Vision-based Control of UAV for Autonomous Firefighting

Ankit Mehra
SMME
Indian Institute of Technology, Mandi
Himachal Pradesh, India
s22020@students.iitmandi.ac.in

Amit Shukla
CAIR
Indian Institute of Technology, Mandi
Himachal Pradesh, India
amitshukla@iitmandi.ac.in

Darshankumar Prajapati
CAIR
Indian Institute of Technology, Mandi
Himachal Pradesh, India
s22053@students.iitmandi.ac.in

Pushkar Kumar
SMME
Indian Institute of Technology, Mandi
Himachal Pradesh, India
s22038@students.iitmandi.ac.in

Ashish Rana
SMME
Indian Institute of Technology, Mandi
Himachal Pradesh, India
s22019@students.iitmandi.ac.in

Tushar Patil
CAIR
Indian Institute of Technology, Mandi
Himachal Pradesh, India
tushar_patil@projects.iitmandi.ac.in

2024 16th International Conference on Computer and Automation Engineering (ICCAE) | 979-8-3503-7005-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCAE59995.2024.10569529

***Abstract*—Urban and industrial environments frequently necessitate the immediate readiness of firefighting personnel to address potential fire emergencies, which can lead to a perpetual shortage of available manpower. To counter this challenge, we propose an integrated approach to autonomous firefighting through the utilization of Unmanned Aerial Vehicles (UAVs). The UAVs serve a dual purpose: providing auxiliary support to conventional firefighting efforts while simultaneously mitigating risks to human life. Our proposed methodology incorporates deep learning-based fire detection. The system combines image feature analysis with data obtained from distance sensor to establish Cartesian coordinates of the identified fire sources. Autonomous vision-based control for multirotor platform is developed, guided by the extracted image features and Cartesian coordinates. This system is designed to facilitate both swift deployment and autonomous operation, while still allowing for manual intervention as necessary. We present experimental validations for the vision-based control of the multirotor platform under outdoor conditions. Further tests are conducted to assess the performance characteristics of the spray subassembly, including spray distance and flow rate, prior to its deployment on aerial platforms. The experimental results are discussed for performance analysis of the approach.**

***Index Terms*—Fire detection, Vision-based control algorithm, Unmanned Aerial Vehicle(UAV), Robotics and Automation, Autonomous firefighting.**

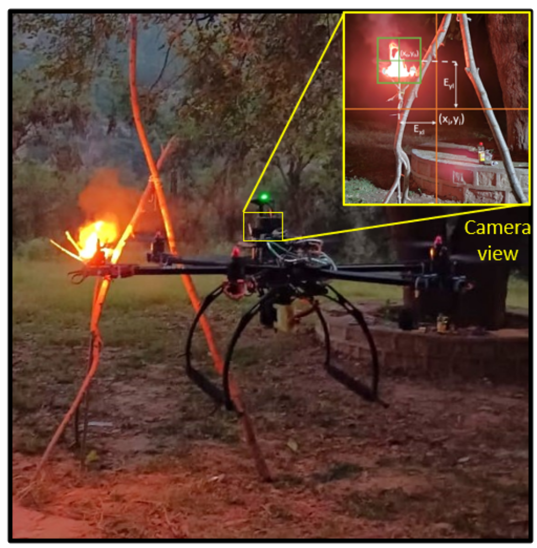

Fig. 1. A firefighting multicopter incorporates a visual and distance sensor with onboard computer leveraging deep learning techniques providing accurate localisation of fire, enhancing control for effective tracking of fires.

I. Introduction

Fire-related incidents present substantial threats to businesses, communities, and operational environments. As rapidly expanding global economy fuels a dramatic increase in the construction of high-rise buildings, the world faces a mounting challenge. Recent history underscores this challenge, as there have been multiple instances of devastating fires, resulting in the tragic loss of human lives and valuable property. For example, on January 26, 2023[1], a major fire engulfed the 22nd floor of a high-rise building in Dadar, Central Mumbai, subsequently spreading to other parts of the structure. This incident serves as a poignant illustration of the risks associated with high-rise fires. On October 6, 2023[2], a substantial fire erupted in a building in Mumbai's Goregaon, claiming at least eight lives, injuring 40 others, and causing the destruction of numerous two-wheelers and cars. The severity of these events underscores the need for innovative firefighting solutions, particularly in scenarios where the fire site is inaccessible to conventional firefighting equipment such as fire trucks.

At present, the firefighting landscape is still predominantly reliant on fire trucks. Fire departments now have introduced robotic systems to assist firefighters in suppression and control operations. The utility of UAVs[3] in firefighting operations is expanding. While they have predominantly served roles in real-time surveillance, data collection, thermal and infrared imaging, and post-fire assessments, notable advancements have been made. Drones like ”Skydio-X2”[4] provide 360° site visibility and live thermal imaging, while ”DJI-Mavic2” and ”Fotokite-Sigma” are deployed for real-time surveillance at accident sites. Notably, ”Walkera-WK1900” [5]and ”Ehang-216F” [6] are multicopter engineered with fire-extinguishing capabilities, allowing them to deploy firefighting agents and suppress flames effectively.

A. Related work and present contribution

Object detection and visual servoying are commonly used techniques in automation systems. Many reaserchers have reported vision based control techniques [7], [8] and object detection techniques [9], [10]. Accurate tracking of fire is essential for the UAV to achieve smooth motion during the extinguishing process. Hence, achieving autonomous fire extinguishment necessitates precise fire tracking, underscoring the imperative integration of a robust real-time object detection system within a vision-based control framework and a proficient spraying mechanism. Autonomous detection of fire in UAV using Single Shot MultiBox Detector (SSD) algorithm [9] and vison based control [7] are individually explored in the literature. However integrating both these systems with spraying subsystem gives us a unique application of extinguishing fire in high rise buildings.

We present an integrated architecture built upon a well-established deep learning model and a feedback controller, our architecture is adaptable for deployment in any UAV equipped with an onboard computer, thereby ensuring scalability and practicality in real-world fire scenarios. To the best of our knowledge this work is the first to present the fire tracking in the image plane using onboard computational unit on a UAV. The paper presents the following sections: ‘Overall Methodology,’ explaining the utilized hardware, the object detection model, communication and control strategy, followed by the section of ‘Experimental Results and Analysis,’ unveiling the outcomes of our experiments followed by ‘Conclusion’.

B. Nomenclature

E_x^I, E_y^I Error in x and y direction in image plane.
E_x^B Error in x direction in body frame.
d_x Distance by LiDAR in x direction in body frame.
ϕ, θ, ψ Roll, Pitch and Yaw angles of the drone.
V_x^I, V_y^I Velocity in image plane.
V_x^B, V_y^B, V_z^B Velocity in body frame.
V_x^W, V_y^W, V_z^W Velocity in world frame.
x^V, y^V Position of fire in Virtual frame.

II. Overall Methodology

The onboard camera continuously captures image frames, utilizing real-time image processing and object detection techniques, the system identifies the presence of fires in the frames. Upon fire detection, the vision based control[11] [12] strategy is engaged to calculate and execute the necessary adjustments in the multicopter’s position and orientation. The multicopter autonomously aligns itself with the detected fire’s location, ensuring optimal proximity for effective fire suppression. The control algorithm is defined in Fig. 3.

A. Hardware desciption

We have used two platforms to perform experiments, a small UAV to test the control algorithm and a bigger UAV to test the spraying. The platform we used for spraying has been chosen keeping in mind the time required for suppressing and control operations, if we wish to do the operation using a UAV we require longer flight time and adequate amount of fire suppressant. We have prepared a subassembly that will be mounted on our UAV to perform the spraying of fire suppresants. Selection of subassembly has been done considering various factors like spray distance, flow rate, operation time, weight of the subassembly. Pump based subassembly was selected for the opearation, it comprises of two Hobbywing 8L brushless water pump, adjustable spray pattern nozzle, EFT 16 L tank, silicon connecting tubes, carbon fibre base platform for mounting. Our control test platform comprises a visual sensor, distance sensor, an onboard computational unit as shown in Fig. 2.

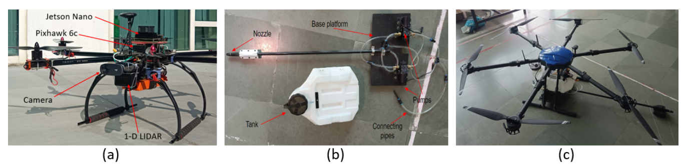

Fig. 2. (a) Control test platform (b) Pump Subassembly (c) Multicopter with subassembly.

B. Detection using Deep Learning

Classical vision methods traditionally rely on handcrafted features and engineered algorithms, while deep learning techniques leverage the power of neural networks to automatically learn hierarchical representations from data. The ability of autonomously extracting relevant features and patterns, enables enhanced adaptability to complex visual scenarios, making deep learning a better choice for object detection[13]. For the detection of fire accurately we are using YOLOv8 [14] object detection architecture, it is a deep learning model specifically designed for real-time and high-accuracy object detection tasks. Most of the datasets available[15] for fire are limited, they either have images of big fire accidents like gas industry explosion or very small fire like candles. For our application we need images of medium size fire in different luminance conditions. To address this need we created a custom dataset by taking multiple videos in different luminance conditions and extracted frames from it, a total of 400 frames were extracted. We annotated them using ‘labelImg’, following are the details of annotations.

1) <filename> Image name.
2) <path> Path of the location of image
3) <source> The dataset name and source of image.

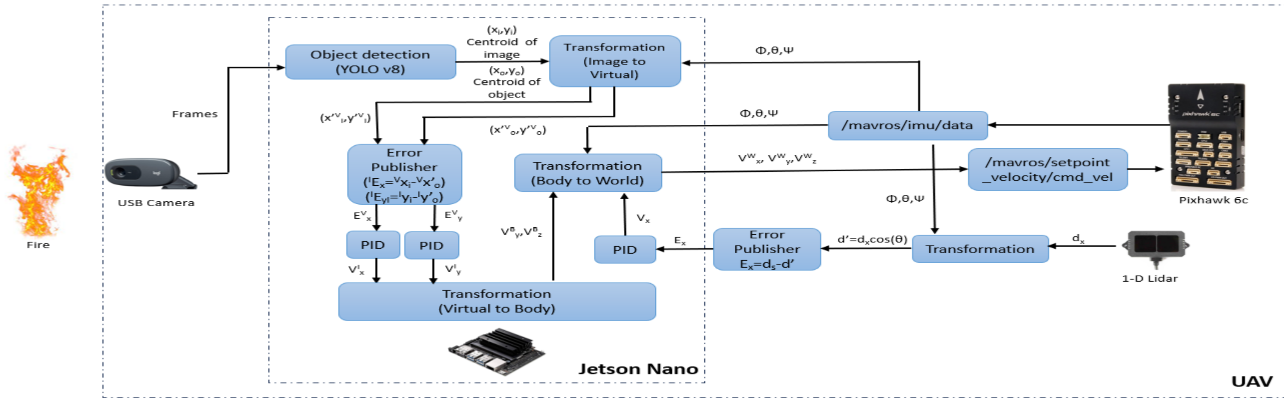

Fig. 3. Control algorithm.

4) <size> Consist of three fields <width>, <height>, <depth> denoting the width, height and channel number of image.
5) <name> The class of object. Our class "Fire".
6) <bndbox> Indicates the position of the object in the image.

We mixed our custom dataset and dataset of fire accident and trained the object detection model A total of 1342 images from different sources were used to train the object detection model. Samples of the dataset are available in our github repository[16]. To increase model robustness and generalisation we have used multiple augmentations [17] like 'Color Jitter', 'Random Brightness', 'Random Sharpen', which introduces variation in color, contrast, brightness, sharpness. To add a little blur and smokiness we have used 'Random Fog' which introduces blurring with a random fog coefficient and makes it more robust to smokey conditions. We have reduced the model filter size by $1/4^{\text{th}}$ at each layer, which reduces the number of parameters and, in turn, reduces the model size without a significant decrease in performance, resulting in faster inference time. The model was also converted from pytorch (.pth) to (.onnx) for faster inference and cross platform compatibility. In light of the inherent computational time required for frame processing, which resulted in a marginal temporal delay. To enhance the system's real-time performance, we adopted a frame-skip strategy, selectively omitting every second frame in the input sequence, thereby feeding the model with every third frame. This strategic adjustment improved the real-time responsiveness of the system while simultaneously ensuring the retention of critical data.

C. Communication

The Jetson Nano is the central hub for camera, Lidar, and Pixhawk integration. The Lidar employs UART communication, facilitated by a USB to TTL converter and a Python script for data access. The camera interfaces directly with the Jetson Nano through USB, utilizing OpenCV for frame retrieval and forwarding to the object detection model. Pixhawk, responsible for actuator and sensor coordination, communicates via ROS using the MAVROS package from PX4. This enables seamless topic-based communication with sensors and actuators. Communication between sensors, onboard computer and pixhawk is shown in Fig. 4.

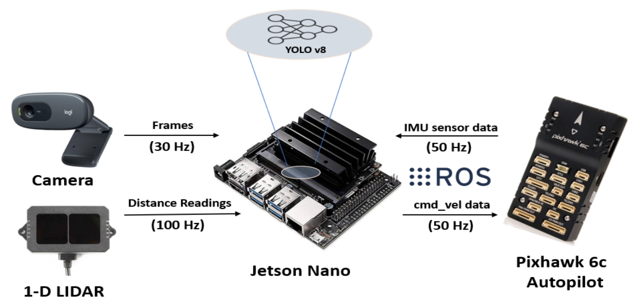

Fig. 4. System overview of sensing and computing components.

D. Vision based Control Strategy

Subsequent to the fire detection model's identification of fire within the image frame, we undertake a series of geometric operations. In cases where multiple bounding boxes are detected, our methodology prioritizes the box with the largest area for further analyses. Considering that we have prior knowledge of the image frame's dimensions, we can determine the centroid of the image, denoted as (x^I,y^I). Furthermore, following the acquisition of bounding box coordinates, we proceed to calculate the centroid of the bounding box, represented as (x^O,y^O). The convergence of centroids of image and bounding box become imperative for accurately aligning the UAV with fire. The lateral motion in the y_B and the altitude control in z_B are intricately guided by the calculated errors between the two centroids. Transformation between multiple entities is shown in Fig. 5.

The longitudnal motion, specifically in the x_B direction toward the fire, is modulated by the lateral error. This error, in turn, is derived from the contrast between the distance determined through LiDAR measurements and the experimentally ascertained spray distance. We applied transformation shown in Fig 6(a), using the current pose angles on the LiDAR data

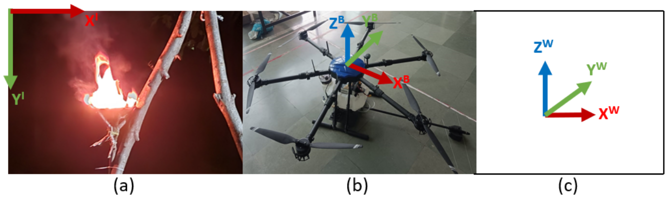

Fig. 5. (a) Image frame (b) Body frame (c) World frame.

to correct the error generated due to the pitching of drone for forward motion.

Similar problem will happen during motion in y_B leading to jerky motion of the UAV due to incorrect error values fed to the PID controller. Due to this tilt in the image plane we won't get the actual coordinates of fire, leading to jerky motion of the UAV due to incorrect error values fed to the PID controller. This problem is solved using the transformations of coordinates of the fire detected in actual frame on the initial horizontal image frame as shown in Fig 6(b).

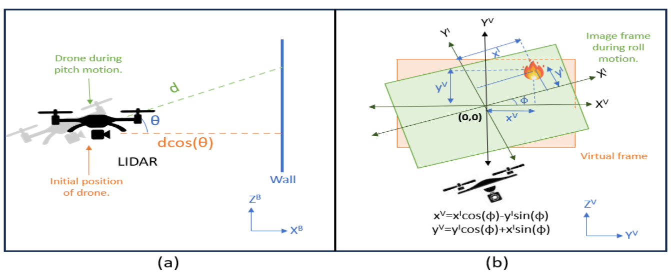

Fig. 6. Transformation from Body frame to Virtual frame.

$$x^V = x^I cos(\phi) - y^I sin(\phi) \tag{1}$$

$$y^V = y^I cos(\phi) + x^I sin(\phi) \tag{2}$$

Further, simple moving error technique is used to mitigate abrupt and unwanted fluctuations within the data using unweighted mean of ten previous data points. The number ten was determined experimentally. This ensures smoothed error value, facilitating precise and responsive control of UAV.

$$e_{n+1}(t) = \frac{1}{n} \sum_{i=n-9}^{n} e_i(t) \tag{3}$$

The refined error data is subsequently input into three distinct PIDs for controlling v_x, v_y, v_z.

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} \tag{4}$$

The optimal K_p, K_i, K_d are achieved by Ziegler-Nichols method.

E. Spraying Strategy.

Following the precise alignment of the UAV with the fire source, spraying subassembly will receive a signal from AUX PWM output of pixhawk 6c. This signal actuates the commencement of the spraying operation. We can continuously monitors the area encompassed by the bounding box surrounding the fire source, this area serves as a critical indicator of the fire's extinguishment status. Currently integration of pump subassembly with the UAV has been tested manually.

III. Experimental results and analsys

A. Object Detection results

In order to confirm the effectiveness of our object detection model, we subject it to rigorous testing using videos captured by our UAV under varying luminance conditions. Fig. 7. showcases a selection of samples from these test scenarios.

Fig. 7. Validation results of the model in different luminance conditions.

The comparison of original error, as obtained from the object detection node, and the subsequent application of the moving average technique, is graphically depicted in the Fig 8.

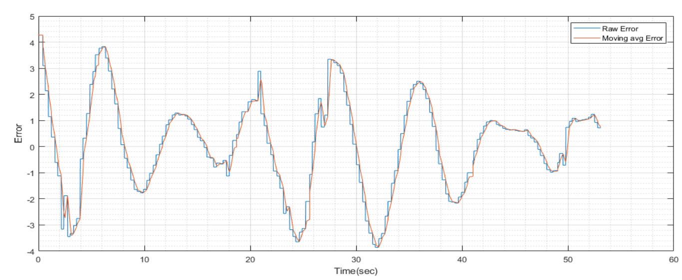

Fig. 8. Comparison of raw error and moving avg error wrt time.

B. Vision based control of the UAV.

To validate a vision strategy and control strategy for UAV motion control, multiple outdoor experiment has been performed. The real-time path of the UAV from one of the experiment as it aligns with the fire is demonstrated in Fig. 9.

v_x^B, v_y^B and v_z^B and their corresponding error are recorded during the experiment as show in the Fig. 10.

C. Spraying after integration with the UAV.

The evaluation of the spraying subassembly's performance was carried out manually post its integration with the UAV platform. Two BLDC motors are connected parallely for increased flow rate. Current flow rate is 15 L/min. The range of the spray is 10 m. Flight time of the UAV is 12 min with 10 kg payload. Spraying is shown in Fig. 11.

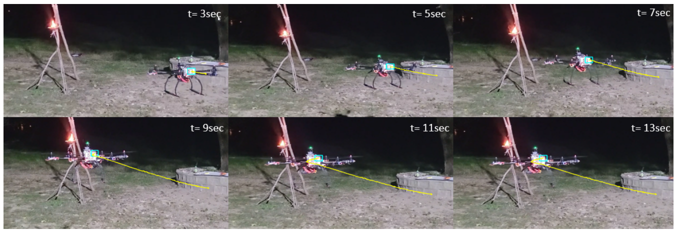

Fig. 9. Trajectory of the UAV while correcting the error in $y^B z^B$ plane.

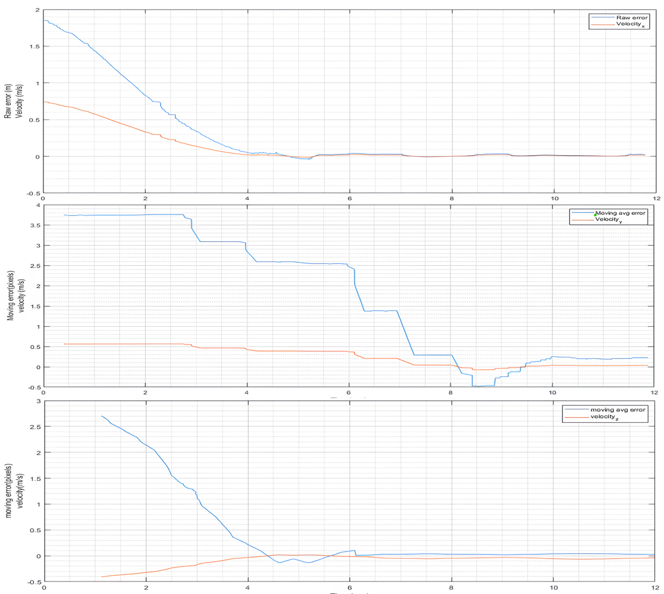

Fig. 10. Error and velocity of UAV in x,y,z directions.

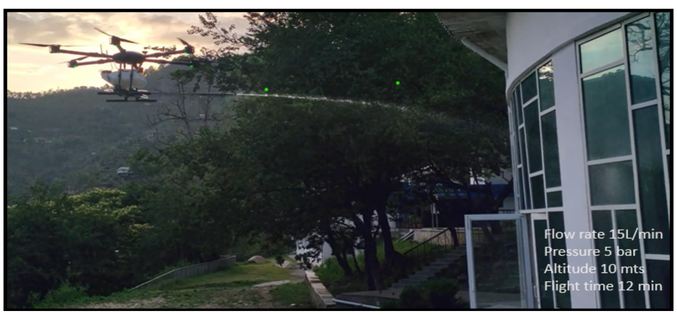

Fig. 11. Manual spraying on a building by UAV.

IV. Conclusion

This work develops a method for the vision-based localization and control of a UAV in firefighting situations. Detection of fire is done using a deep learning technique. Precise 3-D cartesian coordinates of the centroid of fire is obtained using onboard sensors, contributing to the localization of fire. A PID based control strategy is developed for continuous tracking of fire and also maintaining a particular distance from the fire for ensuring accurate spraying. Additionally, the performance of the spraying subassembly has been independently verified. In our forthcoming research endeavors, we aim to delve into the assessment of fire intensity and prioritize multiple identified fires accordingly. Additionally, we will estimate the requisite time for extinguishment and required quantity of extinguishing material. Subsequent to these analyses, we intend to enhance and refine our control strategy in alignment with the acquired insights.

References

[1] I. Today, "Major fire breaks out in high-rise building in mumbai," 2023, https://www.indiatoday.in/cities/mumbai/story/mumbai-news-fire-in-highrise-residential-building-dadar-no-casualties-2326949-2023-01-27.

[2] I. Today., "8 dead as massive fire engulfs 7-storey building in mumbai." https://www.indiatoday.in/cities/mumbai/story/mumbai-goregaon-fire-death-toll-blaze-fire-fighting-operations-2445241-2023-10-06, 2023.

[3] O. Alon, S. Rabinovich, C. Fyodorov, and J. R. Cauchard, "Drones in firefighting: A user-centered design perspective," in *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*, ser. MobileHCI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3447526.3472030

[4] S. X2, "Autonomous monitoring drone," 2023, https://www.skydio.com/skydio-x2.

[5] W. firefighting UAS, "Wk-1900 firefighting drone," 2023, https://en.walkera.com/index.php?id=wk1900.

[6] K. Wang, Y. Yuan, M. Chen, Z. Lou, Z. Zhu, and R. Li, "A study of fire drone extinguishing system in high-rise buildings," *Fire*, vol. 5, no. 3, 2022.

[7] X. Liang, H. Wang, Y.-H. Liu, W. Chen, and Z. Jing, "Image-based position control of mobile robots with a completely unknown fixed camera," *IEEE Transactions on Automatic Control*, vol. 63, no. 9, pp. 3016–3023, 2018.

[8] J. Quenzel, M. Splietker, D. Pavlichenko, D. Schleich, C. Lenz, M. Schwarz, M. Schreiber, M. Beul, and S. Behnke, "Autonomous fire fighting with a uav-ugv team at mbzirc 2020," in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2021, pp. 934–941.

[9] A. Q. Nguyen, H. T. Nguyen, V. C. Tran, H. X. Pham, and J. Pestana, "A visual real-time fire detection using single shot multibox detector for uav-based fire surveillance," in *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*, 2021, pp. 338–343.

[10] R. Arora, M. Khosravi, S. Enayati, and H. Pishro-Nik, "Deployment of a uav-based fire detection system," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1–6.

[11] D. Burschka and G. Hager, "Vision-based control of mobile robots," in *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164)*, vol. 2, 2001, pp. 1707–1713 vol.2.

[12] Z. He, R. Iyer, and P. Chandler, "Vision-based uav flight control and obstacle avoidance," in *2006 American Control Conference*, 2006, pp. 5 pp.–.

[13] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh, "Deep learning vs. traditional computer vision," in *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1*. Springer, 2020, pp. 128–144.

[14] J. Terven and D.-M. Cordova-Esparza, "A comprehensive review of yolo: From yolov1 to yolov8 and beyond," 04 2023.

[15] ONNX, "Dataset available for fire detection." https://www.kaggle.com/datasets/phylake1337/fire-dataset, 2023.

[16] A. Mehra, "Custom dataset for fire in different luminance conditions," 2023, https://github.com/ankitmehra31/Firefighting_data.git.

[17] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.