

Position Uncertainty Reduction in Visual-Inertial Navigation Systems Using Multi-ML Error Compensation

Tarafder Elmi Tabassum, Ivan Petrunin, Zeeshan A. Rana, Cranfield University

BIOGRAPHIES

Tarafder Elmi Tabassum is a researcher at Cranfield University focusing on computer vision and sensor fusion. She is currently working as a part-time research assistant at Cranfield University where she is involved in developing sensor fusion algorithms for e-scooter localisation. She received a master's degree from The University of Manchester. Her PhD research at Cranfield involves alternative navigation solutions using visual sensors for UAV navigation. Her work includes integrating Kalman filters (KF) and machine learning (ML) algorithms, aiming to enhance robustness to ensure safe navigation under complex environments in the absence of global positioning satellite systems (GNSS).

Dr. Ivan Petrunin received an M.Sc. degree in design and manufacturing of electronic equipment from the National Technical University of Ukraine, in 1998, and a Ph.D. degree in signal processing for condition monitoring from Cranfield University, U.K., in 2013. He is currently a Reader in digital signal processing with the Centre for Autonomous and Cyber-Physical Systems, Cranfield University. His current research interests include applied signal processing for autonomous systems with particular emphasis on Position, Navigation and Timing (PNT), sensor technologies, perception, data and information fusion, and decision-making for Cyber-Physical Systems with support from AI-based techniques.

Dr. Zeeshan A. Rana has a background in Mechanical engineering with several years of industrial experience in various industries. He received his PhD in Computational Fluid Dynamics and Aerodynamics investigating turbulent mixing in super/hypersonic flows within the framework of implicit large eddy simulations (ILES) from Cranfield University. He has research interest in the area of Computational Aerodynamics. He is working on various projects related to Computational Fluid Dynamics of aerospace vehicles, high-performance motorsports, external and internal aerodynamics performance, and optimisation at low, transonic and supersonic Mach numbers. He has a keen interest in complex hypersonic flows and applications of Artificial Intelligence in Computational Aerodynamics.

ABSTRACT

In the absence of signals from global navigation satellite systems (GNSS), visual-inertial navigation systems (VINS) are usually utilized in urban air mobility (UAM) applications which require reliable navigation in complex urban areas. This paper focuses on improving vision-based alternatives to GNSS navigation solutions with position uncertainty correction for safe and uninterrupted flights. The novel contribution introduces multiple machine-learning aided hybrid visual-inertial odometry (multi-ML hybrid VIO) that utilizes gated recurrent unit (GRU) based error compensators to enhance positioning within complex environments by reducing the impacts of various sources of uncertainty. Unlike state-of-the-art systems that lack evidence of demonstrating performance enhancement with position uncertainty correction within VIO architectures, the proposed framework simultaneously reduces position uncertainty and improves accuracy. Furthermore, training and testing datasets are generated using MATLAB incorporating unreal engine simulation environment for UAVs to replicate complex scenarios including environmental conditions, illumination variations, weather effects and flight dynamics where traditional VIO systems tend to fail. The proposed hybrid VIO architecture has been validated under combinations of complex scenarios including various sources of uncertainty such as sensor noise, feature tracking error, environmental dynamics, weather effects and lighting conditions for extended flights. The comparison results have demonstrated reduction in horizontal positioning RMSE errors: 1.7m for VIO with VO error compensation 2.18m for VIO with KF error compensation, 1.4m for multi-ML hybrid VIO. Furthermore, it demonstrates generalization ability over seen and unseen fault scenarios that indicates performance improvement of 89% in 3D position compared to VIO with VO error compensation, VIO with KF error compensation, and ESKF-based VIO. Additionally, experimental results demonstrate overall horizontal position uncertainty reduction by 79% for test

1 and 22% for test 2. Finally, this work represents a step forward in improving the safety and effectiveness of UAV navigation by providing vision-based alternative to GNSS solution for uninterrupted flights.

1. INTRODUCTION

Urban air mobility (UAM) represents a revolutionary approach to transportation within urban environments, leveraging the vertical dimension to reduce ground traffic congestion and enhance mobility. Central to the success of UAM in the deployment of unmanned ariel vehicles (UAVs) which requires sophisticated navigation systems to operate safely and efficiently in densely populated areas. Failure to correctly perform the navigation task could potentially lead to catastrophic accidents such as crashes and collisions. Given the limitations of GNSS-based navigation systems, current position, navigation, and timing (PNT) solutions are increasingly providing independent of GNSS techniques.

The vision-based navigation system for UAVs is widely used as an alternative to GNSS solutions while navigating in urban space. State-of-the-art visual-inertial navigation systems frequently experience performance degradation due to various environmental fault factors including illumination variation, motion variation, low texture, and limited field-of-view [1]. These complexities hinder accurate system modelling because of the inherent non-linearity, creating a need for more advanced techniques to enhance system performance. Additionally, these fault factors impact the overall system performance by introducing uncertainty in data association. For instance, matching visual features across frames under lighting variation, or featureless environments introduces ambiguity and uncertainty. Sources like sensor noise, feature tracking errors, data association errors, and feature location errors contribute to aleatoric uncertainty. On the other hand, calibration errors, imperfect models, and model assumptions introduce epistemic uncertainty. Unlike epistemic uncertainty, aleatoric uncertainty cannot be fully eliminated but can be statistically characterised by proposing advanced fusion algorithms.

In classical filter-based VIO navigation systems, VIO feature trackers are frequently found to be brittle under challenging scenarios and rapid motion, particularly due to the effects of motion blur. Researchers working on classical methods typically focus on modelling feature tracking errors and aim to improve positioning performance [2]. However, the utilization of such error models is limited to specific scenarios and lacks evidence of effectively mitigating visual errors to ensure safe navigation under complex urban environments [3]. One of the drawbacks of filter-based VIO fusion using the variants of Kalman filters (KF) is that the linear assumption model and gaussian noise may not hold under dynamic conditions leading to divergence in the state estimation [4], [5], [6].

Advancement in VIO systems using data-driven approaches has been explored due to their ability to learn model complexity that can establish the relationship between sensor measurements and vehicle ego-motion. Despite their success, data-driven VIO methods lack robustness when facing challenges such as UAV operation in low-texture environments, non-ideal image conditions (e.g. blur, distortions, extreme lighting, overexposures) or a significant presence of dynamic objects in the scene (Almalioglu et al., 2022; Li & Waslander, 2020). Additionally, data-driven approaches are influenced by several factors such as, the need for large training datasets for CNN architectures, performance degradation in scenarios that deviate from training data, and inefficient uncertainty estimation associated with position estimation. Achieving accurate positioning with confidence while reducing uncertainty in such architectures remains an open challenge.

Some studies proposed data-driven VIO algorithms with uncertainty awareness yet lack evidence on validating their approach under visually degraded scenarios (e.g. texture-less environments, rapid motion, low-light) resulting in loss of feature tracking or divergence after 100 frames iterations [9], [10], [11]. A recent comparative analysis of techniques including CNN-LSTM-based VO, [12] claims robustness in predicting accurate ego-motion under rainy weather conditions yet face challenges in replicating more realistic fault conditions e.g. raindrops, and reflection. These conditions introduce fault factors like feature detection inability, loss of feature track, inconsistent tracking and resultant frequent failures or divergence in tracking. Other approaches such as Bayesian techniques [10], [13] and RCNN with mixture density network [11], also extract predictive position but lack evidence under the possible source of visual uncertainty such as lighting variation, motion variation, no-texture environments, and repetitive fault conditions. Thus, the predictive accurate position while reducing uncertainty for filter-based VIO and data-driven VIO systems under such scenarios remains challenging due to the performance degradation from unaddressed uncertainty sources.

The KF-inspired NN-aided architectures [14], [15] have gained attention to enhance performance under complex visual scenarios by adapting Neural Networks (NN) models (e.g. Bi-LSTM, LSTM) to mitigate sensor noise, model imperfections, corrupted measurements, and model assumptions. For instance, during limited visibility conditions e.g. foggy weather error compensation of KF measurements used in [15] suffers from performance degradation as inherent nonlinearity affects the model assumptions. Besides, authors have explored hybrid architecture [16] for reducing model uncertainty by predicting K-gain using gated recurrent unit (GRU) to correct the state estimation and covariance model of EKF. However, these approaches are constrained to linear systems, limiting their effectiveness in more complex scenarios. Therefore, overall performance relies on both model's capabilities and the characteristics of the navigation environment. Attention should be given to mitigate data and measurement errors. Hypothesis analysis suggests combining two error compensation approaches within hybrid architectures to mitigate sensor and model measurement errors in the VIO system, aiming to improve overall performance under diverse scenarios. In this study, error state Kalman filter (ESKF) is employed within the hybrid architecture for its ability to linearize only the error states that improves 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 relies on sigma points to approximate distribution, potentially introducing instability. As a result, ESKF is considered more reliable choice for handling nonlinear dynamics.

Therefore, to address the limitations of state-of-the-art systems providing performance enhancement while reducing uncertainty under complex scenarios, this paper proposes a novel approach that combines multiple error compensators within VIO systems. To the best of the authors' knowledge, this is the first attempt to combine multiple ML-based error correction techniques within the VINS system. The proposed hybrid fusion approach integrates a combination of error compensation techniques: one GRU compensate for VO measurement errors by updating the ESKF filter, while another GRU compensate for model measurement errors to correct model inaccuracies. In this research, GRU is utilized 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 on the proposed approach using MATLAB incorporated with unreal engine simulated photorealistic datasets replicating possible uncertainty sources (e.g. feature tracking errors, data association errors, sensor noise, feature location errors) present in urban environments for extended flights. The contributions of this work are highlighted as follows:

- 1) A combination of GRU-based multiple error compensation and ESKF for the hybrid VIO framework forming error compensation by multi-ML hybrid VIO which was not proposed in other literature known to the authors. The algorithm aims to reduce position uncertainty by leveraging sensor error compensator and measurement model error compensator that simultaneously correct model and data uncertainty under complex urban environments.
- 2) Systematic evaluation of the proposed uncertainty reduction mechanism under the possible source of uncertainty such as feature tracking errors, feature location errors, data association errors and sensor noise.
- 3) 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.

2. PROPOSED NOVEL HYBRID VIO FRAMEWORK

This section presents a high-level overview of the proposed novel framework of multi-ML hybrid VIO. In this architecture, monocular camera-generated images are used to estimate VO position using feature-based VO algorithm [2]. Using the GRU-based error compensation method, VO errors arising from possible sources of data uncertainty within the navigation environment are corrected by predicted increments before fusion. Corrected VO positions are then fed into the ESKF filter with IMU-estimated raw linear acceleration and angular velocity for fusion to estimate the position of VIO. Lastly, another GRU is utilized to compensate for ESKF measurement errors arising due to the uncertainty present in the measurement model. The predicted measurement position increment will improve hybrid VIO positioning performance by reducing model and data uncertainty simultaneously. Figure 1 presents a high-level diagram of the proposed novel hybrid VIO framework. This paper investigates and explores the performance of the proposed novel hybrid VIO framework by utilizing the position uncertainty compensation mechanism. However, estimating confidence interval to evaluate integrity for mentioned uncertainty sources is out of the scope of the paper. This study focuses on reducing the effects of various sources of uncertainty simultaneously improving performance while navigating under complex environment.

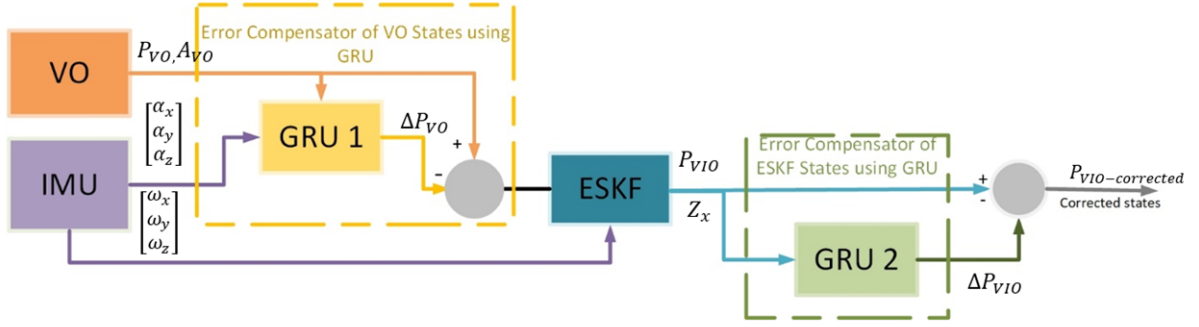


Figure 1: Proposed combined hybrid architecture for VIO position uncertainty reduction.

2.1 ML Module Input and Output Parameters

It is crucial to select input and output parameters for all the ML modules since these parameters directly determine the training efficiency and navigation accuracy. In this work, the selected two ML-based error compensators have different input and output parameters based on their usage within the architecture. The first kind of model is applied to compensate for sensor errors associated with observation noise within data due to several fault factors present in navigation environments. Since visual errors are prominent within VINS systems, VO sensor error compensation mechanisms are conducted by predicting increments of VO using GRU. 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 of dependencies between the estimated position and the actual position, the GRU can learn typical error patterns. Finally, the predicted increments are compensated using the following equation,

$$P_{VO/GRU} = P_{VO} - \Delta P_{VO} \quad (1)$$

Another kind of ML model is applied to compensate for measurement model errors associated with the imperfection of the model due to manual tuning and nonlinearity presence within data. 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-

$$P_{VIO/GRU} = P_{VIO} - \Delta P_{VIO} \quad (2)$$

2.2 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 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 movements. The structure and operation of cells in GRU are 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 hidden layers unit in the network introduces a gaining unit and its structure is adopted from the literature (Geragersian et al., 2022). 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.

2.3 Multi-ML-aided Hybrid VIO

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

$$x = [p \ v \ q \ \alpha_b \ \omega_b]^T \quad (3)$$

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 .

$$x_k = x + \delta x \quad (4)$$

Notably, the frame of references 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.

$$\begin{aligned} \delta \hat{x} &= F_x(x, u_m) \cdot \delta x \\ P &= F_x P_{k-1} F_x^T + F_i Q_i F_i^T \end{aligned} \quad (5)$$

The transition Matrix F_x for the error state as follows-

$$F_x = \left. \frac{\partial f}{\partial \delta x} \right|_x = \begin{bmatrix} I & I\Delta t & 0 & 0 & 0 \\ 0 & I & -C_{ni}[a_m - a_{b_k}]_x \Delta t & -C_{ni}\Delta t & 0 \\ 0 & 0 & C_{ni}^T\{((\omega_m - \omega_{b_k})\Delta t)\} & 0 & -I\Delta t \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix} \quad (6)$$

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 .

Featured-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. Once the GRU 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

$$\begin{aligned} K &= \hat{P}_{k-1} H^T (H \hat{P}_{k-1} H^T + R)^{-1} \\ \delta \hat{x} &= K(y_k - h(\hat{x}_t)) \\ \hat{P} &= (I - KH) \hat{P}_{k-1} \end{aligned} \quad (7)$$

Finally, ESKF measurement position is corrected using equation (2) and estimate improved multi-ML hybrid VIO position.

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 proposes 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 our previous works which was used to implement 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 specification of ICM 20649 model operating 100 Hz sample rate used in our 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 are simulated using ‘US city’ scene available in the ‘simulation 3D scene configuration’ block within Navigation Toolbox and UAV toolbox in MATLAB. This paper investigated only navigation environment of a bright sunny day with 40% floating clouds. However, training datasets are collected by replicating weather effects and lighting conditions. The setting is adopted based on MathWorks user interface described on the website. The flight trajectories are designed and controlled through QgroundControl software connected through MAVLink connectivity.

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 fog, rain, sunny, cloud 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 2.

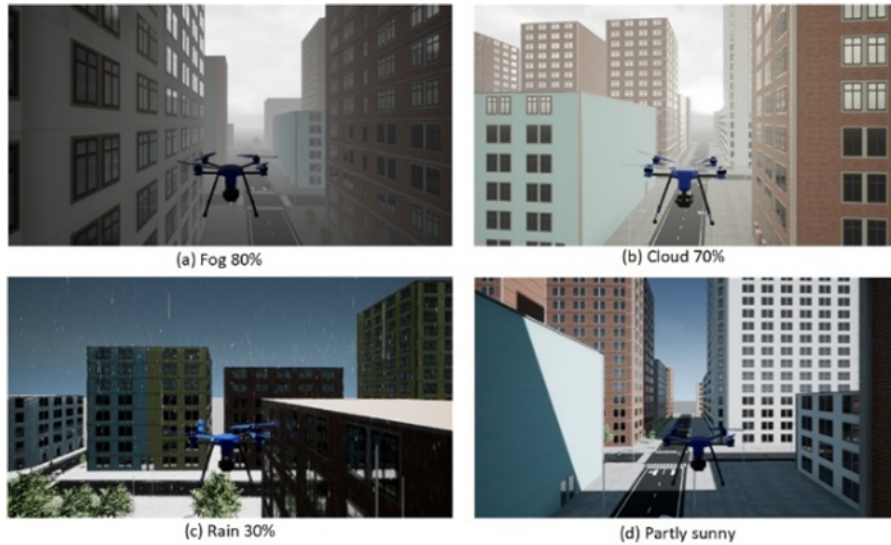


Figure 2: Examples of simulated environmental conditions: (a) limited field of view during 80% fog, (b) texture-less wall during 70% cloud, (c) illumination changes during 30% rain and (d) shadows and texture-less wall during partly sunny used in testing and training datasets.

3.2 Training and Testing Process

The GRU modules are trained with the mentioned 12 trajectories using the input and output data structure described in section 2.1. Notably, this research has utilized similar GRU architectures for two different purposes within the hybrid VIO architecture. To balance the training efficiency and prediction accuracy, the GRU model comprised two GRU layer containing 256 and 36 cells respectively with the sequence length input. A fully connected layer is used to provide the output from the ML model. The batch size used is 128, the dropout rate is set at 0.5, the activation function is ‘tanh’ and recurrent activation function is ‘sigmoid’. The ADAMAX optimizer is employed during the training to optimize the ML model weights and minimize the discrepancies between the predicted and actual errors. Furthermore, the loss function is calculated using mean squared error (MSE).

During the testing phase, ML models operate in the prediction phase, remaining active throughout the experiment. During this phase, the train GRU-based 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 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 predicted position error from KF estimated position to improve overall hybrid VIO positioning performance.

4. PERFORMANCE EVALUATION

To evaluate the performance and adaptability of the proposed novel hybrid VIO framework, two experiments are carried out using two different testing datasets with a variation of fault conditions, shown on Figure 3.

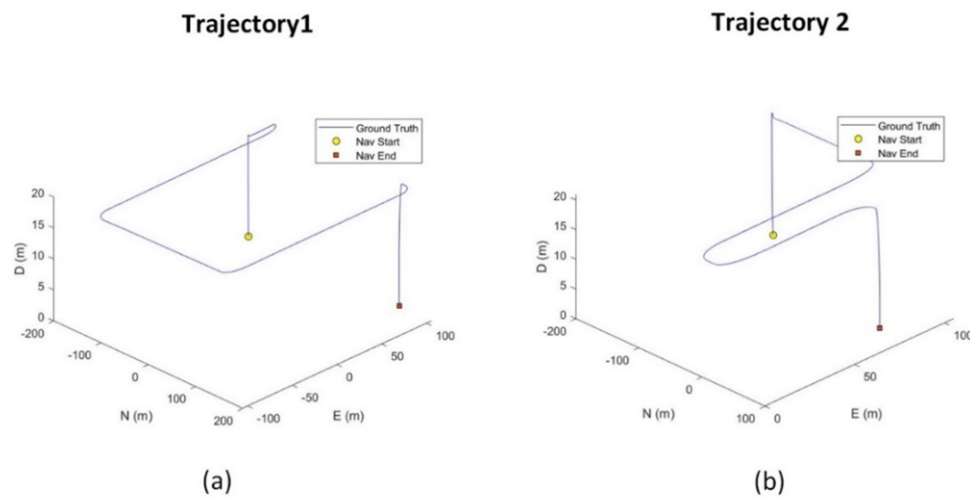


Figure 3: 3D view of test trajectories used for experiments (a) trajectory 1 for test 1, (b) trajectory 2 for test 2.

These trajectories are generated using MATLAB incorporated unreal engine simulator with a ‘US city’ scenario which replicate urban environments with various fault condition including illumination variation, motion variation, altitude variation, less-texture, and flight dynamics. These fault conditions contribute to various sources of uncertainty that are aimed to be reduced by employing the proposed novel multi-ML hybrid VIO framework and evaluate the model adaptability under such scenarios. Comparative analysis is conducted to evaluate the efficiency of proposed multi-ML hybrid VIO benchmarking with four references including three VIO architectures such as GRU-based VO error compensator aided ESKF-hybrid VIO (VIO with VO error compensation), GRU-based KF error compensator aided ESKF-hybrid VIO (VIO with KF error compensation), ESKF-based VIO and feature-based VO (VO).

4.1 Test 1

The purpose of this experiment is to evaluate performance and resilience abilities of the proposed algorithm in managing combined fault conditions and mitigation of arisen source of uncertainty associated with trajectory 1. During the experiment distance of 701m is covered over a duration of 5 min. This experiment is conducted under sunny weather with combined fault conditions including turn with feature-less walls, turn with sky-view, shadow of objects, illumination variation, take-off with altitude variation.

Figure 4 illustrates absolute position error comparison based on the proposed multi-ML hybrid VIO, VIO with VO error compensation, VIO with KF error compensation, ESKF-based VIO and VO under sunny day. Throughout the flight, shadows on surface and shadows on buildings are the common constant fault conditions that has been observed to be mitigated by the proposed algorithm compared to other reference systems. During landing phrase after 2200s, the errors in N, E, D position are observed higher. It is worth noting that the performance degradation arises due to combined fault conditions including UAV turn with sky view and shadows of buildings. These condition lead to illumination variation and less-texture fault factors, resulting in feature extracting error, feature location error and feature tracking error. A source of aleatoric or data uncertainty, such as

feature extraction errors due to limited features available when UAV is turning or landing with a sky-facing view, can result in performance degradation in the D coordinate. This issue is mitigated with a reduction of 96% in the multi-ML hybrid VIO, 94.2% in VIO with VO error compensation, and 93.8% in VIO with KF error compensation. The position error highlights how erroneous data generated by various fault conditions presence in the navigation environments effects performance of VIO architectures. This degradation cannot be mitigated by compensating either sensor or model measurements alone due to multiple sources of uncertainty impacting both. For instance, the VIO with KF error compensation method showed less improvement compared to VIO with the VO error compensator. This suggests that if data sources are corrupted, it negatively impact model performance. Additionally, without sufficient knowledge about the unmodeled dynamics in KF, the KF error compensation ML model is less effective learning the fault patterns, resulting in poor performance compared to VO error compensation technique. Thus, integrating multiple error compensators to simultaneously correct sensor and model measurements reduces the position uncertainty and improves performance to ensure safe navigation.

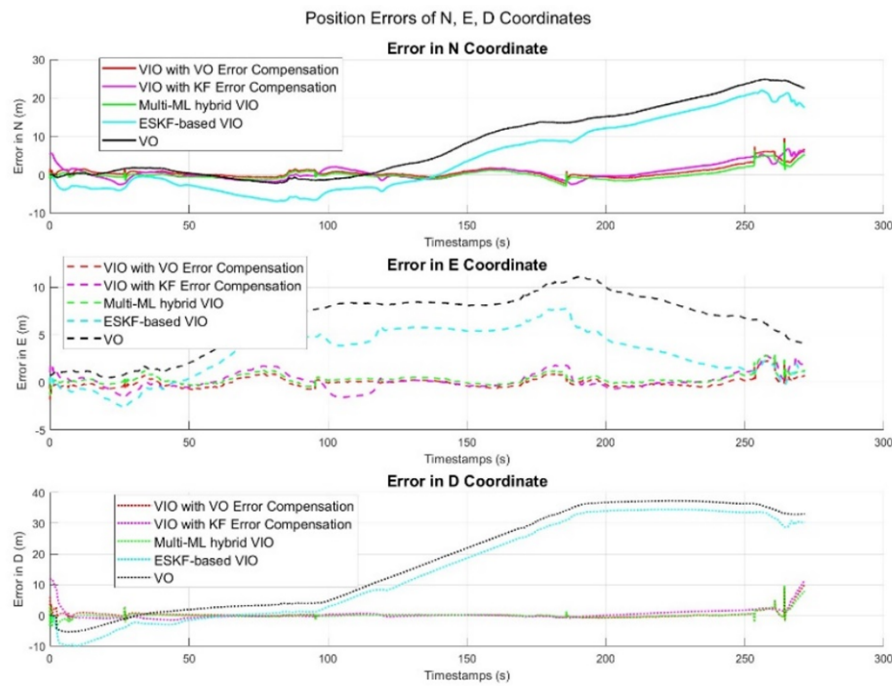


Figure 4: Position error in presence of combined fault conditions for trajectory 1

Figure 5 illustrates the probability distribution of horizontal errors to provide insight into the ability of improving performance while reducing the effects of possible source of uncertainty by the proposed hybrid framework. The distributions VO and ESKF-based VIO reference systems showed a very broad error distribution, indicating significant position uncertainty. However, the hybrid architectures including VIO with VO error compensation, VIO with KF error compenstaion and Multi-ML hybrid VIO showed significant reduction in horizontal position uncertainty with most error concentrated below 8 m. The sharp peak in error distribution of multi-ML hybrid VIO implies that most errors within smaller range compared to other two methods. Thus, the proposed hybrid algorithm is the most effective at reducing position uncertainty by mitigating the effects of data and model uncertainty simultaneously.

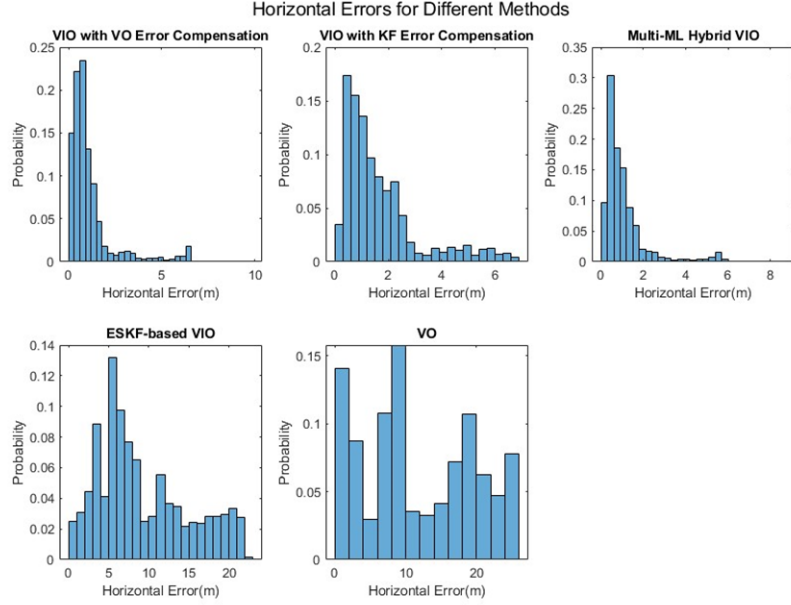


Figure 5: Histogram of horizontal position error for trajectory 1.

Table 1 summarises the performance of the proposed framework compared to the reference algorithms concerning RMSE and 95th percentile for trajectory 1. It is observed that multi-ML hybrid VIO outperforms compared to the fusion reference systems. Horizontal RMSE is improved by 90% for multi-ML hybrid VIO, 88% for VIO with VO error compensation, 85% for VIO with KF error compensation, 23% for ESKF-based VIO. Additionally, the VIO with KF error compensation algorithm demonstrates less effectiveness than the VIO with VO error compensation methods due to inadequate information of unmodeled dynamics created by corrupted VO data. However, incorporating VO error compensation and KF error compensation mechanisms within hybrid VIO architecture effectively improves performance and reduces Horizontal position error by 87% (95th percentile). The system demonstrates the capability of correcting erroneous measurements e.g. corrupted sensor data, imperfect models, and enhancing performance, reducing D coordinate position error (RMSE) to 1.0 m.

TABLE I
POSITIONING PERFORMANCE FOR TRAJECTORY 1 FOR TEST 1.

Techniques	RMSE (m)					95 th Percentile (m)			
	N	E	D	3D	Horizontal	3D	Improvement	Horizontal	Improvement
VO	12.10	7.02	23.69	27.56	14.07	43.47	-	24.80	-
ESKF-based VIO	8.02	2.26	11.72	24.19	10.8	39.02	10%	20.45	17%
VIO with VO error compensation	0.95	0.78	1.22	2.00	1.75	4.71	89%	4.16	83%
VIO with KF error compensation	0.99	0.91	1.28	2.13	2.18	5.85	86%	5.04	79%
Multi-ML Hybrid VIO (proposed)	1.1	0.50	1.0	1.67	1.40	3.75	91%	3.18	87%

Figure 6 illustrates the behaviour of 3D trajectories obtained from the proposed framework with reference VIO architectures. These trajectories are observed under a combination of aleatoric and epistemic uncertainty. Importantly, the entire trajectory is designed with known and observed fault conditions, identical to those used in the training datasets, and occurring at the same location. Notably, our proposed multi-ML hybrid VIO follows closely to the ground trajectory with only minor drifting during landing, resulting 91% improvement in overall positioning compared to VO. Two noticeable deviations in the trajectory caused by feature tracking errors and feature extraction error in less-texture environments while landing due to high mismatched characterization of

underlying dynamics. Despite these, the system effectively overcomes these errors, allowing the trajectory to eventually align with the ground truth.

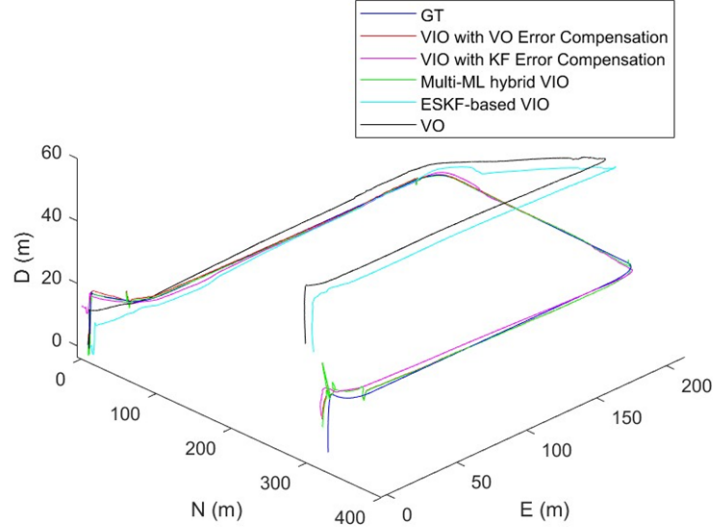


Figure 6: 3D trajectory comparison of the proposed novel multi-ML hybrid VIO framework with four reference systems.

4.2 Test 2

The objective of this experiment is to evaluate generalization 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 includes diverse scenarios, including UAV takes off from different street in the city environments with variation of unknown fault conditions i.e. diagonal rotation during landing and altitude variation while takes-off. The experiment is conducted under sunny weather condition covering distance of 738 m within 4 min.

Figure 7 and Figure 8 illustrate the comparison of position errors and histogram of horizontal errors using proposed multi-ML hybrid VIO, VIO with VO error compensation, VIO with KF error compensation, ESKF-based VIO and VO for trajectory 2 under sunny weather conditions. Performance deviations are noticeable during landing phrases due to the navigation environment variability and presence of unknow fault conditions. However, the hybrid architectures effectively take off from unseen location, showcasing its adaptability.

Notably, performance deviation in the N position is observed throughout the flight which is caused by prominent fault conditions such altitude variation and texture-less environment during flight that led to feature tracking errors. The VIO with KF error compensation improved the performance compared to traditional ESKF-based VIO and VO. Since ESKF-based VIO experienced increase in error due to the inability of traditional Kalman filter to fully compensate unmodeled dynamics where uncertainty and noise degrade performance, reflecting in the wider error distribution of Figure 8. In contrast, VIO with VO error compensation mechanism performs well by compensating VO errors before fusion that allows correction for compounded errors in sensor data. Thus, in turn, helps the KF measurement model avoid dealing with high dynamics present in the sensor data, resulting in more accurate performance that skewed the horizontal error distribution. Additionally, the hybrid approaches have demonstrated error reduction by concentrating peak of the horizontal error distribution up to 5m which provide an insight into the adaptability of the hybrid models to unseen location and faults.

Nevertheless, there are higher degradation after 200 s in N & E direction due to diagonally landing among all the reference systems including proposed framework. These erroneous behaviours arise from the uncertainty sources such as corrupted sensor measurements, feature tracking errors, which lead to mismatched characterization of underlying dynamics of sensor measurements and predictions. Since, VIO with KF error compensation depends entirely on the ESKF output, the error distribution tail leads up to 15m with most of the errors skewed to 4m. The integrated multi-ML hybrid VIO framework demonstrated improvement of 71% and

47% towards N & E coordinates compared to sensor source VO system. It is observed in Figure 8 that the error distribution of the multi-ML hybrid VIO algorithm shows most significant reduction in horizontal position uncertainty, with most errors concentrated suggesting that it is the most effective at minimizing errors compared to other hybrid approach. Additionally, source of aleatoric or data uncertainty e.g. feature tracking error due to diagonally rotation and altitude variation leading to performance degradation during the landing phase has been mitigated with a reduction of 95% or multi-ML hybrid VIO, 93% for VIO with VO error compensation, 90% for VIO with KF error compensation in D position. Finally, it should be mentioned that throughout the flight shadows on the surface and shadows on buildings are the common constant fault conditions which has been effectively mitigated by proposed algorithm. Thus, integrating multiple error compensators to simultaneously correct sensor and model measurements reduces the position uncertainty and improves adaptable ability to ‘out of distribution’ problem, ensuring safe navigation.

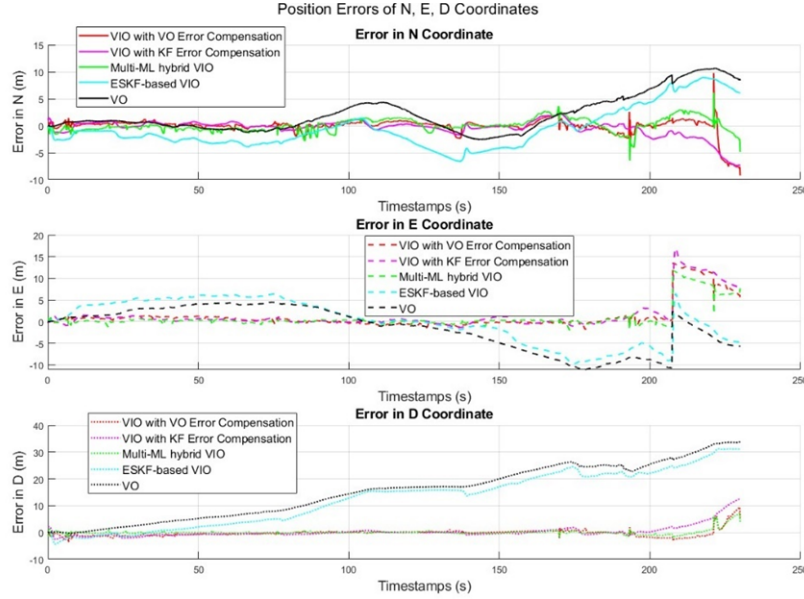


Figure 7: Position error for ‘out of the distribution’ experiment using trajectory 2

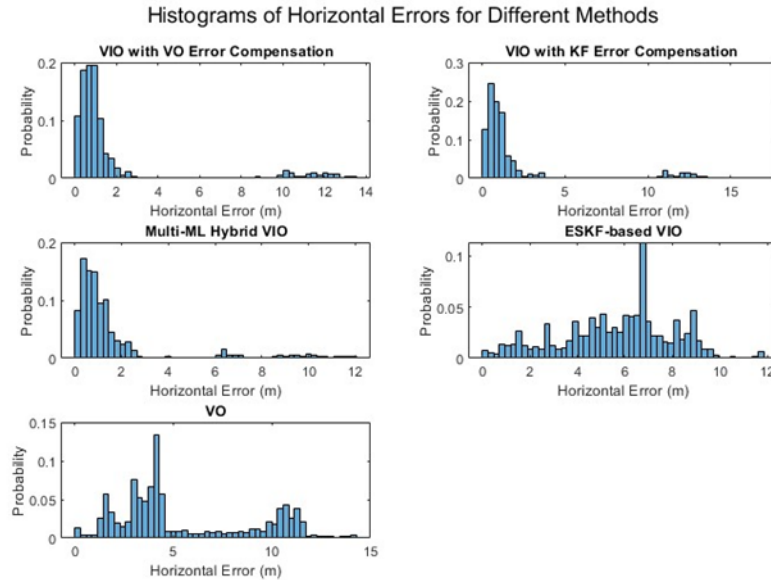


Figure 8: Histogram of horizontal position error for trajectory 2

Table 2 summarises the performance of the proposed framework compared to the reference algorithms concerning RMSE and 95th percentile for trajectory 2. It is observed that multi-ML hybrid VIO outperforms compared to the fusion reference systems. Horizontal RMSE is improved by 89% for multi-ML hybrid VIO, 84%

for VIO with VO error compensation, 85% for VIO with KF error compensation, 13% for ESKF-based VIO. As mentioned previously, test 2 trajectory is designed with high flight dynamics and combined seen and unseen scenarios compared to test 1. The erroneous data sources are coming from sensors which have impact on the model predicting capability. Due to outliers presents in VO arising from feature extraction and tracking incapability during motion variance, altitude variation, rotation and inefficient features, ESKF model suffered to perform linearization under such complex dynamics which can be observed from 95th percentile estimation. However, our proposed sensor and model compensation mechanism successfully mitigate the inherent nonlinearity (position uncertainty) introduced in the ESKF-based VIO fusion model with improvement of 22% in horizontal axis (95th percentile). The hybrid models have demonstrated adaptability in predicting the unknown patterns of faults arising from different sources of uncertainty and showed improvement of 2.81m in horizontal axis. Since, significant degradation observed in ESKF-based VIO, the hybrid approach VIO with KF error compensator has provided limited improvement. Finally, applying our multi-ML hybrid VIO is able to correct such erroneous measurements and reduce overall uncertainty.

TABLE II
POSITIONING PERFORMANCE FOR TRAJECTORY 2 FOR TEST 2.

Techniques	RMSE (m)					95 th Percentile (m)			
	N	E	V	3D	Horizontal	3D	Improvement	Horizontal	Improvement
VO	4.04	5.13	18.11	27.56	6.53	33.20	-	11.36	-
ESKF-based VIO	3.7	4.83	16.00	24.19	6.15	29.63	40%	9.01	21%
VIO with VO error compensation	1.37	3.43	1.25	3.9	3.70	11.85	64%	11.32	0.4%
VIO with KF error compensation	1.76	3.46	1.96	4.31	4.16	11.34	66%	12.12	-
Multi-ML Hybrid VIO (proposed)	1.19	2.74	1.01	3.13	2.81	9.33	72%	8.88	22%

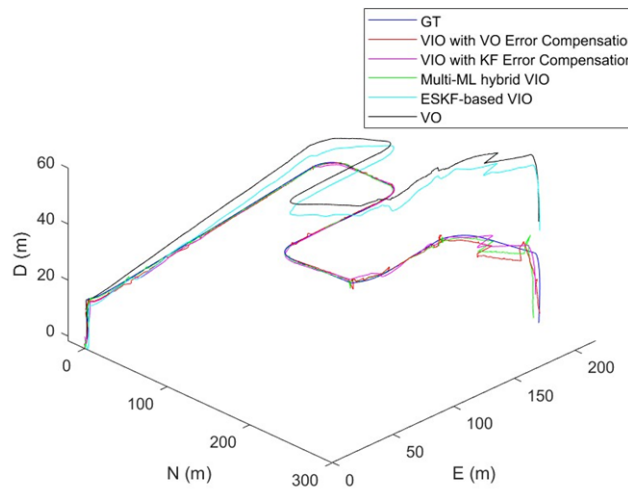


Figure 9: 3D trajectories of proposed multi-ML hybrid VIO for trajectory 2 when tested on a scenario different from training datasets.

Figure 9 illustrates the behaviour of the 3D trajectory obtained from comparing the proposed framework with other reference systems. It is observed that the proposed multi-ML hybrid VIO is able to recover from the derivation caused by altitude variation during flight take-off from different street and able to generalize effectively across seen and unseen scenarios. However, deviation in performance observed during landing and leading to drifting due to the fault factors that was not included in the training dataset as discussed above. These behaviours cause up to 17.15 m drifting in traditional ESKF-based VIO which is reduced to 1.01 m using the proposed framework that shows model adaptability in unseen scenarios and faults.

4.3 Comparison to Literature

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

Table 3 shows the comparative results for MH_05 sequence in EUROCC 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 degree of freedoms and image blur. However, flight dynamics and images within the dataset was consistent following a pattern throughout the flight. Notably, the comparative analysis only included hybrid VIO architectures in SOTA. Our proposed method has effectively mitigated erroneous measurements arises from various sources of uncertainty and showed improvement in 54cm and 33cm for 3D and horizontal positioning respectively (95th percentile). SOTA architectures suffered from performance degradation above 1m in 3D position due to insufficiency in adapting the dynamics presents in the data while using all the series of EUROCC datasets for training their models. For instance, presence of rapid motion and motion blur in the flight lead to feature tracking errors and feature location errors in VO. Integrating multiple error compensation within hybrid VIO architecture effectively mitigates the effects of such errors and improve position approximately 0.48 m in 3D position.

TABLE III
COMPARISON WITH SOTA METHODS THAT USED MH05_DIFFICULT SEQUENCE FROM EUROCC DATASET.

Techniques	RMSE (m)				95 th Percentile (m)	
	N	E	V	3D	Horizontal	3D
End-to-End VIO [7]	-	-	-	1.96	-	-
Self-Supervised VIO [9]	-	-	-	1.02	-	-
VIO with VO error compensation	0.24	0.33	0.25	0.66	0.46	0.73
VIO with KF error compensation	0.34	0.46	0.16	0.78	0.54	0.83
Multi-ML Hybrid VIO (proposed)	0.20	0.22	0.18	0.48	0.33	0.54

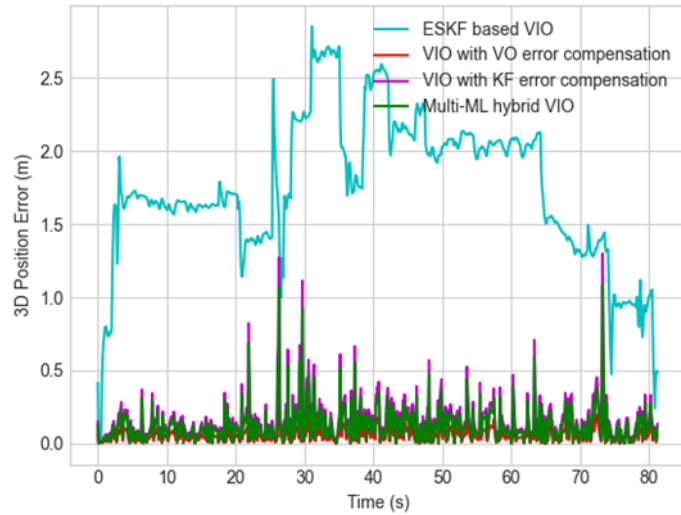


Figure 8: The 3D position error estimated using MH05_difficult seq. from EUROCC dataset with multiple sources of aleatoric uncertainty.

Figure 8 presents the 3D position error of the proposed multi-ML hybrid VIO for processing MH05_difficult dataset. The VIO with KF compensation algorithm has demonstrated least improvement among all the hybrid architectures as the correction mechanism relies on ESKF measurements. The ESKF-based VIO reference has experienced performance degradation due to nonlinear dynamics of MAV (rapid motion) and environmental factors (blur and darkness). However, KF error compensation mechanism has managed to mitigate the

nonlinearities e.g. non-gaussian noise, high-frequency vibrations, MAV dynamics and navigation environment complexity, reducing the error up to 83cm (95th percentile). In contrast, VIO with VO error compensation has performed better by compensating error before fusion, allowing the ESKF model to linearise only corrected sensor data. Integrating combination of VO and KF compensations mechanism significantly mitigate the data and model uncertainty arising from such fault sources and improved overall performance. The key finding is that the position error in RMSE is reduced up to 78% compared to hybrid SOTA methods i.e. end-to-end VIO.

5. CONCLUSION

This study demonstrated a proposed multi-ML hybrid VIO fusion framework using sensor data from IMU and camera to significantly reduce position uncertainty and improve performance. The framework incorporates GRU-based error compensation models that are trained to account for both epistemic and aleatoric uncertainty. To prove the benefits of the proposed approach, a series of experiments are conducted to evaluate the system performance under visual sources of uncertainty such as feature extraction error, feature tracking error, data association error, mismatch in dynamic and imperfect modelling. The testing and training datasets include realistic failure modes and fault factors presents in urban environments. The experimental results indicate that the proposed framework captures the increased position uncertainty in instances where sensor data is corrupted, or the model is imperfect. For example, test 2 demonstrates the frameworks' ability to handle uncertainty arising from the limited visibility of features, reflection, shadow, and flight dynamics. Despite more complex flight dynamics and noisy sensor in test 2, position uncertainty is reduced by 64% for VIO with VO error compensation, 66% for VIO with KF error compensation and 72% for multi-ML hybrid VIO (95th percentile). VIO with VO error compensation and VIO with KF error compensation perform well in both tests, but its effectiveness is varying based on navigation scenarios and flight dynamics. In less complex environment, the improvement is significant, with 81% and 80% reduction of position uncertainty respectively. However, performance degradation is observed in navigation under unseen fault conditions e.g. test 2. Nevertheless, the presence of outliers due to limited visibility and rapid angular motions introduce uncertainty in both data and model, which has been effectively minimized error distribution, with a peak between 0-4 m in horizontal axis. Importantly, while the framework can compensate for introduced errors, it may not repair or predict lost features, which remains a limitation in scenarios where critical features are completely missing. Test 1 exhibits better overall performance compared to test 2, as it involves less complicated scenarios and faults most closely aligned with those present in training data. Among the hybrid architectures VIO with VO error compensation outperforms VIO with KF error compensation. This is because the KF error prediction relies on ESKF model, which struggles to adapt to the complex dynamics and unable to linearise error states, leading to performance degradation. In contrast, compensating VO measurements before fusion provides more accurate input data for ESKF, allowing it to focus on linearising less complex noise and inaccuracies remaining in the data. Finally, 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 87% in test 1 and 54% in test 2 when compared to traditional ESKF-based VIO. In summary, the combined performance and uncertainty gain estimates reveal the proposed framework is capable of reducing position uncertainty under possible realistic complex urban environments, ensuring safe and uninterrupted flights for UAM application. To further enhance performance, incorporating another sensor may help reducing vertical drift. Future work will be focusing on predicting uncertainty from model-based error compensation module using Monto Carlo sampling, providing an additional level of confidence to ensure operational safety.

REFERENCE

- [1] 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.
- [2] 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.
- [3] 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.

- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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.



DSSPACE

<https://dspace.org/>

Position uncertainty reduction in visualInertial navigation systems using multi-ML error compensation

Tabassum, Tarafder Elmi; Petrunin, Ivan; Rana, Zeeshan A.

2024-10-09

Institute of Navigation

<https://doi.org/10.33012/2024.19790>

Tabassum TE, Petrunin I, Rana ZA. (2024) 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), 16-20 Sep 2024, Baltimore, Maryland, pp. 1741-1755