Title Title Title

XXX, YY and ZZ
Department of Computer Science and Mathematics, Lebanese American University, Beirut, Lebanon
Emails: {....}@lau.edu.lb

***Abstract*—.......**

....

....

***Index Terms*—Keyword1, keyword2,....**

I. Introduction

..... [1] and [2]

Unlike existing work, ...

The contributions of the work can be summarized as follows:

- We consider
- We
- We then propose

The rest of the paper is organized as follows. The literature review is presented in Section II. The system model is provided in Section III. .... Conclusions are drawn in Section VI.

II. Literature Review

....

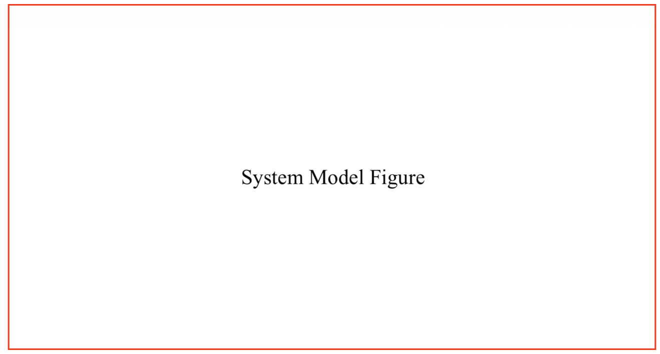

Fig. 1. Network composed of ......

.....

III. System Model

As shown in Fig 1, the proposed system models a UAV operating in a GPS-denied indoor environment containing both static and dynamic obstacles. The UAV is equipped with an RGB-D camera and an IMU that provide synchronized depth, visual, and inertial data streams. The collected sensory data are processed through three main modules: ORB-SLAM3 for visual-inertial localization and mapping, YOLOv5 for real-time semantic object detection, and a dynamic object prediction filter for estimating future trajectories. The outputs of these modules are fused into a semantic-dynamic map that represents both the static structure and the predicted motion of dynamic entities, forming the foundation for proactive UAV navigation and safe landing decisions.

A. ORB-SLAM3

ORB-SLAM3, the most advanced algorithm, which demonstrates robust real-time performance across diverse environments, including various indoor, outdoor, and multi-scale scenarios. It serves as the visual–inertial odometry core of the system, responsible for estimating the UAV's six-degree-of-freedom (6-DoF) pose and reconstructing a geometric map of the environment. It fuses information from the RGB-D camera and IMU to achieve robust localization in GPS-denied indoor environments.

B. YOLOv5

YOLOv5 serves as the semantic perception module of the system, performing real-time detection and segmentation of objects within the UAV's visual field. The YOLOv5-Seg variant produces pixel-level masks for dynamic objects, enhancing spatial precision over standard bounding box detection. By distinguishing between dynamic and static regions, the system labels features from moving entities as dynamic and those from static structures as stationary. The detected objects and their segmentation masks are then passed to the dynamic prediction filter for further motion analysis and trajectory estimation.

C. Dynamic Object Prediction Filter

After YOLOv5 detects and segments dynamic objects, the Extended Kalman Filter (EKF) initializes and tracks the state of each detected object—typically including its position and velocity—using the bounding box coordinates and class information. The EKF extends the standard Kalman Filter to handle nonlinear state transition and observation models, making it suitable for predicting object motion in complex indoor environments where trajectories may not be strictly linear.

The EKF operates in two alternating stages: *prediction* and *update*. During the prediction step, the filter estimates the next state of the object $\hat{x}_{k|k-1}$ based on its previous state $\hat{x}_{k-1|k-1}$ and the motion model $f(\cdot)$:

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}, u_{k-1}) \tag{1}$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \tag{2}$$

where $F_k = \frac{\partial f}{\partial x}\big|_{\hat{x}_{k-1|k-1}}$ is the Jacobian of the motion model, Q_k is the process noise covariance, and $P_{k|k-1}$ represents the predicted covariance estimate.

When a new detection from YOLOv5 is received, the update step corrects the predicted state using the observation model $h(\cdot)$:

$$\tilde{y}_k = z_k - h(\hat{x}_{k|k-1}) \tag{3}$$

$$S_k = H_k P_{k|k-1} H_k^T + R_k \tag{4}$$

$$K_k = P_{k|k-1} H_k^T S_k^{-1} \tag{5}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k \tag{6}$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \tag{7}$$

where $H_k = \frac{\partial h}{\partial x}\big|_{\hat{x}_{k|k-1}}$ is the Jacobian of the observation model, R_k is the measurement noise covariance, and K_k is the Kalman gain.

This iterative process allows the EKF to continuously refine the estimated positions and velocities of dynamic objects while effectively handling sensor noise and detection uncertainties caused by occlusions, lighting variations, or rapid motion. The predicted trajectories are fused into the semantic-dynamic map, allowing the UAV to anticipate future movements of obstacles and adjust its path in real time to ensure safe, collision-free navigation.

D. Semantic-Dynamic Map Fusion

The Semantic-Dynamic Map Fusion module enhances the UAV's environmental understanding by integrating geometric, semantic, and dynamic information into a unified representation of the surroundings. The geometric map from ORB-SLAM3 provides structural context, while semantic detections from YOLOv5 add information about object types and spatial relationships. Incorporating predicted trajectories from the EKF enables the UAV to track moving objects and anticipate their interactions with static elements in the environment. This fusion allows the UAV to identify and avoid both static and dynamic obstacles, plan safe and efficient paths, and make real-time navigation decisions based on current environmental conditions. By recognizing objects such as humans and predicting their motion, the UAV gains contextual awareness crucial for safe operation in crowded or changing indoor spaces. The combined use of geometric and semantic data ensures robustness to environmental variations and supports adaptive, multi-modal perception for tasks like object recognition, obstacle avoidance, and scene understanding—resulting in a more intelligent and resilient navigation system.

IV. Proposed Approach

In this section, we propose .....

V. Performance Results and Analysis

In this section, we present first the simulation setup and system parameters. We then provide the performance evaluation of our proposed approach.

A. Simulation Setup

....

TABLE I
SYSTEM PARAMETERS

Parameter	Value
A	....
B	....
C	....

B. Performance Results and Analysis

......

VI. CONCLUSION

REFERENCES

[1] M. Dai, N. Huang, Y. Wu, J. Gao and Z. Su, "Unmanned-Aerial-Vehicle-Assisted Wireless Networks: Advancements, Challenges, and Solutions," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4117–4147, 2023.

[2] B. Choudhury, V. K. Shah, A. Ferdowsi, J. H. Reed and Y. T. Hou, "AoI-minimizing Scheduling in UAV-relayed IoT Networks," *in Proceedings of the IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems*, 2021.