and 98% depending on the network structure (Adedeji and Wang, 2019; Aral et al., 2018; Bircanoğlu et al., 2018; Ozkaya and Seyfi, 2019; Ruiz et al., 2019). The reported results of all these studies were obtained from the classification of one single item in each image, and some of them exceed the performance achieved with the methodology implemented for this study (Chapters 3,4,5). They rely on the shape characteristics to recognise and classify the item and infer the material.

In contrast to those studies, this work analyses and classifies materials patch by patch, where each patch contains surface details from the items but includes no context information. As this study classifies the material exclusively from the information presented per patch, the approach allows its application to mixed waste materials streams at any stages, from dense mixed waste streams to shredded outputs. This type of analysis is not possible with the approach of classification of waste by items followed by others.

An image of waste materials would typically contain multiple material classes, sometimes overlapping each other. One other study adapted the waste objects database by Yang and Thung (2016) to create artificial images of overlapping waste objects (Awe et al., 2017). Deep convolutional networks were trained to classify these images simulating mixed waste materials, obtaining a mean average accuracy of 68%, to classify the materials into only three categories: landfill, recycling, and paper. However, and like the studies referenced before, this makes the assumption that the items still preserves its shape characteristics.

The work presented here applies and extends the methodology developed in previous studies (Chapters 3, 4, and 5) for the classification of mixed materials into 16 different categories typically found in a waste stream. An experiment determines the effects of non-uniform illumination over the field of view, which assumes that lighting cannot always be controlled in a waste plant, or in other applications of this technology. This work also assesses the effects of classifying different quantities of waste materials, as this leads to material samples

overlapping on top of the field of view. Finally, the average processing time required to analyse an image is measured to determine the potential of the methodology to be deployed as a real-time tool.

6.2 Materials and methods

In this section, the methodology is described in detail, including the required materials, the steps, and procedures required by the experimental design, and the applied data analysis.

6.2.1 Material samples

Waste materials from 16 classes were used for training and testing purposes in this study. These classes are typical materials found in waste streams, including light and heavy fractions (Séverin et al., 2010). For testing purposes, 3 different mixtures were used, each containing a different composition of the same 16 waste material classes, containing samples not included in the training dataset. The household cardboard boxes were pulled apart and flattened so that their volumes were reduced to a minimum. The resulting material samples were all considered as 2-dimensional flat samples. The green cotton fabric material used as a field of view background was included as a separate class but not considered waste material.

The surface area of each of the training and testing samples was measured from pictures taken of the individual samples using the ImageJ tool (Schindelin, 2012; Schneider, 2012). The training group contained 5.2 m^2 of material samples, the test mixture 1 contained a total of 0.5m^2 of mixed materials, the test mixture 2 was comprised of 2.2 m^2 of mixed materials, and the test mixture 3 included 5 m^2 of mixed samples. For some experiments, the materials samples in the light fraction of the test groups were cut into smaller rectangular pieces with longer sides between 3.5 cm and 5 cm long, like the shredded fraction of waste streams and SRF outputs (Séverin et al., 2010). The detailed total areas of each of the material classes in each group can be seen in Table 6-1. The relative area composition in each fraction can be seen in Table 6-2. Each material sample was

weighed, calculating the area density of each material class and the relative weight composition in each group (Table 6-3).

Table 6-1 - Area composition of each of the material groups used for training and testing – the material classes indicate whether they belong to the light (L) or heavy (H) fractions. The shredded fractions (not shown) share the composition of the light fractions.

Material class (fraction)	Area (m^2)					
	Training	Mix 1	Mix 2 (all)	Mix 2 (light)	Mix 3 (all)	Mix 3 (light)
Office paper (L)	0.440	0.031	0.249	0.249	0.499	0.499
Packaging paper (L)	0.500	0.033	0.210	0.210	0.427	0.427
Red tissue paper (L)	0.355	0.026	0.317	0.317	0.423	0.423
Hygienic and kitchen towel (L)	0.828	0.072	0.104	0.104	0.311	0.311
Glossy magazine (L)	0.505	0.031	0.249	0.249	0.499	0.499
Old newspaper (L)	0.605	0.053	0.204	0.204	0.612	0.612
Corrugated cardboard (L)	0.405	0.022	0.072	0.072	0.422	0.422
Household food and drink boxes (L)	0.405	0.041	0.313	0.313	0.517	0.517
Plastic carrier bags (L)	0.205	0.056	0.067	0.067	0.424	0.424
Black plastic trays (L)	0.255	0.025	0.126	0.126	0.244	0.244
White polystyrene foam (L)	0.085	0.010	0.038	0.038	0.070	0.070
Wood (H)	0.110	0.045	0.036	0	0.106	0
Synthetic textile towels (L)	0.355	0.029	0.166	0.166	0.337	0.337
Red bricks (H)	0.075	0.008	0.022	0	0.030	0
Stones (H)	0.025	0.000	0.012	0	0.030	0
Stainless steel pipes (H)	0.073	0.007	0.044	0	0.044	0

Table 6-2 – Relative area composition of each of the material groups used for training and testing expressed as a percentage to the total area of each mixture – the material classes indicate whether they belong to the light (L) or heavy (H) fractions. The shredded fractions (not shown) share the composition of the light fractions

Material class (fraction)	Area % composition				
	Mix 1	Mix 2 (all)	Mix 2 (light)	Mix 3 (all)	Mix 3 (light)
Office paper (L)	6.4%	11.2%	11.8%	10.0%	10.4%
Packaging paper (L)	6.6%	9.4%	9.9%	8.6%	8.9%
Red tissue paper (L)	5.4%	14.2%	15.0%	8.5%	8.8%
Hygienic and kitchen towel (L)	14.7%	4.6%	4.9%	6.2%	6.5%
Glossy magazine (L)	6.4%	11.2%	11.8%	10.0%	10.4%
Old newspaper (L)	10.8%	9.2%	9.7%	12.3%	12.8%
Corrugated cardboard (L)	4.6%	3.2%	3.4%	8.5%	8.8%
Household food and drink boxes (L)	8.4%	14.0%	14.8%	10.4%	10.8%
Plastic carrier bags (L)	11.4%	3.0%	3.2%	8.5%	8.9%
Black plastic trays (L)	5.0%	5.6%	5.9%	4.9%	5.1%
White polystyrene foam (L)	2.1%	1.7%	1.8%	1.4%	1.5%
Wood (H)	9.3%	1.6%	0.0%	2.1%	0.0%
Synthetic textile towels (L)	5.8%	7.4%	7.8%	6.8%	7.0%
Red bricks (H)	1.6%	1.0%	0.0%	0.6%	0.0%
Stones (H)	0.0%	0.6%	0.0%	0.6%	0.0%
Stainless steel pipes (H)	1.5%	2.0%	0.0%	0.9%	0.0%

.

Table 6-3 - Area density and relative weight composition of each material class expressed as a percentage to the total weight of each mixture. The material classes indicate whether they belong to the light (L) or heavy (H) fractions. The shredded fractions (not shown) share the composition of the light fractions

Material class (fraction)	Area density (g / m^2)	Weight composition % w/w				
		Mix 1	Mix 2 (all)	Mix 2 (light)	Mix 3 (all)	Mix 3 (light)
Office paper (L)	79.90	0.2%	0.6%	5.0%	0.7%	4.1%
Packaging paper (L)	73.91	0.2%	0.4%	3.9%	0.5%	3.2%
Red tissue paper (L)	32.92	0.1%	0.3%	2.6%	0.2%	1.4%
Hygienic and kitchen towel (L)	40.99	0.2%	0.1%	1.1%	0.2%	1.3%
Glossy magazine (L)	124.80	0.2%	0.9%	7.8%	1.0%	6.4%
Old newspaper (L)	41.02	0.1%	0.2%	2.1%	0.4%	2.6%
Corrugated cardboard (L)	739.66	1.1%	1.5%	13.3%	5.1%	31.9%
Household food and drink boxes (L)	354.56	0.9%	3.1%	27.8%	3.0%	18.8%
Plastic carrier bags (L)	22.82	0.1%	0.0%	0.4%	0.2%	1.0%
Black plastic trays (L)	431.95	0.7%	1.5%	13.7%	1.7%	10.9%
White polystyrene foam (L)	347.94	0.2%	0.4%	3.3%	0.4%	2.5%
Wood (H)	12,124.78	38.5%	11.6%	0.0%	20.2%	0.0%
Synthetic textile towels (L)	460.30	0.8%	2.1%	19.1%	2.6%	16.0%
Red bricks (H)	107,062.32	53.9%	65.9%	0.0%	53.5%	0.0%
Stones (H)	11,481.46	0.0%	3.9%	0.0%	5.7%	0.0%
Stainless steel pipes (H)	6515.50	2.9%	7.5%	0.0%	4.5%	0.0%

6.2.2 Image capture

The images used for training and testing purposes were captured using 1 visible sensor cameras and 2 light sources. It uses camera C5 and lights C1 and C2 from the setup described by a previous work (Chapter 3). The camera and the light sources were pointing at the target from a focal distance of 1,095 cm. The camera focal length is set to 5.98 mm and the lens aperture set to f/8, which resulted in a field of view of 109.8 cm by 73.3 cm with a depth of field of 57.2 cm in front of the focal plane.

The sample materials were placed on top of a green textile material that covered the entire field of view, representing a conveyor belt used in material recovery facilities. Each material sample was photographed with a camera placed at a zenith angle of 0 degrees in spherical coordinates, looking at the samples directly from above the target. The target area was illuminated using 2 light sources placed at a zenith angle of 30 degrees and azimuth angles of 0 and 90 degrees. The light sources illuminated the area with a beam diameter of 55 cm, which did not cover uniformly the entire field of view. The edges of the field of view were illuminated with 15% of the intensity at the centre of the target area.

The camera image resolution was set to 4,896 by 3,264 pixels, resulting in a spatial, or object resolution of 113 pixels per inch (PPI). The sensitivity of each camera was set to ISO 80 to minimise sensor noise. Each training material sample was photographed 16 times, each using a different exposure time in the range between 1/20 and 2 seconds. Each test material sample was photographed once with an exposure time determined automatically by the camera sensor for every image.

6.2.3 Image processing

The images were augmented to create the training databases, through transformations implemented in C++ (“OpenCV library 3.1.0,” 2015), with Lanczos resampling over an 8 by 8 pixels neighbourhood (Kylberg and Sintorn, 2016). The images were rotated around their centre 11 times with an angular separation of 30 degrees to each other. Each rotated image was then divided into patches of 256 by 256 pixels, overlapping 25% in each dimension with the neighbouring patches (Chapter 3).

A hue, saturation, and intensity (HSI) colour detection ensured that the training and testing images of any class containing any green colour content matching the green background material were discarded. Similarly, the HSI filter was used to select images containing only green background material for training and testing that background material class. All selected images were converted into 3-channel grayscale version with each channel containing the same values (Chapter 4).

6.2.4 Deep learning

The framework used to train and test the neural networks was Caffe version 1.00 (Jia et al., 2014), using the AlexNet network architecture (Krizhevsky et al., 2012). Training was implemented in parallel on 2 nVidia K80 GPUs, whilst testing and classifications were performed on one nVidia K80 GPU. The final layer of the network (fc8) was set to classify 21 materials, including the green background and the 16 material classes used in this study. Transfer learning was implemented in each training exercise, reusing the weight values obtained from previous experiments (Bell et al., 2015; Vrancken et al., 2019). The Adam optimisation algorithm was used to adjust the learning rate dynamically, with parameters $\alpha = 10^{-5}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ (Kingma and Ba, 2014). The dropout factor was set to 0.5 to reduce overfitting and to allow the network to better generalise its predictions.

A vector of 4,096 features to represent each patch in the training databases was calculated by applying an AlexNet network previously trained with the MINC database. The features vector was obtained from the fc7 network layer. These feature vectors were used to calculate the cosine distance between each of the patches in the training database. Patches that were considered too similar were removed from the databases, and the remaining unique patches were used to generate training databases. For some experiments, the unique training patches were upsized by a factor of 2 to 512 by 512 pixels using Lanczos resampling with a kernel of 8 by 8 pixels.

The experiments used different training datasets of unique patches sized 256 by 256 pixels or 512 by 512 pixels. Each patch fed to the network for training is randomly cropped to 227 by 227 pixels to prevent replicate input data and to reduce overfitting. Considering the 113 PPI object resolution, a patch of 227 by 227 pixels represents an area of 5.1 by 5.1 cm when working with the original 256 by 256 patches, or 2.6 by 2.6 cm when training with the upsized 512 by 512 pixels patches.

The training with datasets created using patches 256 by 256 pixels was performed for 20 epochs, so that each patch was used at least 20 times to train the network. The accuracy of the trained networks was checked every 4 epochs to stop the training as soon as the test accuracy stopped improving. Similarly, using patches upsized to 512 by 512 pixels, the number of training epochs was 80 and accuracy checks were performed every 16 epochs. Each training exercise was repeated 3 times.

6.2.5 Classification performance and composition analysis

Test databases, created using unique patches from images of material samples not included in the training database, were used for the accuracy tests during training. The trained networks were also used to analyse the images of the mixed test material, determining the composition based on the classification of individual patches. For some experiments, the images of the test mixtures were upsized by a scale of 2, matching the upsized scaling used for the patches used to train the network. The resizing operations were performed with Lanczos resampling using an 8 by 8 pixels kernel size.

The trained networks were used to determine the material of non-overlapping image patches sized 227 by 227 pixels. The fc8 layer output of the trained network is interpreted as the probability distribution of the tested patch of being one of 21 classes. For some experiments, the class with the highest probability was interpreted as the predicted material for each patch in the testing database. For other experiments, the probability of each class was considered towards the composition of each patch. For a given image of mixed waste material, every patch was analysed, and the composition was calculated as the ratio of quantity of each identified material to the total quantify of non-background materials. This analysis was repeated 3 times, using the 3 different trained neural networks, to calculate the mean composition and standard deviation (SD) for each material class in each image.

Each mixed materials sample was photographed multiple times and the area composition analysis was applied to each test image. This study can present

more than one material class in each patch, so the waste composition was determined using two alternative methods:

- *Method A*: 100% of the area of each patch was classified as the class with the highest (top-1) probability output from the network
- *Method B*: the full probability output vector from the network (top-21) were used as the area proportion of each of the 21 classes to the composition of each patch

The area composition for each class in each image was calculated from the composition of the individual patches from that image with Equation (6-1):

$$E_{ij} = \frac{1}{L_j} * \sum_{l=1}^{L_j} P_{ijl} \quad \textbf{(6-1)}$$

Where L_j is the number of patches analysed in the image j, and P_{ijl} is the proportion of the patch l from image j classified as class i.

For Method A, P_{ijl} is either 1 or 0, denoting the classification of the patch as class i or not, respectively. For Method B, P_{ijl} is a number between 0 and 1, denoting partial classification of the patch as class i.

The weight composition for each class in each image was estimated from the estimated area composition and the measured area density values (Table 6-3) with Equation (6-2):

$$W_{ij} = E_{ij} * L_j * A_{patch} * d_i \quad \textbf{(6-2)}$$

Where E_{ij} is the estimated percentage composition of material class i in the image j, L_j is the number of patches analysed in the image j, A_{patch} is the area of each patch, and d_i is the area density for class i.

The performance of the methodology was measured in terms of speed, or time to complete a composition analysis, and the bias, accuracy and precision based on multiple measurements of the composition of each of the material classes. The bias was indicated by the mean error (ME), the accuracy by the mean absolute error (MAE), and the precision was reported with the standard deviation (SD), calculated with Equations (6-3), (6-4) and (6-5):

$$ME_i = \frac{1}{n} * \sum_{j=1}^{n} (E_{ij} - C_i) \quad \textbf{(6-3)}$$

$$MAE_i = \frac{1}{n} * \sum_{j=1}^{n} \left|E_{ij} - C_i\right| \quad \textbf{(6-4)}$$

$$SD_i = \sqrt{\frac{1}{n} * \sum_{j=1}^{n} (E_{ij} - C_i)^2} \quad \textbf{(6-5)}$$

Where i is a material class, with i<=21 , j is the image number, n is the total number of images of the mixed waste captured in this experiment, E_{ij} is the estimated percentage composition of material class i in the image j, and C_i is the true percentage composition of material class i in the mixed waste considered for the experiment, calculated from the areas from Table 6-1.

The overall performance for a single experiment to classify a test mixture was indicated by the minimum and maximum values of ME, the average MAE and the average SD values, calculated across the 20 different material classes in consideration. The performance measures for each material class were also scaled, as a percentage of the true composition, except when the material class was not present in the analysed mixture.

6.2.6 Experimental design

A series of three experiments was designed to evaluate the performance of the methodology on mixed waste streams. The outputs show the effects of varying quantities and fractions in the waste stream, of using non-uniform illumination, and to assess the limitations imposed by the methodology spatial resolution. In every experiment, trained deep networks were used to estimate the composition of the different fractions of the prepared test material mixtures. Each test mixture was analysed multiple times. Each time, all the materials in the mixture were

randomly spread out on the target area, and a picture was taken for further analysis. A total of 20 images were captured for mixture 1 whilst 30 pictures were taken of each of the fractions from mixtures 2 and 3.

The first experiment used a deep neural network previously trained to recognize 21 material classes (Chapter 5), including the 16 types of material samples used in this study. This network previously achieved an average accuracy per class of 88.5% with images with object resolutions from 50 to 289 PPI (Chapter 5), which includes the resolution of 113 PPI used in this study. However, the network was trained using uniform lighting, and the images in this study are affected by an illumination fall-off towards the edges of the field of view, as explained in Section 6.2.2. Also, this study used images of mixed materials, whilst the previous results were obtained with test patches containing only one material class (Chapter 5).

The second and third experiments assessed the effects and limitations of resolution of the methodology. Both experiments used new patches obtained from images of waste captured with different exposure levels (Section 6.2.2). The resulting databases were balanced with an average number of 60,000 unique patches per class. The training databases were used to fine-tune the trained network from the first experiment. Experiment 2 used the original size of the patches, whilst experiment 3 upscaled these patches with a factor of 2. As described in Section 6.2.4, the training and classification was therefore performed with patches representing 5.1 by 5.1 cm object size for the second experiment, and 2.6 by 2.6 cm on the third experiment.

The trained networks of each of the three experiments were used to analyse the composition of the multiple images of each mixture (Table 6-1). The analysed images have a size of 4,896 by 3,264 pixels, captured with a 113 PPI spatial resolution, thus representing a field of view of 0.8 m^2. For some of the experiments, the test images were upscaled by a factor of 2. The images were divided into test patches sized 227 by 227 pixels, each representing an area of 26 cm^2 when the original image size is used, or 6.5 cm^2 when using an upscaled image. A total of 294 test patches were obtained from each original image, covering 95% from the central region, whilst 5% from the outside borders was

ignored. When working with upscaled images, 1,204 test patches were analysed, covering 97% of the central region, and 3% from the outside borders was ignored.

6.3 Results

In this section, the results obtained by each of the experiments are presented and described.

6.3.1 Experiment 1

Experiment 1 used a previously trained network to classify the images of mixed materials using both classification Methods A and B, using the top-1 and all probabilities, respectively, to determine the composition of the test images. The results show no significant differences between Methods A and B in terms of bias, accuracy, and precision performances for the classification of any material class in any of the test mixtures.

The magnitude of errors, indicated by the MAE, increased from a mean of 3.3% when classifying mixture 1 to 5.8% with all the fractions in mixture 3 (Table 6-4). The mean standard deviation increased from 1.1% when classifying mixture 1 to 3% when classifying all the fractions in mixture 3, though this value is reduced to 1.9% and 0.9% when classifying the light and shredded fractions, respectively.

Similarly, the range of bias, or difference between the maximum and minimum ME values obtained, increased from 20.7% in mixture 1 to 42.5% in mixture 3. The material class most underestimated in the mixture 1 was *wood* samples with a ME=-8.1%, whilst *red tissue paper* in the shredded fraction of mixture 2 resulted in ME=-15%. *Household cardboard boxes* were the most overestimated materials in mixture 1, with ME=12.6%, but the same class resulted in ME=33.6% for the shredded fraction of mixture 3.

Table 6-4. Overall scaled performance indicators obtained with the trained network from Experiment 1, using classification methods A and B, considering the top-1 and all probabilities, respectively, to determine the content of each test patch.

Mixture (fraction)	Method A				Method B			
	ME		MAE	SD	ME		MAE	SD
	min	max	mean	mean	min	max	mean	mean
1 (all)	-8.1%	12.6%	3.3%	1.1%	-8.0%	12.8%	3.2%	0.9%
2 (all)	-13.7%	12.2%	3.9%	2.5%	-13.7%	12.1%	3.8%	2.2%
2 (light)	-14.5%	14.1%	3.6%	1.7%	-14.6%	15.5%	3.6%	2.0%
2 (shredded)	-15.0%	28.9%	6.0%	0.7%	-15.0%	28.1%	5.3%	0.7%
3 (all)	-8.9%	16.2%	5.8%	3.0%	-7.7%	11.4%	3.4%	2.6%
3 (light)	-8.1%	13.7%	3.7%	1.9%	-8.1%	13.9%	3.5%	1.7%
3 (shredded)	-8.8%	33.6%	5.8%	0.9%	-8.8%	32.8%	5.3%	0.8%

6.3.2 Experiment 2

Experiment 2 trained the deep network with images using different exposure levels. No significant difference was found when comparing the compositions for any of the test mixtures when comparing the two classification methods. When using the classification from Method A, the magnitude of errors for all the fractions of mixture 3 was of 3.6%, compared to 1.9% obtained for mixture 1 (Table 6-5). A similar increase was obtained in the standard deviation, from a mean SD=0.9% for mixture 1 to 3% for all the fractions in mixture 3. In terms of bias, the range of values between the maximum and minimum ME values obtained with method A increased from 15.3% in mixture 1 to 16% in all fractions of mixture 3 (Table 6-5).

Table 6-5. Overall scaled performance indicators obtained with a network trained with multiple exposure levels in Experiment 2, using classification methods A and B, considering the top-1 and all probabilities, respectively, to determine the content of each test patch.

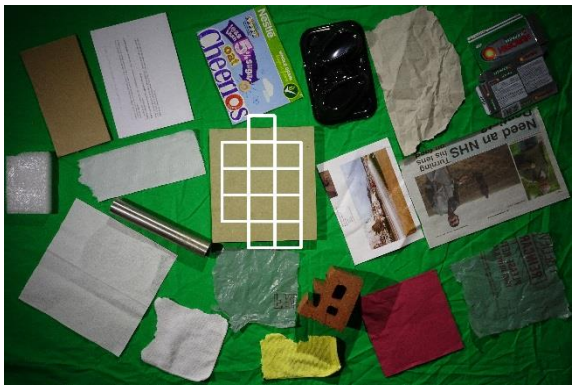
	Method A				Method B			
Mixture (fraction)	ME		MAE	SD	ME		MAE	SD
	min	max	mean	mean	min	max	mean	mean
1 (all)	-4.8%	10.4%	1.9%	0.9%	-4.9%	10.3%	1.9%	0.8%
2 (all)	-13.2%	5.0%	3.3%	2.7%	-13.3%	12.6%	3.9%	2.9%
2 (light)	-13.0%	7.8%	2.8%	2.2%	-14.2%	9.5%	3.3%	2.8%
2 (shredded)	-14.6%	12.1%	3.6%	0.8%	-14.7%	16.3%	3.7%	0.9%
3 (all)	-7.6%	9.6%	3.6%	3.0%	-7.4%	8.4%	3.3%	2.8%
3 (light)	-8.0%	6.4%	2.9%	2.8%	-6.9%	11.7%	2.9%	2.2%
3 (shredded)	-10.4%	17.4%	4.0%	0.9%	-10.1%	13.0%	4.2%	0.8%

An example classification of the test mixture 1 shows that the network trained in Experiment 1 correctly classified six patches from a *wood* sample in central area of the image (Figure 6-1a), equivalent to 0.015 m^2 or 34% of the true area of the *wood* sample in this mix. When the wood sample was in a corner of the image, the same network failed to correctly recognize a single patch (Figure 6-1b). In both images, two and three patches, respectively, with *tissue paper* contents were wrongly identified as *wood* material.

The same images classified using the network from experiment 2 and method A, resulted in 12 patches from the *wood* sample placed in the centre being correctly classified, equivalent to 0.03m^2, or 69% of the true area of *wood* in the sample (Figure 6-1c). With the *wood* sample in the corner, 8 patches were correctly classified, equivalent to 0.021 m^2 or 46% of the true area of the *wood* sample (Figure 6-1d). In the same image, one patch from the *green background* was wrongly identified as *wood* material.

a) Experiment 1, mixture 1, with wood in the centre

b) Experiment 1, mixture 1, with wood in the top-left corner



c) Experiment 2, mixture 1, with wood in the centre

d) Experiment 2, mixture 1, with wood in the top-left corner

Figure 6-1. Results from experiments 1 and 2 to recognize the percentage of wood in mixture 1, using classification method A

The *household cardboard boxes* class was the most overestimated when estimating composition of all the fractions in any of the mixed material samples, with ME=10.4% for mixture 1, and ME=9.6% for mixture 3. *Wood* material was the most underestimated in mixture 1 by an average of 4.8%, but *red tissue paper* was the most underestimated when analysing any of the fractions from mixtures 2 and 3, except the shredded fraction of mixture 3. In this case, the most underestimated class was *office paper* with ME=-10.4%.

A visual analysis of the classification of cardboard material, which included the most overestimated class (*household cardboard boxes*), indicated that the network correctly identified the *corrugated* and *household cardboard boxes*, except for some areas towards the edges of these materials (Figure 6-2). On the other hand, it wrongly classified patches as *cardboard* when they contained

edges of other materials, except in one case where a patch entirely covered by red tissue paper (Figure 6-2a).

a) Experiment 2, mixture 1, corrugated cardboard (top-left corner)

b) Experiment 2, mixture 1, household boxes (top centre and top-right corner)

Figure 6-2. Results from experiment 2 to recognize cardboard material classes in mixture 1, using classification method A

6.3.3 Experiment 3

The deep network in Experiment 3 was trained using upscaled versions of the patches used in Experiment 2. The classification Methods A and B were applied to estimate the composition of images upscaled by the same scale as the training patches. The resulting spatial, or object resolution of each training and test patch is the same as Experiments 1 and 2 (113 PPI), but each patch in this experiment represents a square of 2.6 by 2.6 cm instead of 5.1 by 5.1 cm.

Similar to the previous two experiments, no significant difference was found between the classification results from Methods A and B (Table 6-6). The magnitude of errors for all the fractions of Mixture 3 was of 3%, compared to 1.7% obtained for Mixture 1. A similar increase was obtained in the standard deviation, from a mean SD=0.6% for Mixture 1 to 2.7% for all the fractions in Mixture 3. The range between the maximum and minimum ME values was reduced from 13.6% in Mixture 1 to 23.2% in all fractions from Mixture 2, but the range for all fractions in Mixture 3 was reduced to 13.1%.

Table 6-6. Overall scaled performance indicators obtained with the trained network from Experiment 3, using classification methods A and B, considering the top-1 and all probabilities, respectively, to determine the content of each test patch.

	Method A				Method B			
Mixture (fraction)	ME		MAE	SD	ME		MAE	SD
	min	max	mean	mean	min	max	mean	mean
1 (all)	-4.5%	9.1%	1.7%	0.6%	-4.6%	9.5%	1.7%	0.6%
2 (all)	-11.8%	11.4%	3.5%	2.7%	-11.8%	10.9%	3.5%	2.6%
2 (light)	-13.0%	7.8%	3.1%	2.5%	-12.9%	7.4%	3.1%	2.4%
2 (shredded)	-13.2%	13.6%	3.6%	0.6%	-13.2%	13.5%	3.5%	0.6%
3 (all)	-6.2%	6.9%	3.0%	2.7%	-6.2%	7.0%	3.0%	2.7%
3 (light)	-5.6%	8.5%	2.7%	2.1%	-5.6%	8.6%	2.6%	2.1%
3 (shredded)	-9.9%	14.6%	3.6%	0.7%	-7.1%	14.5%	3.6%	0.7%

The most overestimated class in mixture 1 was *household cardboard boxes* with ME=9.1% (Figure 6-3). When analysing all the fractions in Mixture 2, the most overestimated materials were the *corrugated cardboard* with ME=11.4%, whilst in the case of all the fractions of Mixture 3, the most overestimated material were the *glossy magazines* (ME=6.9%). The most underestimated materials in Mixture 1 was *raw and processed wood* with ME=-4.5% (Figure 6-3), R*ed tissue* was the most underestimated material in all other fractions from Mixtures 2 and 3, except for the shredded fraction of Mixture 3, where the most underestimated material was *office paper* with ME=-9.9%. The scaled mean average error on Mixture 1 was 33.7%.

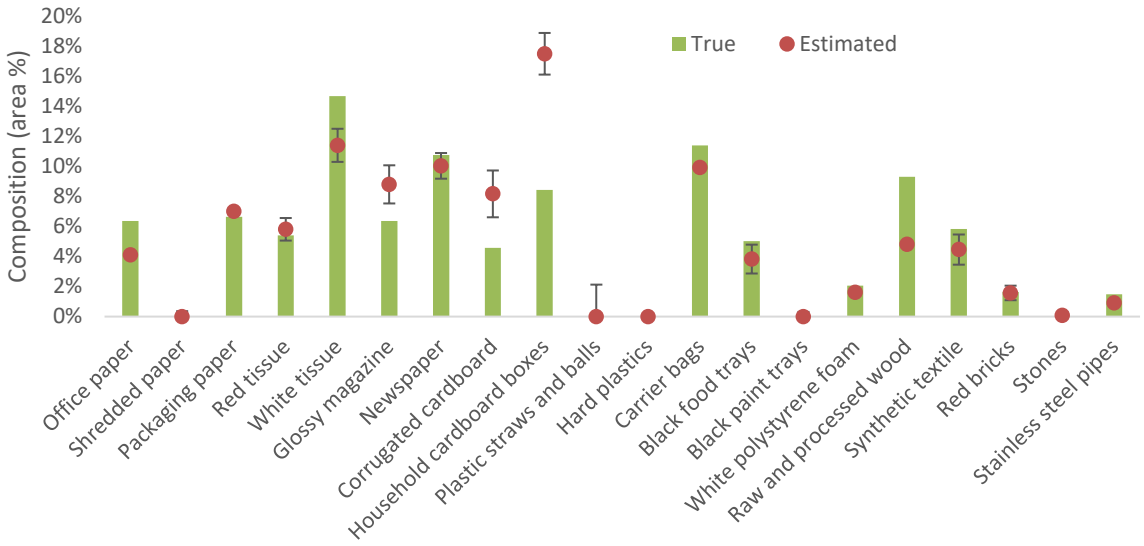


Figure 6-3 – True area (%) composition and the estimates from the analysis of Mixture 1, using the network from Experiment 3 and method A. The vertical bars indicate the standard deviation.

The estimated weight composition of Mixture 1 was calculated from the estimated area composition (Figure 6-4). The most over-estimated class was the *red bricks* (ME=12.5%), whilst the *raw and processed wood* class was the most under-estimated (ME=-15.5%). Both classes of cardboard (*corrugated* and *household boxes*) were over-estimated with ME=1.3% and ME=1.6%, respectively. The scaled mean average error of 42.6%, based on the estimated weight composition of Mixture 1, was higher than the scaled MAE of 33.7% based on the estimated area composition of the same mixture.

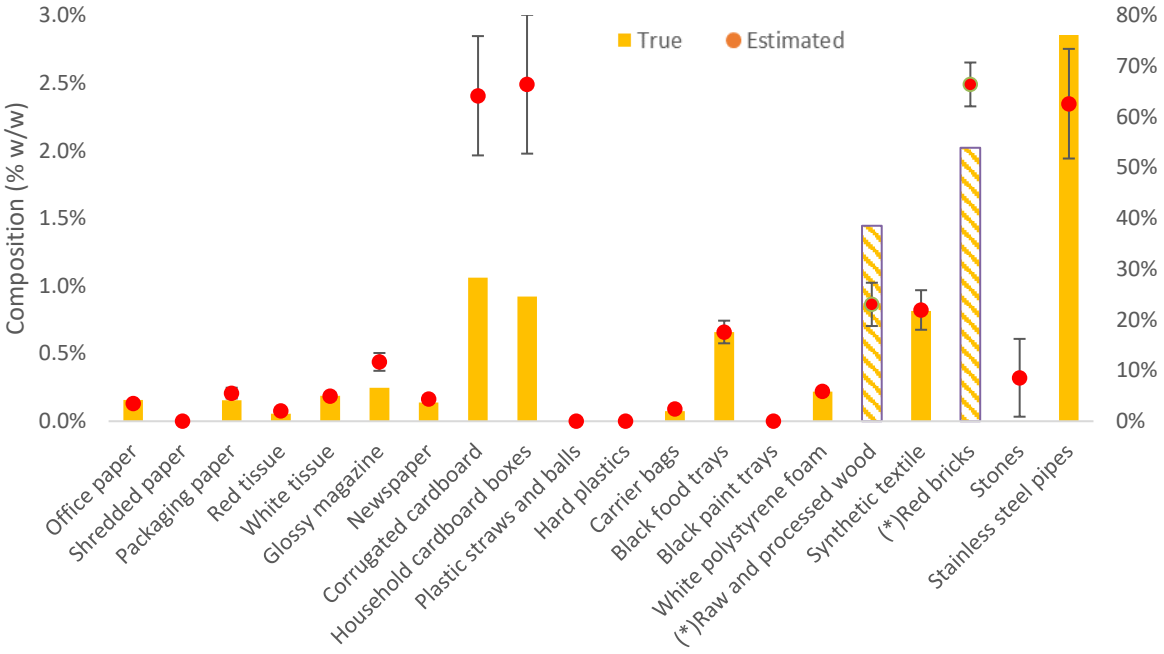


Figure 6-4. True weight % composition and the estimates from the analysis of mixture 1, using the network from experiment 3 and method A. The vertical bars indicate the standard deviation. (*) Note that the wood and brick compositions are shown on a different scale indicated by the secondary axis due to their higher values.

The analysis of an image from Mixture 1 with a *wood* sample in the central area resulted in the correct classification of 59 patches, representing an area of 0.0383 m^2 or 85% of its true size (Figure 6-5a). Four patches in the same image were incorrectly identified as *wood* material, with an equivalent area of 0.0026 m^2. When the same *wood* sample is placed in the corner (Figure 6-5b), 34 patches were correctly classified as *wood*, representing 0.0221 m^2 or 48% of the true size, whilst 6 patches were classified incorrectly as *wood*, equivalent to an area of 0.0039 m^2.

a) Experiment 3, mixture 1, with wood in the centre

b) Experiment 3, mixture 1, with wood in the top-left corner

c) Experiment 3, mixture 1, plastic film

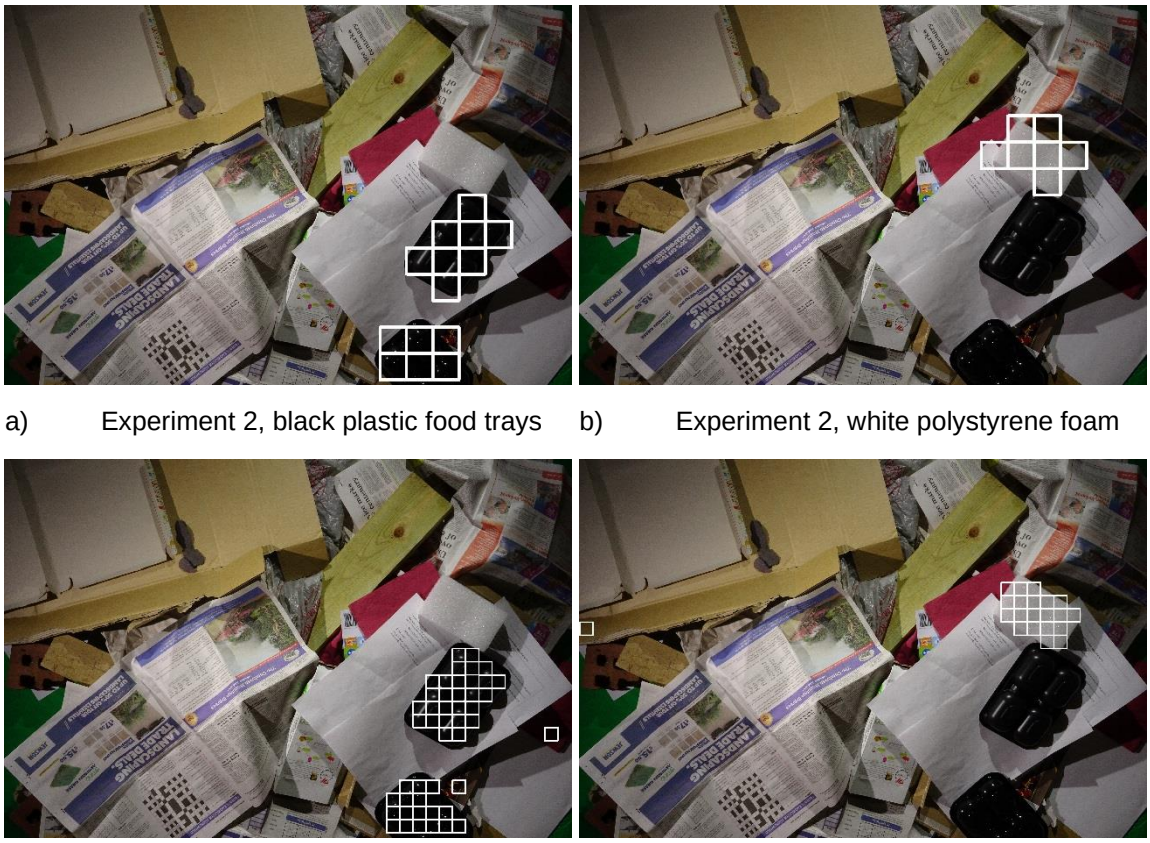
d) Experiment 3, mixture 1, newspaper

Figure 6-5 – Visualisation of partial results from experiments 3 to recognize the percentage of materials in mixture 1, using classification method A

On a different example image from Mixture 1, 90 patches were correctly identified as *film plastic* (Figure 6-5c), representing an area of 0.0584 m^2 or 104.7% of the true size of the samples in the mixture, whilst 3 patches were classified incorrectly, equivalent to 0.0019 m^2. Another example image from Mixture 1, analysed to determine the composition of *newspaper*, shows that 81 patches were classified correctly, corresponding to 0.0526 m^2 or 100% of the true size (Figure 6-5d). The same analysis shows 22 patches wrongly classified as *newspaper* material, equivalent to area of 0.0143 m^2.

A visual analysis illustrated how the patches from Experiments 2 and 3 can determine the composition of mixed waste. No patches were wrongly identified as either white polystyrene or black plastic tray by the network trained in Wxperiment 2 (Figure 6-6a,b). However, the network missed some areas with

black plastic tray material. With the network trained in Experiment 3, two patches from very dark areas in this example image were falsely identified as black plastic food tray (Figure 6-6c), whilst 1 patch corresponding to cardboard was classified as white polystyrene foam (Figure 6-6d).

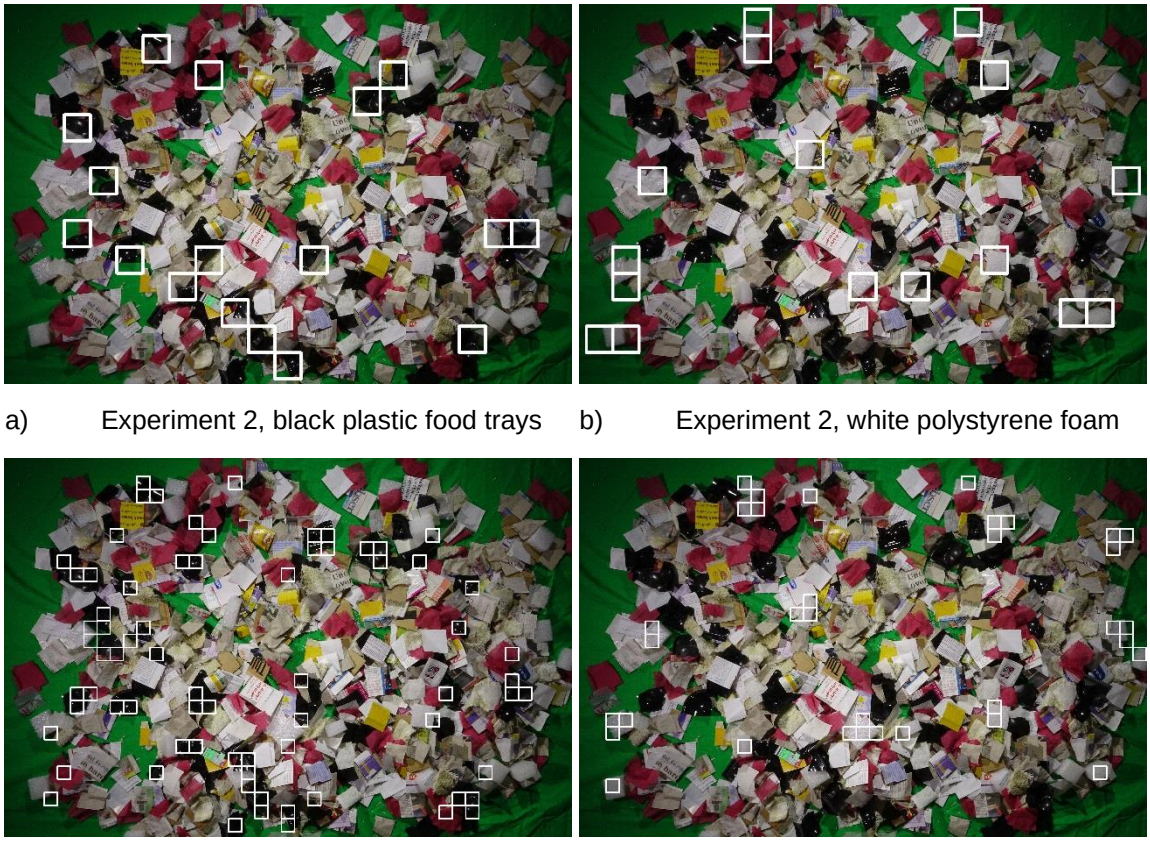


a) Experiment 2, black plastic food trays b) Experiment 2, white polystyrene foam

c) Experiment 3, black plastic food trays d) Experiment 3, white polystyrene foam

Figure 6-6 – Visualisation of partial results from experiments 2 and 3 to recognize the black plastic food trays and white polystyrene foam in an image of all fractions from mixture 2, using classification method A

Further visual analysis of the classification of *polystyrene foam* and *black plastic trays* in the shredded fraction of Mixture 3 illustrated how the network was able to partially recognize the visible materials in the image. With the network trained in Experiment 2, the method correctly classified 17 patches with *black plastic* material, which correspond to an area of 0.0441 m^2, with no wrong classifications (Figure 6-7a).



a) Experiment 2, black plastic food trays b) Experiment 2, white polystyrene foam

c) Experiment 3, black plastic food trays d) Experiment 3, white polystyrene foam

Figure 6-7 – Visualisation of partial results from experiments 2 and 3 to recognize the black plastic food trays and white polystyrene foam in an image of all fractions from mixture 3, using classification method A

The classification results from this example image and the network from Experiment 3 showed 41 patches (0.0656 m^2) identified as *black plastic tray* material, with 6 incorrect classifications (Figure 6-7c). Each correctly classified patch covered an area which partially or completely included *black plastic tray*. Both networks failed to recognize every patch with *black plastic tray* material visible. Similar results were observed when both networks were used to classify *white polystyrene foam* material (Figure 6-7b,d), but in this case no patches were incorrectly classified.

The speed performance was measured for the composition analysis using the networks from Experiments 2 and 3. With an input image resolution of 4,896 by 3,264 pixels, and a patch size of 227 by 227 pixels, the number of patches analysed by the Experiment 2 were 294 per image. In the case of Experiment 3,

the resolution of the input image was upscaled by a factor of 2, which resulted in 1,204 patches analysed per image. The composition analysis in Experiment 2 took an average of 445 ms per image, which increased to 1,923 ms per image for the composition analysis in Experiment 3.

6.4 Discussion

Two classification methods were implemented to determine the composition of images with mixed materials. The analysis was performed by patches, where each patch could have a mixture of different materials. Method A classified each patch as one material, corresponding to the one with the highest probability determined by the trained network. Method B considered each of the probabilities as the relative composition ratio, and every patch was described with the partial composition of each of the possible 21 classes. However, this methodology did not show a significant difference to the results from Method A. Similar classification errors occurred with either method in the presence of mixed materials. This is because the patches represent small areas, and they usually include a single material type. The conclusions that follow are based on the analysis of the results from Method A, but the same applies to the results from Method B.

Experiment 1 reused a deep neural network trained in a previous study using uniform lighting to analyse the composition of mixed materials from images captured using non-uniform illumination. The results on all the different fractions showed an average mean absolute error (MAE) of 4.5%. Experiment 2, which trained the network using images obtained different exposure levels, resulted in an average MAE of 3.2% across all the mixtures classified. The difference in accuracy between Experiment 1 and 2 clearly indicates the negative effect of different illumination levels on the classification performance. The relative drop of performance agrees with the results from a previous study (Chapter 4), which measured the classification rates for images of different brightness levels. The illumination intensities across the field of view changed the brightness of the material samples depending on their location in the image. This was asserted by a visual analysis and comparison of the classification results of patches from the

same samples, captured with different brightness levels but classified with the networks from Experiments 1 and 2, as the exampled presented in Figure 6-5.

When classifying test Mixture 1, the 0.5 m^2 of material samples are spread out in the 0.8 m^2 field of view, as they would appear on a conveyor belt after going through a vibratory feeder. Therefore, every sample is made entirely visible to the camera, and the area composition of the captured images is equivalent to the true composition of the mixtures. The best overall performance on this test mixture was obtained with the network trained in Experiment 3, which analysed patches with an equivalent patch size of 2.6 by 2.6 cm.

The scaled results from the analysis of Mixture 1 in Experiment 3, calculated only for the 16 materials classes included in the Mixture as the proportion of the error to the true area, showed a scaled mean average error of 33.7%. This compares negatively to the results from a previous study (Chapter 5), which resulted in an average accuracy per class of 87.9%. However, this previous study classified patches containing only one visible material, whilst this work classified patches with mixed materials. In the case of Mixture 1, with no sample overlapping, the patches from the edges of the material samples mostly include green background material, and in some cases a second or third class of waste.

The negative effects of classifying patches with mixed materials were clearly visualised in Figure 6-6. The convolutional neural networks were trained using patches containing no background and were instead completely covered by only one material class. In the classification experiments in this work, the trained network recognized the edges against the background, or contrast generated by combination of different materials, as features or characteristics from other materials, not necessarily present in the patch. Based on visual analysis of the results, such as the example observations from Figure 6-6, it was evident that for Mixture 1, the network identified most of the partial patches as *household food and drink boxes*. These cardboard items are typically painted with different colours and with large shapes and irregular visual features. Also, the back sides of these cardboard boxes are visually similar to *corrugated cardboard*, which was also over-estimated. This would explain the over-estimation of the cardboard

classes observed in all three experiments when classifying the Mixture 1, as the example shown in Figure 6-3.

The images for all the experiments were captured using the same spatial resolution (113 PPI), which meant that reducing the size of the patches from 5.1 by 5.1 cm (experiments 1 and 2) to 2.6 by 2.6 cm (experiment 3) reduced the amount of visual information in each patch. However, this also increased the resolution of the methodology, providing more data points from the same image in order to estimate its composition. Thus, the measured MAE of 1.7% and standard deviation (SD) of 0.6% from Experiment 3 are smaller than those from Experiment 2 (MAE=1.9% and SD=0.9%).

The results from Experiment 3 indicate that the accuracy and precision of the methodology improved when training and testing patches a quarter of the size of those used in Experiments 1 and 2. This suggests that these smaller patches contain enough visual information to extract as identifying features to classify the waste materials. Using a higher quantity of smaller patches also increases the proportion of patches containing exclusively one material, which based on previous results, could be classified with a rate of 87.9% (Chapter 5).

Using even smaller patch sizes, for example 1.3 by 1.3 cm would further increase the methodology resolution and generate a higher proportion of patches with only one material class, improving the classification performance. However, this would also reduce the information available for training and classifying, which might be compensated by acquiring images with higher spatial resolution. For example, this could be achieved by imaging a smaller field of view or by using a camera sensor with a higher resolution, or a combination of both (Chapter 5). However, using more patches would also require more time to complete a composition analysis. An alternative solution would be to generate training patches covered only partially with materials, and using different materials as backgrounds.

In contrast to Mixture 1, the materials in Mixtures 2 and 3, which contain 2.2 m^2 and 5 m^2 of samples, respectively, overlapped each other on top of the 0.8 m^2 field of view. As expected, the accuracy performance of the methodology decreased as the number of patches with only one visible material was reduced.

This was evidenced by the overall results obtained when using the networks trained in Experiment 1, 2 and 3 to classify all the fractions in the test mixtures. Each of the networks performed worst on Mixtures 2 and 3 than on Mixture 1, indicated by an average increase of the mean average error value by a factor of 1.75, which indicates the negative effect of mixed materials on the accuracy of the methodology.

A comparison of results from all three experiments to analyse all fractions in Mixtures 2 and 3 and all fractions in Mixture 1, show that the mean standard deviation increases by an average of 3.4. This indicates a loss of precision when the quantities of waste increased. As the materials overlap each other, some materials would be over-represented in the captured images whilst the rest of the samples are not proportionally visible, or even completely invisible. For example, a large sample of corrugated cardboard could cover fall on top of the mix, covering most of the field of view and blocking all the materials underneath. On a different picture of the same mixture, this same sample of corrugated cardboard could be under the rest of the materials. Therefore, the variability of the composition analysis is affected by the overlapping and spreading of the different material samples in this methodology, which is indicated by the increasing standard deviation values observed.

However, the precision of the methodology improved when classifying the shredded fraction of the test mixtures. The shredded fractions in these experiments are composed of light materials with particles of similar sizes, between 3.5 cm and 5 cm each side. These shredded materials constitute a homogeneous mixture (Séverin et al., 2010; Wagland et al., 2011), which prevents larger samples covering smaller ones such as those present in the heterogeneous light fractions. This increases the visibility of all the materials proportionally to their composition and reduces the variability in the images of these mixtures. Experiments 2 and 3 resulted in an average SD of 2.5% and 2.3%, respectively, when classifying the light fraction of Mixtures 2 and 3. When using the same trained networks from Experiments 2 and 3 to analyse the composition of the shredded fraction of Mixtures 2 and 3, the average SD values

obtained were 0.9% and 0.7%, respectively, which correspond to an average reduction factor of 0.33.

When estimating the weight compositions of mixture 1 (Figure 6-4) based on the area composition measured in experiment 3, the scaled mean average error increased from 33.7% to 42.6%. This loss of accuracy is explained by the range of area densities values of the different classes of materials considered (Table 6-6). For example, the *red brick* class was one of the most accurate area compositions with ME=-0.1% and MAE=0.2% (Figure 6-3). When calculating the weight composition using equation (6-2), these error values increased to ME=12.5% and MAE=12.5%, resulting in *red bricks* being the most over-estimated class by weight in the mixture. Similarly, the *wood* class was under-estimated in the area composition with ME=-4.5 and MAE=4.5%. But these metrics were affected by the high density of *wood*, indicated by the estimated ME=-15.5% and MAE=15.5% based on the weight composition, resulting in *wood* becoming the most under-estimated class by weight in the mixture.

The accuracy of the methodology to classify the shredded fractions, similar to a typical SRF or RDF output, compares negatively to the results by Peddireddy et al. (2015) obtained using human perception. Their average MAE of 4.4%, calculated from their results, is lower than the average MAE of 8.8% based on the estimated weight composition obtained in this study using the network from Experiment 3. This is no evidence that this method is inferior, though, because this study classified the waste materials into 21 classes, compared to only 5 classes in their work. Furthermore, the types of materials used in their study have a range of relative densities between 1 (*film plastics*) and 4.1 (*wood*), whilst the range of relative densities in this study is between 1 (*carrier bags plastic*) and 4,692 (*red bricks*). As discussed above, this wide range of densities can accentuate small errors in area composition that lead to higher estimation errors in the weight composition.

Additionally, the accuracy of the estimated weight composition compares favourably to the results obtained in another study using human perception to classify unprocessed waste materials (Wagland et al., 2012). The average MAE

of 2.4 and 3.2% obtained in Experiment 3 from the weight composition analysis of Mixtures 2 and 3, respectively, were both lower than the overall MAE=5.1% calculated from their results. This indicates that the methodology developed in this study can classify solid waste materials more accurately than previous efforts, even when undertaking a more complex classification task, using more classes of materials and a higher range of densities. Moreover, the methodology from this work can be automated and implemented in real-time, as it is able to provide a composition analysis of an area of 0.8 m^2 every 0.5 to 2 seconds. In contrast to any manual inspection, it can analyse waste streams continuously without performance degradation due to human.

6.5 Conclusions

This is a new application of computer vision and deep learning towards the characterisation of mixed solid waste as it can be found in MRF, in either upstream waste flows or in shredded outputs prepared for energy recovery. This research evaluates the performance of the methodology developed previously (Chapters 3, 4, and 5) on a new type of task. The steps include the mixing and spreading of mixed materials on the field of view, capturing the images using non-uniform illumination, using deep networks for image classification, and determining the composition of the waste.

The results from this work are not directly comparable to any previous results, as this study estimated area composition, instead of counting number of patches or items. However, the average accuracy per class, measured through the average MAE per class, exceeds the industry expectations while providing more detail analysis than current industry practices.

The results quantified the effects of analysing increasing quantities of materials to total areas greater than the field of view of the camera. The mean error and the standard deviation increased by factors of 1.75 and 3.4, respectively, due to the materials overlapping each other. However, when the same samples were shredded, the standard deviation was reduced with a factor of 0.33, which indicates an improvement of the precision of the methodology. The overall accuracy and precision performance of the presented methodology was better

than previously obtained results to analyse both heterogeneous and shredded, or homogeneous, fractions of mixed waste. Finally, the results indicated that smaller patches could improve both accuracy and precision performance, but with a speed performance trade-off, suggesting that augmenting the training dataset to represent the partially covered patches would be better.

This study extends previous research results, by automating the waste analysis while providing operationally acceptable performance. The results clearly show the potential of this methodology to waste composition analysis. The precision measured in this study, combined with the speed of the process make a strong case for further investigations towards the deployment of this technology for automated real-time analysis of waste materials and waste derived fuels. With new legislation requiring more data and with more granularity, this methodology could provide a solution to a pressing industry need.

6.6 References

Adedeji, O., Wang, Z., 2019. Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network. Procedia Manufacturing, The 2nd International Conference on Sustainable Materials Processing and Manufacturing, SMPM 2019, 8-10 March 2019, Sun City, South Africa 35, 607–612. https://doi.org/10.1016/j.promfg.2019.05.086

Aral, R.A., Keskin, Ş.R., Kaya, M., Hacıömeroğlu, M., 2018. Classification of TrashNet Dataset Based on Deep Learning Models, in: 2018 IEEE International Conference on Big Data (Big Data). Presented at the 2018 IEEE International Conference on Big Data (Big Data), pp. 2058–2062. https://doi.org/10.1109/BigData.2018.8622212

Awe, O., Mengistu, R., Sreedhar, V., 2017. Final Report: Smart Trash Net: Waste Localization and Classification [WWW Document]. URL http://cs229.stanford.edu/proj2017/final-reports/5226723.pdf (accessed 11.8.19).

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Bircanoğlu, C., Atay, M., Beşer, F., Genç, Ö., Kızrak, M.A., 2018. RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks, in: 2018 Innovations in Intelligent Systems and Applications (INISTA). Presented at the 2018 Innovations in Intelligent Systems and Applications (INISTA), pp. 1–7. https://doi.org/10.1109/INISTA.2018.8466276

David McKinnon, Jak Fazakerley, Rolph Hultermans, 2017. Waste Sorting Plants: Extracting Value from Waste [WWW Document]. URL http://www.iswa.org/index.php?eID=tx_iswaknowledgebase_download&documentUid=5071 (accessed 11.8.19).

DEFRA, 2014. Materials facilities: how to report on mixed waste sampling [WWW Document]. URL https://www.gov.uk/government/publications/materials-facilities-how-to-report-on-mixed-waste-sampling (accessed 11.8.19).

DEFRA, 2018. 25 Year Environment Plan [WWW Document]. URL https://www.gov.uk/government/publications/25-year-environment-plan (accessed 10.3.19).

DEFRA, 2019. Consultation on reforming the UK packaging producer responsibility system [WWW Document]. URL https://consult.defra.gov.uk/environmental-quality/consultation-on-reforming-the-uk-packaging-produce/ (accessed 11.9.19).

DEFRA, 2021. Extended Producer Responsibility for Packaging [WWW Document]. URL https://consult.defra.gov.uk/extended-producer-responsibility/extended-producer-responsibility-for-packaging/ (accessed 11.8.21).

EU Commission, 2018. EUR-Lex - 52018DC0029 - EN - EUR-Lex [WWW Document]. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A29%3AFIN (accessed 12.18.19).

EU Council, 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 [WWW Document]. URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:FULL:EN:PDF (accessed 10.19.16).

EU Council, 1999. Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste. Official Journal of the European Communities L 182/1, 1–19.

EU Parliament, 2018. Circular economy package: Four legislative proposals on waste [WWW Document]. URL http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2018)614766 (accessed 4.24.18).

EU Parliament, 2017. The circular economy package: new EU targets for recycling [WWW Document]. URL https://www.europarl.europa.eu/news/en/headlines/society/20170120STO59356/the-circular-economy-package-new-eu-targets-for-recycling (accessed 12.18.19).

Fasting, M.S., 2019. Industry 4.0 in Waste Management (MSc Thesis). Norwegian University of Science and Technology.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv:1408.5093 [cs].

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

Kylberg, G., Sintorn, I.-M., 2016. On the influence of interpolation method on rotation invariance in texture recognition. EURASIP Journal on Image and Video Processing 2016. https://doi.org/10.1186/s13640-016-0117-6

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M., 2014. Industry 4.0. Business & Information Systems Engineering 6, 239–242. https://doi.org/10.1007/s12599-014-0334-4

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Mittal, G., Yagnik, K.B., Garg, M., Krishnan, N.C., 2016. SpotGarbage: Smartphone App to Detect Garbage Using Deep Learning, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '16. ACM, New York, NY, USA, pp. 940–945. https://doi.org/10.1145/2971648.2971731

Nasrullah, M., Vainikka, P., Hannula, J., Hurme, M., Karki, J., 2014. Mass, energy and material balances of SRF production process. Part 1: SRF produced from commercial and industrial waste. Waste Management 34, 1398–1407. https://doi.org/10.1016/j.wasman.2014.03.011

OpenCV library 3.1.0 [WWW Document], 2015. URL https://docs.opencv.org/3.1.0/ (accessed 5.2.18).

Ozkaya, U., Seyfi, L., 2019. Fine-Tuning Models Comparisons on Garbage Classification for Recyclability. ArXiv abs/1908.04393.

Peddireddy, S., Longhurst, P.J., Wagland, S.T., 2015. Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste Management 40, 9–13. https://doi.org/10.1016/j.wasman.2015.03.015

Ruiz, V., Sánchez, Á., Vélez, J.F., Raducanu, B., 2019. Automatic Image-Based Waste Classification, in: Ferrández Vicente, J.M., Álvarez-Sánchez, J.R., de la Paz López, F., Toledo Moreo, J., Adeli, H. (Eds.), From Bioinspired Systems and Biomedical Applications to Machine Learning, Lecture Notes in Computer

Science. Springer International Publishing, Cham, pp. 422–431. https://doi.org/10.1007/978-3-030-19651-6_41

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y

Sakr, G.E., Mokbel, M., Darwich, A., Khneisser, M.N., Hadi, A., 2016. Comparing deep learning and support vector machines for autonomous waste sorting, in: Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International. IEEE, pp. 207–212.

Schindelin, J., 2012. Fiji: an open-source platform for biological-image analysis.

Schneider, C.A., 2012. NIH Image to ImageJ: 25 years of image analysis.

Séverin, M., Velis, C.A., Longhurst, P.J., Pollard, S.J.T., 2010. The biogenic content of process streams from mechanical–biological treatment plants producing solid recovered fuel. Do the manual sorting and selective dissolution determination methods correlate? Waste Management, IWWG Task Group on “Thermal Treatments” 30, 1171–1182. https://doi.org/10.1016/j.wasman.2010.01.012

Sharan, L., Rosenholtz, R., Adelson, E.H., 2014. Accuracy and speed of material categorization in real-world images. Journal of vision 14, 1–24. https://doi.org/10.1167/14.9.12

Vrancken, C., Longhurst, P., Wagland, S., 2019a. Deep learning in material recovery: Development of method to create training database. Expert Systems with Applications 125, 268–280. https://doi.org/10.1016/j.eswa.2019.01.077

Vrancken, C., Longhurst, P., Wagland, S., 2017. Critical review of real-time methods for solid waste characterisation: Informing material recovery and fuel production. Waste Management 61, 40–57. https://doi.org/10.1016/j.wasman.2017.01.019

Vrancken, C., Wagland, S., Longhurst, P., 2019b. Balanced training raw data. 10.17862/cranfield.rd.9968051.

Wagland, S.T., Dudley, R., Naftaly, M., Longhurst, P.J., 2013. Determination of renewable energy yield from mixed waste material from the use of novel image analysis methods. Waste Management 33, 2449–2456. https://doi.org/10.1016/j.wasman.2013.06.021

Wagland, S.T., Kilgallon, P., Coveney, R., Garg, A., Smith, R., Longhurst, P.J., Pollard, S.J.T., Simms, N., 2011. Comparison of coal/solid recovered fuel (SRF) with coal/refuse derived fuel (RDF) in a fluidised bed reactor. Waste Management 31, 1176–1183. https://doi.org/10.1016/j.wasman.2011.01.001

Wagland, S.T., Veltre, F., Longhurst, P.J., 2012. Development of an image-based analysis method to determine the physical composition of a mixed waste material. Waste Management 32, 245–248. https://doi.org/10.1016/j.wasman.2011.09.019

Yang, M., Thung, G., 2016. Classification of Trash for Recyclability Status (Report). Stanford University, http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf (accessed 11.08.19)

CHAPTER 7

DISCUSSION

7 Discussion

The overall aim of this research project was to develop and evaluate a material classification method to enable automated and real-time solid waste composition analysis. The research programme has developed a waste material classification method which has then been validated for the specific application in waste management processes. The classification method was developed in this context as a technique for composition analysis of mixed waste materials. The following sections discuss the key findings and their significance in a wider context of waste management. The main limitations of the research methods are examined with further applications for this new knowledge considered. Recommendations for future research work are then outlined.

7.1 Contribution to knowledge

This original work gained an understanding of the required waste information and investigated solutions to the specific challenges of waste material recognition through deep learning. The findings from this work contribute to knowledge in three main areas: data capture, data processing, and applied deep learning methods (Figure 7-1). The developed classification and composition analysis methods present a flexible and robust alternative to existing solutions. This is a significant development for the waste industry as it offers a technology to retrieve key information in real-time from multiple stages of waste management.

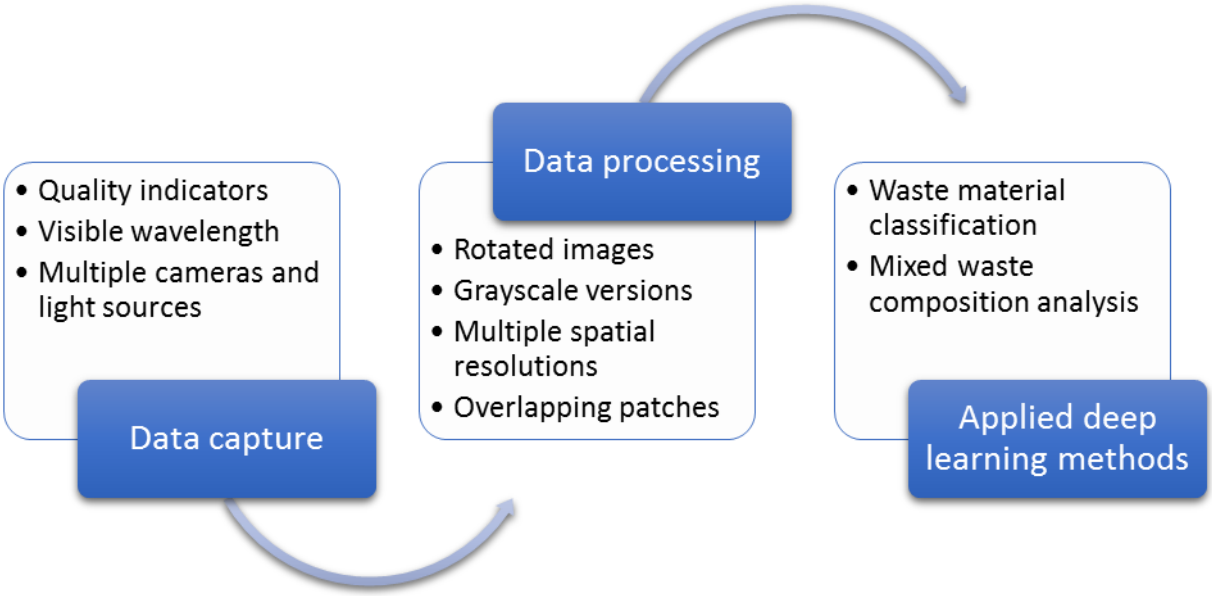


Figure 7-1. Contributions to knowledge

7.1.1 Data capture

Chapter 2 applied a process analytical technology (PAT) approach to material recovery facilities. Consistent with other industries, the successful application of the PAT in an MRF could increase the material recovery levels and ensure the quality of the outputs. The literature review identified the critical quality attributes or **quality indicators**, which are the data required to improve existing waste processes and inform the stakeholders. The review also shows which critical process parameters affect the operation and safety of the different mechanical and biological processes involved in an MRF. These are important observations as they focus the research and development efforts to monitor these indicators in real-time and assist device manufacturers to incorporate them into their designs.

A comparison of advanced sensors and their potential uses in an MRF shows that optical technologies can determine and provide all the identified quality indicators (Chapter 2). Optical sensors can monitor input, intermediate, and output waste streams to provide feedback information in real-time to assist and control each of the mechanical and biological processes implemented in an MRF, increasing the overall efficiency and consistency in output quality. Among optical sensors, **visible wavelength** cameras were selected as the lowest cost and best

performing technology, as they offer information about waste materials that can improve most of the different processes at the different stages in a material recovery facility.

Even though some of the quality indicators or parameters cannot be extracted directly from the images of waste, they can either be estimated indirectly, or complemented with a moisture sensor as previously demonstrated by Wagland et al. (2013, 2012). A visible sensor offers an advantage over near-infrared sensors to identify black objects and to perform any colour-based classifications. Furthermore, a camera-based solution is less complex and typically cheaper than other optical-based equipment such as laser induced breakdown spectroscopy or hyperspectral imaging technologies. These findings determined and justified the choice of a visible wavelength camera to acquire images of waste in the experimental works.

The design of a camera-based solution to be deployed in an MRF would depend on the specific stage or process to monitor and control. The composition, and sizes and shapes of the waste materials change throughout the different MRF stages, from the tipping hall where unprocessed waste is received to the different fractions being extracted by the processes. The methods used by the different processes and the way the material is transported between them can also affect their visible presentation. Finally, it may not always be possible to control the illumination on the field, for example in open air locations or rooms with big windows. Therefore, the methodology developed makes a limited number of assumptions regarding the visual aspect of waste materials and attempts to generalise the classification task for different representations of waste materials.

The developed image acquisition method uses a hardware setup with **multiple cameras and light sources** to capture a diverse set of images for training the deep neural networks. The method resembles the bidirectional texture function techniques which are used to synthesize and represent surface materials based on thousands of images from different camera and light positions (Haindl and Filip, 2013). But the results from the experiments indicate that the implemented combination of five cameras and six illuminations, which generates 30 images

per material sample, contain enough descriptive information to train a deep neural network for the classification of waste materials.

The observations from Chapter 3 indicate that the five installed cameras offer sufficient alternative views of the same material samples to generalise for different positioning of the waste samples on a conveyor belt. The results using five cameras were better but not significantly different than using four cameras, therefore no attempts were made to use a higher number of cameras. Similarly, using six illuminations increase the diversity of the training dataset, including different surface features from each waste sample, either through shadows projected from the different angles, or a lack of shadows when the sample is flooded in light from all angles. Increasing the combination of cameras and illuminations would provide more images. But this would require a larger capital investment to setup the capture hardware, longer processing times, and the findings from Chapter 3 indicate they would not offer a significant gain in performance.

In general, the more data is available to train a deep network, the better the classification performance. However, the experiments in this study assume that it is not always possible to obtain large quantities of material samples. The findings from Chapter 3 indicate that the combination of five cameras and six illuminations can represent a significant relative improvement of 18.4% of the classification rates. A data capture setup with multiple cameras and light sources is the most important factor to increase the diversity and quality of the input images, as the performance gains cannot be matched with data augmentation through image processing. This hardware setup increases the capital investment and requires a dedicated space, compared to a single camera with no controlled lighting, but it ensures that the materials are better described by the captured images. As a trade-off, the findings from Chapter 3 show that the hardware setup compensates for the lack of waste material samples when no further data processing is applied.

The contributions made to data capture in this work focused on classifying waste materials using cameras with using visible wavelength sensors. The same

principles could be applied using different wavelength sensors to retrieve different and complementary information. For example, infrared cameras could capture details that are not present in the visible spectrum which would add information representing the materials. The recommendations to use multiple cameras to capture different angles of view and using multiple light sources to enhance the surface features still apply.

7.1.2 Data processing

In the context of an MRF, the image of a sample can be affected by many external factors, including the arbitrary rotation of materials on top of the conveyor belt. Therefore, waste materials are better represented by adding **rotated images** to the training dataset. Whilst rotation is a common data augmentation technique in deep learning (Taylor and Nitschke, 2017), the angular rotation is usually limited to a certain range and the typical implementation uses random rotations angles within that range. The observations from Chapter 3 indicate that a rotation angle of 30 degrees between training images leads to increased classification performance.

The results from the Chapter 4 experiments demonstrate that removing the colour content and training the network to classify **grayscale versions** of the images have a positive impact on the performance. Even though colour-based classification of materials has been used extensively before for the recognition of objects and materials, these findings support the notion that colour does not add information regarding the waste materials. Furthermore, as the colours are removed and grayscale images are used instead, the trained network becomes more resilient to changes to the colour hue and saturation of waste materials. A colour sensor would not be required, and the captured images would require less storage space, reducing the size of the training dataset to one third of the colour dataset.

The spatial resolution of an image is measured by the number of captured pixels representing a certain distance of the visualised target area. This is determined by the sensor size and resolution, the distance from the sensor to the target area, and the lens characteristics. In the context of an MRF, this parameter could be

controlled when implementing the solution on a fixed position above a conveyor belt. The spatial resolution would change if the optical system is installed on a moving platform, for example a on a grab crane looking at input waste, or a fixed location looking at increasing heights of output materials. The work discussed in Chapter 5 increased the diversity of the training dataset by generating **multiple spatial resolutions** of the captured images. As expected, the overall performance of the network over a range of spatial resolutions was increased. Furthermore, the classification rates were better than using a single, matching resolution for training and testing.

The acquired and processed images are subdivided into smaller patches which are used for network training and for classification purposes. Each patch represents a smaller area of the sample, allowing the network to learn from surface features visible in different positions of the material samples. This is further improved by two extra steps to increase the size and diversity of the dataset, first to extract **overlapping patches**, and the second to extract random, and even smaller crops, from each patch. Random cropping of patches is common practice in deep learning and its benefits have been explained and exploited since the seminal work by Krizhevsky et al. (2012). Overlapping patches has been used before in traditional approaches for texture recognition (Cavalin et al., 2013; Cimpoi et al., 2016; Hafemann et al., 2014) and for material classification using deep learning (Bell et al., 2015). The findings from Chapter 3 indicate that a 25% overlap between patches leads to the highest classification performance.

The variability refers to the difference in the visual appearance of waste materials. The intra-class variability is low for some material classes when the all the samples are similar, for example white office paper. Other materials such as plastic, cardboard, and textile, can have high intra-class variability as one sample can present characteristics such as colours, drawings or writings which are non-existent or different to another sample, or even to another region of the same material sample. In the context of waste materials in an MRF, the variability is further increased because the samples can be placed at any rotation angle on

the conveyor belt. This rotation can be around a vertical axis perpendicular to the conveyor belt, but also around horizontal axis as the materials can be lifted or supported by other materials underneath. Their appearance is also affected by the illumination and the spatial resolution, determined by the environment and the imaging sensor and lenses used.

A low intra-class variability in the training dataset would lead to the network not learning to differentiate the material features from enough representative images, which leads to failure to recognize and classify materials on a conveyor belt. On the other hand, a high intra-class variability poses a difficulty for training a deep neural network to classify the materials, as the model tries to learn the common features in a class. The data capture methodology and the different strategies described above increase the number and diversity of images in the training datasets by using multiple cameras and multiple illuminations (Chapter 3), different illumination levels (Chapter 6), image rotations (Chapter 3), and multiple resolutions (Chapter 6). This increases the intra-class variabilities in the training dataset to generalise the neural network training for the classification of waste materials. The proposed technique controls the range of **intra-class variability** in two ways. Firstly, it ensures a minimum cosine distance between any two training images, to keep only representative and diverse images in the training dataset. Secondly, it splits the classes of materials into several subclasses to reduce the intra-class variability, even if the final classification is between only classes.

The contributions made to data processing in this work focused on increasing the diversity of the images to represent and classify waste materials in different situations. The same principles and recommendation apply to images captured using optical sensors sensitive to other wavelengths. The developed strategies involve converting the colour images into grayscale levels, which makes it compatible to images captured in the infrared. The recommendations to control the intra-class variability, to rotate and include multiple scale resolutions of the images, and to overlap the patches to create the training database still apply.

7.1.3 Deep learning methods

The process monitoring and analytical tools required for the implementation of PAT were developed in Chapters 3 to 6 in the form of material classification and composition analysis methods. The combination of the image capture hardware setup and image processing to augment the training dataset was developed into a technique for training deep neural networks for **waste material classification**. The outcome is a robust classification model, based on a trained deep neural network which generalises to the various presentations of waste. The network can classify waste with variabilities in material rotation, colour, and scale within the ranges of values expected in the context of a waste material classifier or composition analyser.

The technique applied in this research uses a high-resolution camera to capture an image of the waste material and subdivides the image into smaller patches. A trained deep neural network classifies each patch, and a majority voting system determines the class of the visualised material, assuming one sample per image. In the example discussed in Chapter 4, the deep neural network performed better than expected manual sorting rates (Tchobanoglous and Kreith, 2002) and automatic methods using hand-crafted filters (Rahman et al., 2014) for identifying paper and cardboard materials.

Based on the findings from Chapter 6, the classification method performs fast enough to classify every waste material sample moving. In the context of an MRF, waste materials can be individually classified and sorted accordingly, either manually by a worker or by an automatic method using optical sensors. To assist in this, a vibratory feeder is typically combined with an accelerating conveyor belt to separate and spread the materials on the conveyor belt. However, even if more than one sample are visible in the image, a simple image segmentation routine could be implemented to provide a classification of each material.

The test results from Chapter 5 indicate the neural network can classify images captured with a range of spatial resolutions. Whilst this requires extra storage and training time, there is no negative effect to the classification rates. More importantly, the resulting neural network offers flexibility to its implementation and

deployment in an MRF, even without knowing a-priory where the camera is to be installed. To illustrate the tested range of spatial scales, setting this same camera at a distance of 1.5 metres and adjusting the lens to its shortest focal length offers a field of view area of 3.5 m^2. This large area could be analysed for example to classify the input materials at the waste reception space. Furthermore, an area of 0.6 m^2 covering the width of 1-metre wide conveyor belt could be analysed when this camera is placed at any distance between 0.7 and 20 metres by adjusting its focal length. This would allow the camera to be mounted at the ceiling of the facility if necessary.

The developed methodology for **mixed waste composition analysis** uses a visible wavelength sensor camera to capture a high-resolution image of the mixed materials which are divided into small patches. A trained network classifies each patch into different material classes and the composition of the mixture is estimated based on the individual classification results. The method was implemented and tested in Chapter 6 to analyse the composition of mixed materials on a target area representing a 1-metre-wide conveyor belt. The best results were obtained by subdividing the image into 1,204 patches, each representing a 2.6 by 2.6 cm area with a spatial resolution of 113 PPI.

The accuracy metrics indicate that the method performs better with low quantities of mixed material that are fully spread out on the conveyor belt. In this case, the composition is automatically determined with an average scaled error of 34%, which is in range with the expected error rate incurred by a worker performing manual sorting (Tchobanoglous and Kreith, 2002). Another study using deep learning (Awe et al., 2017) obtained a higher accuracy (calculated as mean average precision) of 68.3%, but it was based on the classification of simulated images of waste objects into only three classes of waste. This is to be expected because, as discussed in Chapter 4, the classification performance increases when the number of target classes is reduced, as similar classes are grouped into bigger classes, thus decreasing the number of classification errors.

The precision of the methodology is negatively affected when analysing the composition of larger quantities of mixed waste as the network can only classify

the visible materials from the top layer. However, the precision improves when analysing shredded materials like those in a SRF mixture, as observed previously in study by Peddireddy et al. (2015). The mixed shredded mixture is homogeneous and therefore not affected by materials hiding under other, large samples. The precision for large quantities of shredded materials is not significantly different to the precision observed for low quantities of materials spread out with no overlapping. The precision of the methodology could be further improved by averaging the compositions over a number of samples so that the standard deviation of the mean is lower than a specified error margin.

For the device to be able to provide complete information, it should be able to analyse every sample of material flowing between two processes. The implementations of the methodology in Chapter 6 provide a composition analysis of an area of 0.8 m^2 every 0.5 to 2 seconds, equivalent to 0.4 to 1.6 m^2 analysed per second. Considering a manual sorting station with a conveyor belt just over 1 metre wide moving at a speed of 0.5 metres per second, the method would be able to analyse all the flowing material. Furthermore, in a MRF the conveyor belts between two process units are in average 10 metres long and that they move at speeds up to 3 metres per second (Cimpan et al., 2016; Pichtel, 2005), which translates to a minimum travel time of over 3 seconds. The implementation tested in Chapter 6 could provide real-time feedback in time to control these processing units and allowing for flow diversion or stoppages if necessary.

The evidence from the successive experiments provided independent validation of the trained deep network models. The experiment 1 in Chapter 3 validated the basic data capture and image processing methodology developed in Chapter 2 by testing it with a new set of data. Likewise, the experiment 7 in Chapter 3 and the experiments 1 to 5 in Chapter 4 used separate sets of data to test and validate the method developed in Chapter 3. A qualitative analysis validated the results from Chapter 4 by providing a visual understanding of the decisions made by the model. In experiment 1 of Chapter 5, the classification model developed in Chapter 4 is applied and validated with different sets of data. Furthermore, the performance of the composition analysis method developed in Chapter 5 was

measured in terms of accuracy, precision, and bias when applied to different types and quantities of mixed waste materials. These performance parameters provided evidence that the composition analysis method is fit-for-purpose as it achieves the target values defined as previous performance results in this application area (Peddireddy et al., 2015; Wagland et al., 2012).

7.2 Applications in the wider context

The implementation of the proposed solutions in an MRF creates an opportunity to increase the transparency of waste management by informing the stakeholders. The real-time composition enables down-the-line consumers of RDF or SRF to prepare and adjust their processes accordingly and ensure they meet their regulatory requirements. Regulators could monitor waste segregation levels in real-time instead of the periodic manual sampling to update their recommendations and reduce the environmental impact of waste disposal. Governments can use this data to track the quantity and detailed composition of waste going into landfill which could affect the decision for the development of new technologies, for example for landfill mining. Finally, the proposed methodology provides an opportunity to adjust the gate or tipping fees according to the composition of waste, instead of the current values set solely by the quantity of received waste.

Though the methodology developed in this work focused on material recognition, the same principles can be applied for the recognition of dangerous objects. Over 30% of waste treatment site fires are likely caused by items such as lithium batteries, aerosols, gas cylinders, and flammable liquids (WISH Forum, 2017). The implementation of the methodology at different stages to identify dangerous wastes could prove to be an affordable solution to prevent the common occurrence of accidents and fires. For example, the real-time information could be used to start the appropriate safety control measures, such as stopping the processes and starting a localised water sprinkler system.

The applications of the developed strategies and methodologies offer opportunities outside the scope of a material recovery facility. For example, when implemented at homes or offices, either on “smart bins” using custom illumination

and visible sensors, or on a mobile phone application, they would assist in the correct segregation of waste materials. The solution can be deployed for the transport of waste from one location to another. Bin lorries could use the method to analyse composition of waste during collection and use it to define the appropriate local destination. Similarly, international shipments of waste could be monitored on departure and arrival to increase the transparency of transactions.

Application of this work also creates an opportunity for the waste industry and governments to develop and implement a scheme for extended producer responsibility. This could be achieved by collecting information about the waste materials and objects from the moment they become waste until they reach the end of their cycle, and they are derived for example for recycling or for landfill. Identifying each waste object and matching it to the manufacturer or producer could enable stakeholders to better understand the incurred costs and determine liabilities and future responsibilities.

It can be further speculated that the knowledge gained through this work could be applied outside the context of waste management. The classification and composition analysis methods can be implemented for quality assurance and process control of products based on their surface features. The data strategies developed to recognise material of different scales could be applied to identify impurities in a microscopic level. On the other hand, the composition analysis could be used for example to analyse satellite images, to detect and count specific structures. The food industry could further develop this method to detect not only impurities but the quality of fresh produce.

7.3 Limitations

The hardware setup used five identical cameras with the same settings, but other cameras models and lenses combinations could provide lower or higher qualities of images. Different camera qualities and JPEG compression settings could affect the visible level of details and features, and the networks would need to be trained to recognized those. The methodology in this study was developed so that it could be potentially implemented and deployed with low investment costs. Therefore, the selected camera is a compact, entry-level, and off-the-shelf unit

with a sensor similar to those used in mobile phones, so that the classification could be applied to images capture with those.

The waste material used in the experiments was limited to small quantities to investigate the effects on the performance of the method. The collected waste materials reflect those found in commercial and industrial sources and did not include organics. For the experiments that required it, the material samples were cut to smaller pieces using scissors and guillotines and then mixed manually to create the shredded mixtures resembling SRF. The shredded fraction looks different in an MRF as the size reduction processes produce materials with jagged edges. However, as discussed above, the current implementation of the methodology misclassifies the material edges so the performance would not be affected. The number of samples used for testing the methodologies was also limited and they do not completely reflect the variety of waste materials. Nevertheless, the samples for training and for testing were carefully split into two different groups to ensure a proper test of the generalisation of the trained networks.

The deep learning methodology was limited to one artificial network architecture (AlexNet). Its choice was based on its proven accuracy to classify images and the number of academic publications using it for benchmark comparisons. Deep learning is a very active field and most of the academic research is focused on the development of network architectures. However, the emphasis of this work was on data capture and augmentation, and the recommendations can be transferred and applied to any network architecture.

7.4 Further work

Further research is required to extend and improve the methodology in classification accuracy and speed performance. A set of experiments based on the Chapter 5 methodology but with a higher resolution camera would increase the effective resolution to increase the method accuracy. A review of existing deep network architectures should be conducted to choose a comparison in terms of classification and speed performance. A custom network architecture would be worth exploring and developing to take advantage of the fact that the

methodology only requires one grayscale channel. On the other hand, even though the proposed method recommended using grayscale only, the fact is that the colour information is lost and there might be some examples where the colour channels would facilitate the classification of these materials, so an investigations to combine grayscale and colour images to train the networks is recommended.

Further applied research is needed to establish whether the composition analysis could replace or supplement the current mixed waste composition analysis. Current regulations require waste facilities to sample and test, using laboratory techniques, the composition of mixed waste materials. A minimum quantity of both incoming and target materials needs to be tested according to an annual sampling schedule to ensure representative results. The analysis of incoming materials should report the composition based on paper, glass, metal, plastic and others, whilst purity is reported for the target materials.

Future studies should consider that the appearance of waste materials in an MRF might be different to the material samples used in this work. A prototype deployment in a material recovery facility is recommended to validate the methodology with actual waste material streams. New material samples should be collected from the incoming, intermediate and output stages of the facility, segregated and used for training the deep neural network using the strategies developed in this work. This step would ensure that the network is trained to recognize waste materials which are not necessarily clean, as they have been previously mixed and exposed to other materials, dirt, and environmental variables such as sun and moisture.

A long-term study is recommended for testing to continuously analyse the compositions of the input and output materials, and to analyse the composition of the samples before they are sent to a laboratory for composition analysis. The performance of spot analysis would be validated by the laboratory results, and the methodology could be adjusted accordingly for the real-time analysis, for example by applying a bias correction. The results over 1-year could provide many comparison points with laboratory values to validate the methodology and provide new data regarding accuracy and precision. Furthermore, another study

could be implemented to determine the number of samples required to estimate a mean composition analysis with the required precision level for the facility.

7.5 References

Awe, O., Mengistu, R., Sreedhar, V., 2017. Final Report: Smart Trash Net: Waste Localization and Classification [WWW Document]. URL http://cs229.stanford.edu/proj2017/final-reports/5226723.pdf (accessed 11.8.19).

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Cavalin, P.R., Kapp, M.N., Martins, J., Oliveira, L.E., 2013. A multiple feature vector framework for forest species recognition, in: Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM, pp. 16–20.

Cimpan, C., Maul, A., Wenzel, H., Pretz, T., 2016. Techno-economic assessment of central sorting at material recovery facilities – the case of lightweight packaging waste. Journal of Cleaner Production 112, Part 5, 4387–4397. https://doi.org/10.1016/j.jclepro.2015.09.011

Cimpoi, M., Maji, S., Kokkinos, I., Vedaldi, A., 2016. Deep Filter Banks for Texture Recognition, Description, and Segmentation. International Journal of Computer Vision 1–30. https://doi.org/10.1007/s11263-015-0872-3

Hafemann, L.G., Oliveira, L.S., Cavalin, P., 2014. Forest species recognition using deep convolutional neural networks, in: 2014 22nd International Conference on Pattern Recognition. IEEE, pp. 1103–1107.

Haindl, M., Filip, J., 2013. Visual Texture: Accurate Material Appearance Measurement, Representation and Modeling. Springer Science & Business Media.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105.

Peddireddy, S., Longhurst, P.J., Wagland, S.T., 2015. Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste Management 40, 9–13. https://doi.org/10.1016/j.wasman.2015.03.015

Pichtel, J., 2005. Waste Management Practices: Municipal, Hazardous, and Industrial. CRC Press.

Rahman, M.O., Hussain, A., Basri, H., 2014. A critical review on waste paper sorting techniques. International Journal of Environmental Science and Technology 11, 551–564. https://doi.org/10.1007/s13762-013-0222-3

Taylor, L., Nitschke, G., 2017. Improving Deep Learning using Generic Data Augmentation. arXiv:1708.06020 [cs, stat].

Tchobanoglous, G., Kreith, F., 2002. Handbook of solid waste management. New York : McGraw-Hill, 2002.

Wagland, S.T., Dudley, R., Naftaly, M., Longhurst, P.J., 2013. Determination of renewable energy yield from mixed waste material from the use of novel image analysis methods. Waste Management 33, 2449–2456. https://doi.org/10.1016/j.wasman.2013.06.021

Wagland, S.T., Veltre, F., Longhurst, P.J., 2012. Development of an image-based analysis method to determine the physical composition of a mixed waste material. Waste Management 32, 245–248. https://doi.org/10.1016/j.wasman.2011.09.019

WISH Forum, 2017. Reducing fire risk at waste management sites [WWW Document]. URL https://wishforum.org.uk/?page_id=33 (accessed 12.9.19).

CHAPTER 8

CONCLUSIONS

8 Conclusions

This research study identified the data acquisition technology and used known principles to develop data processing strategies and methodologies for the classification and composition analysis of solid waste materials to answer the following research question:

Given the current state of available technologies, is it possible to develop and define a methodology for the classification of images of waste to analyse the composition of solid waste automatically and accurately in material recovery facilities?

The answer is formulated through the completion of the following research objectives:

1) To identify the state of the art in classification of materials using advanced sensor technologies and assess the potential impact and limitations of using visible sensor cameras in the context of waste management

The comprehensive literature review in Chapter 2 identified the different material characteristics, or quality indicators, that have a direct impact on the design and operation of the different processing units in a waste material recovery facility. A comparison of different advanced sensor technologies indicated that optical sensors could provide all these quality indicators. Furthermore, visible sensor cameras were identified as the solution offering the highest flexibility as they were demonstrated to be able to extract, or indirectly estimate all the quality indicators except for moisture.

Traditionally, waste material recovery facilities have used manual sorting, combined with mechanical and biological process to sort and separate recyclable materials (Tchobanoglous and Kreith, 2002). In recent years, advanced sorters were implemented using near infrared technology to classify items, for example for the separation of plastics by polymers (Huth-Fehre et al., 1995; WRAP, 2011). The main use case and advantage of these NIR sensor-based solutions compared to visible sensors is that can identify for example different plastic polymers. However, the literature review revealed that near infrared-based

solutions have important limitations, that could be overcome with the use of a solution based on visible wavelength sensor cameras:

- Classify different types of waste using the same material, such as PET bottles vs PET trays. A visible camera could easily determine whether the item is a bottle or not.
- Bottles with a sleeve made of a different material than the main bottle body. The NIR sensor would detect the label material and assign that composition to the entire bottle. A visible camera could identify the item type, such as a clear PET bottle, whether it has a sleeve on it or not, and correctly account for its composition
- Multi-layered items, for example drink cartons or food pouches, as the NIR can only see the top layer and assumes the entire item is made from the same material. A visible light sensor could identity the carton and estimate its correct composition.
- Distinguish waste items previously used for food vs non-food. As the industry moves towards circularity, regulations allow packaging producers to recycle food packaging into food packaging, with strict controls of contamination in the form of non-food packaging, as its previous contents could he a health hazard. This distinction could not be made by a NIR sensor, but a visible sensor camera could identify the item type, using recognisable features.
- Detect contaminated items that would not be acceptable for recycling, for example non-empty bottles, or greasy papers. These are visible characteristics but could not be possible using a NIR sensor.

2) To design, develop and test an image acquisition method to capture images of waste samples with representative variability to create a high-quality dataset for training a deep learning model for the classification of images of waste

The image capture methodology developed in Chapter 3 and used on Chapters 4 to 6 is based on bidirectional texture function techniques used for texture synthesis and representation. The test results from Chapter 3 indicate that a

combination of five cameras and five light sources produce enough variations in angle of views and illuminations to create a training dataset for training deep networks with state-of-the-art classification performance. The methodology was further extended in Chapter 6 to use 16 different exposure levels to generalise the classification in the context of a waste recovery facility.

The creation of high-quality datasets for training and testing is key to deep learning techniques. However, waste presents challenges as the main physical properties of items can change significantly, which makes it difficult to capture a representative dataset. The proposed methodology uses high resolution cameras, to capture enough details so that materials can be recognised based on surface features. Even if the items are crushed, broken up or deformed, the surface details will remain and would be captured by the developed methodology.

Furthermore, as the develop method is in essence a controlled data capture method, it simplifies and reduces the time and potential costs of labelling the images. For example, if images of waste would be photographed on conveyor belts in MRF, each picture would have to be manually analysed to identify and label each item present in the pictures. This contrasts with the developed methodology, where the waste material in the images can be known a priori, enabling automatic labelling. The pipeline implemented requires only manual intervention in the preparation of the physical samples in front of the cameras. The rest of the steps could be fully automated, from image capture using multiple illuminations, to image enhancement to data augmentation and annotation. This could accelerate the adoption of such a technology to be able to scale this solution for the waste industry.

3) To investigate the importance of coloured content in the images for the recognition of waste and critically evaluate the recognition performance when using greyscale images to create a solution to typical operational challenges

Previous investigations into the visual classification of materials depended on the presence of many features in the images, such as context information and colour

content (Bell et al., 2015; Sharan et al., 2013). In the waste industry, the waste items and materials are mixed and shipped and mechanically processed along the supply chain until they are either recovered or disposed of. Their general appearance changes throughout that process, for example because of colour transfer between items, or the colours are degraded due to their exposure to the environment, or the producer changes the packaging. In any case, the colour of the waste items is not constant and cannot be taken for granted.

The results from Chapter 4 proved that the identification of materials based on surface features is possible without colours. In contrast to previous analysis (Sharan et al., 2013), where the lack of colour information resulted in significantly worst classification performances, we have demonstrated that classification of greyscale patches achieves similar or better performance than using colour. This indicates than other surface features, like texture and reflectance, are enough to compensate for the lack of colour info, so that the model can learn and classify the different classes of materials. Furthermore, using grayscale patches makes the solution more robust to changes to colour changes, thus becoming more relevant for the waste industry.

4) *To understand the effects of different spatial scales and apply scale jittering as a data augmentation technique to generalise the solution to different operational conditions in waste management*

This work uses scale jittering and deep learning techniques to create a scale-invariant classifier. Previous investigations use this same technique with images of random resolutions over a given range (Simonyan and Zisserman, 2014). The technique developed in Chapter 5 used a series of spatial resolutions calculated from different scenarios and applications in the waste industry. The model achieved state of the art performance, classifying materials with an accuracy of 88.5%.

The results obtained in Chapter 5 demonstrate it is possible to train and deploy a single robust solution that can handle every industrial situation, from analysing waste on a tipping floor (low spatial resolution) to analysing waste items on a slow conveyor belt at the outputs (high spatial resolution). The model not only can classify the materials on a range of resolutions, but it can do so better than when trained with a single resolution matching the resolution of the tested patches.

5) *To formulate, test and define an end-to-end pipeline of image acquisition, image processing and data augmentation techniques to increase the diversity of images and improve the generalisation of deep learning models for the classification of waste in the field*

Training a deep learning model requires good quality training data. The classification of waste based on images is a daunting task, due to the infinite ways that a single material can be presented to a camera. Any different in illumination, pose, rotations, decolouration, or shape deformation would have to be considered and included in a representative training dataset.

Despite these challenge, deep learning has been applied in recent and concurrent years for the recognition of waste (Adedeji and Wang, 2019; Aral et al., 2018; Bircanoğlu et al., 2018; Dewulf, 2017; Ozkaya and Seyfi, 2019; Ruiz et al., 2019; Sakr et al., 2016; Yang and Thung, 2016). These different models and solutions recognise waste items, as an extension of object recognition in images. While the performance they achieved is high, with accuracies from 83% to 98%, their solutions use a limited number of classes defined by objects and items, and could not work in environments where the shape of those objects and items is completely changed, for example after a shredder process for MRF preparing waste derived fuels.

The solution developed in this thesis considered the industrial requirements and needs, which includes the analysis of shredded materials to produce waste derived fuels such as RDF and SRF. The proposed methodology analyses small patches, to avoid the reliance on context and on defined edges or shape

information about the items. It uses grayscale version of the images, which are captured with enough resolution to enable the classification based on texture and or other surface details, while being robust to changes to the waste colour. It adds multiple resolutions to the training dataset to create a scale-invariant classifier. The obtained results show a state-of-the-art performance with classification accuracy of 88.5% and was proved to be robust to changes to colour, lighting, and resolution, which are key to implement it in the waste industry.

6) *To synthesise the findings from objectives 1 to 5 into a methodology to automatically estimate the composition of mixed waste using image-based classification, and to appraise its practical uses in the waste industry*

The work in this thesis presents an end-to-end pipeline, developed on Chapters 3 through 5, to generate a high-quality dataset of small patches for training a model that can recognise waste based exclusively on material features, without context information and without relying on the general shape of the items. Mixed waste could thus be classified, to estimate its waste composition with errors of 3.1% points when analysing waste items, or 3.6% points when analysing shredded waste. The achieve performance is better the obtained in previous research using visual inspection (Wagland et al., 2012), while enabling automatic classification of images of waste.

The solution implemented in Chapter 6 was demonstrated to be robust to expected conditions and situation in the waste industry supply chain. This means it can be deployed to automatically analyse the composition of waste in any location in a MRF, to provide data that could be used to improve material recovery processes (Vrancken et al., 2017) and supporting the implementation of future legislation depending on high volumes of waste composition data (DEFRA, 2021).

The innovative work presented in the thesis proved that it is possible to use currently available technologies to develop a methodology for the classification

of images of waste to analyse the composition of solid waste automatically and accurately in material recovery facilities. The results and observations provide new data and understanding of the processes and challenges of waste material classification. It makes contributions to knowledge in the fields of image capture and processing, deep learning, and material recognition (Chapter 7).

The developed method is robust, and more flexible and affordable than current automated solutions. Whilst current mechanical and biological treatment processes are based on physical properties of the materials, real-time composition information could be incorporated in new designs as part of a feedback control system. The developed method provides real-time information that could be used outside a material recovery facility to increase transparency and new waste management applications in the industry. Stakeholders in the waste industry could learn from the new information to create new policies and regulations to further reduce the environmental impact of waste.

8.1 References

Adedeji, O., Wang, Z., 2019. Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network. Procedia Manufacturing, The 2nd International Conference on Sustainable Materials Processing and Manufacturing, SMPM 2019, 8-10 March 2019, Sun City, South Africa 35, 607–612. https://doi.org/10.1016/j.promfg.2019.05.086

Aral, R.A., Keskin, Ş.R., Kaya, M., Hacıömeroğlu, M., 2018. Classification of TrashNet Dataset Based on Deep Learning Models, in: 2018 IEEE International Conference on Big Data (Big Data). Presented at the 2018 IEEE International Conference on Big Data (Big Data), pp. 2058–2062. https://doi.org/10.1109/BigData.2018.8622212

Bell, S., Upchurch, P., Snavely, N., Bala, K., 2015. Material recognition in the wild with the materials in context database, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3479–3487. https://doi.org/10.1109/CVPR.2015.7298970

Bircanoğlu, C., Atay, M., Beşer, F., Genç, Ö., Kızrak, M.A., 2018. RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks, in: 2018 Innovations in Intelligent Systems and Applications (INISTA). Presented at the 2018 Innovations in Intelligent Systems and Applications (INISTA), pp. 1–7. https://doi.org/10.1109/INISTA.2018.8466276

DEFRA, 2021. Extended Producer Responsibility for Packaging [WWW Document]. URL https://consult.defra.gov.uk/extended-producer-responsibility/extended-producer-responsibility-for-packaging/ (accessed 11.8.21).

Dewulf, V., 2017. Application of machine learning to waste management: identification and classification of recyclables. https://doi.org/10.13140/RG.2.2.26805.50406

Ozkaya, U., Seyfi, L., 2019. Fine-Tuning Models Comparisons on Garbage Classification for Recyclability. ArXiv abs/1908.04393.

Ruiz, V., Sánchez, Á., Vélez, J.F., Raducanu, B., 2019. Automatic Image-Based Waste Classification, in: Ferrández Vicente, J.M., Álvarez-Sánchez, J.R., de la Paz López, F., Toledo Moreo, J., Adeli, H. (Eds.), From Bioinspired Systems and Biomedical Applications to Machine Learning, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 422–431. https://doi.org/10.1007/978-3-030-19651-6_41

Sakr, G.E., Mokbel, M., Darwich, A., Khneisser, M.N., Hadi, A., 2016. Comparing deep learning and support vector machines for autonomous waste sorting, in: Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International. IEEE, pp. 207–212.

Sharan, L., Liu, C., Rosenholtz, R., Adelson, E.H., 2013. Recognizing Materials Using Perceptually Inspired Features. Int J Comput Vis 103, 348–371. https://doi.org/10.1007/s11263-013-0609-0

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs].

Tchobanoglous, G., Kreith, F., 2002. Handbook of solid waste management. New York : McGraw-Hill, 2002.

Vrancken, C., Longhurst, P., Wagland, S., 2017. Critical review of real-time methods for solid waste characterisation: Informing material recovery and fuel production. Waste Management 61, 40–57. https://doi.org/10.1016/j.wasman.2017.01.019

Wagland, S.T., Veltre, F., Longhurst, P.J., 2012. Development of an image-based analysis method to determine the physical composition of a mixed waste material. Waste Management 32, 245–248. https://doi.org/10.1016/j.wasman.2011.09.019

Yang, M., Thung, G., 2016. Classification of Trash for Recyclability Status (Report). Stanford University, http://cs229.stanford.edu/proj2016/report/ThungYang-ClassificationOfTrashForRecyclabilityStatus-report.pdf (accessed 11.08.19).

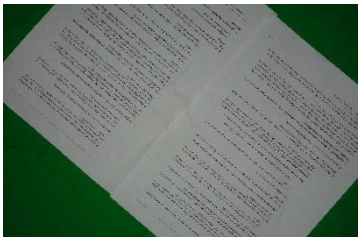
APPENDICES

Appendix A

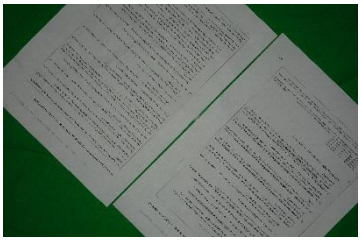
This section includes data to supplement the information provided in some of the chapters.

A.1 . Supporting information for Chapter 3

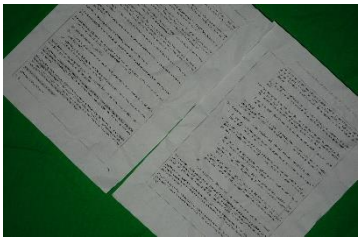
This section shows the images of the material samples used in Chapter 3, captured with camera C5 and illumination L5, from which the training and testing patches were extracted.



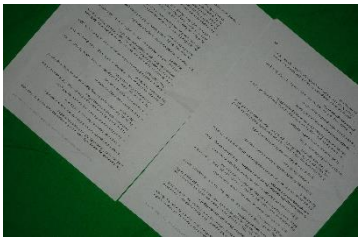
a4printed_1.jpg



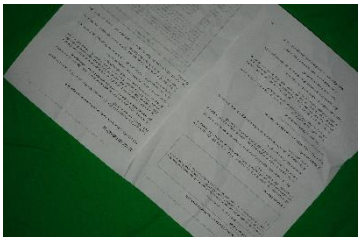
a4printed_10.jpg



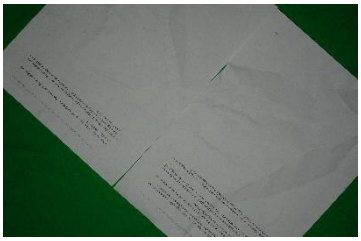
a4printed_11.jpg



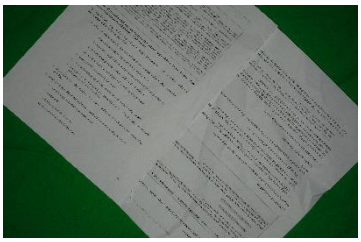
a4printed_2.jpg



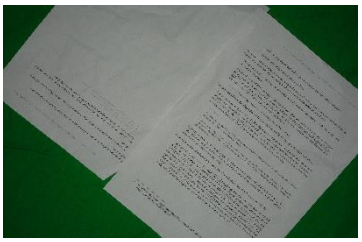
a4printed_3.jpg



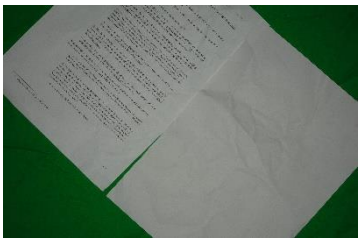
a4printed_4.jpg



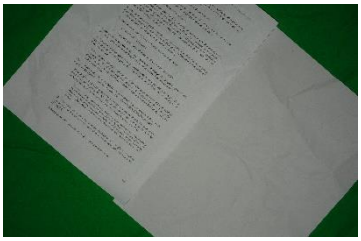
a4printed_5.jpg



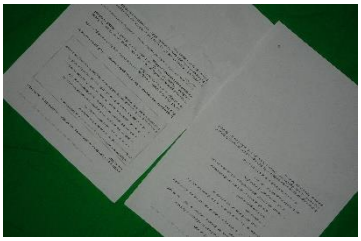
a4printed_6.jpg



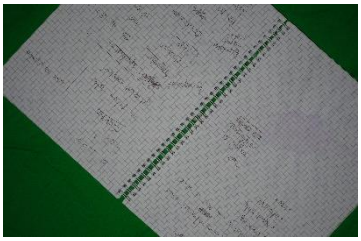
a4printed_7.jpg



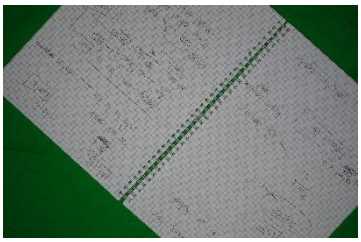
a4printed_8.jpg



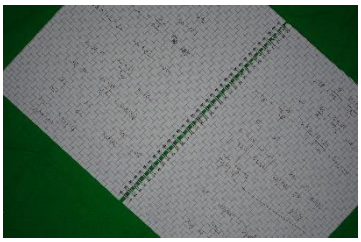
a4printed_9.jpg



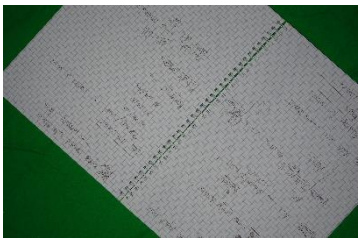
paper_notepad1.jpg



paper_notepad2.jpg



paper_notepad3.jpg



paper_notepad4.jpg

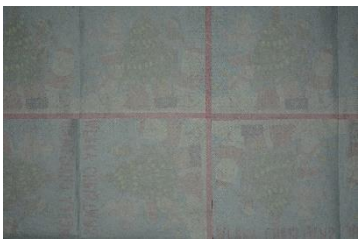
paper_shredded_2.jpg

paper_shredded_3.jpg

paper_brownpacking.jpg

paper_brownpacking2.jpg

paper_brownpacking5.jpg



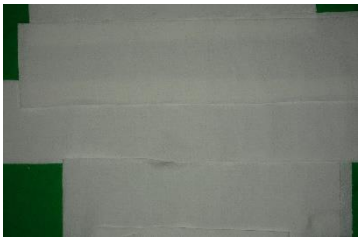
paper_tissue_party_napkin2_back.jpg

paper_tissue_party_napkin2_front.jpg

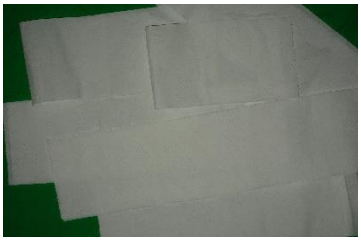
paper_tissue_party_napkin3_back.jpg

tissue_red.jpg

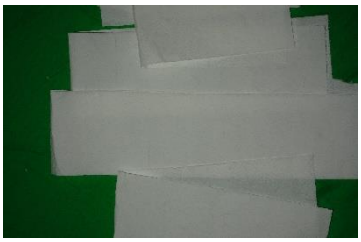
tissue_red_2.jpg



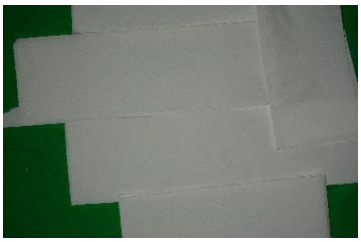
paper_tissue_hygienic_7_back.jpg



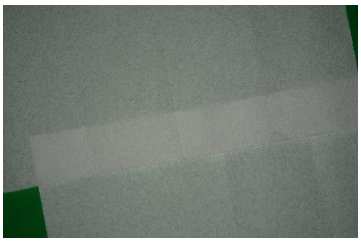
paper_tissue_hygienic_7_front.jpg



paper_tissue_hygienic_8_back.jpg



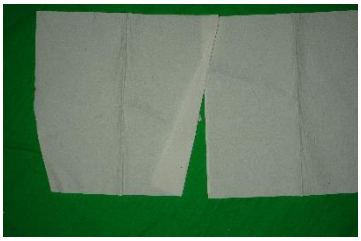
paper_tissue_hygienic_8_front.jpg



paper_tissue_kitchen_towel_2020.jpg

tissue_hygienic1.jpg

tissue_hygienic2.jpg



tissue_whitetowel_1.jpg

tissue_whitetowel_2.jpg

magazinelifespan_page10.jpg

magazinelifespan_page12.jpg

magazinelifespan_page14.jpg

magazinelifespan_page18.jpg

magazinelifespan_page20.jpg

magazinelifespan_page6.jpg

magazinelifespan_page8.jpg

magazinevitamin_page10.jpg

magazinevitamin_page12.jpg

magazinevitamin_page14.jpg

registermag_p10.jpg

registermag_p12.jpg

registermag_p14.jpg

registermag_p16.jpg

registermag_p18.jpg

registermag_p20.jpg

registermag_p4.jpg

registermag_p6.jpg

registermag_p8.jpg

whichmag_p22.jpg

whichmag_p24.jpg

whichmag_p26.jpg

whichmag_p28.jpg

whichmag_p30.jpg

whichmag_p32.jpg

whichmag_p34.jpg

whichmag_p36.jpg

whichmag_p38.jpg

whichmag_p40.jpg

whichmag_p42.jpg

newspaper_1.jpg

newspaper_2.jpg

newspaper_3.jpg

newspaper_4.jpg

newspaper_5.jpg

newspaper_6.jpg

newspaper_7.jpg

newspaper_8.jpg

Figure A-1 - Images of paper samples used for training, as captured with camera C5 and illumination L5

amazonboxmedium_back.jpg

amazonboxmedium_front.jpg

amazonenvelope_smallbox_back.jpg

amazonenvelope_smallbox_front.jpg

amazonenvelopelarge_back.jpg

amazonenvelopelarge_front.jpg

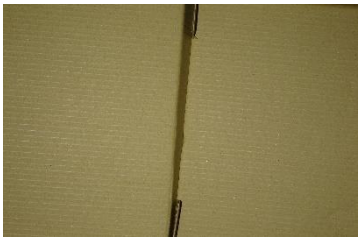
brownbox1_back.jpg

brownbox1_front.jpg

browncard_back.jpg

browncard_front.jpg

cardboard_brown_arteck_front.jpg



cardboard_brown_swimtime_back.jpg

cardboard_brown_swimtime_front.jpg

tassimoboard_back.jpg

tassimoboard_front.jpg

almondmilk_lightbluetissuebox_back.jpg

almondmilk_lightbluetissuebox_front.jpg

braunrefills_nescafegold_back.jpg

braunrefills_nescafegold_front.jpg

cardboard_alpro_chocolate_twinings_camomile_back.jpg

cardboard_alpro_chocolate_twinings_camomile_front.jpg

cardboard_cheerios_clipper_earlgrey_back.jpg

cardboard_cheerios_clipper_earlgrey_front.jpg

cardboard_fishgoujon_quornfillets_back.jpg

cardboard_fishgoujon_quornfillets_front.jpg

cardboard_jordans_bosh_back.jpg

cardboard_jordans_bosh_front.jpg

cardboard_like_nuggets_strawberry_yogurt_back.jpg

cardboard_like_nuggets_strawberry_yogurt_front.jpg

cardboard_tesco_cones_peppermint_tea_back.jpg

cardboard_tesco_cones_peppermint_tea_front.jpg

cereals_1_2_back.jpg

cereals_1_2_front.jpg

farmanimals_clippersleep_back.jpg

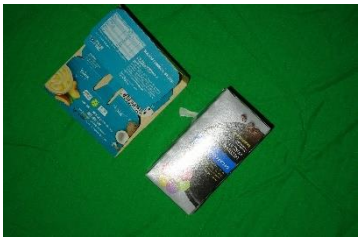
farmanimals_clippersleep_front.jpg

happymonkeysorangemango_tassimorning_back.jpg

happymonkeysorangemango_tassimorning_front.jpg

kiev_chocpie_back.jpg

kiev_chocpie_front.jpg



lordecaf_nutridelipeach_back.jpg

lordecaf_nutridelipeach_front.jpg

montezuma_tescoveggiefingers_back.jpg

montezuma_tescoveggiefingers_front.jpg

oatcheescake_kellogsoriginal_back.jpg

oatcheescake_kellogsoriginal_front.jpg

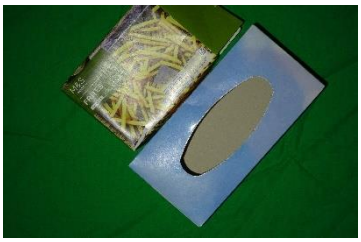
passata_lentilcrispbread_back.jpg

passata_lentilcrispbread_front.jpg

pizza_kirstys_back.jpg

pizza_kirstys_front.jpg

sainsburytissue_msfrites_back.jpg



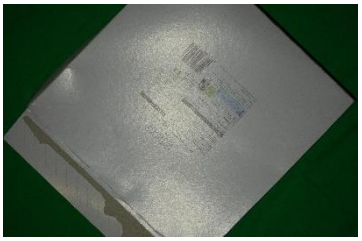
sainsburytissue_msfrites_front.jpg

shortcrust_bakingpaperbox_back.jpg

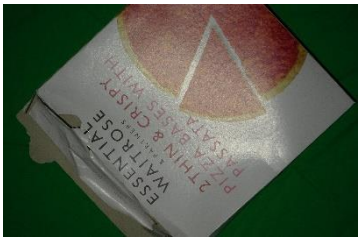
shortcrust_bakingpaperbox_front.jpg

tescomash_tastelinguine_back.jpg

tescomash_tastelinguine_front.jpg



waitrosepizzabase_back.jpg



waitrosepizzabase_front.jpg

Figure A-2 - Images of cardboard samples used for training, as captured with camera C5 and illumination L5

validation_magazinevitamin_page20.jpg

validation_magazinevitamin_page22.jpg

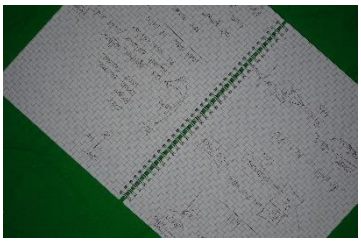
validation_paper_brown_filler.jpg

validation_paper_crumpled.jpg

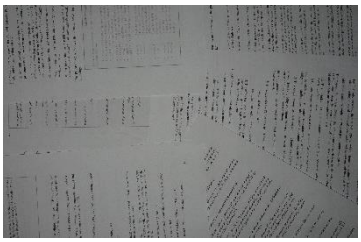
validation_paper_magazine1.jpg

validation_paper_newspaper1.jpg

validation_paper_newspaper2.jpg



validation_paper_notepad5.jpg

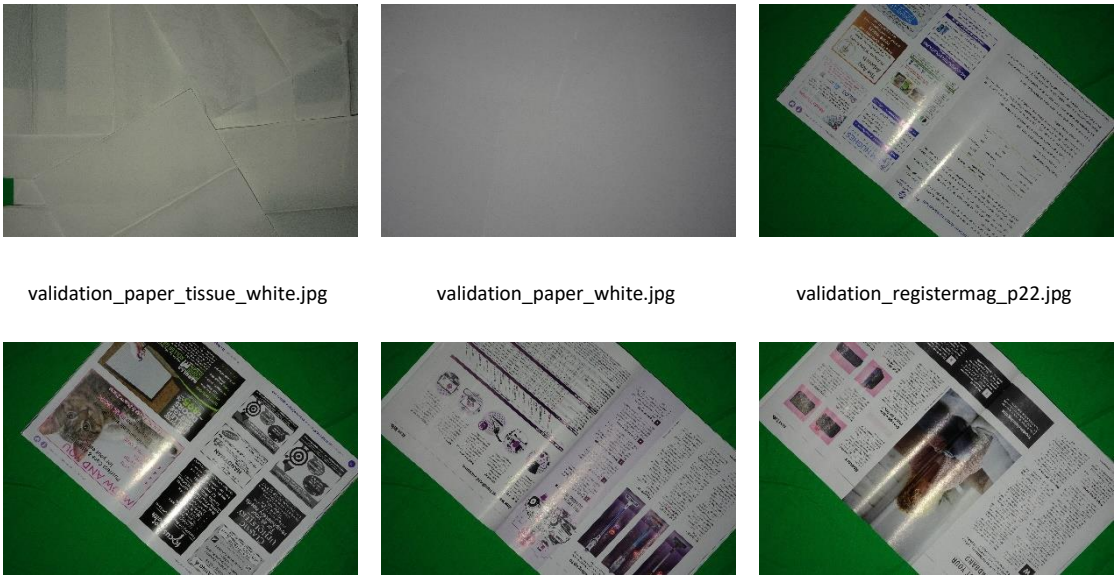


validation_paper_printed.jpg

validation_paper_shredded.jpg

validation_paper_tissue_hygienic.jpg

validation_paper_tissue_party_napkin3_front.jpg



validation_paper_tissue_white.jpg | validation_paper_white.jpg | validation_registermag_p22.jpg

validation_registermag_p24.jpg | validation_whichmag_p44.jpg | validation_whichmag_p46.jpg

Figure A-3 - Images of paper samples used for testing, as captured with camera C5 and illumination L5

validation_amisacrispbread_lmon_back.jpg

validation_amisacrispbread_lmon_front.jpg

validation_bettychoc_twinningsleep_eatnatmaple_back.jpg

validation_bettychoc_twinningsleep_eatnatmaple_front.jpg

validation_cardboard_brown_arteck_back.jpg

validation_cardboard_brown_workpro_front.jpg

validation_cardboard_corrugated1.jpg

validation_cardboard_corrugated2.jpg

validation_happymonkeystrawberry_tassimoxl_back.jpg

validation_happymonkeystrawberry_tassimoxl_front.jpg

Figure A-4 - Images of cardboard samples used for training, as captured with camera C5 and illumination L5

A.2 . Supporting information for Chapter 4

Table A-1 - Confusion matrix when using the Coloured version of the patches. Bold indicates correct predictions. Colour shades indicate the proportion of the number of a predicted class relative to the number of samples for that row.

	Paper white	Paper white printed	Paper white (crumpled, shredded)	Paper brown filler	Paper tissue red	Paper tissue (white, hygienic)	Paper magazine	Paper newspaper	Cardboard corrugated	Cardboard household
Paper white	**1,972**	83	7	0	0	0	4,029	2	0	380
Paper white printed	246	**3,000**	8	0	0	17	2,605	87	0	471
Paper white (crumpled, shredded)	98	900	**6,982**	6	0	46	839	984	3	1,458
Paper brown filler	0	0	7	**4,147**	0	1,606	3	3,655	401	2,547
Paper tissue red	0	0	0	0	**5,992**	0	0	1	5	389
Paper tissue (white, hygienic)	75	19	228	130	0	**11,528**	194	287	3	111
Paper magazine	41	292	54	0	0	47	**5,514**	1,132	0	5,247
Paper newspaper	17	593	143	33	0	499	1,550	**18,513**	136	4,183
Cardboard corrugated	0	0	1	52	0	577	133	492	**17,632**	461
Cardboard household	10	30	39	12	29	38	465	927	18	**28,643**

Table A-2 - Confusion matrix when using the greyscale version of the patches. Bold indicates correct predictions. Colour shades indicate the proportion of the number of a predicted class relative to the number of samples for that row.

	Paper white	Paper white printed	Paper white (crumpled, shredded)	Paper brown filler	Paper tissue red	Paper tissue (white, hygienic)	Paper magazine	Paper newspaper	Cardboard corrugated	Cardboard household
Paper white	**4,864**	350	50	0	1	78	570	0	1	559
Paper white printed	484	**4,161**	5	0	1	23	1,038	463	22	236
Paper white (crumpled, shredded)	106	624	**7,158**	36	5	306	467	1,342	11	1,261
Paper brown filler	0	0	18	**10,375**	25	24	0	577	1,021	329
Paper tissue red	0	0	20	60	**5,266**	40	10	338	417	239
Paper tissue (white, hygienic)	123	80	508	451	67	**9,601**	98	522	705	419
Paper magazine	94	567	158	10	4	22	**5,349**	1,315	20	4,786
Paper newspaper	17	592	488	243	154	272	407	**19,093**	390	4,010
Cardboard corrugated	0	1	6	60	232	31	4	279	**18,436**	297
Cardboard household	67	83	196	176	71	115	438	2,306	322	**26,438**

Table A-3 - Confusion matrix when using a model trained with a balanced training dataset, with 5,000 patches per class, to classify 10 classes. Bold indicates correct predictions. Colour shades indicate the proportion of the number of a predicted class relative to the number of samples for that row

	Paper white	Paper white printed	Paper white (crumpled, shredded)	Paper brown filler	Paper tissue red	Paper tissue (white, hygienic)	Paper magazine	Paper newspaper	Cardboard corrugated	Cardboard household
Paper white	**6461**	4	1	0	0	3	2	0	0	1
Paper white printed	963	**4754**	1	0	1	11	661	20	4	18
Paper white (crumpled, shredded)	537	1067	**6779**	91	21	703	730	688	12	689
Paper brown filler	0	0	19	**11215**	144	179	17	146	512	135
Paper tissue red	2	0	1	116	**5924**	69	8	48	188	35
Paper tissue (white, hygienic)	775	36	312	351	302	**10049**	108	172	348	120
Paper magazine	293	1871	286	46	13	32	**6145**	1079	21	2542
Paper newspaper	227	3613	848	605	392	907	1166	**13964**	312	3633
Cardboard corrugated	74	4	3	381	1628	714	104	127	**16201**	113
Cardboard household	568	731	791	908	395	885	2476	3882	561	**19016**