



Tarafder Elmi Tabassum

ML-Enhanced Visual Inertial Navigation System for UAVS

School of Aerospace Transport and Manufacturing
Aerospace

PhD Thesis
Academic Year: 2021 - 2024

Supervisor: Ivan Petrunin
Associate Supervisor: Zeeshan Rana
December 2024



School of Aerospace Transport and Manufacturing
Aerospace

PhD Thesis

Academic Year 2021 - 2024

Tarafter Elmi Tabassum

ML-Enhanced Visual Inertial Navigation System for UAVs

Supervisor: Ivan Petrunin
Associate Supervisor: Zeeshan Rana
December 2024

This thesis is submitted in partial fulfilment of the requirements for
the degree of PhD Aerospace

(NB. Remove if the degree award is based solely on examination of the thesis)

© Cranfield University 2024. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

Academic integrity declaration

I declare that:

- the thesis submitted has been written by me alone.
- the thesis submitted has not been previously submitted to this university or any other.
- that all content, including primary and/or secondary data, is true to the best of my knowledge.
- that all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

Abstract

Urban Air Mobility (UAM) applications, including passenger taxis, cargo transport and aerial surveys, are significantly impacted by GNSS vulnerabilities in urban environments highlighting the requirement of accurate and robust alternative navigation solutions like Visual-Inertial Navigation Systems (VINS) to ensure reliable and safe operation. However, ensuring robustness and high accuracy in the VINS system requires the development of advanced methodologies to address the limitations of current state-of-the-art solutions, which struggle to perform effectively in low-light environments, adverse weather conditions, low-texture areas, and during rapid motion variation. Therefore, the primary aim of this thesis is to develop a robust and resilient vision-based alternative solution to GNSS outage by mitigating visual faults, enabling safe and reliable navigation under challenging urban environments for UAM systems.

This thesis designs an advanced methodology to address visual faults by exploring hybrid architectures in VINS systems to improve robustness in complex visually degraded urban environments, ensuring operational safety during GNSS outages. To enhance robustness and reliability, a hybrid VINS system incorporating multiple error compensation (Multi-ML Hybrid VIO) is proposed that simultaneously addresses and mitigates visual failure mode effects arising from weather conditions, lighting effects, low-texture environments and flight dynamics to ensure integrity. Furthermore, uncertainty prediction for visual failure modes is proposed to address aleatoric and epistemic uncertainty, providing compensation for accumulated errors that improve accuracy and robustness. The evaluation of different hybrid architectures proposed during this study indicates that Multi-ML Hybrid Visual Inertial Navigation Systems (VIO) solution outperforms in mitigating visual failure modes and addressing various sources of visual uncertainty, ensuring navigation integrity. Training and testing results obtained from simulated complex urban prototypes demonstrate the remarkable performance of the proposed solution under diverse visually degraded scenarios and its seamless integration within the multi-sensor navigation system during GNSS outage.

Overall, the findings in this thesis demonstrate a robust and reliable vision-based alternative navigation solution tailored for UAM applications.

Keywords:

Complex Urban Environment; Machine Learning; Error Compensation; Uncertainty; Hybridization; Urban Air Mobility.

Acknowledgements

I would like to extend my gratitude to all those who supported me throughout my PhD studies.

Firstly, I am incredibly grateful to supervisors, Dr. Ivan Petrunin and Dr. Zeeshan Rana, for granting me this exceptional opportunity as a Ph.D candidate and for their invaluable guidance throughout the study process. Their continuous support and supervision have been crucial in shaping my research and academic development. I am particularly appreciative of Dr. Ivan Petrunin's insightful suggestions and ideas that have significantly contributed to the refinement of my paper writing and successful submission. His expertise and mentorship have played a vital role in my growth as a researcher. Once again, I express my deep appreciation to Dr. Ivan Petrunin and Dr. Zeeshan Rana for their unwavering support and mentorship during my PhD journey.

Moreover, I would like to express my sincere gratitude to my parents who supported me unconditionally throughout my academic journey. I would like to dedicate this thesis to my father. He has always been my biggest supporter and admirer. My parents' belief in my abilities has been a constant source of inspiration, and I am deeply thankful for their sacrifices and encouragement throughout my PhD study.

Furthermore, I would like to convey my utmost appreciation and understanding to my best friends Anika, Parisa, Shorna, Sharmin, Mohua, Mumu, Ehsana and Moon. Their unwavering support, love and understanding have been instrumental in my academic journey. Additionally, I expressed my gratitude to Anirudh and Huw to solve my problems and guide me right direction. Lastly, special thanks to my younger brother and sister from another mother, Farhad and Jafnah for cheering me up and tolerating my anxiety periods.

Table of Contents

Academic integrity declaration.....	i
Abstract	ii
Acknowledgements	iv
List of Figures.....	viii
List of Tables	xii
List of Abbreviations	xiv
1 Introduction.....	16
1.1 Background and Motivation	16
1.2 Research Aim and Objectives.....	22
1.3 Contribution	23
1.4 Publications	24
1.5 Thesis Outline	25
2 Integrating GRU with a Kalman Filter to Enhance Visual Inertial Odometry Performance in Complex Environments	28
2.1 Introduction	28
2.2 Related Works	31
2.2.1 Kalman Filter for VIO.....	31
2.2.2 Hybrid Fusion Enhanced by AI	33
2.2.3 FMEA in VIO	36
2.3 Proposed Fault-Tolerant Navigation System	38
2.3.1 Failure Mode and Effect Analysis (FMEA).....	39
2.3.2 Fault-Tolerant Federated Navigation System Architecture.....	41
2.3.3 Proposed GRU-aided ESKF VIO Integration (Sub-Filter 1).....	42
2.3.4 EKF Based GNSS/IMU Integration (Sub-Filter 2).....	45
2.3.5 Federated Filter for Multi-Sensor Fusion	46
2.4 Experimental Setup.....	48
2.5 Test and Results	51
2.5.1 Experiment 1—Dense Urban Cynon	52
2.5.2 Experiment 2—Semi-Structured Urban Environment.....	55
2.5.3 Performance Evaluation Based on Zone Categories	58
2.5.4 Performance Comparison with Other Datasets	65
2.6 Conclusions	67
3 A Comparative Analysis of Hybrid Sensor Fusion Schemes for Visual– Inertial Navigation.....	68
3.1 Introduction	68
3.2 Hybrid Visual-Inertial Navigation Architectures	70
3.2.1 Method 1	72
3.2.2 Method 2	73
3.2.3 Method 3	74
3.3 Experimental Setup and Data Generation	75

3.3.1 Experimental Setup	75
3.3.2 Training Datasets	77
3.3.3 Evaluation metrics	79
3.4 Performance analysis	80
3.4.1 Single Fault Condition	81
3.4.2 Combined Fault Conditions with Weather Effects	85
3.4.3 Complex Fault conditions with un-seen scenarios	93
3.5 Conclusion	96
4 Improving Generalization Ability using GRU-AE based Error Compensation in Hybrid Visual Inertial Navigation Systems	98
4.1 Introduction	98
4.2 Proposed Novel Hybrid VIO Framework	102
4.2.1 GRU-AE VO Error Correction.....	103
4.2.2 GRU-based ESKF Model Error Compensation	105
4.2.3 Machine Learning Model: GRU	106
4.2.4 Multi-ML-aided Hybrid VIO	108
4.3 Experimental Setup and Data Generation	109
4.3.1 Training Datasets	111
4.3.2 Training and Testing Process.....	112
4.4 Performance Evaluation.....	114
4.4.1 Proposed Multi-layer GRU-architecture based on Performance Evaluation under different trajectory shapes	115
4.4.2 Nominal Fault Conditions	117
4.4.3 Weather effects	120
4.4.4 Lighting Effects.....	124
4.4.5 Unseen scenario with unknown fault conditions.....	130
4.4.6 Comparison with literature.....	133
4.5 Conclusion	135
5 Uncertainty Prediction and Model Correction by Bayesian GRU in Visual Navigation Applications	138
5.1 Introduction	138
5.2 Proposed Uncertainty Prediction and Error Correction Mechanism	141
5.2.1 Bayesian GRU	142
5.2.2 Uncertainty Prediction in Hybrid VIO	143
5.2.3 Hybrid Multi-Sensor Navigation System	144
5.3 Uncertainty Inspection in Visual Navigation	147
5.4 Experimental Setup & Data Collection	149
5.4.1 Training Methodology:.....	150
5.4.2 Testing Phase	152
5.4.3 Uncertainty Evaluation Metrics	152
5.5 Performance Evaluation.....	153
5.5.1 Uncertainty Estimation in Vision-based Alternative Navigation	155

5.5.2 Multi-Sensor Navigation System Performance.....	166
5.6 Conclusion	167
6 General Discussion	170
7 Conclusion and Future Work.....	174
Reference.....	176

List of Figures

Figure 1-1 Illustration of the relationship between, true position and the required integrity level by the Stanford diagram [33].....	18
Figure 1-2 The workflow of the thesis.....	26
Figure 2-1 Published articles on the hybrid machine learning usage in KF	34
Figure 2-2 Fault-tolerant multi-sensor aided with FMEA framework.....	39
Figure 2-3 Fault tree analysis for feature-based VIO.....	39
Figure 2-4 The architecture of the proposed federated fault-tolerant multi-sensor navigation system.....	41
Figure 2-5 Illustrative diagram of GRU model.	45
Figure 2-6 Experimental simulation setup in MATLAB.	49
Figure 2-7 Overall visual navigation error in the presence of multiple failure modes for Experiment 1.....	53
Figure 2-8 Visual navigation error along each axis in the presence of multiple failure modes for Experiment 1.....	54
Figure 2-9 The 3D-position error in the presence of multiple failure modes for Experiment 2.	56
Figure 2-10 Position error along each axis in the presence of multiple failure modes for Experiment 2.....	57
Figure 2-11 Feature tracking and Feature mismatch error failure modes due to tall-shadowed buildings in experiment 1.....	59
Figure 2-12 Position errors of Zone 1 (shadowed building) estimated in each axis for experiment 1 (a). Position errors of Zone 1 (shadowed building) estimated in each axis for experiment 2 (b).	60
Figure 2-13 Motion dynamics, feature tracking and feature mismatch error failure modes due to the tall, shadowed buildings in experiment 2.....	61
Figure 2-14 Position errors of Zone 2 (shadowed building and UAV turns) estimated in each axis for experiment 1 (a). Position errors of Zone 2 (shadowed building and UAV turn) estimated in each axis for experiment 2 (b).	61
Figure 2-15 Feature tracking, feature extraction error, and feature mismatch failure modes for dark and well-illuminated tall buildings in experiment 2.	62
Figure 2-16 Position errors in Zone 3 (UAV turn and illuminated tall buildings) estimated in each axis for experiment 2.	63
Figure 2-17 Feature extraction error due to the plain wall in experiment 1.....	64

Figure 2-18 Position errors of Zone 4 (UAV turn, shadowed buildings, and featureless wall buildings) estimated in each axis for experiment 1.	64
Figure 2-19 The 3D position error estimated using a MH05 seq EUROC Dataset with a motion blur failure mode.....	66
Figure 3-1 Categories of hybrid NN-KF techniques according to [80]	69
Figure 3-2 The high-level architecture of adopted techniques implemented with VIO systems.	71
Figure 3-3 ESKF Error compensation structure for forward propagation.....	74
Figure 3-4 Kalman gain kk predictor structure for forward propagation.	75
Figure 3-5 Examples of simulated environmental conditions used in testing and training datasets.	77
Figure 3-6 Examples of trajectories used for training purposes.....	79
Figure 3-7 3D view of test trajectories used for experiments.....	81
Figure 3-8 3D trajectory comparison of method 1, method 2, and method 3 algorithms under single fault condition.....	82
Figure 3-9 Example of feature tracking errors and feature mismatch failure modes due to the presence of white rooftop view at 115s.....	82
Figure 3-10 Position error statistics for adopted three hybrid architectures using trajectory 1.....	83
Figure 3-11 Position error in the presence of a single fault condition for Trajectory 1	85
Figure 3-12 Position error statistics for adopted three hybrid architectures under weather effects using trajectory 2.	87
Figure 3-13 Position error along with each axis in the presence of a single fault condition for trajectory 2 under weather effects	89
Figure 3-14 Trajectory comparison of hybrid algorithms under diverse weather conditions for trajectory 2.....	92
Figure 3-15 3D trajectories of hybrid algorithms under complex fault conditions with unseen scenarios for trajectory 3.	94
Figure 3-16 Boxplot summarizing the position error statistics for adopted three hybrid architectures under weather effects using trajectory 3.....	94
Figure 4-1 Proposed combined hybrid architecture with the denoising mechanism for VIO position uncertainty reduction.....	103
Figure 4-2 Architecture of GRU-AE VO Error Correction Model.....	105
Figure 4-3 Architecture of KF Measurement Error Compensator	106

Figure 4-4 High level multi-Layer GRU Architecture used for error compensation modules in Multi-ML VIO	107
Figure 4-5 MATLAB incorporating Unreal Engine Simulation Environment for Data Correction	111
Figure 4-6 Fault Scenarios for Training Datasets	111
Figure 4-7 Examples of simulated environmental conditions: (a) limited field of view during 80% fog, (b) texture-less wall during 70% cloud, (c) illumination changes during 30% rain and (d) shadows and texture-less wall during partly sunny used in testing and training datasets.....	112
Figure 4-8 3D view of testing trajectories for experiments (a) Trajectory 1, (b) Trajectory 2, (c) Trajectory 3, (d) Trajectory 4, (e) Trajectory 5, (f) Trajectory 6	115
Figure 4-9 4-layer GRU architecture	116
Figure 4-10 Horizontal error comparison highlighting the impact of incorporating additional layers in GRU architecture within Multi-ML hybrid VIO.....	117
Figure 4-11 (a) Position error plot in the presence of single fault condition for trajectory 1; (b) 3D trajectory comparison of the proposed multi-ML hybrid VIO framework with four reference systems for trajectory 1.	118
Figure 4-12 Horizontal error distribution comparison for trajectory 1	120
Figure 4-13 Position error and 3D trajectory comparison under weather effects for trajectory 2 (a) Position error under foggy; (b) Position error under rainy; (c) 3D trajectory under foggy; (d) 3D trajectory under rainy.....	121
Figure 4-14 Error distribution under weather effects for trajectory 2 (a) Foggy weather; (b) Rainy weather.....	123
Figure 4-15 Position error and 3D trajectory comparison under lighting effects for trajectory 3 (a) Position error under shadow; (b) Position error under sunny; (c) 3D trajectory under shadow; (d) 3D trajectory under sunny	125
Figure 4-16 Error distribution under lighting effects for trajectory 2 (a) shadow, (b) sunny.....	127
Figure 4-17 Position error and horizontal error distribution comparison for different shape of trajectory under sunrise light conditions (a) position error for trajectory 5; (b) position error for trajectory 6; (c) horizontal error distribution for trajectory 5; (d) horizontal error distribution for trajectory 6.	128
Figure 4-18 Position error and 3D trajectory comparison for trajectory 4 under unseen scenario with unknow fault conditions.....	131
Figure 4-19 Error distribution in horizontal positioning for trajectory 4 under unseen scenario with unknown fault conditions.....	132

Figure 4-20 3D position error estimated using MH05_difficult seq. from EUROC dataset with multiple sources of aleatoric uncertainty.....	134
Figure 5-1 Robust fault-tolerant multi-sensor fusion framework with vision-based A-PNT uncertainty prediction to enhance integrity in the absence of GNSS	142
Figure 5-2 Fault propagation in feature-based visual navigation representing sources of uncertainty.....	148
Figure 5-3 3D view of testing trajectories for experiments (a) Trajectory 1, (b) Trajectory 2, (c) Trajectory 3, (d) Trajectory 4, (e) Trajectory 5, (f) Trajectory 6, (g) Trajectory 7, (h) Trajectory 8	154
Figure 5-4 Position error comparison for trajectory 7 and trajectory 8 under unseen scenario and seen scenarios with complex dynamics.....	156
Figure 5-5 Uncertainty comparison based on 95th percentile UI and 99th percentile UI on trajectory 1, trajectory 2, trajectory 3, and trajectory 5...	161
Figure 5-6 Uncertainty comparison based on 95th percentile UI and 99th percentile UI on trajectory 4, trajectory 6, trajectory 7, and trajectory 8...	163
Figure 5-7 Uncertainty estimation in EuRoc MH05- difficult dataset.....	165
Figure 5-8 3D error comparison on integrated multi-sensor navigation system for trajectory 6 and trajectory 7	166

List of Tables

Table 2-1 Common faults in the visual positioning based on a state-of-the-art review.	40
Table 2-2 Specification of the ICM 20649 IMU model.	50
Table 2-3 RMSE Comparison on the performance of two experiments.....	55
Table 2-4 Maximum position error comparison for two experiments.	58
Table 2-5 The RMSE of 3D position errors of GNSS/faulted-IMU/faulted-VO solution with faulted-GNSS/faulted-IMU/faulted-VO.	65
Table 2-6 Comparison with state-of-the-art methods that used MH_05 seq. EUROC Dataset with motion blur failure mode.....	66
Table 3-1 Specification of ‘Simulation 3D Scene Configuration’ block in MATLAB 2022a.....	76
Table 3-2 Characteristics of training trajectories	78
Table 3-3 The defined experiments covering scenarios	80
Table 3-4 Positioning performance for Trajectory 1 under single fault conditions.	84
Table 3-5 Positioning performance for Trajectory 2 under sunny weather conditions.	90
Table 3-6 Positioning performance for Trajectory 2 under foggy weather conditions.	90
Table 3-7 Positioning performance for Trajectory 2 under rainy weather conditions.	90
Table 3-8 Positioning performance for Trajectory 3 under unseen scenarios. .	95
Table 4-1 Test trajectory description	114
Table 4-2 RMSE and 95 th percentile Comparison on the performance of single fault conditions for trajectory 1	119
Table 4-3 RMSE and 95 th percentile comparison on the performance under foggy weather effects for trajectory 2	122
Table 4-4 RMSE and 95 th percentile comparison on the performance under rainy weather effects for trajectory 2	123
Table 4-5 RMSE and 95 th percentile comparison on the performance under lighting effects for trajectory 3.....	126
Table 4-6 RMSE and 95 th percentile comparison on the performance under sunrise light conditions for trajectories 5 and 6.....	129

Table 4-7 RMSE and 95th percentile comparison on the performance under unseen scenario with unknown fault conditions for trajectory 4	133
Table 4-8 Comparison with SOTA methods that used MH05_difficult sequence from EuRoc dataset.....	134
Table 5-1 Comparison between Bayesian hyperparameter optimisation based on training trajectories	151
Table 5-2 Test trajectory description	153
Table 5-3 RMSE and 95th percentile comparison on trajectory 7 and 8 under sunny conditions.....	157
Table 5-4 Uncertainty estimation for the testing trajectories.....	158
Table 5-5 Comparative analysis of uncertainty estimation with SOTA architecture	165
Table 5-6 RMSE comparison results on Trajectory 6 and 7 using integrated multi-sensor navigation system	167

List of Abbreviations

AL	Alert Limit
ANN	Artificial Neural Network
APNT	Alternative Positioning, Navigation and Timing
CKF	Cubature Kalman Filter
DNN	Deep Neural Network
ESA	European Space Agency
ESKF	Error-State Kalman Filter
EKF	Extended Kalman Filter
EU	Epistemic Uncertainty
FMEA	Failure Mode and Effects Analysis
FTA	Fault Tree Analysis
GRU	Gated Recurrent Unit
GNSS	Global Navigation Satellite System
HAU	Heteroscedastic Aleatoric Uncertainty
IMU	Inertial Measurement Unit
INS	Inertial Measurement System
IR	Integrity Risk
KF	Kalman Filter
LSTM	Long Short-Term Memory
LiDAR	Light Detection and Ranging
ML	Machine Learning
MSCKF	Multi-state Constraint Kalman Filter
MSE	Mean Square Error
MCD	Monte Calo Dropout
NN	Neural Network
NLOS	No-line-of-sight
OR	Out of Range
PNT	Positioning, Navigation and Timing
PF	Particle Kilter
PE	Position Error
PL	Protection Level
RAIM	Receiver Autonomous Integrity Monitoring

RCTM	Residual Chi-square Test Mothed
RMSE	Root Mean Square Error
UAVs	Unmanned Aerial Vehicles
UKF	Unscented Kalman Filter
UI	Uncertainty Interval
UAM	Urban Air Mobility
UWB	Ultra-Wide Band
VINS	Visual Inertial Navigation Systems
VIO	Visual Inertial Odometry
VO	Visual Odometry
VSLAM	Visual Simultaneous Localisation and Mapping

1 Introduction

1.1 Background and Motivation

Urban Air Mobility represents an emerging air transportation system designed for passengers and cargo in urban environments, facilitating by advanced technologies and integrated within multimodal transportation systems [1], [2]. The vision for UAM involves widespread utilization in urban and suburban areas, complementing existing transportation systems and supporting the decarbonization of the transportation sector [3]. Unmanned Aerial Vehicles (UAVs) deployed for UAM applications often required to navigate through urban space, including flying between buildings. In such scenarios, Global Navigation Satellite System (GNSS) proves to be an unreliable navigation system due to its vulnerability to environmental obstructions such as multipath effects, no-line-of-sight (NLOS), jamming, signal blockage, atmospheric effects [4], [5], [6]. To address these challenges, GNSS is often integrated with the Inertial Navigation System (INS) or Inertial Measurement Unit (IMU) using fusion algorithms. However, challenges persist as errors in these systems accumulate over time [7], [8], [9], [10], [11]. For instance, while INS utilises GNSS signals to provide accurate positioning through continuous dead reckoning, it struggles during extended GNSS outages or when signals are intermittent or noisy. Alternative approaches such as using external sensors like lasers [12], Lidar [13], [14], UWB [15], [16], [17] have also been explored. However, as pointed out by the authors [18], [19], [20], [21], these methods are prone to unbounded errors from generated measurements, leading to solutions that either diverge or become corrupted by noise. Unlike Light Detection and Ranging (LiDAR) or Ultra-Wide Band (UWB) systems, vision-based navigation systems have emerged as a promising, cost-effective, and lightweight alternative solution to address the limitations of GNSS and navigate the complexities of challenging urban environments, making them preferred choice for UAM systems.

Current research trends in autonomous UAV navigation point towards the growing use of visual sensors with Inertial Measurement Unit (IMU) incorporating advanced sensor fusion techniques to enhance positioning performance in urban

areas. By leveraging visual information, UAVS can navigate through urban areas, where GNSS may be obstructed or unreliable [22], [23], [24], [25]. Visual Inertial Navigation System (VINS) has been extensively investigated and deployed on robotic applications, encompassing notable examples like VINS-Mono [26], MSCKF [27], [28], ORB-SLAM3 [29], and open-VINS [30]. However, these state-of-the-art solutions exhibit performance degradation under fault conditions such as illumination, motion, blur, field of view, and texture environments. These limitations hinder the reliability of existing solutions for UAM applications, which require safety assurance with high accuracy, particularly for navigating urban space, including flying between buildings or above rooftops. An extensive review of vision-based navigation solutions highlights the requirement of exploring advanced fault mitigation techniques in the VINS system to enhance accuracy and robustness in such complex scenarios. Despite advances, challenges remain in ensuring robustness and integrity for autonomous driving in UAM systems, emphasising the critical need for further research in this domain [31].

In a safety context, Position, Navigation and Timing (PNT) solutions are categorised by accuracy, integrity, continuity, and availability within the performance-based navigation solution [32]. While accuracy is the initial stage of the PNT solution, the trustworthiness of the system is specified based on the level of accuracy as integrity. Continuity is the simultaneous success rate of accuracy and integrity requirements over a flight period. Finally, the navigation system is available only if the above three requirements are achieved. Integrity reflects the trustworthiness of the navigation solutions and is defined by four key parameters: Alert Limit (AL), Integrity Risk (IR), Time to Alert (TIA), and Protection Level (PL), with the latter ensuring position errors remain within defined statistical bounds [31]. Based on the Stanford diagram [33] as illustrated in Figure 1-1, safe mission operation is achieved if AL or PL is above PE. Further, correct integrity algorithm implementation is achieved if PL is always above PE. This means that the system continuously guarantees to alert the mission controller if the information is correct. While GNSS-based integrity-assured systems are extensively studied, challenges remain in ensuring the integrity of vision-based navigation solutions.

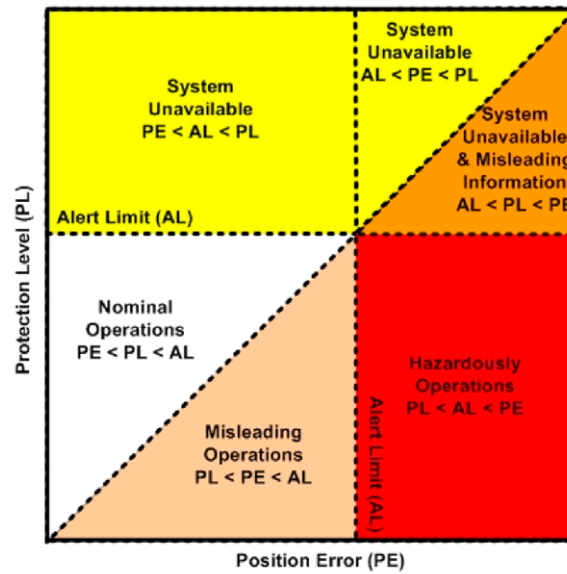


Figure 1-1 Illustration of the relationship between, true position and the required integrity level by the Stanford diagram [33].

Since integrity failures can cause collisions and accidents in UAM applications, the tolerance of the integrity Alert limit (AL) is set to $\leq 10\text{m}$ while the integrity risk is at $\leq 10^{-7}$ by the European Space Agency (ESA) [31]. While maintaining the previously defined availability of the navigation system, attention should be given to ensuring robustness and reliability to avoid hazardous outcomes, which can enhance integrity. A comprehensive review [32], [34] on integrity challenges for PNT solutions has highlighted the requirement to address robustness in the presence of fault conditions while designing navigation algorithms to ensure integrity. The review suggests that prioritising robustness and integrity may come at the expense of accuracy but has been shown to enhance reliability in complex environments, ultimately reducing the risk of crashes and collisions.

To maintain both accuracy, robustness and integrity requirements, a multi-sensor navigation system in challenging urban environments is essential, where a vision-based Alternative PNT (APNT) solution can be integrated to ensure safe navigation. The traditional centralised multi-sensor navigation system imposes the limitation of assuring safe navigation under a prominent visually degraded environment during GNSS outages [35]. For achieving the fault tolerance capability in integrated multi-sensor systems, the decentralised filtering design,

especially using the federated architecture, is considered to be a promising solution [36], [37]. However, this further requires assessing the reliability and robustness of APNT solutions as they serve as the core solution during GNSS outages. Current developments in VINS systems still focus on the mitigation and detection of conventional faults or predefined fault models of large biases, which require attention. Due to the complexity in resolving the inherent nonlinear behaviour of core visual sensors, developments were only focused on mitigating IMU faults in VINS systems [38], [39], [40], [41], [42], [43] involving filter-based and data-driven approaches. Another approach has endeavoured in literature to integrate the CNN-LSTM-aided EKF-based VIO [44], [45] approach, aiming to predict measurements from either Visual Odometry (VO) or IMU. These attempts have encountered difficulties in adequately presenting the complexity of dynamic scenes due to the insufficient variety of datasets and the selection of mitigation mechanisms. Attention should be given to addressing visual failure modes while designing navigation solutions to achieve significant improvement. Zhu et. al. [31] have highlighted the requirement of investigation of possible fault conditions or failure modes occurring in visual navigation while designing a robust and high-integrity PNT solution for safety-critical applications. To facilitate this, there is a need to reexamine and reevaluate current state-of-the-art multi-sensor navigation systems to achieve fault tolerance capabilities.

Given the constraints of current state-of-the-art PNT solutions for UAM application, the development of integrity-enhanced robust vision-based PNT solutions facilitating uninterrupted navigation in an urban complex environment faces substantial challenges and research gaps, which are addressed by this research as follows-

- The development of multi-sensor navigation systems, specifically requiring fault-tolerance capabilities that ensure operational continuity even if one of the sensors fails to perform. While centralised multi-sensor navigation systems have explored integrity monitoring methodologies [35], there is a lack of evidence ensuring operational continuity under visually degraded scenarios. However, this work provides a direction for considering failure modes in navigation systems while designing integrity-enhanced

solutions. Another approach [36], [39], [46] explores federated architectures within multi-sensor systems, proposing an adaptive positioning algorithm to eliminate issues of GNSS signal interruption by employing VO large bias. However, faults arising from less-features, environment complexity, motion dynamics and high non-linear characteristics have not been fully taken into account. Besides, the fault detection mechanism employed on such architectures operates based on statistical models, which can result in false alerts due to the dependency on pre-defined Gaussian error models. In contrast, error distribution in such complex navigation systems is non-Gaussian in the real world. Thus, there is a need to consider realistic fault conditions to design a fault mitigation mechanism to enhance the PNT solution. The motivation of this component is to address the possible visual failure modes that result in various fault conditions in visually degraded urban environments.

- The Research gap emerges in developing an appropriate vision-based navigation solution that is robust under various fault conditions arising from UAV flight dynamics and visually degraded scenarios. Current solutions, including filter-based [26], [47], data-driven [43], [48], [49] and hybrid architectures [41], [45] lack robustness under realistic maximum fault conditions. For example, challenges related to routing remote locations, handling speed and multiple manoeuvres remain largely unexplored in state-of-the-art vision-based solutions. Current approaches to removing visual errors involve the integration of predefined fault models like feature extraction error, photometric noise, and geometric error, which may not be proven effective under other faults. For example, outlier rejection method like RANSAC fails to reduce fault effects when there is a mismatch and lack of features in the data [50], [51], [52]. Furthermore, challenges remain in visual navigation while operating in different weather conditions, with safety assurance by mitigating visual errors. For instance, raindrops, glare and limited visibility remain challenging to existing VO solutions due to additional noise induced in the data [53], [54]. Thus, Attention should be given to mitigate visual errors while developing the vision-based solution

to ensure robust and safe navigation. The motivation of this component is to develop robust fusion solutions that account for possible failure modes, facilitating uninterrupted operation during different flight paths with combinations of fault conditions.

- Another gap lies in the evaluation of existing solutions on datasets that encompass such scenarios, apart from EuRoc Dataset [55] which is recorded indoor and suffers from poor synchronisation between sensor outputs. Additionally, fusion solutions utilising this dataset often prioritise achieving high accuracy, necessitating its division into six small sets due to the presence of high dynamics and dark lighting conditions, which frequently lead to feature tracking failure [26], [29], [41].
- The ongoing development of vision-based APNT solutions presents significant challenges in establishing confidence measures to ensure integrity. Quantifying error arising from visual sources of uncertainties poses an ongoing challenge, particularly in the context of enhancing integrity through visual navigation [31], [51], [56]. Existing studies on error quantification often rely on predefined uncertainty models and Gaussian representation of errors [57], [58], [59] which lack evidence of providing reliable confidence measures across various fault scenarios. ML-aided solutions incorporating into visual-based systems also struggle to provide confidence measures for epistemic (associated with the model) and aleatoric (associated with input data) uncertainty. For example, popular ML methods like Bayesian LSTM utilising Monte Carlo Dropout (MCD) [60], Bayesian Variational Autoencoder [61], [62] and RCNN with mixture density network [63], employed on visual systems also lack validation under visually degraded scenarios (e.g. texture-less environments, rapid motion, low-light. Furthermore, existing solutions often lack statistical rigour, particularly when dealing with non-Gaussian error distributions, where 3σ uncertainty interval (UI) cannot ensure error bound. The ongoing motivation is to bridge this gap by developing an advanced solution to predict uncertainty for non-Gaussian error distribution within vision-based

PNT systems to provide confidence measures for safe and reliable navigation during GNSS outage in complex urban environments.

Therefore, the research challenges and gaps in the domain of vision-based APNT solutions for UAM applications emphasise the requirement of developing advanced methodologies in autonomous operation in urban environments. The motivation for addressing these gaps is rooted in the need to establish a robust vision-based navigation solution that provides confidence measures for operational safety under complex environments to ensure integrity. Furthermore, bridging these gaps aligns with the goals of UAM applications, facilitating uninterrupted and reliable flight in urban areas.

1.2 Research Aim and Objectives

The research aims to develop robust advanced methodologies to mitigate visual failure modes in vision-based navigation systems to ensure safe operation under challenging urban environments.

Based on the outlined challenges and gaps in safe UAV navigation for UAM applications, the specific objectives of the thesis can be expressed as follows-

- **O1. State-of-the-art Review**

Review the state-of-the-art vision-based navigation algorithms, analyse challenges in sensor fusion algorithms incorporating VO, and evaluate existing public datasets.

- **O2. Visual failure modes mitigation within hybrid VIO**

Develop a hybrid solution to mitigate visual failure modes identified through FMEA on VINS applications and evaluate its effectiveness within a fault-tolerant multi-sensor navigation system to enhance performance under multiple visual failure modes.

- **O3. Error Compensation within Hybrid VIO mitigating visual failure modes**

Develop error compensation mechanisms within hybrid VIO architectures to evaluate the effectiveness and adaptability of proposed methodologies

for mitigating visual failure modes under adverse scenarios like weather conditions, lighting conditions, and flight dynamics to ensure robustness.

- **O4. Uncertainty prediction of visual failure modes to ensure integrity**

Develop an uncertainty prediction technique for complex visual failure modes arising from various visual sources of uncertainties, providing confidence measures and enabling position correction to achieve an accurate and robust solution while ensuring integrity.

- **O5. Evaluation of the effectiveness of different hybrid VIO solutions in the mitigation of visual failure modes**

Evaluate the resilience of different hybrid architectures in vision-based APNT solutions, incorporating GRU and ESKF to mitigate visual failure modes and gain insight into the behaviour of hybrid models in complex urban environments with external combinations of faults.

1.3 Contribution

This work makes contributions to advancing alternative PNT techniques for navigating under urban environments for safety-critical UAM applications. The detailed contributions are listed as follows-

- [1] The novel GRU-based enhancement of the ESKF VIO system for VO error correction is proposed to mitigate FMEA-identified visual failure modes within the fault-tolerant multi-sensor navigation system under GNSS outages. This research lays the validation of utilising failure modes of visual navigation to design system architecture, a step forward towards ensuring integrity in UAM applications navigating in urban space.
- [2] The novel KF error compensation mechanism in hybrid VIO is proposed with a combination of GRU and ESKF, aiming to mitigate visual failure modes arising from various fault factors like environmental complexity and flight dynamics.
- [3] A novel robust Multi-ML hybrid VIO is proposed with a combination of multiple ML-aided error compensation and ESKF, aiming to reduce position uncertainty. This integrates one GRU to compensate for VO measurement errors by updating the ESKF filter, while another GRU

compensates for model measurement errors to correct model inaccuracies. The aim is to correct aleatoric and epistemic uncertainty under complex urban environments simultaneously.

- [4] A novel denoising mechanism using an autoencoder for GRU-based VO error compensators is proposed to minimise outliers and improve the integrated solution's generalisation ability in the presence of a scarcity of features, reflections, and noise artefacts.
- [5] Systematically evaluated the resilience of various hybrid VIO architectures proposed in this study to address complex visual failure modes, including feature extraction errors, feature tracking errors, feature location errors and data association errors.
- [6] A novel Bayesian GRU is proposed within Multi-ML hybrid VIO to predict model errors with uncertainty measures of complex visual failure modes, which is utilised to correct the overall position. The algorithm aims to estimate a robust position with uncertainty estimation, balancing robustness and integrity under complex urban environments.
- [7] Systematically evaluated integrated fault-tolerant hybrid GNSS/IMU/VO multi-sensor navigation system leveraging the Multi-ML hybrid vision-based APNT solution developed in this study, demonstrating its effectiveness and ability to provide accurate positioning with confidence measures under GNSS outage scenarios.

1.4 Publications

The subsequent sections contain a compilation of journal and conference papers that have been either submitted or published, all of which have direct relevance to the scope of this Ph.D. research.

Journal Papers

- **J1.** Tabassum, T. E., Xu, Z., Petrunin, I., & Rana, Z. A. (2023). Integrating GRU with a Kalman Filter to Enhance Visual Inertial Odometry Performance in Complex Environments. *Aerospace*, 10(11), 923. <https://doi.org/10.3390/aerospace10110923> (Published)

- **J2.** Tabassum, T. E., Petrunin, I., & Rana, Z. A. A Comparative Analysis of Hybrid Fusion Schemes for Visual–Inertial Navigation (**Under Review - Manuscript ID: TIM-24-04121 on IEEE Transactions on Instrumentation and Measurement**)
- **J3.** Tabassum, T. E., Petrunin, I., & Rana, Z. A. Improving Generalization Ability using GRU-AE based Error Compensation in Hybrid Visual Inertial Navigation Systems (**To be Submitted to Institute of Navigation Journal**)
- **J4.** Tabassum, T. E., Petrunin, I., & Rana, Z. A. Uncertainty Prediction and Model Correction by Bayesian GRU in Visual Navigation Applications (**To be Submitted to IEEE Transactions on Mechatronics**)

Conference Papers

- **C1.** Elmi Tabassum, T., Petrunin, I., Rana, Z., (2022). Failure Mode Analysis (FMA) for Visual-Based Navigation for UAVs in Urban Environment. UKRAS22 Conference “Robotics for Unconstrained Environments” Proceedings, 28-29. doi: 10.31256/Mx7Zx4J (**Published**)
- **C2.** Tabassum TE, Petrunin I, Rana Z. (2022) A fault tolerant multi-sensor fusion navigation system for drone in urban environment. In: POSNAV 2022: Positioning and Navigation for Intelligent Transport Systems, 3-4 November 2022, Berlin, Germany (**Published**)
- **C3.** Tabassum, Tarafder Elmi, Petrunin, Ivan, Rana, Zeeshan A., "Position Uncertainty Reduction in VisualInertial Navigation Systems Using Multi-ML Error Compensation," Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2024), Baltimore, Maryland, September 2024, pp. 1741-1755. <https://doi.org/10.33012/2024.19790> (**Published**)

1.5 Thesis Outline

This thesis adopts a paper-based format, with chapters (from 2 to 5) corresponding to a distinct paper that has been either published or submitted. As a result, every chapter stands as a self-contained representation of

comprehensive review, methodology, results, discussion, and conclusions. This format facilitates a coherent presentation of the research, allowing each chapter to delve into a distinct scope of the study.

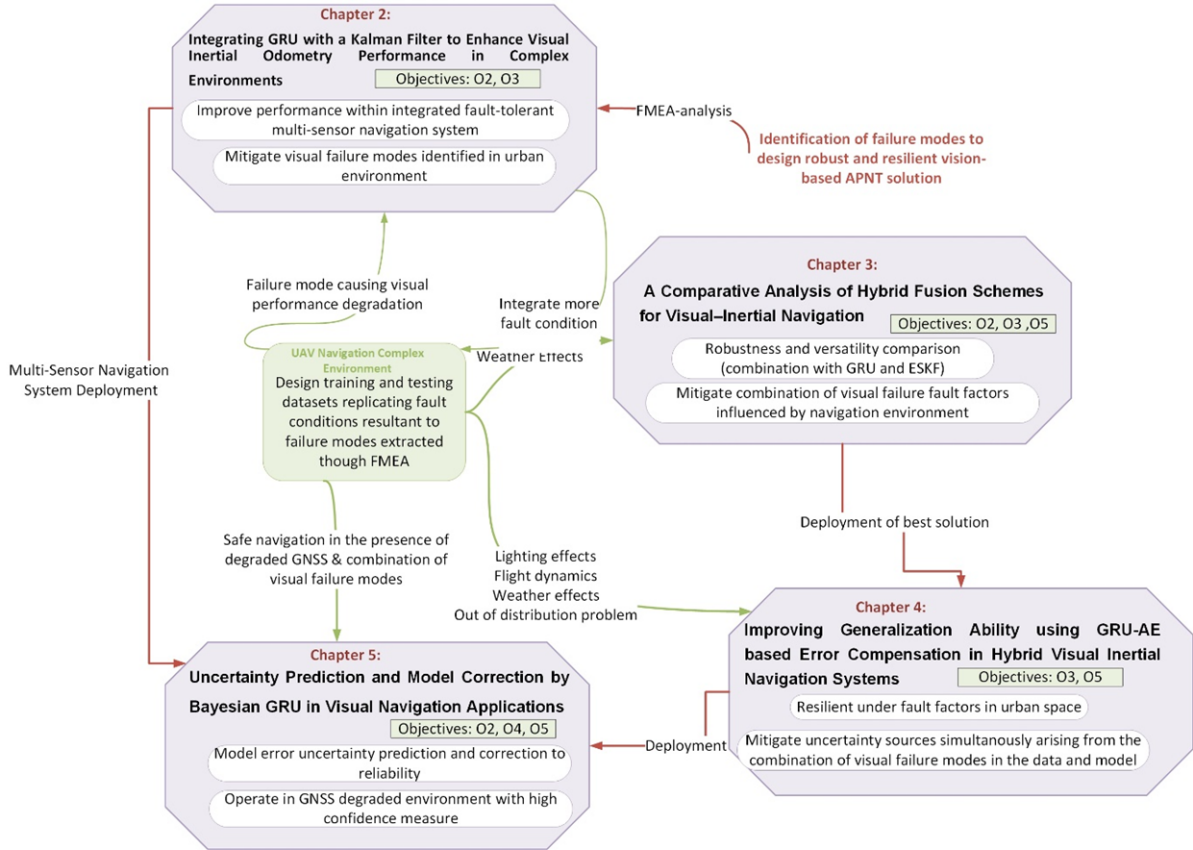


Figure 1-2 The workflow of the thesis

The thesis presents a structured approach toward addressing key challenges in vision-based navigation systems for UAVS. Chapter 1 introduces background and motivation, research objectives of the study, with their interconnections and relevance to the subsequent chapters and related publication, illustrated in Figure 1-2. Chapter 2 focuses on mitigating visual failure modes using FMEA-supported hybrid VIO methods to improve performance in the urban environment. Building on this foundation, chapter 3 explores advanced error compensation techniques within hybrid VIO architectures, comparing them with other hybrid architectures to evaluate their resilience and adaptability under adverse weather conditions and flight dynamics. Extending this work, chapter 4 proposes an advanced methodology that integrates error compensation techniques into hybrid VIO,

reducing both data and model position uncertainty while improving generalisation across various visual failure modes. Chapter 5 introduces techniques for uncertainty predictions and compensation of complex failure modes, ensuring accurate and robust visual navigation. The developed visual navigation solution integrates into multi-sensor navigation systems to evaluate performance enhancements during GNSS outage to ensure operational continuity.

Finally, the concluding Chapters 6 and 7 bring the thesis to a close, offering a comprehensive overview and discussion of the research's outcomes and suggesting potential avenues for further exploration.

2 Integrating GRU with a Kalman Filter to Enhance Visual Inertial Odometry Performance in Complex Environments

2.1 Introduction

In recent years, unmanned aerial vehicles (UAVS) have gained attention with the evolution of technologies such as artificial intelligence and computer vision, which have effectively broadened pathways for diverse applications and services. UAVS have been utilised in many civil applications, such as aerial surveillance, package delivery, precision agriculture, search and rescue operations, traffic monitoring, remote sensing, and post-disaster operations [64]. The increasing demand for commercial UAVS for such applications has highlighted the need for robust, secure, and accurate navigation solutions. However, achieving accurate and reliable UAV positioning in complex environments, including overpasses, urban canyons, illumination variability, etc., has become increasingly challenging.

Although Global Navigation Satellite Systems (GNSS) have become one of the most popular navigation systems in recent decades, the utilisation of GNSS remains suspicious due to its vulnerability to satellite visibility, interference of jamming and spoofing, as well as environmental effects such as multipath, building mask, ionospheric and tropospheric delays. Furthermore, the effects will lead to sharp deteriorations in the positioning precision and GNSS availability [65]. The inertial navigation system (INS) facilitates the provision of high-frequency and continuous position, velocity, and attitude information, which makes the integration of INS with GNSS prevalent in most navigation architecture designs. However, the drift error generated from INS accumulates over time, which will result in divergent positioning output. The impact of INS drifting on GNSS/INS fusion performance in the case of long-term GNSS outages has been explored widely [42], [66], [67]. Nevertheless, more sensor types are still in demand to provide more resilient and accurate positioning resolutions under complex operation scenarios.

The vision-based navigation system is a promising alternative for providing reliable positioning information without radio frequency interference effects during GNSS outages. Visual odometry VO is frequently employed as a critical component in the vision-based navigation system due to its efficient deployment with low computational complexity in contrast to Visual Simultaneous Localisation and Mapping (VSLAM). The Visual Inertial Navigation Systems (VINS) have been fully explored by researchers, encompassing notable examples like VINS-Mono [26], MSCKF [68], ORB-SLAM3 [29], and open-VINS [30]. As a common solution for improving navigation performance in terms of accuracy, integrity, update rate, and robustness through adding sensor types with GNSS, the VINS navigation system with multiple integrated sensors presents a higher possibility of existing multiple faults, noise, and sensor failures within the system. It was discovered that purely VO-enabled navigation presents performance degradation caused by factors such as illumination, motion blur, field of view, moving objects, and texture environment [31]. As a result, there is a need to explore fault-tolerant designs in the VINS navigation systems to mitigate the fault impact on the visual systems.

For achieving the fault tolerance capability in integrated multi-sensor systems, the decentralised filtering design, especially using federated architecture, has become popular in recent years. Dan et al. [39] proposed an adaptive positioning algorithm based on a federated Kalman filter combined with a robust Error-Estate Kalman Filter (ESKF) with adaptive filtering for the UAV-based GNSS/IMU/VO navigation system to eliminate issues of GNSS signal interruption and lack of sufficient feature points while navigating in indoor and outdoor environments. However, most studies using ESKF only measure VO faults by adding position errors, whilst faults coming from visual cues like scarcity of features caused by environment complexity and motion dynamics or high non-linearity characteristics have not been fully taken into account. Therefore, there is a gap in detecting and identifying faults and threats with consideration of visual cues in the GNSS/IMU/VO navigation system.

Current state-of-the-art fault-tolerant GNSS/IMU/VO navigation systems encounter more difficulties when operating in complex scenarios due to the

challenges of identifying visual failure modes and correcting VO errors. As a structured and systematic fault identification method, Failure Mode and Effect Analysis (FMEA) is capable of identifying various fault types, defects, and sensor failures based on predicted and measured values as they occur instantly or shortly after they occur. FMEA is commonly used to assess risks to improve the reliability of complex systems by identifying and evaluating potential failures with the provision of occurrence likelihood, severity of impact, and detectability, as well as prioritizing high-risk failure modes. However, researchers working on VIO discussed several faults caused by navigation environment or sensor error individually [31], [45], [69], [12,13] but identifying failure modes is a gap. Moreover, despite their extracted failure modes, systematic faults have not been discovered with only the consideration of single or specific combined faults like motion blur, rapid motion, and illumination variation.

To resolve the inherent non-linearity in the visual navigation system, AI has been employed with a Kalman filter to enhance the ability to model temporal changes in sensor data. Nevertheless, AI has the disadvantages of training time and predicted value inevitability, containing errors that can be partially resolved by simplifying the neural network, for example, with a Gradient Recurrent Unit (GRU) suppressed with ESKF fusion, thus enabling ESKF to better handle scenarios with verifying level of uncertainty and dynamic sensor conditions.

To implement a fault-tolerance navigation system against the visually degraded environment, this chapter proposes a GRU-aided ESKF VIO algorithm that conducts FMEA on the VINS system to identify failure modes and then assists the architecture of a fault-tolerant multi-sensor system where the AI-aided ESKF VIO integration is used as one of the sub-filters to correct identified visual failure modes. The major contributions of this chapter are highlighted as follows:

1. The proposition of an FMEA-supported fault-tolerant federated GNSS/IMU/VO integrated navigation system. The FMEA execution on an integrated VINS system contributes to enhancing the system's design, with a focus on accurate navigation during GNSS outages.

2. The proposition of a GRU-based enhancement of ESKF for predicting increments of positions to update measurements of ESKF, aiming to correct visual positioning errors, leading to more accurate and robust navigation in challenging conditions.
3. Performance evaluation of GRU-aided ESKF-based VIO within the fault-tolerant GNSS/IMU/VO multi-sensor navigation system. Training datasets for the GRU model are selected to replicate the failure modes extracted with fault conditions from FMEA. The verification is simulated and benchmarked on the Unreal engine, where the environment includes complex scenes of sunlight, shadow, motion blur, lens blur, no-texture, light variation, and motion variation. The validation dataset is grouped into multiple zone categories in accordance with single or multiple fault types due to environmental sensitivity and dynamic motion transitions.
4. The performance of the proposed algorithm is compared with the state-of-the-art End-to-End VIO and Self-supervised VIO by testing similar datasets on the proposed algorithm.

The remaining part of the chapter is organised as follows. Section 2.2 discusses the existing systems designed based on a hybrid approach; Section 2.3 introduces the proposed GRU-aided KF-based federated GNSS/INS/VO navigation System; Section 2.4 discusses the experimental setup; Section 2.5 discusses the roast test, and result analysis comparison with state-of-the-art systems, and the conclusion is presented in Section 2.6.

2.2 Related Works

2.2.1 Kalman Filter for VIO

A Kalman Filter (KF), along with its variations, is a traditional method that can efficiently fuse VO and IMU information. Despite its effectiveness, the KF faces certain challenges that can impact its performance. One of the main challenges in navigation applications arising from VO vulnerability is the interruption in updating KF observation, leading to a gradual decline in system performance over time. Moreover, if the error characteristics are non-Gaussian and cannot be

fully described within the model, the KF may struggle to provide an accurate estimation.

Some studies aim to improve the fusion robustness against non-linear natures from high system dynamics and complex environments, variants of Kalman filter such as the Extended Kalman Filter (EKF) [47], [70], [71], [72], Multi-state Constraint Kalman Filter (MSCKF) [27], [28], [73], [74], [75], Unscented Kalman Filter (UKF) [76], Cubature Kalman filter (CKF) [77], [78], and Particle Kalman Filter (PF) [79], have been proposed and evaluated. One challenge in the EKF-based VIO navigation system is handling significant non-linearity during brightness variations [47], [71] and dynamic motion [71], which will cause feature-matching errors and degrade the overall performance. When these feature-matching errors occur, the EKF's assumption about linear system dynamics and Gaussian noise may no longer hold, leading to suboptimal states and even filter divergences. To improve VIO performance under brightness variation and significant non-linearities, MSCKF was proposed and evaluated in complex environments such as insufficient light [27], [28], [74], [75], texture missing [27], [28] and camera jitter [27] specifically characterised by blurry images [75]. However, in the MSCKF algorithm, the visual features are mostly treated as separate states, so the process is delayed until all visual features are obtained. Another approach is proposed to handle additive noise: VIO uses camera observation to update the filter, allowing for avoiding over-parameterisation and helping to reduce growth caused by UKF [76]. In most studies like [77], [78] researchers have not focused on evaluating their proposed algorithms among complex scenarios such as light variation, rapid motion, motion blur, overexposures, and field of view, which deteriorate the accuracy and robustness of the state estimation.

Nevertheless, in GNSS-denied scenarios, UAV navigation predominantly depends on VIO, so the challenges remain unresolved using KF variations only for VIO applications. Therefore, the use of ESKF-based VIO for performance improvement is highlighted through alleviating challenges by managing parameter constraints, mitigating singularity and gimbal lock concerns,

maintaining parameter linearization, and supporting sensor integration in this chapter.

2.2.2 Hybrid Fusion Enhanced by AI

A number of artificial techniques, such as neural networks, including Deep Neural Networks (DNN), Artificial Neural Networks (ANN), and Reinforcement Learning (RL), have been studied for sensor fusion applications to formulate hybrid fusion solutions. Kim et al. [80] conducted a detailed review of KF with AI techniques to enhance the capabilities of the KF and address its specific limitations in various applications. The recent survey [81] summarizes detailed reviews of the GNSS/INS navigation system that utilises Artificial Neural Networks (ANN) in combination with the Kalman Filter. This survey highlights the advantages of hybrid methods that leverage ANNs to mitigate INS performance degradation during GNSS vulnerability in specific conditions of aerial underwater vehicles [82], [83], [84] and Doppler underwater navigation[85]. It was found that one advantage of the hybrid fusion scheme is providing the ability to intercorporate a priori knowledge about the level of change in the time series, enhancing the system's adaptability to varying environments and conditions. Additionally, the ANN error predictor proves to be successful in providing extremely precise adjustments to standalone INS when GNSS signals are unavailable, ensuring continuous and accurate navigation. This survey motivates further exploration and development of hybrid fusion-based navigation schemes to maintain robust performance in challenging GNSS-degradation environments.

Figure 2-1 presents an overview of publications on NN-assisted navigation applications over the 2018–2023 period and categorizes them into KF performance degradation following the category rule defined by Kim et al. [80] It is concluded that most studies focus on updating the state vector or measurements of KF, ignoring issues arising from imperfect models. Hereby, estimating pseudo-measurements during GNSS outages to update KF measurements using AI is suggested by [80]. In certain scenarios, particularly when dealing with high nonlinear sensors or navigating in complex environments, updating state vectors directly by predicting sensor increments using NN in

measurement steps be-comes critical regarding the sensor's complexity and the challenging nature of the environment.

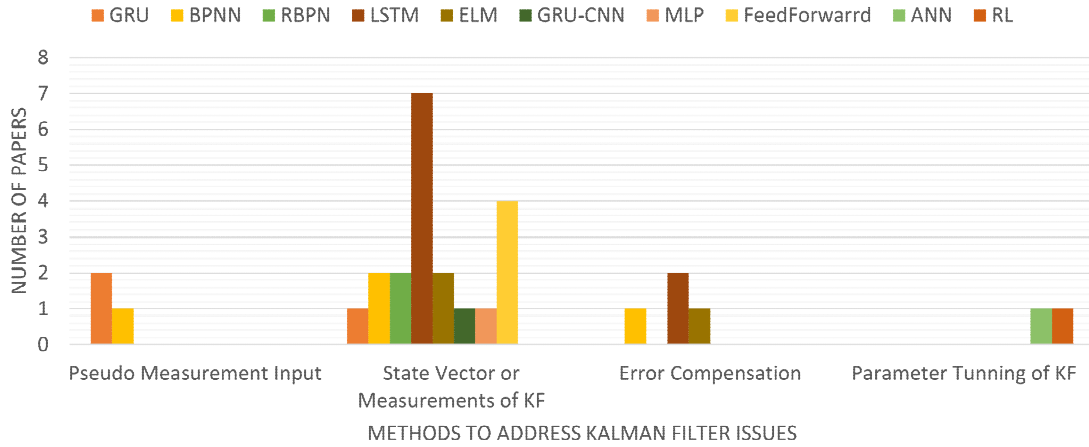


Figure 2-1 Published articles on the hybrid machine learning usage in KF

To overcome KF drawbacks in navigation, hybrid methods combining AI approaches, especially machine learning (ML) algorithms, have become promising due to the accurate prediction of INS and visual sensor errors from diverse training datasets. As depicted in Figure 1, most publications belong to the category of ‘State Vector or Measurements of KF’, meaning most studies apply ML to predict and compensate for state vector error or measurements in KF. For instance, Zhang et al. [86] used RBFN to predict states in the prediction step. However, studies like [86] predict absolute state vectors instead of vector increments using NN, which increases model complexity and requires a more extensive training process. Studies from [44], [82], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97] adopted vector increments of the sensor observations and predictions during KF prediction, whilst most of the work only works on GNSS/INS navigation during GNSS outages, aiming to improve INS efficiency in urban settings and situations [11], [84], [85], [92], [94], [95][31,38,39,41,42,46].

Other studies [98], [99] corresponding to the error compensation category, employ ML to compensate for the navigation performance error with KF, but sensor errors such as the non-linear error model of INS, are excluded.

The category called pseudo measurement input applies NN for predicting pseudo-range errors when the occurrence of a shortage of satellite numbers to update measurement steps in CKF [100] and the adaptive EKF [11], [101]

The category of parameter tuning of KF aims to enhance the performance using RL by predicting the covariance matrix, such as using AKF in [101]. However, the prediction of the covariance matrix relies on changeable factors like temperature, which is difficult to predict and model. Therefore, using NN for parameter tuning in KF is not considered in this chapter.

Regarding the selection of NN types, Radial Basis Factor Neural Networks (RBFNs), Backpropagation Neural Networks (BPNNs), and Extreme Learning Machines (ELMs) are powerful learning algorithms and more suitable for static data or non-sequential problems, but neglect the information of historical data. Additionally, navigation applications are inherently time-dependent and dynamic, making them challenging to model using these learning algorithms. Other studies [83], [91] have proved RBFN has less complexity than BPNN and multi-layer backpropagation networks, yet haven't considered any dynamic changes over time. Other studies were presented using simple neural networks on SLAM applications by predicting state increments in diverse scenarios.

The authors in [88], [90] did not account for the temporal variations in features that can significantly impact the performance, given that basic neural networks are sensitive to such changes. Kotov et al. [89] compared NNEKF-MPL and NNEKF-ELM, demonstrating that NNEKF-MPL performs better when the vehicle exhibits a non-constant systematic error. However, the aforementioned NN methods do not take temporal information contained within historical data into training, making those methods insufficient for addressing navigation applications' dynamic and time-dependent characteristics.

Some studies have proved the advantages of using time-dependent recurrent NN architectures like Long Short-Term Memory (LSTM) in VIO. VIIONET [49] used LSTM to process high-rate IMU data that concentrated on feature vectors from images processed through CNN. Although adopting IMUs can facilitate the mitigation of IMU dynamic errors, the compensation of visual cues in VIO affected

by complex environmental conditions is more critical. DynaNet [102] was the first to apply LSTM-aided EKF to show hybrid mechanism benefits in improving motion prediction performance on VO and VIO without sufficient visual features. Furthermore, this study suggested unresolved challenges of motion estimation tasks, multi-sensor fusion under data absence, and data prediction in visual degradation scenarios. Subsequently, researchers aim to learn VO positions from raw image streams using CNN-LSTM [38], [49], [103]. With the utilisation of CNN-LSTM, some aimed to reduce IMU errors by predicting IMU dynamics in complex lighting conditions [103], [104].

Two recent studies attempted to use CNN-LSTM-based EKF VIO [44], [45] to evaluate visual observation in dynamic conditions. Still, the evaluation of the algorithm's performance is insufficient due to the lack of sufficient datasets for training the CNN model. One of the common drawbacks of using DL-based visual odometry is the necessity of huge training datasets. The DL model requires vast amounts of diverse data to generalise well and provide reliable results in the real world. Suppose the CNN-LSTM model is not trained properly due to the insufficient variety of datasets. In that case, it may struggle to capture the complexity of dynamic scenes, leading to subpar performance and unreliable VIO results [44], [45]. Thus, by leveraging feature-based techniques, I can simplify the VIO structure where ESKF facilitates the simplification by providing a robust mechanism for handling uncertainty and noise in the data.

This study chooses GRU due to its efficiency in processing time-varying sequences, offering advantages over other recurrent NNs like LSTM. Some literature has used GRU to predict state increments and update the state vector of KF [94], [99] for GNSS/INS systems during GNSS outages. However, the fusion of GRU with GNSS/INS/VIO remains unexplored as of yet.

2.2.3 FMEA in VIO

FMEA can provide a detailed description of potential faults in VIO that can affect the whole system, leading to positioning errors in complex environments. In 2007, Bhatti et al. [105], [106] carried out FMEA in INS/GPS integrated systems to categorise potential faults with their causes, characteristics, impact on users, and

mitigation methods. They discussed how this advanced fault analysis could help to improve positioning performance. Du et al. [31], [106] reviewed GNSS precise point positioning vulnerability and assisted the researchers in examining failure modes to enhance performance. Current developments in visual navigation research apply conventional fault analysis, which has given a reason to adapt the existing GNSS fault analysis approach to solve the crucial problems raised by specific characteristics of visual sensors. Zhai et al. [13] have investigated visual faults in visual navigation systems and suggested that identifying potential faults in such systems could prevent users from being threatened by large visual navigation errors. However, the proposed analysis could not consider all of the potential threats and faults contributing to positioning errors in such systems. A recent work by Brandon et al. [45] tested their proposed Deep VO-aided and EKF-based VIO on four individual faults by corrupting images of the EU-ROC dataset [58] with methods like overshoot, blur, and image skipping and showed their proposed algorithm's performance compared to other state-of-the-art systems.

However, this study aims to investigate FMEA usage on VIO to give an overview of the faults that occur while navigating complex environments experiencing disruptions in system performance. Here, carefully analysing the characteristics of the failure modes presented in the VIO assists the research in recognising the essentials of error compensators while designing a fault-tolerant federated architecture that combines GNSS, VO, and IMU.

Nevertheless, current state-of-the-art systems include advanced techniques for fault detection and mitigation, including Interacting Multiple Models (IMMs) [107], [108], hypothesis tests [109], [110] and Mahalanobis distances [111], [112], [113]. IMM and hypothesis testing have drawbacks of utilizing predefined model and assumptions about system behaviour that may not always hold in complex, realistic scenarios, leading to detection errors when faced with unexpected faults or change in the system. These can lead to missed detection when working with noisy sensor data due to manual tuning. The Mahalanobis distance techniques for system model error and abnormal measurements are sensitive to the distribution and correlations of multivariate data. When dealing with high-

dimensional data, it may lead to difficulty estimating the covariance matrix accurately.

In contrast, learning-based approaches are used in error correction that learn complex patterns and representations in the data. These approaches can generalise different fault types and scenarios, provided they have been trained on a diverse dataset [81], while classical methods may struggle with new or unforeseen fault patterns. Thus, I have adopted the learning-based method as the error compensator in my research.

Therefore, to mitigate failure modes arising from diverse, complex scenarios in the urban environment, this chapter proposes a GRU-aided ESKF-based VIO, aiming to enhance visual positioning. Later, this proposed system contributes to one of the sub-systems in fault-tolerant federated multi-sensor navigation systems to mitigate overall position errors and increase reliability in multiple fault conditions arising from IMU, GNSS, and VO sensor errors. Ultimately, my proposed fault-tolerant system aims to provide a reliable and effective solution for navigation in complex conditions where VO and GNSS/INS systems may face limitations.

2.3 Proposed Fault-Tolerant Navigation System

To correct visual positioning errors that arise from multiple systematic faults when navigating in urban areas, FMEA is executed at the first step to identify and analyse systematic failure modes according to the extracted fault tree model. The failure modes are prioritised based on potential impact and likelihood of occurrence, enabling the anticipation and mitigation of visual positioning errors. With the FMEA outcome, the hybrid GRU-aided ESKF VIO algorithm is discussed, as well as the algorithm implementation following the federated multi-sensor framework. The overall fault-tolerant multi-sensor system aided by FMEA is shown in Figure 2-2.

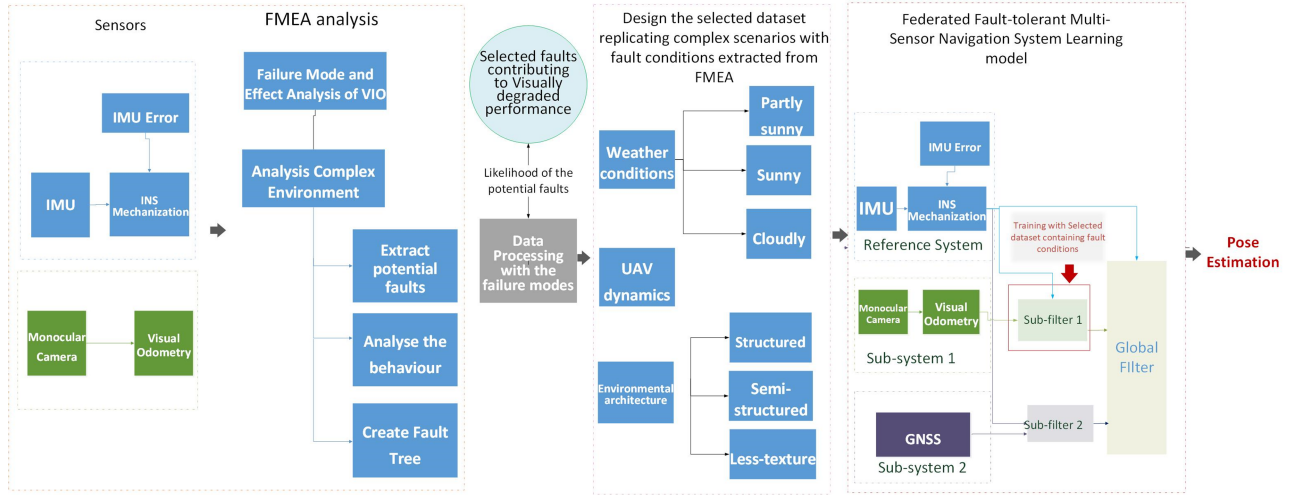


Figure 2-2 Fault-tolerant multi-sensor aided with FMEA framework.

2.3.1 Failure Mode and Effect Analysis (FMEA)

The implementation of FMEA on vision-based navigation systems enables the breaking down of high-level failure events into lower-level events along with allocating risks. Referring to error sources in every domain from the literature review [31], the fault tree model shown in Figure 2-3 is extracted.

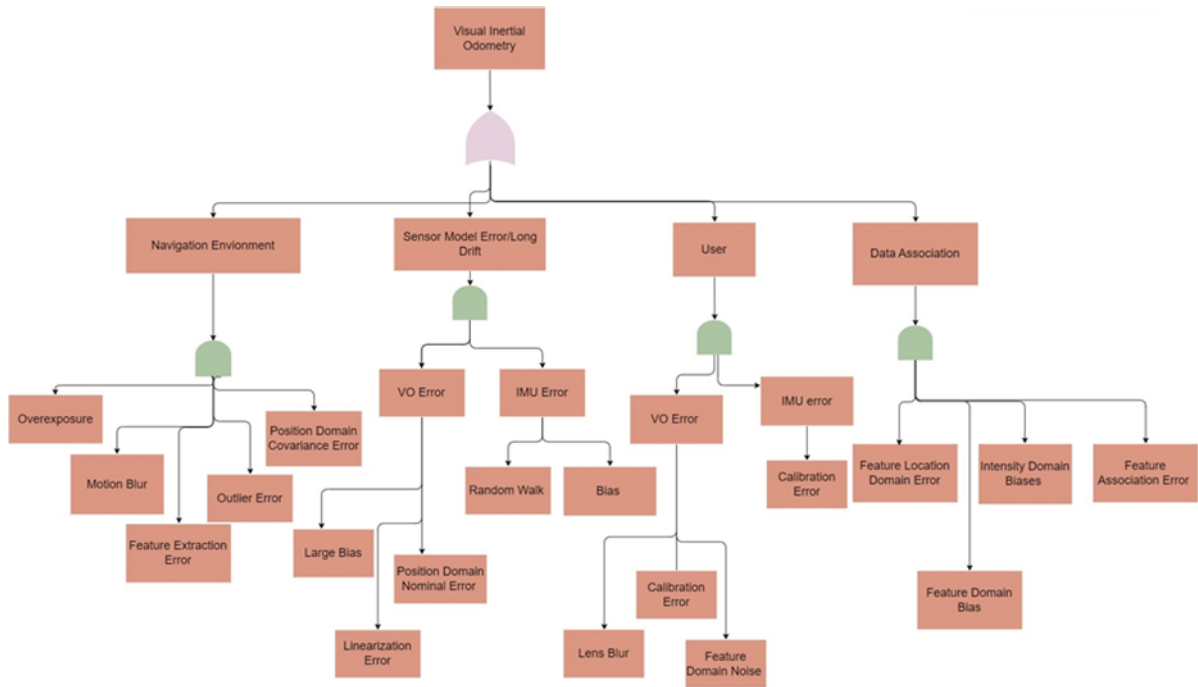


Figure 2-3 Fault tree analysis for feature-based VIO.

The preliminary conclusion of FMEA is that the error occultation in the camera presents a higher likelihood of the possibility during feature extraction due to the presence of multiple faults resulting in position errors in the whole system. Specifically, two major fault events, i.e., navigation environment errors and data association errors [31], show higher possibilities of faults in visual systems. Table 2-1 lists reviews of common error sources over the navigation environment and data association error events, along with their visual faults targeted for mitigation in the context of VIO.

- One common fault in the navigation environment fault event is the feature extraction error that contains deterministic biases that frequently lead to position errors.
- Another common fault in the data association failure event is the feature association error that occurs during matching 2D feature locations with 3D landmarks.
- The sensor model error/long drift failure events represent errors generated by sensor dynamics, including VO error and IMU error types.
- User failure events stand for the errors created during user operations that are normally relevant to user calibration mistakes.

Table 2-1 Common faults in the visual positioning based on a state-of-the-art review.

Error Sources	Fault Event	References	Error Effect
Feature Tracking Error		[45]	Motion blur
Outlier Error	Navigation environment error	[52]	Overexposure
Feature Extraction Error		[12,45]	Rapid Motion
Feature Mismatch		[10,12]	Overshoot
Feature Domain Bias	Data Association error	[52]	Lighting Variation

With the extension of fault events from the state-of-the-art reviews and proposition of thorough error analysis, i.e., Figure 2-3 this study aims to mitigate feature extraction errors occurring in failure modes linked to navigation environment and data association events through a fault-tolerant GNSS/IMU/VO

navigation system. The hybrid integration of VIO holds great potential for achieving precise and reliable navigation performance in complex conditions.

2.3.2 Fault-Tolerant Federated Navigation System Architecture

The utilisation of a federated architecture proves advantageous in implementing fault-tolerant multi-sensor systems, as it is known for its robustness in handling faults. This chapter proposes a federated architecture-based integrated multi-sensor fusion scheme based on the IMU, VO, and GNSS combination. The overall architecture of my method is shown in Figure 2-4.

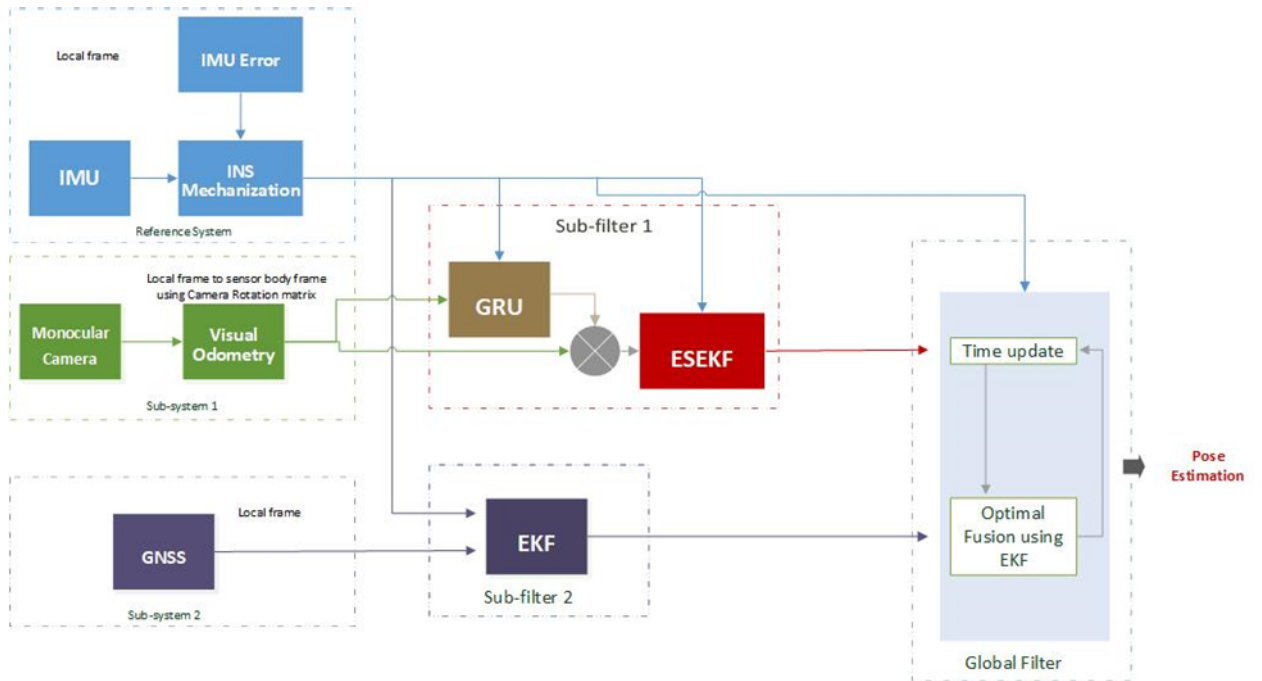


Figure 2-4 The architecture of the proposed federated fault-tolerant multi-sensor navigation system.

Two sub-filters exist in the proposed architecture: a hybrid GRU-aided ESKF IMU/VO sub-filter and an EKF-based traditional GNSS/IMU integration sub-filter. The output of the two sub-filters is merged together with a global EKF to generate the ultimate position estimations. The former sub-filter of the hybrid GRU-aided ESKF IMU/VO attempts to compensate for VO errors, while the latter sub-filter of EKF-based GNSS/IMU integration aims to correct errors from GNSS and IMU.

2.3.3 Proposed GRU-aided ESKF VIO Integration (Sub-Filter 1)

2.3.3.1 ESKF VIO Fusion

In the tightly coupled GRU-aided ESKF VIO integration, pose measurements generated by VO are fused with the linear acceleration generated by the accelerometer and angular velocity generated by the gyroscope. The extraction of visual features by VO is firstly adapted to produce relative pose measurements for ESKF updates [83]. The system filter uses GRU-predicted VO increments to correct the corresponding VIO states to obtain the corrected position.

The state vector x of the proposed GRU-aided ESKF VIO selects the following states:

$$x = [P_x, P_y, P_z, V_x, V_y, V_z, A_x, A_y, A_z, \Delta\alpha_{b_x}, \Delta\alpha_{b_y}, \Delta\alpha_{b_z}, \Delta\omega_{b_x}, \Delta\omega_{b_y}, \Delta\omega_{b_z}] \quad 2-1$$

where, P_x, P_y and P_z denote the position, V_x, V_y and V_z denote velocity, A_x, A_y and A_z denote attitude, $\Delta\alpha_x, \Delta\alpha_y$ and $\Delta\alpha_z$ denote accelerometer bias and $\Delta\omega_x, \Delta\omega_y$ and $\Delta\omega_z$ denote gyroscope bias in the x-axis, y-axis and z-axis.

The widely used, tightly coupled VIO is based on a Kalman filter [50], [83], [114]. The system dynamic model and measurement model are:

$$x_k = \Sigma_k x_{k-1} + Q_k \quad 2-2$$

$$z_k = H_k x_k + \Pi_k \quad 2-3$$

where x_k and x_{k-1} represent the system vector at k and $k - 1$ epoch; Σ_k and Q_k represent the state transition matrix and system processing noise; z_k and H_k represent the measurement vector and measurement matrix; and Π_k represents the measurement noise.

Differing from conventional KF, ESKF uses an error-state representation, which has several benefits in terms of computational efficiency, numerical stability and prevention of singularities or gimbal lock issues. The details of the nominal and error states are included in the state-of-the-art research.

In order to improve positioning in diverse visual conditions, the predicted error measurements $P_{VO/GRU}$ are added to the error-state in a measurement update of ESKF. After the error-state update, the nominal state is updated with corrected error-states using the appropriate compositions:

$$x_k = x_{k-1} \oplus \widehat{\partial X}_k \quad \mathbf{2-4}$$

In addition, the updated estimated state vector by correcting error states are:

$$\hat{P}_k = \hat{P}_{k|k-1} + \widehat{\partial P}_k + \Delta P_{VO|k-1} \quad \mathbf{2-5}$$

$$\hat{V}_k = \hat{V}_{k|k-1} + \widehat{\partial V}_k \quad \mathbf{2-6}$$

$$\hat{A}_k = \hat{A}_{k|k-1} + \widehat{\partial A}_k \quad \mathbf{2-7}$$

$$\hat{\alpha}_k = \hat{\alpha}_{k|k-1} + \widehat{\partial \alpha}_k \quad \mathbf{2-8}$$

$$\hat{\omega}_k = \hat{\omega}_{k|k-1} + \widehat{\partial \omega}_k \quad \mathbf{2-9}$$

where, $\hat{P}_k$ and $\hat{P}_{k|k-1}$ denote the nominal position vector at k and $k - 1$ epoch; $\widehat{\partial P}_k$ denotes the measured error-state position; $\Delta P_{VO|k-1}$ denotes the VO predicted increments vector at $k - 1$ epoch; $\hat{V}_k$ and $\hat{V}_{k|k-1}$ denote the nominal velocity at k and $k - 1$; $\widehat{\partial V}_k$ denotes the measured error-state velocity; $\hat{A}_k$ and $\hat{A}_{k|k-1}$ denote the nominal attitude at k and $k - 1$; $\widehat{\partial A}_k$ denotes measured error-state attitude; $\hat{\alpha}_k$ and $\hat{\alpha}_{k|k-1}$ denote nominal acceleration bias at k and $k - 1$; $\widehat{\partial \alpha}_k$ denotes the measured error-state acceleration bias; $\hat{\omega}_k$ and $\hat{\omega}_{k|k-1}$ denote the nominal gyroscope bias at k and $k - 1$; and $\widehat{\partial \omega}_k$ denotes measured error-state gyroscope bias. After the error integrates into the updated nominal state, the error-state variables need to reset, which has been adopted from Zhonghan et al. [115].

The proposed VIO measurement update process is illustrated in the following equations:

$$K_k = \hat{\Sigma}_{k|k-1} H_k^T (H_k \hat{\Sigma}_{k|k-1} H_k^T + \Pi)^{-1} \quad \mathbf{2-10}$$

$$x_k = x_{k|k-1} + K_k(z_k - H_k x_{k|k-1}) \quad 2-11$$

$$\Sigma_k = (1 - K_k H_k) \Sigma_{k|k-1} \quad 2-12$$

where K_k represents Kalman gain, Σ_k and $\hat{\Sigma}_{k|k-1}$ represent the measurement covariance matrix at k and $k - 1$ epoch.

2.3.3.2 GRU-aided VIO

The VIO sub-filter uses the proposed ESKF-based tightly coupled integration strategy with a GRU model, which works during GNSS outages. The GRU consists of an update gate z_t that controls the extent of the impact on the current state force by the previous state and a reset gate to determine the forgetfulness degree of the hidden state information h_t . The details of the GRU propagation formula and architecture are adapted from Geragersian et al. [10]. The formula for the GRU architecture is presented below:

$$z_t = \sigma(W^z X_t + U^z h_{t-1}) \quad 2-13$$

$$r_t = \sigma(W^r X_t + U^r h_{t-1}) \quad 2-14$$

here, W^z and W^r are the weight matrix of input for hidden layers and reset gate respectively; U^z and U^r are the weight matrix of the hidden state for hidden layers and reset gate, respectively; σ is the activation matrix.

The GRU model is trained with multiple trajectories containing complex scenarios that facilitate failure modes extracted using FMEA analysis so that it can predict VO error during flight under diverse conditions.

The GRU output equation is formulated by:

$$P_{VO/GRU} = P_{VO} - \Delta P_{VO} \quad 2-15$$

where P_{VO} denotes position error from VO, ΔP_{VO} is position error deviations, and $P_{VO/GRU}$ represents the predicted position increments.

The GRU model is shown in Figure 2-5.

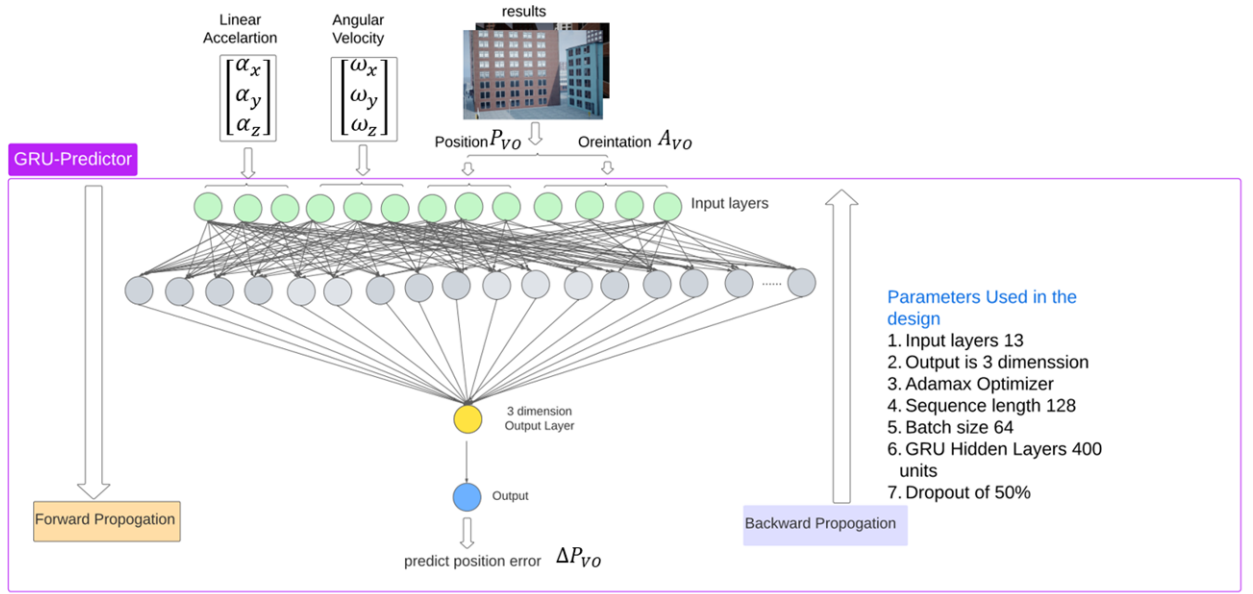


Figure 2-5 Illustrative diagram of GRU model.

Two data sources from IMU and VO are generated and used to gather positioning and attitude data for training each trajectory. Additionally, the VO position P_{VO} and orientation φ_{VO} covering multiple complex environments by UAV and IMU measured angular velocity α_{INS} and linear acceleration ω_{INS} are used to calculate inputs of GRU. The output of the GRU model is the positioning error ΔP_{VO} generated by VO. When the GNSS signal is unavailable, IMU/VO operates to estimate the position of the UAV where the GRU module operates in predicting mode that predicts the position error ΔP_{VO} , which is to be updated to the measurements vector in the ESKF module. When VO diverges, the GRU block predicts visual errors for error correction.

2.3.4 EKF Based GNSS/IMU Integration (Sub-Filter 2)

The tightly coupled architecture is implemented in the GNSS/MU integrated sub-filter of the proposed fault-tolerant multi-sensor navigation system. The GPS measurement position and IMU measurement acceleration and angular velocity proceed to estimate the state vector, including position velocity and attitude, using traditional EKF-based fusion filtering. Optimal states of the state vector from traditional EKF can be obtained through prediction and observation update and

is discussed in Mitchel et al. [86]. The generic observation equation for EKF can be written as:

$$z_{k_{GNSS/IMU}} = H_{k_{GNSS/IMU}} X_{k_{GNSS/IMU}} + \Lambda_{GNSS/IMU} \quad \mathbf{2-16}$$

where, $H_{k_{GNSS/IMU}}$ represents the observation matrix; $z_{k_{GNSS/IMU}}$ represents the observation vector; $X_{k_{GNSS/IMU}}$ represents the observation state vector, and $\Lambda_{GNSS/IMU}$ represents the observation noise matrix.

2.3.5 Federated Filter for Multi-Sensor Fusion

The proposed federated GNSS/IMU/VO multi-sensor navigation system uses VIO and GNSS/IMU integrated systems as the switching criterion. However, the global filter integration was conducted using the EKF approach to fuse the data generated by sub-filters. The detailed description and the state equations are the same as in the GNSS/INS sub-filter. In order to reduce computational complexity, the state equation of the GNSS/IMU sub-filter is the same as the global filter. The fusion resolution of the federated filter is as follows:

$$P_{k|global} = \sum_{i=1}^r P_{il} \quad \mathbf{2-17}$$

$$X_{k|global} = P_{k|global}^{-1} \left(\sum_{i=1}^r P_{il} X_{il} \right) \quad \mathbf{2-18}$$

where r is the number of sub-filters; $P_{k|global}$ and P_{il} are the covariance vector of the i th sub-filter and the global filter; and $X_{k|global}$ and X_{il} are the estimated states of i th sub-filter and the global filter. The global state estimation and the covariance vector are obtained by fusing the sub-filter estimated position, thus yielding a global solution.

The pseudo-code of the proposed fault-tolerant federated multi-sensor navigation system algorithm is presented in A 2-1.

A 2-1: Algorithm of GNSS/IMU/VO Multi-Sensor Navigation System

Input: P_{VO} , α_{xyz} , ω_{xyz} , P_{GNSS} // P_{VO} denotes the VO estimated position; α_{xyz} denotes the linear acceleration; ω_{xyz} denotes the angular velocity; and P_{GNSS} denotes the GNSS estimated position

Output: P_{VIO} , $P_{GNSS/IMU}$, $P_{GNSS/IMU/VO}$ // P_{VIO} denotes the sub-filter 1 estimated position; $P_{GNSS/IMU}$ denotes the sub-filter 2 estimated position; and $P_{GNSS/IMU/VO}$ denotes the global filter estimated position

Initialize:

1. Initial the values for VIO and GNSS/IMU sub-filters: $x_k = x_0$;// x_0 denotes ground truth from UAV.
2. Initial values for global filter EKF: $x_{k|global} = x_0$;// x_0 denotes the ground truth from UAV.

for $n = 1, 2, 3, \dots$ **do**

// n denotes number of iterations

Prediction Phase:

3. The sub-filter time update model has been derived in the equation 2-2, which is similar for both sub-filters and global, as using the EKF variant.
4. Covariance vector is defined as, $\Sigma_{k|k-1} = F_{k|k-1}\Sigma_{k-1}F_{k|k-1}^T + L_{k|k-1}Q_{k-1}L_{k|k-1}^T$.

Measurement Phase for sub-filters:

5. Generate the $P_{VO/GRU}$ predicted increments and updated the following measurement steps of the VIO sub-filter.
 6. Estimate the states using the updated equation (2-4)-(2-9)
 7. Calculate the Kalman gain, updated VIO states, and the covariance matrix using equations (2-10)-(2-12)
-

8. Corrected $X_{K_{VIO}} = x_{k|k-1_{VIO}} + K_{k_{VIO}} (z_{k_{VIO}} - H_{k_{VIO}} x_{k|k-1_{VIO}})$.
9. Start from step 7 to estimate the Kalman gain using Equation 10 for the GNSS/MU integration.
10. Update the observation equation 2-16.
11. Estimate the updated states and covariance matrix of the GNSS/IMU filter using equations 2-11 and 2-12.
12. The proposed VIO integrated sub-filter and GNSS/INS integration using EKF are running in parallel to estimate the position for each filter.

Measurement update for Global filter:

13. Calculate the Kalman gain:

$$K_{global} = \hat{\Sigma}_{global} H_{k_{global}}^T \left(H_{k_{global}} \hat{\Sigma}_{k|k-1_{global}} H_{k_{global}}^T + \Pi_{global} \right)^{-1}$$

14. Update the states using the traditional EKF equation 2-16.
15. Obtain $P_{k|global}$ and $X_{k|global}$ using equation 2-17 and 2-18.

End for

2.4 Experimental Setup

To verify the proposed fault-tolerant navigation system performance under complex environments, a GNSS/IMU/VO simulator is built on Unreal Engine with UAV unity dynamic models integrated into urban scenarios in MATLAB 2022a. The sensors implemented in the simulator include an IMU block, GNSS receivers, and a front-facing monocular camera model generated by the Navigation Toolbox and UAV Toolbox. The choice of a monocular camera is beneficial in my application due to its advantage of being less expensive, simpler to implement compared to a stereo camera, and lightweight to fit into a drone. The complex simulated environment is generated using Unreal Engine 5.0.3, utilising the default 'US city' urban scenario provided in the 'Simulation 3D Scene Configuration' block. The scenario was selected based on the chapter's

objectives, as it enables the validation of the proposed solution against visual failure modes arising from various fault factors present in an urban environment. The simulated scenario used in the experiments is a bright sunny day with 20% floating clouds. The dataset is acquired with these sensors mounted on a quadrotor, as shown in Figure 2-6.

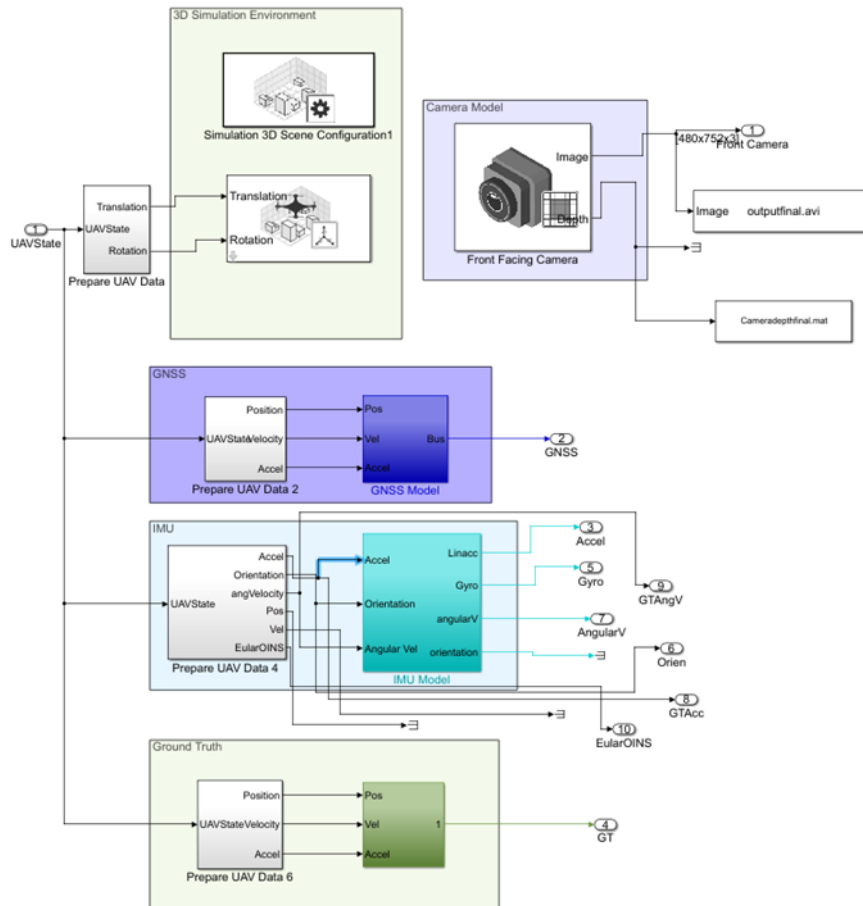


Figure 2-6 Experimental simulation setup in MATLAB.

The simulator consists of four blocks. The first block is the 3D simulation environment block, which is designed to simulate the 3D urban environment using ‘US city’ scenario available in MATLAB. This block integrates camera and UAV-based quadrotor models. Notably, the simulated scenario does not correspond to any specific real-world city but rather represents a generic urban layout, allowing for consistent testing across different scenes. The second block is GNSS integration with a quadrotor consisting of GNSS and the quadrotor dynamics. The third block is the IMU block, interfaced with the quadrotor block. The fourth block

is the ground truth from the quadrotor dynamics to provide true quadrotor trajectories.

The IMU selects the ICM 20649 model with the specifications provided in Table 2-2. The experimental data are collected with sampling rates of 10 Hz, 100 Hz, and 100 Hz for the camera, IMU, and GNSS, respectively. The random walk error $[0.001, 0.001, 0.001] \text{ m/s}^2$ in the IMU accelerometer results in a position error growing at a quadratic rate. Varying sample rates among the sensors influence the accuracy and temporal alignment of the data fusion. In particular, the higher frequency IMU and GNSS provided detailed motion data and frequent position updates, while the camera offered less frequent but complementary measurements. To address this mismatch and enable synchronisation, linear interpolation is applied to downsample IMU and GNSS data and maintain a unified rate of 10 Hz. This ensures consistent temporal alignment across all the sensors, reduces fusion errors and enhances reliable navigation performance even in dynamic conditions.

Table 2-2 Specification of the ICM 20649 IMU model.

Parameter	Value
Accelerometer Bias Stability	0.014 m/s^2
Gyroscope Bias Stability	0.0025 rad/s
Angle Random Walk	$0.00043633 \text{ m/s}^2 * \sqrt{\text{Hz}}$
Velocity Random Walk	$0.0012356 \text{ m/s}^2 * \sqrt{\text{Hz}}$

The GNSS model is initialized by injecting two common failure modes of random walk and step error that will most likely occur in an urban environment, leading to a multipath effect.

The camera model has specifications, including a 1109 focal length, 640×360 optical center, and 720×1280 image size. Regarding the calibration of the camera and extraction of extrinsic and intrinsic parameters of the simulated front-facing camera, the coordinate conversion matrix from world coordination to pixel coordination is denoted by camera instincts matrix k :

$$k = \begin{bmatrix} 1109 & 0 & 640 \\ 0 & 1109 & 360 \\ 0 & 0 & 0 \end{bmatrix}$$

For urban operation scenarios surrounded by buildings, the visual data of tall buildings are captured by a camera for VIO to provide positioning information. Meanwhile, the satellite availability is obstructed by buildings, causing a GNSS outage. The MATLAB simulator connects to the QGroundControl software to generate real-time flight trajectories for the data collection and save it into text file format. The QGroundControl uses a MAVLink communicator to connect the base station of the UAV block in the Simulator [116]. The integration of MAVLink with MATLAB/Simulink is adopted into the UAV package delivery example.

Regarding the training GRU models, 10 trajectories covering more than 100,000 samples from each sensor are used for training. The sensor blocks, including IMU, the GPS provided by the UAV Toolbox, and the Navigation Toolbox, operate in the local frame. To ensure compatibility and effective data fusion using the proposed algorithm, a crucial step in the data pre-processing phase involves converting the data from the local frame to the sensor body frame. This transformation is essential for aligning the sensor data with the algorithm's requirements and the system's operational frame reference.

The general performance evaluation method uses root mean square error (RMSE) formulated by:

$$RMSE_{overall} = \sum_{i=1}^N \sqrt{\frac{((P_{xPre} - P_{xGT})^2 + (P_{yPre} - P_{yGT})^2 + (P_{zPre} - P_{zGT})^2)}{N}} \quad 2-19$$

here, $P_{xPre}, P_{yPre}, P_{zPre}$, are the predicted position generated by proposed algorithm in x-, y- and z-axis, respectively; $P_{xGT}, P_{yGT}, P_{zGT}$ are the ground truth generated from UAV in x-, y- and z-axis, respectively. The number of samples is represented using capital N.

2.5 Test and Results

In order to evaluate the performance of the proposed GRU-aided ESKF-based VIO, two trajectories corresponding to experiments 1 and 2 are selected from the package delivery experiment in an urban environment. Both experiments are carried out under sunlight conditions, introducing common fault scenario

shadows, lighting variations, motion blur, no-texture, and motion variation consistently present throughout the flight duration. For both experiments, a consistent fault condition is injected, i.e., the shadow of tall buildings on a sunny day during the flight. According to the fault types and number, the flying regions in the experiments are categorized into four distinct zones that encompass single faults, multiple faults, and combined faults. The faults arising from environmental sensitivity and dynamic motion transitions under previously estimated two major failure events in visual systems using FMEA analysis results in Section 2.3.1 have been discussed.

2.5.1 Experiment 1—Dense Urban Cynon

The purpose of this experiment is to validate the effectiveness of the proposed GRU-aided VIO in managing specific failure modes in complex conditions. This experimental environment includes a combination of tall and short buildings. During the experiment, a distance of 235 m was covered within a time span of 55 s. Figure 2-7 Overall visual navigation error in the presence of multiple failure modes for Experiment 1. shows the accumulated 3D visual position error. The proposed GRU-aided VIO is able to reduce position error by 86.6% compared with an ESKF-based VIO reference system. The maximum error in Figure 2-7 reduces from 7.5 m to 1.9 m with the GRU adoption. The significant drop in error for the proposed approach is primarily due to its ability to learn and compensate for error patterns arising due to complex failure modes. By leveraging temporal dependencies from IMU input and VO position, the GRU-aided ESKF VIO can predict and correct for drift and failures arising from feature extraction errors, feature tracking errors and feature location errors.

When the UAV takes off with rapid and sudden changes in waypoints, the Dynamic Motion Transitions failure mode occurs under the condition of jerking movements that consistently result in feature tracking errors and feature association error failure modes in VO. Moreover, the other dynamic motion transition failure mode is incorporated, i.e., turning to replicate complex environmental conditions, thereby emulating real flight scenarios. The UAV takes off speed sets to 20 m/s while its speed will increase up to 50 m/s at 7 s.

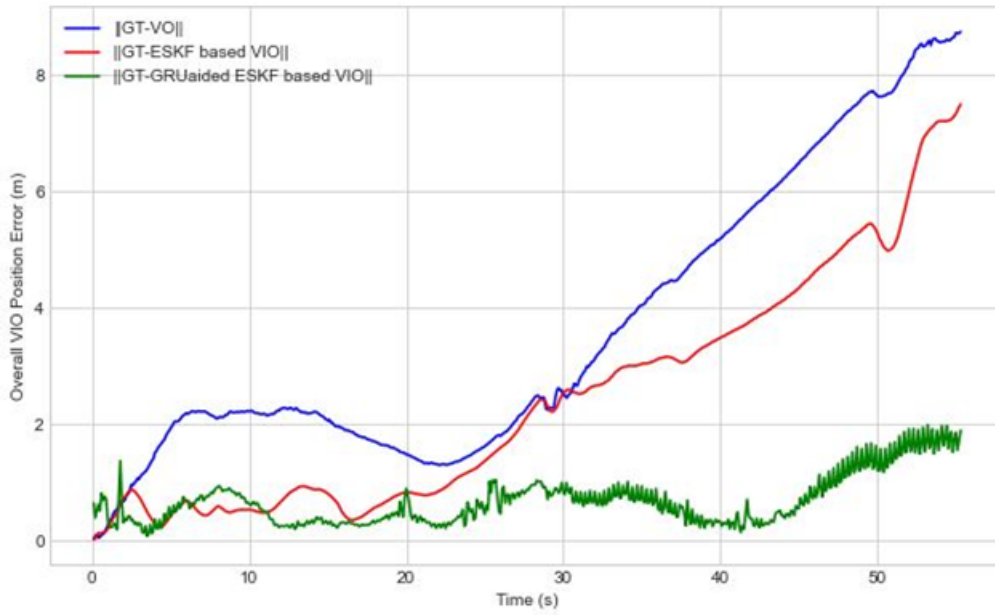


Figure 2-7 Overall visual navigation error in the presence of multiple failure modes for Experiment 1.

Figure 2-8 shows the RMSE position error of the proposed GRU-aided VIO system and benchmarked with two references of VO- and ESK-based VIO position error. The RMSE position error in the x-axis under NED coordination is relatively lower compared to y- and z-axis. It is worth noting that the maximum position error of the ESKF-based VIO reference system is 3.227 m, 5.6 m, and 4.1 m in the x, y, and z-axes.

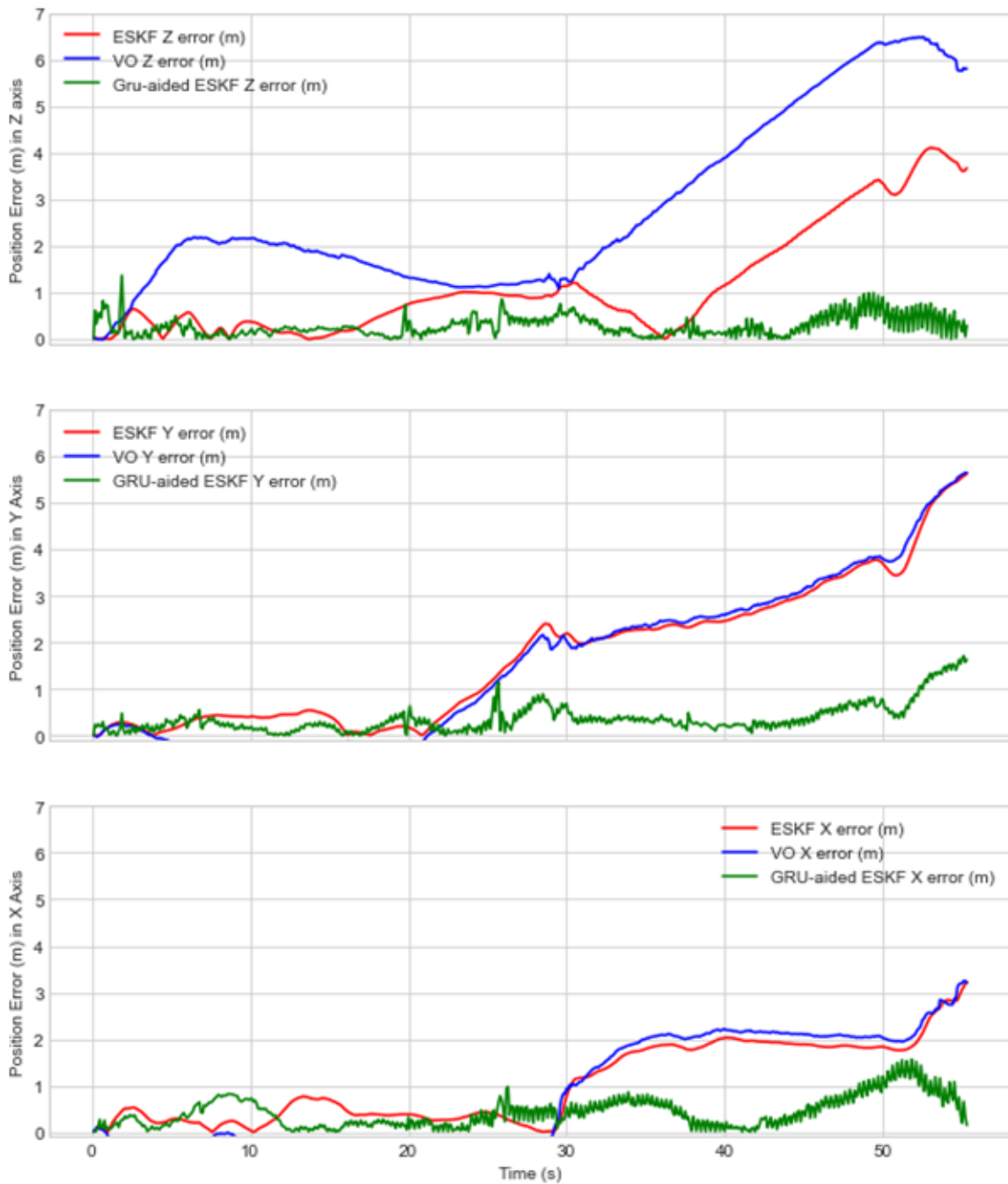


Figure 2-8 Visual navigation error along each axis in the presence of multiple failure modes for Experiment 1.

In the y-axis, the position error axis increases at 22 s due to the shadow of another building creating a variation of light. At 27 s, the UAV encounters a turn facing a plain wall because of a lack of textures, leading to drift, inaccuracies and failure of visual odometry. ESKF-based VIO showed relatively poor performance along the y-axis during diagonal motion due to cross-axis coupling and multiple

failure modes due to featureless plain wall, sunlight variation, and shadow of tall buildings, leading to feature degradation and tracking features. The loss of visual features results in insufficient information for ESKF to estimate the position accurately. In Figure 2-7 and Figure 2-8, it is shown that VIO based on ESKF fails to mitigate visual positioning error due to non-linear motion, lack of observable features, non-Gaussian noise, and uncertain state estimation, leading to non-linearization error propagation. However, it was found that the proposed solution can mitigate position error by 60.44%, 78.13% and 77.13% in the x-, y- and z-axes, respectively. The maximum position error is decreased from 1.5 m, 1.6 m, and 1.2 m.

When analysing the z-axis performance in Figure 2-8, at around 37 s, the ESKF-based VIO position error starts increasing due to a dark wall shadowed from another tall building, causing variation of light in the frame leading to feature association error. ESKF performance degrades due to multiple factors present in the scenario, and error accumulates over time. After applying the proposed GRU-aided ESKF, the VIO fusion method is able to reduce position error at 37 s from 2.1 m to 0.6 m by predicting the VO error indicated with details in Table 2-3.

Table 2-3 RMSE Comparison on the performance of two experiments.

Experiment	Method	RMSE(m)-x-axis	RMSE(m)-y-axis	RMSE(m)-z-axis	RMSE(m)-Overall
1	VO	1.4	2.3	3.4	4.3
	ESKF VIO	1.3	2.2	1.6	3.1
	GRU-aided ESKF VIO	0.5	0.4	0.3	0.7
2	VO	6.6	2.9	10.4	12.6
	ESKF VIO	4.7	2.5	8.6	10.1
	GRU-aided ESKF VIO	0.8	0.9	0.5	1.3

2.5.2 Experiment 2—Semi-Structured Urban Environment

This experiment aims to measure GRU-aided ESKF VIO performance under environments of tall buildings with open space parking areas. In this experiment, the UAV encounters two turns, which means changing two waypoints at 60 s and 110 s, leading to motion variation causing feature association error described as ‘Dynamic Motion Transitions’.

Figure 2-9 shows the accumulated 3D position error of the proposed GRU-aided ESKF VIO system and is benchmarked with two references of VO- and ESKF-based VIO position error. In Figure 2-9, the position error in the first few seconds is negative due to multiple factors, such as shadows of tall buildings, trees on plain surfaces, shadowed buildings, and lighting variations due to sunlight, motion blur, and rapid motion. The maximum position error is 19.1 m at 110 s due to combinational failure modes such as dark wall, rapid motion and motion blur. The proposed GRU-aided ESKF VIO is able to mitigate 86.62% of overall position error. The maximum error is reduced from 19.1 m to 6.8 m.

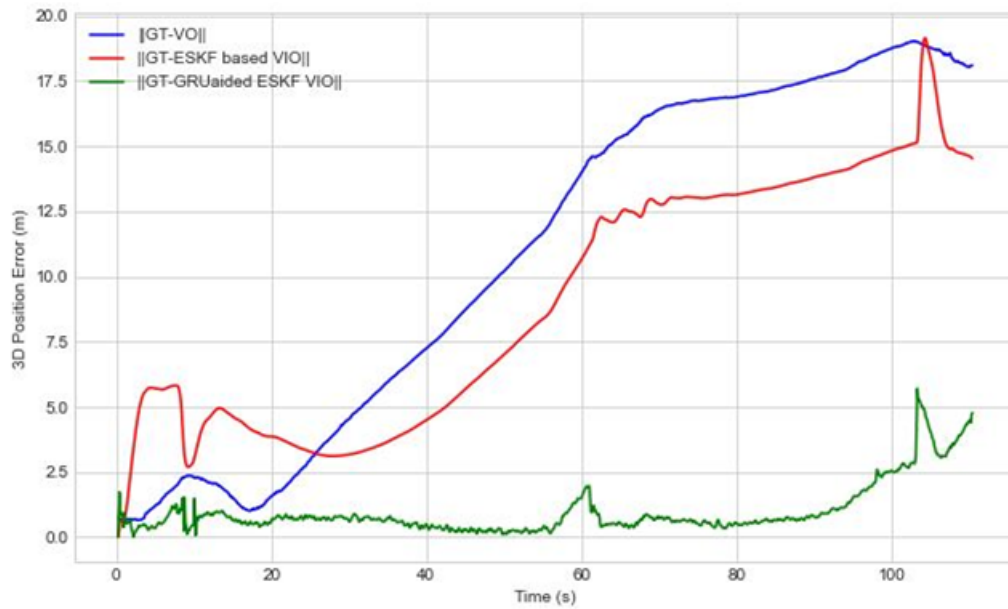


Figure 2-9 The 3D-position error in the presence of multiple failure modes for Experiment 2.

Figure 2-10 shows the VIO position RMSE error in the separate x-, y- and z-axes. Table 3 indicates that position error RMSE of ESKF is 4.7 m, 2.5 m, and 8.6 m in the x-, y- and z-axes, respectively. The proposed GRU-aided ESKF VIO has a remarkable improvement in terms of position error, and the specific values along the x-, y- and z-axes are 0.8 m, 0.9 m, and 0.5 m, respectively. Notably, the reference systems exhibit a significant increase in errors along the z-axis and x-axis, primarily due to shadows induced failures in feature extraction and feature tracking. These lead to increased drift in VO, which ESKF-VIO is unable to

effectively compensate for. Due to cross-axis coupling, the y-axis faces a larger estimated position error than others. During the time interval of 57–65 s, UAV takes a turn and passes through a parking area where buildings are under limited field of view. In this case, the feature extraction and tracing processes encounter challenges and lead to position estimation errors. The proposed solution has shown excellent performance improvement in the presence of failure modes of feature extraction error and feature tracking error, where the position error is decreased by 62.86% in comparison to the reference systems.

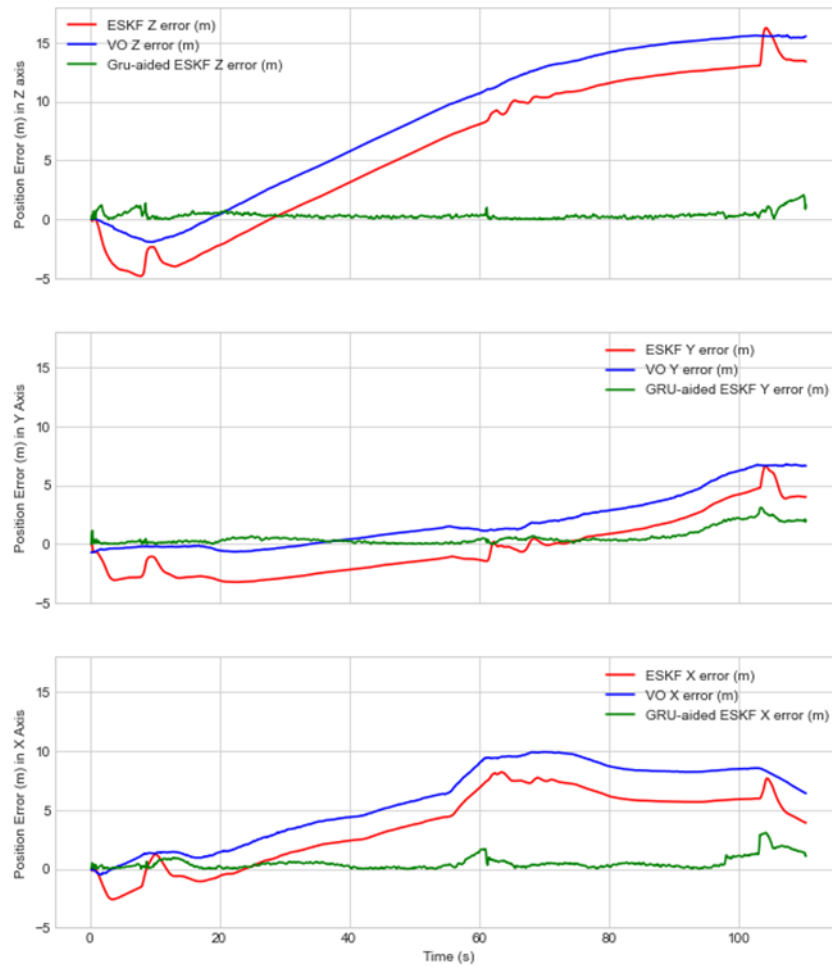


Figure 2-10 Position error along each axis in the presence of multiple failure modes for Experiment 2.

When analysing the z-axis performance in Figure 2-10, the proposed GRU-aided ESKF VIO outperforms the reference ESKF fusion with respect to a reduction in the z-axis error by 93.46%. According to Table 2-3, the reference ESKF has

shown the worst performance of 8.6 m RMSE compared to the proposed method of 0.5 m RMSE. It is noted that at 104 s during the landing phase, the UAV turns around and encounters a black wall. This leads to higher performance errors because the VO system struggles to extract enough features in the complex scene with poor lighting [6,8,9]. The GRU-aided ESKF VIO demonstrates improvement compared to the traditional ESKF-based approach, resulting in a remarkable reduction of error of 93.45%. The maximum position error in the z-axis due to dark scene at 110 s is reduced from 16.2 m to 3.0 m. During the experiment, a distance of 459 m was covered within a time span of 1 min 50 s.

Table 2-4 indicates the maximum position error comparison for two of the experiments. By integrating fault-tolerant mechanisms, the approach achieves more accurate position estimation, even in challenging situations with limited visual cues. The fault-tolerant GRU-aided ESKF VIO architecture shows robustness over a number of realistic visual degradation scenarios.

Table 2-4 Maximum position error comparison for two experiments.

Experiment	Method	Maximum Error in x-axis (m)-	Maximum Error in y-axis (m)-	Maximum Error in z-axis (m)-	Maximum Error (m)-	3D
1	ESKF VIO	3.2	5.6	4.1	7.5	
	GRU-aided ESKF VIO	1.5	1.6	1.2	1.9	
2	ESKF VIO	8.2	6.6	16.2	19.1	
	GRU-aided ESKF VIO	5.0	4.4	3.0	6.8	

2.5.3 Performance Evaluation Based on Zone Categories

To further evaluate the successful rate when mitigating failure modes from experiments 1 and 2, as detailed in Section 2.3.1, the fault zones are extracted in the above two experiments.

Zone 1 indicates building shadow as the single fault triggering feature matching and feature tracking error failure modes within the time interval. In experiment 1, between the time interval of 33–49 s, UAV passes through shadow buildings that distort visual features, causing incorrect matches and tracking errors when they

move into or out of shadows, as shown in Figure 2-11. In addition, introducing the sudden change in lighting may be misinterpreted as IMU acceleration and rotation.

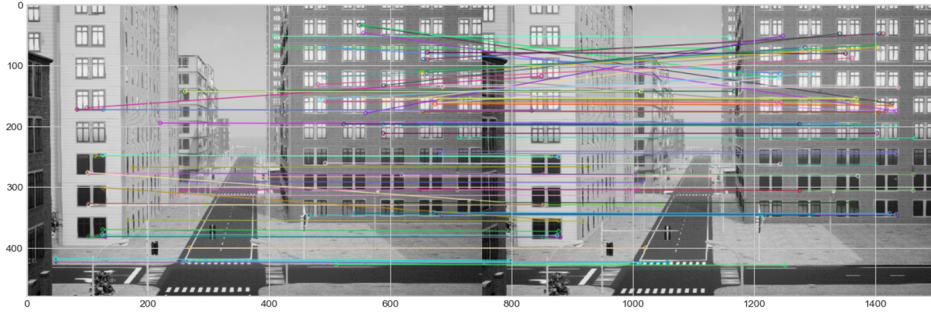


Figure 2-11 Feature tracking and Feature mismatch error failure modes due to tall-shadowed buildings in experiment 1.

Figure 2-12 a,b depict the zone 1 region to show performance comparisons of the proposed algorithm with the reference algorithms in the presence of two failure modes. The maximum position errors for experiment 1 in zone 1 along x-, y- and z-axes are reduced by 52.38%, 81.57%, and 73.17%, respectively. In experiment 2, a single fault is encountered twice during the time interval of 13–44 s and 70–106 s. Maximum position error in the time interval of 13–44 s in x-, y- and z-axes are reduced by 93.33%, 75%, and 85%, respectively. Hence, the proposed solution proves to be robust over two failure modes.

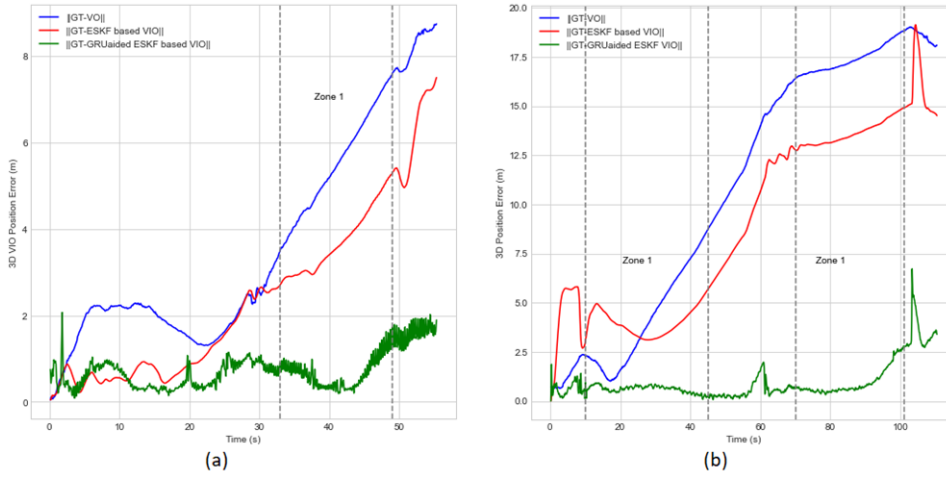


Figure 2-12 Position errors of Zone 1 (shadowed building) estimated in each axis for experiment 1 (a). Position errors of Zone 1 (shadowed building) estimated in each axis for experiment 2 (b).

Zone 2 includes multiple faults, including turning manoeuvre and shadow of tall buildings that are present in both experiments. When a UAV makes a turn, the motion dynamics change rapidly. This leads to challenges in estimating camera motion and orientation estimations, causing tracking errors. In the meantime, visual distortion also causes feature extraction errors and feature mismatch errors due to inconsistent lighting, as shown in Figure 2-13. The combination of both conditions adds complexity to the environment, exacerbating the existing challenges in traditional ESKF-based VIO. The proposed algorithm is able to mitigate these failure modes and shows robustness in such complex scenes compared with traditional VIO systems. The algorithm is able to mitigate motion dynamics and feature extraction error, reducing feature matching error by 20%, 20%, and 50% at the time interval of 13–24 s in experiment 1 and 62.5%, 40%, and 90% at the time interval of 45–70 s with respect to the x-, y- and z-axes, as shown in Figure 2-14.

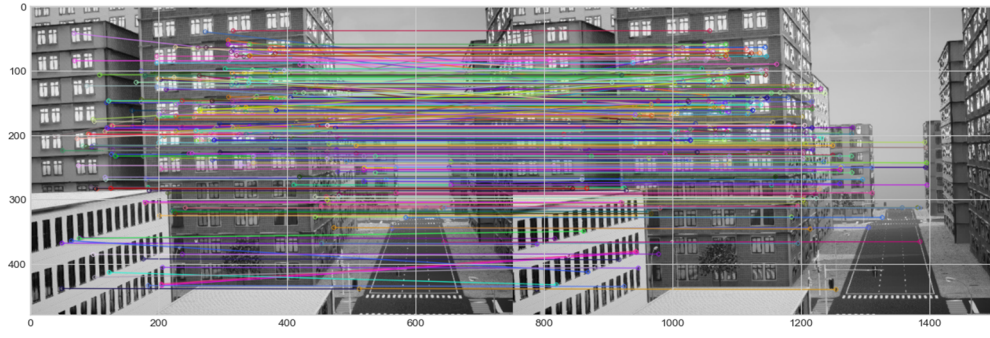


Figure 2-13 Motion dynamics, feature tracking and feature mismatch error failure modes due to the tall, shadowed buildings in experiment 2.

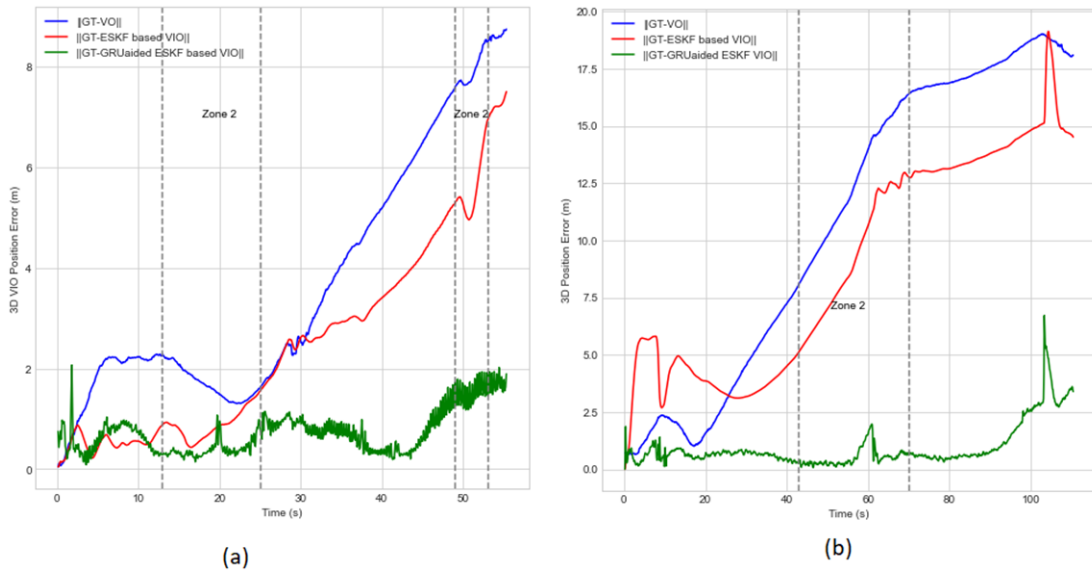


Figure 2-14 Position errors of Zone 2 (shadowed building and UAV turns) estimated in each axis for experiment 1 (a). Position errors of Zone 2 (shadowed building and UAV turn) estimated in each axis for experiment 2 (b).

In Zone 3, multiple faults are combined together, including the turning manoeuvre, shadows from the tall buildings, variations in lighting, areas of darkness and sunlight shadows. Zone 3 only exists as one of the most complex conditions in experiment 2. As observed in Zone 1 and Zone 2, the turning behaviour and shadow of the tall buildings introduce changes in motion dynamics that make the position estimation and feature tracking challenging for traditional, ESKF-based VIO. Additionally, the presence of both dark and well-lit areas within the scene created abrupt changes in illumination.

Figure 2-14Figure 2-15 presents one demonstration image from the mounted front-facing camera in a UAV when passing through an illuminated and shaded area. The shadows caused by direct sunlight also create sharp pixel contrast between illuminated and shaded areas. These sudden lighting changes and the combination of multiple fault conditions amplify the challenges posed by each individual fault, making the overall VIO performance more susceptible to tracking errors, feature mismatches, and feature extraction error failure modes.

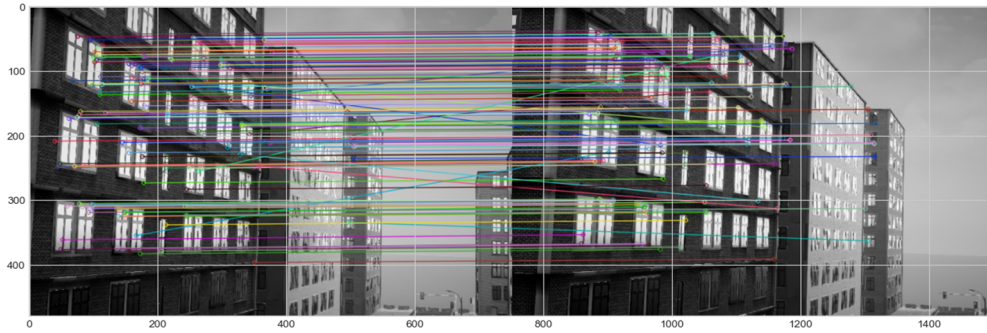


Figure 2-15 Feature tracking, feature extraction error, and feature mismatch failure modes for dark and well-illuminated tall buildings in experiment 2.

Figure 2-16 shows that the GRU-aided ESKF VIO architecture reduces maximum position error in experiment 2 at the time interval of 107–114 s from 32.6%, 81.327%, and 64.397% in x -, y - and z -axes. Therefore, the GRU-aided fusion algorithm can perform without interruption when the UAV navigates in illuminated and shaded areas and has shown robustness in the presence of multiple failure modes and moving features amidst dynamic lighting.

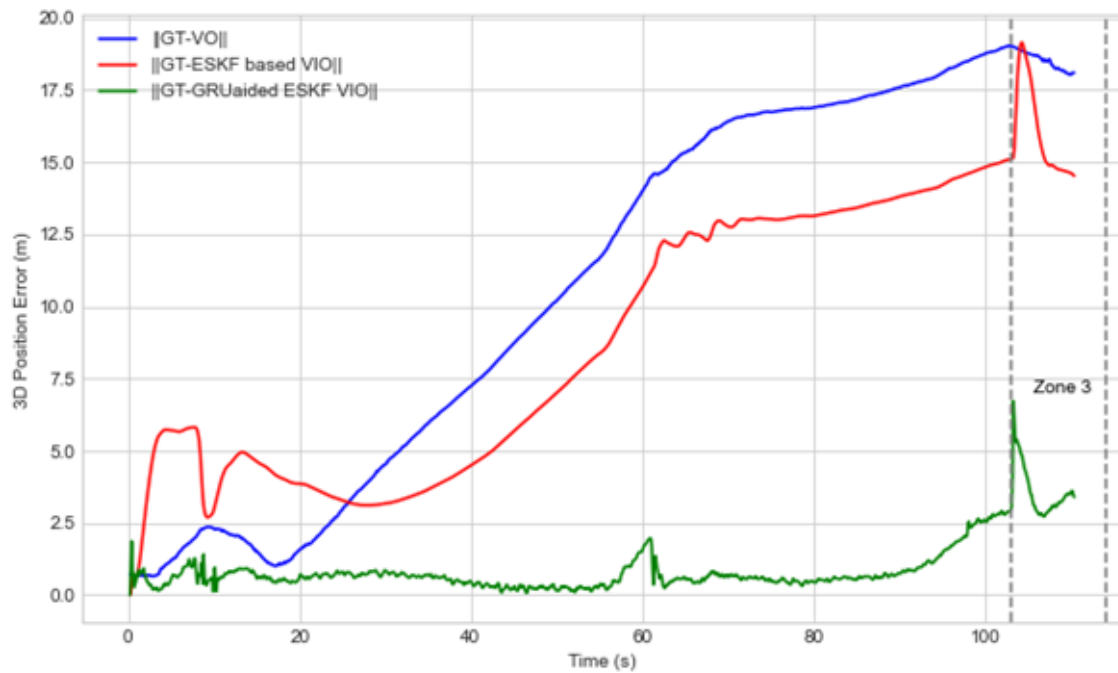


Figure 2-16 Position errors in Zone 3 (UAV turn and illuminated tall buildings) estimated in each axis for experiment 2.

Zone 4 consists of a combination of complex faults, including navigation environmental error and data association fault events. The fault events consist of turning manoeuvres, building shadows, the presence of featureless blank walls and variation in lighting. In experiment 1, the UAV encountered a plain wall at 27–32 s of its flight, resulting in a feature extraction error due to the lack of distinctive features on the wall shown in Figure 2-17. As a result, the feature extraction process failed, leading to a lack of identifiable features to track and match in consecutive frames. Such lack of features caused the VO to lose its frame-to-frame correspondence, resulting in the inability to accurately estimate the UAV's motion in this specific time of 27–32 s. Figure 2-18 shows the increment of position error caused by the mentioned disruption. The ESKF algorithm performance is heavily affected, leading to incremental tracking errors and loss of tracking when dealing with a featureless wall. Figure 2-18 shows that the algorithm has effectively reduced the maximum position error by 42.1%, 63.3%, and 60.12% in x-, y- and z-axes, respectively.

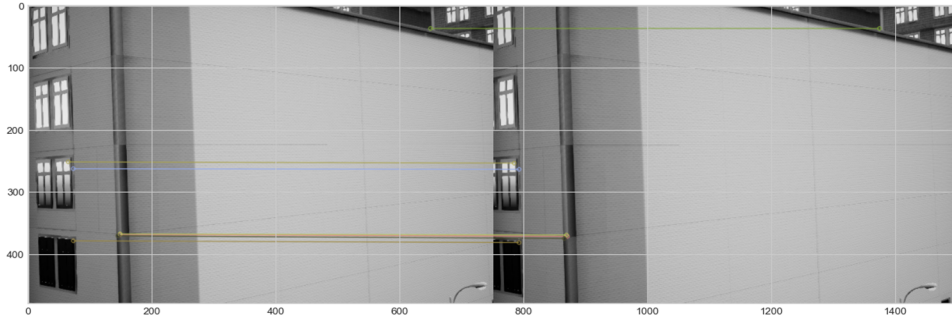


Figure 2-17 Feature extraction error due to the plain wall in experiment 1.

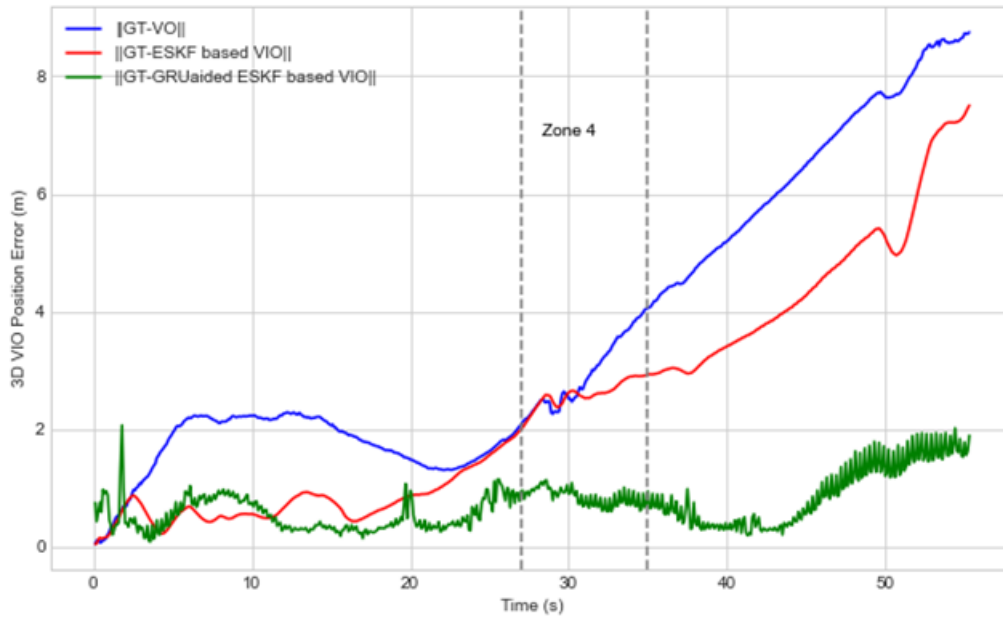


Figure 2-18 Position errors of Zone 4 (UAV turn, shadowed buildings, and featureless wall buildings) estimated in each axis for experiment 1.

To evaluate the performance of the proposed fault-tolerant federated multi-sensor navigation system, the experiment is conducted using the Experiment 1 dataset with GNSS condition applied (faulted-GNSS and without fault GNSS) adopted from [117]. Table 2-5 shows the performance comparison results of the proposed GRU-aided ESKF VIO with faulted GNSS and no-fault GNSS. The results indicate that the hybrid approach enables mitigating the overall position error even with faulted GNSS conditions compared to [39], [117]]. Therefore, it is approved that the FMEA-assisted fault-tolerant multi-sensor navigation system

facilitates positioning performance in the presence of multiple faults covering all the sensor faults in diverse, complex environments.

Table 2-5 The RMSE of 3D position errors of GNSS/faulted-IMU/faulted-VO solution with faulted-GNSS/faulted-IMU/faulted-VO.

Algorithms	3D RMSE Position Error (m)
Faulted-VO/GNSS/IMU	1.2
Faulted-VIO-ESKF/GNSS/IMU	0.7
Faulted-GRU-aided-ESKF-VIO/GNSS/IMU	0.09
Faulted-VO/Faulted-GNSS/IMU	1.5
Faulted-ESKF-VIO/Faulted-GNSS/IMU	1.0
Faulted-GRU-aided-ESKF-VIO/Faulted-GNSS/IMU	0.2

2.5.4 Performance Comparison with Other Datasets

This chapter selects the EUROC dataset, specifically MH05_difficult, for benchmarking with other algorithms since EUROC is commonly applied by other researchers [12,45]. The EUROC sequence dataset was collected in an indoor machine hall with light variation. MH05_difficult dataset contained black and white images, which they referred to as dark and shadowed environments and captured with rapid motion. The dataset is captured in a customized way that has several limitations, including manipulated images, customized blur, and brightness.

Figure 2-19 presents the magnitude of 3D position error of running my proposed GRU-aided ESKF VIO to process MH05_difficult. The key finding is that the position error in RMSE is reduced by 67.32%. The maximum error is reduced from 2.81 m to 1.5 m.

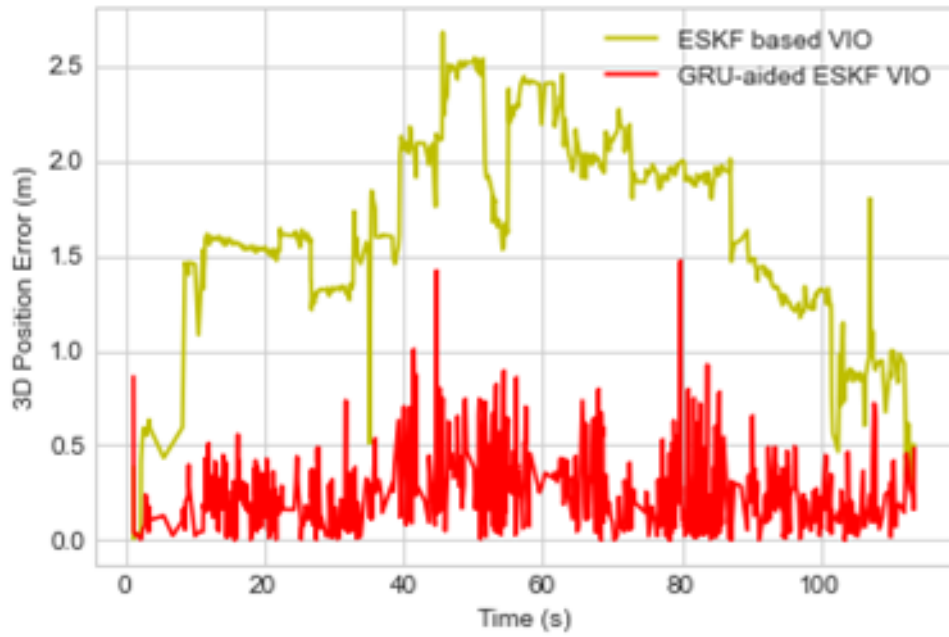


Figure 2-19 The 3D position error estimated using a MH05 seq EUROC Dataset with a motion blur failure mode.

Table 2-6 highlights the comparison results in terms of the accumulated RSME with state-of-the-art systems, i.e., End-to-End VIO, Self-supervised VIO [44], [45]. It is worth mentioning that the EUROC dataset does not provide tight synchronization between IMU and images, which is a primary requirement of using RNN-based VIO.

Table 2-6 Comparison with state-of-the-art methods that used MH_05 seq. EUROC Dataset with motion blur failure mode.

Algorithms	RMSE (m)-of MH_05 seq	Improvement
End-to-end VIO [45]	1.96 m	66.32%
Self-supervised VIO [12]	1.48 m	55.4%%
Proposed GRU-aided ESKF VIO	0.66 m	-----

By cross-checking with the work of Brandon et al. [12], the proposed GRU-aided ESKF VIO confirms the robust improvement in the presence of motion blur failure mode.

2.6 Conclusions

Aiming to provide fault-tolerant VIO navigation solutions against complex environments, this study proposed a hybrid federated navigation system framework aided by FMEA for enabling fault tolerance and GRU fused with ESKF-VIO to mitigate visual positioning errors.

Through simulations, the main advantages of the GRU and ESKF hybrid algorithm are summarized as follows: (1) A high-efficiency recurrent neural cell with simple architecture, namely the GRU, was chosen to predict the position error during visual degradation. Benefiting from the proper selection of the Kalman filter performance enhancement method, such as updating the state vector by predicting errors using the AI method, the proposed algorithm possesses superior navigation accuracy under complex conditions. (2) The FMEA analysis helps to prioritize anticipated failure modes such as feature extraction error, feature tracking error, and motion dynamics, enabling us to mitigate position error caused by these failure modes before they lead to operation failure. (3) The mitigation of feature extraction failure modes, which can subsequently lead to feature association errors. Via demonstrations, it is found that multiple factors or faults within the navigation environment and the UAV's dynamics reduce the impact of those failures.

This approach represents a significant step towards improving the robustness and reliability of VIO, particularly in complex and dynamic environments where feature extraction error, feature tracking error, and feature mismatch are critical for accurate navigation. With the correction of the VIO, the fault-tolerant multi-sensor performance is demonstrated to be improved under diverse, complex urban environments in terms of robustness and accuracy at different time scales, enabling uninterrupted and seamless flight operations.

3 A Comparative Analysis of Hybrid Sensor Fusion Schemes for Visual–Inertial Navigation

3.1 Introduction

Visual Inertial Navigation Systems (VINS) have garnered significant interest among researchers as a prominent alternative solution due to GNSS vulnerability in urban environments. While fusion techniques such as filter-based VIO and data-driven VIO have generated reasonable results throughout their evaluation, they still prove to be fragile in environments with complexity characterised by blur [118], low texture [27], [44], [50], illumination variation [28], [53], [74], [119], [120] and weather conditions [53]. Besides, modelling the effects can be challenging, and the lack of feature tracking led to divergence in the optimisation process, which traditional fusion techniques may not eliminate. Moreover, classical VIO typically necessitate parameter tuning, vast datasets for training and calibration for effective operation. The rise in machine learning has enabled the direct processing of high-dimensional data and learning underlying models that can enhance state estimation. Nevertheless, the utilisation of data-driven VIO is prone to the fact that performance usually drops in scenarios that considerably differ from the training dataset [121], [122].

Hybrid integrations allow for improved handling of complex scenarios by evaluating the level of epistemic uncertainty, adjusting to changing sensor conditions and identifying fault patterns. Strategies such as enhancing the model's architecture, refining its parameters, and expanding training datasets are being reviewed for machine learning based VIO, with hybrid approaches offering potential solutions to minimise these requirements. Several hybrid (NN-KF) VIO methods have provided the advantages of using time-dependent recurrent NN models like Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU) and Bi-directional LSTM (Bi-LSTM) in VIO as state-predictors [44], [49] and error compensators [123], [124]. The authors of these studies have investigated VIO performance at different fault conditions arising while navigating complex environments, yet the generalization ability of such systems needs to be explored. Kim et. al [80] have reviewed hybrid NN-KF techniques in a broad range

of applications and categorised them according to the role of NN applied to KF, as follows- 1) prediction of tuning parameters using NN, 2) error compensation using NN, and 3) update observation model of KF using NN as illustrated in Figure 3-1.

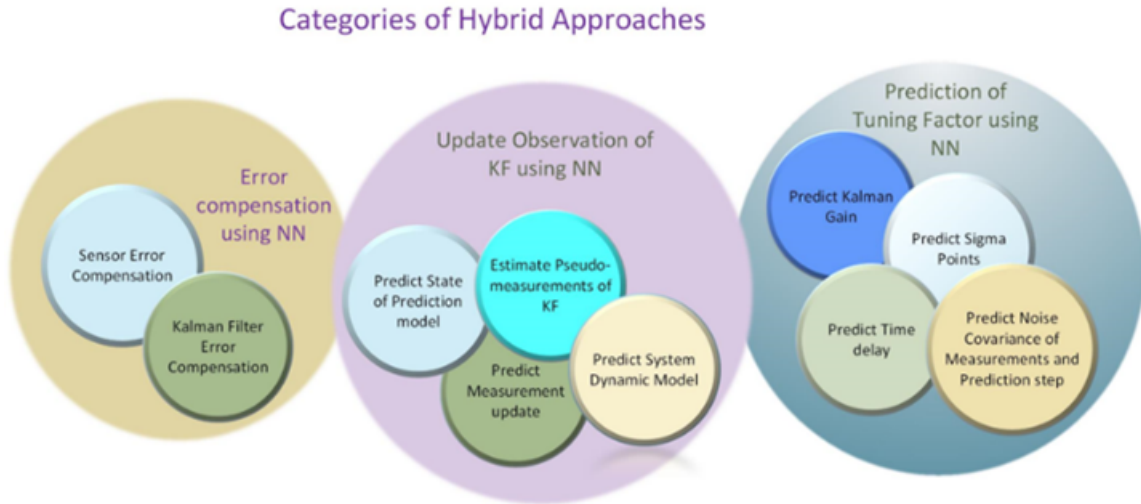


Figure 3-1 Categories of hybrid NN-KF techniques according to [80]

Authors of [16], and [17] have outlined recommendations concerning improving accuracy, robustness, and generalization ability for the systems utilizing hybrid NN-KF approaches to mitigate sensor error, imperfections in system models, and nonlinearity in dynamic conditions. However, to the best of the author's knowledge, studies on quantitative differences between various types of hybrid architectures do not provide any evidence on exploring performance differences and generalization ability in a systematic way that would allow robust conclusions on the pros and cons of each architecture. Therefore, this work proposes a systematic study of the performance differences in three types of hybrid NN-KF approaches implemented in VIO-based navigation systems for UAVs in complex urban scenarios. Three different hybrid approaches for VIO fusion are adopted based on mostly utilized methods [80] including VO error compensation using NN, KF error compensation using NN, and prediction of Kalman gain using NN. The models are designed as a combination of GRU and ESKF where GRU is utilized for different purposes for each hybrid approach to enhance the efficiency of the Kalman filter. In addition, the formulation is extended by providing

robustness analysis on these methods using MATLAB incorporated with Unreal Engine simulated photorealistic datasets replicating possible fault conditions present in urban environments, including weather effects. The main contributions of this work are highlighted as follows -

1. Comparison of mostly adopted NN-KF architectures within VIO systems to evaluate their performance across different environmental conditions such as fog, sun, and rain alongside various fault conditions like illumination changes, no-texture, reflections, motion variation and altitude variation.
2. Performance evaluation of each hybrid VIO algorithm's ability to mitigate various failure modes, including feature extraction errors, feature tracking errors, feature mismatch, feature location errors and data association errors.
3. Assessment of the generalisation ability of each algorithm to adapt and perform effectively in unseen scenarios, providing insights into their robustness and versatility.

The remainder of this chapter is structured as follows: section 3.2 includes three hybrid visual-inertial navigation architectures with brief descriptions. Section 3.3 discusses experiment setups for data generation to train and test the hybrid models. Section 3.4 discusses the test and performance analysis of the selected algorithms under diverse complex scenarios with comparison. Finally, the conclusion is presented in the section 3.5.

3.2 Hybrid Visual-Inertial Navigation Architectures

This section provides a brief overview of the hybrid architectures implemented in VINS systems. In line with the research objective of assessing the strengths and limitations of mostly utilized hybrid VIO systems in complex environments including the influence of weather effects, three hybrid VIO fusion architectures are as follows-

1. GRU-aided ESKF-VIO-VO error compensator (method 1)
2. GRU-aided ESKF-VIO-KF error compensator (method 2)

3. GRU-aided ESKF-VIO Kalman gain predictor (method 3)

These architectures are illustrated in Figure 3-2. The explored hybrid ESKF-based VIO fusion scheme is well-known for its ability to minimise linearization errors, mitigate singularity, handle sensor delays, and efficiently model errors. The GRU is implemented for augmentation of ESKF, chosen due to its effectiveness in processing time-varying sequences compared to other NN.

In this chapter, each hybrid fusion scheme integrates pose measurements generated by VO with linear acceleration from the accelerometer and angular velocity from the gyroscope. Feature-based VO is preferred for the research due to its simplicity and reliability, contrasting with data-driven VO, which requires more computational resources and extensive training data.

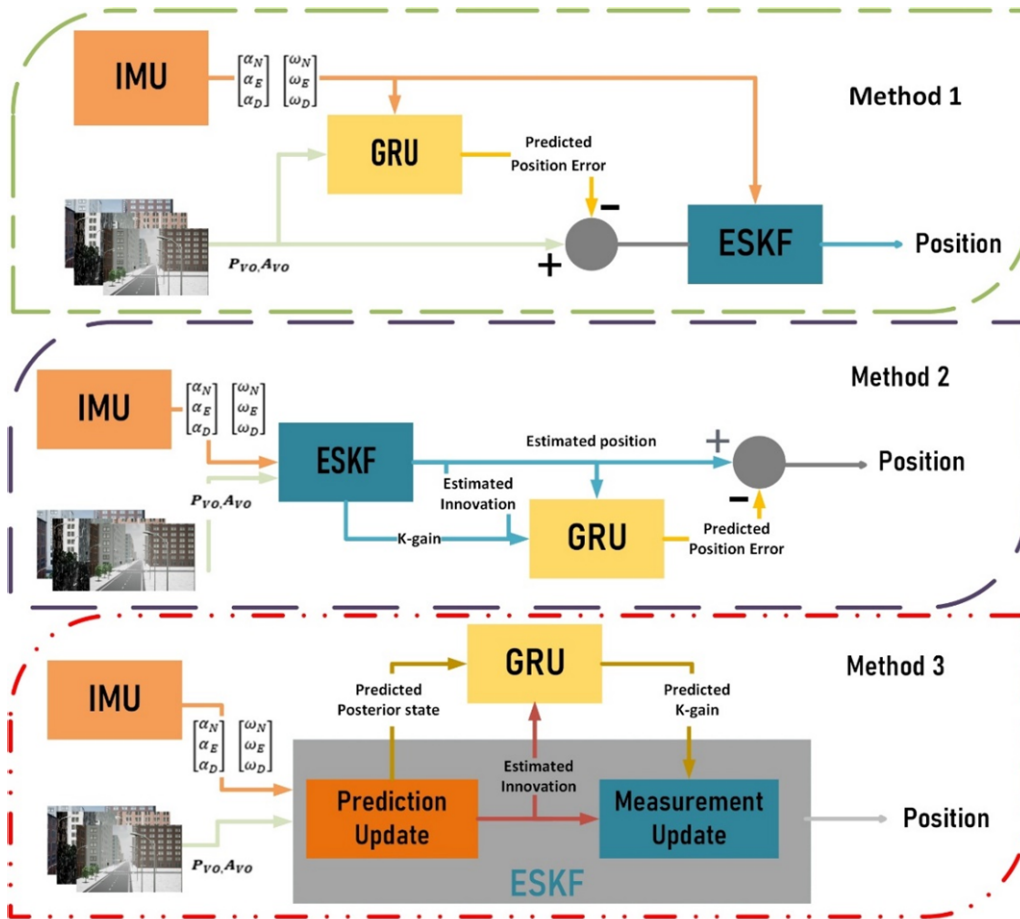


Figure 3-2 The high-level architecture of adopted techniques implemented with VIO systems.

The three hybrid fusion architectures each offer distinct theoretical advantages tailored to different types of sensors and model degradation. Method 1 focuses on learning and correcting VO drift and corrupted output before fusion, making it particularly effective in environmental conditions where visual degradation, like shadows, low availability of features, and weather effects, cause failure. Nonetheless, it is limited in scope compared to others; it does not account for induced noise and fusion properties. Method 2, which applies corrections to Kalman filter estimation, addresses fusion inconsistencies arising from unmodeled dynamics or predefined models, making it suitable for high-dynamic scenarios with complex motion. In contrast, Method 3 leverages the GRU to adaptively predict Kalman gain, allowing the filter to dynamically adjust confidence in sensor inputs. This approach is especially beneficial in environments where sensor reliability changes over time, such as during rapid lighting changes, occlusion or motion blur. Despite this, it poses a greater risk in terms of filter consistency and stability, as it alters the core filter structure and can diverge if not well-trained. Method 2 introduces potential instability and reduced interpretability and lacks the ability of Method 3 in managing time-varying output reliability. Detailed explanations about each method are provided in the next sections.

3.2.1 Method 1

The concept of compensating sensor error using NN is utilised to enhance the performance of KF in this integration. The visual sensor has prominent fault factors contributing to performance degradations in traditional KF-fusion techniques. Tabassum et. al [123] proposed tightly coupled GRU-based VO error compensation aided with ESKF to fuse corrected VO data with IMU as part of a fault-tolerant FMEA-supported federated multi-sensor navigation system. The fusion aims to enable uninterrupted operation of UAVS under various fault conditions, including illumination variation, motion variation, low-texture environment, and flight dynamics. The VO position is corrected by utilizing GRU-predicted VO increments before integrating with IMU data using ESKF fusion. Such architectures could be beneficial to correct imperfect sensors, parameters

or errors of the sensor caused by its characteristics (VO drift for longer flights). The presence of outliers, which features tracking errors and illumination change, mostly occurs when weather changes, and such a design could improve system robustness under complex scenarios.

3.2.2 Method 2

Compensating the errors in the result of KF using errors predicted by GRU is adopted in this hybrid fusion architecture. This fusion scheme is beneficial to correct modelled errors estimated by KF by predicting unmodelled error using NN. Villalonga Torres et al. [124] have proposed Bi-LSTM-based error correction mechanisms to correct the measurements of multiple EKF filters as a modification in federated architecture for an alternative to the GNSS multi-sensor navigation system. The results indicated that the error compensation scheme demonstrates relatively better results when evaluated under different weather and lighting conditions, particularly for straight trajectory flights. Thus, this chapter has adopted the KF error compensation strategy to design the method 2 hybrid architecture for conducting a comparative study to evaluate the robustness of hybrid architecture under fault scenarios including weather effects. In the filtering methods of VIO, a semi-tightly coupled architecture is implemented using ESKF. The state vector of ESKF-based VIO selects the following states-

$$x = [P_N, P_E, P_D, V_N, V_E, V_D, A_N, A_E, A_D] \quad 3-1$$

Where P_N, P_E , and P_D denote position, V_N, V_E , and V_D denote velocity and A_N, A_E , and A_D denote attitude in the N, E, and D position. The fusion strategy is utilized based on state models described in [70]. The updated fusion measurement states update process is illustrated in the following equations-

$$x_k = x_{k|k-1} \oplus \widehat{\delta x_k} \quad 3-2$$

$$x_k = x_{k|k-1} + k_k(z_k - H_k x_{k|k-1}) \quad 3-3$$

$$\hat{x}_k = x_k - \epsilon_k \quad 3-4$$

Where, x_k and $x_{k|k-1}$ represents the systems vector at k epoch and k-1 epoch; $\widehat{\partial x_k}$ represents the measured error-states; k_k denotes the Kalman gain, z_k denotes the innovation at k epoch; $\hat{x}_k$ denotes method 2 measured state vector and ϵ_k represents predicted unmodeled KF measurements increments at k epoch. The forward propagation of the GRU model is adopted using similar data structures in [124] illustrated in Figure 3-3.

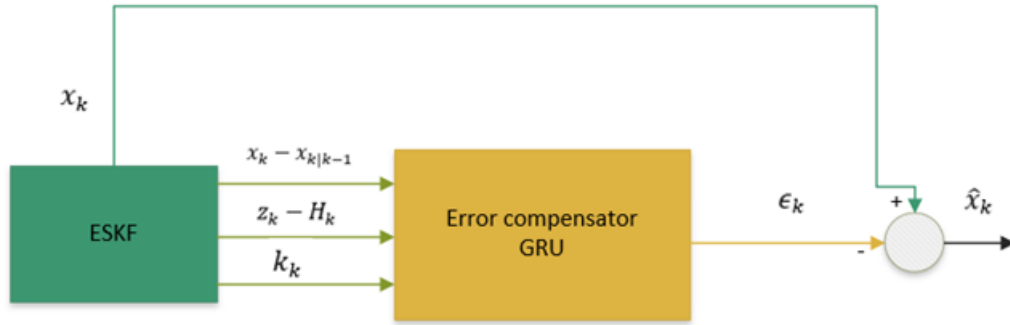


Figure 3-3 ESKF Error compensation structure for forward propagation

The inputs are composed of ESKF, estimating the difference between the state prediction and previous state estimation, $x_k - x_{k|k-1}$, the difference between measurement and observation estimation, $z_k - H_k$ and ESKF estimated gain, k_k . The predicted ESKF measurement increments are the output of the model. Regarding GRU architecture, input layers comprise 13 units, with 400 hidden units, a sequence length of 128, a batch size of 64 and a dropout rate of 0.5.

3.2.3 Method 3

The influence of Kalman gain on noise statistics of KF in a dynamic environment inspired the exploration of the hybrid fusion architecture, where GRU is employed to predict Kalman gain, aiming to mitigate state nonlinearity. KalmanNet [126] is a popular hybrid architecture that uses the Extended Kalman filtering fusion approach. In situations with complex system dynamics and limited available knowledge, especially during heavy rainfall obscuring environmental features, employing an NN-based K-gain prediction is advantageous. This approach enables the updating of KF measurements, overcoming the limitations imposed by parametric models, such as assuming Gaussian noise. Hence, this chapter

has adopted this phenomenon on the VIO system and compared it with the other two hybrid methods. The GRU-based K-gain prediction follows the similar architecture states in the previous section. The input layers consist of the difference between measurements and observations, $z_k - H_k$; the difference between posterior state and prior state estimate, $x_k - x_{k|k-1}$ and the output layer is predicted Kalman gain, k_k . The high-level forward propagation structure is illustrated in Figure 3-4. Later, the predicted k_k is updated using equation 3-3 which estimates the VIO position.

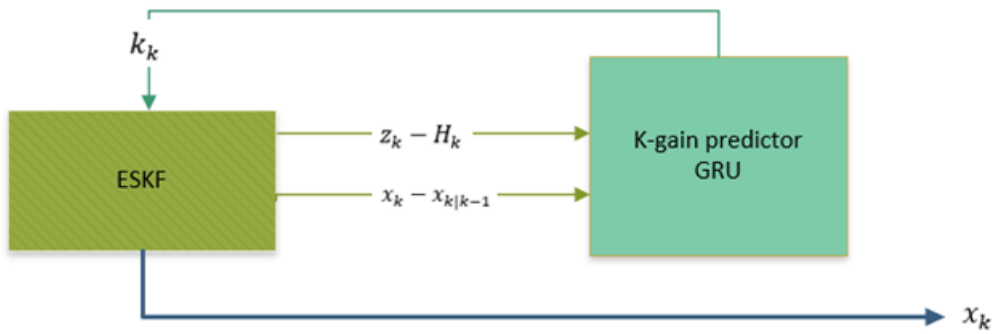


Figure 3-4 Kalman gain k_k predictor structure for forward propagation.

3.3 Experimental Setup and Data Generation

The section outlines the experiments designed to evaluate the strengths and limitations of the proposed comparative study on three hybrid approaches. The experimental setup is described first, which is used to generate training and testing datasets to conduct the numerical analysis and performance evaluation on each hybrid VIO.

3.3.1 Experimental Setup

The proposed comparative analysis is implemented and evaluated in MATLAB, incorporating the Unreal engine to simulate the data collection process for UAVS [123]. The simulated synthetic datasets encompass a variety of challenging fault conditions, including illumination variation, rapid motion, motion variance, flight dynamics, and less-textured environments under different weather conditions. The experimental setup used for the proposed comparative study is based on the

simulation setup used in [123] with default configuration parameters for the monocular camera and IMU sensor. Nevertheless, to simulate synthetic weather conditions, including partly sunny, fog, cloud, and rain ‘Simulation 3D Scene Configuration’ block in MATLAB 2022a has been updated with the parameters listed in Table 3-1.

Table 3-1 Specification of ‘Simulation 3D Scene Configuration’ block in MATLAB 2022a

Scene Selection	Weather Selection			
US City	Sunny	Sun altitude 90	Sun Azimuth 180	
	Cloud	Sun altitude 90	Sun Azimuth 180	
		Cloud opacity 40	Cloud speed 80	
	Fog	Fog Density 80	Sun altitude -45	
			Sun Azimuth 270	
	Rain	Rain Density 30	Sun altitude 90	Sun Azimuth 180
			Cloud opacity 40	Cloud speed 80

In urban operation scenarios surrounded by buildings, VIO utilises a camera to capture visual features from environmental structures (tall buildings, lampposts, trees), which in turn provides position information. The MATLAB simulator interfaces with QgroundControl software to generate multiple trajectories that combine different failure modes and fault conditions, providing control over the selection of single and combined faults. This capability allows for the analysis of the strengths and limitations of different hybrid approaches implemented on VIO systems under complex weather effects and lighting conditions, which is the focus of the chapter. Figure 3-5 illustrates different fault conditions in different weather effects used in this chapter.

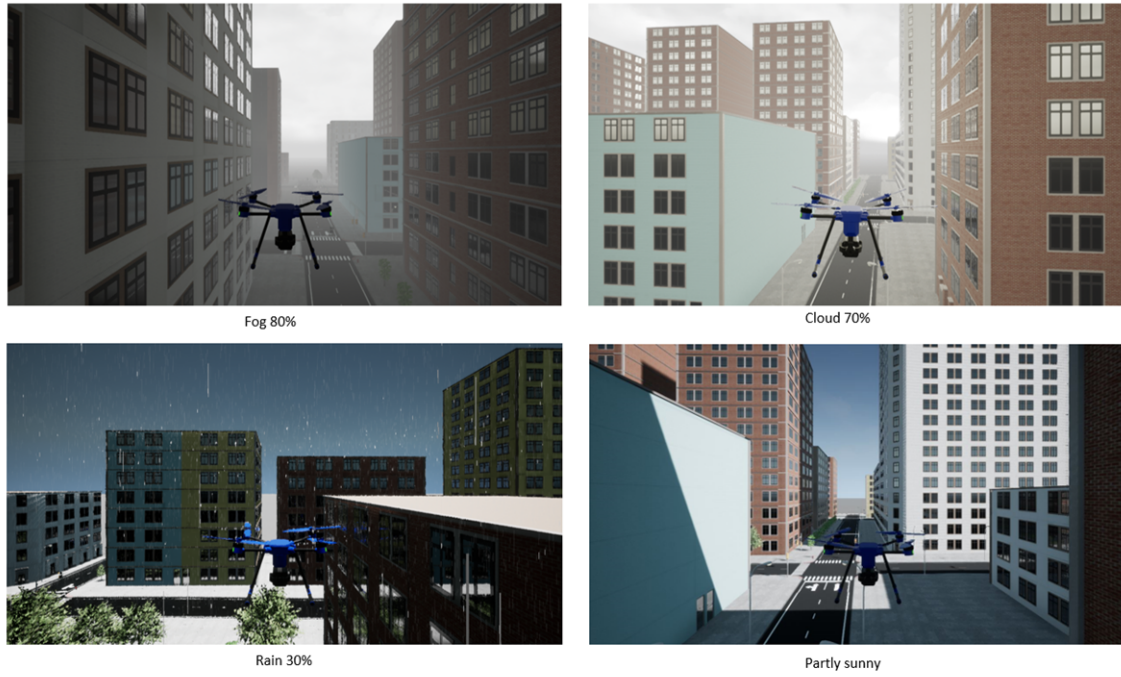


Figure 3-5 Examples of simulated environmental conditions used in testing and training datasets.

3.3.2 Training Datasets

Regarding the training of the GRU model, 12 sets of trajectories are utilized, encompassing environmental fault conditions, flight dynamics, and lighting variation, e.g. sunset, evening, midday, and sunrise. Each trajectory spans 15 minutes and features various weather conditions such as fog, sunny, rain, and clouds. A summary of the training trajectories containing fault conditions is given in Table 3-2. Since the chapter utilises three different hybrid methods, the GRU model employed on each hybrid VIO is trained separately using the data format explained in Section 3.2. For simplicity, the pre-trained GRU model used in [123] is adopted for the method 1 algorithm. Additionally, the GRU model utilised in each hybrid VIO is trained separately with 12 distinct datasets, each tailored to the mentioned weather and lighting conditions, following the structures

explained in Section 3.2.

Table 3-2 Characteristics of training trajectories

Training Sequence	Turn followed by the featureless wall.	Skyview with turn	White Rooftop view	Shadow of building on a building	Illumination change	Flying nearby objects (buildings)	Shadow of buildings and trees	Take off with altitude variation.
1	√	√		√	√		√	√
2			√	√	√	√	√	√
3		√		√	√	√	√	√
4				√	√	√	√	
5				√	√		√	
6	√		√	√	√	√	√	√
7		√		√		√	√	
8				√	√		√	
9	√	√	√	√			√	√
10			√	√		√	√	
11	√			√	√	√	√	√
12		√	√	√		√	√	

The extraction of visual features using an ORB-feature detector by VO is firstly adapted to produce relative pose measurements for the fusion strategy [123]. The IMU measurements utilized in this work have been adopted from previous work All the demonstration experiments for training the GRU model [10], utilise the Adamax optimiser and Mean Squared Error (MSE) loss function. These experiments are conducted in Python, leveraging relevant libraries compatible with TensorFlow for implementation. Figure 3-6 illustrates examples of training trajectories used to train the GRU model for each hybrid algorithm.

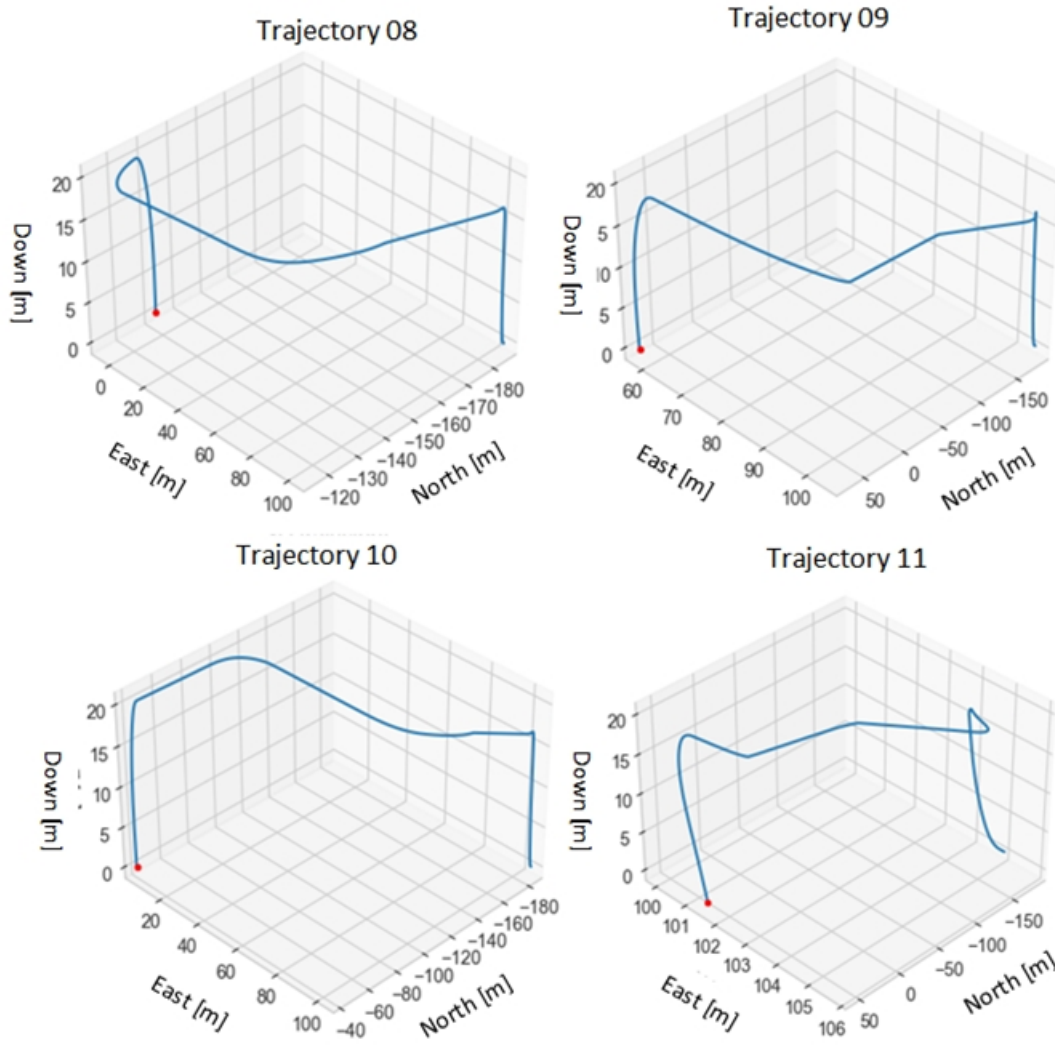


Figure 3-6 Examples of trajectories used for training purposes.

3.3.3 Evaluation metrics

In order to assess the performance of the proposed comparative analysis, the Root Mean Square errors (RMSE), horizontal errors, and 95th percentile are calculated and compared with the three hybrid VIO algorithms. To analyse the behaviour of each algorithm and error distribution over failure modes, the whole trajectories visualization plots have been utilized with ground truth. The RMSE calculation for each axis is as follows-

$$RMSE = \sum_N^{i=1} \sqrt{\frac{(\hat{P}_I - P_i)^2}{N}} \quad 3-5$$

Where, $\hat{P}_I$ and P_i are the predicted value and ground truth along with N, E, D position.

RMSE for 3D position as follows-

$$RMSE_{3D} = \sum_{N=1}^i \sqrt{\frac{((P_{N_I} - P_{N_i})^2 + (P_{E_I} - P_{E_i})^2 + (P_{D_I} - P_{D_i})^2)^2}{N}} \quad 3-6$$

RMSE for horizontal error as follows-

$$RMSE_{Horizontal} = \sum_{N=1}^i \sqrt{\frac{((P_{N_I} - P_{N_i})^2 + (P_{E_I} - P_{E_i})^2)^2}{N}} \quad 3-7$$

Where, P_{N_I}, P_{E_I} and P_{D_I} are predicted NED position and P_{N_i}, P_{E_i} and P_{D_i} are the ground truth positioning values from MATLAB incorporated Unreal Engine simulator in NED coordinates.

3.4 Performance analysis

To assess the strengths and limitations of the hybrid VIO algorithms encompassing the purpose of the research, three types of experiments are conducted using the three testing datasets, shown in Figure 3-7. These trajectories are generated using MATLAB incorporated Unreal engine simulator with US city scenario which recreates urban environments with various fault conditions including illumination variation, motion variation, altitude variation, no-texture, and weather effects. The details of the three experiments covering fault scenarios with three specific trajectories are described in Table 3-3.

Table 3-3 The defined experiments covering scenarios

# Trajectory	Scenarios	Fault type
1	Single fault condition	1. Shadows of objects 2. White Rooftop 3. Turn
2	Combined fault conditions with weather effects	1. Turn with features-wall. 2. Turn with sky view. 3. Shadows of objects 4. Take off with different altitudes

3	Complex fault with unseen scenarios	<ol style="list-style-type: none"> 1. Turns with featureless walls. 2. Turns with sky-view. 3. Shadows of objects 4. Diagonal rotation 5. Altitude variation 6. Take-off with different altitudes.
---	-------------------------------------	--

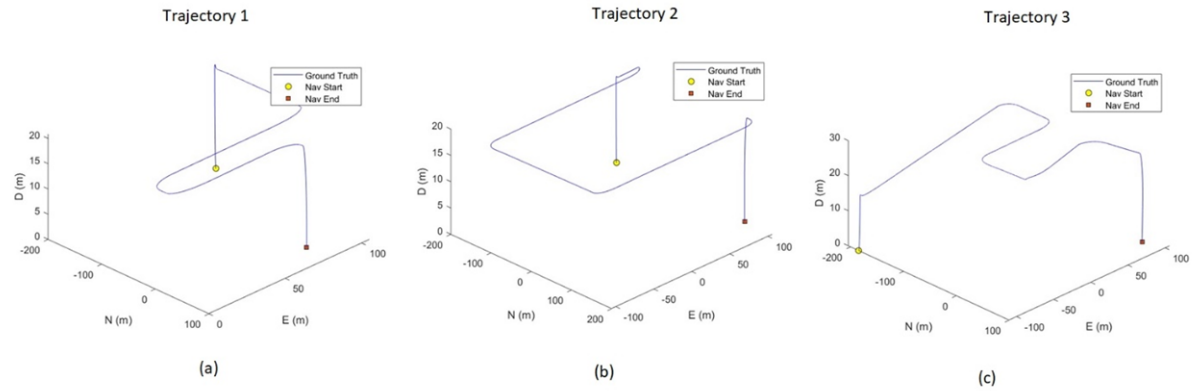


Figure 3-7 3D view of test trajectories used for experiments

The performance analysis has been categorized based on navigation scenario types and explained in the following sections.

3.4.1 Single Fault Condition

The purpose of this experiment is to evaluate the effectiveness and limitations of three hybrid VIO systems followed by method 1, method 2 and method 3 in managing single-fault condition and mitigation of failure modes associated with at a time with trajectory 1. During the experiment, distance of 702 m is covered over a duration of 152 s.

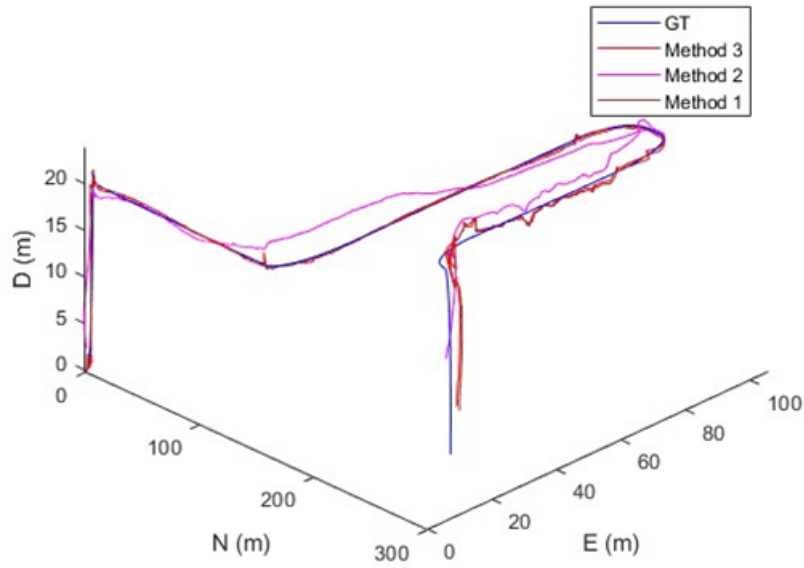


Figure 3-8 3D trajectory comparison of method 1, method 2, and method 3 algorithms under single fault condition.

Figure 3-8 shows the behaviour of 3D trajectories obtained from method 1, method 2, and method 3 hybrid architectures. These result trajectories are observed under single fault conditions where the presence of a white rooftop view, as illustrated in Figure 3-9 leads to deviation in the performance after the third turn. Notably, white rooftop triggering feature extraction errors and feature tracking errors failure modes within the time interval of 113-140 s. The method 1 effectively mitigates the failure modes and follows closely to the ground trajectory.

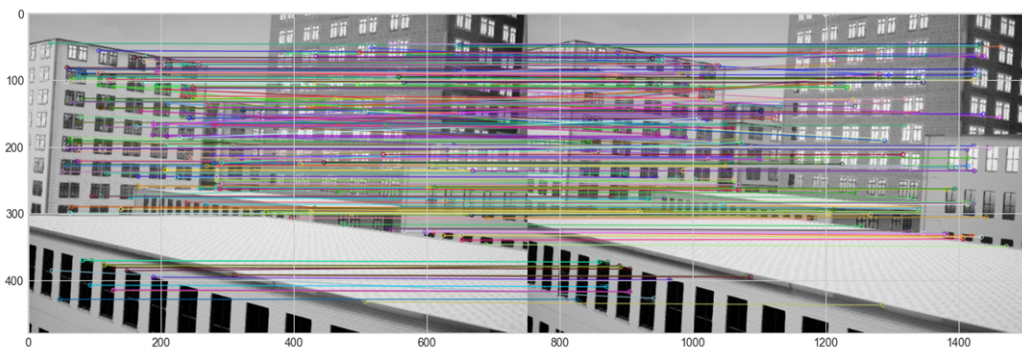


Figure 3-9 Example of feature tracking errors and feature mismatch failure modes due to the presence of white rooftop view at 115s

Figure 3-10 illustrates the error distribution of each hybrid algorithm relative to ground truth trajectory 1. It is notable that the error along the E coordinate significantly influences overall performance, primarily due to the presence of the featureless rooftop view. This effect, however, appears to be mitigated effectively by the method 1 algorithm compared to other hybrid approaches. The efficacy of the VO-error compensation mechanism is particularly noteworthy in this context as it addresses faults before fusion, thereby preventing the ESKF model from being significantly impacted by the substantial nonlinearity introduced by sensor anomalies. Conversely, the performance of the K-gain prediction mechanism in the E position appears to be less favourable, presumably due to its inability to sufficiently mitigate the imperfections in the model caused by corrupted sensor measurements.

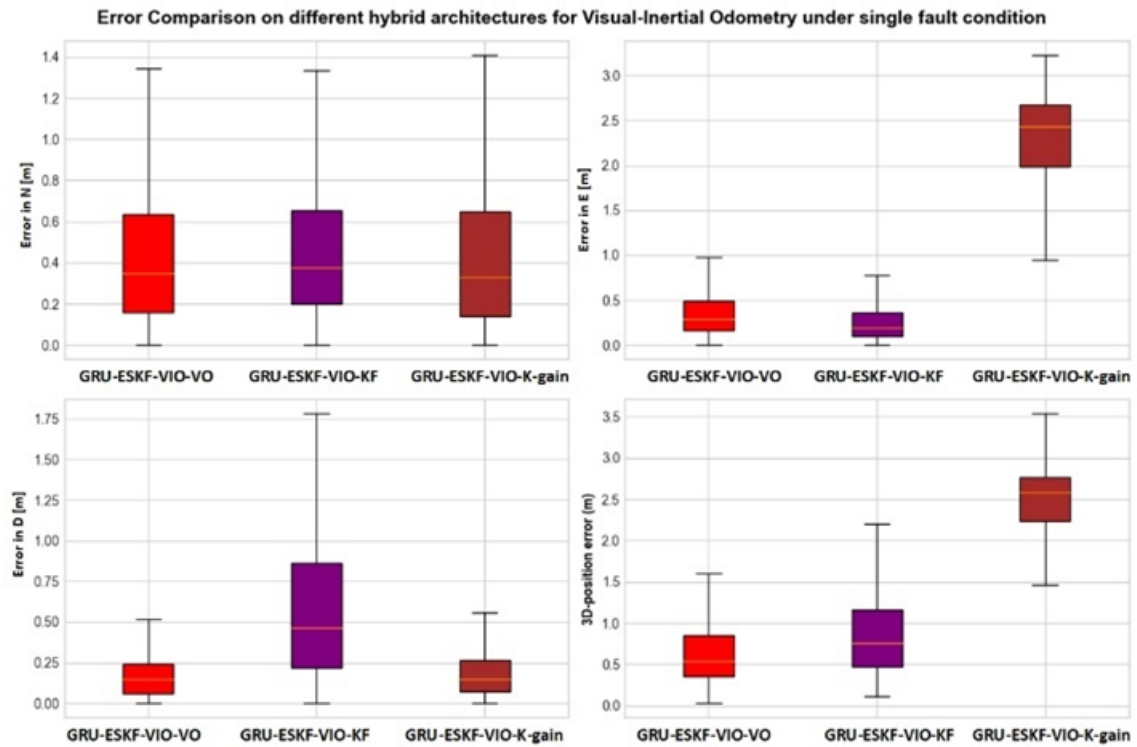


Figure 3-10 Position error statistics for adopted three hybrid architectures using trajectory 1.

Table 3-4 summarizes the qualitative results of the hybrid architectures with respect to two reference systems for easy comparison. It is noticed that the influence of minimizing drift with the VO measurement error compensator

mechanism is prominent and showed 77.4% and 75.5% improvements in horizontal positioning (95th percentile) compared to the VO and ESKF-VIO reference systems respectively. The prominent failure modes occurring in the horizontal axis when the UAV takes turns include feature tracking errors, feature mismatch errors and data association errors exacerbated by factors such as abrupt changes in the environment, occlusion of visual features and rapid motion. The hybrid algorithms are effectively mitigating these failure modes and improving overall horizontal positioning performance by 85.5%, 70.2%, and 84.1% for method 1, method 2, and method 3 compared to ESKF-VIO respectively.

Table 3-4 Positioning performance for Trajectory 1 under single fault conditions.

Techniques	RMSE (m)				95 th Percentile (m)				Horizontal RMSE (m)	Horizontal 95 th Percentile (m)
	N	E	D	3D	N	E	D	3D		
VO	9.7	4.03	13.36	16.24	19.5	7.49	13.55	23.65	10.53	13.94
ESKF-VIO	8.02	2.26	11.72	14.24	18.14	4.85	11.83	21.43	9.27	12.82
Method 1	0.95	0.95	0.65	1.43	1.35	2.31	2.71	4.22	1.32	3.14
Method 2	0.99	0.77	1.12	1.66	3.2	2.8	3.82	4.85	2.7	4.27
Method 3	1.08	2.28	0.72	2.03	1.46	2.20	3.82	5.16	1.30	3.24

Figure 3-11 illustrates the RMSE in position estimation for each algorithm with two reference systems of VO and ESKF-VIO. The observed performance variance indicates at 60 s, coinciding with the UAV's initial turn, attributable to the limited field of view. This effect persists until the second turn at 85 s, where objects (buildings, lampposts, and trees) are situated far away. The presence of such fault conditions triggers feature extraction errors and feature location errors failure modes which are effectively mitigated by hybrid algorithms. Furthermore, The RMSE position errors from 135-153s are higher where 95th percentile positions are reduced to 81.7%, 20%, 47.2% for VO-error compensation mechanisms, 10%, 17%, 35.6% for KF-error compensation, 17.9%, 34.2%, 59.4% for K-gain predictor concerning ESKF-based VIO fusion along with N, E, and D position. It is observed that the Method 1 algorithm has shown better

performance under single-fault conditions while mitigating failure modes including feature extraction errors, feature tracking errors, feature mismatch, data association errors and feature location errors.

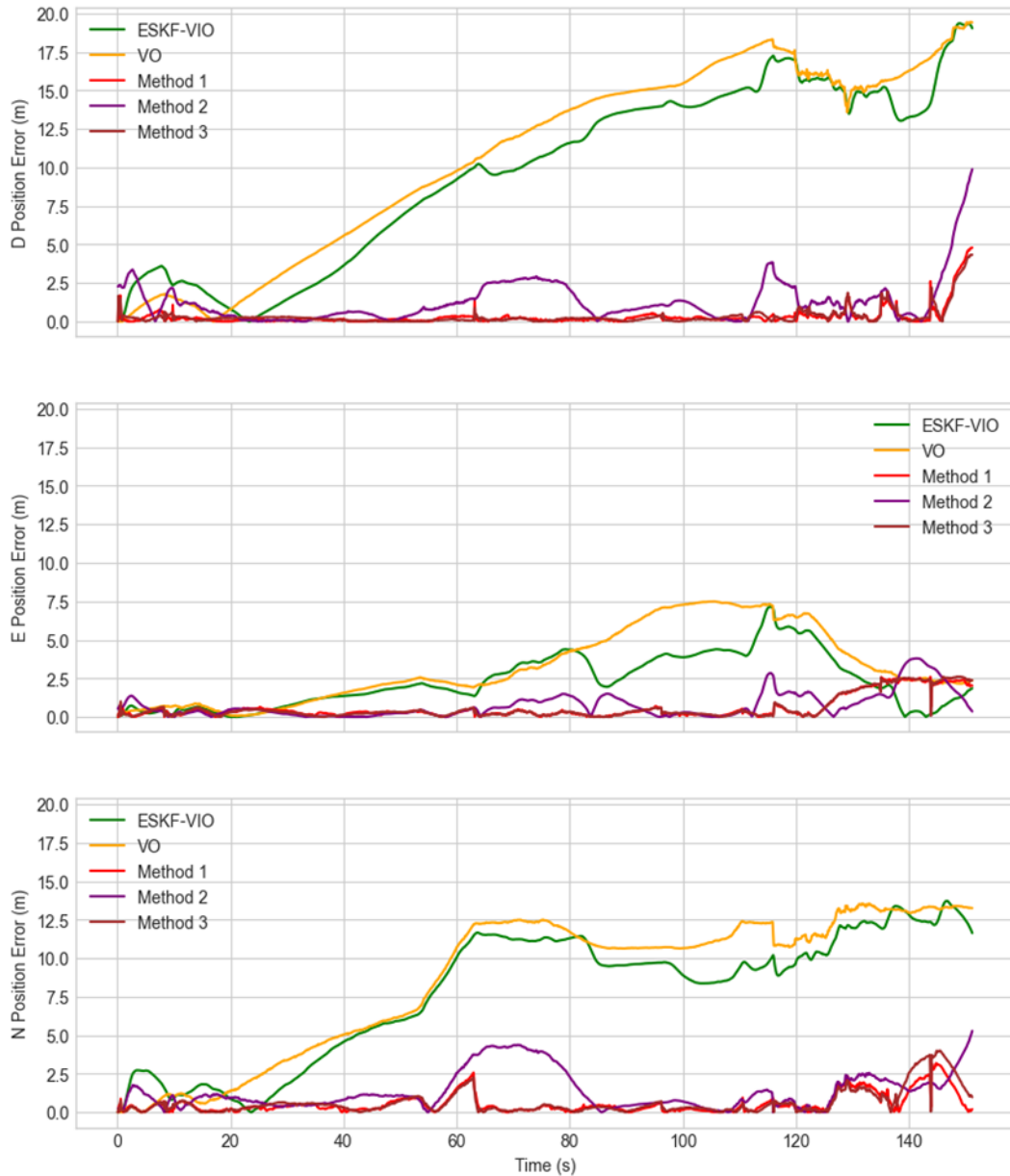


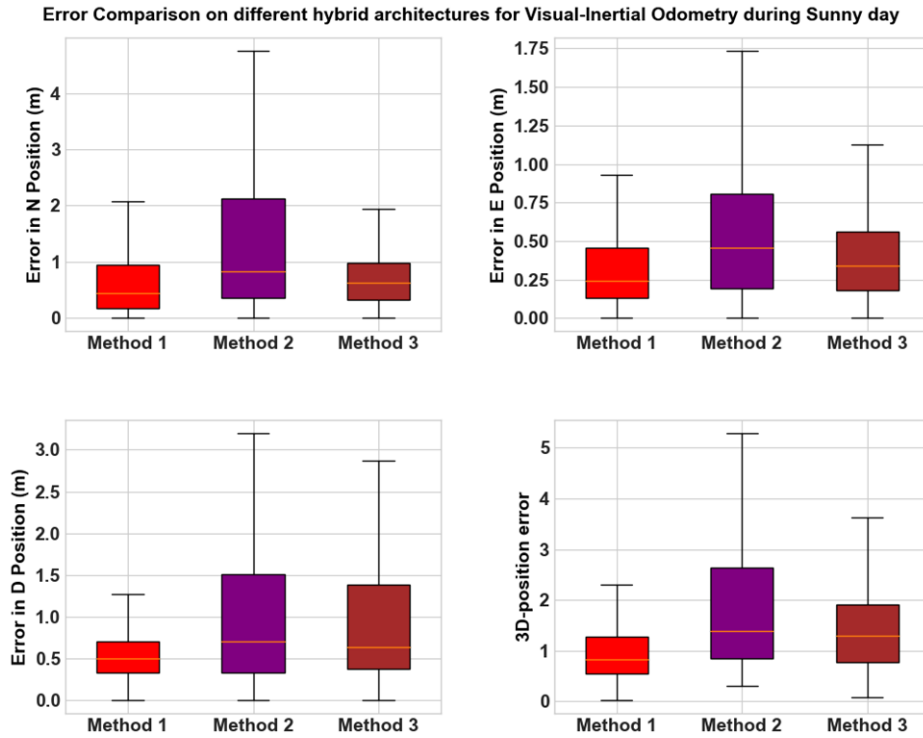
Figure 3-11 Position error in the presence of a single fault condition for Trajectory 1

3.4.2 Combined Fault Conditions with Weather Effects

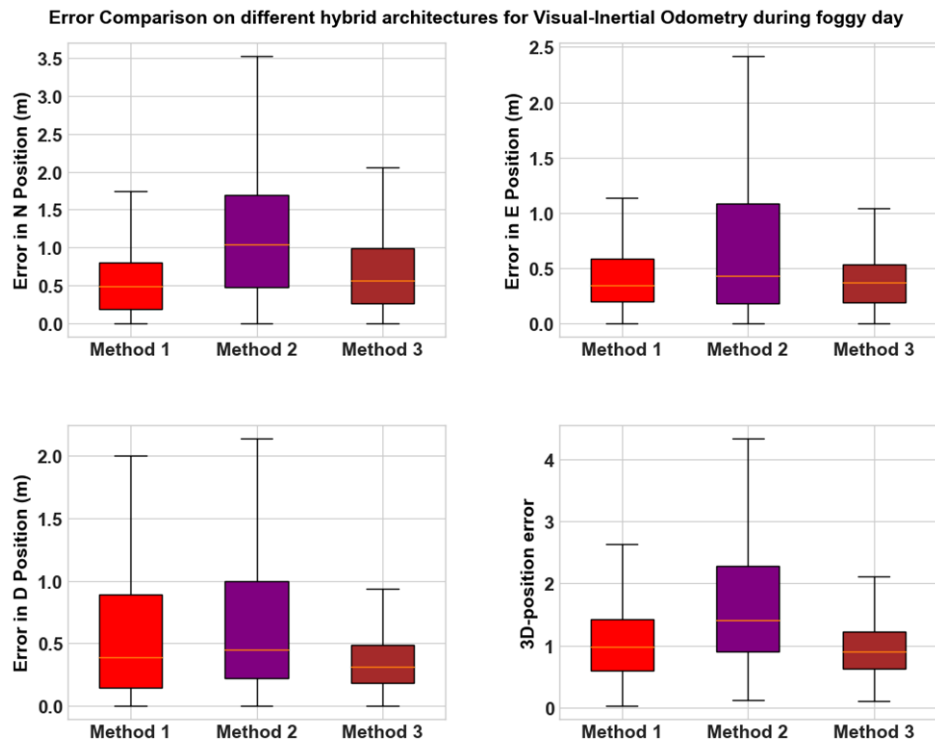
The purpose of this experiment is to evaluate the effectiveness and limitations of three hybrid VIO systems followed by method 1, method 2, and method 3 in

managing combined fault conditions and associated failure modes alongside weather effects using trajectory 2. The experimental environment includes a combination of tall and short buildings with shadows of the environmental objects (e.g. buildings, trees), turn with features-wall, turn with skyview, shadow of tall buildings on building and take-off with different altitudes. Three experiments are conducted with similar trajectories under three weather conditions such as partly sunny, rain and fog. During the experiment distance of 966 m was covered within a duration of 303 s.

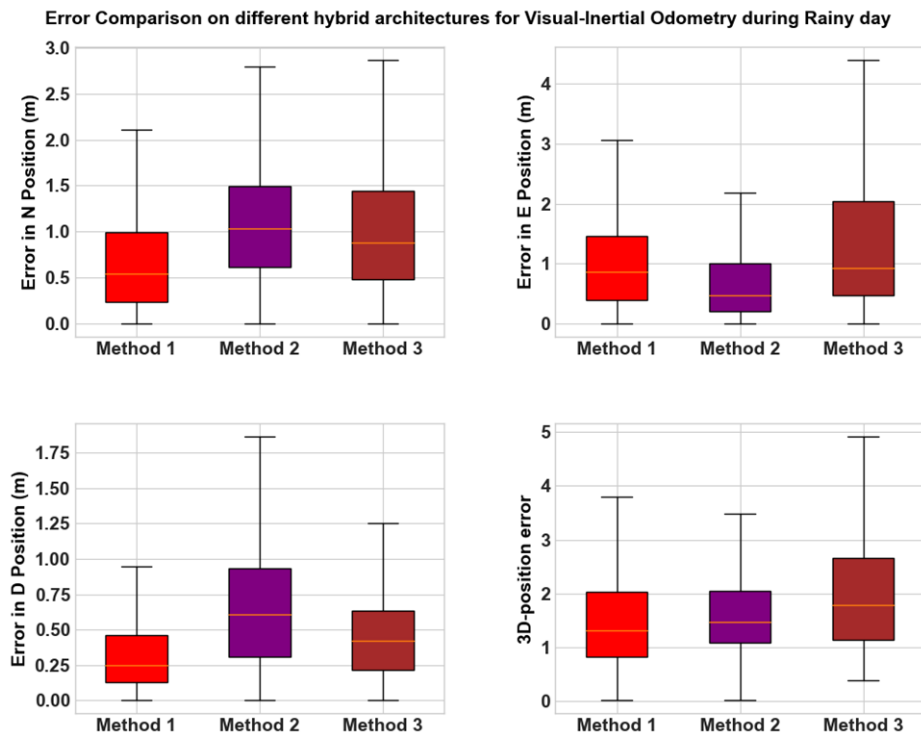
Figure 3-12, the boxplots summarising the position errors of the three hybrid approaches, including method 1, method 2, and method 3, are presented, illustrating how the results vary for different approaches.



(a)



(b)



(c)

Figure 3-12 Position error statistics for adopted three hybrid architectures under weather effects using trajectory 2.

During partly sunny weather conditions, it is observed that the presence of increased noise measurements during extended flight and shadows (illumination variation) with combined fault conditions limit the K-gain predictor and KF-error compensator mechanism approach, providing the erroneous position compared to the VO-error compensator approach. These erroneous measurements resulting from multiple failure modes occurring simultaneously, such as feature location errors, feature tracking errors and feature mismatch errors are not only influenced by the illumination changes but also other fault factors. For instance, when the UAV turns near the featureless wall and Skyview trigger feature tracking errors and feature extraction errors failure modes. These additional failure modes contribute to the performance reduction observed in ESKF-VIO as illustrated in Figure 3-13. However, the hybrid approaches effectively mitigate the effects introduced by multiple failure modes: method 1 by 96%, method 2 by 91.1%, method 3 by 92.3%.

Challenging scenarios, such as foggy weather conditions significantly contribute to visual failure modes like feature extraction errors and feature location error failure modes due to limited visibility to environmental features. These condition leads to model uncertainty which is improved by learning the K-gain predictor and achieving better overall 3D position performance in 31.5% enhancement. Furthermore, Rain results in water droplets accumulating on the camera lens, blurring visual features, lighting variation and reflections leading to obstruction of feature tracking and matching. It is noticed that the presence of increased noise measurements due to these factors limits the K-gain predictor and KF measurement error compensation techniques, proving improvement in the overall position for the VO-error compensation mechanism. However, the method 1 mechanism has shown better performance in the vertical axis during foggy and rainy weather conditions, removing the drift while landing compared to other approaches.

Figure 3-13 illustrates RMSE position error comparison based on the adopted three hybrid architectures and benchmark with two reference systems of VO and ESKF-based VIO under sunny, foggy, and rainy weather conditions. During the

landing phase, notably in foggy and rainy weather conditions, the RMSE errors in the E and D positions are observed higher. It is worth noting that the performance degradation is noticed after 250 s due to combined fault conditions including UAV turn with sky view and shadow of buildings on a building. These condition leads to illumination variation and less-texture environmental fault conditions, resulting in feature extraction errors, feature location errors and feature tracking errors failure modes. However, fault conditions induced by weather effects such as reflection, overexposure, less visibility, and lighting variation further complicate the navigation environment. For instance, The KF error compensator mechanism exhibits an improvement of 3% and 2% at horizontal (95th percentile) to VO error compensation and K-gain prediction mechanism under foggy weather conditions. Additionally, the drift observed while landing after 280 s is also reduced up to 2.25 m at the D position (95th percentile).

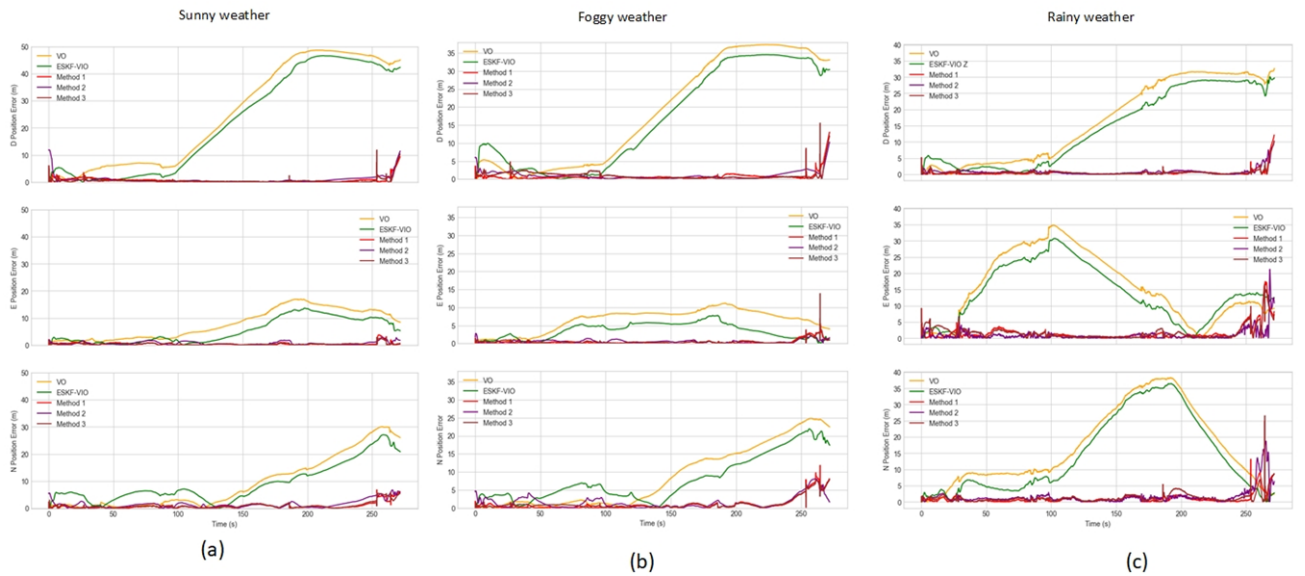


Figure 3-13 Position error along with each axis in the presence of a single fault condition for trajectory 2 under weather effects

Table 3-5 Positioning performance for Trajectory 2 under sunny weather conditions.

Techniques	RMSE (m)				95 th Percentile (m)				Horizontal RMSE (m)	Horizontal 95 th percentile (m)
	N	E	D	3D	N	E	D	3D		
VO	12.10	7.20	23.69	27.56	11.35	10.50	24.09	43.47	14.08	24.80
ESKF-VIO	10.08	4.11	21.60	24.19	34.52	6.79	20.42	39.01	10.89	20.45
Method 1	1.29	0.68	1.22	1.86	2.00	0.86	4.09	4.39	1.47	4.36
Method 2	1.28	0.51	1.28	2.13	1.79	1.06	2.25	3.00	2.18	5.03
Method 3	1.44	0.56	1.18	1.95	1.27	0.85	3.67	4.59	1.56	4.05

Table 3-6 Positioning performance for Trajectory 2 under foggy weather conditions.

Techniques	RMSE (m)				95 th Percentile (m)				Horizontal RMSE (m)	Horizontal 95 th percentile (m)
	N	E	D	3D	N	E	D	3D		
VO	13.16	9.47	30.97	34.95	48.26	16.94	28.32	56.09	16.21	30.96
ESKF-VIO	11.38	6.94	28.60	31.56	46.61	12.72	23.94	50.61	13.33	25.76
Method 1	2.29	0.62	1.58	2.35	1.87	1.37	7.03	7.51	2.13	6.32
Method 2	2.92	0.41	1.48	3.31	4.20	1.00	8.24	6.39	2.33	5.36
Method 3	1.82	0.76	1.29	2.87	1.46	1.65	5.36	9.40	2.05	6.01

Table 3-7 Positioning performance for Trajectory 2 under rainy weather conditions.

Techniques	RMSE (m)				95 th Percentile (m)				Horizontal RMSE (m)	Horizontal 95 th percentile (m)
	N	E	D	3D	N	E	D	3D		
VO	20.54	19.01	20.48	34.69	36.22	10.50	24.09	43.47	27.99	39.57
ESKF-VIO	18.54	16.50	18.17	30.76	34.45	6.79	20.42	39.01	24.82	36.253
Method 1	1.53	2.70	1.27	3.45	0.74	0.47	2.99	6.28	3.25	6.64
Method 2	3.11	2.15	1.27	3.93	4.49	6.61	4.17	6.43	3.60	10.29
Method 3	1.99	2.94	3.35	3.80	1.64	6.05	7.67	7.75	3.29	7.03

Table 3-5, Table 3-6 and Table 3-7 summarise the performance of the three hybrid architectures compared to the reference systems concerning RMSE, 95th percentiles, horizontal RMSE and horizontal 95th percentile error under weather effects. It is observed that there is a variation in performance at each algorithm under different fault conditions. During sunny weather, the KF measurement error mechanism can improve overall 3D performance (95th percentile) by 25% and 6% for method 1 and method 3 respectively. However, K-gain predictor mechanisms have shown poor performance due to dependency on assumptions and predictive models, sensitivity noise and limitations to adapt complex fault conditions by the predefined model. For instance, in foggy weather conditions characterized by limited visibility, overexposure, and contrast, the overall performance exhibits degradation to 13.40 m in the 3D position (95th percentile). Furthermore, the method 1 and method 2 algorithms overcome such complexity in environments by directly mitigating sensor errors during rainy weather and measurement errors during sunny and foggy weather. It is observed that compensation of sensor measurements or model measurements is beneficial under this fault condition. Under different weather conditions, it is noted that the prominent failure modes are feature tracking errors and feature extraction errors due to lighting variation, reflection on the surface, speed of raindrops and fewer features extracted from the sky. Such effects introduce erroneous measurements and resulting performance degradation in all hybrid architectures during the landing phase as discussed. However, the VO error compensation technique shows relatively improved performance, resulting in 27.8%, 30.29%, and 28.95% improvement at the D position (95th percentile) for VO error compensation, KF error compensation, and K-gain prediction mechanism respectively during foggy weather. As per principle, K-gain prediction solely relies on the predefined predictive model which results in the imperfection of the model. Method 3 can mitigate the fault conditions mentioned above, yet the method 1 performs better, boasting a 3D position error (RMSE) of 2.35 m. Compared to the performance sunny weather conditions, the drift is higher position error of 6.59 m, 6.00 m, and 6.99 m (95th percentile) for method 1, method 2, and method 3 respectively.

Figure 3-14 illustrates the behaviour of 3D trajectories obtained from method 1, method 2, and method 3 hybrid architectures. These trajectories are evaluated under combined fault conditions across three weather conditions. Notably, deviations in performance are observed at each turn due to the presence of less features. Moreover, during the landing phase, performance degradation is encountered due to complex fault conditions computed by variations in weather. For instance, rainy weather introduces variation in lighting along with raindrops, resulting in failures in feature tracking and extraction simultaneously. An important conclusion about these results can be drawn by comparing the 3D trajectories. The method 1 algorithm shows better performance in mitigating these failure modes along with following the ground truth trajectory 2 under these diverse weather conditions. Furthermore, method 2 indicates better performance in reducing drift and effects of combined associated failure modes e.g. feature tracking errors, feature location errors and data association errors during landing.

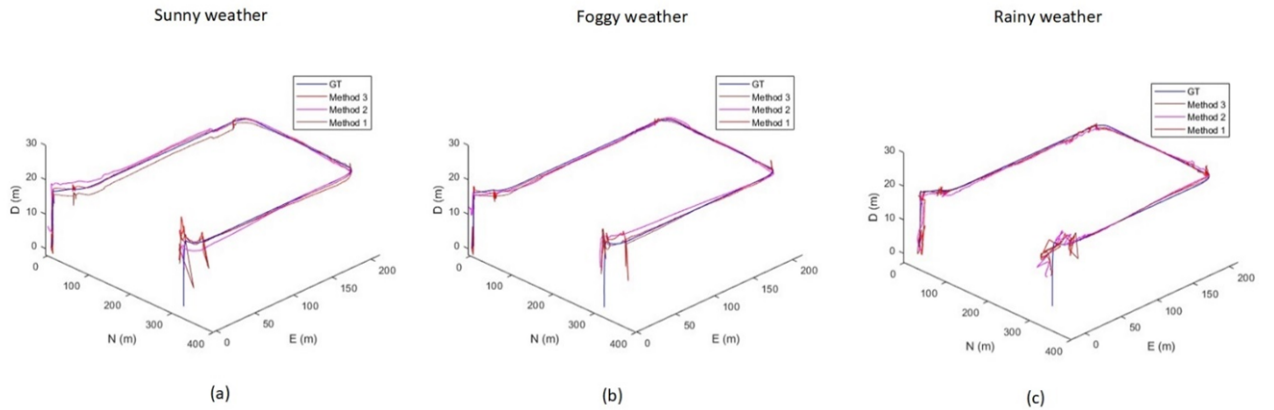


Figure 3-14 Trajectory comparison of hybrid algorithms under diverse weather conditions for trajectory 2.

In conclusion, the above analysis underscores the importance of the robustness of hybrid VIO systems with the mitigation of complex and combined visual failure modes, particularly under challenging conditions. The ability of method 1 and method 2 algorithms to adapt to varying fault scenarios and environmental conditions highlights their potential for enhancing navigation accuracy and reliability.

3.4.3 Complex Fault conditions with un-seen scenarios

The objective of this experiment is to evaluate the generalization ability of three hybrid VIO systems in managing combined fault conditions, both seen and unseen scenarios within trajectory 3. The experimental setup comprises a diverse scenario, including the combination of tall and short buildings with shadows of the environmental objects (e.g. buildings, trees), turns with featureless walls, turns with sky-view, shadow of tall buildings on building, diagonal rotation, altitude variation, and take off with different altitudes. The experiment is conducted with trajectory 3 under sunny weather conditions. The flight path deviates from training datasets as it takes a different road, albeit encountering similar patterns throughout the flight. During the experiment distance of 738 m was covered within a duration of 230 s.

Figure 3-15 presents the behaviour of 3D trajectories obtained from three hybrid approaches. It is observed that the hybrid algorithms demonstrate the ability to generalize effectively across seen and unseen scenarios, closely following the ground truth trajectory. During the takeoff, all hybrid approaches align well with the ground trajectory, successfully adapting to the unseen scenario. However, deviation in performance is observed during the landing phase, particularly due to the diagonally rotated landing of the UAV to an unseen scenario. Notably, among all the hybrid approaches, the method 1 exhibit better performance in mitigating the fault conditions compared to others.

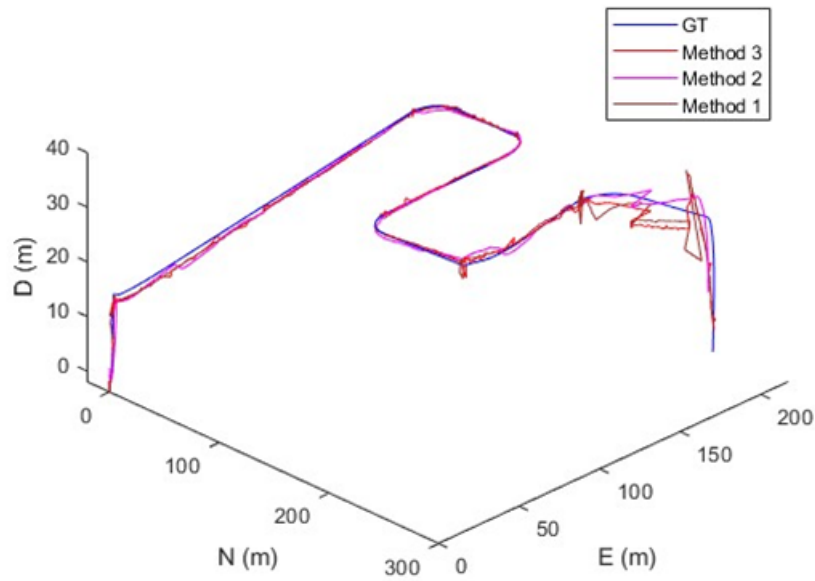


Figure 3-15 3D trajectories of hybrid algorithms under complex fault conditions with unseen scenarios for trajectory 3.

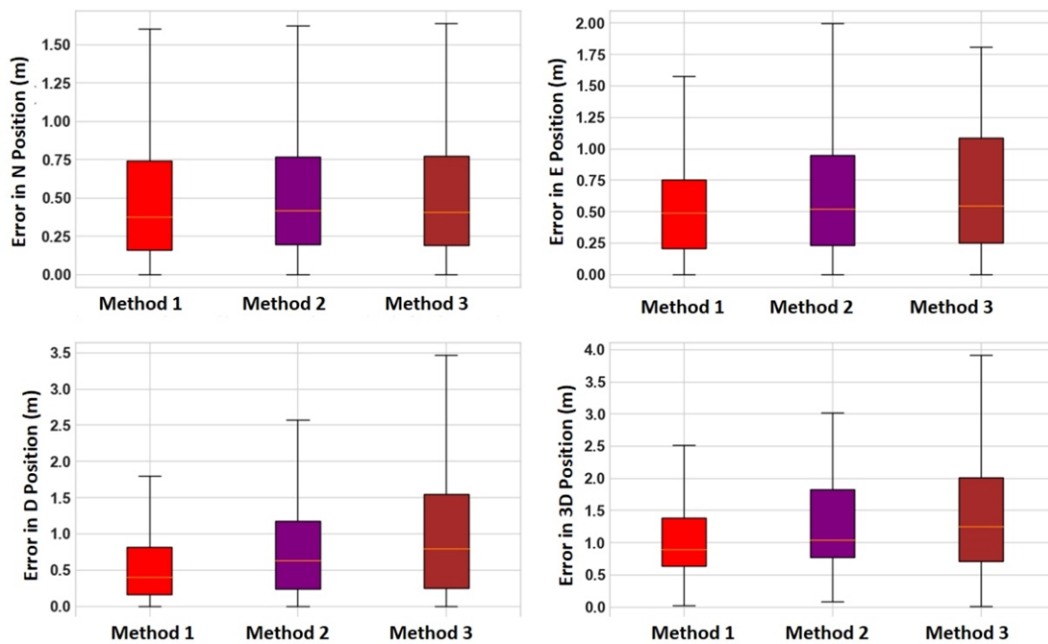


Figure 3-16 Boxplot summarizing the position error statistics for adopted three hybrid architectures under weather effects using trajectory 3.

In Figure 3-16, the boxplots summarizing the position error of the three hybrid approaches including method 1, method 2, and method 3 are presented,

illustrating how the results vary for different approaches with unseen scenarios. It is observed that the presence of increased noise measurements during extended flight and sunny weather conditions with combined fault conditions limit the K-gain predictor and KF-error compensator mechanism approach, providing the erroneous position compared to the VO-error compensator approach. Complex unseen scenarios like diagonally rotation fault conditions contribute to likely feature extraction errors and feature tracking errors, resulting in model uncertainty which is improved by learning fault patterns in sensor data and achieving better D position performance to 1.28 m RMSE error. However, the performance improvement was shown by the VO error compensator in the 3D position by 19.86% compared to the ESKF- VIO reference system. The N position acts after takeoff when the UAV enters into an unseen scenario. Results from Table 3-8 indicate KF error compensator algorithm can minimize the erroneous measurements introduced after flight takeoff in the horizontal axis due to that turn with a featureless wall and variation of altitude at each turn. It is noticed that the presence of increased noise measurements indicates imperfection in the model, thereby limiting the effectiveness of the K-gain predictor technique compared to error compensation techniques. The findings address noise-included imperfections in measurements and highlight that the error compensation technique works better, particularly in scenarios involving variations in altitude and no-texture environments.

Table 3-8 Positioning performance for Trajectory 3 under unseen scenarios.

Techniques	RMSE (m)				Horizontal RMSE (m)
	N	E	D	3D	
VO	4.04	5.13	18.11	19.25	6.55
ESKF-VIO	3.77	4.85	16.00	17.15	6.12
Method 1	1.30	3.36	1.28	3.83	3.76
Method 2	0.76	3.46	2.46	4.31	3.44
Method 3	1.40	3.36	1.98	4.16	3.94

3.5 Conclusion

This study systematically investigates the performance differences among three types of hybrid NN-KF approaches implemented in VIO-based navigation systems. The evaluation encompasses three different navigation scenarios including single fault conditions, combined fault conditions with weather effects and complex fault conditions with unseen scenarios to provide insight into their robustness under multiple visual failure modes. It has been discovered that hybrid approaches effectively address the failure modes resulting from fault factors encountered in navigation environments like illumination variation, motion variation, altitude variation, limited field of view, and no-texture environments. The evaluation under different weather conditions exacerbates these faults by introducing additional fault conditions like shadows, limited visibility, contrast, overexposure, reflections, and speed of droplets. The method 1 algorithm exhibits better performance regarding mitigating fault conditions in sunny and rainy weather. However, the method 2 algorithm effectively mitigates the performance degradation caused by UAV turns in less feature environments, thereby reducing feature extraction errors with feature tracking errors failure modes. During rainy weather conditions performance improved by 90% and 91% for the VO error compensator and KF error compensator mechanism. Notably, the KF measurement error compensators work better than VO error compensators and K-gain predictor approaches in certain scenarios due to their ability to adjust the filter's state estimates by mitigating model errors. This adjustment allows the algorithm to better handle the specific combinational failure mode types occurring during UAV turns in the no-texture environment or diagonal rotation and landing drift. Additionally, the method 3 algorithm shows relatively poor performance among other hybrid approaches in terms of mitigating failure modes arising from different fault conditions. Integrating NN to predict K-gain attributes to challenges like imperfect model assumptions, bias and model uncertainty introduced by NN and imperfect predictive model due to the influence of noises under multiple fault conditions. In the experiment of evaluating the generalization ability of the hybrid systems under seen and unseen scenarios, the VO error compensation technique exhibits better performance with 19.86%

enhancement compared to other approaches. Conversely, out of all the hybrid approaches, not one method performed consistently well out of the evaluated fault conditions. Challenges remain, particularly in scenarios where multiple fault conditions converge such as diagonal rotation, landing drift or when navigating featureless environments under rainy and foggy weather conditions. Additional diversity in training datasets including complex fault scenarios to enhance the prediction accuracy of the learning-based GRU algorithm can be the forward step towards enhancing the adaptability of the system. Incorporating combinational hybrid architectures may complement each other's limitations and enhance the overall performance under practical operating conditions.

4 Improving Generalization Ability using GRU-AE based Error Compensation in Hybrid Visual Inertial Navigation Systems

4.1 Introduction

Urban Air Mobility (UAM) is revolutionising transportation by utilising the vertical dimension to ease ground traffic congestion and improve mobility within urban environments. It promises a wide range of services, such as emergency medical evacuations, air taxis, rescue operations and parcel delivery, which enhance the movement of people and goods in diverse contexts. To satisfy these, Unmanned Aerial Vehicles (UAVS) must operate safely and efficiently, often flying close to buildings in densely populated areas. However, such scenarios present significant navigation challenges due to the vulnerability of Global Navigation Satellite Systems (GNSS), which are prone to signal outages and multipath effects in urban canyons[2], [6]. These limitations emphasise the need for advanced positioning, navigation and timing (PNT) solutions that operate independently of GNSS. Furthermore, operational safety has high importance in UAM applications, with navigation safety being a critical prerequisite. Ensuring navigation safety requires robust systems capable of maintaining accuracy and promptly mitigating anomalies that could otherwise lead to catastrophic failures, such as crashes and collisions[1]. Overcoming these challenges with alternative PNT solutions for safe UAV operation will play a crucial role in gaining public acceptance and ensuring the sustainability of UAM applications.

In vision-based navigation systems, there are significant gaps in ensuring the robustness and reliability of these vision-based navigation solutions under fault conditions like environmental factors, lighting variations, flight dynamics and motion blur. Additionally, these fault conditions impact overall system performance by introducing uncertainty in the input data, e.g. ambiguity from feature mismatch errors[50] and the impact of non-Gaussian noise on predefined models [118]. Thus, vision-based navigation systems designed for safety-critical applications require careful consideration of all possible sources of uncertainty to ensure integrity and robustness.

Researchers working on classical visual navigation systems, considering improving integrity, typically focus on modelling feature tracking errors and aim to improve positioning performance [50]. However, the utilisation of such error models is limited to specific scenarios and lacks evidence of effectively mitigating visual errors to ensure safe navigation under diverse, complex urban environments [51]. In addition, error propagation as discussed in [31] highlights the impact of unmodeled noise arising from uncertainty sources, feature extraction errors, feature tracking errors, and feature location errors. These factors, including photometric noise, geometric errors and association errors, need to be considered as they contribute significantly to overall position error in VO, particularly in complex scenarios. Few studies [47], [77], [127] have proposed incorporating the photometric error model into the KF-based VIO system to minimise the effects of feature tracking errors, demonstrating improved performance in handling speed variations and low-texture scenarios while providing confidence boundaries. Nonlinearity in image intensities, caused by changes in lighting, shadows, reflection or exposures, leads to model mismatch due to linear assumptions of photometric error models [127]. These nonlinear intensity variations can introduce non-Gaussian noise into the data, which classical methods are not well-equipped to handle, often leading to divergence in state estimation. In contrast, data-driven VIO systems have been explored due to their ability to learn model complexity that can establish the relationship between sensor measurements and vehicle ego-motion. While effective in some cases, these methods lack robustness in low-texture environments, adverse image conditions (e.g. blur, distortions, extreme lighting) [43], [44], [49] and dynamic scenes. A notable limitation of end-to-end VIO architectures and CNN-based VO is that additional sources of uncertainty arise when the test and training data are mismatched, leading to unmodelled outliers. Furthermore, current state-of-the-art involving such architectures face challenges including reliance on large datasets [44], performance degradation in out-of-distribution test scenarios and insufficient uncertainty estimation [128], [129]. Some studies considering predicting position with uncertainty measure lack of considering all possible sources of visual uncertainty with adequate minimising outliers, leaving long tail

issues unresolved. Therefore, there is a need to incorporate the noise removal mechanism alongside addressing potential uncertainty sources within VIO systems to enhance robustness, resilience while reducing induced position uncertainty to ensure integrity.

To address shortcomings of these architectures, hybrid architectures incorporating Kalman filters (KF) with neural networks (NN) have gained attention to enhance performance under complex visual scenarios by adapting Neural Networks (NN) models (e.g. Bi-LSTM, LSTM, GRU) to mitigate sensor noise, model imperfections, corrupted measurements, and model assumptions [45], [48], [123], [124], [130]. However, during lighting variation, limited visibility and distortion in images, hybrid VIO with camera sensor measurement correction [45] and KF model measurement correction [124] suffers from performance degradation due to unmodelled outliers introducing nonlinearity that affects model assumptions. To mitigate sensor and model measurement errors, the combination of error compensation mechanisms was presented in my previous work [131] to enhance performance and reduce position uncertainty arising from sources of uncertainty like feature extraction errors, feature tracking errors and feature location errors in complex scenarios. However, the approach struggles to provide a confidence boundary due to induced outliers under corrupted sensor measurements when feature-based VO fails to extract any features, resulting in a tight distribution with long tails.

To address unmodelled outliers in non-Gaussian error distribution, providing position uncertainty reduction to ensure integrity under complex scenarios, this study extends the previous work by proposing a GRU-Autoencoder (GRU-AE) denoising mechanism for VO error Correction within Multi-ML Hybrid VIO to mitigate outlier effects. To the best of the author's knowledge, this is the first attempt to combine denoising-aided error correction with model error correction within VIO, providing the confidence boundary for integrity. The proposed hybrid fusion approach leverages the combined capabilities of GRU and Autoencoder to address data uncertainty and outliers for updating the ESKF filter, while another GRU compensates for model measurement errors to correct model

inaccuracies. In this research, ESKF is integrated into the hybrid architecture for its ability to linearise only the error states, enhancing stability and robustness, particularly in highly nonlinear systems. This makes ESKF more effective compared to other variants of KF, such as Unscented Kalman filters (UKF), which rely on sigma points to approximate distribution, potentially introducing instability [42]. As a result, ESKF is considered a more reliable choice for handling nonlinear dynamics. In addition, GRU is employed due to its efficiency in processing time-varying sequences, offering advantages over other NNs like LSTM and Bi-LSTM. In addition, the formulation is extended by exploring position uncertainty mitigation with denoising mechanism on the proposed approach using MATLAB incorporated with unreal engine simulated photorealistic datasets replicating possible uncertainty sources (e.g. feature tracking errors, feature exaction errors, feature location errors) and sources of unmodelled noise (e.g. limited visibility, darkness, extreme lighting) present in urban environments for extended flights. The contributions of this work are highlighted as follows:

- 1) A novel error compensation methodology is proposed with a combination of multiple error compensators forming Multi-ML hybrid VIO, which simultaneously corrects data uncertainty and unmodelled outliers arising from corrupted sensor measurements in complex urban environments.
- 2) A novel autoencoder-based denoising mechanism is proposed, which integrates into the GRU-based VO error compensator to enhance the generalising ability and denoise the sensor measurement outliers. The algorithm aims to reduce position uncertainty and minimise long tails in error distribution.
- 3) Systematic evaluation of the proposed uncertainty reduction with denoising mechanism under various sources of uncertainty, including feature tracking errors, feature location errors, data association errors and sensor noise, unmodelled noise from fault conditions such as limited visibility, darkness, extreme lighting.
- 4) Evaluation of performance under 'out of distribution' problems for the proposed novel hybrid VIO framework, aiming to estimate its ability

to adapt and perform effectively in unseen scenarios and flight dynamics.

The remainder of this chapter is structured as follows: section 4.2 includes the proposed hybrid visual-inertial navigation framework with brief descriptions. Section 4.3 discusses experiment setups for data generation to train and test the hybrid models. Section 4.4 discusses the testing performance analysis of the proposed algorithm under diverse complex scenarios, providing insight into the generalization ability. Finally, the conclusion is presented in the section 4.5.

4.2 Proposed Novel Hybrid VIO Framework

This section presents a high-level overview of the proposed novel framework that integrates the GRU-Autoencoder (GRU-AE) denoising mechanism for sensor error correction within Multi-ML hybrid VIO. In this architecture, images captured by a monocular camera are processed using the feature-based VO algorithm [50] to estimate VO position. The GRU-AE mechanism is employed to denoise and correct VO measurements by predicting position increments affected by data uncertainty and unmodelled outliers within complex navigation environments. The denoised and corrected VO positions are subsequently fused with raw IMU data, including linear acceleration and angular velocity, within the ESKF filter to estimate the VIO position. To further enhance positioning performance and reduce predefined model uncertainty, another GRU is utilised to compensate for ESKF measurement errors. The predicted measurement position increment will improve hybrid VIO positioning performance by reducing model and data uncertainty simultaneously. Figure 4-1 presents a high-level diagram of the proposed enhanced hybrid VIO framework. This chapter investigates and explores the performance of the proposed enhanced hybrid VIO framework, which integrates the position uncertainty compensation with the corrupted signals denoising mechanism, providing insight into the confidence interval to ensure integrity under the aforementioned uncertainty sources. The study focuses on mitigating the effects of various sources of uncertainty while simultaneously enhancing robustness and resilience for navigating under complex environments.

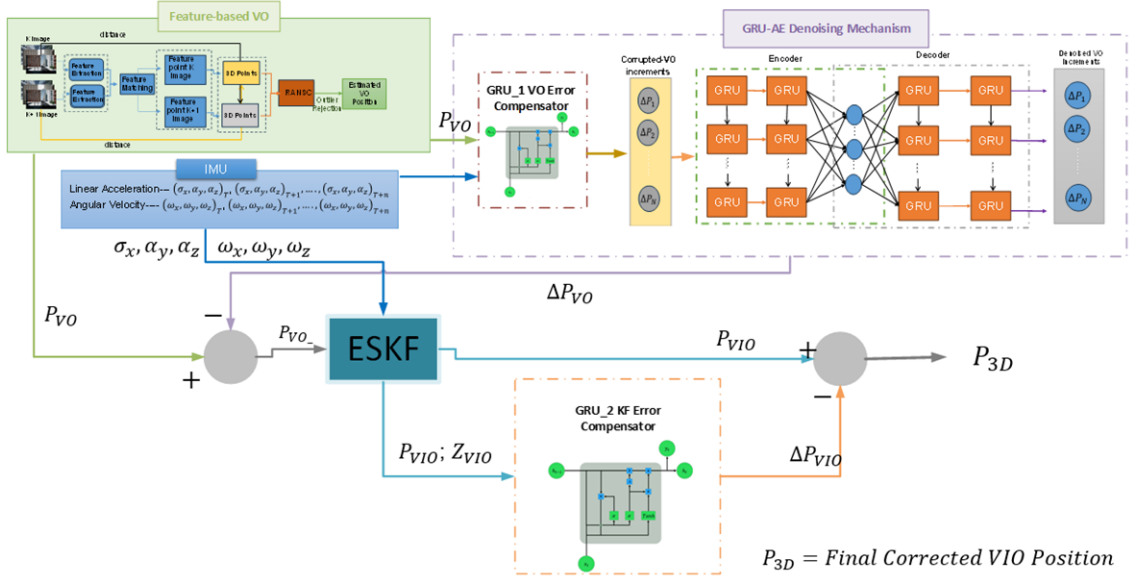


Figure 4-1 Proposed combined hybrid architecture with the denoising mechanism for VIO position uncertainty reduction

4.2.1 GRU-AE VO Error Correction

The motivation for integrating denoising mechanisms within the sensor error correction model is to address and mitigate the abnormal signals induced due to corrupted VO estimation in complex scenarios (e.g. no-feature to extract, darkness, change of flight dynamics). For this purpose, the denoising autoencoder is integrated at the output of the GRU-based error compensator to filter corrupted signals of the predicted increments. In practice, challenges arise to validate the lighting invariance assumption and output integrity of the end-to-end architecture [53], [60]. This research, therefore, focuses on feature-based visual positioning, and statistical FMEA analysis conducted in [123] highlights that visual fault conditions are major contributors on a large scale. The VO error compensation method in the hybrid VIO architecture demonstrated promising results in mitigating failure modes in urban environments. An extensive review of error space propagation within the feature-based VO algorithm, as conducted in [31], reveals that error arises during feature extraction and feature matching stages between consecutive frames. These errors introduce geometric errors and photometric noise, which negatively impact overall performance. Moreover, if the

input lacks features available from the environment, the VO error compensator may struggle, leading to large outliers and resulting in a narrow error distribution with long tails. Therefore, denoising of the predicted GRU-based error compensation method is needed to reduce outlier effects utilising auto-encoder capabilities of replicating underlying data patterns, thereby enhancing robustness while reducing data uncertainty of the compensation process.

Autoencoder is the type of artificial neural network model that comprises an encoder and decoder. The purpose of autoencoders is to achieve feature recognition by learning a latent space representation of the input. Denoising autoencoders (e.g. image denoising) have been applied to VIO applications for removing IMU noises [49] and image noises [132], [133]. General architecture of autoencoders [134] is considered for this model, where the encoder compresses the input x into a compressed latent space representation h and remap it into a decoded output $\hat{x}$, can be described as-

$$h = f(Wx + b) \quad 4-1$$

$$\hat{x} = f(\hat{W}x + \hat{b}) \quad 4-2$$

Where, W and $\hat{W}$ are the weights of encoder and decoder; b and $\hat{b}$ are the bias of encoder decoder, f is an activation function. Given a training data set $\{x_i\}_{i=1}^N$, overall training process of the encoder-decoder structure yields-

$$\hat{x}_i = f(\hat{W}(Wx_i + b) + \hat{b}) \quad 4-3$$

The denoising mechanism eliminates noise by leveraging the feature learning ability of the hidden layers. When the output is noisy, it forces to removal of random noise during the learning process, thereby improving generalising ability and preventing overfitting during training. Figure 4-2 illustrates the high-level architecture of the GRU-AE error compensation mechanism. The GRU-AE model is utilized for the time series data from two input sources VO and IMU that includes position from VO, IMU linear acceleration in body frame within NED coordinate, $\alpha_N, \alpha_E, \alpha_D$ and IMU angular velocity in the body frame within NED

coordinate, $\omega_N, \omega_E, \omega_D$. The model combines both GRU layers for sequential modelling and autoencoder layers for feature extraction and reconstruction. The model predicts position error ΔP_{VO} between VO measurements and ground truth. Having the advantage of utilising the canonical relationship between inertial frame and UAV dynamics, the model can effectively learn implicitly to track the vehicle's movement by observing the sequences of IMU linear acceleration and angular velocity. By analysing dependencies between the estimated position and the actual position, the GRU-AE can learn typical error patterns and denoise abnormal signals based on reconstructed loss. Finally, the predicted denoised increments are compensated using the following equation,

$$P_{VO/GRU} = P_{VO} - \Delta P_{VO} \quad 4-4$$

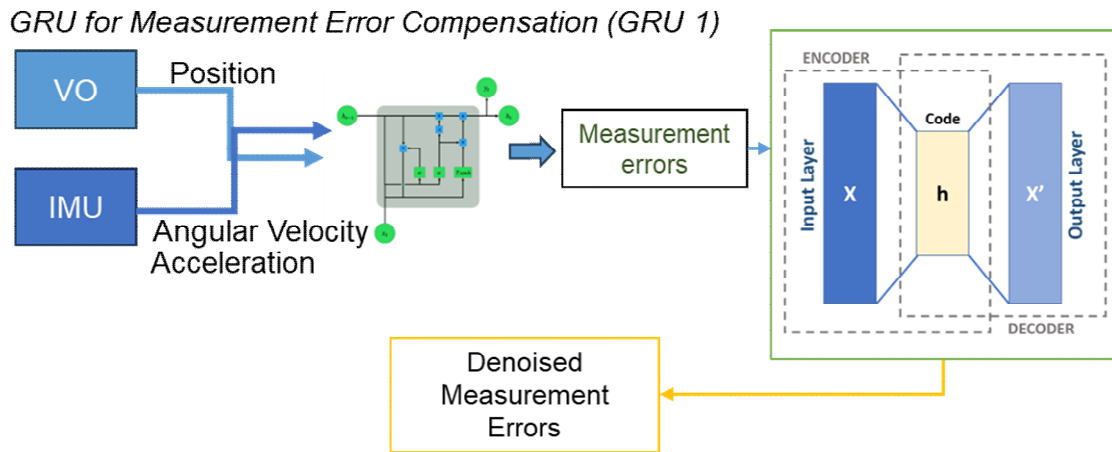


Figure 4-2 Architecture of GRU-AE VO Error Correction Model

4.2.2 GRU-based ESKF Model Error Compensation

Another GRU model is applied to compensate for measurement model errors associated with the imperfection of the model due to manual tuning and nonlinearity presence within the data. Figure 4-3 illustrates the high-level diagram of the GRU-based model compensator. The model measurement errors are compensated by predicting KF measurement increments using GRU. The ΔP_{VIO} model predicts position error between ESKF measured position and ground truth with the input parameters ESKF estimated position and innovation, i.e. the difference between the predicted observations and observed value in the NED

frame. The GRU-based model error compensator can effectively learn from the information about the state-evaluation process and uncertainty of state estimate. Finally, the predicted model increments are compensated using the following equation after fusion-

$$P_{VIO/GRU} = P_{VIO} - \Delta P_{VIO} \quad 4-5$$

GRU for Estimation Error Compensation (GRU 2)

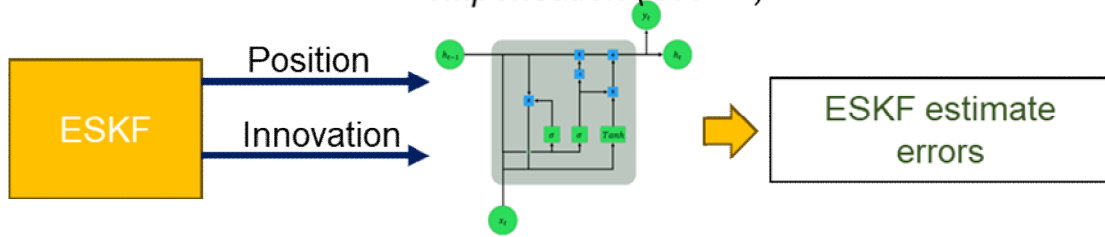


Figure 4-3 Architecture of KF Measurement Error Compensator

4.2.3 Machine Learning Model: GRU

The multi-ML hybrid VIO uses the features discussed in the previous section to reduce position uncertainty associated with incorrect sensor data and the imperfect measurement model. GRU is designed to handle sequential data and can maintain a memory of past states, making it well-suited for learning the temporal dependencies in vehicle operation, with the cell structures different from other recurrent neural networks (RNNs). It is a variant of LSTM that merges the input gates and forgetting gates along with merging the memory state and hidden state. GRU has much fewer parameters than LSTM, and the training speed is faster. The calculation process is as follows-

$$\begin{aligned} z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\ r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\ \tilde{h}_t &= \tanh(W \cdot [r_t * h_{t-1}, x_t]) \\ h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \end{aligned}$$

The greater the value of the update gate, the more information is brought in from the previous moment. The smaller the value of the reset gate, the less information from the previous moment is written.

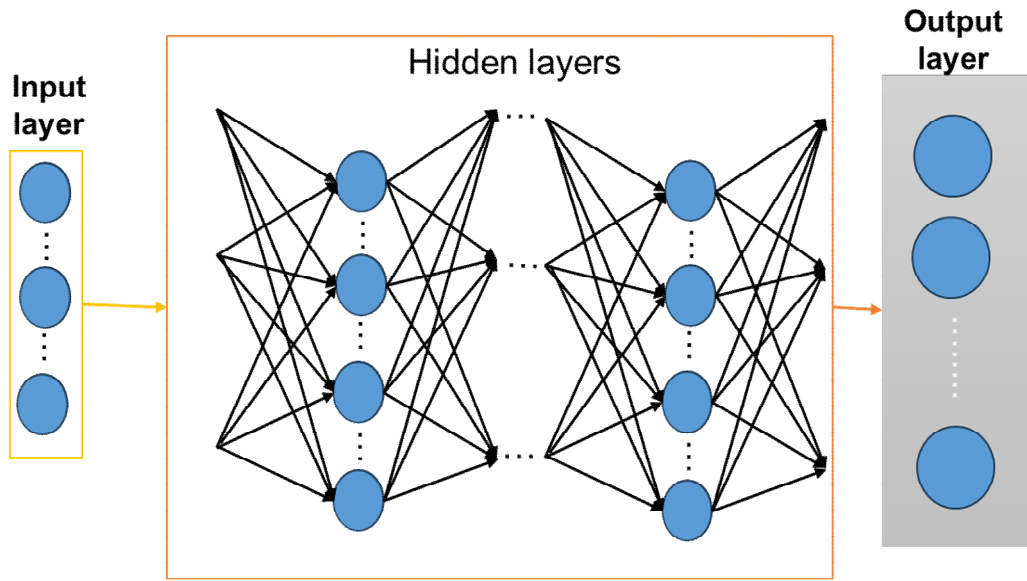


Figure 4-4 High level multi-Layer GRU Architecture used for error compensation modules in Multi-ML VIO

Recent work [126] on hybridisation has utilised a cascade of three layers of GRU to learn patterns of data effects in different observation and predicted states within a predefined Kalman filter, providing an accurate prediction of Kalman gain. However, integrating such a hybrid architecture in VIO application often struggles to minimise increased noise measurement caused by the inherent nonlinearity VIO process. This issue becomes more pronounced under high dynamics, where VO sequences and IMU data exhibit strong nonlinearity and noise. Furthermore, the forgetting problem in sequential learning architectures further complicates the process of maintaining accurate predictions over extended time dependencies and struggles with generalisation over unknown flight patterns. The hyperparameters that have the greatest impact on the performance include the number of GRU layers, the number of neurons and the step size. Too many neurons lead to overfitting and degradation of generalisation performance, while insufficient layers of neurons cannot fully extract features from input and output data. To address these challenges, this work proposes to integrate four layers of GRU for error compensation within the multi-ML hybrid VIO designed to effectively capture temporal dependencies and handle high dynamics in the system, as shown in Figure 4-4. Based on extensive analysis, four layers of GRU are sufficient to extract the hidden information. The GRU architecture, combined

with denoising mechanisms in the form of VO-AE correction, is deigned to integrate time-dependent patterns in IMU data and VO measurements. By learning these sequential dependencies through connected layers, it aims to mitigate the induced noise and spikes in the predicted data, thereby enhancing the generalisation ability of the system. In contrast, the GRU-aided KF correction focuses on leveraging time-dependent patterns in ESKF measurements and innovation. Using the same architecture, it predicts increments that compensate for errors in the predefined ESKF model, ensuring robust performance in dynamic and noisy environments.

4.2.4 Multi-ML-aided Hybrid VIO

The proposed novel architecture is designed with a combination of multiple error corrections and ESKF to estimate improved positioning performance under dynamic conditions. Two error corrections are utilised for different purposes within the architecture. The ESKF formulation is based on the previous work. In this formulation, the UAV state x is defined as-

$$x = [p \ v \ q \ \alpha_b \ \omega_b]^T \quad 4-6$$

Where, $p = [p_N, p_E, p_D]$ - represents the position in the platform; $v = [v_N, v_E, v_D]$ - platform's velocity; $q = [q_N, q_E, q_D]$ - represents orientation in the form of quaternion; $\alpha_b = [\alpha_{b_N}, \alpha_{b_E}, \alpha_{b_D}]$ - represents accelerometer drift and $\omega_b = [\omega_{b_N}, \omega_{b_E}, \omega_{b_D}]$ - gyroscope drift accordingly.

The main concept of the ESKF fusion strategy consists of the separation of the state vector into two parts: nominal state x , and error state δx .

$$x_k = x + \delta x \quad 4-7$$

Notably, the frame of reference used to develop navigation equations for state estimation of ESKF follows the north-east-down (NED) frame format.

The raw IMU linear acceleration and angular velocity are used within the prediction step to propagate corrected state estimate and state covariance.

$$\delta \hat{x} = F_x(x, u_m) \cdot \delta x$$

4-8

$$P = F_x P_{k-1} F_x^T + F_i Q_i F_i^T$$

The transition Matrix F_x for the error state as follows-

$$F_x = \frac{\partial f}{\partial \delta x} \Big|_x = \begin{bmatrix} I & I\Delta t & 0 & 0 & 0 \\ 0 & I & -C_{ni}[a_m - a_{b_k}]_x \Delta t & -C_{ni}\Delta t & 0 \\ 0 & 0 & C_{ni}^T\{((\omega_m - \omega_{b_k})\Delta t)\} & 0 & -I\Delta t \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix} \quad 4-9$$

Where, F_x and F_i represent the Jacobians of $f()$ with respect to the error and perturbation vectors; C_{ni} represents the rotation matrix of altitude q .

Once the GRU-AE compensated corrected VO position becomes available, the observation is incorporated to continuously update the filter. The process involves calibrating accelerometer and gyroscope drifts. The observation is expressed as follows-

$$y_k = h(x_t) + v$$

Where, y_k -represents measurement vector, $h()$ - represents system state and v -represents which Gaussian noise with Covariance R .

The measurement update equations are derived as

$$K = \hat{P}_{k-1} H^T (H \hat{P}_{k-1} H^T + R)^{-1}$$

4-10

$$\delta \hat{x} = K(y_k - h(\hat{x}_t))$$

$$\hat{P} = (I - KH)\hat{P}_{k-1}$$

Finally, ESKF measurement position is corrected using equation 4-5 and estimate an improved multi-ML hybrid VIO position.

4.3 Experimental Setup and Data Generation

This section outlines the experimental setup to evaluate the performance of the proposed Multi-ML hybrid VIO. The experimental setup is described first, which is used to generate training and testing datasets to conduct numerical analysis

and performance evaluation in the presence of multiple sources of uncertainty under complex environments. The experimental setup used in this research has been implemented in MATLAB, incorporating the Unreal Engine to simulate data collection for UAVS. The simulated datasets involve diverse challenging conditions, including illumination variation, rapid motion, altitude variation, no-texture, and weather effects for shorter and extended flights. The experimental architecture, presented in Figure 4-5 is adopted from one of the previous works, which was used to implement the fault-tolerant federated multi-sensor GNSS/VO/IMU navigation system. The choice of monocular camera is beneficial for the UAM application due to its advantage of being less expensive, simpler to implement and lightweight to fit into a UAV with a specification of 1109 (pixels) focal length, 640x360 (pixels) optical centre and 720x1280 (pixels) image size, operating in 10 Hz sample rate. The simulated IMU data has the specification of the ICM 20649 model operating 100 Hz sample rate used in the previous work. During the post-processing step, linear interpolation has been carried out to synchronise timestamps between both sensors. For urban operation scenarios surrounded by buildings, the visual data of tall buildings is simulated using 'US city' scene available in the 'simulation 3D scene configuration' block within the Navigation Toolbox and UAV toolbox in MATLAB. This chapter investigated the effect of different lighting conditions in navigation environments (e.g. bright sunny, partly sunny, foggy and rainy). However, training datasets are collected by replicating weather effects and lighting conditions based extensive review of the visual environment fault tree, as shown in Figure 4-6. The setting is adopted based on the MathWorks user interface described on the website. The flight trajectories are designed and controlled through QgroundControl software connected through MAVLink connectivity.

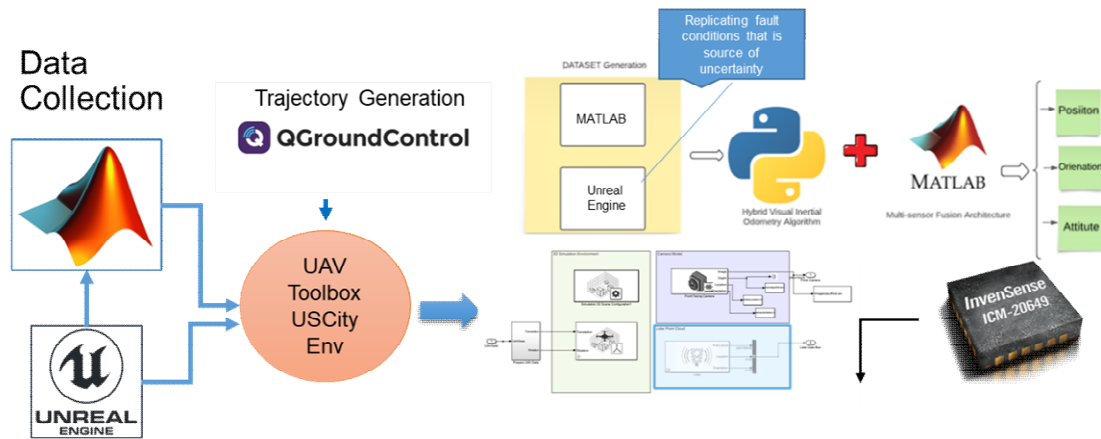


Figure 4-5 MATLAB incorporating Unreal Engine Simulation Environment for Data Correction

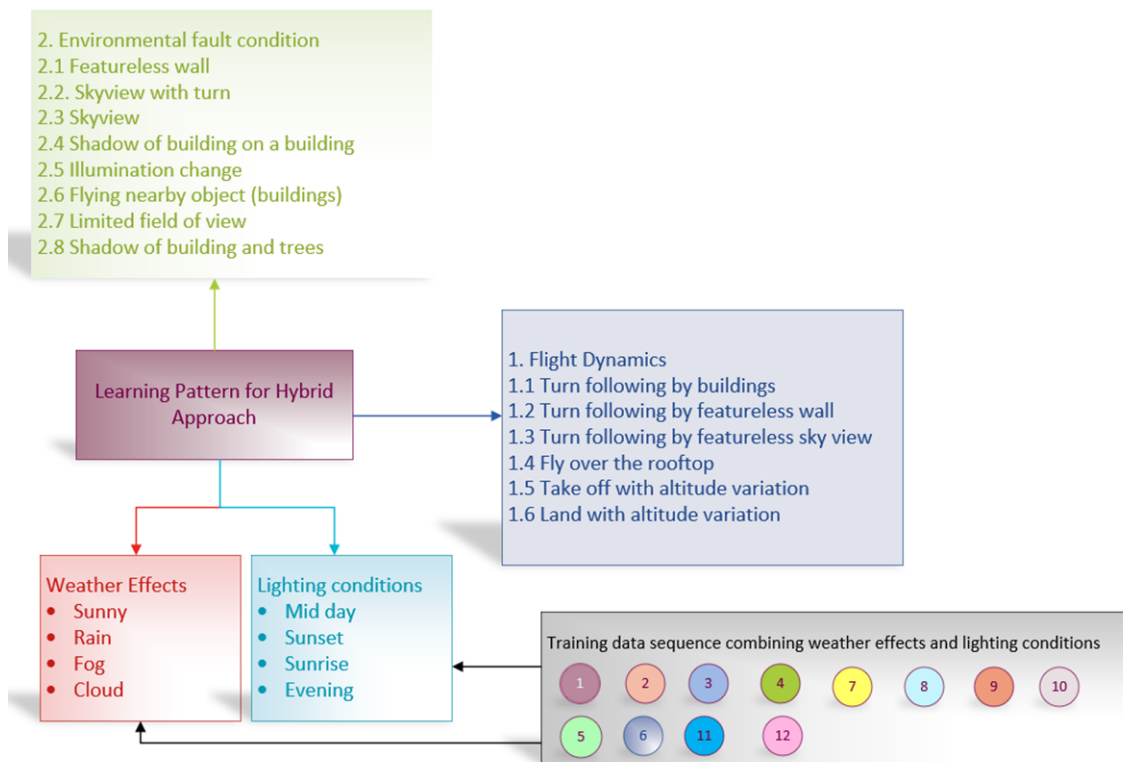


Figure 4-6 Fault Scenarios for Training Datasets

4.3.1 Training Datasets

The training dataset includes 12 trajectories combining fault conditions that can trigger multiple sources of data and model uncertainty within the navigation system, involving imperfect modelling, sensor noise, data association errors, feature tracking errors and feature location errors. Within these trajectories, a set

of various weather conditions are presented including sunny, cloud, rain, fog and various lighting conditions like dusk, sunrise, afternoon, and evening. This diversity allows the model to learn and adapt to different fault conditions, improving overall performance and generalization ability in the presence of multiple fault conditions. The testing dataset is designed with a combination of seen and unseen fault scenarios. Examples of training and testing datasets are presented in Figure 4-7.

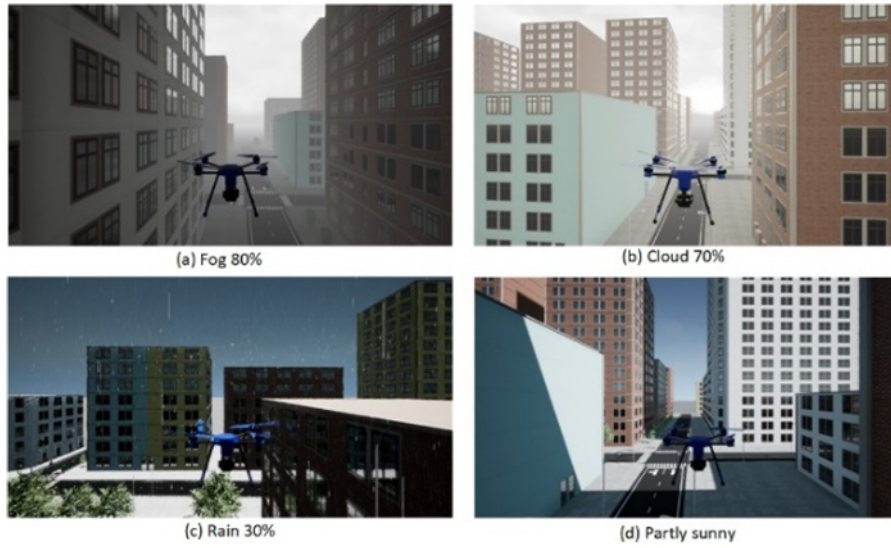


Figure 4-7 Examples of simulated environmental conditions: (a) limited field of view during 80% fog, (b) texture-less wall during 70% cloud, (c) illumination changes during 30% rain and (d) shadows and texture-less wall during partly sunny used in testing and training datasets.

4.3.2 Training and Testing Process

The ML modules are trained with the mentioned 12 trajectories in different weather and lighting conditions, forming 48 datasets (e.g. sunny, foggy, rain, midday) using the input and output data structure described in Section 4.2. Notably, this research has utilised similar GRU architectures for two different purposes within the hybrid VIO architecture. The autoencoder model that is used to denoise GRU increment prediction is trained using only sunny weather conditions for the 12 trajectories, with an assumption of good lighting conditions, affecting VO relatively less than other weather effects. To balance the training efficiency and prediction accuracy, the GRU-AE model comprised four GRU

layers containing 50, 350, 350 and 50 cells respectively, with the sequence length input, optimised through hyperparameter tuning from Keras library 'RandomSearch'. A fully connected layer is used to provide the output from the ML model. The dense layer in the autoencoder compresses the GRU output into smaller latent representation (128) with 'relu' activation function applied. The batch size used is 128, the dropout rate is found at 0.1, the activation function is optimised for each hidden layer shown on Figure 4-4. The ADAMAX optimiser is employed during the training to optimise the ML model weights and minimise the discrepancies between the predicted and actual errors. Furthermore, the loss function is calculated using mean squared error (MSE), and the learning rate is set at 0.001. In contrast, the GRU-based ESKF error compensator model is trained using 48 training datasets, described previously, incorporating weather and lighting effects. To balance training efficiency, prevent model overfitting and improve prediction accuracy, the training dataset was split into subsets to learn GRU flight manoeuvring patterns. Hyperparameters are tuned using Keras library 'RandomSearch' optimisation method with the multi-layer architecture incorporating similar architecture described previously. The loss function 'MSE' and optimiser 'ADAMAX' are set similarly to the GRU-AE model.

During the testing phase, ML models operate in the prediction phase, remaining active throughout the experiment. During this phase, the train GRU-AE error correction model predicts VO position error based on the given navigation frame linear acceleration, angular velocity, VO position and VO current orientation. For evaluation purposes, the updated position error prediction is deducted from the actual VO position estimated from the feature-based VO algorithm. Lastly, the trained GRU-based KF measurement error compensator predicts KF estimated position error based on KF predicted position and observation differences from the measurement model in the current navigation frame. The updated position is estimated by deducting the predicted position error from the KF estimated position to improve overall hybrid VIO positioning performance.

4.4 Performance Evaluation

To evaluate the performance and adaptability of the proposed novel hybrid VIO framework, six experiments are carried out using six different testing datasets (four different trajectory shapes), with a variation of fault conditions, weather and lighting effects. In addition, the learning and prediction capability of the proposed GRU-architecture has been validated based on the performance under different shapes of trajectory. Table 4-1 represents the description of testing scenarios, trajectory types and fault conditions.

Table 4-1 Test trajectory description

# Trajectory	Testing Scenario	Lighting condition	Fault type
1	Nominal Single faults	Sunny	<ol style="list-style-type: none"> 1. Shadows of objects 2. White Rooftop 3. Turn
2	Weather effects	Fog, Rain	<ol style="list-style-type: none"> 1. Turn with features-wall. 2. Turn with sky view. 3. Shadows of objects 4. Take off with different altitudes
3	Lighting Effects	Sunny, partly sunny	<ol style="list-style-type: none"> 1. Sharp turns with sky-view 2. Vibration while turning 3. Altitude variation in the flight
4	Unseen Scenario with unknown faults	Sunny	<ol style="list-style-type: none"> 1. Turns with featureless walls. 2. Turns with sky-view. 3. Shadows of objects 4. Vibration with sharp turn 5. Altitude variation 4. Take-off with different altitudes.
5	Lighting Effects	Sunrise	<ol style="list-style-type: none"> 1. Altitude variation among waypoints 2. Turn with sky-view 3. Vibration with waypoint changes 4. Take-off from unknown location
6	Lighting Effects	Sunrise	<ol style="list-style-type: none"> 1. Sharp and curved turns 2. Turn with sky-view 3. Vibration with waypoint changes 4. Feature-less wall 5. Take-off from unknown location

Figure 4-8 represents the testing trajectory to evaluate the proposed approach with the denoising mechanism performance in terms of robustness under different weather conditions and adaptability over unknown scenarios and fault factors. Three types of testing scenarios (e.g. weather effects, unknown scenarios with unknown faults, lighting effects) have been selected to evaluate

performance enhancement with position uncertainty minimization using denoising based hybrid VIO-

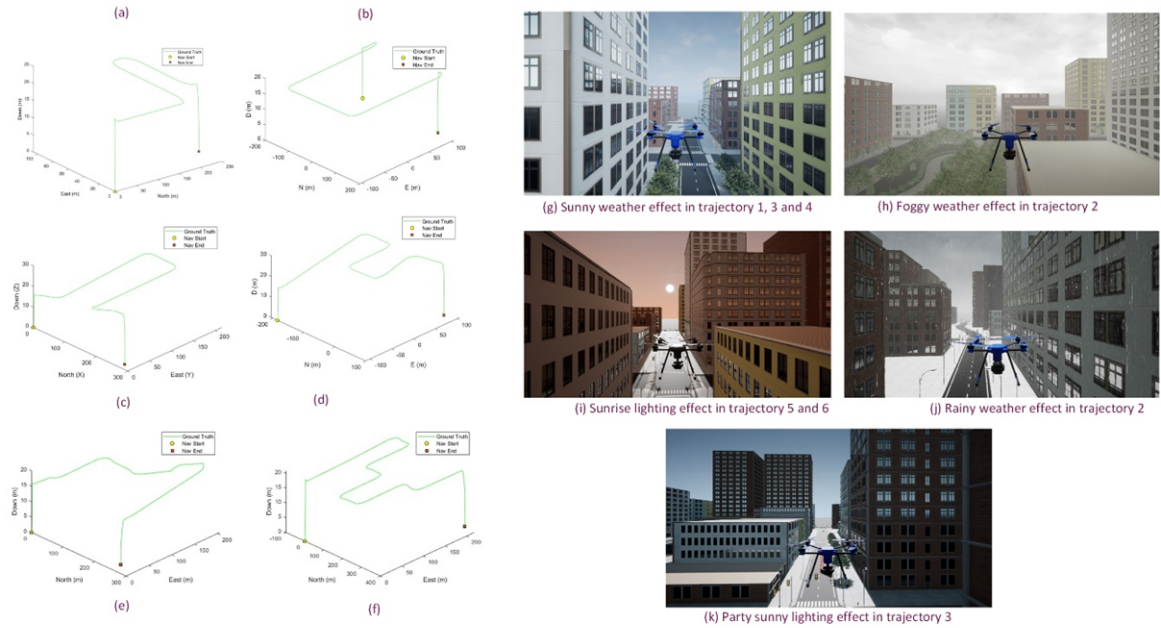


Figure 4-8 3D view of testing trajectories for experiments (a) Trajectory 1, (b) Trajectory 2, (c) Trajectory 3, (d) Trajectory 4, (e) Trajectory 5, (f) Trajectory 6

4.4.1 Proposed Multi-layer GRU-architecture based on Performance Evaluation under different trajectory shapes

The experiment aims to assess the effectiveness of utilising multi-layer sequential GRU for error compensation within Mult-ML hybrid VIO. This approach is designed to enhance performance and generalisation ability, particularly under complex flight patterns. To evaluate its effectiveness, three tests are conducted on trajectory 1, trajectory 2, trajectory 3 and trajectory 4, which is selected for their diverse flight characteristics such as narrow shapes, altitude variation between flights, and multiple maneuvers. In order to evaluate the impact on performance, five sets of Multi-ML hybrid VIO models were created using the similar training methodology described in section 4.3.2.

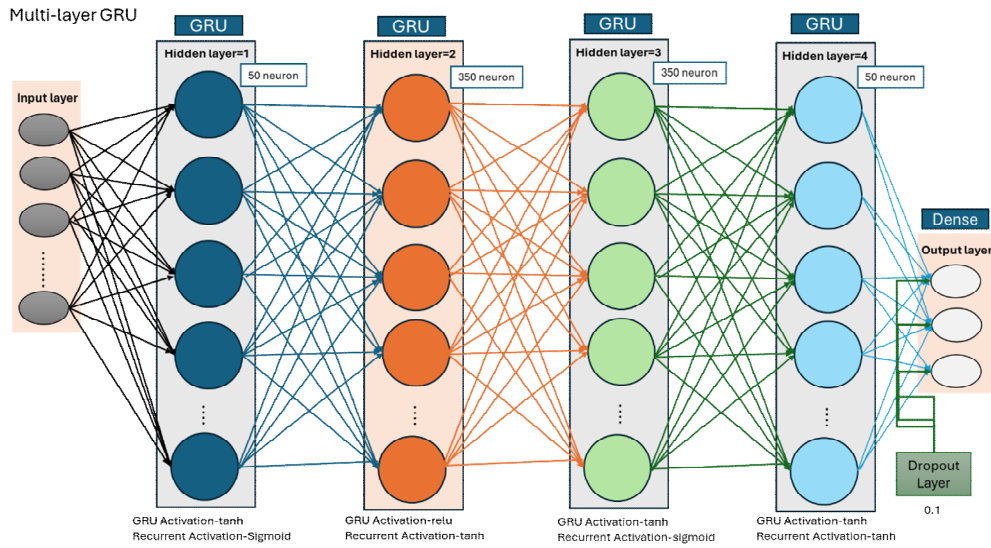


Figure 4-9 4-layer GRU architecture

Figure 4-9 illustrates the detailed architecture of the proposed 4-layer GRU used for error correction in this chapter. This has been designed based on extensive analysis using the testing trajectories 1-4, with the number of layers determined by the model's effectiveness in error compensation and mitigation of failure modes under different fault conditions like lighting variation, flight dynamics and low-texture environments. The hyperparameters such as activation functions, neurons and dropouts have been selected through a systemic optimization approach using 'RandomSearch' method as stated in Section 4.3.2. This method explores the activation function space during training to identify the most effective configuration for capturing nonlinear dynamics and high motion variability present in the training datasets.

Figure 4-10 illustrates horizontal error comparison across the testing trajectories for the Multi-ML hybrid VIO system, analysing the influence of varying the number of GRU layers. The comparison was used to optimise the GRU architecture and identify the optimal layer for effective error correction. The results indicate that increasing layers in both GRU-based error compensations significantly improve overall performance by reducing large outliers, particularly in the landing phases. Furthermore, the proposed multi-layer architecture effectively minimizes the effects of different flight dynamics (e.g. curved turns and altitude variation) where

single layers, 2-layers and 3 layers GRU architecture struggle to maintain stability.

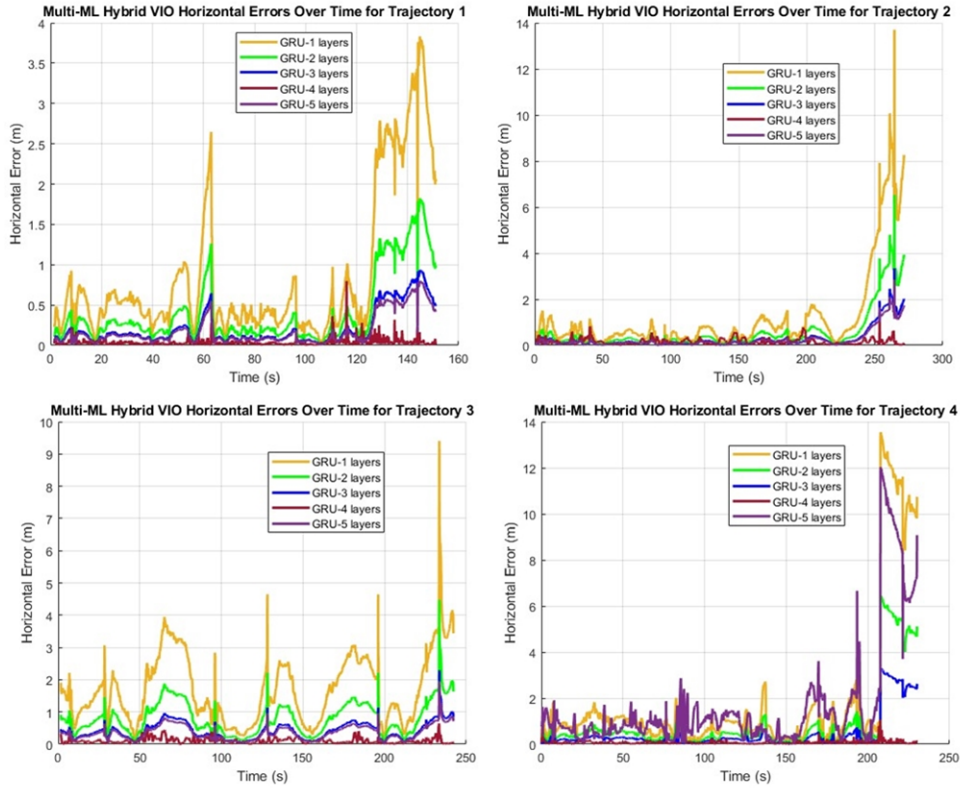


Figure 4-10 Horizontal error comparison highlighting the impact of incorporating additional layers in GRU architecture within Multi-ML hybrid VIO

4.4.2 Nominal Fault Conditions

The experiment aims to evaluate the performance of the proposed hybrid algorithm and its ability to reduce uncertainty while navigating in good weather conditions with single fault factors (e.g. shadows of buildings, smooth turns, flying above rooftops) in an urban environment. During the test, the UAV encounters one turn long between buildings and one curved turn over the rooftop, with two waypoint changes at 60s and 110s, including motion variation that results in feature extraction error and feature location errors. These faults are identified as a “source of uncertainty” in the input data that cause performance deviations in feature-based VO algorithm. The experiment covers a distance of 702m in a duration of 152 s.

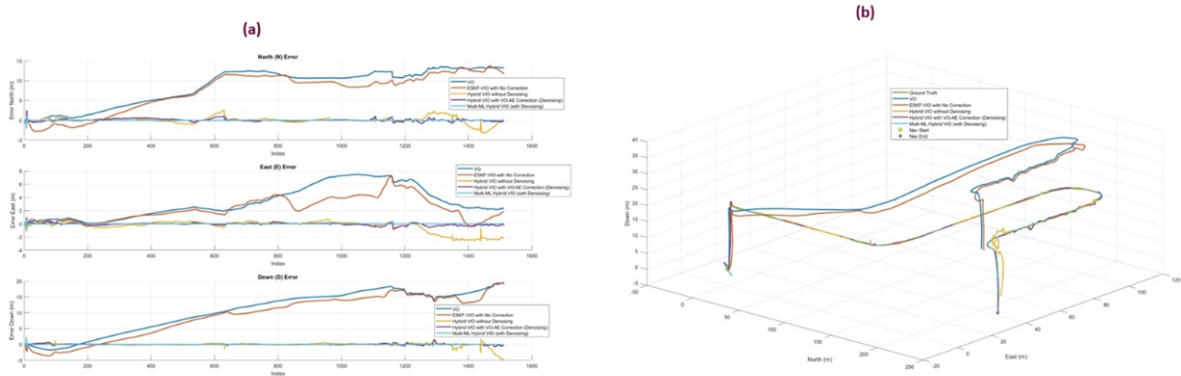


Figure 4-11 (a) Position error plot in the presence of single fault condition for trajectory 1; (b) 3D trajectory comparison of the proposed multi-ML hybrid VIO framework with four reference systems for trajectory 1.

Figure 4-11 illustrates position error and 3D trajectory comparison based on the proposed Multi-ML hybrid VIO, Hybrid VIO with VO-AE Correction, Hybrid VIO without denoising, ESKF VIO with no correction and VO. By the end of the flight, after 110s, the errors in N, E, D positions are observed to be higher. It is worth noting that the performance degradation is attributed to fault conditions encountered while flying above the rooftop, where less features are available for extraction. These conditions lead to feature extracting error, feature location error and feature tracking error, identified as the source of aleatoric or data uncertainty. These challenges result in unmodelled outliers and increased nonlinearity, which ESKF VIO with no correction and hybrid VIO without the denoising mechanism fail to mitigate due to induced outliers affecting the pre-defined ESKF model. This degradation cannot be mitigated by compensating either sensor or model measurements alone due to multiple sources of uncertainty impacting both. Consequently, model and data uncertainty arise simultaneously, impacting performance. Table 4-2 represents RMSE comparison for the experiment and showed that above mentioned fault conditions is mitigated with improvement of 99% in the proved multi-ML hybrid VIO, 97% in Hybrid VIO with VO-AE correction, and 90%% in Hybrid VIO with no denoising across 3D position. It has been observed from Figure 4-11(b) that the proposed algorithm with uncertainty correction and denoising mechanism demonstrated higher performance

enhancement, following similar trajectory path to ground truth, while other reference systems resulted in a few meters deviation from the ground trajectory.

Table 4-2 RMSE and 95th percentile Comparison on the performance of single fault conditions for trajectory 1

Algorithms	RMSE (m)					95th percentile (m)				
	N	E	D	Horizontal	3D	N	E	D	Horizontal	3D
VO	9.72	4.04	12.37	10.55	16.29	13.31	7.31	19.01	13.94	23.65
ESKF VIO with No Correction	8.81	2.87	11.21	9.27	14.55	12.77	5.57	17.06	12.83	21.43
Hybrid VIO without Denoising	0.84	0.91	0.68	1.27	1.47	1.69	0.43	0.48	2.88	3.21
Hybrid VIO with VO-AE Correction	0.31	0.18	0.19	0.38	0.45	0.57	0.19	0.35	0.68	0.78
Multi-ML Hybrid VIO (Denoising)	0.09	0.05	0.07	0.11	0.14	0.06	0.036	0.046	0.11	0.24

Figure 4-12 illustrates error distribution comparison of horizontal errors, providing insight into the ability of denoising corrupted VO signals in feature-less environments and reducing position uncertainty, while improving overall performance through the proposed framework. The distribution of VO and ESKF VIO with no correction showed a wide error distribution, indicating significant position uncertainty. This challenge has been mitigated by 79% with VO-AE correction in hybrid VIO compared to VO. However, for better performance enhancement and reducing the effects of sources of uncertainty, both sensor and model measurements need to be considered. The integration of both VO and KF correction within hybrid VIO simultaneously reduces the model and data uncertainty, resulting in minimisation of outliers. As shown in Table 4-2, the representation of 95th percentile error decreases from 23.65 m to 0.24 m in 3D axis,. These results demonstrate the proposed solution ability to address the long tails issue, enhancing performance under different sources of uncertainty. This

approach provide robust safety assurance boundaries, ensuring reliable operation under nonimal fault conditions.

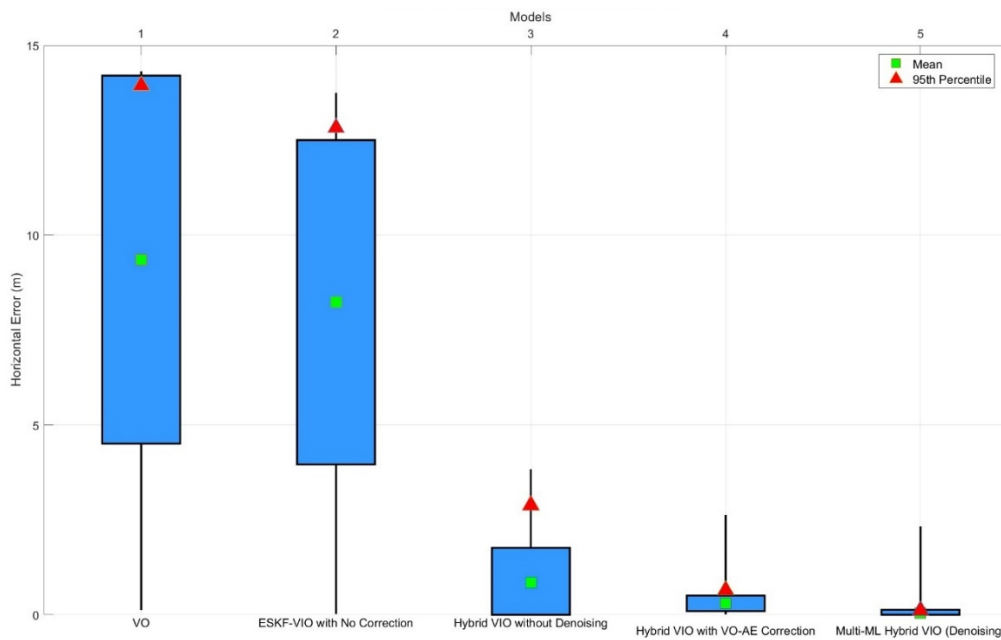


Figure 4-12 Horizontal error distribution comparison for trajectory 1

4.4.3 Weather effects

In order to analyse the robustness of the proposed hybrid VIO algorithm with position uncertainty reduction under combinational fault faction influenced by different weather effects, 5 mins flight covering 966m is conducted within a location familiar to the trained models. However, the fault factors influencing the system are unknown to the models. This test aims to analyse the algorithm's performance in terms of position uncertainty reduction and to assess how the fault factors contribute to the overall performance, while other aspects, such as weather conditions and flight path, remain within the scope of the trained model.

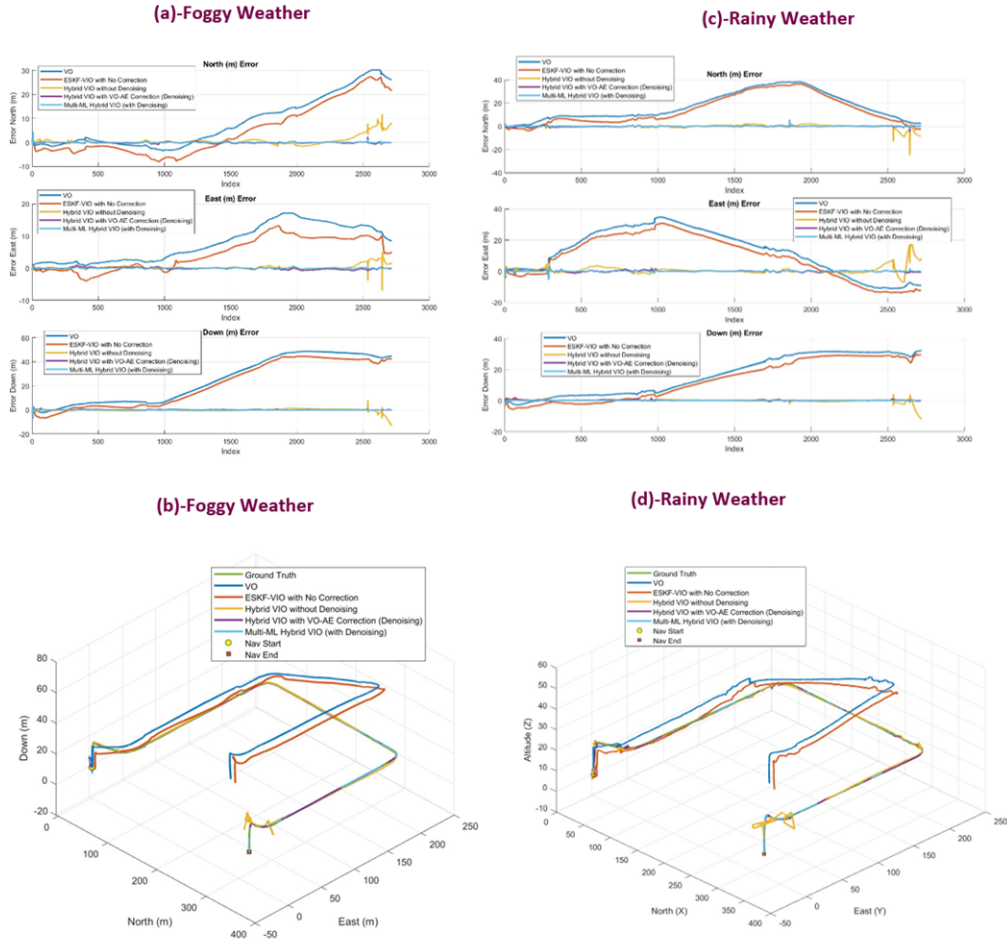


Figure 4-13 Position error and 3D trajectory comparison under weather effects for trajectory 2 (a) Position error under foggy; (b) Position error under rainy; (c) 3D trajectory under foggy; (d) 3D trajectory under rainy

Figure 4-13 illustrates the position error and 3D trajectory comparison across trajectory 2 under foggy and rainy weather conditions. The proposed solution demonstrates significant performance improvement in minimising noise artefacts such as raindrops, lens flare, reflection, and limited visibility during the flights. For example, adverse weather conditions (e.g. limited visibility in foggy) introduce additional uncertainties in data, such as feature tracking error, data association error, where ESKF VIO RMSE horizontal error reaches 13.33m, and hybrid VIO without denoising reduces to 2.13m. Integrating denoising mechanism in hybrid VIO system effectively such degradation and achieves 23 cm horizontal RMSE, enhancing performance while minimising the effects of data and model uncertainty. Similarly, during rainy weather conditions, horizontal positioning error

risers to 24.82 m in ESKF-VIO and 3.29m in hybrid VIO (without denoising). The performance degradation in ESKF-VIO, hybrid VIO (without denoising) and hybrid VIO with VO-AE correction cannot be mitigated by compensating either sensor or model measurements alone due to multiple sources of uncertainty impacting both. Additionally, the hybrid VIO (without denoising) showed less improvement compared to the hybrid VIO-AE correction due to additional noise remaining in the data. This suggests that if data sources are corrupted and noisy, it negatively impact model performance. Furthermore, without sufficient knowledge about the unmodeled dynamics in KF, the KF error compensation ML model is less effective in learning the fault patterns, resulting in poor performance compared to the VO error compensation technique only. Notably, hybrid VIO with VO-AE correction and Multi-ML hybrid VIO closely follow the flight trajectory, providing insight into the system robustness while flying in adverse weather conditions with high flight dynamics.

Table 4-3 RMSE and 95th percentile comparison on the performance under foggy weather effects for trajectory 2

Algorithms	RMSE(m)					95 th Percentile(m)				
	N	E	D	Horizontal	3D	N	E	D	Horizontal	3D
VO	13.14	9.57	30.96	16.21	34.95	48.26	16.94	28.32	30.96	56.09
ESKF VIO with No Correction	11.53	6.62	28.02	13.33	31.56	46.61	12.72	23.94	25.76	50.61
Hybrid VIO without Denoising	1.98	0.78	1.49	2.13	2.35	5.82	2.10	1.12	6.34	7.51
Hybrid VIO with VO-AE Correction	0.28	0.30	0.31	0.41	0.58	0.32	0.36	0.42	0.70	0.84
Multi-ML Hybrid VIO (Denoising)	0.19	0.12	0.21	0.23	0.32	0.31	0.20	0.19	0.50	0.64

Table 4-3 and Table 4-4 summarise the performance proposed framework compared to reference systems concerning RMSE and 95th percentile for trajectory 2. It is observed that the multi-ML hybrid VIO outperforms compared to the fusion reference systems. Horizontal RMSE is improved by 98% under foggy and rainy weather conditions for Multi-ML hybrid VIO compared to VO. In contrast, the Horizontal 95th percentile is improved by 99% under foggy and rainy weather conditions for the proposed solution compared to ESKF-VIO. Based on reviewing the effects of adverse weather conditions on flight dynamics, the proposed solution demonstrates robustness capabilities with a reduction of

uncertainty, while performance in terms of accuracy is higher in nominal fault conditions.

Table 4-4 RMSE and 95th percentile comparison on the performance under rainy weather effects for trajectory 2

Algorithms	RMSE(m)					95 th Percentile(m)				
	N	E	D	Horizontal	3D	N	E	D	Horizontal	3D
VO	20.54	19.01	20.48	27.99	34.69	36.22	10.50	24.09	39.57	43.47
ESKF VIO with No Correction	18.54	16.50	18.17	24.82	30.76	34.45	6.79	20.42	36.25	39.01
Hybrid VIO without Denoising	1.81	2.70	1.27	3.29	3.45	1.86	0.61	0.60	7.27	6.64
Hybrid VIO with VO-AE Correction	0.43	0.51	0.32	0.67	0.74	0.51	0.70	0.63	1.12	1.95
Multi-ML Hybrid VIO (Denoising)	0.27	0.36	0.09	0.43	0.45	0.35	0.40	0.06	0.69	0.85

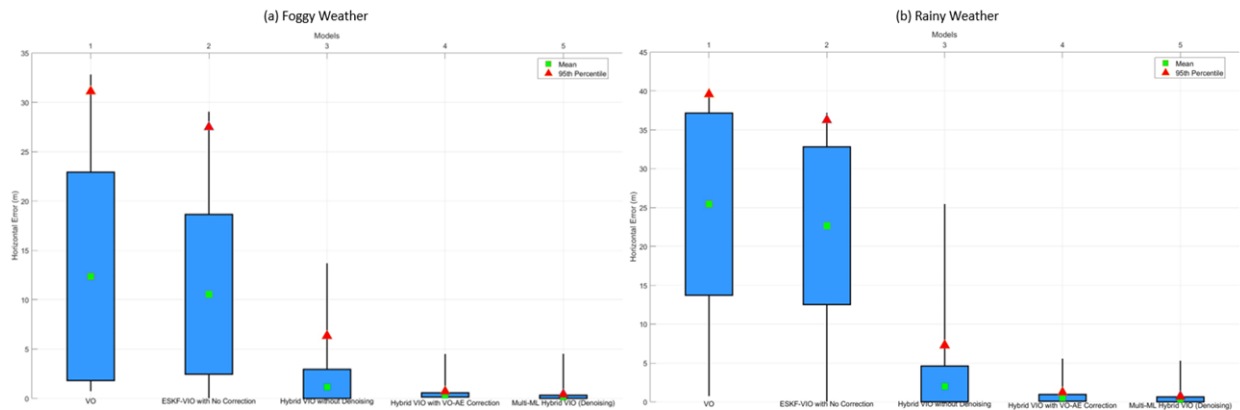


Figure 4-14 Error distribution under weather effects for trajectory 2 (a) Foggy weather; (b) Rainy weather

Figure 4-14 represents the horizontal error distribution in foggy and rainy weather conditions using box plots. This summarises the performance of different reference systems along with the proposed solution, providing insight into their ability to improve performance and reduce outliers, thereby addressing uncertainties in adverse weather conditions. It is observed that the proposed solution outperforms in both scenarios, providing the best performance by narrowing down the distribution and is capable of reducing uncertainty and noise effectively. Hybrid VIO without denoising improved performance, but noise artefacts from rain are not fully mitigated, while hybrid VIO with VO-AE correction

minimised rain-induced noise, improving accuracy and lowering the 95th percentile.

4.4.4 Lighting Effects

In this section, two types of experiments are conducted to analyse the impact of additional lighting effects under environmental and flight dynamics fault conditions. The results demonstrate that the denoising mechanism within hybrid VIO achieves the best performance by effectively reducing noise caused by various light effects. This approach enhances performance while reducing outliers, with error distribution remaining within the 95th percentile confidence interval and exhibiting no long tails.

4.4.4.1 Lighting Effects on the same trajectory

The purpose of this experiment is to evaluate the performance, robustness and resilience abilities of the proposed algorithm in managing combined fault conditions with variation of lights (e.g. bright sunlight, shadows) and mitigation of the arisen source of uncertainty associated with trajectory 3. Trajectory 3 is composed of environmental and flight fault factors such as featureless building, turn with sky-view, and altitude variation during flight. During the experiment distance of 935m is covered over a duration of 4.2 minutes. This experiment is conducted different time zone on a sunny day.

Figure 4-15 illustrates position error and 3D trajectory comparisons for the proposed solution against four reference systems under two scenarios of sunny days: partly sunny (shadows) and bright sunny (sunny) conditions. Throughout the flight, fault factors such as take-off location and altitude variation remain unknown to the model, as they are included in training datasets. The proposed solution demonstrates superior performance by mitigating fault conditions arising from possible sources of uncertainty (e.g. feature tracking errors due to altitude variation in the flight). It achieves enhancement of 98% under sunny conditions and 99% under shadow effects compared to ESKF VIO with no correction applied. The comparison between hybrid VIO without denoising and hybrid VIO with denoising indicates that applying the denoising mechanism to VO

increments significantly improves robustness and performance under more complex lighting conditions, as shown on Figure 4-14 (b) and (d). For instance, after 180s hybrid VIO without denoising algorithm struggles to maintain accuracy, introducing large deviations in NED coordinates under bright sunny conditions. In contrast, hybrid VIO with denoising mechanism mitigates these errors, achieving the reduction of 93%, 95% and 97% in N, E, D coordinates, respectively. Furthermore, under shadow effects, both hybrid VIO without denoising and hybrid VIO with VO-AE correction show reduced performance due to illumination variation throughout flight. These variations adversely affect sensor data and the predefined ESKF model prediction, introducing additional uncertainties in the systems. In contrast, the Multi-ML hybrid VIO (denoising) solution mitigates these additional data and model uncertainties, achieving a 99% improvement in 3D position.

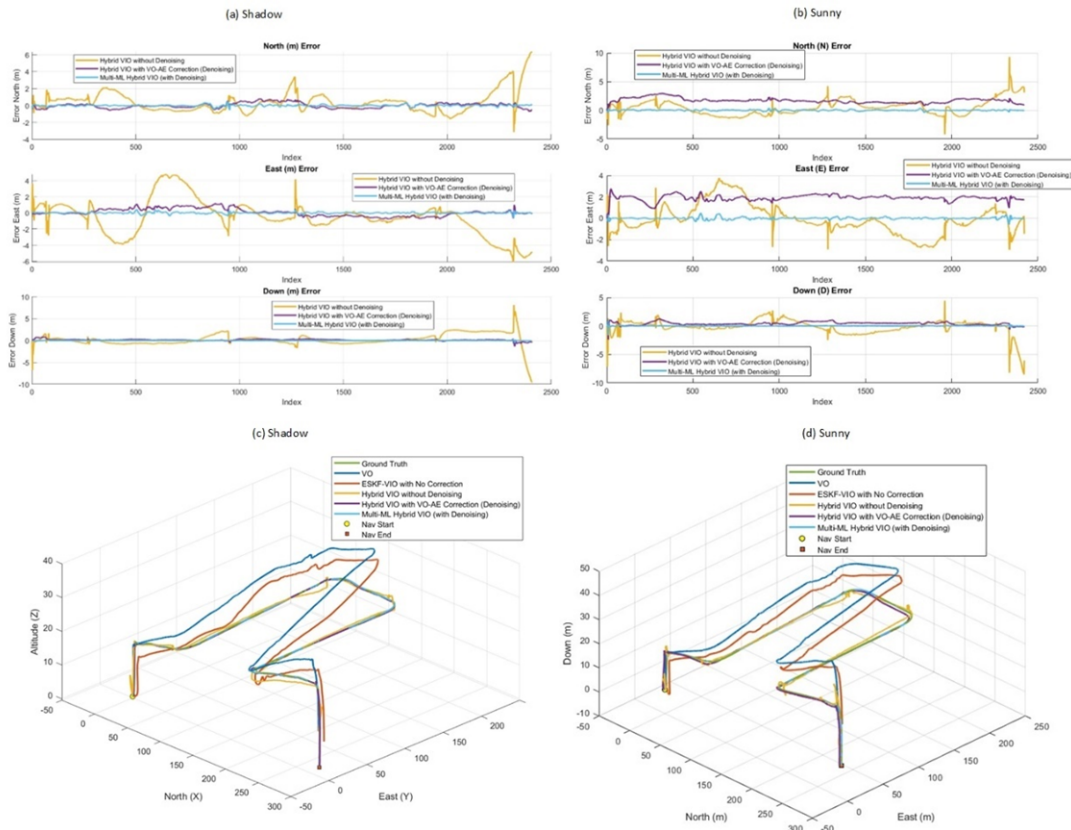


Figure 4-15 Position error and 3D trajectory comparison under lighting effects for trajectory 3 (a) Position error under shadow; (b) Position error under sunny; (c) 3D trajectory under shadow; (d) 3D trajectory under sunny

Table 4-5 RMSE and 95th percentile comparison on the performance under lighting effects for trajectory 3 summarises the performance of the proposed framework concerning RMSE and 95th percentile for trajectory 3. The results indicate that denoising of sensor output within the hybrid VIO architecture demonstrates superior performance in both horizontal and 3D position. This approach effectively minimises drift issue in visual navigation, achieving accuracy below sub-meter with VO-AE correction alone and 5 cm accuracy when multiple corrections are applied, even under varying lighting conditions along in similar flight path, with high confidence (95th percentile).

Table 4-5 RMSE and 95th percentile comparison on the performance under lighting effects for trajectory 3

Algorithms	Shadow									
	RMSE(m)					95 th Percentile(m)				
	N	E	D	Horizontal	3D	N	E	D	Horizontal	3D
VO	4.03	12.89	9.12	13.51	16.30	8.34	20.79	16.40	20.91	24.16
ESKF VIO with No Correction	2.84	10.55	6.19	10.93	12.56	5.04	17.91	12.45	18.14	22.85
Hybrid VIO without Denoising	1.29	2.23	1.46	2.57	2.96	2.96	3.98	0.31	5.35	7.58
Hybrid VIO with VO-AE Correction	0.30	0.46	0.21	0.55	0.58	0.55	0.86	0.31	0.88	0.96
Multi-ML Hybrid VIO (Denoising)	0.09	0.09	0.02	0.13	0.14	0.15	0.14	0.01	0.28	0.28
Algorithms	Sunny									
	RMSE(m)					95 th Percentile(m)				
	N	E	D	Horizontal	3D	N	E	D	Horizontal	3D
VO	7.19	6.74	12.39	9.86	15.83	11.45	18.51	16.40	15.29	23.49
ESKF VIO with No Correction	5.40	5.45	9.07	7.67	11.88	9.54	8.74	14.77	11.42	18.51
Hybrid VIO without Denoising	1.40	1.39	1.30	1.98	2.37	3.11	2.92	1.57	3.47	3.69
Hybrid VIO with VO-AE Correction	1.76	1.91	0.48	2.60	2.64	2.67	2.28	0.86	3.34	3.45
Multi-ML Hybrid VIO (Denoising)	0.09	0.09	0.05	0.13	0.14	0.14	0.14	0.04	0.28	0.28

Figure 4-16 represents error distribution comparison of horizontal position under different lighting scenarios for the same shape of trajectory. It is observed that the error distribution for both hybrid VIO with VO-AE correction and the proposed hybrid VIO remain within 95th percentile error boundary with no long outliers. This demonstrates the effectiveness of denoising mechanisms in handling variation of lighting scenarios such as shadows, glare, and sudden brightness changes, which can introduce noise and artefacts in sensor data. By filtering predicted VO

increments using the autoencoder, the proposed solution enhances position accuracy and robustness, even under challenging lighting conditions.

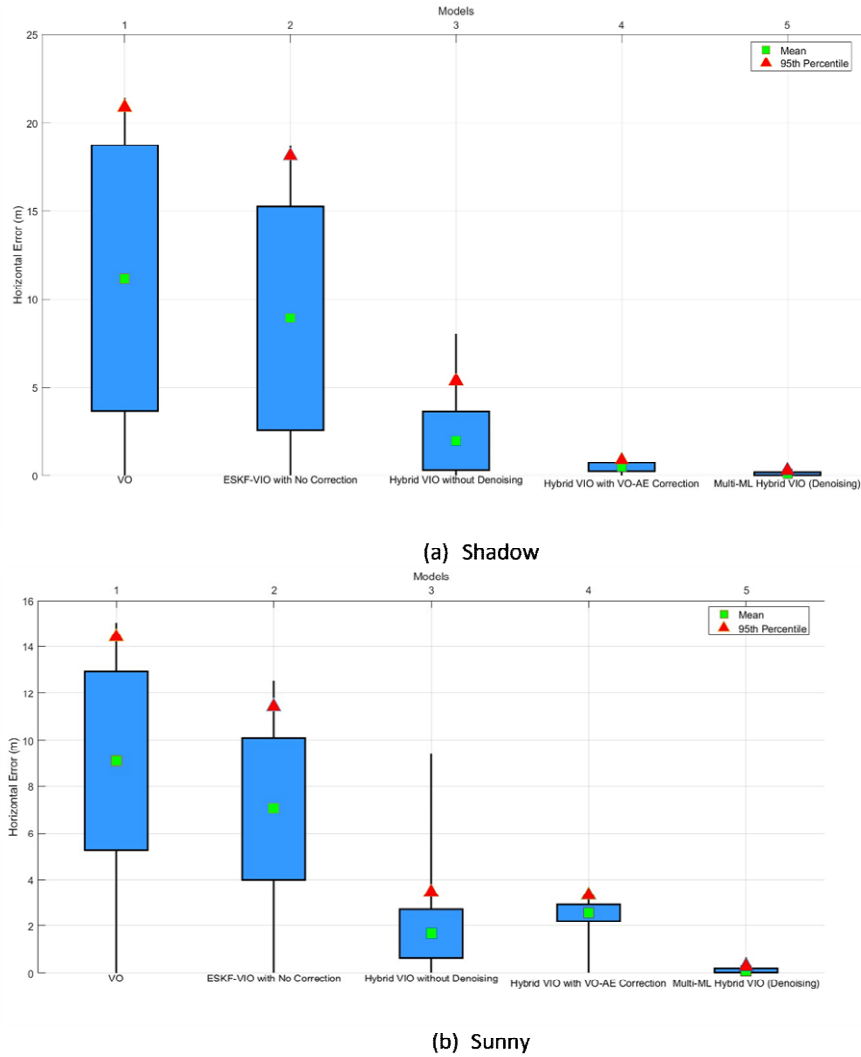


Figure 4-16 Error distribution under lighting effects for trajectory 2 (a) shadow, (b) sunny

4.4.4.2 Lighting Effects on Different Trajectory Shapes

The objective of this experiment is to evaluate the generalisation ability of the proposed multi-ML hybrid VIO framework in managing lighting effects for different trajectory shapes with seen and unseen flight take-off locations with altitude variation within the flight. The testing datasets include diverse scenarios, including UAVS taking off from different streets in the city environments with

variation of unknown fault conditions, i.e. altitude variation during flights between waypoints. The experiment is conducted in sunrise light conditions, covering 858 m with 3.40 minutes for trajectory 5 and 1349 m with 6 minutes for trajectory 6.

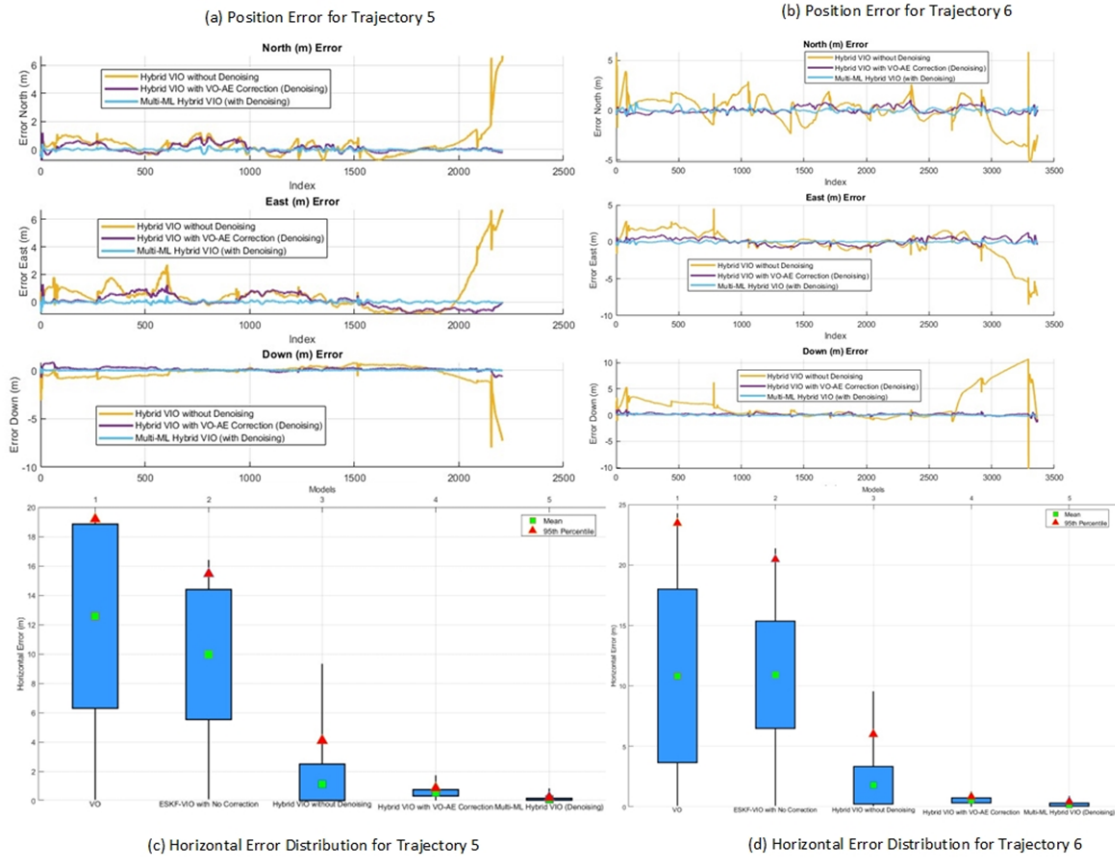


Figure 4-17 Position error and horizontal error distribution comparison for different shape of trajectory under sunrise light conditions (a) position error for trajectory 5; (b) position error for trajectory 6; (c) horizontal error distribution for trajectory 5; (d) horizontal error distribution for trajectory 6.

Figure 4-17 (a) & (b) illustrates compares position error for each coordinate and horizontal error distribution for trajectories 5 and 6 under sunrise lighting effects. Noticeable performance deviations are observed during landing phrases in each coordinate, due to the navigation environment variability and presence of unknown fault conditions in trajectory 5 when using hybrid VIO without denoising. This results in long outliers in the horizontal error distribution as presented in Figure 4-17. However, the hybrid architectures incorporating the denoising mechanism effectively mitigate noise caused by various fault factors such as

glare, low-light, shadow, reflection and overexposed areas. This demonstrates their adaptability and robustness under sunrise lighting effects for both trajectories by narrowing down the error distribution with no tails. Additionally, rapid lighting variation and shadows introduce misalignments in tracking visual features, particularly in complex trajectory shapes with additional flight dynamic fault factors (e.g. turn with vibrations, turn with altitude variation). These challenges exacerbate errors in sensor corrections for hybrid VIO without denoising, as shown from Figure 4-17 (b). The proposed denoising enables Multi-ML hybrid VIO to successfully mitigate these issues by reducing induced data and model uncertainty, ensuring high accuracy at the centimetre level and reliability by providing a horizontal error boundary within the 95th percentile under sunrise lighting conditions.

Table 4-6 RMSE and 95th percentile comparison on the performance under sunrise light conditions for trajectories 5 and 6

Trajectory 5										
Algorithms	RMSE(m)					95 th Percentile(m)				
	N	E	D	Horizontal	3D	N	E	D	Horizontal	3D
VO	5.65	9.31	15.07	14.04	22.87	12.45	16.70	26.79	19.17	32.94
ESKF VIO with No Correction	5.65	9.30	15.07	10.89	18.56	9.61	13.85	23.12	15.44	27.78
Hybrid VIO without Denoising	1.00	1.45	0.97	1.76	2.01	1.21	4.03	0.55	4.08	4.24
Hybrid VIO with VO-AE Correction	0.29	0.48	0.21	0.56	0.61	0.63	0.80	0.31	0.84	0.91
Multi-ML Hybrid VIO (Denoising)	0.08	0.07	0.02	0.11	0.11	0.11	0.11	0.03	0.20	0.20
Trajectory 6										
VO	5.95	11.53	34.01	12.98	36.41	13.92	6.59	61.97	23.45	66.36
ESKF VIO with No Correction	5.83	10.23	30.83	11.78	33.01	12.78	10.89	58.78	20.45	62.41
Hybrid VIO without Denoising	1.53	1.80	3.46	2.26	4.19	1.79	1.94	8.71	6.00	10.41
Hybrid VIO with VO-AE Correction	0.33	0.48	0.26	0.57	0.63	0.60	0.74	0.47	0.83	0.89
Multi-ML Hybrid VIO (Denoising)	0.19	0.12	0.10	0.22	0.24	0.28	0.17	0.16	0.42	0.45

Table 4-6 compares the performance of the proposed framework compared to the reference algorithms concerning RMSE and 95th percentile for trajectories 5 and 6. Difficulties arise from changing of waypoint during the flight, which leads

to increased vibrations and feature tracking errors. The comparison highlights the significant influence of flight dynamics on the performance of evaluated systems, including VO, ESKF VIO with no correction, hybrid VIO without denoising, hybrid VIO with VO-AE correction and Multi-ML hybrid VIO (denoising). Errors in 3D position increase to 22.87 m for trajectory 5 and 36.42 m for trajectory 6 using feature-based VO, reflecting the impact of dynamic conditions on feature tracking and feature extraction steps. Notably, hybrid VIO without a denoising mechanism struggle to mitigate drift error in trajectory 6 due to induced additional noise by high dynamics, underscoring the importance of denoising mechanism for maintaining robustness and reducing uncertainty in such conditions. The 95th percentile error for trajectory 5 is reduced from 32.94m to 0.57m using hybrid VIO with VIO-AE correction and further reduced to 0.24m by compensating for ESKF model errors within Multi-ML hybrid VIO in 3D position compared to feature-based VIO. Similarly, for trajectory 6, the 95th percentile error decreases from 6636m to 0.89m with VO-AE correction and is further reduced to 0.45m with Multi-ML hybrid VIO compared to VO. These results provide strong evidence of the robustness and enhanced generalisation ability of the proposed Multi-ML, which effectively handles the simultaneous challenges of complex flight dynamics and varying lighting conditions in complex urban environments, ensuring high accuracy and reliability. Thus, the proposed approach ensures high confidence in mitigating fault factors while enhancing integrity, even with corrupted sensor measurements and complex flight dynamics.

4.4.5 Unseen scenario with unknown fault conditions

The objective of this experiment is to evaluate the generalisation ability of the proposed multi-ML hybrid VIO framework in managing combined fault conditions with seen and unseen scenarios, including unseen fault conditions. The testing datasets include diverse scenarios, including UAV takeoffs and landings from different streets in the city environments with variation of unknown fault conditions, i.e. rotation with vibration during landing and altitude variation during turn. The experiment is conducted under sunny weather conditions, covering a distance of 738 m within 4 min.

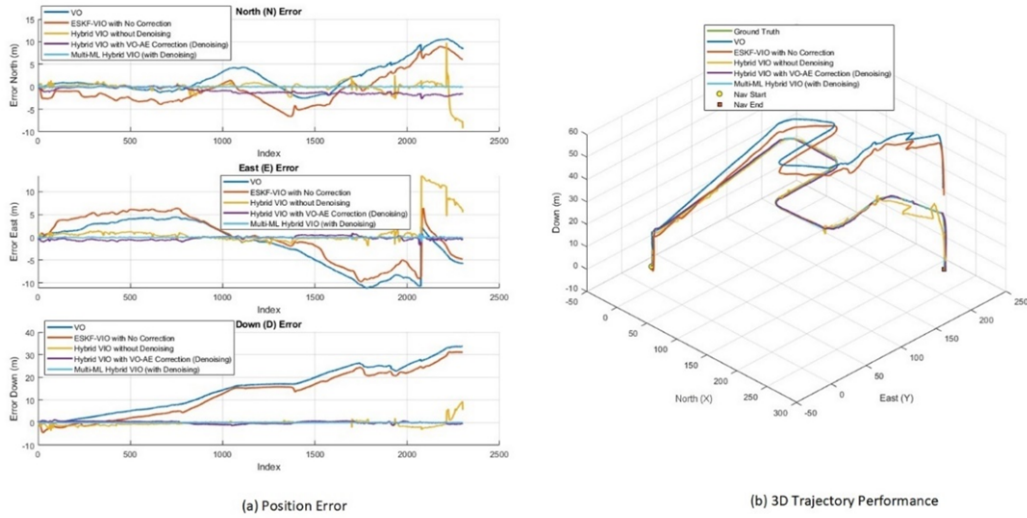


Figure 4-18 Position error and 3D trajectory comparison for trajectory 4 under unseen scenario with unknow fault conditions

Performance deviations are noticeable during landing phrases due to the navigation environment variability and presence of unknown fault conditions, as shown in Figure 4-18. Position error and 3D trajectory comparison for trajectory 4 under unseen scenario with unknown fault conditions. However, the hybrid architectures successfully adapt to take off from unseen locations, demonstrating their adaptability, whereas ESKF VIO struggles, with errors reaching up to 6m in E coordinate at 100s. Notably, deviations in the N & E position are observed throughout the flight, primarily caused by prominent fault conditions such as altitude variations and texture-less environment, which lead to feature tracking errors. ESKF-based VIO experiences increased errors due to the inability of the traditional Kalman filter to fully compensate for unmodeled dynamics, as uncertainty and noise degrade its performance. This is reflected in the wider error distribution shown in Figure 4-19. In contrast, hybrid VIO without denoising performs better by compensating for VO errors before fusion, allowing correction of compounded errors in sensor data. However, it introduces larger outliers despite a narrower distribution. After 200s, performance degradation is observed in N & E direction across VO, ESKF-VIO and hybrid VIO without denoising reference systems due to rotation with vibration during landing. These erroneous behaviours arise from the uncertain sources, such as corrupted sensor

measurements, feature tracking errors, which lead to mismatched characterisation of the underlying dynamics of sensor measurements and predictions. Hybrid VIO with VO-AE correction significantly reduces long tails, achieving horizontal 95th percentile errors of 1.92m. The integrated multi-ML hybrid VIO framework demonstrates even greater improvement, with a reduction of 98% and 94% towards N & E coordinates compared to ESKF-VIO. As shown in Figure 4-18 the multi-ML hybrid VIO algorithm shows the most significant reduction in horizontal position, indicating its efficiency to reduce uncertainty in sensor and model measurements. Additionally, the source of aleatoric or data uncertainty, e.g. feature tracking error due to rotation with vibration and altitude variation, leading to performance degradation during the landing phase, is mitigated. Multi-ML hybrid VIO achieves a 92% reduction, a 90% reduction for VIO with VO-AE correction in horizontal RMSE, and 99% reduction for D position for Multi-ML hybrid VIO compared to hybrid VIO without denoising. Thus, integrating denoising mechanism within multi-ML improves its adaptable ability to ‘out of distribution’ problems, ensuring safe navigation under challenging conditions.

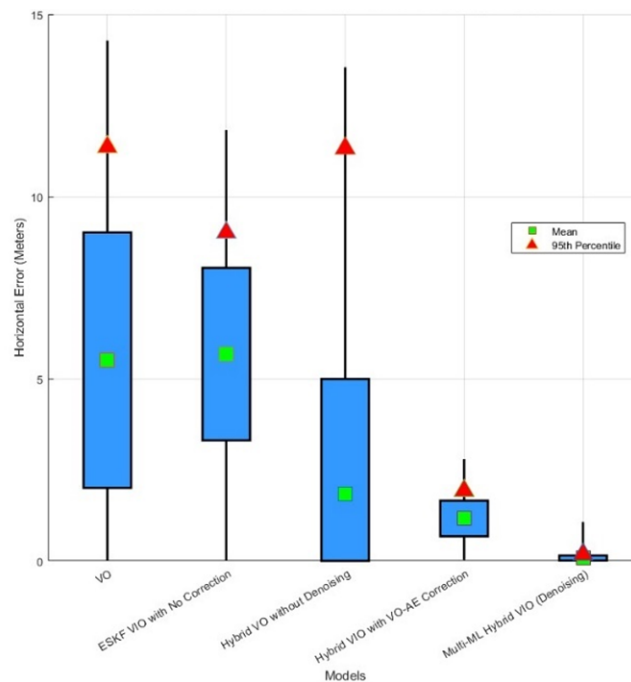


Figure 4-19 Error distribution in horizontal positioning for trajectory 4 under unseen scenario with unknown fault conditions

Table 4-7 RMSE and 95th percentile comparison on the performance under unseen scenario with unknown fault conditions for trajectory 4

Trajectory 4										
Algorithms	RMSE(m)					95 th Percentile(m)				
	N	E	D	3D	Horizontal	N	E	D	Horizontal	3D
VO	4.04	5.13	18.11	27.56	6.53	9.87	4.20	31.50	11.36	33.20
ESKF VIO with No Correction	3.7	4.83	16.00	24.19	6.15	7.79	6.05	28.26	9.01	29.63
Hybrid VIO without Denoising	1.37	3.43	1.25	3.9	3.70	1.09	11.29	0.88	11.32	11.85
Hybrid VIO with VO-AE Correction	1.18	0.44	0.52	1.37	1.27	1.89	0.78	0.86	1.92	1.97
Multi-ML Hybrid VIO (Denoising)	0.06	0.08	0.03	0.11	0.10	0.13	0.13	0.05	0.18	0.19

4.4.6 Comparison with literature

In order to show an empirical comparison to the closest related methods in the literature, the chapter selects EUROC dataset, especially MH05_difficult, for benchmarking to evaluate the performance of the proposed novel multi-ML hybrid VIO framework. EUROC dataset was collected with MAV flying in indoor environments, leveraging complex scenarios, i.e. darkness, illumination variation, rapid motion, and blur. The EUROC datasets MH01-MH04 series are used to train the model.

Table 4-8 presents the comparative results for MH_05 sequence in EUROC dataset. It is observed that the proposed framework constantly outperforms existing methods in terms of pose estimation accuracy. Difficulties arise from harsh motion dynamics, increased degrees of freedom and image blur. However, flight dynamics and images within the dataset were consistent following a pattern throughout the flight. Notably, the comparative analysis only included hybrid VIO architectures in SOTA. The proposed method has effectively mitigated erroneous measurements arising from various sources of uncertainty and showed improvement in 12cm and 16cm for 3D and horizontal positioning, respectively (95th percentile). In contrast, SOTA architectures suffered from performance degradation above 1m in 3D position due to insufficiency in adapting to the dynamics present in the data while using all the series of EUROC datasets for training their models. Additionally, many SOTA architectures evaluated their solutions by splitting the datasets and reinitializing the starting points as described in [26], [30], [49] While this approach helps achieving centimetre level

accuracy by addressing challenges like presence of rapid motion and motion blur in the flight that lead to feature tracking errors and feature location errors, it struggle to maintain confidence boundaries required for uninterrupted flight integrity. In contrast, this study evaluates the proposed solution on the full dataset without a split and demonstrates significant improvement, reducing 3D error to 7cm. Thus, integrating multiple error compensation with VO denoising mechanism within hybrid VIO architecture effectively mitigates the effects of such errors and enhance robustness to ensure safety.

Table 4-8 Comparison with SOTA methods that used MH05_difficult sequence from EuRoc dataset.

Techniques	RMSE (m)				95 th Percentile (m)	
	N	E	V	3D	Horizontal	3D
IR-VIO[135]	-	-	-	0.21	-	-
Self-VIO [43]	-	-	-	0.29	-	-
Hybrid VIO without Denoising	0.20	0.22	0.18	0.48	0.33	0.54
Hybrid VIO with VO-AE Correction	0.18	0.12	0.10	0.21	0.27	0.39
Multi-ML hybrid VIO (denoising)	0.04	0.05	0.03	0.07	0.12	0.16

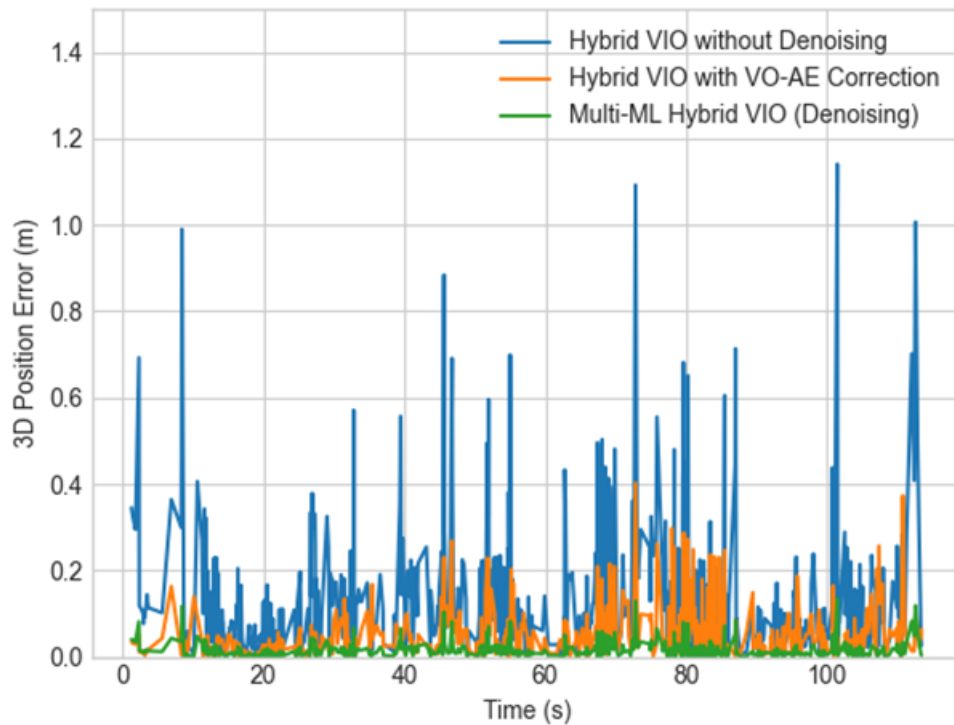


Figure 4-20 3D position error estimated using MH05_difficult seq. from EUROC dataset with multiple sources of aleatoric uncertainty.

Figure 4-20 presents the 3D position error of the proposed multi-ML hybrid VIO for processing MH05_difficult dataset. The hybrid VIO without denoising algorithm has demonstrated the least improvement among all the hybrid architectures due to the presence of complex flight dynamics and harsh lighting conditions, where feature-based VO struggles. Since the corrected VO position is fused ESKF within hybrid VIO, its performance suffers due to the nonlinear dynamics of MAV (rapid motion) and environmental factors (blur and darkness). These additional noise artefacts render the predictions less reliable. By incorporating denoising mechanism within hybrid VIO, the framework mitigates the nonlinearities, e.g. non-Gaussian noise, high-frequency vibrations, MAV dynamics and navigation environment complexity. This enhancement reduces the horizontal error to 27cm (95th percentile). Furthermore, incorporating VO-AE correction and KF model correction within Multi-ML hybrid VIO allows the ESKF model to linearise only the corrected sensor data, significantly improving performance by compensating additional model uncertainty and reducing large outliers. The key finding is that the proposed solution reduces position error in RMSE by up to 96% compared to hybrid SOTA methods i.e. end-to-end VIO.

While the study does not provide a full experimental comparison of computational requirements and response time across all evaluated methods, it provides a general discussion on the computational characteristics of the proposed approaches with SOTA methods. Due to the limited availability of open-source VIO algorithms, quantitative benchmarks for computational cost and real-time performance were not included in the evaluation. However, since the primary focus of this work is on fusion frameworks, it is worth noting that learning-based VIO like self-VIO rely on CNN-based architectures, which require 27ms per frame for fusion[43]. In contrast, the proposed GRU-based approach demonstrates a significantly lower processing latency, requiring only 3ms for GRU interface and approximately 1.92ms per frame for fusion.

4.5 Conclusion

This study presents a proposed multi-ML hybrid VIO incorporating denoising mechanism for a fusion framework using sensor data from IMU and camera to

significantly reduce position uncertainty and enhance robustness. The framework incorporates error compensators, including denoising-based VO-AE correction and ESKF error correction, trained to address both epistemic and aleatoric uncertainties. Furthermore, the proposed error compensation models utilise multi-layer GRU architectures to handle high flight dynamics and various sources of uncertainty. A series of experiments were conducted to validate the proposed approach under diverse conditions, including adverse weather, challenging light and high flight dynamics. These scenarios introduce visual sources of uncertainties such as feature extraction error, feature tracking error, data association error, mismatch in dynamics, unmodelled noise and imperfect modelling. The testing and training datasets included realistic failure modes and fault factors representative of urban environments. Notably, the multi-layers GRU demonstrated enhancement in performance in complex flight dynamics, achieving an 89% improvement in 3D position for challenging trajectory (trajectory 4) compared to simpler flight path (trajectory 1). Experimental results highlight the proposed framework's ability to capture and mitigate position uncertainty while removing additional induced unmodelled noise, even when sensor data is corrupted or the model is imperfect. For example, experiments evaluating the effect of weather conditions on the same trajectory revealed the frameworks' robustness against uncertainty caused by the limited visibility of features, reflection, shadow, and flight dynamics. The proposed solution achieved an improvement of 89% in foggy and 92% in rainy conditions compared to hybrid VIO without denoising. Despite additional noise from raindrops, glare and reflections, the proposed approach outperformed in foggy conditions. In less complex environments with nominal fault conditions, the framework demonstrated a significant reduction in position uncertainty by 91% compared with hybrid VIO without denoising. Performance degradation was observed under varying trajectory shapes e.g. lighting effects (sunrise) of different trajectory shapes. However, VO-AE correction mechanism effectively minimised uncertainty and smoothed error distribution, eliminating peaks along the horizontal axis under variation of lighting effects (sunny, shadow, sunrise). While denoising mechanisms eliminate outliers and unmodelled noise in harsh lighting

conditions, its effectiveness may be less pronounced under moderate light conditions, increasing confidence in its robustness and resilience under maximum fault scenarios. Comparison with reference systems demonstrates superior performance in conditions featuring maximum fault factors, even those absent during training. Lastly, the proposed multi-ML hybrid VIO demonstrates superior adaptability and performance in urban environments with various fault factors, effectively reducing aleatoric uncertainty within seen and unseen scenarios. It improves overall performance by 96% in 3D position when compared to traditional ESKF-based VIO. In summary, the combined performance and uncertainty gain estimates indicate the proposed framework's capability to ensure robust, safe and uninterrupted navigation using the vision-based A-PNT solution in realistic, complex urban environments, critical for UAM application. Future work will focus on enhancing the A-PNT framework by integrating additional sensors to mitigate vertical drift in high-flight dynamic scenarios. Furthermore, the prediction of uncertainty using Monte Carlo sampling within the model-based error compensation module will be explored to provide an additional level of confidence and error bounding to ensure operational safety.

5 Uncertainty Prediction and Model Correction by Bayesian GRU in Visual Navigation Applications

5.1 Introduction

Position, Navigation and Timing (PNT) solutions must be accurate and have high integrity, robustness, continuity and availability to meet safety requirements set by the International Civil Aviation Organisation (ICAO) [106]. Integrity is a critical performance metric for UAVS operating in low-altitude urban environments with stronger error turbulence. It reflects the trustworthiness of the navigation solutions and is defined by four key parameters: Alert Limit (AL), Integrity Risk (IR), Time to Alert (TIA), and Protection Level (PL), with the latter ensuring position errors remain within defined statistical bounds [31].

A significant challenge in urban navigation is ensuring the reliability of core sensors like the Global Navigation Satellite System (GNSS) for PNT. In urban canyons, GNSS signals are frequently affected by factors such as no-line-of-sight (NLOS) and multipath effects, necessitating the integration of alternative sensors with GNSS to achieve accurate positioning. Research focuses on integrating GNSS/IMU/VO in multisensory navigation systems to improve accuracy and integrity of PNT solutions [35], [39], [136]. This approach involves accessing and mitigating IMU errors such as bias drift and noise alongside GNSS errors caused by multipath effects, signal blockages and atmospheric disturbances. However, the impact of drift error in GNSS/INS fusion performance becomes particularly significant during prolonged GNSS outages, leading to performance degradation. This underscores the importance of carefully considering integrity in designing robust and resilient vision-based A-PNT solutions within GNSS/IMU/VO systems for safe navigation during GNSS outage.

Current state of art explores various methods such as Receiver Autonomous Integrity Monitoring (RAIM), residual chi-square test method (RCTM), KF-based integrity monitoring to improve navigation integrity for GNSS/IMU systems which works by classifying GNSS or IMU signals as being faulty or not [34], [35], [137], [138], [139]. However, these methods are only effective in situations where most

signals are not erroneous and may not be suitable for slow-growing failures and harsh environmental effects, as corrupted or noisy signals from VO are integrated over time. A common approach to improving navigation integrity involves incorporating fault detection in feature-based VO algorithms. However, these methods often struggle to bound errors arising from various sources of measurement uncertainties in urban environments [50], [51], [52]. The reliance on predefined error models for visual uncertainties, such as feature tracking errors, feature extraction errors, feature location errors, and data association errors, has shown significant limitations. These approaches are prone to performance degradation in harsh environments, including low light, darkness, and high flight dynamics. While many studies focus on mitigating individual sources of uncertainty, such as feature extraction error [56], [58][59], feature matching error [140], feature association error [57], they fail to address error bounds under multiple or combined fault factors. This oversight can introduce noise into the system, further compromising the performance of fusion models.

To address these challenges, hybrid architectures have been integrated into VIO systems to compensate for errors in corrupted VO sensors and Kalman filter (KF) model measurements using machine learning (ML), thereby enhancing performance while accounting for various sources of uncertainties under complex urban environments [123], [131]. However, quantifying errors remains an open issue, particularly when employing ML methods within hybrid VIO frameworks to enhance integrity. In addition, many potential solutions involving ML-based VO are being investigated to provide the confidence measure for epistemic (associated with the model) and aleatoric (associated with input data) uncertainty. For example, Costante et. al [60] proposed a Bayesian LSTM utilising Monte Carlo Dropout (MCD) to provide pose prediction with confidence measures that incorporate both model (epistemic) and observation noise variance (aleatoric). Nevertheless, this approach lacks validation under visually degraded scenarios (e.g. texture-less environments, rapid motion, low-light), resulting in low confidence and degraded performance when insufficient iterations and evaluation are performed. Other methods within VO utilised Bayesian Variational Autoencoder [61], [62] and RCNN with mixture density network [63], extracted

predictive position with confidence measures, but failed to provide statistical guarantees under complex scenes. In addition, state-of-the-art systems continue to focus on advanced training methods to improve robustness with a high confidence measure. However, they often lack statistical rigour, particularly when dealing with non-Gaussian error distributions, where 3σ uncertainty interval (UI) cannot ensure error bound. Attention should be given to predicting uncertainty for non-Gaussian error distributions within hybrid VIO to provide confidence measures for safe and reliable navigation during GNSS outage in complex urban environments. Therefore, this study extends the previous work on vision-based A-PNT [131] solution with a proposition of estimating the uncertainty of model-based error using Bayesian GRU with MCD, aiming to provide a confidence measure for safe navigation. Thereby, the hybrid architecture combines GRU-AE-assisted VO correction and Bayesian GRU-aided ESKF measurement correction with ESKF that estimates robust and accurate position with error uncertainty prediction to enhance navigation integrity. To the best of the author's knowledge, this is the first attempt to estimate uncertainty using Bayesian GRU from model compensators that also address data uncertainty for VIO applications. In this research, Bayesian GRU is utilised due to its efficiency in processing time-varying sequences with less computational complexity, making it well-suited for faster and accurate uncertainty quantification in real-time navigation systems. It offers advantages over other Bayesian LSTM and Bayesian Bi-LSTM models, which involve more complex architectures. The proposed Bayesian-based model compensator exhibits Bayesian characteristics primarily through MCD, providing a balance between computational efficiency and uncertainty quantification. In addition, the study integrates an enhanced fault-tolerant hybrid GNSS/IMU/VO multi-sensor navigation system, enabling prediction of confidence measures for A-PNT solutions to ensure high integrity during navigation in complex urban areas. Furthermore, the formulation is extended by adversarial training methodology for ML models to improve robustness of the proposed approach using MATLAB incorporated with unreal engine simulated photorealistic datasets replicating fault factors that produce possible uncertainty sources (e.g. feature tracking errors, data association errors,

sensor noise, feature location errors) in urban environments for various adverse flight scenarios. The contributions of this work are highlighted as follows:

- [1] A novel Bayesian GRU is proposed within Multi-ML hybrid VIO to predict model errors with uncertainty measures of complex visual failure modes, which is utilised to correct overall position. The algorithm aims to estimate a robust position with uncertainty estimation, balancing robustness and integrity under complex urban environments.
- [2] Systematic evaluation of the proposed approach under complex visual failure modes that produce possible uncertainty sources such as feature tracking errors, feature location errors, data association errors and sensor noise.
- [3] Performance evaluation of integrated fault-tolerant hybrid GNSS/IMU/VO multi-sensor navigation system leveraging proposed APNT solution, aiming to estimate its ability to perform effectively and provide the confidence measure under GNSS outage scenarios.

The remaining part of the chapter is organised as follows. Section 5.2 introduced the proposed uncertainty-aware hybrid VIO A-PNT solution, incorporating it within hybrid GNSS/IMU/VO navigation system; Section 5.3 uncertainty inspection in visual navigation; Section 5.4 discusses the experimental setup and data collection; Section **Error! Reference source not found.** discusses the performance evaluation, and result analysis comparison with state-of-the-art systems, and the conclusion is presented in Section 5.6.

5.2 Proposed Uncertainty Prediction and Error Correction Mechanism

This section presents a high-level overview of the multi-sensor hybrid GNSS/IMU/VO navigation system incorporating the proposed hybrid VIO with uncertainty prediction, aiming to predict model error uncertainty and correct overall position to ensure robust and safe navigation under complex urban environments.

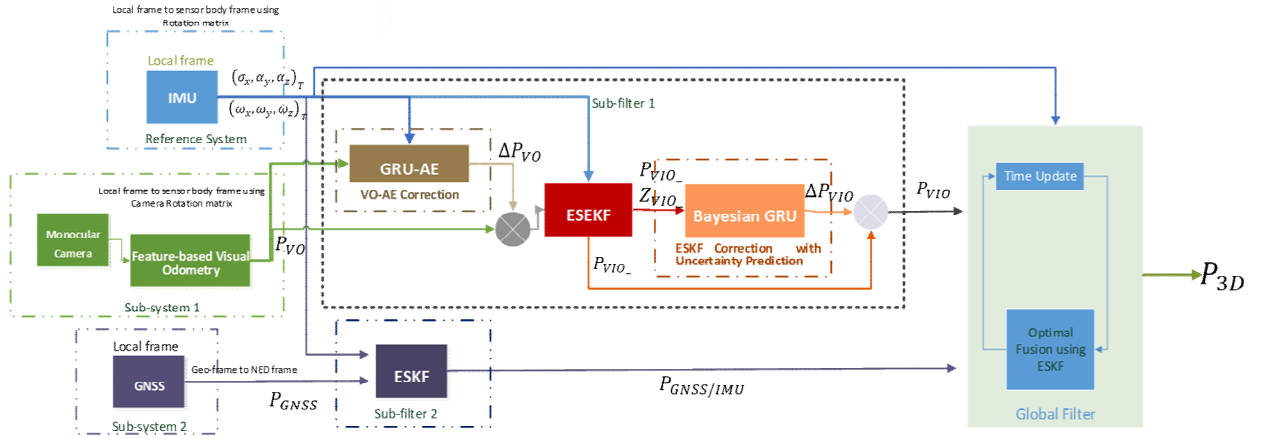


Figure 5-1 Robust fault-tolerant multi-sensor fusion framework with vision-based A-PNT uncertainty prediction to enhance integrity in the absence of GNSS

Two sub-filters exist in the hybrid GNSS/IMU/VO architecture: an uncertainty-aware hybrid VIO sub-filter and an ESKF-based traditional GNSS/IMU integration sub-filter. The output of the two sub-filters is merged together with a global ESKF to generate the ultimate position estimations. The former sub-filter of the uncertainty-aware hybrid VIO attempts to provide a reliable and robust VIO position with error uncertainty estimation, while the latter sub-filter of EKF-based GNSS/IMU integration aims to estimate position from GNSS/IMU. The motivation for introducing the integrated hybrid GNSS/IMU/VO fusion in this study is to evaluate the accuracy and reliability of the proposed A-PNT solution that addresses possible fault conditions by managing uncertainty effectively to ensure high integrity in the complex urban environment.

5.2.1 Bayesian GRU

Bayesian-GRU refers to the implementation of the GRU neural network with Bayesian inference to allow for the quantification of output uncertainty. In contrast, GRU is designed to handle sequential data and can maintain a memory of past states, making it well-suited for learning the temporal dependencies in vehicle movements. It is a variant of LSTM that merges the input gates and forgetting gates along with merging the memory state and hidden state. GRU has much fewer parameters than LSTM, and the training speed is faster. The hidden layers unit in the network introduces a gain unit, and adding additional layers

provides efficiency to learn complex vehicle dynamics in the data. The chapter introduces a 4-layer GRU architecture with varying configurations, one of which integrates the Bayesian framework to predict uncertainty of the predictive ESKF model error estimation. Bayesian principles into GRU replaces point estimates of model's weight with probability distributions, enabling a more comprehensive representation of uncertainty with predictions. This approach enhances the model robustness and improves generalisation capabilities. Abdar et al. [141] reviewed and compared various uncertainty quantification methods, highlighting the advantages of MCD over other sampling methods in the Bayesian Interface. Their analysis emphasised MCD's balance of accuracy and computational efficiency. In contrast, methods like MCMC can be computationally intense, while VI may often require careful specification of the approximating family [136].

Machine learning model predictions are generally affected by heteroscedastic aleatoric uncertainty (HAU) and EU. This study considers an adverse training methodology to handle HAU uncertainty by employing VO-AE error compensator to correct input VO data within hybrid VIO. Detailed explanations are provided in section 5.2.2. and 5.3. For the EU, the predicted ESKF model compensator output uncertainty is quantified using the MCD sampling method, which predicts the mean and 95th percentile confidence interval of the network prediction to ensure a reliable confidence measure for the hybrid model. Given that the error distribution of the hybrid VIO is non-Gaussian, relying solely on the mean and variance of the predicted output to define uncertainty boundaries is not realistic.

5.2.2 Uncertainty Prediction in Hybrid VIO

The uncertainty-aware Hybrid VIO incorporates multiple error compensation mechanisms, such as GRU-AE correction to compensate VO sensor and Bayesian GRU-based model correction to compensate ESKF model with integration of ESKF. Here, GRU-AE is adopted from section 4.2.1, remaining unchanged for the adaptability of input and output parameters for VO sensor correction. The ΔP_{VO} model predicts position error between VO measurements and ground truth with the input parameters IMU linear acceleration in body frame within NED coordinate, $\alpha_N, \alpha_E, \alpha_D$ and IMU angular velocity in the body frame

within NED coordinate, $\omega_N, \omega_E, \omega_D$. Having the advantage of the utilizing canonical relationship between inertial frame and UAV dynamics, the model can effectively learn implicitly to track the vehicle's movement by observing the sequences of IMU linear acceleration and angular velocity. Additionally, VO estimated position and orientation are also other input parameters for the GRU-based VO error compensator. By analysing dependencies between the estimated position and the actual position, the GRU can learn typical error patterns.

The Bayesian GRU model is applied to compensate and predict the uncertainty of the ESKF measurement model errors associated with the imperfection of the model due to manual tuning and nonlinearity presence within the data. The model measurement errors are compensated by predicting KF measurement increments. The ΔP_{VIO} the model predicts position error between the ESKF-measured position and ground truth with the input parameters ESKF estimated position and innovation i.e. the difference between the predicted observations and observed value in the NED frame. The Bayesian GRU model error compensator can effectively learn from the information about the state-evaluation process and the uncertainty of the state estimate. Additionally, incorporating MCD dropout facilities prediction of model uncertainty in predicted increments utilises 500 samples to provide reliability of estimated mean and 95th percentile UI, capturing the model's predictive uncertainty to ensure accurate confidence intervals. Notably, increments of ESKF with VO-AE correction are selected as ground truth in these cases for evaluation of the prediction of the Bayesian GRU model.

5.2.3 Hybrid Multi-Sensor Navigation System

The fault-tolerant hybrid multi-sensor navigation system utilises uncertainty-aware hybrid VIO as one of the sub-filters, while ESKF-based GNSS/IMU integrates into another sub-filter. The GNSS/IMU fusion sub-filter is adopted from the literature [142]. The multi-sensor system employs a federated architecture designed to maintain confidence in its operation even if one of the sub-filters becomes corrupted. In this chapter, I evaluate the trust measures of the vision-based APNT component within this architecture, particularly in scenarios where

GNSS fails or diverges. The system is expected to ensure that errors remain bounded within predicted confidence measures, thereby maintaining integrity. The state vector x of the global ESKF of the multi-sensor system selects the following states:

$$x = \left[P_x, P_y, P_z, V_x, V_y, V_z, A_x, A_y, A_z, \Delta\alpha_{b_x}, \Delta\alpha_{b_y}, \Delta\alpha_{b_z}, \Delta\omega_{b_x}, \Delta\omega_{b_y}, \Delta\omega_{b_z} \right] \quad 5-1$$

where, P_x, P_y and P_z denote the position, V_x, V_y and V_z denote velocity, A_x, A_y and A_z denote attitude, $\Delta\alpha_x, \Delta\alpha_y$ and $\Delta\alpha_z$ denote accelerometer bias and $\Delta\omega_x, \Delta\omega_y$ and $\Delta\omega_z$ denote gyroscope bias in the x-axis, y-axis and z-axis. In this context, IMU biases are extracted from global ESKF, where IMU serves as the reference system to provide accurate update to the ESKF prediction. The fusion resolution of the federated filter is followed by equation 2-18. The pseudo-code of the integrated fault-tolerant hybrid multi-sensor navigation is presented in A 5-1.

A 5-1: Algorithm of Hybrid GNSS/IMU/VO Multi-Sensor Navigation System

Input: $P_{VO}, P_{VIO}, P_{VIO_}, \alpha_{xyz}, \omega_{xyz}, P_{GNSS} // P_{VO}$ denotes the VO estimated position; P_{VIO} denotes the ESKF-VIO with VO-AE Correction estimated position; $P_{VIO_}$ denotes uncertainty-aware hybrid VIO estimated position; α_{xyz} denotes the linear acceleration; ω_{xyz} denotes the angular velocity; and P_{GNSS} denotes the GNSS estimated position

Output: $P_{VIO_}, UI, P_{GNSS/IMU}, P_{GNSS/IMU/VO} // P_{VIO_}$ denotes the sub-filter 1 estimated position; UI denotes uncertainty intervals; $P_{GNSS/IMU}$ denotes the sub-filter 2 estimated position; and $P_{GNSS/IMU/VO}$ denotes the global filter estimated position

Initialize:

1. Initial the values for hybrid uncertainty-aware VIO and GNSS/IMU sub-filters: $x_k = x_0$;// x_0 denotes ground truth from UAV.
 2. Initial values for global filter ESKF: $x_{k|global} = x_0$;// x_0 denotes the ground truth from UAV.
-

for $n = 1, 2, 3, \dots$ **do**

// n denotes number of iterations

Prediction Phase:

3. The sub-filter time update model has been derived in equation 2-2, which is similar for both sub-filters and global as using the ESKF variant.
4. Covariance vector is defined as, $\Sigma_{k|k-1} = F_{k|k-1}\Sigma_{k-1}F_{k|k-1}^T + L_{k|k-1}Q_{k-1}L_{k|k-1}^T$.

Measurement Phase for sub-filters:

5. Generate the $P_{VO/GRU}$ using equation 4-4 incorporating VO-AE correction predicted increments and updated the following measurement steps of the ESKF VIO sub-filter.
6. Estimate the states using the updated equation (2-4)-(2-9)
7. Calculate the Kalman gain, Corrected VIO states incorporating VO-AE correction only, and the covariance matrix using equations (2-10)-(2-12)
8. Corrected $X_{KVIO} = x_{k|k-1_{VIO}} + K_{k_{VIO}}(z_{k_{VIO}} - H_{k_{VIO}}x_{k|k-1_{VIO}})$.
9. Generate the $P_{VIO/GRU}$ using incorporating Bayesian GRU predicted increments of ESKF model with model uncertainty UI and updated the uncertainty-aware hybrid VIO position in sub-filter.
10. Start from step 2-4 to estimate the Kalman gain using Equation 4-7 for the GNSS/IMU integration.
11. Update the observation equation 4-7.
12. Estimate the updated states and covariance matrix of the GNSS/IMU filter using equations 4-7.
13. The proposed uncertainty-aware hybrid VIO integrated sub-filter and GNSS/INS integration using ESKF are running in parallel to estimate the position for each filter.

Measurement update for Global filter:

14. Calculate the Kalman gain:

$$K_{global} = \hat{\Sigma}_{global} H_{k_{global}}^T \left(H_{k_{global}} \hat{\Sigma}_{k|k-1_{global}} H_{k_{global}}^T + \Pi_{global} \right)^{-1}$$

15. Update the states using the traditional ESKF equation 4-7.

16. Obtain $P_{k|global}$ and $X_{k|global}$ using equation 2-17 and 2-18.

End for

5.3 Uncertainty Inspection in Visual Navigation

In this section, the basics of feature-based visual navigation error sources are briefly introduced. Figure 5-2 illustrates fault propagation in feature-based visual navigation, resulting in various sources of uncertainty. Images contain huge information provided by intensity values of all pixels and their spatial relationship, where feature points are detected in order to extract the most important geometric information for positioning. The corresponding error is photometric error and geometric error in such an intensity domain raising due to the unavailability of sufficient features. This leads to a feature extraction error in 2D coordinates. On top of such 2D error, feature location error occurs while matching 2D coordinates in the 3F points to estimate camera pose. It should be noted that faults caused by such error sources may occur independently in different domains besides fault propagation. For instance, during the feature extraction step, a combination of other error sources may be predominant, resulting in combinations of erroneous output. Thus, the error sources in all different phases in visual navigation must be taken into consideration in order to ensure integrity. Besides, Ochella et. al [143] categorised uncertainty models into four, including present uncertainty, future uncertainty, modelling uncertainty and prediction method uncertainty. It is suggested to take the prediction of model uncertainty while designing Bayesian networks, as it accounts for the combination of present, future and modelling uncertainty.

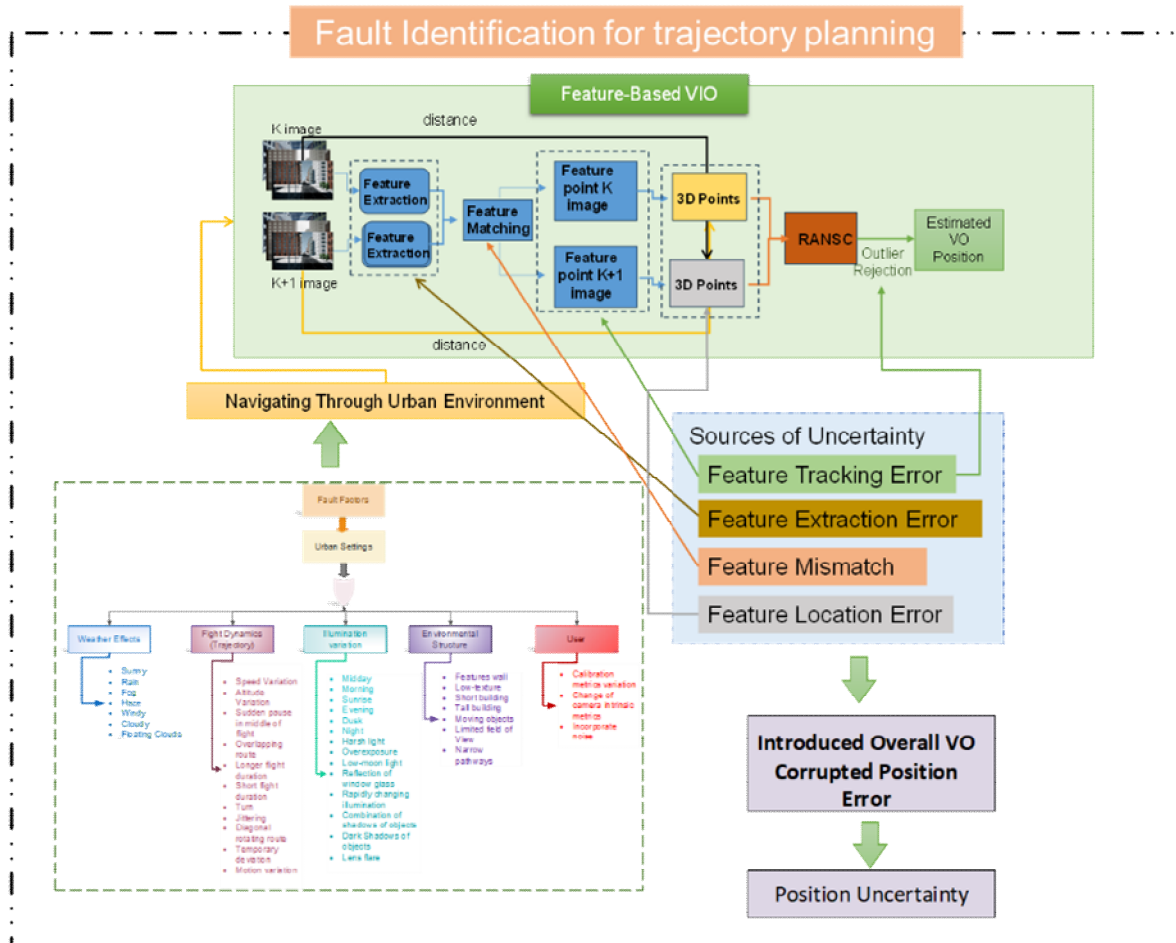


Figure 5-2 Fault propagation in feature-based visual navigation representing sources of uncertainty

Evaluating the adaptability and robustness of ML algorithms across diverse scenarios without sacrificing accuracy remains challenging, compounded by the need for large training datasets. Based on an extensive review using FMEA in visual navigation systems on section 2.2.3, the research has categorised possible factors in urban environments, including weather effects, flight dynamics, illumination variations, environmental structure and user interface. The fault factors arising from various sources of data (aleatoric) uncertainty contribute to position uncertainty. Additionally, the autoencoder in the VO error compensator has capabilities to minimise artefact noise arising from raindrops, featureless scenarios, glare and reflection that often leads to long outliers. The objective is to minimise aleatoric uncertainty in the proposed APNT solution by structuring training datasets for ML-based compensators, enabling them to be fault-tolerant

across diverse error sources to ensure integrity. Here, the GRU-AE error VO error compensator has been implemented to reduce aleatoric uncertainty. Subsequently, the Bayesian GRU is employed to predict residual uncertainty affecting the model performance and predict the increments of the ESKF model. These predicted errors are then compensated, resulting in an overall reduction in positional uncertainty with estimation of UI with a 95th confidence interval. Thus, the proposed uncertainty-aware Multi-ML hybrid VIO aim to enhance the overall vision-based navigation system in UAM applications.

5.4 Experimental Setup & Data Collection

This section outlines the experimental setup to evaluate the performance of the proposed multi-ML hybrid VIO. The experimental setup is described first, which is used to generate training and testing datasets to conduct numerical analysis and performance evaluation in the presence of multiple sources of uncertainty under complex environments for visual navigation. The experimental setup used in this research has been implemented in MATLAB, incorporating the Unreal Engine to simulate data collection for UAVS. The simulated datasets involve diverse challenging conditions, including illumination variation, rapid motion, altitude variation, no-texture, and weather effects for shorter and extended flights. The experimental architecture is adopted from one of the previous works, which was used to implement a fault-tolerant federated multi-sensor GNSS/VO/IMU navigation system [123]. The choice of monocular camera is beneficial for the UAM application due to its advantage of being less expensive, simpler to implement and lightweight to fit into a UAV with a specification of 1109 (pixels) focal length, 640x360 (pixels) optical centre and 720x1280 (pixels) image size, operating in 10 Hz sample rate. The simulated IMU data has the specification of the ICM 20649 model operating 100 Hz sample rate used in the previous work. The simulated GPS block is configured with specifications representative of low-cost receiver accuracy parameters, including horizontal positioning accuracy of 2.104m and vertical positioning accuracy of 3.461m [144]. During the post-processing step, linear interpolation has been carried out to synchronise timestamps between both sensors. In contrast, for urban operation scenarios

surrounded by buildings, the visual data of tall buildings are simulated using ‘US city’ scene available in the ‘simulation 3D scene configuration’ block within the Navigation Toolbox and UAV Toolbox in MATLAB. The simulated weather effects and fault conditions are configured from the previous setup, explained in section 3.3.1. This chapter investigated only the navigation environment of sunny, foggy and sunrise environments. However, training datasets are collected by replicating weather effects and lighting conditions. The setting is adopted based on the MathWorks user interface described on the website. The flight trajectories are designed and controlled through QgroundControl software connected through MAVLink connectivity.

5.4.1 Training Methodology:

The training dataset includes 12 trajectories combining fault conditions like flight dynamics, environmental structures and illumination variations that can trigger multiple sources of data and model uncertainty within the navigation system involving imperfect modelling, sensor noise, data association errors, feature tracking errors and feature location errors. Within these trajectories, a set of various weather conditions is presented, including fog, rain, sunny, cloudy and various lighting conditions like dusk, sunrise, afternoon, and evening. Thus, 12 different shapes of trajectories are formulated in various weather effects and lighting conditions, forming 96 datasets for training each ML model. This diversity allows the model to learn and adapt to different fault conditions, improving overall performance and generalisation ability in the presence of multiple fault conditions. In the proposed uncertainty-aware Multi-ML hybrid VIO is configured with the combination of multiple ML models that are utilised for different purposes within the system. The denoising autoencoder is trained on 12 trajectories under sunny weather conditions, with an assumption of having minimal fault effects compared to others. This training setup is designed to enhance the denoising mechanism’s ability to handle erroneous signals with long outliers. Furthermore, the GRU-based VO error predictor model is trained on 96 datasets using a random search hyperparameter optimisation technique, allowing it to learn the complex flight dynamics by splitting the datasets. The GRU layers, neurons, learning rate,

activation function and dropout remain the same as evaluated and described in chapter 4. In this study, the proposed Bayesian optimisation is optimised and trained on similar 96 trajectories covering weather effects and lighting conditions using the Bayesian optimisation tuner from the TensorFlow Keras python library. This method systemically explores the activation function space during training to identify the most optimal configuration for capturing nonlinear dynamics and high motion variability present in the training datasets. This hyperparameter tuner is selected based on a review of other hyperparameter tuning methods and an accuracy comparison conducted on similar datasets. To avoid overfitting and review performance, 12 trajectories with sunny weather are utilised for the training to find the best match of hyperparameters. The activation function, number of neuron units and dropout rate are initially fluctuating with changes of trajectory. However, after training on 7 trajectories, the optimal configuration is identified, requiring no further adjustments. Table 5-1 denotes an example of changes of configuration values using the Bayesian optimisation method.

Table 5-1 Comparison between Bayesian hyperparameter optimisation based on training trajectories

#Trajectory	Activation function	Recurrent Activation function	Neurons	Dropout	Accuracy (%)
1	Layer-1: tanh Layer-2: relu Layer-3: relu Layer-4: tanh	Layer-1: tanh Layer-2: sigmoid Layer-3: relu Layer-4: tanh	Layer-1:300 Layer-2: 150 Layer-3: 80 Layer-4:50	Layer-1:0.0 Layer-2:0.3 Layer-3: 0.2 Layer-4:0.1	84.8
2	Layer-1: relu Layer-2: tanh Layer-3: tanh Layer-4: relu	Layer-1: sigmoid Layer-2: tanh Layer-3: relu Layer-4: relu	Layer-1:250 Layer-2:138 Layer-3: 72 Layer-4: 40	Layer-1:0.2 Layer-2:0.1 Layer-3:0.0 Layer-4:0.3	87.4
3	Layer-1: sigmoid Layer-2: tanh Layer-3: tanh Layer-4: sigmoid	Layer-1: sigmoid Layer-2: tanh Layer-3: sigmoid Layer-4: relu	Layer-1: 300 Layer-2:300 Layer-3: 150 Layer-4: 100	Layer-1:0.3 Layer-2:0.0 Layer-3:0.1 Layer-4:0.1	89.1
4	Layer-1: relu Layer-2: tanh Layer-3: relu Layer-4: tanh	Layer-1: relu Layer-2: sigmoid Layer-3: tanh Layer-4: relu	Layer-1: 200 Layer-2:178 Layer-3:120 Layer-4:48	Layer-1:0.4 Layer-2:0.3 Layer-3:0.2 Layer-4:0.2	89.3
5	Layer-1: tanh Layer-2: relu Layer-3: tanh Layer-4: relu	Layer-1: relu Layer-2: sigmoid Layer-3: tanh Layer-4: relu	Layer-1:350 Layer-2:350 Layer-3:100 Layer-4:40	Layer-1:0.2 Layer-2:0.2 Layer-3:0.4 Layer-4:0.0	93.7
6	Layer-1: relu Layer-2: sigmoid Layer-3: tanh Layer-4: relu	Layer-1: relu Layer-2: sigmoid Layer-3: tanh Layer-4: relu	Layer-1:50 Layer-2:128 Layer-3:128 Layer-4:50	Layer-1:0.0 Layer-2:0.2 Layer-3:0.3 Layer-4:0.2	97.6
7	Layer-1: relu Layer-2: relu	Layer-1: relu Layer-2: sigmoid	Layer-1:50 Layer-2:128	Layer-1:0.1 Layer-2:0.3	95.1

	Layer-3: tanh Layer-4: tanh	Layer-3: tanh Layer-4: relu	Layer-3:128 Layer-4:50	Layer-3:0.2 Layer-4:0.2	
--	--------------------------------	--------------------------------	---------------------------	----------------------------	--

Finally, the fully connected 4 sequential layers are utilized in proposed Bayesian GRU to provide output predicted increments with uncertainty estimation using 500 Monte Carlo samples. The batch size used is 128, the Monte Carlo dropout rates are set at 0.1, 0.3, 0.2 and 0.2, the activation function are including 'relu', 'relu', 'tanh' and 'tanh' and recurrent activation function are 'relu', 'sigmoid', 'tanh' and 'relu'. The ADAMAX optimiser is employed during the training to optimize the ML model weights and minimise the discrepancies between the predicted and actual errors. Furthermore, the loss function is calculated using mean squared error (MSE).

5.4.2 Testing Phase

The testing dataset is designed with a combination of seen and unseen fault scenarios, which are different from the training datasets. During the testing phase, ML models operate in the prediction phase, remaining active throughout the experiment. During this phase, the trained GRU-AE VO error compensator predicts VO position error based on the given navigation frame linear acceleration, angular velocity, VO position and VO current orientation. For evaluation purposes, the updated position error prediction is deducted from the actual VO position estimated from the feature-based VO algorithm. Lastly. The trained Bayesian GRU KF measurement error compensator predicts KF estimated position error with uncertainty based on KF predicted position and observation differences from the measurement model in the current navigation frame.

5.4.3 Uncertainty Evaluation Metrics

A good uncertainty estimation model should provide uncertainty intervals (UI) that are as small as possible, at the same time including the ground truth value. In this study, the Bayesian GRU is predicting model error uncertainty, provided that the ground truth value is the error associated with a part of the uncertainty-aware Multi-ML hybrid VIO with only GRU-AE correction applied. Thereby, ground truth is estimated by subtracting the ESKF estimated position with GRU-AE correction

from the actual test trajectory to evaluate the behaviour of capturing EU uncertainty within the model. The updated position is estimated by deducting the predicted position error from the KF estimated position to improve overall uncertainty-aware Multi-ML hybrid VIO positioning performance. The mean and 95th percentile confidence boundary is selected to evaluate the model capturing of uncertainty in the error distribution of target data. The comparison is made by computing the percentage of ground truth value fault out of the 95th percentile UI (out of range value percentage- OR%) and 99% percentile UI OR%.

5.5 Performance Evaluation

To evaluate the performance and adaptability of the proposed novel hybrid VIO framework, six experiments are carried out using eight different testing datasets (eight different trajectory shapes), with a variation of fault conditions, weather and lighting effects. Table 4-1 represents the description of testing scenarios, trajectory types and fault conditions.

Table 5-2 Test trajectory description

# Trajectory	Testing Scenario	Lighting condition	Fault type
1	Nominal Single faults	Sunny	1. Shadows of objects 2. White Rooftop 3. Turn
1	Weather effects	Fog	1. Turn with features-wall. 2. Turn with sky view. 3. Shadows of objects 4. Take off with different altitudes
3	Lighting Effects	partly sunny	1. Sharp turns with sky-view 2. Vibration while turning 3. Altitude variation in the flight
4	Unseen Scenario with unknown faults	Sunny	1. Turns with featureless walls. 2. Turns with sky-view. 3. Shadows of objects 4. Vibration with sharp turn 5. Altitude variation 6. Take-off with different altitudes.
5	Lighting Effects	Sunrise	1. Altitude variation among waypoints 2. Turn with sky-view 3. Vibration with waypoint changes 4. Take-off from unknown location
6	Lighting Effects	Sunrise	1. Sharp and curved turns 2. Turn with sky-view 3. Vibration with waypoint changes

			4. Feature-less wall 5. Take-off from unknown location
7	Flight Complexity with an unseen location	Sunny	1. Turns with featureless walls. 2. Turns with sky-view. 3. Shadows of objects 4. Repeated turn 5. Multiple turns
8	Flight Complexity	Sunny	1. Turns with featureless walls. 2. Turns with sky-view. 3. Shadows of objects 4. Curved turn 5. Multiple turns

Figure 5-3 represents testing trajectory to evaluate proposed approach with denoising mechanism performance in terms of robustness under different weather conditions and adaptability over unknown scenarios and fault factors. four types of testing scenarios (e.g. weather effects, unknown scenarios with unknown faults, complexity of flight dynamics, lighting effects) have been selected to evaluate performance enhancement with position uncertainty minimisation using uncertainty-aware Multi-ML hybrid VIO-

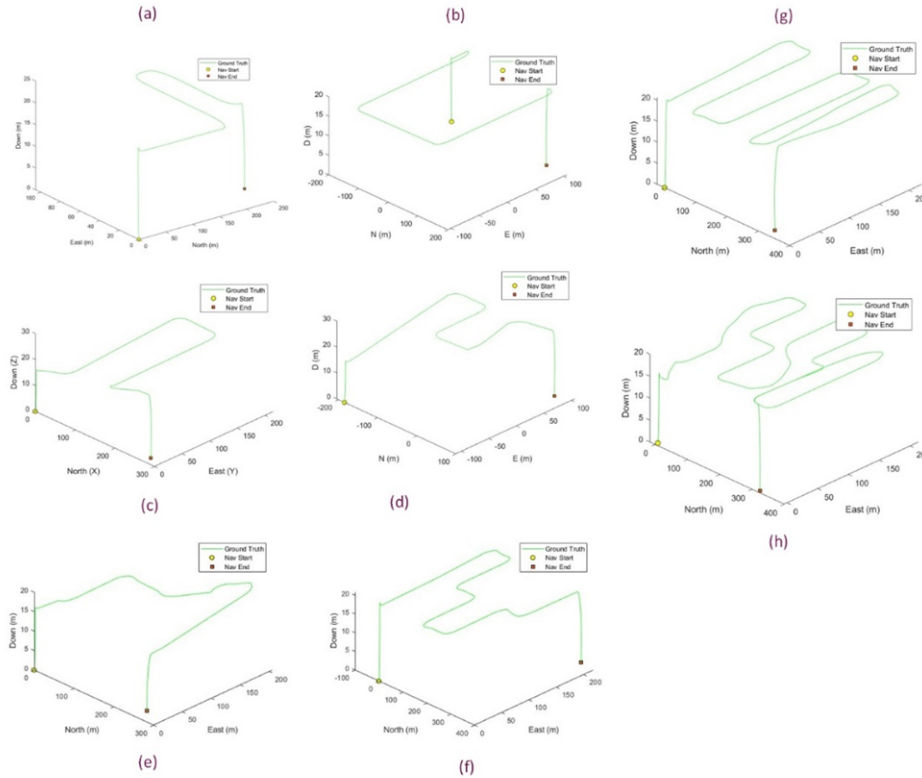


Figure 5-3 3D view of testing trajectories for experiments (a) Trajectory 1, (b) Trajectory 2, (c) Trajectory 3, (d) Trajectory 4, (e) Trajectory 5, (f) Trajectory 6, (g) Trajectory 7, (h) Trajectory 8

5.5.1 Uncertainty Estimation in Vision-based Alternative Navigation

5.5.1.1 Fusion Experiment over Complex Survey Trajectory

In this section, the experiment aims to test whether the proposed Bayesian GRU within uncertainty-aware-VO can effectively improve trajectory accuracy in the presence of complex flight dynamics. This also evaluates the model's adaptability to unknown scenarios. For this purpose, two trajectories are chosen based on high dynamics and navigation environments complexity, incorporating trajectory one starting from the new location, which is not included in the training data. Both of the tests are conducted in sunny weather conditions. Trajectory 7 is composed of environmental and flight fault factors such as featureless buildings, turns with sky-view, shadows of buildings, and take-off from different streets in the city environment. During the experiment distance of 1632 is covered over a duration of 9.1 minutes. This experiment is conducted different time zone on a sunny day.

Figure 5-4 illustrates the comparison of position errors of each coordinate using the proposed uncertainty-aware Multi-ML hybrid VIO and Multi-ML hybrid VIO (denoising), which is proposed in chapter 4, evaluating performance enhancement with the proposed Bayesian GRU between trajectory 7 and trajectory 8. Performance enhancements are noticeable all over the flight in both trajectories compared to the Multi-ML-hybrid VIO (denoising) method. However, both hybrid architectures successfully adapt to take off from unseen locations in trajectory 7, demonstrating their adaptability, whereas ESKF VIO struggles, with errors reaching up to 35.91m drift in 3D position. In contrast, comparative analysis observed deviation in performance in N coordinate using Multi-ML-hybrid VIO (denoising), which is mitigated by the proposed Bayesian GRU, providing insight into its ability to mitigate complex flight dynamics, including curve turns, altitude variation, and loops within flight covering low-texture environment. Table 5-3 summarises the comparison of RMSE and 95th percentile values for trajectories 7 and 8. Deviations in the D position for trajectory 7 are observed throughout the flight in Multi-ML hybrid VIO (denoising), primarily caused by prominent fault conditions such as multiple turns with sky-view in the new location, which lead to feature extraction errors. Thereby, the model uncertainty affecting overall

performance should be taken into consideration to improve robustness. This provides insight into the initial concept of the chapter, stating that the VIO model should be designed to predict and compensate model uncertainty while compensating data uncertainty simultaneously. However, the proposed solution has demonstrated enhancement in performance with 99.1%, 99.6% 99.7% in N, E, D coordinates compared with VO for trajectory 7. In contrast, comparison between both hybrid approaches has demonstrated enhancement of 64% for trajectory 7 and 68% for trajectory 8 in 3D coordinates. This demonstrates its robustness capabilities under more complex flight dynamics.

The integrated Bayesian GRU with multi-ML hybrid VIO framework demonstrates greater improvement in mitigating drift error in D coordinate, with a reduction of 68% and 65% for trajectory 7 and 8 compared to Multi-ML hybrid VIO (denoising). However, trajectory 7 outperforms due to the presence of less complexity compared to trajectory 8, highlighting the model's generalisation ability as flight includes unseen scenarios.

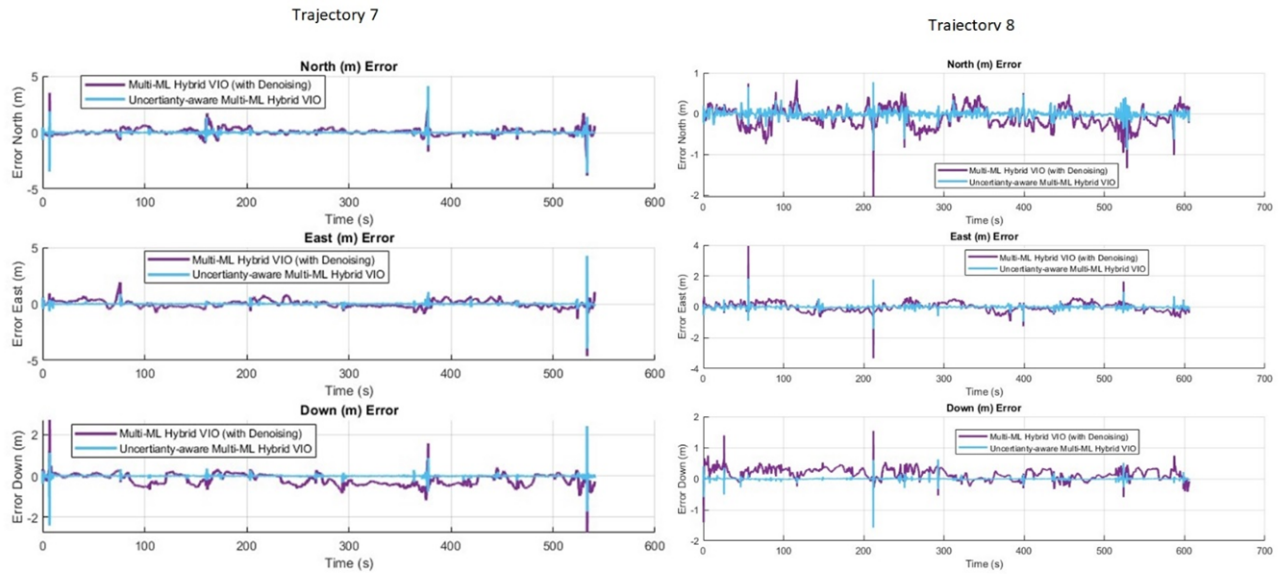


Figure 5-4 Position error comparison for trajectory 7 and trajectory 8 under unseen scenario and seen scenarios with complex dynamics

Table 5-3 RMSE and 95th percentile comparison on trajectory 7 and 8 under sunny conditions

Trajectory 7										
Algorithms	RMSE(m)					95 th Percentile(m)				
	N	E	D	3D	Horizontal	N	E	D	Horizontal	3D
VO	13.28	25.82	24.89	39.58	29.04	10.94	6.07	46.62	53.41	66.02
ESKF VIO with No Correction	11.69	23.71	24.30	35.91	26.44	14.36	8.69	43.65	49.63	61.68
Muli-ML Hybrid VIO (Denoising)	0.31	0.32	0.32	0.55	0.45	0.53	0.55	0.18	0.73	0.85
Uncertainty-aware Multi-ML Hybrid VIO	0.12	0.10	0.07	0.17	0.16	0.08	0.07	0.04	0.15	0.16

Trajectory 8										
Algorithms	RMSE(m)					95 th Percentile(m)				
	N	E	D	3D	Horizontal	N	E	D	Horizontal	3D
VO	9.80	7.60	21.08	24.46	13.83	22.36	13.48	34.90	23.15	37.64
ESKF VIO with No Correction	4.39	5.28	18.42	20.64	11.52	20.7	10.52	29.4	18.53	32.15
Muli-ML Hybrid VIO (Denoising)	0.25	0.27	0.24	0.38	0.38	0.26	0.43	0.42	0.59	0.65
Uncertainty-aware Multi-ML Hybrid VIO	0.07	0.09	0.04	0.12	0.12	0.11	0.10	0.02	0.21	0.21

The performance trend that emerges from the results confirms that uncertainty-aware Multi-ML hybrid VIO allows for compensation for the epistemic uncertainty arising from complex flight dynamics (e.g. rotation, curved turns, altitude variation).

5.5.1.2 Uncertainty Estimation on various fault factors

In this section, a set of experiments is designed to prove the quality of uncertainty prediction intervals provided by the uncertainty-aware Multi-ML hybrid VIO. A quantitative insight about the estimated uncertainties is summarised in Table 5-4 for eight flight test trajectories involving different scenarios and fault conditions. Estimation of OR% is determined based on comparing GT error with the 95th percentile and 99th percentile UI. It is highlighted that UI for trajectory 1 demonstrates the least performance compared to all the trajectories, where mean values remain smaller, indicating under-confidence in the uncertainty estimated.

In contrast, trajectory 4 outperforms among all proving lowest 95th percentile UI and 99th percentile UI intervals. However, comparison between 95th percentile UI and 99th percentile indicates that the proposed solution captures uncertainty in model well providing approximately reduction in 8%, 7%, 4% across N, E, D coordinate for OR% compared to 99th percentile UI.

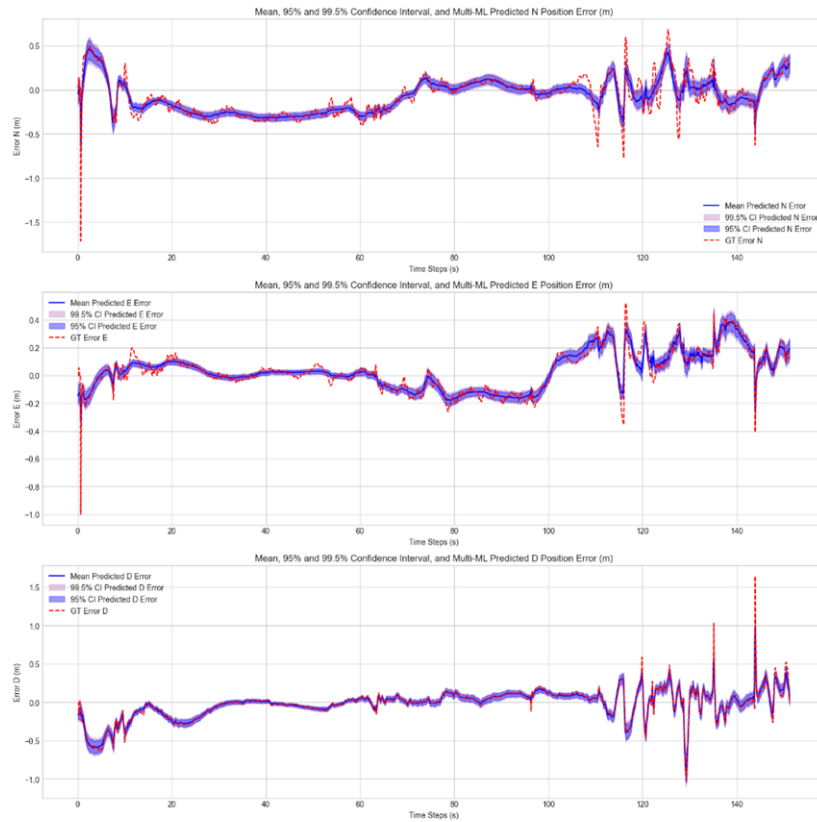
Table 5-4 Uncertainty estimation for the testing trajectories

# Trajectory	Mean (m)			UI-95 th percentile (OR%)			UI-99 th percentile (OR%)		
	N	E	D	N	E	D	N	E	D
1	0.16	0.11	0.12	36.84	33.80	6.75	27.71	24.27	4.03
2	0.20	0.25	0.27	4.42	25.18	19.33	3.20	18.52	13.14
3	0.31	0.38	0.19	22.43	21.19	6.36	14.67	11.67	3.86
4	0.72	0.54	0.42	11.47	13.82	2.13	6.17	8.82	1.22
5	0.17	0.41	0.18	24.35	11.54	4.30	16.39	6.97	2.49
6	0.26	0.39	0.19	17.94	16.43	37.61	9.39	8.23	20.94
7	0.21	0.28	0.18	27.60	26.94	10.52	17.49	29.19	3.72
8	0.20	0.20	0.19	18.69	18.17	4.94	12.33	11.95	3.25

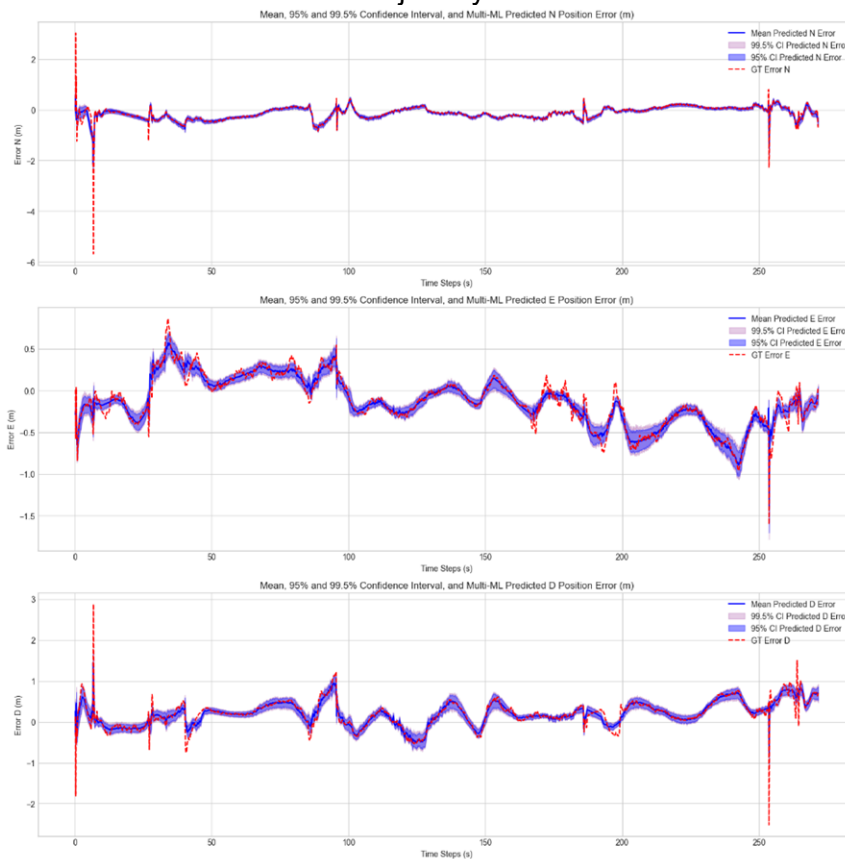
Figure 5-5 illustrates the uncertainty boundaries of trajectory 1, trajectory 2, trajectory 3 and trajectory 5. This result states a qualitative overview of the proposed Bayesian GRU within Multi-ML VIO effectively provides error boundaries based on the 95th percentile UI and 99th percentile UI. The uncertainty estimation model provides UI, ensuring ground truth errors mostly remain within the boundaries, providing safety assurance. Four trajectories are selected to evaluate the model under less fault conditions while characterised by lighting variations: sunny (trajectory 1), shadow (trajectory 3), sunrise (trajectory 5), fog (trajectory 2). Comparative analysis highlights the model's efficiency in providing uncertainty measures with high confidence, even in challenging conditions like limited visibility, reflection, lighting variation, and darkness. However, after 100s in trajectory 1, where the UAV flies over a rooftop with no detectable features, the

model UIs indicate a reduction in confidence with OR% 36.84%, 33.80%, 6.75% across N, E, D codenames. This reduction reflects the model adaptability by providing approximately 75% assurance of error boundaries.

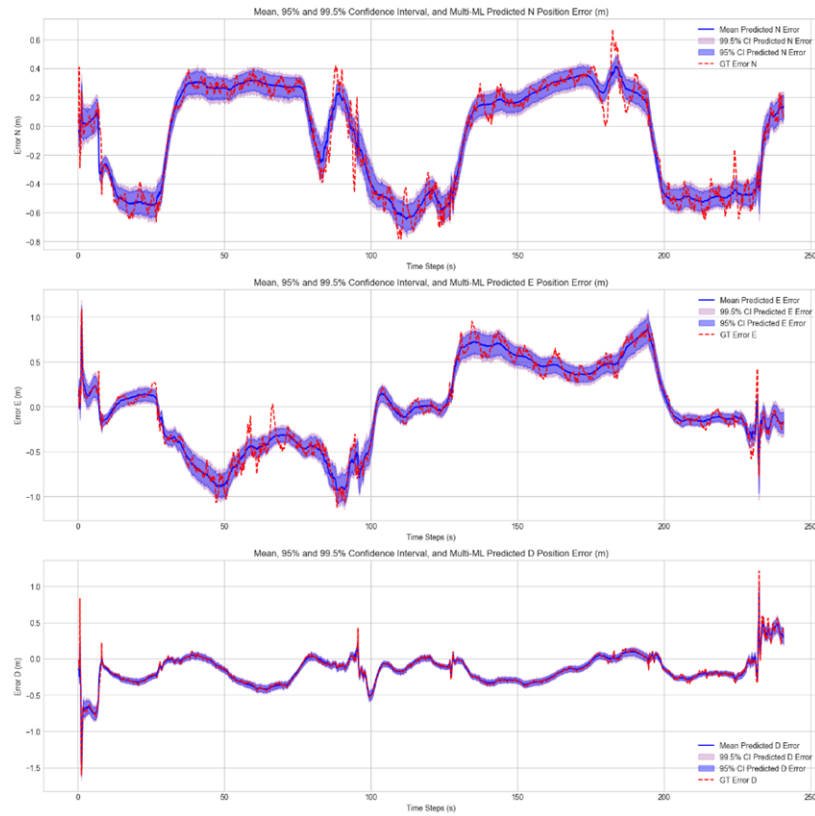
Figure 5-6 illustrates qualitative uncertainty boundaries in the presence of complex flight dynamics with combined fault conditions on trajectory 4, trajectory 6, trajectory 7 and trajectory 8. This comparison represents model generation ability over seen and unseen scenarios with high confidence measures. Trajectory 4 and trajectory 7 are selected as the most complex testing datasets in terms of evaluating generalisation ability over unknown locations covering fault conditions, altitude variation, rotation, turns with complex scenarios like less-features. The uncertainty-aware Multi-ML VIO ensures the confidence boundary in such complex flight dynamics and provides estimation of the lowest 95th percentile UI, including 11.47%, 13.82%, 2.13% in OR% for trajectory 4 and 27.60%, 26.94%, 10.52% in OR% for trajectory 7. Additionally, the results of trajectory 8 confirm that the proposed model is able to capture increased uncertainty due to complexity (altitude variability, curved turns, loop in the flight), resulting in degradation in Multi-ML VIO (denoising), which is shown in the previous section. In contrast, the model demonstrates lower OR% values in D coordinate of trajectory 6, which is related to combinational faults such as darkness, reflection, and multiple turns. The model captures the effects of this data variability on the predefined ESKF model and provides a maximum error boundary. Thus, the solution improves performance in such scenarios by predicting error UI, providing integrity-enhanced robust navigation.



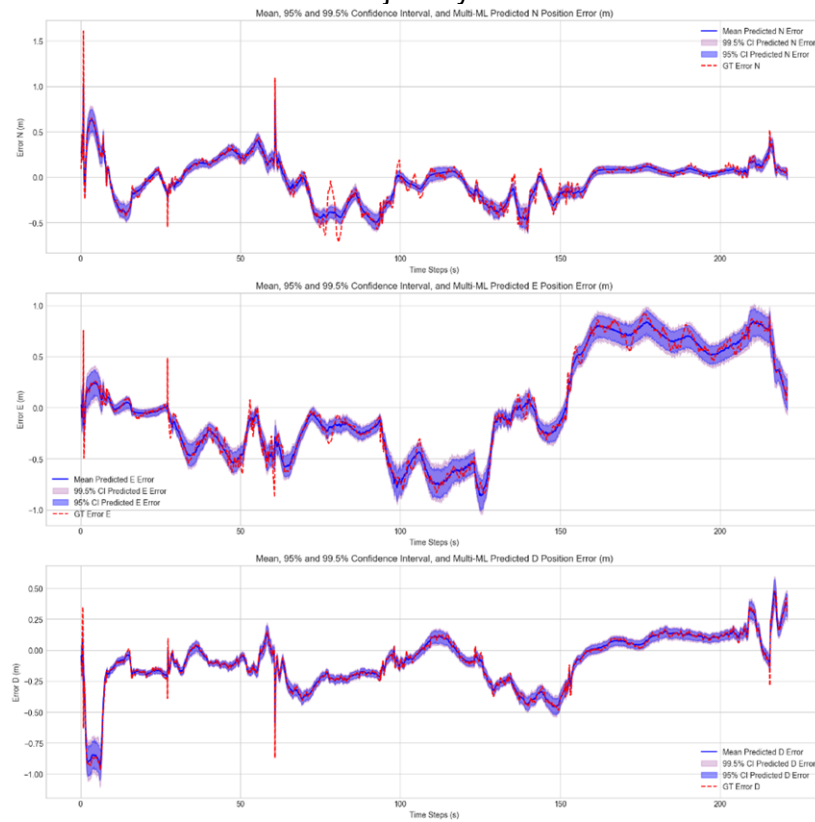
Trajectory 1



Trajectory 2

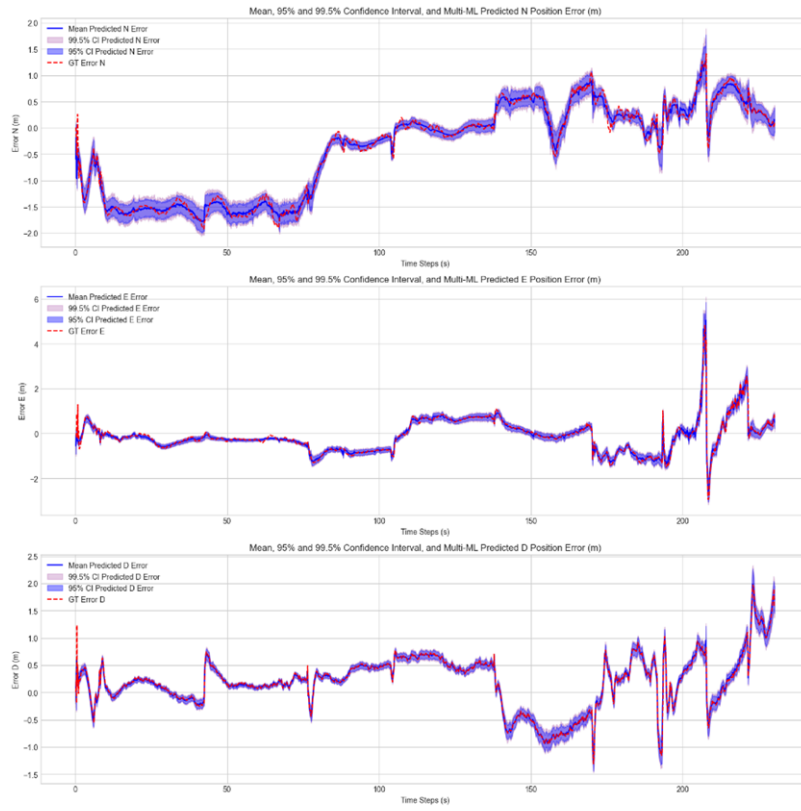


Trajectory 3

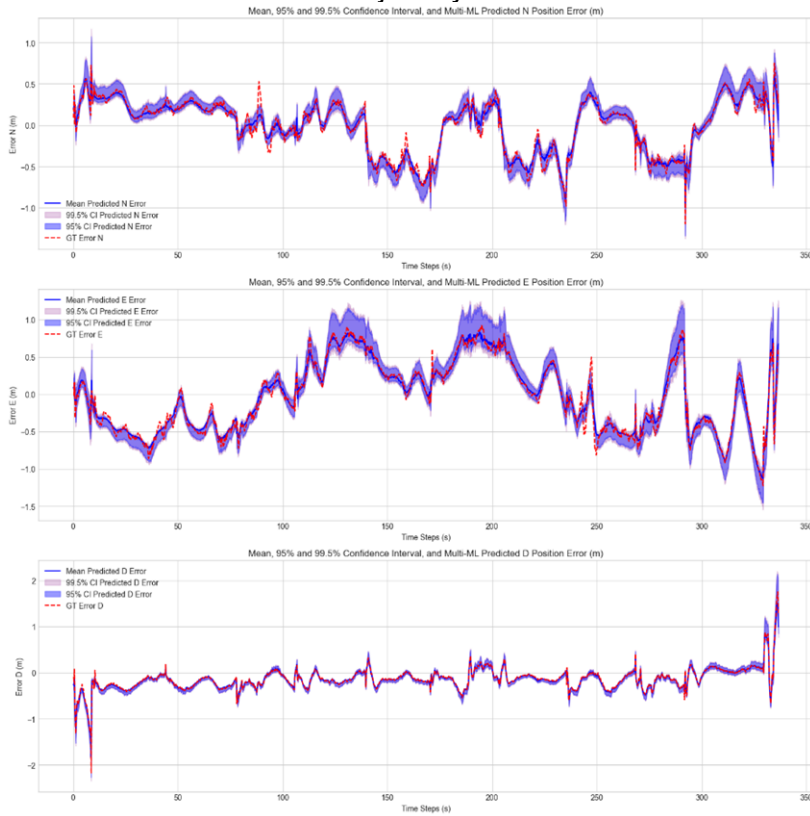


Trajectory 5

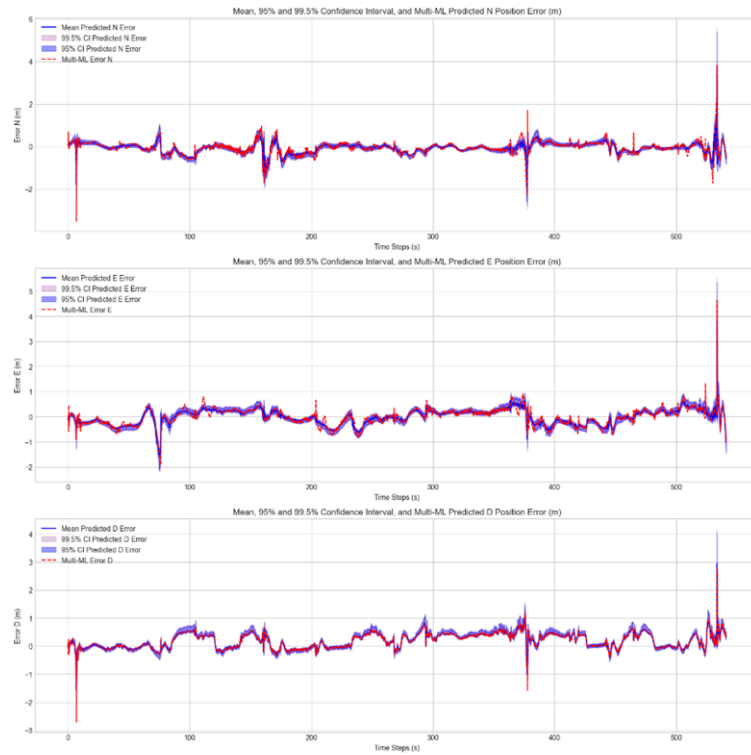
Figure 5-5 Uncertainty comparison based on 95th percentile UI and 99th percentile UI on trajectory 1, trajectory 2, trajectory 3, and trajectory 5.



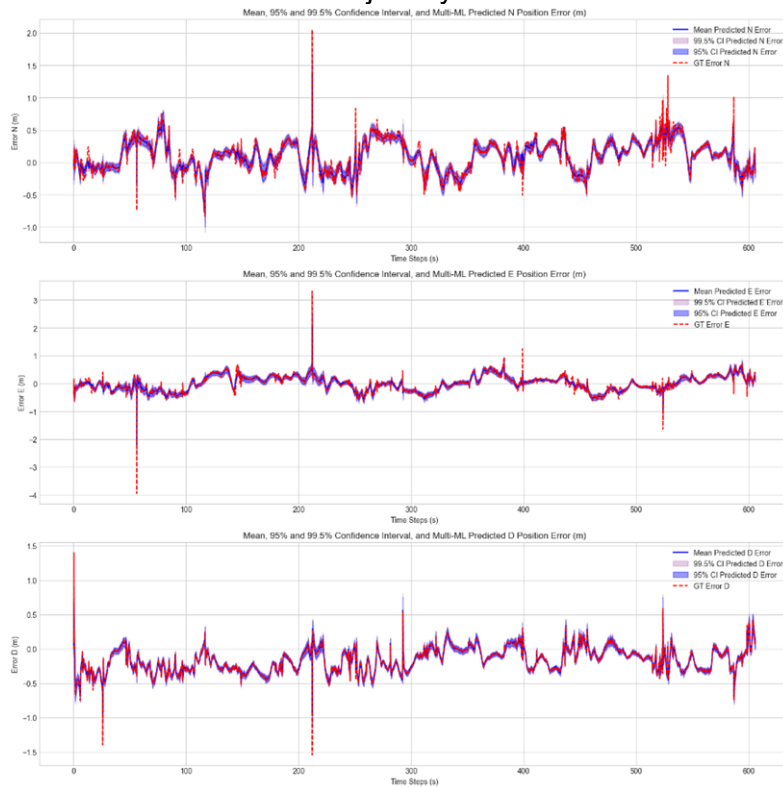
Trajectory 4



Trajectory 6



Trajectory 7



Trajectory 8

Figure 5-6 Uncertainty comparison based on 95th percentile UI and 99th percentile UI on trajectory 4, trajectory 6, trajectory 7, and trajectory 8.

5.5.1.3 Comparison with literature

In order to show an empirical comparison on uncertainty estimation to the closest related methods in the literature, the chapter selects EUROC dataset, especially MH05_difficult, for benchmarking to evaluate the performance of the proposed novel multi-ML hybrid VIO framework. The EUROC dataset was collected with MAV flying in indoor environments, leveraging complex scenarios i.e. darkness, illumination variation, rapid motion, and blur. The EUROC datasets MH01-MH04 series are used to train the model. There is a gap in evaluating the performance of models using the entire EuRoc datasets due to the complexity of the data (darkness, rapid motion, vehicle dynamics). Unlike existing solutions concerning uncertainty in VO systems, where EuRoc sequences were split into 100 test samples, this study evaluates the performance of on entire datasets involving 2273 frames, 22212 samples in IMU data. Linear interpolation is used to synchronise both sensors for fusion.

Figure 5-7 represents uncertainty estimation for MH_05 sequence in EUROC dataset. Difficulties arise from harsh motion dynamics, increased degree of freedom and image blur. The proposed model is able to capture the uncertainty, providing an uncertainty boundary in each coordinate. The GT error is estimated similar way in the previous section. The existing solution explored 3D position for uncertainty boundaries by replacing observations with Gaussian noise and estimating 3σ boundaries. However, the nonlinearity in such datasets is not Gaussian, where 3σ boundaries cannot provide confidence measures. Thus, this study has utilised the 95th percentile and 99th percentile UI for evaluation. Table 5-5 represents the comparative analysis only includes hybrid VIO architectures in SOTA. The proposed method has effectively mitigated erroneous measurements arising from various sources of uncertainty, with a reduction of mean error to 0.04, 0.05, 0.03 across N, E, D coordinates. Additionally, a comparison between the UI of the 95th percentile and the 99th percentile has demonstrated that approximately 10% of the data remain out of range compared to GT.

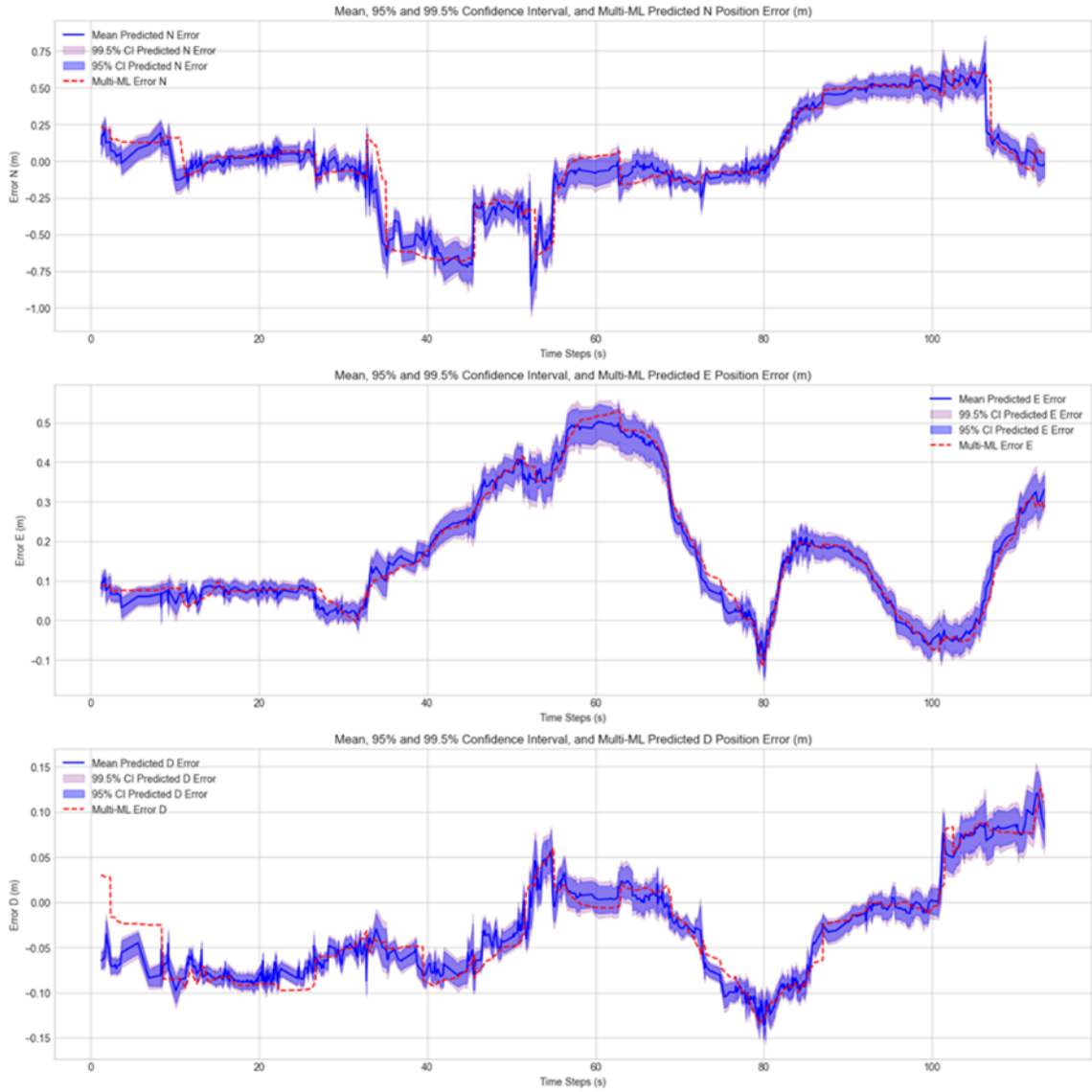


Figure 5-7 Uncertainty estimation in EuRoc MH05- difficult dataset

Table 5-5 Comparative analysis of uncertainty estimation with SOTA architecture

Algorithm	Mean (m)			UI-95 th percentile (OR%)			3 σ (OR%)	UI-99 th percentile (OR%)		
	N	E	D	N	E	D	3D	N	E	D
UA-VO [60]	-	-	-	-	-	-	21.95	-	-	-
Proposed solution	0.04	0.05	0.03	15.89	13.67	29.68	-----	10.71	8.50	19.95

5.5.2 Multi-Sensor Navigation System Performance

This section evaluates the performance of the proposed uncertainty-aware Multi-ML hybrid VIO integration with the multi-sensor navigation system, illustrating on Figure 5-1 under GNSS degraded environments. The experiment is conducted on two testing datasets, trajectory 9 and trajectory 6, under sunny weather conditions with GNSS-fault conditions applied.

Figure 5-8 illustrates the performance comparison on the integrated multi-sensor solution with proposed solutions, which indicate overall performance enhancements by reducing GNSS fault effects. The integrated system resulted in position enhancements of 96% compared with faulted-GNSS and 91% compared with GNSS/IMU ESKF fusion (sub-filter 1) in 3D position for trajectory 6, where no-texture walls, shadows with fewer turns are prominent fault conditions. In contrast, performance deviation was observed in trajectory 7 multi-sensor navigation system compared to trajectory 6 due to complex flight dynamics (turns and rotation) that accumulated GNSS and IMU error, resulting least improvement in sub-filter 2. However, the integrated multi-sensor navigation system mitigates errors accumulated with flight dynamics and improves overall performance.

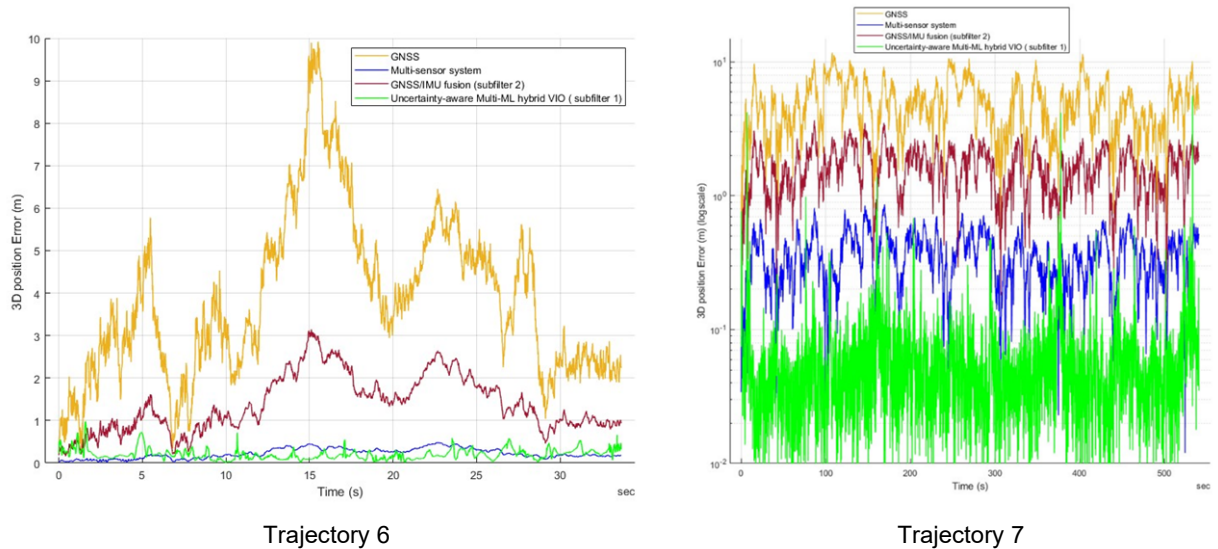


Figure 5-8 3D error comparison on integrated multi-sensor navigation system for trajectory 6 and trajectory 7

Table 5-6 illustrates RMSE comparison results on two trajectories, providing insight into the performance enhancement achieved with the integration of proposed vision-based APNT solutions within multi-sensor navigation system architecture. It is observed that challenges arise in the performance of faulty GNSS signal with the change of flight patterns. Thereby, flight dynamics remain challenging conditions while operating in urban areas due to flying between narrow or wide pathways. The integrated navigation solution demonstrates performance enhancement and increases accuracy in D coordinate by 97% for trajectory 7 and 99% for trajectory 6, mitigating driving issues in urban areas.

Table 5-6 RMSE comparison results on Trajectory 6 and 7 using integrated multi-sensor navigation system

Algorithms	Trajectory 7				
	N (m)	E (m)	D(m)	Horizontal (m)	3D(m)
GNSS (faulted)	2.76	3.91	3.51	4.91	5.36
GNSS/IMU fusion	1.95	1.11	2.84	2.25	3.96
Uncertainty-aware Multi-ML hybrid VIO	0.12	0.10	0.07	0.17	0.16
Multi-sensor system	0.14	0.17	0.09	0.22	0.23
	Trajectory 6				
	N (m)	E (m)	D(m)	Horizontal (m)	3D(m)
GNSS (faulted)	3.01	3.15	3.45	2.75	5.79
GNSS/IMU fusion	1.01	2.29	1.85	2.65	3.79
Uncertainty-aware Multi-ML hybrid VIO	0.15	0.10	0.09	0.19	0.21
Multi-sensor system	0.10	0.14	0.04	0.11	0.12

5.6 Conclusion

This study presents a proposed uncertainty-aware multi-ML hybrid VIO integrated into the multi-sensor fusion framework that significantly addresses data and model uncertainty and provides a confidence interval of bounded errors to

enhance navigation integrity. The framework simultaneously predicts uncertainty using Bayesian GRU and measurement error increments, that is utilised to obtain overall corrected, robust and integrity-enhanced VIO position. Furthermore, the proposed model utilises Monte Carlo sampling method with Bayesian hyperparameter optimisation to handle the predictive uncertainty of flight dynamics, weather effects, lighting effects and various sources of uncertainty induced by additional noise. A series of experiments were conducted to validate the proposed vision-based APNT approach under diverse conditions, including adverse weather, challenging light and high flight dynamics. The solution is able to capture introduced visual sources of uncertainties such as feature extraction error, feature tracking error, data association error, mismatch in dynamics, unmodelled noise and imperfect modelling and provide UI with 95th percentile confidence boundary. The testing and training datasets included realistic failure modes and fault factors representative of urban environments. Notably, it demonstrated superiority in capturing complex flight dynamics in unseen locations (trajectory 4 and trajectory 7) by proving the highest confidence measure with 2.13% in D OR% compared to trajectory 1 with minimal fault conditions. Additionally, trajectory 1 results highlight the proposed framework's ability to capture uncertainty and reduce position error while removing additional induced unmodelled noise, even when sensor data is corrupted or the model is imperfect. Additionally, experiments evaluating the effect of weather conditions and lighting effects revealed the frameworks' ability to capture uncertainty caused by the limited visibility of features, reflection, shadow, and flight dynamics, illustrated with performance in trajectory 5, 6 and 2. Thus, the approximation of OR% in these trajectories remains within 80% confidence intervals, providing safety assurance while navigating in a time of day, including weather effects (fog). Lastly, the proposed APNT solution integrates multi-sensor navigation, provides assurance with increased under-faulted-GNSS conditions, and assures 97% enhancement under complex flight dynamics in urban areas. In summary, the study provides an integrity-enhanced robust vision-based framework that ensures uninterrupted and safe navigation under complex urban environments, enhancing a step forward to employ in safety-critical UAM applications. Future

work will focus on enhancing the A-PNT framework by integrating additional sensors to mitigate vertical drift in high-flight dynamic scenarios.

6 General Discussion

In chapters (2 to 5), the details of various workflows, e.g. strategy on visual failure mode mitigation, vision-based APNT fusion, multi-sensor fusion, uncertainty inspection, etc related to enhancing integrity and robustness of vision-based navigation for UAVS under urban environments have been thoroughly described. The focus of this particular chapter is to explore and uncover any information that has not been discussed previously in earlier chapters. Furthermore, an analysis is conducted to evaluate the completeness of all objectives, ensuring all aspects of robust visual navigation for UAVS with safety assurance are covered.

All research objectives, which include visual failure mode mitigation mechanisms, developing robust and resilient vision-based fusion frameworks under complex scenarios for UAM applications with safety assurance by providing the confidence boundary to ensure uninterrupted flight during GNSS outage, have been successfully met. Details can be provided as follows:

- a) In the background or related work section of each chapter, the state-of-the-art works for each module, e.g. visual failure mode mitigation, evaluation of superior hybrid approaches in visual fault mitigation, robust vision-based APNT fusion framework, and uncertainty inspection to ensure integrity of APNT solutions within multi-sensor navigation, are described. The comprehensive review includes highlighting the advantages of proposed solutions compared to existing approaches. For example, proposed APNT solutions may include fault mitigation strategies by designing training datasets with complex fault conditions that ensure robustness and quantify tight error boundaries for safe navigation.
- b) A strategic approach for mitigating visual failure mode is proposed, utilising an FMEA-supported hybrid multi-sensor framework that integrates proposed VO error compensation within hybrid VIO as one of the sub-filters. The solution can effectively handle failure modes arising from navigation environments, such as feature extraction errors, feature tracking errors, and data association errors in urban environments. Therefore, this approach effectively bridges the gap during GNSS outage

- in challenging environments, enabling new research direction with the utilisation of failures in vision navigation to ensure operational continuity.
- c) A novel proposition of error compensation methodology within hybrid VIO is developed, with the aim of evaluating error compensation mechanisms in hybrid architectures, mitigating visual failure modes. This comparative analysis provides insight into model efficiency, robustness and adaptation in a diverse range of fault conditions arising from weather effects, lighting conditions, illumination variation and flight dynamics. In challenging conditions, this approach successfully addresses the gap in visual fault mitigation using error compensation mechanisms.
 - d) A new approach for uncertainty prediction of complex visual failure modes is proposed, providing confidence measures and compensating epistemic uncertainty to ensure robust and safe navigation in urban areas. In GNSS denied environments with visually degraded scenarios, this approach successfully mitigates visual uncertainty sources and improves performance while integrated into the multi-sensor system.
 - e) Systematic evaluation of proposed error compensation mechanisms in the thesis study on different hybrid VIO architectures in the mitigation of visual failure modes is conducted. Following this comparative analysis, the best solution is integrated into the multi-sensor navigation system that enhances performance during GNSS outage.

In conclusion, the stated objectives in Chapter 1 have been achieved. Through diligent efforts and effective strategies, the desired outcomes have been achieved. To provide additional insight into the perspective of the research, which is focused on only visual fault mitigation methodology to ensure the integrity of overall systems during GNSS outages. The exploration of additional external APNT sensors, e.g. radar, lidar, UWB, has not been explored in this study due to the challenges associated with sensor capabilities. However, some approaches utilised a barometer for vertical take-off and landing [145] due to the VO scaled factor error and drift. This study has effectively mitigated vertical error and drift, providing centimetre accuracy in high-dynamics flight paths. Thus, there is no need to explore this sensor.

Then, it is crucial to evaluate the capabilities of each machine learning model while integrated into real-time applications. From the post-processing point of view, each GRU-based compensation method interface time of approximately 0.2ms for training and 4.5minutes for training for each dataset. In contrast, Bayesian GRU incorporating MCD interface time approximately 5hours for training. Notably, GRU hyperparameter tuning took approximately 44.5mins for each dataset. The overall interface time for both error correction activated is 30ms for testing. To facilitate the deployment, the proposed algorithm was developed and run on a workstation configuration of 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz 2.30 GHz. The proposed approaches in this study can be implemented on UAVS for real-time implementation.

The adaptability of each module within the framework can be ensured. The generalisation ability of the modules proposed outperforms the state-of-the-art system. For instance, authors in [43], [44] stated the necessity of vast training to evaluate their solutions under degraded environments, where they utilised 11 sets of Euroc datasets for training and testing on the same model. In contrast, authors in [41], [43], [60] split the testing trajectories (100 frames each iteration) to achieve high accuracy. Some studies focusing on filter-based methods struggle to evaluate their solution using Euroc MH05-difficult dataset. However, the proposed error compensation mythologies outperformed utilising only four sequences for training, while tested on the most complex dataset in Euroc. Additionally, the proposed Bayesian GRU with MCD outperformed when evaluated with the same dataset compared to the state-of-the-art solution[60] where the uncertainty interval is wider, reflecting a degraded prediction of the confidence interval.

In addition to its capability, the application and supported services can be extended to cater to other stakeholders. This means that the system has the capability to accommodate the needs of various users and entities beyond its initial scope. For instance, the Milton Keynes city council intend to expand and enhance their smart city infrastructure, which can provide a possibility for future development of this approach. Additionally, the UK civil aviation

authority has selected 6 new projects covering autonomous drone delivery. Amazon Prime Air and Royal Mail are finalising their trials to ensure autonomous package delivery, where they struggle with ensuring flight safety. This finding and the investigations carried out throughout this research can be seamlessly applied to their specific scenario, offering valuable insight and a solution. Moreover, the research supports the notion of continuous improvement and advancement, as it can serve as a basis for further revision and enhancement of optimised outcomes and address specific needs for the smart city. Furthermore, the proposed vision-based algorithm integrated within the multi-sensor system assures navigation safety during GNSS outage under diverse urban scenarios like weather conditions, lighting effects and flight dynamics.

In conclusion, all objectives outlined in this research have been accomplished, including the implementation of visual failure mode mitigation methodologies that can operate in different scenarios with operational safety tailored to applications such as aerial survey, cargo drones and passenger taxis.

7 Conclusion and Future Work

This thesis focuses on developing robust and accurate vision-based APNT solutions for UAVS to ensure safe and secure operation in urban environments. At the beginning of the work, visual failure modes causing state-of-the-art system failure are addressed using error compensation mechanisms in hybrid VIO. Notably, the sensor error compensation mechanism outperforms compared to KF-based model error compensation due to the inheritance of nonlinearity in data, such as aleatoric uncertainty, hindering predefined ESKF model performance. Again, other hybrid approaches like K-gain prediction struggle due to increased noise in the sensor data. The potential visually degraded scenarios, like weather effects, lighting conditions and UAV dynamics effects, are reviewed and addressed with the proposition of combinational error compensation mechanisms within hybrid VIO that improve performance with the prediction of complex visual uncertainty. Throughout the studies, various proposed error compensation mechanisms are investigated to ensure robustness that enhances integrity within the overall solution. In order to assess the performance of the proposed methods, it has been integrated into multi-sensor navigation and evaluated during GNSS outage.

It is observed that VO error correction mechanisms significantly contribute to visual fault mitigation, while KF measurement correction performance depends on this. Thereby, this study suggests employing sensor measurement error correction within the navigation system if it has been integrated with other combinations of sensors.

In future work, there are several avenues to explore and improve upon the current research. Firstly, refining and expanding the proposed methodologies can enhance their performance and applicability. This could involve incorporating additional data sources to enhance the predictive capabilities. Additionally, conducting extensive field experiments and gathering more real-world data can provide valuable insights and validate the effectiveness of the proposed approaches in diverse operational environments. However, further investigation on visual integrity can provide safety assurance during wind turbulence. Furthermore,

exploring integration technologies such as Internet of Things (IOT) and cloud computing can potentially enhance capabilities while integrating into the real drone. Lastly, collaboration with industry partners and stakeholders can facilitate the development the deployment and implementation of the developed solution in real-world scenarios, enabling practical impact in the field of urban air mobility.

Reference

- [1] L. A. Garrow, B. J. German, and C. E. Leonard, "Urban air mobility: A comprehensive review and comparative analysis with autonomous and electric ground transportation for informing future research," *Transp Res Part C Emerg Technol*, vol. 132, Nov. 2021, doi: 10.1016/J.TRC.2021.103377.
- [2] H. Pak *et al.*, "Can Urban Air Mobility become reality? Opportunities and challenges of UAM as innovative mode of transport and DLR contribution to ongoing research," *CEAS Aeronaut J*, pp. 1–31, May 2024, doi: 10.1007/S13272-024-00733-X/FIGURES/7.
- [3] G. Brunner, B. Szebedy, S. Tanner, and R. Wattenhofer, "The urban last mile problem: Autonomous drone delivery to your balcony," in *2019 International Conference on Unmanned Aircraft Systems, ICUAS 2019*, 2019. doi: 10.1109/ICUAS.2019.8798337.
- [4] Y. Chang, Y. Cheng, U. Manzoor, and J. Murray, "A review of UAV autonomous navigation in GPS-denied environments," *Rob Auton Syst*, vol. 170, p. 104533, Dec. 2023, doi: 10.1016/J.ROBOT.2023.104533.
- [5] D. A. Grejner-Brzezinska, C. K. Toth, T. Moore, J. F. Raquet, M. M. Miller, and A. Kealy, "Multisensor Navigation Systems: A Remedy for GNSS Vulnerabilities?," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1339–1353, Jun. 2016, doi: 10.1109/JPROC.2016.2528538.
- [6] S. Bijjahalli, R. Sabatini, and A. Gardi, "GNSS Performance Modelling and Augmentation for Urban Air Mobility," *Sensors 2019, Vol. 19, Page 4209*, vol. 19, no. 19, p. 4209, Sep. 2019, doi: 10.3390/S19194209.
- [7] S. Hosseinyalamdary, "Deep Kalman filter: Simultaneous multi-sensor integration and modelling; A GNSS/IMU case study," *Sensors (Switzerland)*, vol. 18, no. 5, May 2018, doi: 10.3390/s18051316.
- [8] B. GAO, W. LI, G. HU, Y. ZHONG, and X. ZHU, "Mahalanobis distance-based fading cubature Kalman filter with augmented mechanism for

- hypersonic vehicle INS/CNS autonomous integration,” *Chinese Journal of Aeronautics*, vol. 35, no. 5, pp. 114–128, May 2022, doi: 10.1016/J.CJA.2021.08.035.
- [9] I. A. Ali, M. Mahmoud, I. Alaa, A. Wassal, A. Noureldin, and A. Eldieb, “Tuning of The Error Covariance Parameters in EKF-Based INS/GPS Systems: A Practical Approach,” 2017. [Online]. Available: <https://www.researchgate.net/publication/316753131>
- [10] P. Geragersian, I. Petrunin, W. Guo, and R. Grech, “An INS/GNSS fusion architecture in GNSS denied environments using gated recurrent units,” *AIAA Science and Technology Forum and Exposition, AIAA SciTech Forum 2022*, 2022, doi: 10.2514/6.2022-1759.
- [11] Y. Tang, J. Jiang, J. Liu, P. Yan, Y. Tao, and J. Liu, “A GRU and AKF-Based Hybrid Algorithm for Improving INS/GNSS Navigation Accuracy during GNSS Outage,” *Remote Sens (Basel)*, vol. 14, no. 3, Feb. 2022, doi: 10.3390/rs14030752.
- [12] W. I. Liu, Z. Li, and Z. Zhang, “Error Modelling and Optimal Estimation of Laser Scanning Aided Inertial Navigation System in GNSS-Denied Environments,” *The Journal of Navigation*, vol. 72, no. 3, pp. 741–758, May 2019, doi: 10.1017/S037346331800084X.
- [13] Z. Qiu, J. Lv, D. Lin, Y. Yu, Z. Sun, and Z. Zheng, “A Lidar-Inertial Navigation System for UAVs in GNSS-Denied Environment with Spatial Grid Structures,” *Applied Sciences 2023, Vol. 13, Page 414*, vol. 13, no. 1, p. 414, Dec. 2022, doi: 10.3390/APP13010414.
- [14] J. Tang *et al.*, “LiDAR Scan Matching Aided Inertial Navigation System in GNSS-Denied Environments,” *Sensors 2015, Vol. 15, Pages 16710-16728*, vol. 15, no. 7, pp. 16710–16728, Jul. 2015, doi: 10.3390/S150716710.
- [15] Y. Xianjia, L. Qingqing, J. P. Queralta, J. Heikkonen, and T. Westerlund, “Cooperative UWB-based localization for outdoors positioning and

- navigation of UAVs aided by ground robots,” in *ICAS 2021 - 2021 IEEE International Conference on Autonomous Systems, Proceedings*, 2021. doi: 10.1109/ICAS49788.2021.9551177.
- [16] H.-Y. Lin and J.-R. Zhan, “GNSS-denied UAV indoor navigation with UWB incorporated visual inertial odometry,” *Measurement*, vol. 206, p. 112256, Jan. 2023, doi: 10.1016/j.measurement.2022.112256.
 - [17] W. Jiang, Z. Cao, B. Cai, B. Li, and J. Wang, “Indoor and Outdoor Seamless Positioning Method Using UWB Enhanced Multi-Sensor Tightly-Coupled Integration,” *IEEE Trans Veh Technol*, vol. 70, no. 10, pp. 10633–10645, Oct. 2021, doi: 10.1109/TVT.2021.3110325.
 - [18] X. Ye, F. Song, Z. Zhang, and Q. Zeng, “A Review of Small UAV Navigation System Based on Multisource Sensor Fusion,” *IEEE Sens J*, vol. 23, no. 17, pp. 18926–18948, Sep. 2023, doi: 10.1109/JSEN.2023.3292427.
 - [19] N. Gyagenda, J. V. Hatilima, H. Roth, and V. Zhmud, “A review of GNSS-independent UAV navigation techniques,” *Rob Auton Syst*, vol. 152, Jun. 2022, doi: 10.1016/J.ROBOT.2022.104069.
 - [20] D. J. Yeong, G. Velasco-hernandez, J. Barry, and J. Walsh, “Sensor and sensor fusion technology in autonomous vehicles: A review,” Mar. 02, 2021, *MDPI AG*. doi: 10.3390/s21062140.
 - [21] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, “Multisensor data fusion: A review of the state-of-the-art,” *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013, doi: 10.1016/j.inffus.2011.08.001.
 - [22] B. Li, C. Mu, and B. Wu, “A survey of vision based autonomous aerial refueling for unmanned aerial vehicles,” in *ICICIP 2012 - 2012 3rd International Conference on Intelligent Control and Information Processing*, 2012. doi: 10.1109/ICICIP.2012.6391480.
 - [23] A. Al-Kaff, D. Martín, F. García, A. de la Escalera, and J. María Armingol, “Survey of computer vision algorithms and applications for unmanned aerial vehicles,” 2018. doi: 10.1016/j.eswa.2017.09.033.

- [24] Y. Xu, J. Li, L. Pan, N. Jing, C. Du, and J. Wu, "Vision-based UAVs aerial image localization: A survey," in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, GeoAI 2018*, 2018. doi: 10.1145/3281548.3281556.
- [25] Y. Lu, Z. Xue, G. S. Xia, and L. Zhang, "A survey on vision-based UAV navigation," *Geo-Spatial Information Science*, vol. 21, no. 1, 2018, doi: 10.1080/10095020.2017.1420509.
- [26] T. Qin, P. Li, and S. Shen, "VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator," Aug. 2017, doi: 10.1109/TRO.2018.2853729.
- [27] G. Li, L. Yu, and S. Fei, "A Binocular MSCKF-Based Visual Inertial Odometry System Using LK Optical Flow," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 100, no. 3–4, pp. 1179–1194, Dec. 2020, doi: 10.1007/S10846-020-01222-Z/METRICS.
- [28] F. Ma, J. Shi, Y. Yang, J. Li, and K. Dai, "ACK-MSCKF: Tightly-coupled ackermann multi-state constraint kalman filter for autonomous vehicle localization," *Sensors (Switzerland)*, vol. 19, no. 21, Nov. 2019, doi: 10.3390/s19214816.
- [29] C. Campos, R. Elvira, J. J. G. Rodriguez, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874–1890, Dec. 2021, doi: 10.1109/TRO.2021.3075644.
- [30] P. Geneva, K. Eickenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation," *Proc IEEE Int Conf Robot Autom*, pp. 4666–4672, May 2020, doi: 10.1109/ICRA40945.2020.9196524.
- [31] C. Zhu, M. Meurer, and C. Günther, "Integrity of Visual Navigation—Developments, Challenges, and Prospects," *NAVIGATION: Journal of the*

- Institute of Navigation*, vol. 69, no. 2, p. 69, Jun. 2022, doi: 10.33012/NAVI.518.
- [32] Y. Du, J. Wang, C. Rizos, and A. El-Mowafy, "Vulnerabilities and integrity of precise point positioning for intelligent transport systems: overview and analysis," Dec. 01, 2021, *Springer*. doi: 10.1186/s43020-020-00034-8.
 - [33] M. Tossaint *et al.*, "The Stanford - ESA Integrity Diagram: Focusing on SBAS Integrity," Sep. 29, 2006. Accessed: Dec. 05, 2024. [Online]. Available: <http://www.ion.org/publications/abstract.cfm?jp=p&articleID=6963>
 - [34] U. I. Bhatti, W. Y. Ochieng, and S. Feng, "Integrity of an integrated GPS/INS system in the presence of slowly growing errors. Part I: A critical review," *GPS Solutions*, vol. 11, no. 3, pp. 173–181, Jul. 2007, doi: 10.1007/s10291-006-0048-2.
 - [35] H. Jiang, T. Li, D. Song, and C. Shi, "An Effective Integrity Monitoring Scheme for GNSS/INS/Vision Integration Based on Error State EKF Model," *IEEE Sens J*, vol. 22, no. 7, pp. 7063–7073, Apr. 2022, doi: 10.1109/JSEN.2022.3154054.
 - [36] Z. Yue, B. Lian, C. Tang, and K. Tong, "A novel adaptive federated filter for GNSS/INS/VO integrated navigation system," *Meas Sci Technol*, vol. 31, no. 8, p. 085102, May 2020, doi: 10.1088/1361-6501/AB78C2.
 - [37] J. H. Choi, Y. W. Park, J. Kim, T. S. Choe, and J. B. Song, "Federated-filter-based unmanned ground vehicle localization using 3D range registration with digital elevation model in outdoor environments," *J Field Robot*, vol. 29, no. 2, pp. 298–314, Mar. 2012, doi: 10.1002/ROB.21416.
 - [38] A. Yusefi, A. Durdu, M. F. Aslan, and C. Sungur, "LSTM and filter based comparison analysis for indoor global localization in UAVs," *IEEE Access*, vol. 9, pp. 10054–10069, 2021, doi: 10.1109/ACCESS.2021.3049896.

- [39] J. Dai, X. Hao, S. Liu, and Z. Ren, "Research on UAV Robust Adaptive Positioning Algorithm Based on IMU/GNSS/VO in Complex Scenes," *Sensors (Basel)*, vol. 22, no. 8, Apr. 2022, doi: 10.3390/S22082832.
- [40] N. Gu, F. Xing, and Z. You, "Visual/Inertial/GNSS Integrated Navigation System under GNSS Spoofing Attack," *Remote Sens (Basel)*, vol. 14, no. 23, p. 5975, Nov. 2022, doi: 10.3390/rs14235975.
- [41] M. F. Aslan, A. Durdu, A. Yusefi, and A. Yilmaz, "HVIONet: A deep learning based hybrid visual-inertial odometry approach for unmanned aerial system position estimation," *Neural Networks*, vol. 155, pp. 461–474, Nov. 2022, doi: 10.1016/J.NEUNET.2022.09.001.
- [42] J. Liao, X. Li, X. Wang, S. Li, and H. Wang, "Enhancing navigation performance through visual-inertial odometry in GNSS-degraded environment," *GPS Solutions*, vol. 25, no. 2, Apr. 2021, doi: 10.1007/s10291-020-01056-0.
- [43] Y. Almalioglu, M. Turan, R. U. Saputra, P. P. B. De Gusmão, A. Markham, and N. Trigoni, "SelfVIO: Self-supervised deep monocular Visual-Inertial Odometry and depth estimation," *Neural Networks*, vol. 150, pp. 119–136, 2022, doi: 10.1016/j.neunet.2022.03.005.
- [44] C. Li and S. L. Waslander, "Towards End-to-end Learning of Visual Inertial Odometry with an EKF," *Proceedings - 2020 17th Conference on Computer and Robot Vision, CRV 2020*, pp. 190–197, May 2020, doi: 10.1109/CRV50864.2020.00033.
- [45] B. Wagstaff, E. Wise, and J. Kelly, "A Self-Supervised, Differentiable Kalman Filter for Uncertainty-Aware Visual-Inertial Odometry," *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, vol. 2022-July, pp. 1388–1395, 2022, doi: 10.1109/AIM52237.2022.9863270.
- [46] A. Ben Afia, A. Escher, and C. Macabiau, "A Low-cost GNSS/IMU/Visual monoSLAM/WSS Integration Based on Federated Kalman Filtering for Navigation in Urban Environments," 2015.

- [47] M. Bloesch, M. Burri, S. Omari, M. Hutter, and R. Siegwart, "Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback," *International Journal of Robotics Research*, vol. 36, no. 10, pp. 1053–1072, Sep. 2017, doi: 10.1177/0278364917728574/ASSET/IMAGES/LARGE/10.1177_0278364917728574-FIG20.JPEG.
- [48] X. QIU, H. ZHANG, and W. FU, "Lightweight hybrid visual-inertial odometry with closed-form zero velocity update," *Chinese Journal of Aeronautics*, vol. 33, no. 12, pp. 3344–3359, Dec. 2020, doi: 10.1016/J.CJA.2020.03.008.
- [49] M. F. Aslan, A. Durdu, and K. Sabanci, "Visual-Inertial Image-Odometry Network (VIIONet): A Gaussian process regression-based deep architecture proposal for UAV pose estimation," *Measurement*, vol. 194, p. 111030, May 2022, doi: 10.1016/J.MEASUREMENT.2022.111030.
- [50] Y. Fu, S. Wang, Y. Zhai, and X. Zhan, "Visual odometry errors and fault distinction for integrity monitoring," *AerSy*, vol. 3, no. 4, pp. 265–274, Dec. 2020, doi: 10.1007/S42401-020-00062-X.
- [51] Y. Fu, S. Wang, Y. Zhai, X. Zhan, and X. Zhang, "Measurement Error Detection for Stereo Visual Odometry Integrity," *NAVIGATION: Journal of the Institute of Navigation*, vol. 69, no. 4, Dec. 2022, doi: 10.33012/NAVI.542.
- [52] S. Wang, X. Zhan, Y. Fu, and Y. Zhai, "Feature-based visual navigation integrity monitoring for urban autonomous platforms," *Aerospace Systems*, vol. 3, no. 3, pp. 167–179, Sep. 2020, doi: 10.1007/S42401-020-00057-8/FIGURES/7.
- [53] Y. X. Tan, M. B. Prasetyo, M. A. Daffa, D. S. Nitin, and M. Meghjani, "Evaluating Visual Odometry Methods for Autonomous Driving in Rain," *IEEE International Conference on Automation Science and Engineering*, vol. 2023-August, Sep. 2023, doi: 10.1109/CASE56687.2023.10260549.

- [54] Y. Almalioglu, M. Turan, N. Trigoni, and A. Markham, "Deep learning-based robust positioning for all-weather autonomous driving," *Nat Mach Intell*, vol. 4, no. 9, pp. 749–760, Sep. 2022, doi: 10.1038/S42256-022-00520-5.
- [55] M. Burri *et al.*, "The EuRoC micro aerial vehicle datasets," *International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, Sep. 2016, doi: 10.1177/0278364915620033/ASSET/IMAGES/LARGE/10.1177_0278364915620033-IMG2.JPEG.
- [56] C. Li and S. L. Waslander, "Visual Measurement Integrity Monitoring for UAV Localization," *2019 IEEE International Symposium on Safety, Security, and Rescue Robotics, SSRR 2019*, pp. 22–29, Sep. 2019, doi: 10.1109/SSRR.2019.8848975.
- [57] C. Zhu, M. Joerger, and M. Meurer, "Quantifying Feature Association Error in Camera-based Positioning," *2020 IEEE/ION Position, Location and Navigation Symposium, PLANS 2020*, pp. 967–972, Apr. 2020, doi: 10.1109/PLANS46316.2020.9109919.
- [58] C. R. Lee and K. J. Yoon, "Confidence analysis of feature points for visual-inertial odometry of urban vehicles," *IET Intelligent Transport Systems*, vol. 13, no. 7, pp. 1130–1138, Jul. 2019, doi: 10.1049/IET-ITS.2018.5196.
- [59] C. Zhu, C. Steinmetz, B. Belabbas, and M. Meurer, "Feature Error Model for Integrity of Pattern-based Visual Positioning," *Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2019*, pp. 2254–2268, Sep. 2019, doi: 10.33012/2019.16956.
- [60] G. Costante and M. Mancini, "Uncertainty estimation for data-driven visual odometry," *IEEE Transactions on Robotics*, vol. 36, no. 6, pp. 1738–1757, Dec. 2020, doi: 10.1109/TRO.2020.3001674.
- [61] A. C. Stutts, D. Erricolo, T. Tulabandhula, and A. R. Trivedi, "Lightweight, Uncertainty-Aware Conformalized Visual Odometry," *IEEE International*

Conference on Intelligent Robots and Systems, pp. 7742–7749, 2023, doi: 10.1109/IROS55552.2023.10341924.

- [62] J. Song, H. G. Jo, Y. Jin, and S. J. Lee, “Uncertainty-Aware Depth Network for Visual Inertial Odometry of Mobile Robots,” *Sensors* 2024, Vol. 24, Page 6665, vol. 24, no. 20, p. 6665, Oct. 2024, doi: 10.3390/S24206665.
- [63] N. Kaygusuz, O. Mendez, and R. Bowden, “MDN-VO: Estimating Visual Odometry with Confidence,” *IEEE International Conference on Intelligent Robots and Systems*, pp. 3528–3533, 2021, doi: 10.1109/IROS51168.2021.9636827.
- [64] M. Y. Arafat, M. M. Alam, and S. Moh, “Vision-Based Navigation Techniques for Unmanned Aerial Vehicles: Review and Challenges,” *Drones*, vol. 7, no. 2, Feb. 2023, doi: 10.3390/DRONES7020089.
- [65] Y. Alkendi, L. Seneviratne, and Y. Zweiri, “State of the Art in Vision-Based Localization Techniques for Autonomous Navigation Systems,” *IEEE Access*, vol. 9, pp. 76847–76874, 2021, doi: 10.1109/ACCESS.2021.3082778.
- [66] A. Ben Afia, A.-C. Escher, and C. Macabiau, “A Low-cost GNSS/IMU/Visual monoSLAM/WSS Integration Based on Federated Kalman Filtering for Navigation in Urban Environments A Low-cost GNSS/IMU/Visual monoSLAM/WSS Integration Based on Federated Kalman Filtering for Navigation in Urban A Low-cost GNSS/IMU/Visual monoSLAM/WSS Integration Based on Kalman Filtering for Navigation in Urban Environments,” 2015. [Online]. Available: <https://enac.hal.science/hal-01284973>
- [67] Y. D. Lee, L. W. Kim, and H. K. Lee, “A tightly-coupled compressed-state constraint Kalman Filter for integrated visual-inertial-Global Navigation Satellite System navigation in GNSS-Degraded environments,” *IET Radar, Sonar and Navigation*, vol. 16, no. 8, pp. 1344–1363, Aug. 2022, doi: 10.1049/rsn2.12265.

- [68] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," *Proc IEEE Int Conf Robot Autom*, pp. 3565–3572, 2007, doi: 10.1109/ROBOT.2007.364024.
- [69] Y. Zhai, Y. Fu, S. Wang, and X. Zhan, "Mechanism Analysis and Mitigation of Visual Navigation System Vulnerability," *Lecture Notes in Electrical Engineering*, vol. 773 LNEE, pp. 515–524, 2021, doi: 10.1007/978-981-16-3142-9_49.
- [70] L. Markovic, M. Kovac, R. Milijas, M. Car, and S. Bogdan, "Error State Extended Kalman Filter Multi-Sensor Fusion for Unmanned Aerial Vehicle Localization in GPS and Magnetometer Denied Indoor Environments," *2022 International Conference on Unmanned Aircraft Systems, ICUAS 2022*, pp. 184–190, Sep. 2021, doi: 10.1109/ICUAS54217.2022.9836124.
- [71] X. Xiong, W. Chen, Z. Liu, and Q. Shen, "DS-VIO: Robust and Efficient Stereo Visual Inertial Odometry based on Dual Stage EKF," May 2019, [Online]. Available: <http://arxiv.org/abs/1905.00684>
- [72] F. Fanin and J.-H. Hong, *Visual Inertial Navigation for a Small UAV Using Sparse and Dense Optical Flow*. 2019. doi: 10.0/Linux-x86_64.
- [73] K. Sun *et al.*, "Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight," *IEEE Robot Autom Lett*, vol. 3, no. 2, pp. 965–972, Apr. 2018, doi: 10.1109/LRA.2018.2793349.
- [74] Y. Yang, P. Geneva, K. Eickenhoff, and G. Huang, "Visual-Inertial Odometry with Point and Line Features," *IEEE International Conference on Intelligent Robots and Systems*, pp. 2447–2454, Nov. 2019, doi: 10.1109/IROS40897.2019.8967905.
- [75] Z. Wang, B. Pang, Y. Song, X. Yuan, Q. Xu, and Y. Li, "Robust Visual-Inertial Odometry Based on a Kalman Filter and Factor Graph," *IEEE Transactions on Intelligent Transportation Systems*, Jul. 2023, doi: 10.1109/TITS.2023.3258526.

- [76] O. Omotuyi and M. Kumar, "UAV Visual-Inertial Dynamics (VI-D) Odometry using Unscented Kalman Filter," in *IFAC-PapersOnLine*, Elsevier B.V., Nov. 2021, pp. 814–819. doi: 10.1016/j.ifacol.2021.11.272.
- [77] X. Sang *et al.*, "Invariant Cubature Kalman Filtering-Based Visual-Inertial Odometry for Robot Pose Estimation," *IEEE Sens J*, vol. 22, no. 23, pp. 23413–23422, Dec. 2022, doi: 10.1109/JSEN.2022.3214293.
- [78] J. Xu, H. Yu, and R. Teng, "Visual-inertial odometry using iterated cubature Kalman filter," *Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018*, pp. 3837–3841, Jul. 2018, doi: 10.1109/CCDC.2018.8407789.
- [79] Y. Liu *et al.*, "Stereo Visual-Inertial Odometry with Multiple Kalman Filters Ensemble," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 10, pp. 6205–6216, Oct. 2016, doi: 10.1109/TIE.2016.2573765.
- [80] S. Kim, I. Petrunin, and H. S. Shin, "A Review of Kalman Filter with Artificial Intelligence Techniques," in *Integrated Communications, Navigation and Surveillance Conference, ICNS*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICNS54818.2022.9771520.
- [81] D. J. Jwo, A. Biswal, and I. A. Mir, "Artificial Neural Networks for Navigation Systems: A Review of Recent Research," *Applied Sciences 2023, Vol. 13, Page 4475*, vol. 13, no. 7, p. 4475, Mar. 2023, doi: 10.3390/APP13074475.
- [82] N. Shaukat, A. Ali, M. Moinuddin, and P. Otero, "Underwater Vehicle Localization by Hybridization of Indirect Kalman Filter and Neural Network," in *2021 7th International Conference on Mechatronics and Robotics Engineering, ICMRE 2021*, Institute of Electrical and Electronics Engineers Inc., Feb. 2021, pp. 111–115. doi: 10.1109/ICMRE51691.2021.9384844.
- [83] N. Shaukat, A. Ali, M. J. Iqbal, M. Moinuddin, and P. Otero, "Multi-sensor fusion for underwater vehicle localization by augmentation of rbf neural network and error-state kalman filter," *Sensors (Switzerland)*, vol. 21, no. 4, pp. 1–26, Feb. 2021, doi: 10.3390/s21041149.

- [84] L. Vargas-Meléndez, B. L. Boada, M. J. L. Boada, A. Gauchía, and V. Díaz, "A sensor fusion method based on an integrated neural network and Kalman Filter for vehicle roll angle estimation," *Sensors (Switzerland)*, vol. 16, no. 9, Sep. 2016, doi: 10.3390/s16091400.
- [85] Z. Jingsen, Z. Wenjie, H. Bo, and W. Yali, "Integrating Extreme Learning Machine with Kalman Filter to Bridge GPS Outages," in *Proceedings - 2016 3rd International Conference on Information Science and Control Engineering, ICISCE 2016*, Institute of Electrical and Electronics Engineers Inc., Oct. 2016, pp. 420–424. doi: 10.1109/ICISCE.2016.98.
- [86] X. Zhang, X. Mu, H. Liu, B. He, and T. Yan, "Application of Modified EKF Based on Intelligent Data Fusion in AUV Navigation; Application of Modified EKF Based on Intelligent Data Fusion in AUV Navigation," 2019.
- [87] N. Al Bitar and A. I. Gavrilov, "Neural Networks Aided Unscented Kalman Filter for Integrated INS/GNSS Systems," *27th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2020 - Proceedings*, May 2020, doi: 10.23919/ICINS43215.2020.9133878.
- [88] Z. Miljković, N. Vuković, and M. Mitić, "Neural extended Kalman filter for monocular SLAM in indoor environment," *Proc Inst Mech Eng C J Mech Eng Sci*, vol. 230, no. 5, pp. 856–866, Mar. 2016, doi: 10.1177/0954406215586589.
- [89] K. Y. Kotov, A. S. Maltsev, and M. A. Sobolev, "Recurrent neural network and extended Kalman filter in SLAM problem," in *IFAC Proceedings Volumes (IFAC-PapersOnline)*, IFAC Secretariat, 2013, pp. 23–26. doi: 10.3182/20130902-3-CN-3020.00057.
- [90] M. Choi, R. Sakthivel, and W. K. Chung, "Neural Network-Aided Extended Kalman Filter for SLAM Problem," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 1686–1690.

- [91] L. Chen and J. Fang, "A hybrid prediction method for bridging GPS outages in high-precision POS application," *IEEE Trans Instrum Meas*, vol. 63, no. 6, pp. 1656–1665, 2014, doi: 10.1109/TIM.2013.2292277.
- [92] J. Ki Lee and C. Jekeli, "Neural network aided adaptive filtering and smoothing for an integrated INS/GPS unexploded ordnance geolocation system," *Journal of Navigation*, vol. 63, no. 2, pp. 251–267, Apr. 2010, doi: 10.1017/S0373463309990397.
- [93] S. Bi, L. Ma, T. Shen, Y. Xu, and F. Li, "Neural network assisted Kalman filter for INS/UWB integrated seamless quadrotor localization," *PeerJ Comput Sci*, vol. 7, pp. 1–13, 2021, doi: 10.7717/PEERJ-CS.630.
- [94] C. Jiang *et al.*, "A Novel Method for AI-Assisted INS/GNSS Navigation System Based on CNN-GRU and CKF during GNSS Outage," *Remote Sensing 2022, Vol. 14, Page 4494*, vol. 14, no. 18, p. 4494, Sep. 2022, doi: 10.3390/RS14184494.
- [95] D. Xie *et al.*, "A Robust GNSS/PDR Integration Scheme with GRU-Based Zero-Velocity Detection for Mass-Pedestrians," *Remote Sens (Basel)*, vol. 14, no. 2, Jan. 2022, doi: 10.3390/rs14020300.
- [96] Y. Jiang and X. Nong, "A Radar Filtering Model for Aerial Surveillance Base on Kalman Filter and Neural Network," in *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, IEEE Computer Society, Oct. 2020, pp. 57–60. doi: 10.1109/ICSESS49938.2020.9237716.
- [97] Z. Miao, H. Shi, Y. Zhang, and F. Xu, "Neural network-aided variational Bayesian adaptive cubature Kalman filtering for nonlinear state estimation," *Meas Sci Technol*, vol. 28, no. 10, Sep. 2017, doi: 10.1088/1361-6501/aa7d70.
- [98] S. Hosseinyalamdary, "Deep Kalman filter: Simultaneous multi-sensor integration and modelling; A GNSS/IMU case study," *Sensors (Switzerland)*, vol. 18, no. 5, May 2018, doi: 10.3390/s18051316.

- [99] L. Song, Z. Duan, B. He, and Z. Li, "Application of Federal Kalman Filter with Neural Networks in the Velocity and Attitude Matching of Transfer Alignment," *Complexity*, vol. 2018, 2018, doi: 10.1155/2018/3039061.
- [100] D. Li, Y. Wu, and J. Zhao, "Novel Hybrid Algorithm of Improved CKF and GRU for GPS/INS," *IEEE Access*, vol. 8, pp. 202836–202847, 2020, doi: 10.1109/ACCESS.2020.3035653.
- [101] X. Gao *et al.*, "RL-AKF: An adaptive kalman filter navigation algorithm based on reinforcement learning for ground vehicles," *Remote Sens (Basel)*, vol. 12, no. 11, Jun. 2020, doi: 10.3390/rs12111704.
- [102] C. Chen, C. X. Lu, B. Wang, N. Trigoni, and A. Markham, "DynaNet: Neural Kalman Dynamical Model for Motion Estimation and Prediction," *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 12, pp. 5479–5491, Aug. 2019, doi: 10.1109/TNNLS.2021.3112460.
- [103] S. Zuo, K. Shen, and J. Zuo, "Robust Visual-Inertial Odometry Based on Deep Learning and Extended Kalman Filter," *Proceeding - 2021 China Automation Congress, CAC 2021*, pp. 1173–1178, 2021, doi: 10.1109/CAC53003.2021.9727349.
- [104] Y. Luo, J. Hu, and C. Guo, "Right Invariant SE2(3) - EKF for Relative Navigation in Learning-based Visual Inertial Odometry," *2022 5th International Symposium on Autonomous Systems, ISAS 2022*, 2022, doi: 10.1109/ISAS55863.2022.9757054.
- [105] U. I. Bhatti and W. Y. Ochieng, "Failure Modes and Models for Integrated GPS/INS Systems," *The Journal of Navigation*, vol. 60, no. 2, pp. 327–348, May 2007, doi: 10.1017/S0373463307004237.
- [106] Y. Du, J. Wang, C. Rizos, and A. El-Mowafy, "Vulnerabilities and integrity of precise point positioning for intelligent transport systems: overview and analysis," Dec. 01, 2021, *Springer*. doi: 10.1186/s43020-020-00034-8.
- [107] B. Gao, G. Hu, Y. Zhong, and X. Zhu, "Cubature Kalman Filter with Both Adaptability and Robustness for Tightly-Coupled GNSS/INS Integration,"

- IEEE Sens J*, vol. 21, no. 13, pp. 14997–15011, Jul. 2021, doi: 10.1109/JSEN.2021.3073963.
- [108] B. Gao, S. Gao, Y. Zhong, G. Hu, and C. Gu, “Interacting multiple model estimation-based adaptive robust unscented Kalman filter,” *Int J Control Autom Syst*, vol. 15, no. 5, pp. 2013–2025, Oct. 2017, doi: 10.1007/S12555-016-0589-2/METRICS.
- [109] G. Gao, Y. Zhong, S. Gao, and B. Gao, “Double-Channel Sequential Probability Ratio Test for Failure Detection in Multisensor Integrated Systems,” *IEEE Trans Instrum Meas*, vol. 70, 2021, doi: 10.1109/TIM.2021.3072674.
- [110] G. Gao, B. Gao, S. Gao, G. Hu, and Y. Zhong, “A Hypothesis Test-Constrained Robust Kalman Filter for INS/GNSS Integration With Abnormal Measurement,” *IEEE Trans Veh Technol*, vol. 72, no. 2, pp. 1662–1673, Feb. 2023, doi: 10.1109/TVT.2022.3209091.
- [111] B. Gao, G. Hu, Y. Zhong, and X. Zhu, “Cubature rule-based distributed optimal fusion with identification and prediction of kinematic model error for integrated UAV navigation,” *Aerosp Sci Technol*, vol. 109, p. 106447, Feb. 2021, doi: 10.1016/J.AST.2020.106447.
- [112] B. Gao, G. Hu, X. Zhu, and Y. Zhong, “A Robust Cubature Kalman Filter with Abnormal Observations Identification Using the Mahalanobis Distance Criterion for Vehicular INS/GNSS Integration,” *Sensors 2019, Vol. 19, Page 5149*, vol. 19, no. 23, p. 5149, Nov. 2019, doi: 10.3390/S19235149.
- [113] G. Hu, B. Gao, Y. Zhong, L. Ni, and C. Gu, “Robust Unscented Kalman Filtering with Measurement Error Detection for Tightly Coupled INS/GNSS Integration in Hypersonic Vehicle Navigation,” *IEEE Access*, vol. 7, pp. 151409–151421, 2019, doi: 10.1109/ACCESS.2019.2948317.
- [114] A. Roux, S. Changey, J. P. Lauffenburger, and J. Weber, “Mortar Trajectory Estimation by a Deep Error-State Kalman Filter in a GNSS-Denied Environment,” *2023 IEEE/ION Position, Location and Navigation*

- Symposium, PLANS 2023*, pp. 119–127, 2023, doi: 10.1109/PLANS53410.2023.10140013.
- [115] Z. Li and Y. Zhang, “Constrained ESKF for UAV Positioning in Indoor Corridor Environment Based on IMU and WiFi,” *Sensors 2022, Vol. 22, Page 391*, vol. 22, no. 1, p. 391, Jan. 2022, doi: 10.3390/S22010391.
- [116] C. Ramirez-Atencia and D. Camacho, “Extending QGroundControl for Automated Mission Planning of UAVs,” *Sensors 2018, Vol. 18, Page 2339*, vol. 18, no. 7, p. 2339, Jul. 2018, doi: 10.3390/S18072339.
- [117] G. E. Villalobos Hernandez, I. Petrunin, H.-S. Shin, and J. Gilmour, “Robust multi-sensor navigation in GNSS degraded environments,” Jan. 2023, doi: 10.2514/6.2023-2242.
- [118] B. Wagstaff, E. Wise, and J. Kelly, “A Self-Supervised, Differentiable Kalman Filter for Uncertainty-Aware Visual-Inertial Odometry,” *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, vol. 2022-July, pp. 1388–1395, 2022, doi: 10.1109/AIM52237.2022.9863270.
- [119] D. You, J. Jung, and J. Oh, “Enhancing Low-light Images for Monocular Visual Odometry in Challenging Lighting Conditions,” *Int J Control Autom Syst*, vol. 21, no. 11, pp. 3528–3539, Nov. 2023, doi: 10.1007/S12555-023-0378-7.
- [120] A. Yusefi, A. Durdu, M. F. Aslan, and C. Sungur, “LSTM and filter based comparison analysis for indoor global localization in UAVs,” *IEEE Access*, vol. 9, pp. 10054–10069, 2021, doi: 10.1109/ACCESS.2021.3049896.
- [121] N. Mungoli, “Adaptive Feature Fusion: Enhancing Generalization in Deep Learning Models,” Apr. 2023, Accessed: Nov. 03, 2023. [Online]. Available: <https://arxiv.org/abs/2304.03290v1>
- [122] B. Neyshabur, S. Bhojanapalli, D. McAllester, and N. Srebro, “Exploring Generalization in Deep Learning,” *Adv Neural Inf Process Syst*, vol. 2017-December, pp. 5948–5957, Jun. 2017, Accessed: Nov. 03, 2023. [Online]. Available: <https://arxiv.org/abs/1706.08947v2>

- [123] T. E. Tabassum, Z. Xu, I. Petrunin, and Z. A. Rana, "Integrating GRU with a Kalman Filter to Enhance Visual Inertial Odometry Performance in Complex Environments," *Aerospace 2023, Vol. 10, Page 923*, vol. 10, no. 11, p. 923, Oct. 2023, doi: 10.3390/AEROSPACE10110923.
- [124] M. Villalonga Torres, T. Elmi Tabassum, I. Petrunin, and A. Tsourdos, "Hybrid Multi-Sensor Navigation System With Uncertainty Correction for GNSS-Denied Environments," *AIAA SCITECH 2024 Forum*, Jan. 2024, doi: 10.2514/6.2024-2799.
- [125] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan, and Z. Wang, "A review: state estimation based on hybrid models of Kalman filter and neural network," *Systems Science & Control Engineering*, vol. 11, no. 1, p. 2173682, Dec. 2023, doi: 10.1080/21642583.2023.2173682.
- [126] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. G. Van Sloun, and Y. C. Eldar, "KalmanNet: Neural Network Aided Kalman Filtering for Partially Known Dynamics," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1532–1547, 2022, doi: 10.1109/TSP.2022.3158588.
- [127] J. H. Jung, Y. Choe, and C. G. Park, "Photometric Visual-Inertial Navigation with Uncertainty-Aware Ensembles," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2039–2052, Aug. 2022, doi: 10.1109/TRO.2021.3139964.
- [128] R. Nakashima and A. Seki, "Uncertainty-Based Adaptive Sensor Fusion for Visual-Inertial Odometry under Various Motion Characteristics," *Proc IEEE Int Conf Robot Autom*, pp. 3119–3125, May 2020, doi: 10.1109/ICRA40945.2020.9197397.
- [129] F. Crocetti, M. L. Fravolini, G. Costante, and P. Valigi, "Data-driven and uncertainty-aware robust airstrip surface estimation," *Neural Comput Appl*, vol. 35, no. 26, pp. 19565–19580, Sep. 2023, doi: 10.1007/S00521-023-08779-4/FIGURES/10.

- [130] Y. Liu, C. Zhao, and M. Ren, "An Enhanced Hybrid Visual-Inertial Odometry System for Indoor Mobile Robot," *Sensors (Basel)*, vol. 22, no. 8, Apr. 2022, doi: 10.3390/S22082930.
- [131] T. E. Tabassum, I. Petrunin, and Z. A. Rana, "Position Uncertainty Reduction in Visual-Inertial Navigation Systems Using Multi-ML Error Compensation," pp. 1741–1755, Sep. 2024, doi: 10.33012/2024.19790.
- [132] H. Wang, H. Deng, and H. Zhang, "Lightweight image denoising method for feature-based visual SLAM in γ radiation environments," *Proc Inst Mech Eng G J Aerosp Eng*, Oct. 2024, doi: 10.1177/09544100241293001/ASSET/IMAGES/LARGE/10.1177_09544100241293001-FIG10.JPEG.
- [133] L. Cao, J. Ling, and X. Xiao, "Study on the Influence of Image Noise on Monocular Feature-Based Visual SLAM Based on FFDNet," *Sensors* 2020, Vol. 20, Page 4922, vol. 20, no. 17, p. 4922, Aug. 2020, doi: 10.3390/S20174922.
- [134] M. Abboush, C. Knieke, and A. Rausch, "GRU-Based Denoising Autoencoder for Detection and Clustering of Unknown Single and Concurrent Faults during System Integration Testing of Automotive Software Systems," *Sensors* 2023, Vol. 23, Page 6606, vol. 23, no. 14, p. 6606, Jul. 2023, doi: 10.3390/S23146606.
- [135] Z. Song, X. Zhang, T. Li, S. Zhang, Y. Wang, and J. Yuan, "IR-VIO: Illumination-Robust Visual-Inertial Odometry Based on Adaptive Weighting Algorithm With Two-Layer Confidence Maximization," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 4, pp. 1920–1929, Aug. 2023, doi: 10.1109/TMECH.2023.3277031.
- [136] P. Geragersian, I. Petrunin, W. Guo, and R. Grech, "A Hybrid Deep Learning Approach for Robust Multi-Sensor GNSS/INS/VO Fusion in Urban Canyons," *Proceedings of the 36th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2023*, pp. 2624–2643, Sep. 2023, doi: 10.33012/2023.19271.

- [137] J. Lee, M. Kim, D. Min, S. Pullen, and J. Lee, "Navigation Safety Assurance of a KF-Based GNSS/IMU System: Protection Levels Against IMU Failure," *NAVIGATION: Journal of the Institute of Navigation*, vol. 70, no. 4, Dec. 2023, doi: 10.33012/NAVI.612.
- [138] E. Shytermeja, A. Garcia-Pena, and O. Julien, "Proposed architecture for integrity monitoring of a GNSS/MEMS system with a Fisheye camera in urban environment," *Proceedings of 2014 International Conference on Localization and GNSS, ICL-GNSS 2014*, Oct. 2014, doi: 10.1109/ICL-GNSS.2014.6934179.
- [139] J. Zhao and D. Song, "Integrity monitoring method for GNSS/IMU integrated navigation system of UAV," *Acta Aeronautica et Astronautica Sinica*, vol. 45, no. 7, Apr. 2024, doi: 10.7527/S1000-6893.2023.28943.
- [140] C. Yang, A. Vadlamani, A. Soloviev, M. Veth, and C. Taylor, "Feature matching error analysis and modeling for consistent estimation in vision-aided navigation," *Navigation*, vol. 65, no. 4, pp. 609–628, Dec. 2018, doi: 10.1002/NAVI.265.
- [141] M. Abdar *et al.*, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Information Fusion*, vol. 76, pp. 243–297, Dec. 2021, doi: 10.1016/j.inffus.2021.05.008.
- [142] S. W. Cheng and Y. H. Huang, "A Computationally Efficient GNSS/INS Design of Multirotor based on Error-state Kalman Filter," *2023 62nd Annual Conference of the Society of Instrument and Control Engineers, SICE 2023*, pp. 907–913, 2023, doi: 10.23919/SICE59929.2023.10354209.
- [143] S. Ochella, F. Dinmohammadi, and M. Shafiee, "Bayesian neural networks for uncertainty quantification in remaining useful life prediction of systems with sensor monitoring," *Advances in Mechanical Engineering*, vol. 16, no. 7, Jul. 2024, doi: 10.1177/16878132241239802/ASSET/IMAGES/LARGE/10.1177_16878132241239802-FIG6.JPEG.

- [144] U. Robustelli, M. Cutugno, and G. Pugliano, "Low-Cost GNSS and PPP-RTK: Investigating the Capabilities of the u-blox ZED-F9P Module," *Sensors* 2023, Vol. 23, Page 6074, vol. 23, no. 13, p. 6074, Jul. 2023, doi: 10.3390/S23136074.
- [145] S. A. Negru, P. Geragersian, I. Petrunin, and W. Guo, "Resilient Multi-Sensor UAV Navigation with a Hybrid Federated Fusion Architecture," *Sensors* 2024, Vol. 24, Page 981, vol. 24, no. 3, p. 981, Feb. 2024, doi: 10.3390/S24030981.