Multimedia Tools and Applications (2023) 82:29115–29133
https://doi.org/10.1007/s11042-023-14551-6

The influence of micro-expressions on deception detection

Suleyman Yildirim[1] · **Meshack Sandra Chimeumanu**[2] · **Zeeshan A. Rana**[3]

Received: 16 December 2020 / Revised: 24 August 2021 / Accepted: 31 January 2023 /
Published online: 16 March 2023
© The Author(s) 2023

Abstract

Facial micro-expressions are universal symbols of emotions that provide cohesion to interpersonal communication. At the same time, the changes in micro-expressions are considered to be the most important hints in the psychology of emotion. Furthermore, analysis and recognition of these micro-expressions have pervaded in various areas such as security and psychology. In security-related matters, micro-expressions are widely used to detect deception. In this research, a deep learning model that interprets the changes in the face into meaningful information has been trained using The Facial Expression Recognition 2013 dataset. Necessary data is also obtained through live stream or video stream by detecting via computer vision and evaluating with the trained model. Finally, the data obtained is transformed into graphic and interpreted to determine whether the people are trying to deceive or not. The deception classification accuracy of the custom trained model is 74.17% and the detection of the face with high precision using the computer vision methods increased the accuracy of the obtained data and provided it to be interpreted correctly. In this respect, the study differs from other studies using the same dataset. In addition, it is aimed to facilitate the deception detection which is performed in a complex and expensive way, by making it simple and understandable.

Keywords Micro-expressions · Deception detection · Machine learning

✉ Suleyman Yildirim
suleyman.yildirim@cranfield.ac.uk

Meshack Sandra Chimeumanu
meshacksandra@gmail.com

Zeeshan A. Rana
zeeshan.rana@cranfield.ac.uk

[1] Digital Aviation Research and Technology Centre (DARTeC), Cranfield University, Cranfield, MK43 0AL, UK

[2] School of Aerospace, Transport and Manufacturing (SATM), Cranfield University, Cranfield, MK43 0AL, UK

[3] Centre for Aeronautics, Cranfield University, Cranfield, MK43 0AL, UK

1 Introduction

Facial micro-expressions, which play an important role in interpersonal communication, contain important information about people's emotions, thoughts and mental states. In 1872, Charles Darwin stated in his book "Expression of Emotions in Humans and Animals" [9] that some innate emotions in humans and animals emerge as micro-expressions, and that these expressions are perceived in the same sense throughout the world, and that their micro-expressions are universal. This was the basis of the study of micro-expressions and micro-expression analysis has become a common research topic for many behavioural scientists. Examination of micro-expressions has an important place in many fields such as verbal and nonverbal communication, psychological research, human-computer interaction, medicine, virtual reality, computer vision and intelligent systems. As the need arises for deception detection in many areas such as crime investigation, airport security, national security, interrogations and interviews [20, 25, 28, 41], detecting deception efficiently, effectively and reliably would be beneficial for society.

The polygraph has been used widely to detect deception however some studies [23, 48] showed it can be misled as it requires human ability to perform the tests therefore it can be biased [4]. As the polygraph method requires physiological responses of blood pressure, heartbeat, skin conductivity, muscle tremor and respiration while answering questions [50], it becomes infeasible in terms of large-scale applications. Given the difficulties and limitations associated with the use of the polygraph method, it is critical to detect deception without human intervention. Therefore, researchers have been proposed learning-based solutions to address the deception detection problem using different modalities such as video [7], text [19] and speech [30, 37].

In this study, it is aimed to detect deception by analysing video modalities that are collected from human faces such as facial expressions [49]. To test the custom trained deception detection model's performance, several experiments were conducted. According to answers which have been given by participants, the deception detection model has been tested and has achieved 74.17% correct classification accuracy of deception.

2 Related work

In literature, micro-expressions have been used to determine the mental status of patients with schizophrenia [2, 22] to provide security against thieves and malicious people [1, 5] and to detect pain for newborn babies [34, 35]. Some studies use micro-expressions to detect drivers' fatigue and sleepiness [27], to predict safe driving behaviour [31], and to create automatic music playlists [13, 26].

In 1872, Darwin claimed that there were innate feelings of birth, and he claimed the universality and continuity of micro-expressions [51]. In 1971, Ekman and Friesen [52] stated that together with their unique micro-expressions, each of them had six basic emotions with distinctive content. In 1978, to determine micro-expressions using facial anatomy-based action units (AUs), they developed FACS (Facial Actions Coding Systems) [17] and defined 46 different action units for each independent muscle activity of the face. FACS is an objective method based [14] on the movement of facial muscles and different combinations of action units have been observed from 7,000.

In another study, Ekman et al. [12] proposed an automated facial analysis system that would classify face actions. In their work, they created a database of more than 1100 images with 24 different subjects using more than 150 action units or combinations and compared

three different approaches to classifying facial actions in images: a holistic spatial analysis based on basic components of greyscale images.

Yacoob and Davis [51] presented another approach to the representation and analysis of facial dynamics to recognize micro-expressions from image sequences. This approach is based on qualitative monitoring of major parts of the face and flows calculations at high-density slope points. In their study, optical flow calculation was used to determine the direction of movements caused by micro-expressions. A moderate symbolic representation was calculated in spatial and temporal motion areas. Since the approach they provide is based on the statistical characterisation of the movement patterns in the designated regions of the face, rather than to follow the individual point movements, they have developed a region viewer for the rectangles that include the facial features. Because each rectangle surrounds a related property, the movements of other facial features do not interfere with flow calculations within the region.

To be able to see the differences between Closed Question Method (CQM) and photographic questionnaire, Delmas et al. [10] conducted a study. Instead written questionnaire, to improve the classification of facial cues, they used photographic questionnaire method. Based on Facial Actions Coding Systems, 54 standardised photos used to determine whether a particular expression was more or less present during deception.

Porter et al. [42] investigated emotional leakage in both upper and lower faces as a result of emotional intensity. Disgusting, sad, frightening and happy images were showed to participants and their expressions were recorded as a response to images. Results were showed emotional leakage lasted longer in both upper and lower faces and supported the inhibition hypothesis.

Contrast to current 3D model trend, 2D appearance-based method was embraced by Su and Levine [45] which 2D invariant features were extracted from nine separate facial regions that are related to 3D characteristics. Extracted cues used to form a facial pattern vector to train a Random Forest into deceptive and truthful categories.

Micro-expressions are also being used by clinicians [29] to judge the location and severity of pain rather than patient's self-report. Even though, genuine pain-related micro-expressions in intensity and frequency are resembled, patients can be successful at dissimulating. To discriminate genuine and deceptive facial pain, a research is examined which 120 judges assigned to 4 different groups that are control, corrective feedback, deception training and deception training plus feedback. Videotaped micro-expressions including neutral expressions, faked pain, masked pain and genuine pain were showed to judges. The participants rated unpleasantness and pain intensity for each category. Participants' detection accuracy were increased by the help of immediate corrective feedback.

In spite of the fact that the use of facial expressions in detecting deception is a debatable subject, several studies have been presented that facial expressions can be used to discriminate between the truth and lies. Zhang et al. presented in their study, facial expressions can be used to detect the deception using deception indicators by identifying the differences between involuntary and stimulated expressions [53]. While Tian et al. proposed [47] a method that uses the orientation of face and the intensity of facial expressions to detect the deception, Owayjan et al. [38] and Thannoon et al. [46] utilised geometric based features which have been derived from facial expressions to find clues for detection. Only using mouth motion, wrinkle occurrence, eye blink and eyebrow motion indicators Su et al. designed [44] a system based on only facial expressions for deception detection. Crockett et al. presented an automated method [8] which extracts non-verbal information from facial behaviours using 36 channels hierarchical neural networks and reports the gender-based effect on deception clues. Perez et al. proposed an approach [39] by analysing text, video

modalities and achieved 75% classification accuracy using Random Forest, Decision Tree algorithms. Later on, their study [40] has been reached 82% classification accuracy by expanding it to Support Vector Machines. Moreover, 96% accuracy has been obtained on the same dataset as Perez et al. used by Gogate et al. [24] using Convolutional Neural Networks and combining the text, voice, video modalities. Chebbi et al. [7] trained their KNN classifier using real-life trial data by fusing all three modalities and it has achieved 97% classification accuracy.

As the introduction to deception detection and methods for detecting lies is presented in Section 1, the rest of the paper is organised as follows: the literature review is investigated in Section 2. Section 3 is the definition of micro-expressions, the dataset used for developing a machine learning model. Section 4 covers the structure of the machine learning model, a brief explanation of conducted experiment's procedure and questionnaire session, video clips and the micro-expression analysis. The results and the discussion can be found in Section 5. Finally, the conclusion is presented in Section 6.

3 Facial expressions

Facial expressions are an aspect of human behaviour that has been recognised as the most salient and influential aspect of human communication. The term 'facial expression' has been reserved by researchers to define certain reoccurring movements of the facial muscle that relay thoughts, emotions or behaviour [21]. A study done by [36] reveals that about 55% cues regarding emotions are conveyed through facial expressions. The face has been known to express various thoughts and emotions and there is an evidence to show that these emotions are universally expressed and interpreted in the same way across every culture [15]. This 'universal' recognition and interpretation across various cultures suggests that those emotions and their associative facial expressions might be determined genetically as opposed to being learnt socially. However this claim has been refuted by some controversies [43].

3.1 Micro and macro expressions

Facial expressions can be categorised into two types [16] namely macro-expressions and micro-expressions. It is the duration of the facial expression and not intensity that differentiates the macro from micro-expressions; expressions that last longer are macro expressions while brief expressions are called micro-expressions. Macro-expressions generally last between three quarter of a second to two seconds and such muscle movements are said to occur simultaneously at multiple parts of the individuals face. Macro-expressions are easily perceived by humans in conversations and do not accurately represent the emotional state of an individual as it could easily be faked. Hence, it is imperative to investigate a form of emotional state that occurs through muscular movement which is the micro-expression that can neither be repressed nor faked.

3.2 Nonverbal features

In this study, nonverbal facial features such as eye and mouth movements have been used. As the video clips originate from interviews which are based on interactions with participants who tell truth or lies, facial expressions have been annotated by this principle. Particularly, facial displays have been focused on during the annotation period since previous studies [11] stated their correlation with deceptive actions.

3.3 Dataset

The dataset used is the Facial Expression Recognition (FER-2013) dataset [6] which is an open-source dataset that was created by Pierre-Luc Carrier and Aaron Courville and was shared publicly for a competition at Kaggle prior to ICML2013 (International Conference on Machine Learning). The dataset consists of about 35,887 greyscale, 48X48 sized facial images with labelled seven emotions.

The labels of the emotions in the dataset are:

- 0: 4593 images – Angry
- 1: 5121 images – Fear
- 2: 547 images – Disgust
- 3: 8989 images – Happy
- 4: 4002 images – Surprise
- 5: 6077 images – Sad
- 6: 6198 images – Neutral

3.4 Deep learning model

The Convolutional Neural Network (CNN) was used because of its adequacy in properly classifying images. The model classifies the images into one of the expressions that are universally available and are present in the FER-2013 dataset-Angry, Fear, Disgust, Happy, Surprise, Sad, Neutral. The development of the model and the classification step consists of several phases as shown in the diagram above. Overall, there are four main phases in the classification process:

- Data Splitting : This process was already done by the FER-2013 dataset as the dataset comes split and already labelled.
- Model Training and Generation : This consists of the Convolutional layers, pooling layers, activation function, softmax and fully connected layers.
- Model Evaluation : The training model is evaluated using the validation data set from the FER-2013 dataset.
- Classification of real time images/saved videos : The model is then used for classification of real time images or pre-saved videos to detect for possible deception.

4 Methodology

4.1 Training custom model

Deep Learning is a sub-domain of Machine Learning which is inspired by the structure of the neural networks. Deep learning has several kinds of algorithms such as Convolutional Neural Networks, Recurrent Neural Networks and Artificial Neural Networks. But, particularly, Convolutional Neural Networks contributed to computer vision more than any other algorithm. Convolutional Neural Networks applications range from image segmentation, video and image recognition, image classification.

- Convolutional Layer : The convolutional layer is the first layer of the CNN structure and it is used to extract the features from the input image. By sliding the filter of size

MxM, the mathematical operation is performed on the input image and it produces a feature map that gives us general knowledge about the input such as edges and corners.

- Pooling Layer : Generally, the pooling layer comes after the convolutional layer and its primary goal is to reduce the size of the feature map which is created by the previous convolutional layer. To do so, the pooling layer decreases the connection between layers.
- Fully Connected Layer : The fully connected layer is usually placed before the dense layers to make the connections of neurons between two different layers and forms the last layers.
- Dense Layer : The dense layer mostly places at the end of the CNN architecture to receive inputs from all previous layers and perform the vector multiplications.
- Weights : Every neuron in the neural network creates an output value by applying the particular function to input values that are received from the previous layer. The vector of weights and the bias determine the function that is applied. As the weight and bias are representing particular features such as the shape of the input, they are being called filters.
- Dropout : To reduce the over-fitting in neural networks by preventing the adaptations on the input data, the dropout is widely used as a regularisation method. During the training process, dropout function randomly omits both visible and hidden units.
- Softmax : As the input values can be zero, negative and positive, the softmax function turns the vector of real values into another vector of real values which sums to 1. Basically, transforms the values into between 1 and 0 to enable the probabilities to be interpreted.

To capture the particular facial features efficiently which are used for training, an automated system needs to be employed as doing all this by hand is not feasible. To address this issue, nonverbal features have been defined and an automated feature extraction method is employed as shown in Fig. 1. What makes CNN an effective method is its ability to enable the application of different filters to input data.

While CNN extracts facial features from real-time stream, it stores into windows as follows: $\mathbb{R}^{h \times w \times f \times c}$ where h is the height, w is the width, f is the number of frames and c is the number of channels in the input image; $\mathbb{R}^{k_h \times k_w \times k_d \times k_m \times c}$ where k_h is the kernel's height, k_w is the kernel's width, k_d is the kernel's depth, k_m is the number of feature maps and c is the number of the channels in the input image; $\mathbb{R}^{p_h \times p_w \times p_d}$ where p_h is the pooling height, p_w is the pooling width, p_d is the pooling depth.

After extracting features from each frame and storing them into window vectors, the custom-designed CNN is fed by that information from each window vector. RMSprop optimiser algorithm has been used during the training process. RMSprop is one of the adaptive learning methods to fight dramatically reducing learning rates and it is quite fast as well as being so popular. RMSprop adapts the learning rate so the model can reach to saddle point faster as possible regardless of how small the gradients are.

4.2 Participants

Fifteen participants were enrolled for this experiment. Amongst the participants were seven females and eight males all within the age range of 25-33 years old. Majority of the participants were students while a little number were young professionals. The participants were not paid for the study and were encouraged to opt out when and if they decided to.

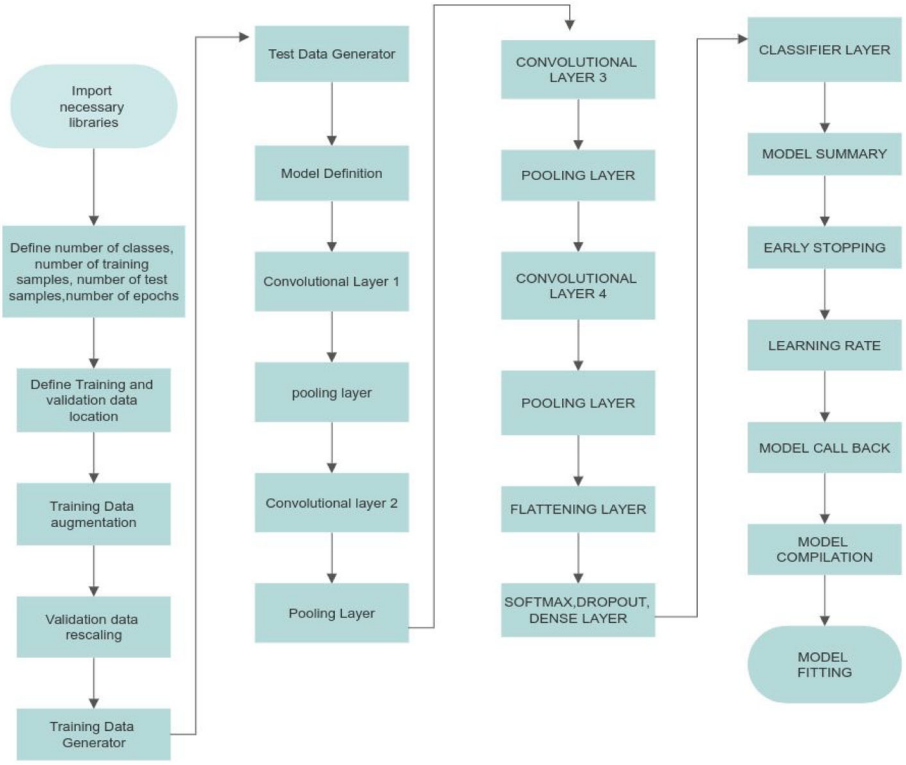


Fig. 1 Model development structure

4.3 Procedure

The participants were interviewed individually by an interviewer who they were not familiar with. Each participant had three sessions and were required to answer the same questions during the three sessions. The procedure for the sessions was as follows:

- Session 1 : The participants were asked to answer all questions sincerely.
- Session 2: The participants were required to answer some questions correctly and answer others falsely.
- Session 3: The participants were required to answer all questions falsely.

Each participant was given a basket with three folded papers inside the basket. The content of the papers were the sessions above and the instructions for each session. The participant was required to pick a paper from his/her basket before each session and was required to answer the questions based on the instructions on the paper. The participants were not required to inform the interviewer how they chose to answer during the sessions i.e. the interviewer was not allowed to see the content of the paper before and after the session. Therefore, the video clips were not tagged as the interviewer had no idea if the participant was being sincere, partly sincere or outright deceptive. The interview was conducted in a room with a camera facing the participant. The participants were asked to sit uprightly and to look directly into the camera. The participants were asked to answer Yes or No to

the questions asked and were encouraged to be as relaxed as possible. Each session lasted for a maximum of five minutes and the participants were asked fifteen questions in total. Five questions served as baseline questions while ten questions served as control questions. During each interview, the camera was turned on and recorded for the span of the interview.

4.4 Video clips

The video clips recorded from the sessions were later viewed by the interviewer and some observers and they tagged the video clips based on the perception of the sessions i.e Session 1, 2 or 3. The observers were in two categories. One category were people with close relationships with the participants and another category were people not known to the participants. The participants were also encouraged to tag the video clips based on how they responded to the sessions. The observers were required to tag the videos as an attempt to investigate and/or support the studies from various research from scientists that shows that it is difficult to tell when a person in a close relationship with oneself is being deceitful. Hence, the use of observers with close relationships and observers not known to the participants to compare the results and the reasons for their decision in tagging the videos.

4.5 Facial expression analysis

The video clips from the sessions were passed through the developed deception detection model. The micro-expressions appearing on the participants faces were compared to the micro-expressions that appeared in the participants faces when they were asked the baseline questions to ascertain possible deception as deviation from the baseline micro-expression could be an indication of possible deception. The interviewer then tags the video clips based on the analysis from the deception detection model.

The tagging by the observers, participants and indications from the model were compared to ascertain similarities. The observers who had closer relationships with the participants wrongly tagged the clips. The observers with distant relationships with the participants tagged some correctly and some wrongly. This supports the studies from research that shows that people with closer relationships can not tell when a person close to them is being deceitful because they want to believe the best of the individuals closer to them. The interviewers wrongly tagged the clips. The model did a good job in ascertaining and detecting deception and the tagging extracted the data from the model was correct. However, there were some inconsistencies in the interviews as some questions the participants admitted to be truthful were flagged as possible deceit by the model. When probed further, the participants were found to be telling the truth but there was an underlying problem because the participants admitted that some of the questions made them uncomfortable while they did not think some questions were a yes or no question as they would have preferred to explain further.

5 Results

5.1 Confusion matrix

Confusion matrix shows the summary of the prediction results, gives an insight into the errors made and the types of errors being made by the classifier (Fig. 2).

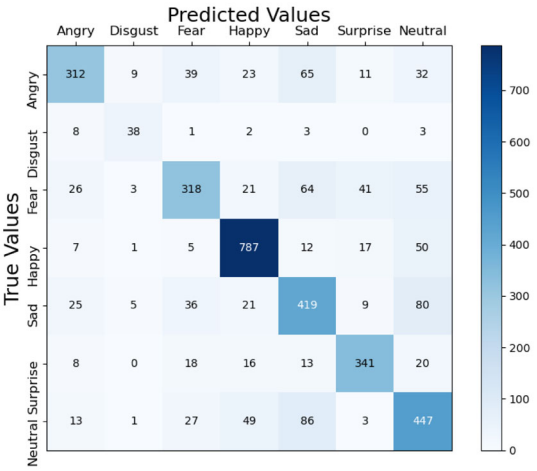


Fig. 2 Confusion matrix for deception detection model

5.2 Classification accuracy

The accuracy of the model shows the proportion of the total number of predictions that were correctly predicted by the model. The deception classification accuracy for this model is 74.17% which is considered as high performing in comparison to other models designed using CNN and the FER-2013 dataset. The model was also evaluated using other matrices as evaluation with only model accuracy is not a good choice with unbalanced classes like the FER-2013 dataset.

5.3 Recall

This is a model's ability to properly identify cases that are relevant within the dataset. From the confusion matrix above, the classes were evaluated using recall. The percentages of the recall is an indication of how well the classes are performing individually. The least performing class using recall is Neutral and this is because the model identifies a large percentage of the dataset as sad. However, Recall should work hand in hand with precision as precision tells the accuracy of the percentage identified. It correlates the effectiveness of the deception detection classifier to differentiate the truth and lies (Table 1).

Table 1 Recall for deception detection model

Micro-expressions	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
Recall	0.78	0.68	0.72	0.86	0.63	0.81	0.65

Table 2 Precision for deception detection model

Micro-expressions	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
Precision	0.64	0.69	0.60	0.90	0.71	0.82	0.71

5.4 Precision

Precision is a model's ability to identify the relevant data points. From the confusion matrix above, the classes were also evaluated using precision. A high precision and low recall is an indication that the model has a high number of false negatives however, the predictions that were positive are indeed positive as seen in the Angry and Fear classes. A high recall and a low precision is an indication that majority of the positive datasets have been correctly recognised on the other hand there may be false positives although it is arguable that the false positives may be negligible in this case as the percentage differences are minimal. It correlates the percentage of the correctly detected deceptions amongst the total number of deceptions (Table 2).

5.5 F1-score

The F1-Score is the harmonic mean of precision and recall values. It's a good way to ascertain the performance of the model. From the table below, the F1-Scores are all closer to 1 than 0 which indicates that it is a good performing model (Table 3).

5.6 Micro-expression detection

The model developed was tested for the detection of micro-expressions. However, the model can not only detect the percentage availability of the most dominant micro-expression but can also detect the percentage availability of other micro-expressions in a given clip or image (Figs. 3, 4, 5, 6, 7 and 8).

5.7 Model validation for deception detection

1. Session 1: In Session 1, the participants were required to be totally sincere. The dominant micro-expression detected from the result of one of the participant is neutral which happens to be the baseline micro-expression as seen in the graphical chart below. Neutral as the dominant micro-expression is an indication that the participant was being entirely sincere as neutral happens to be the baseline micro-expression for the participant. The data was not retrieved on a time basis as the participant was only asked a yes and no question hence didn't talk unless asked a question. When the participant was not interacting, his baseline micro-expression remained neutral.
2. Session 2: In Session 2, the participants were required to answer some questions sincerely and be deceptive in answering some questions. The dominant micro-expressions

Table 3 F1-score for deception detection model

Micro-expressions	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
F1-Score	0.7031	0.6850	0.6545	0.8795	0.6676	0.8150	0.6787

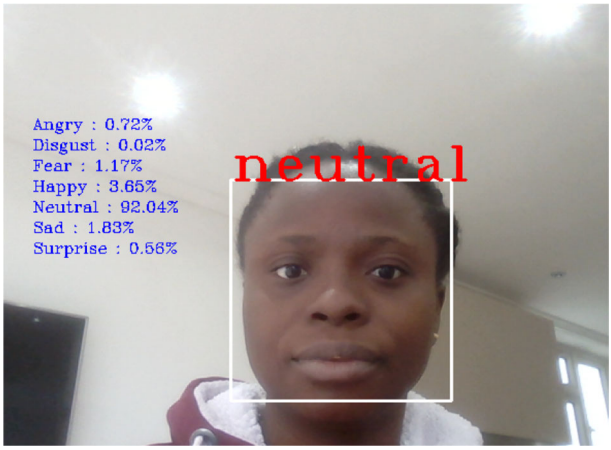


Fig. 3 Detection of the neutral micro-expression

found in the result of one of the participants as shown in the chart below are happy and neutral. Happy was the micro-expression detected when the participant was being deceptive while neutral is the baseline micro-expression. The participant is said to be deceptive when the micro-expression shifts from neutral to happy because the participant was further probed and admitted that he was being deceitful at the points where the micro-expression changed to happy. The data was not retrieved on a time basis as the participant was only asked a yes or no question hence didn't talk unless asked a question. When the participant was not interacting, his baseline micro-expression remained neutral.

3. Session 3: In Session 3, the participants were required to answer all questions falsely. The dominant micro-expression found in the result of one of the participants was sad. It is arguable that the participant was not proud of her attempt to be deceitful and couldn't suppress the emotion even when she tried her best to keep a neutral face.



Fig. 4 Detection of the fear micro-expression



Fig. 5 Detection of the sad micro-expression

From the results of the numerous experiments and validations carried out, it is evident that although the human eye cannot adequately detect deception, the model can accurately detect deception by the aid of micro-expressions. The involuntary nature of micro-expressions cannot be inhibited hence can be spotted using a working model. Participants attempted to conceal their emotions or micro-expressions when answering all questions even when they were being deceitful however, the model could detect the micro-expressions as an indication for deception which supports the argument that micro-expressions can neither be falsified nor suppressed.

The aim of asking the participants to judge and label the clips of other participants was to identify whether it was indeed possible for unskilled people to detect deception. However, low accuracy was expected and received. As there were no obvious emotional cues that can easily be identified by the untrained human eye neither can the participants accurately

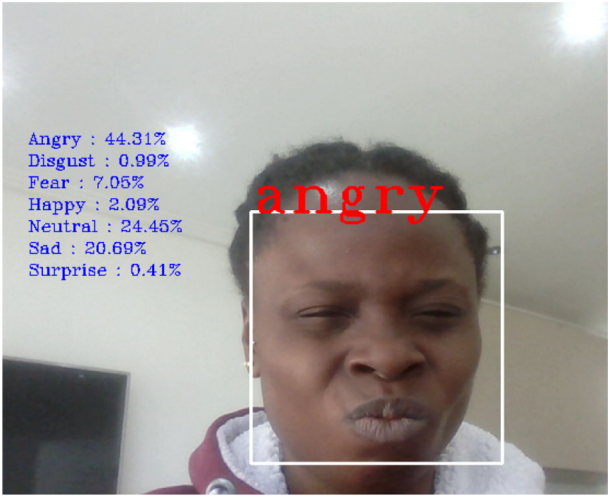


Fig. 6 Detection of the angry micro-expression

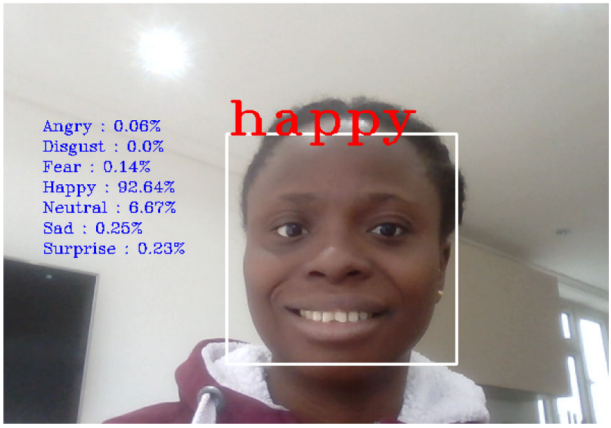


Fig. 7 Detection of the happy micro-expression

identify deception due to their emotional closeness with some of the other participants. This shows that the best and easiest way to detect deception is by using a working deception detection model. However, we must note that it is possible for the model to detect a change in micro-expression in the absence of deception as the participant may just be uncomfortable answering a particular question hence the need to probe further where necessary.

After the video clips were passed through the model and tagged by the model, the participants were also asked to tag their interview sessions to compare result and to check the accuracy of the model. As can be observed from the figures above, micro-expressions can vary according to given answers during different sessions. In Fig. 9, as the participants have given the answers correctly and comfortably, their confidence and comfort are reflected in the micro-expression data that have been obtained during Session 1. Emotion levels seem stable, sudden mood changes are not experienced. Custom trained deception detection model evaluates these changes in micro-expressions. Therefore, model did not indicate



Fig. 8 Detection of the surprised micro-expression

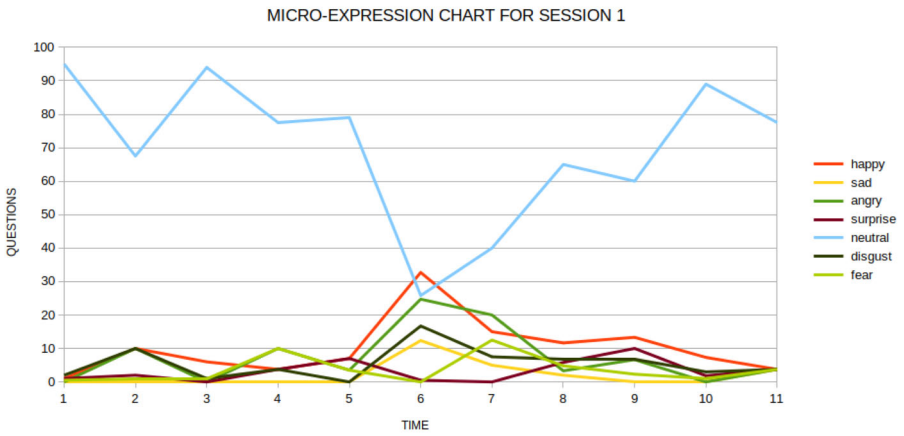


Fig. 9 Sample micro-expressions chart for session 1

deception during Session 1. However, it is seen in Fig. 10 how the discomfort affects sudden emotional changes due to the answers given by the participants. Sudden mood changes may be a indication of deception, custom trained deception detection model indicated the deception several times during Session 2.

5.8 Model validation for deception detection using bill clinton video

A famous video of Bill Clinton was used to ascertain the model's ability to accurately detect deception.

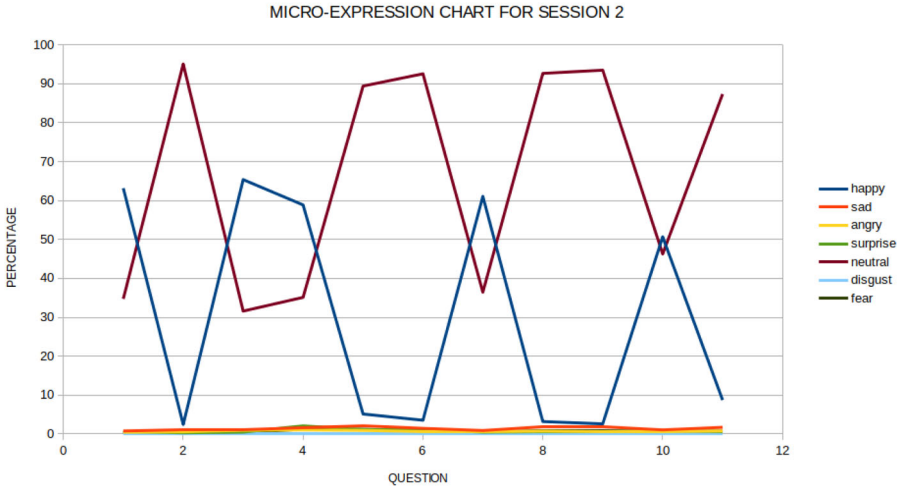


Fig. 10 Sample micro-expressions chart for session 2

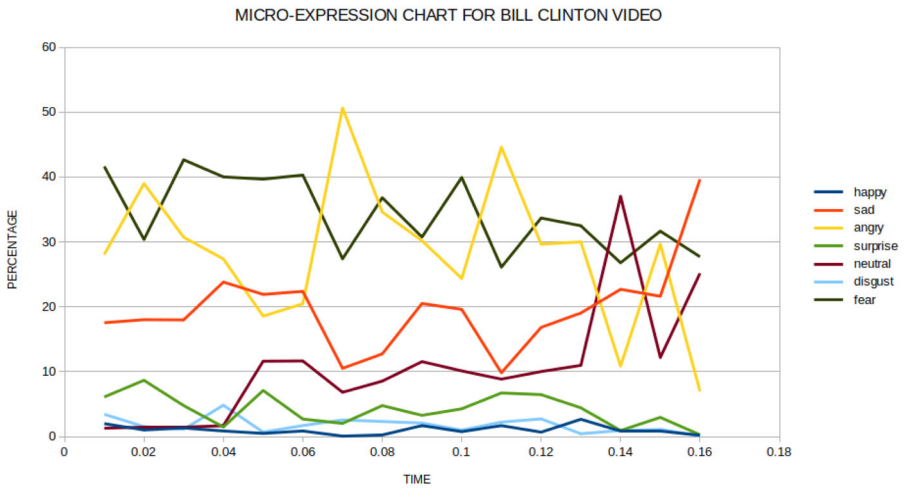


Fig. 11 Micro-expressions chart for deception detection in Bill Clinton's video

In the above chart, Bill Clinton only showed neutral as the most dominant micro-expression at a point where he paused which is an indication that neutral is most likely his baseline expression. Ascertaining the reasons for the plethora of micro-expressions is unnecessary in this case as it was later discovered that Bill Clinton was not totally sincere when he gave the speech.

From the results of the sessions, the accurate detection of micro-expressions and the comparison of the tags obtained from the participants and the performance metrics of the model; it is deducible that the developed model can adequately detect deception using micro-expressions. As the micro-expressions only appear for a short moment and hard to fake, micro-expressions greatly facilitate deception detection. When the video analysed using deception detection model, Bill Clinton's discomfort and sudden mood changes are reflected in the Fig. 11. This can be seen throughout the video. Similar to Session 2, the model indicated the deception many times in the video. Overall, custom trained model can successfully indicate the deception during all sessions.

5.9 Comparison of state-of-the-art solutions

Video, audio and text modalities which were obtained using OpenSmile [18] and OpenFace [3] softwares, have been used to feed the SVM algorithm by Jaiswal et al. [32] and they achieved 78.95% accuracy by fusioning modalities.

Rosas et al. proposed an approach using video and text modalities. By combining bigrams, unigrams, facial features and hand gestures using Decision Tree algorithm [39], they obtained 75% accuracy. Later on, they extended their research [40] and obtained 82% classification accuracy with Support Vector Machine. Karimi et al. developed their Long Short-Term Memory model [33] which achieved 84.16% accuracy, using vocal and visual temporal information.

Gogate et al. combined video, audio and text modalities. To extract the visual temporal and spatial features, they used the CNN algorithm. OpenSmile software has been

Table 4 Comparison of state-of-the-art solutions

Studies	Approach	Modalities	Features	Classifier	Accuracy
Our Study	feature-level	Video	Facial-based Features	CNN	74.17%
Karimi et al. [33]	feature-level	Video+Audio	Vocal and Visual Features	CNN	84.16%
Gogate et al. [24]	decision-level	Video+Audio+Text	Temporal and Spatial Features from Video, Vocal and Text	CNN	92%
Gogate et al. [24]	feature-level	Video+Audio+Text	Temporal and Spatial Features from Video, Vocal and Text	CNN	96%
Jaiswal et al. [32]	feature-level	Video+Text	18 OpenFace AUs 28 Vocal Features	SVM	78.95%
Rosas et al. [39]	feature-level	Video+Text	Bigrams, Facial Features Hand Gestures, Unigrams	DT	75%
Rosas et al. [40]	feature-level	Video+Text	Bigrams, Facial Features Hand Gestures, Unigrams	SVM	82%
Jaiswal et al. [32]	decision-level	Video+Audio+Text	18 OpenFace AUs 28 Vocal Features, Unigrams	SVM	76%
Chebbi et al. [7]	feature-level	Video+Audio+Text	39 Facial-based, 21 Text-based 72 Vocal-based Features	KNN	97%

used to obtain vocal features such as MFCC, pitch, loudness and intensity. Lastly, text features, words forming pronunciations converted to vectors. As a result [24], their feature-level approach achieved 96%, decision-level approach achieved 92% classification accuracy (Table 4).

6 Conclusion

In this study, the deception detection method which employs multiple facial features has been presented. To test the model's performance, a novel dataset covering recordings from human participants for quantitative and qualitative experiments has been collected. The analysis has shown nonverbal behaviours obtained from micro-expressions play a key role in deception detection. A deception detection classifier has been built which is only using non-verbal features, it achieved 74.17% correct classification accuracy. To better understand and put the results in perspective, a task that shows the human ability to identify the deception has been carried out and revealed the difficulty of deception detection as human participants performed worse than the classifier.

Computer vision and machine learning have been combined to study the pseudoscience aspect of deception detection. Future work will focus on developing a fully autonomous deception detection system. To do so, other modalities such as vocal and text will be combined with video features. Video features will be increased using head movements, hand gestures and body posture. While increasing the features, the usage of the computation

power will be taken into consideration to reduce the complexity and enable the real-time deception detection operation.

Acknowledgements Cranfield University's High Performance Technical Computing (HPC) system where the computation of the model was run. This project was supported by Republic of Turkey, Ministry of National Education and Niger Delta Development Commission, Federal Republic of Nigeria through grants awarded to the first and second authors respectively.

Author Contributions Conceptualisation S.Y., M.S.C., Z.A.R.; Methodology S.Y., M.S.C.; Software M.S.C.; Formal Analysis S.Y., M.S.C.; Draft Preparation S.Y., M.S.C.; Review and Editing S.Y., M.S.C., Z.A.R.; Supervision Z.A.R.; All authors have read and agreed to the published version of the manuscript.

Declarations

Ethics approval and consent to participate All procedures performed in studies involving human participants were done so in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Conflict of Interests None of the authors reports any conflict of interest for the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Al-modwahi AAM, Sebetela O, Batleng LN, Parhizkar B (2012) Lashkari, AH. In: Proceeding of world congress in computer science computer engineering, and applied computing. Facial expression recognition intelligent security system for real time surveillance
2. Alvino C, Kohler C, Barrett F, Gur R, Gur R, Verma R (2007) Computerized measurement of facial expression of emotions in schizophrenia. J Neurosci Methods 163(2):350–361
3. Baltrušaitis T, Robinson P, Morency L-P (2016) Openface: an open source facial behavior analysis toolkit. In: 2016 IEEE winter conference on applications of computer vision (WACV). IEEE, pp 1–10
4. Bond CF Jr, DePaulo BM (2006) Accuracy of deception judgments. Personal Soc Psychol Rev 10(3):214–234
5. Butalia DPKA, Ingle MDr (2012) Facial expression recognition for security. Int J Mod Eng Res 2(4):1449–1453. http://www.ijmer.com/papers/Vol2_Issue4/A02414491453.pdf
6. Carrier P-L, Courville A, Goodfellow IJ, Mirza M, Bengio Y (2013) Fer-2013 face database. Universit de Montral
7. Chebbi S, Jebara SB (2021) Deception detection using multimodal fusion approaches. Multimed Tools Appl:1–30
8. Crockett K, O'Shea J, Khan W (2020) Automated deception detection of males and females from non-verbal facial micro-gestures. In: 2020 International joint conference on neural networks (IJCNN). IEEE, pp 1–7
9. Darwin C (2009) The expression of the emotions in man and animals. joe cain and sharon messenger eds
10. Delmas H, Elissalde B, Rochat N, Demarchi S, Tijus C, Urdapilleta I (2019) Policemen's and civilians' beliefs about facial cues of deception. J Nonverbal Behav 43(1):59–90
11. DePaulo BM, Lindsay JJ, Malone BE, Muhlenbruck L, Charlton K, Cooper H (2003) Cues to deception. Psychol Bull 129(1):74

12. Donato G, Bartlett MS, Hager JC, Ekman P, Sejnowski TJ (1999) Classifying facial actions. IEEE Trans Pattern Anal Mach Intell 21(10):974–989
13. Dureha A (2014) An accurate algorithm for generating a music playlist based on facial expressions. Int J Comput Appl 100:33–39
14. Ekman P (1982) Methods for measuring facial action. Handb Methods in Nonverbal Behav Res:45–90
15. Ekman P (1993) Facial expression and emotion. Am Psychol 48(4):384
16. Ekman P (2009) Telling lies: clues to deceit in the marketplace, politics, and marriage (revised edition). WW Norton & Company, New York
17. Ekman P, Friesen WV (1978) Facial action coding systems. Consulting Psychologists Press, Sunnyvale
18. Eyben F, Schuller B (2015) Opensmile:) the munich open-source large-scale multimedia feature extractor. ACM SIGMultimedia Records 6(4):4–13
19. Feng S, Banerjee R, Choi Y (2012) Syntactic stylometry for deception detection. In: Proceedings of the 50th annual meeting of the association for computational linguistics (Volume 2: Short Papers), pp 171–175
20. Forkosch MD (1975) The lie detector and mechanical jurisprudence. Okla L Rev 28:288
21. Frank MG, Stennett J (2001) The forced-choice paradigm and the perception of facial expressions of emotion. J Personal Social Psychol 80(1):75
22. Gaebel W, Wölwer W (1992) Facial expression and emotional face recognition in schizophrenia and depression. Eur Arch Psychiatry Clin Neurosci 242(1):46–52
23. Gannon TA, Beech AR, Ward T (2009) Risk assessment and the polygraph, the use of the polygraph in assessing, treating and supervising sex offenders: a practitioner's guide:129–154
24. Gogate M, Adeel A, Hussain A (2017) Deep learning driven multimodal fusion for automated deception detection. In: 2017 IEEE symposium series on computational intelligence (SSCI). IEEE, pp 1–6
25. Granhag PA, Strömwall LA (2004) The detection of deception in forensic contexts, Cambridge University Press, Cambridge
26. Gupte A, Naganarayanan A, Krishnan M (2016) Emotion based music player-xbeats. Int J Adv Eng Res Sci 3(9):236854
27. Hachisuka S (2013) Human and vehicle: driver drowsiness detection by facial expression. In: Proceedings - 2013 international conference on biometrics and kansei engineering, ICBAKE 2013, pp 320–326
28. Hartwig M, Granhag PA, Strömwall LA, Kronkvist O (2006) Strategic use of evidence during police interviews: when training to detect deception works. Law Hum Behav 30(5):603
29. Hill ML, Craig KD (2004) Detecting deception in facial expressions of pain: accuracy and training. Clin J Pain 20(6):415–422
30. Hirschberg J, Benus S, Brenier J, Enos F, Hoffman S, Gilman S, Girand C, Graciarena M, Kathol A, Michaelis L, Pellom B, Shriberg E, Stolcke A (2005) Distinguishing deceptive from non-deceptive speech. In: 9th European conference on speech communication and technology vol 01, pp 1833–1836
31. Jabon M, Bailenson J, Pontikakis E, Takayama L, Nass C (2011) Facial-expression analysis for predicting unsafe driving behavior. IEEE Pervasive Comput 10(4):84–95
32. Jaiswal M, Tabibu S, Bajpai R (2016) The truth and nothing but the truth: multimodal analysis for deception detection. In: 2016 IEEE 16th international conference on data mining workshops (ICDMW). IEEE, pp 938–943
33. Karimi H, Tang J, Li Y (2018) Toward end-to-end deception detection in videos. In: 2018 IEEE international conference on big data (Big Data). IEEE, pp 1278–1283
34. Lu G, Li X, Li H (2008) Facial expression recognition for neonatal pain assessment. In: 2008 International conference on neural networks and signal processing, pp 456–460
35. Lu G, Yuan L, Li X, Li H (2008) Facial expression recognition of pain in neonates. In: 2008 International conference on computer science and software engineering, vol 1, pp 756–759
36. Mehrabian A (1968) Some referents and measures of nonverbal behavior. Behav Res Methods Instrum 1(6):203–207
37. Newman ML, Pennebaker JW, Berry DS, Richards JM (2003) Lying words: predicting deception from linguistic styles. Personal Soc Psychol Bull 29(5):665–675
38. Owayjan M, Kashour A, Al Haddad N, Fadel M, Al Souki G (2012) The design and development of a lie detection system using facial micro-expressions. In: 2012 2nd international conference on advances in computational tools for engineering applications (ACTEA). IEEE, pp 33–38
39. Pérez-Rosas V, Abouelenien M, Mihalcea R, Burzo M (2015) Deception detection using real-life trial data. In: Proceedings of the 2015 ACM on international conference on multimodal interaction, pp 59–66
40. Pérez-Rosas V, Abouelenien M, Mihalcea R, Xiao Y, Linton C, Burzo M (2015) Verbal and nonverbal clues for real-life deception detection. In: Proceedings of the 2015 conference on empirical methods in natural language processing, pp 2336–2346

41. Porter S, Ten Brinke L (2010) The truth about lies: what works in detecting high-stakes deception? Leg Criminol Psychol 15(1):57–75
42. Porter S, Ten Brinke L, Wallace B (2012) Secrets and lies: involuntary leakage in deceptive facial expressions as a function of emotional intensity. J Nonverbal Behav 36(1):23–37
43. Russell JA (1994) Is there universal recognition of emotion from facial expression? a review of the cross-cultural studies. Psychol Bull 115(1):102
44. Su L, Levine MD (2014) High-stakes deception detection based on facial expressions. In: 2014 22nd International conference on pattern recognition. IEEE, pp 2519–2524
45. Su L, Levine M (2016) Does "lie to me" lie to you? an evaluation of facial clues to high-stakes deception. Comput Vis Image Underst 147:52–68
46. Thannoon HH, Ali WH, Hashim IA (2019) Design and implementation of deception detection system based on reliable facial expression. J Eng Appl Sci 14(15):5002–5011
47. Tian Y-L, Kanade T, Cohn JF (2005) Facial expression analysis. In: Handbook of face recognition. Springer, pp 247–275
48. Vrij A (2000) Detecting lies and deceit: the psychology of lying and implications for professional practice. Wiley, New York
49. Vrij A, Semin GR (1996) Lie experts' beliefs about nonverbal indicators of deception. J Nonverbal Behav 20(1):65–80
50. Vrij A, Edward K, Roberts KP, Bull R (2000) Detecting deceit via analysis of verbal and nonverbal behavior. J Nonverbal Behav 24(4):239–263
51. Yacoob Y, Devis LS (1996) Recognizing human facial expressions from long image sequences using optical flow. IEEE Trans Pattern Anal Mach Intell 18(6):636–642
52. Zhang Z. (1999) Feature-based facial expression recognition: sensitivity analysis and experiments with a multilayer perceptron. Int J Pattern Recognit Artif Intell 13(6):893–911
53. Zhang Z, Singh V, Slowe TE, Tulyakov S, Govindaraju V (2007) Real-time automatic deceit detection from involuntary facial expressions. In: 2007 IEEE Conference on computer vision and pattern recognition. IEEE, pp 1–6

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



https://dspace.org/

The influence of micro-expressions on deception detection

Yildirim, Suleyman; Chimeumanu, Meshack Sandra; Rana, Zeeshan A.

2023-08-01
Springer
https://doi.org/10.1007/s11042-023-14551-6
Yildirim S, Chimeumanu MS, Rana ZA. (2023) The influence of micro-expressions on deception detection. Multimedia Tools and Applications, Volume 82, Issue 19, August 2023, pp. 29115-29133

Downloaded from a DSpace repository. https://dspace.org/