Vision-Based UAV Localization Using a Fisheye Ground-Camera Network for Sensor Package Deployment in Post-Disaster Structure Inspection

Nolan Shute[a], Md Asifuzzaman Khan[a], Joud N. Satme[a], Mark Zheng[a], Korebami Adebajo[a], Austin R.J. Downey[a,b], and Matt Folsom[c]

[a]Department of Mechanical Engineering, University of South Carolina, Columbia, USA
[b]Department of Civil and Environmental Engineering, University of South Carolina, Columbia, USA
[c]Interactive Aptitude LLC, Tyngsboro, MA, USC

ABSTRACT

In the aftermath of structural failure, rapid assessment of damage is essential for understanding failure mechanisms and ensuring safety during inspection. Unmanned aerial vehicles (UAVs) provide a means to perform structural health monitoring in hazardous environments while reducing risk to human operators. However, autonomous navigation of UAVs remains challenging in GPS-denied environments, such as within large or metallic structures, where conventional localization methods are unreliable. Existing approaches rely on onboard sensing or external motion capture systems, which can be limited by cost, accessibility, or environmental constraints. Ground-based vision systems offer an alternative, but achieving accurate, real-time localization from limited viewpoints remains an open challenge. Here we present a ground camera network using upward-facing fisheye cameras to localize a UAV through machine learning–based coordinate mapping. We show that combining stereo vision with regression-based transformation enables reconstruction of UAV trajectories using only ground observations, without reliance on satellite positioning. The proposed system translates image detections into real-world coordinates and enables controlled navigation for tasks such as sensor deployment. Compared to traditional motion capture systems, the method provides a lower-cost and more deployable solution while maintaining strong agreement with reference measurements. The approach demonstrates that accurate localization can be achieved using minimal sensing infrastructure, though performance depends on consistent visibility and detection of the UAV. More broadly, this framework establishes a pathway toward autonomous UAV operation in confined or obstructed environments where conventional positioning systems fail, with applications in post-disaster inspection and structural health monitoring. In experimental validation, the system tracked UAV motion over a 46 s flight and achieved a root mean squared localization error of approximately 10.8 mm, with average errors on the order of 1 cm relative to motion capture reference data.

Keywords: Structural Health Monitoring, Unmanned Aerial Vehicles, Sensor Package Deployment, Ground Camera Network, Camera Packages

1. INTRODUCTION

Catastrophic structural failures require rapid evaluation to determine both the causes and extent of damage. In post-seismic or collapse-prone environments, measuring structural response often requires the placement of sensors within unstable buildings.[1] Manual sensor deployment in such environments exposes personnel to significant safety risks, including falling debris, structural instability, and restricted access. To mitigate these risks, recent research has explored automated sensor placement using unmanned aerial vehicles (UAVs).[2,3] UAV-assisted structural health monitoring enables rapid data collection in hazardous or inaccessible environments while minimizing human exposure and expanding the range of feasible inspection scenarios.[4]

Further author information: (Send correspondence to Austin Downey)
Austin Downey: Email: austindowney@sc.edu

However, for UAV-based systems to operate fully autonomously, they must possess sufficient environmental awareness to safely and accurately navigate complex structural environments. A major limitation arises in GPS-denied environments,[5] such as those within large metallic structures, where conventional satellite-based localization becomes unreliable.[6] In these conditions, navigation approaches based on inertial sensing suffer from accumulated error, while alternative methods such as ultra-wideband ranging are degraded by multipath interference.[7,8] Vision-based localization has therefore emerged as a promising alternative, enabling UAVs to interpret their surroundings using image data;[9] however, onboard vision systems are often constrained by limited field of view and computational burden, motivating the use of external camera networks for precise control in complicated or dangerous flight environments.

Recent work by the authors has explored external vision-based approaches for UAV localization and autonomous sensor deployment. A stereo camera system combined with a You Only Look Once (YOLO)–based object detection algorithm to track UAV position in three-dimensional pixel space was developed, enabling real-time identification within the operational environment.[10] We then introduced a machine learning–based framework to map these pixel coordinates to real-world positions using supervised regression models trained on synchronized camera and motion capture data.[11] While this approach achieved centimeter-level localization accuracy, performance depends on maintaining visibility of the UAV, as forward-facing camera configurations may not fully capture the docking region.

In this work, a ground camera network for UAV localization in GPS-denied environments was developed for post-disaster structural inspection. The system uses upward-looking cameras with fisheye lenses to provide wide-field coverage, improving visibility and robustness to occlusion. By integrating object detection, stereo vision, and machine learning–based mapping, the framework enables accurate position estimation and controlled flight. The contributions are: (1) an upward-looking camera configuration, (2) fisheye-based wide-field tracking, and (3) experimental validation of localization performance. Code, data, and artifacts are publicly available.[12]

2. METHODOLOGY

This section describes the experimental setup and data processing pipeline used for UAV localization and navigation. Relevant data, code, and artifacts related to this paper are available through a public repository.[12]

2.1 Experimental Setup

The overall experimental configuration is shown in Figure 1. The system consists of a ground-based camera network composed of two upward-facing fisheye cameras positioned beneath the UAV operating region. As illustrated in Figure 1(a), the cameras are placed on the ground and oriented vertically to capture the UAV from below during flight. Figure 1(b) shows the corresponding top-down view, highlighting the overlapping fields of view that provide continuous visual coverage of the UAV within the workspace.

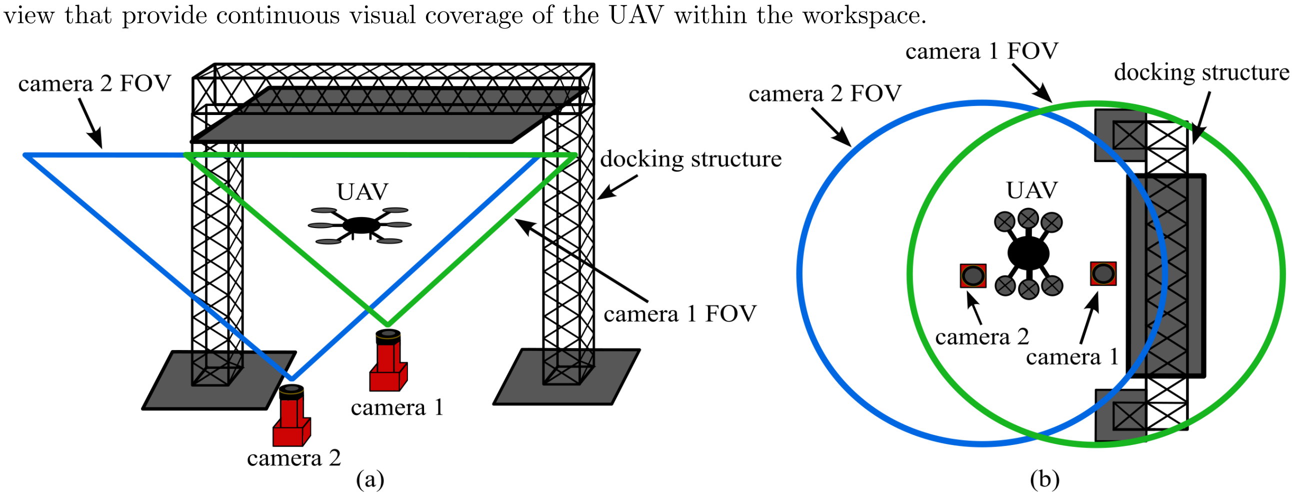
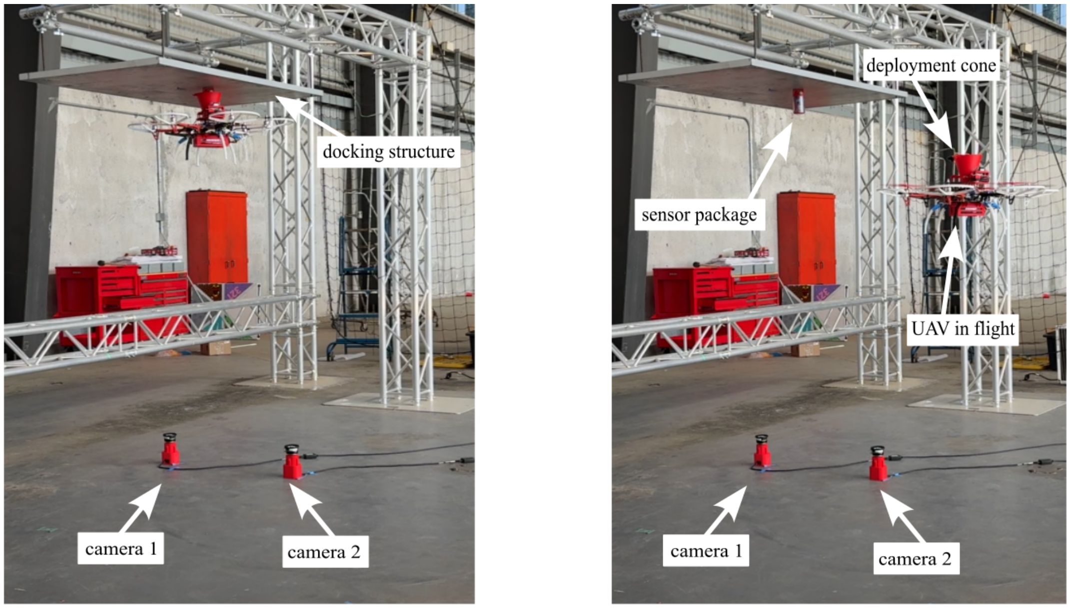
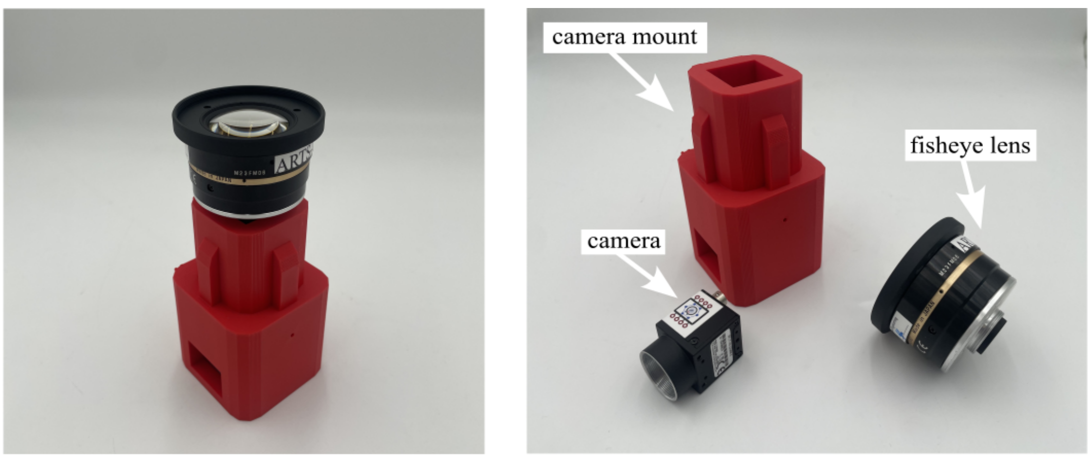


Figure 1. Proposed ground camera network for UAV localization, showing: (a) two upward-facing fisheye cameras tracking the UAV during flight, and; (b) camera positions depicted from an above-ground perspective.



(a) (b)

Figure 2. UAV operation within the experimental setup, showing: (a) ground cameras positioned beneath the UAV capturing flight data, and; (b) successful sensor package docking during flight.



(a) (b)

Figure 3. Camera package design for UAV tracking, showing: (a) assembled camera package, and; (b) labeled components of the system.

The experimental setup is implemented within a UAV flight cage designed to simulate under-structure environments relevant to sensor deployment, as shown in Figure 2. The UAV operates beneath a mock structural surface, performing maneuvers associated with sensor package deployment. As shown in Figure 2(a), the ground cameras are positioned beneath the UAV to capture flight data, while Figure 2(b) illustrates the UAV during a sensor docking operation against the underside of the structure. The cameras are placed on the floor of the cage with a baseline separation of 0.72 m (28.5 in) and are configured to maintain visibility of the UAV throughout its motion. Identical camera models and lenses are used at both positions to ensure consistent imaging characteristics. The camera arrangement provides complementary viewing angles of the UAV during flight, enabling robust detection and localization. A human pilot performs controlled flight maneuvers, including hovering and sensor package deployment,[2,3] to generate representative trajectories within the workspace. This setup enables consistent and repeatable data collection for training and evaluation of the machine learning–based localization framework.

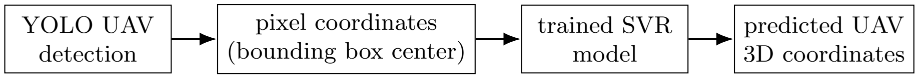


Figure 4. Inference workflow of the proposed UAV localization pipeline.

The camera package shown in Figure 3 consists of a housing 3D-printed in PLA and designed to support a machine vision camera in an upright configuration. Image data were acquired using 2.3 MP monochrome cameras (BFLY-U3-23S6M-C, Teledyne FLIR, USA), selected for their high sensitivity, frame rate capability, and improved performance in low-light conditions. Each camera was equipped with a 6 mm fixed focal length C-mount fisheye lens (M23FM06, Tamron Co., Ltd., Japan), providing a wide field of view for tracking the UAV within the experimental workspace. Image acquisition and processing were performed using the Spinnaker SDK (Teledyne FLIR).

2.2 Data Processing and Localization Framework

Previous work by Khan et al.[11] established a machine learning–based framework for UAV localization by mapping image-derived pixel coordinates to real-world positions using supervised regression. As illustrated in Figure 4, this approach consists of detecting the UAV in image frames, extracting pixel coordinates, and mapping these observations to three-dimensional positions using a trained regression model. In this framework, pixel coordinates obtained from camera observations are paired with ground-truth positions from an OptiTrack motion capture system to train a model that predicts the UAV position in three-dimensional space. This motion capture system uses reflective markers to track object motion and position within an environment and is commonly used as a reference standard in camera-based localization studies.[15] In this study, motion capture system data serve as ground truth for both training and validation of the regression model, enabling direct comparison between predicted and measured UAV trajectories.

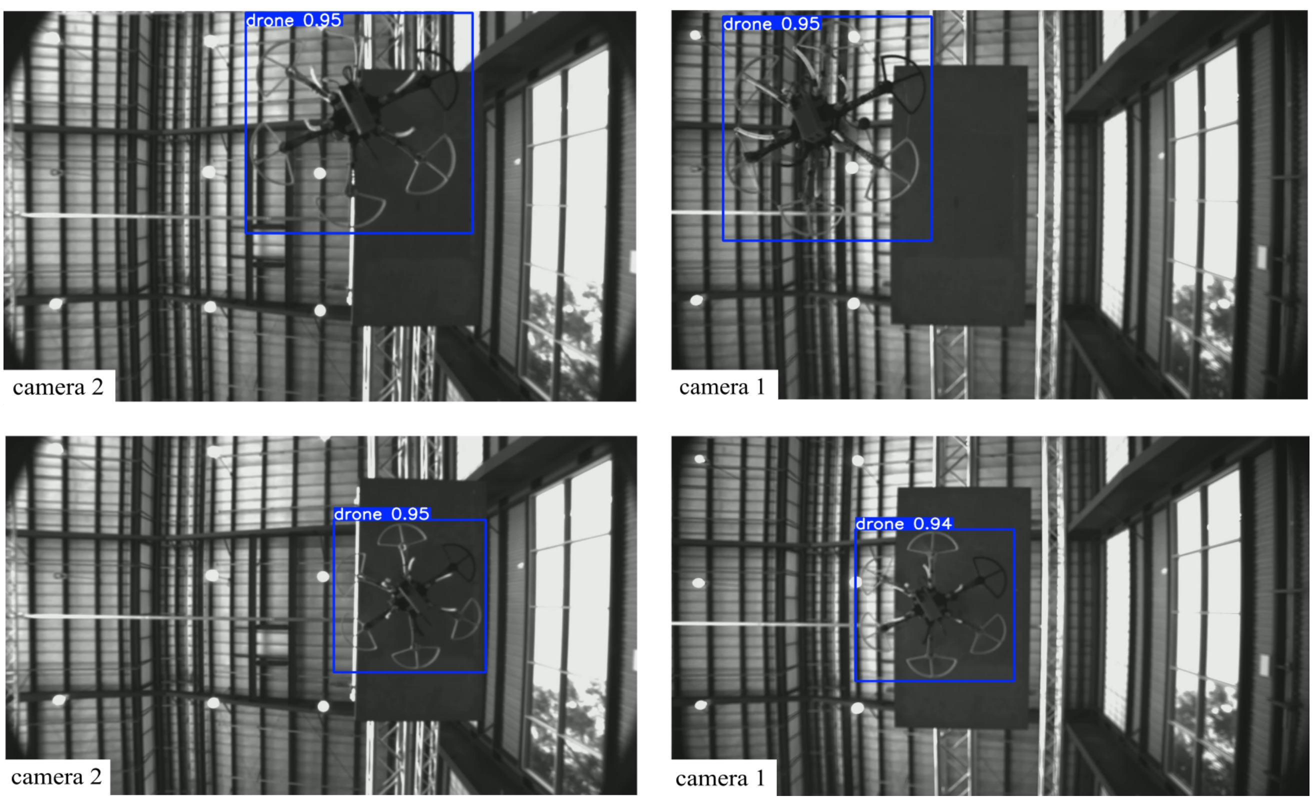


Figure 5. Frames used for training the YOLO model to detect the UAV, with both camera angles employed at two distinct positions.

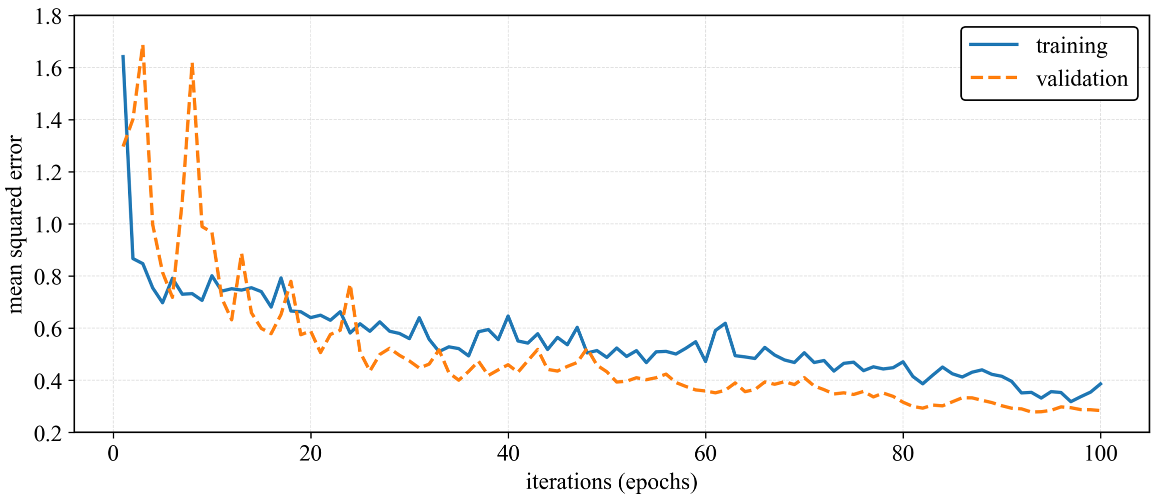


Figure 6. The training plot of the UAV with the corresponding epochs and the mean squared error to show the difference in the training model and validation, where the error was the lowest around the iteration of 70 epochs.

To generate the image-based observations, synchronized video streams from both cameras were processed using a YOLO-based object detection model to identify the UAV in each frame. As shown in Figure 5, the model detects the UAV from a ground-based perspective and produces bounding boxes corresponding to its location in each image. The center of each bounding box is used to extract pixel coordinates representing the UAV position. Frame timestamps were used to align detections across both camera views, producing synchronized pixel coordinate pairs for each time step. The YOLO model was trained using a combination of manual labeling through Roboflow, with approximately 100 labeled frames, and further training implemented in Python using a dataset of approximately 130 images over 100 epochs. The resulting training performance is shown in Figure 6, where a reduction in mean squared error over successive epochs indicates improved detection accuracy and model convergence.

The synchronized pixel coordinates from both camera views are then used as inputs to a supervised regression model. Following the approach of Khan et al.,[10,11] a support vector regression (SVR) model is trained to map image-based observations to real-world 3D coordinates using corresponding motion capture system measurements. Once trained, the model is applied to the full dataset to estimate UAV positions throughout the flight. This process enables reconstruction of the UAV trajectory in three-dimensional space using only image data, providing a vision-based localization framework for operation in GPS-denied environments.

3. RESULTS

Figure 7 presents the SVR model predictions compared to the motion capture system ground truth for the 20% test dataset. The model was trained on 80% of the data and evaluated on the remaining 20%. The predicted coordinates closely follow the reference trajectory across all three axes, demonstrating accurate mapping from image-based observations to real-world positions. The SVR model used hyperparameters $C = 1000$ and $\gamma = 1$. Model performance was quantified using a root mean squared error (RMSE) of 10.81 mm.

The three-dimensional trajectory plots in Figure 8 further illustrate the accuracy of the model, capturing both altitude and lateral motion consistent with the reference data. The predicted trajectory closely follows the motion capture ground truth throughout the flight, demonstrating accurate reconstruction of the UAV motion using the proposed vision-based localization framework. Minor deviations and increased noise are observed near the end of the trajectory in Figure 8(a), where the UAV is docked to the structure during sensor package deployment. In this phase, the UAV experiences limited motion and intermittent occlusion, which introduces variability in the detected pixel coordinates and results in localized fluctuations in the predicted position.

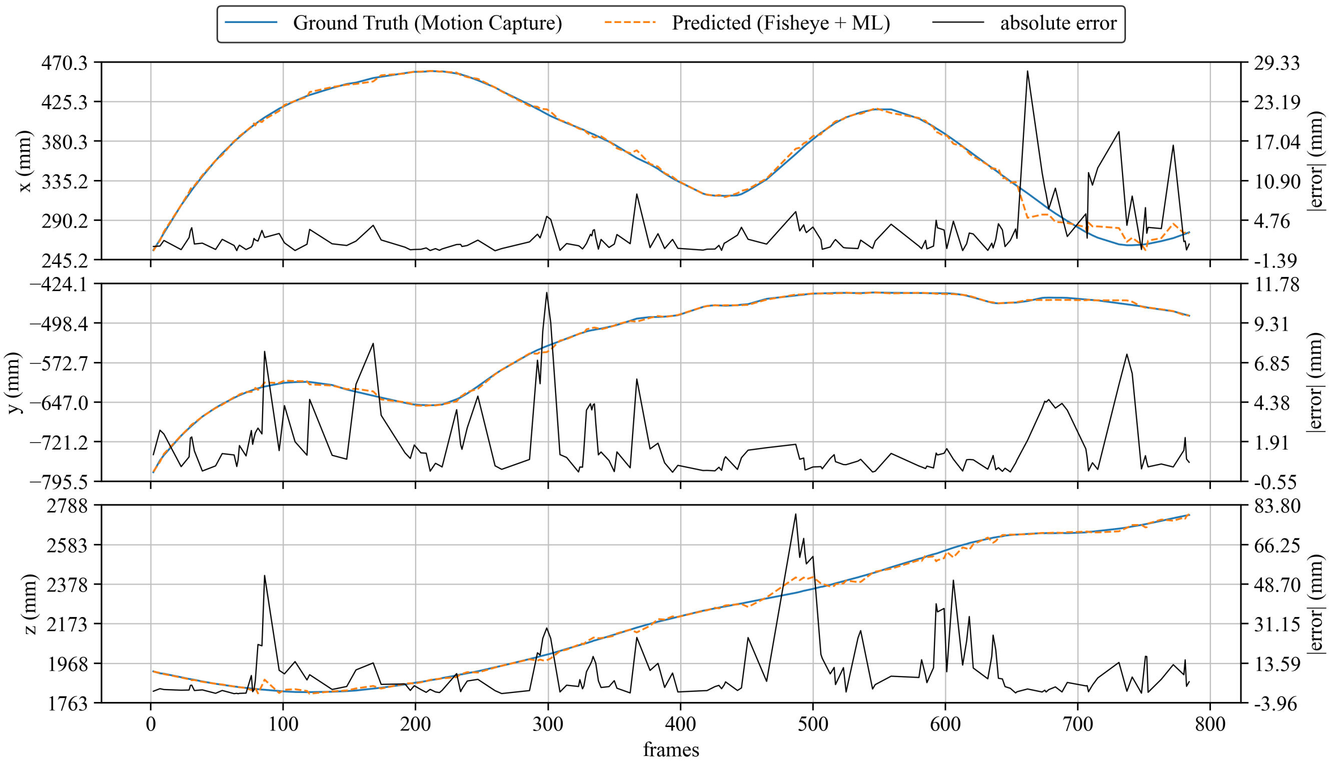


Figure 7. The 20% comparison plot between the SVR predicted coordinates and the motion capture system coordinates, with absolute data plotted alongside the location of the UAV on each axis.

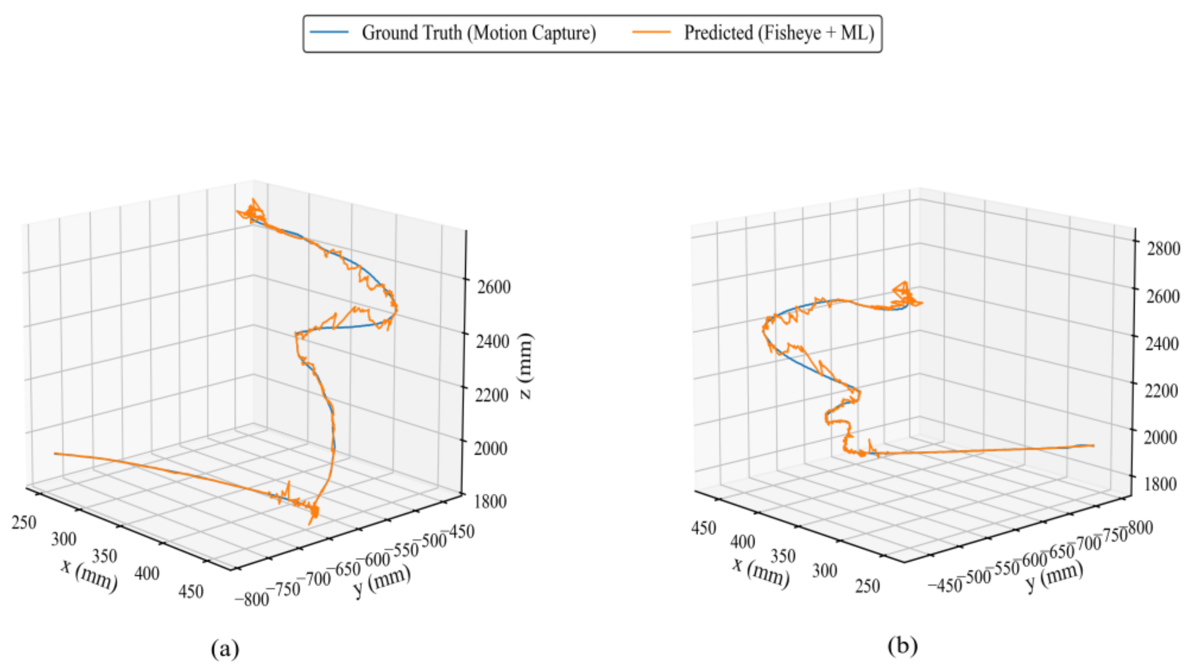


Figure 8. Three-dimensional UAV flight trajectory reconstruction using machine learning, showing: (a) predicted flight path based on image-derived coordinates compared to the motion capture system reference data, and; (b) alternate view of the trajectory illustrating full spatial motion.

4. CONCLUSION

This study developed and evaluated a ground-based fisheye camera network for estimating UAV position in GPS-denied environments, with a primary focus on three-dimensional coordinate mapping and trajectory reconstruction. By combining multi-view image observations with machine learning–based coordinate transformation, the system was able to recover UAV flight trajectories using only ground-based visual data. Experimental results demonstrate that the proposed framework provides accurate and consistent localization performance. The support vector regression (SVR) model showed strong agreement with OptiTrack reference data, achieving a root mean squared error (RMSE) of 10.8 mm, with average errors on the order of 1 cm during a 46 s flight. These results indicate that the proposed pipeline—object detection and regression-based coordinate mapping—can produce accurate and physically consistent trajectory estimates under controlled conditions. However, the results also highlight key limitations. The accuracy of the system is highly sensitive to consistent UAV detection and multi-view visibility; loss of tracking or partial occlusion can degrade localization performance. Additionally, the current implementation does not operate in real time and relies on post-processed data, limiting its immediate applicability to autonomous flight control. While comparison with motion capture system data shows promising agreement, validation remains limited in scope and requires expansion across larger datasets and more varied flight conditions. Overall, this work establishes a proof-of-concept framework for ground-based optical localization of UAVs using fisheye cameras and machine learning. Future work will focus on improving robustness of detection, implementing real-time processing, and extending the system to full 3D localization and closed-loop navigation. With these developments, the approach has the potential to contribute to autonomous UAV operation for structural inspection in GPS-denied environments.

ACKNOWLEDGMENTS

This material is based upon work supported by the Air Force Office of Scientific Research (AFOSR) through award no. FA9550-21-1-0083. This work is also partly supported by the National Science Foundation (NSF) grant numbers CCF - 1956071, CMMI - 2152896, CCF-2234921 and, CPS - 2237696. Additional funding for this work comes from the Office of Naval Research through the award number N000142412727. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force, the National Science Foundation, or the United States Navy. This work was also partially supported, in part, at the University of South Carolina by the State of South Carolina through funding for the Battelle Savannah River Alliance Workforce Development Program.

REFERENCES

[1] López-Castro, B., Haro-Baez, A. G., Arcos-Aviles, D., Barreno-Riera, M., and Landázuri-Avilés, B., "A systematic review of structural health monitoring systems to strengthen post-earthquake assessment procedures," *Sensors* **22**, 9206 (Nov. 2022).

[2] Carroll, S., Satme, J., Alkharusi, S., Vitzilaios, N., Downey, A., and Rizos, D., "Drone-based vibration monitoring and assessment of structures," *Applied Sciences* **11**, 8560 (Sept. 2021).

[3] Satme, J. N., Yount, R., Goujevskii, N., Jannazzo, L., and Downey, A. R. J., "Sensor package deployment and recovery cone with integrated video streaming for rapid structural health monitoring," in [*ASME 2024 Conference on Smart Materials, Adaptive Structures and Intelligent Systems*], *SMASIS2024*, American Society of Mechanical Engineers (Sept. 2024).

[4] Laflamme, S., Blasch, E., Ubertini, F., Liu, Z., Wertz, J., Knott, C., Cherry, M., Lindgren, E., Chang, F.-K., Kumar, A., Poole, J., Worden, K., Downey, A., Wei, J., Musgrave, P. F., Wong, A. S., Quaranta, G., Rosso, M. M., Marano, G. C., Chen, Y., Ardiles-Cruz, E., Soleimani-Babakamali, M. H., Avci, O., Inman, D. J., Taciroglu, E., Dodson, J., Chen, G., Meng, W., Zhu, C., Liu, Z., Zuo, J., Liu, Q., Khan, S., Hu, C., Hu, Z., Cicirello, A., Cross, E., Chatzi, E., Weng, Y., Yuan, J., Wen, S., Han, L., Metaxas, D., Tronci, E. M., Moaveni, B., Chen, Q., Ng, M. S., Hackl, J., Chen, G., Wei, S., Mitoulis, S.-A., Izonin, I., Uva, G., Ruggieri, S., Mao, Z., Kiranyaz, S., Devecioglu, O. C., Gabbouj, M., and Mohammadi, J., "Roadmap: Integrating artificial intelligence in structural health monitoring systems," *Measurement Science and Technology* **37**, 103001 (Mar. 2026).

[5] Al-Jarrah, O. Y., Shatnawi, A. S., Shurman, M. M., Ramadan, O. A., and Muhaidat, S., "Exploring deep learning-based visual localization techniques for UAVs in GPS-denied environments," *IEEE Access* **12**, 113049–113071 (2024).

[6] Kos, T., Markezic, I., and Pokrajcic, J., "Effects of multipath reception on GPS positioning performance," in [*Proceedings ELMAR-2010*], 399–402 (2010).

[7] Petritoli, E., Leccese, F., and Spagnolo, G. S., "Inertial navigation systems (ins) for drones: Position errors model," in [*2020 IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*], 500–504, IEEE (June 2020).

[8] Silva, B. and Hancke, G. P., "IR-UWB-based non-line-of-sight identification in harsh environments: Principles and challenges," *IEEE Transactions on Industrial Informatics* **12**, 1188–1195 (June 2016).

[9] Liu, H., Long, Q., Yi, B., and Jiang, W., "A survey of sensors based autonomous unmanned aerial vehicle (UAV) localization techniques," *Complex and Intelligent Systems* **11** (July 2025).

[10] Zheng, Q. M., Khan, A., N., J., Adebajo, K., Yount, R., and Downey, A. R., "Stereo yolo uav localization and tracking enabling autonomous sensor deployment on critical infrastructure," in [*AIAA SCITECH 2026 Forum*], American Institute of Aeronautics and Astronautics (Jan. 2026).

[11] Khan, A., N., J., Zheng, Q. M., Tadamatla, S., Phillips, E., Adebajo, K., Shute, N., Yount, R., and Downey, A. R., "Machine learning-based UAV localization for under-structure sensor deployment in GPS-denied environments," in [*SPIE Smart Structures and Nondestructive Evaluation*], SPIE (Mar. 2026).

[12] ARTS-Lab, "Paper 2026 Fish-eye ground camera network for UAV localization." https://github.com/ARTS-Laboratory/paper-2026-fisheye-ground-camera-network-UAV-localization (2026). Github.

[13] Himawan, R. W., Baylon, P. B. A., Sembiring, J., and Jenie, Y. I., "Development of an indoor visual-based monocular positioning system for multirotor uav," in [*2023 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES)*], 1–7, IEEE (Oct. 2023).

[14] Andrean, S. Y., Joelianto, E., Widyotriatmo, A., and Adiprawita, W., "Low cost vision-based 3D localization system for indoor unmanned aerial vehicles," in [*2013 International Conference on Robotics, Biomimetics, Intelligent Computational Systems*], 237–241, IEEE (Nov. 2013).

[15] NaturalPoint Inc., "OptiTrack Motion Capture System." https://www.optitrack.com (2024). Accessed: 2026-01-15.

[16] Benesty, J., Chen, J., Huang, Y., and Cohen, I., [*Pearson Correlation Coefficient*], 1–4, Springer Berlin Heidelberg (2009).

[17] Arlot, S. and Celisse, A., "A survey of cross-validation procedures for model selection," *Statistics Surveys* **4** (Jan. 2010).