An Early Forest Fire Detection System Based On Drone and Aerial Sensing

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—This papers presents a drone based early forest fire detection system. Based on multiple areial sensors, thermal images, RGB images and distances of the early forest fire point can be captured from the air. To fully use the captured data for forest fire detecting and confirming, both deep learning based and traditional computer vision algorithms are employed data. Due to the different complexity and computational demand of submodules in this system, the onboard computer and ground station computer works collaboratively as designed to achieve higher efficiency. Integrated using sensor data and implementing a two stage approach of potential early forest fire detection and confirmation strategy, the proposed system achieves a relatively low false alarm rate and has good robustness in the outdoor early forest fire detecting experiments.**

I. Introduction

Forest fire has significant effects on the ecosystem, global and human beings. Directly, the forest fire reduces the forest density, burns the vegetation, and damages the animal resources. However, potential and side effects generated by the forest fire can produce long-term and continuous damages to the environment and climate: the forest fire will lower organic matters and water in the soils, making the ground surface acid hardly for vegetation restoration; on the other hand, burning destroys the forest as the "Lung of Earth" and emits huge amounts of greenhouse gases into the atmosphere, which deteriorate the climate change. The forest fire now gradually becomes a global challenge: as of 9 March 2020, the fires induced by the drought and hot weather in Austria burnt an estimated 18.6 million hectares (46 million acres; 186,000 square kilometres; 72,000 square miles), destroyed over 5,900 buildings (including 2,779 homes) and killed at least 34 people [1]; in China and United States, however, 190kha and 830kha of land has burned so far in 2021 separately [2]. Early forest fire detection and alarm, as a result, can minimize the loss by prevent the forest fire turning into a disaster during the incipient period.

Temporally, forest fire development can be simply divided into 6 phases: (1) incipient phase, (2) growth phase (pre-flashover), (3) flashover, (4) fully developed phase (post-flashover), (5) decay phase, and (6) extinction [3]. In the incipient period of the forest fire, there usually no strong smoke or considerable flames, and the ignited zone could be very small, which may bring about difficulties for the wide-range but rough detection methods.

In the early days of the forest fire monitoring, human observers and watchtowers are employed to acquire the spatio-temporal information of the forest fire. Apparently, watch-towers lacks the flexibility and wardens potentially suffer danger [4]. With the development of electronic and embedded systems, forest fire can be monitored by the pre-placed sensors, including the temperature sensors, smoke detector, etc. These pre-placed sensors can also network with the IoT (Internet of Things) [5].But this solution heavily depends on the communication capability and the power supplement [6]. Implementation of these networks could also be unfeasible when monitoring the primeval forest. Other approach is online monitoring the forest fire using remote sensing. However, remote sensing-based methods also have backwards: satellite images and data usually have wider range but relatively lower resolution [7]; Manned or manually-operated aircraft costs higher human labour. [4]

Drones, on the other hand, have multiple advantages in terms of the early forest fire detection: with onboard computer, decision making, mission planning, motion controlling can be executed onboard automatically; multiple types of sensors can be mounted as the payload of the drones, such as RGB camera, infrared camera, laser rangefinder and multi-spectral camera, etc, to get various information of the forest fire; Due to the high flexibility, drones are less likely to suffer connection failure or the failure of power supplement.

In this work, we present an online early forest fire detection system based on the popular drone platform. Multiple types of sensors are employed to detect multiple features of the fire flame and smoke. The hardware and system architecture are illustrated in Section III. Both deep convolutional neural networks(CNN) and traditional computer vision algorithms are implemented to process the vision images and infrared image, which will be presented in Section IV. The outdoor experiment results are given in Section V. Finally, Section VI summarizes this paper and prospects the future work. For the benefit of community, we make our source code public. Experiments demonstration and code can be found on github.[1]

[1] https://github.com/lee-shun/forest_fire_detection_system

II. Related Works

The drone based platform for forest fire detection mission have advantages comparing with other methods mainly because of the flexibility of the drone as well as multiple types of the payload sensors. Multiple types of sensors have been applied to the forest fire detection mission.

A. RGB Images

RGB camera, due to its low price and intuitive results, is extensively used in forest fire detection mission. Methods and algorithms in the computer vision field can be directly implemented in the forest fire images processing tasks: In 2015 and 2017, Yuan et al. using the RGB, HSV and HSI color space and Ostu segmentation [8] as well as the optical flow [4] [9] to detect, segment and tracking the forest fire flame. In 2019, they implemented the fuzzy logic and EKF (Extended Kalman Filter) to segment the forest fire smoke. [10]. In 2020, Vasconcelos proposed 7 segmentation rules based on the YCbCr and RGB color space to segment the forest fire pixel. [11]

Combining the human-defined features (color, texture, motion features, etc.) and traditional machine learning algorithm(random forest, support vector machine, etc.) usually achieves higher accuracy compared with threshold and rules-based methods. In 2018, Prema using the flame texture and YCbCr color space to define the forest fire features; then an ELM(Extreme Learning Machine) is trained to classify the extracted features [12]. Hossein et al. in 2020 finished multiple features and learning methods cross experiments and compared the metrics of these methods [13].

The deep learning based image processing methods represented by CNN (convolutional neural network) extract more types of features. These deep learning-based methods can be categorised into three main applications: classification, semantic and instance segmentation and object detection [14]. In 2019, Kyrkou et al. using DJI M100 and Odroid-XU4 onboard computer to implement several classification algorithm based on CNN including VGG16, Resnet50 and MobileNet, etc [15]. Barmpoutis et al. proposed a Faster-RCNN and spatial texture analysis based forest fire flame detection method [16]. In 2019, Kinaneva et al. use the DJI M210 as the drone platform and implement deep learning algorithm on it to realize the pre-mentioned three main tasks [17]. Several general and popular object detection and segmentation framework have been applied in the forest fire detection task: Jiao et al. using the DJI drones and Manifold onboard computer to implement YOLOv3 for fire detection [18] [19]. Barmpoutis in 2020 used 360-degree sensor region based method to reduce the false alarm generated by the DeepLabV3+ [20]. Zhang et al. in 2016 use the picture patches and CNN to classify and roughly locate the fire flame position [21].

B. Theraml Images

Thermal images are the direct information that can reflect the abnormal heat distribution of the forest fire. Each pixel in theraml images stands for the temperature of that corresponding point.

Dai et al. use the Gabor and fractal feature combined with simple linear iterative clustering (SLIC) method to segment the smoke from the thermal images [22]. In 2021, Moffatt et al. proposed a DJI M600 drone based forest fire detction, localization and supression system, in which thermal images is used to segment fire location [23].Beside the value of each pixel in the thermal images, motion features of the forest fire can also be used: in 2017, Yuan et al. combined the optical flow and Otsu segmentation to segment and track the fire using theraml images [24]. Sadi et al. using the homography matrix without the object depth to alignment the RGB and thermal images [25].

III. Drone Based Early Forest Fire Detection Platform

A. System Architecture

1) Hardware: For the onboard sensors, the ZenMuse H20T camera is chosen: it embedded wide camera, zoom camera, infrared camera and laser rangefinder in on payload, which significantly simplifies the payload mounting and improves the system reliability. For the experiment convenient, DJI M300 Drone is selected as the platform. The chosen hardware is shown figure 1.

Fig. 1. ZenMuse H20T camera (left) and DJI M300 drone (right)

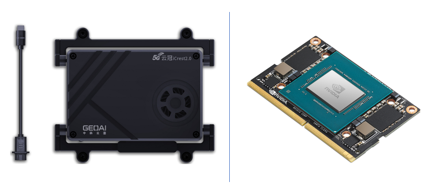


Fig. 2. iCrest2.0 onboard computer (left) and NVIDIA Jetson Xavier NX (right)

Onboard computer is the real-time high lever mission control, image processing and motion control platform. The platform should equip with GPU(s) intending to accelerate the inference of deep learning based image processing. Thus, the ICrest 2.0 onboard computer provided by GEOAI and NVIDIA is chosen for this work; this computer is based on the NVIDIA Jetson Xavier NX, “bringing up to 21 TOPs accelerated computing delivers the horsepower to run modern neural networks in parallel and process data from multiple high-resolution sensors.” With the development of GEOAI, it has the ability to communicate with DJI M300 and H20T

Camera online. Ground station computer should have the ability to monitor DJI M300 mission states and to accelerate a more complex image segmentation algorithm. Nowadays laptop can do this job; in the outdoor experiment, we use Lenovo IdeaPad 15 with Intel i5-6300HQ CPU and NVIDIA GeForce GTX 950M GPU. The outdoor experiment details will be introduced in Section V.

The proposed early forest fire detection system is not tightly coupled with the hardware platform, which mainly benefits from the hardware abstraction of ROS(Robot Operating System). Theoretically, proposed algorithm and software can be embedded in any platforms, sensors with the similar functions as we use.

2) Software: Software is developed based on the ROS (Robot Operating System) and DJI Onboard SDK (software development kit), which provides communication between our software applications and the hardware abstraction. Different applications are designed to run on different hardware as the following figure 3 showing:

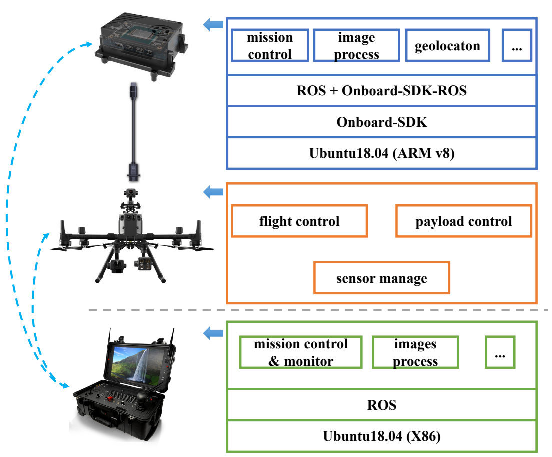


Fig. 3. proposed system software

Trough the APIs (application programming interface) provided by the DJI onboard SDK, our applications running on iCrest 2.0 can access the sensor and autopilot data and send control command to the flight control module and payload control module(such as gimbal driver, etc.) via wire connection. A task manager and image segmentation application is running on the ground station computer simultaneously. The former application is designed to monitor and control the whole detection mission; the latter is an image segmentation application as a “completion” of the onboard computer.

For the benefit of the real-time application, TensorRT developed by NVIDIA is used for optimizing and accelerating the deep learning-based algorithms inference. According to the Pytroch and TensorRT implementation website[1], resnet18 can achieve around 7 times faster after using the TensorRT on NVIDIA Xavier NX.

[1] https://github.com/NVIDIA-AI-IOT/torch2trt

B. System Functions

For the early forest fire detection mission, the proposed system performs the following functions:

- regular patrol and surveillance over the given region of the forest.
- on-board potential early forest fire detection.
- online potential early forest fire region confirmation.
- early forest fire source geolocation.
- early forest fire alarm.

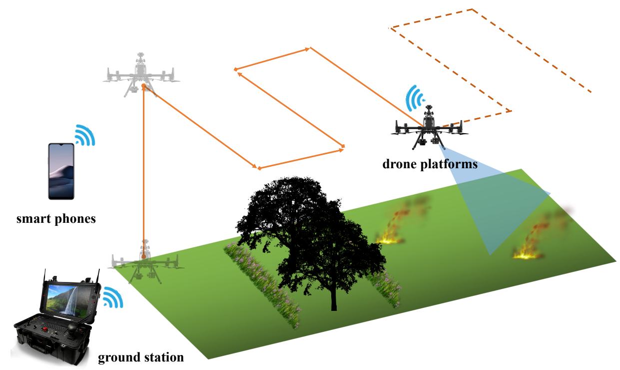


Fig. 4. typical early forest fire detection mission scenario

As shown in figure 4. The early forest fire surveillance and detection mission started with pre-defined path. During this stage, drone will take off and fly along the zigzag path, making sure that each part of the given zone of the forest can be covered. The zigzag path planning should takes the drones power, the region shape and the sensor effectual working distance. Multiple drones can patrol over different forest region according to the mission distribution. Aerial images captured from RGB zoom camera is processed by the Resnet18 based image classification algorithm, after which the current image is classified into 3 classes: normal, smoke and fire; simultaneously, traditional computer vision based segmentation algorithm marks the potential fire source with captured thermal images.

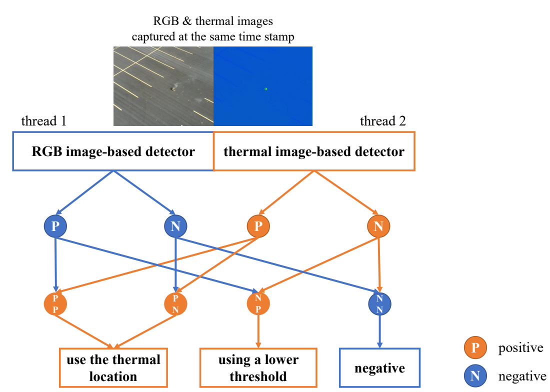


Fig. 5. confirmation request generating process

If the classification algorithm or the thermal image-based algorithm detect the potential forest fire, a confirmation request will be triggered. Otherwise, the drone(s) will keep flying along the pre-defined path until the mission endless or the power alarming triggers returning request.

As shown in figure 5, RGB and thermal images captured at the same time stamp will be fed into the 2 types of detectors: if they both have negative results, no event will be triggered; if both of the positive results or only thermal image-based detector has positive result, the confirmation will be triggered and the gimbal will be rotated; if only the RGB image-based detector predicts positive, a lower threshold will be used to segment the thermal image; At this time, if thermal image-based detector gives positive then the confirmation request will be triggered, otherwise, mark this frame as negative. The segmentation and classification algorithms will be introduced in detail in Section IV.

Once the confirmation is triggered, the drone(s) will interrupt the patrolling mission and hover. Then a gimbal controlling algorithm will rotate the camera facing the potential forest fire point according to the segmentation on the thermal images. The gimbal controlling method will also be presented in Section IV. At the meanwhile, zoom RGB camera on H20T will also be set to "5x" in order to capture the potential forest fire point in detail, as shown in figure 6:

(a) zoom camera at "2×" (b) zoom camera at "5×"

Fig. 6. DJI H20T zoom camera

After that, a Unet-based early forest fire flame and smoke segmentation algorithm running on the ground station will be trigged by drones automatically. Ground station computer, controlling the drones, has higher computing power to run more complex neural networks. The RGB images captured from the H20T will be sent to the ground station by wireless image transmission. If the confirmation is established, the early forest fire region will be located and then the forest fire alarm and fire point geolocation with detection results will be sent to other termination such as smart phones, for the forest rangers or firefighters.

In summary, the proposed early forest fire detection system takes the hardware and sensors limitation, algorithms false alarm, and computation power distribution into consideration, achieving relatively good robustness.

IV. Vision and Thermal Images Based Early Forest Fire Detection Algorithms

A. Resnet18-based Early Forest Fire Classification

Resnet18 is used for the onboard early forest fire classification, through which the RGB images captured by the H20T zoom camera are classified into normal, potential fire and potential smoke.

Resnet [26] series use the identity mapping to address the vanishing gradient and exploding gradient problem generated by the layers becoming deep. In figure 7, green block shows the basic component in the Resnet: layers among the residual block learns residual value instead of the original value, which makes the network more robust while training. Although Resnet has many variations, they can be separated into 3 parts: input, output and middle convolutional module. The following right figure 7 shows the Resnet18 architecture:

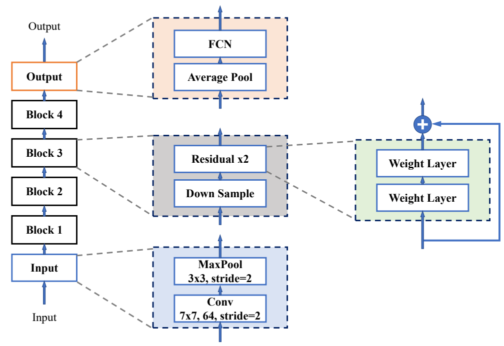


Fig. 7. Resnet18 architecture

Input module consists of 7×7 convolutional layer and a 3×3 maxpooling layer; the middle convolutional layer consists of 4 similar residual blocks. The features map channel accumulates from 64 to 512; finally, after an average pooling and fully connected layer, feature channel changes from 512 to 3, according to 3 classes of an image.

B. Unet-based Early Forest Fire smoke Segmentation

Unet(U-net) [27] structure is widely used in image segmentation, reconstruction and GAN(Generative Adversarial Network). In proposed system, Unet is running on the ground station computer for the early forest fire flame and smoke segmentation.

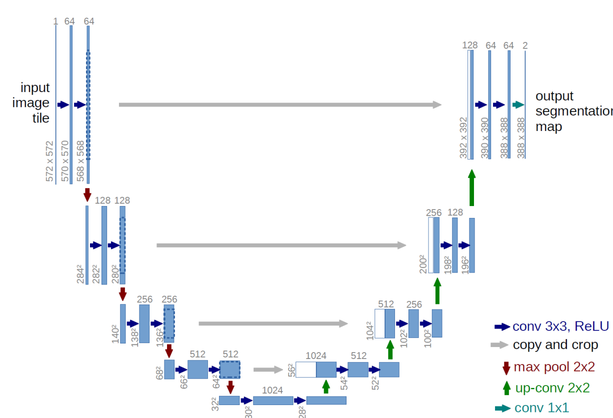


Fig. 8. Unet architecture [27]

The network architecture is as shown in the figure above, which can be described as consisting of a contraction path (left) and an expansion path(right). The contraction path is the same as the traditional convolutional network. It consists of an unpadded convolution with a convolution kernel size of 3×3. After each convolution, the features passes through a ReLU function, and a maximum pool with a size of 2×2, stride of 2. This maximum pooling is the process of downsampling. After downsampling, the number of channels are doubled.

The expansion path is up-convolution with size of 2×2. , the output channels of the upper convolution are half of the original, and then connected in series with the corresponding feature map (after cropping) to obtain channels of the same size as the original, and then pass through two sizes of 3×3 convolution and ReLU function. The corresponding cropping feature map is necessary, because there will be loss of boundary pixels in the process of our convolution. In the last layer, the desired target type is obtained through convolution with a convolution kernel size of 1×1. In this network, there are 23 convolutional layers. The network adopts the common Encoder-Decoder structure, and adds to the original structure the operation of directly intercepting information from the encoder and placing it in the decoder. This operation can effectively retain the edge detail information in the original image and prevent excessive edges.

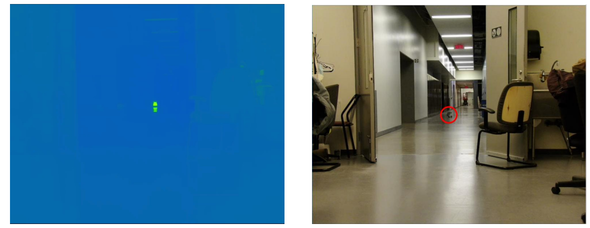
Compared with traditional fully convolutional networks, Unet can achieve relative acceptable precision from less training data.

C. Threshold and Sliding Window-based Thermal Image Segmentation

1) HSV color space: The thermal images captured by the H20T represents the heat radiation at each pixel. Inspired by the segmentation method proposed in [8], we use the HSV color space and threshold to segment the potential heat source; after that, morphological operations are used to clear the noise. Finally, the we use the sliding window to find highest potential heat in every patch.

Different from RGB (Red, Green, Blue) color space, which cannot separate the colors and brightness, HSV color space use the hue, saturation and value to represent a color. Only hue channel is used for describing the color, the other two channels describe the purity and brightness.

In the proposed system, we use the H20T "north pole isothermal" mode to generate the thermal images, as shown in figure 9. Under this mode, with temperature going up, the color of this point will go from blue to red; the brightness of this point will also go up. Please note that thermal images generated directly by H20T are still in the format of RGB.



(a) thermal images generated by the H20T "north pole isothermal" mode

(b) original RGB image(red circle: kettle)

Fig. 9. thermal and RGB images captured by H20T

Using HSV color space makes it possible to find a threshold for three channels, especially for color channel. Conversion between RGB and HSV color space can be defined by equation 1∼4.

$$\begin{array}{lll} R' = R/255 & C_{max} & = max(R', G', B') \\ G' = G/255 & C_{min} & = min(R', G', B') \\ B' = B/255 & \Delta & = C_{max} - C_{min} \end{array} \tag{1}$$

$$H = \begin{cases} 0 & \Delta = 0 \\ 60 \times (\frac{G'-B'}{\Delta} + 0) & C_{max} = R' \\ 60 \times (\frac{B'-R'}{\Delta} + 2) & C_{max} = G' \\ 60 \times (\frac{R'-G'}{\Delta} + 4) & C_{max} = B' \end{cases} \tag{2}$$

$$S = \begin{cases} 0 & C_{max} = 0 \\ \frac{\Delta}{C_{max}} & C_{max} \neq 0 \end{cases} \tag{3}$$

$$V = C_{max} \tag{4}$$

Then a threshold is defined based on the experiment to segment the heat point and background:

$$P_{heat} = \begin{cases} 1 & T_l < H(x,y) < T_u \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

In equation 5, P is the rule to segment the heat point; $H(x, y)$ is the H value at (x, y) and T_l, T_u are the threshold lower and upper boundary.

After the segmentation, usually there will be burr-like and discrete points, around the heat point. The morphological operation can be used as the geometry filter to reduce the noise. With the opening operation (erosion first, then dilation), most of the noise can be suppressed.

2) Sliding Window: As shown in figure 10, the sliding window(in green) traversals every patch of the thermal image, and compare the heat pixels inside. The patch with highest pixel value will be marked as the heat point position on the image. For example, patch 1 is chosen as the location of the heat point in the following figure:

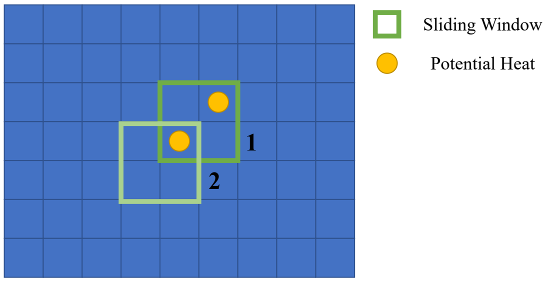


Fig. 10. sliding window generate the patches with heat point pixels

For the benefit of real-time property, the size of sliding window in our proposed system is fixed.

D. Laser Rangefinder-based RGB and Thermal Image Alignment

In order to find the heat point pixel location on the RGB image, the 2 types of images alignment is needed: Assume the coordinate value of a 3D point A in a camera-fixed axis(camera C) is p_A. Pixel value of this point A projected into the camera C pixel-fixed axis is q. The relationship between p and q can be represented as the following equation using the homogeneous coordinate system:

$$p = zK^{-1}q, p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, q = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{6}$$

where, K is the camera intrinsic matrix and z is the depth of the 3D point A.

The same 3D point A in infrared camera-fixed axis and RGB camera-fixed axis satisfy the following equation:

$$\begin{bmatrix} p_{rgb} \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{ir} \\ 1 \end{bmatrix} \tag{7}$$

or using the normal coordinate system:

$$p_{rgb} = Rp_{ir} + t \tag{8}$$

Substitute equation 6 into equation 8:

$$z_{rgb}K_{rgb}^{-1}q_{rgb} = z_{ir}RK_{ir}^{-1}q_{ir} + t \tag{9}$$

After simplifying, corresponding projection from infrared image to RGB image can be expressed by:

$$q_{rgb} = \frac{z_{ir}}{z_{rgb}}K_{rgb}RK_{ir}^{-1}q_{ir} + \frac{1}{z_{rgb}}K_{rgb}t \tag{10}$$

As shown in figure 1, the baseline between the RGB camera and Thermal camera is much shorter than the heat point distance. Thus, $\frac{z_{ir}}{z_{rgb}} \approx 1$. Then equation 10 can be simplified to:

$$q_{rgb} = K_{rgb}RK_{ir}^{-1}q_{ir} + \frac{1}{z}K_{rgb}t \tag{11}$$

where z is the depth of the early forest fire or heat point.

For ZenMuse H20T camera, z is the depth directly measured by the laser rangefinder; R, t K_{rgb} and K_{ir} is camera-based constant. To estimate the unknown alignment parameters, equation 11 can be rewritten to:

$$q_{rgb} = R'q_{ir} + \frac{1}{z}t', R' \in \mathbb{R}^{3\times3}, t' \in \mathbb{R}^{3\times1} \tag{12}$$

where $R' = K_{rgb}RK_{ir}^{-1}$ and $t' = K_{rgb}t$ denotes the unknown parameters of the 2 cameras, which could be estimated by the measurement with Least Squares Method.

E. Gimbal Control and Fire Point Geolocation

1) Gimbal Control Algorithm: As mention in Section III, a gimbal controller will rotate and zoom the H20T camera to focus on the potential early forest fire point with the location in the thermal image after confirmation request is trigged. A PD controller is designed for the rotating task, which is shown in the figure 11: The error is defined as the pixel difference between the desired position and the current position of the fire point. Then a PD controller is designed and tuned to generate the control command for the DJI gimbal controller.

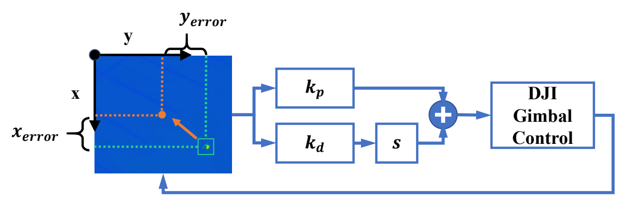


Fig. 11. gimbal control block digram

2) Fire Point Geolocation Algorithm: The geolocation of the early forest fire point can be estimated by the laser rangefinder and gimbal angle.

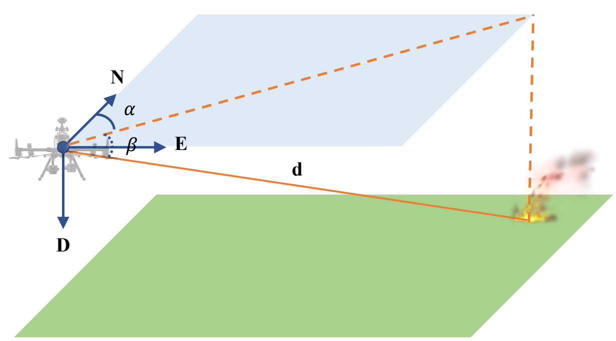


Fig. 12. geolocation of the early forest fire point

In the navigation coordinate system (NED), the relative location can be defined as:

$$\begin{cases} l_N = dcos\beta cos\alpha \\ l_E = dcos\beta sin\alpha \\ l_D = dsin\beta \end{cases} \tag{13}$$

V. EXPERIMENTS

A. Experiments Scenario

1) trajectory:

2) gimbal control:

B. Thermal and RGB Images Processing Results

1) Thermal Images Segmentation and Localization:

2) Resnet18-based Classification Results:

3) Unet-based Segmentation Results:

4) Thermal and RGB Images Alignment:

VI. CONCLUSION

This paper presents a early forest fire detection system based on the aerial sensor data. Both deep learning-based and traditional computer vision algorithms are employed to process the RGB and thermal images efficiently. The proposed strategies for decision making based on multiple aspects including algorithms, sensor data as well as hardware makes implementation more reliable and robust.

However, there are still several improvements can be done in the future work. Due to the incomprehensible of software implementation, laser rangefinder-based geolocation is not flexible enough to work on onboard computer automatically. To solve this, besides of the software improvements, vision odometry based geolocation method could estimate the forest fire point depth based on the images. Resnet18 can only output classification, which does not use the information provided by the RGB images completely even in the ”rough” detection stage. One possible solution is to implement the object detection algorithms to extract the potential forest fire location simultaneously, which can be used as an completion with thermal image-based location method.

REFERENCES

[1] Wikipedia, “2019–20 australian bushfire season,” 2021, https://en.wikipedia.org/wiki/20199320_Australian_bushfire_season. I

[2] GlobalForestWatch, “forest fire data globally,” 2021, https://www.globalforestwatch.org/about/. I

[3] C.-Y. Lee, C.-T. Lin, C.-T. Hong, M.-T. Su *et al.*, “Smoke detection using spatial and temporal analyses,” *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 6, pp. 1–11, 2012. I

[4] C. Yuan, Z. Liu, and Y. Zhang, “Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles,” *Journal of Intelligent & Robotic Systems*, vol. 88, no. 2, pp. 635–654, 2017. I, II-A

[5] S. Abdullah, S. Bertalan, S. Masar, A. Coskun, and I. Kale, “A wireless sensor network for early forest fire detection and monitoring as a decision factor in the context of a complex integrated emergency response system,” in *2017 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*. IEEE, 2017, pp. 1–5. I

[6] P. Dasari, G. K. J. Reddy, and A. Gudipalli, “Forest fire detection using wireless sensor networks.” *International Journal on Smart Sensing & Intelligent Systems*, vol. 13, no. 1, 2020. I

[7] V. Sherstjuk, M. Zharikova, and I. Sokol, “Forest fire-fighting monitoring system based on uav team and remote sensing,” in *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*. IEEE, 2018, pp. 663–668. I

[8] C. Yuan, Z. Liu, and Y. Zhang, “Uav-based forest fire detection and tracking using image processing techniques,” in *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2015, pp. 639–643. II-A, IV-C1

[9] C. Yuan, Z. Liu, and Y. Zhang, “Vision-based forest fire detection in aerial images for firefighting using uavs,” in *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2016, pp. 1200–1205. II-A

[10] C. Yuan, Z. Liu, and Y. Zhang, “Learning-based smoke detection for unmanned aerial vehicles applied to forest fire surveillance,” *Journal of Intelligent & Robotic Systems*, vol. 93, no. 1-2, pp. 337–349, 2019. II-A

[11] J. Vasconcelos Reinolds de Sousa and P. Vieira Gamboa, “Aerial forest fire detection and monitoring using a small uav,” *KnE Engineering*, vol. 5, no. 6, p. 242–256, Jun. 2020. II-A

[12] C. E. Prema, S. Vinsley, and S. Suresh, “Efficient flame detection based on static and dynamic texture analysis in forest fire detection,” *Fire technology*, vol. 54, no. 1, pp. 255–288, 2018. II-A

[13] F. A. Hossain, Y. M. Zhang, and M. A. Tonima, “Forest fire flame and smoke detection from uav-captured images using fire-specific color features and multi-color space local binary pattern,” *Journal of Unmanned Vehicle Systems*, vol. 8, no. 4, pp. 285–309, 2020. II-A

[14] L. P. Osco, J. M. Junior, A. P. M. Ramos, L. A. d. C. Jorge, S. N. Fatholahi, J. d. A. Silva, E. T. Matsubara, H. Pistori, W. N. Gonçalves, and J. Li, “A review on deep learning in uav remote sensing,” *arXiv preprint arXiv:2101.10861*, 2021. II-A

[15] C. Kyrkou and T. Theocharides, “Deep-learning-based aerial image classification for emergency response applications using unmanned aerial vehicles.” in *CVPR Workshops*, 2019, pp. 517–525. II-A

[16] P. Barmpoutis, K. Dimitropoulos, K. Kaza, and N. Grammalidis, “Fire detection from images using faster r-cnn and multidimensional texture analysis,” in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 8301–8305. II-A

[17] D. Kinaneva, G. Hristov, J. Raychev, and P. Zahariev, “Early forest fire detection using drones and artificial intelligence,” in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, 2019, pp. 1060–1065. II-A

[18] Z. Jiao, Y. Zhang, L. Mu, J. Xin, S. Jiao, H. Liu, and D. Liu, “A yolov3-based learning strategy for real-time uav-based forest fire detection,” in *2020 Chinese Control And Decision Conference (CCDC)*. IEEE, 2020, pp. 4963–4967. II-A

[19] Z. Jiao, Y. Zhang, J. Xin, L. Mu, Y. Yi, H. Liu, and D. Liu, “A deep learning based forest fire detection approach using uav and yolov3,” in *2019 1st International Conference on Industrial Artificial Intelligence (IAI)*. IEEE, 2019, pp. 1–5. II-A

[20] P. Barmpoutis, T. Stathaki, K. Dimitropoulos, and N. Grammalidis, “Early fire detection based on aerial 360-degree sensors, deep convolution neural networks and exploitation of fire dynamic textures,” *Remote Sensing*, vol. 12, no. 19, p. 3177, 2020. II-A

[21] Q. Zhang, J. Xu, L. Xu, and H. Guo, “Deep convolutional neural networks for forest fire detection,” in *Proceedings of the 2016 international forum on management, education and information technology application. Atlantis Press*, 2016. II-A

[22] M. Dai, P. Gao, M. Sha, and J. Tian, “Smoke detection in infrared images based on superpixel segmentation,” in *MIPPR 2019: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications*, vol. 11432. International Society for Optics and Photonics, 2020, p. 1143213. II-B

[23] A. Moffatt, N. Turcios, C. Edwards, A. Karnik, D. Kim, A. Kleinman, V. Nguyen, V. Ramos, E. Ranario, T. Sato *et al.*, “Collaboration between multiple uavs for fire detection and suppression,” in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2021, pp. 730–737. II-B

[24] C. Yuan, Z. Liu, and Y. Zhang, “Fire detection using infrared images for uav-based forest fire surveillance,” in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2017, pp. 567–572. II-B

[25] M. Sadi, Y. Zhang, W.-F. Xie, and F. A. Hossain, “Forest fire detection and localization using thermal and visual cameras,” in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2021, pp. 744–749. II-B

[26] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778. IV-A

[27] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241. IV-B, 8