



Proposal of DNN-based Wildfire Detection with UAV

By: Linhan Qiao

Supervisor: Dr. Youmin Zhang

Department of Mechanical, Industrial and Aerospace Engineering

Presented in Partial Fulfillment of the Requirements for the Degree of
Doctorate of Philosophy (Ph.D.) in Mechanical Engineering at

Concordia University

Montreal, Quebec, Canada

December 2021

©Linhan Qiao

Abstract

With the support of specific sensors, it is possible for unmanned aerial vehicle to be conveniently applied in wildfire detection, fire front tracking and prediction tasks. With the encouragement of deep neural network (DNN) technologies, this research focused on image processing and vision-based algorithms to design and develop a mixed air-ground wildfire discovering and monitoring system to support the early wildfire fighting. The proposal report of this research first provide the overview of the related works of this research, and stated the challenges. The detailed strategies to deploy the schemes are outlined after the description of objectives of this research. In specific, they are: DNN-based wildfire detection schemes on-board and on ground work station, DNN enhanced wildfire prediction, filtering-based scheme of wildfire tracking and additionally, DNN-based wildfire image restoration. Finally, parts of these schemes are verified and demonstrated by a series of simulation and experiment.

Keywords— Unmanned aerial vehicle (UAV); Deep neural network (DNN); Wildfire detection; Wildfire prediction; Fire front tracking; Wildfire image restoration

Contents

1 Introduction

Nowadays, to decrease wildfire caused economical loss and ensure human life safety, the concepts of wildfire management are being more and more accepted [?]. Unmanned aerial vehicles (UAV) is thought one of the most efficient equipment for wildfire fighting missions because of their high efficiency and low cost [?]. To keep the wildfire controllable, detecting wildfire early and predicting the spreading of wildfire are both needed. Sometimes, even the real-time situations need to be observed and the emergent fire points need to be tracked.

There are plenty of sensors could equip UAVs to help detecting wildfire during its early periods. Most researchers agree that detecting smoke plume is one of the efficient schemes since years ago [?], and the performance of such scheme is demonstrated [? ? ?]. However, as stated in [?], smoke detection is facing some challenges. During the early period of wildfire, the smoke only appears in small area of the whole image and diffusion of smoke is irregular. Detection of flame under vegetation cover is also limited. Traditional wildfire detection and alarming schemes might not be efficient enough. Under such situation, neural-network-based (NN-based) schemes are getting popular in dealing with such detection tasks because of their less emphasis on features extracting manually and priory experience of experts.

What NN-based schemes emphasis more are the structures designing of the network and tuning in training process. These networks could also be helpful in analyzing wildfire spreading [? ?], or tracking the fire front lines, to enhance the wildfire detection and help people separate the disaster levels and make decisions more efficiently.

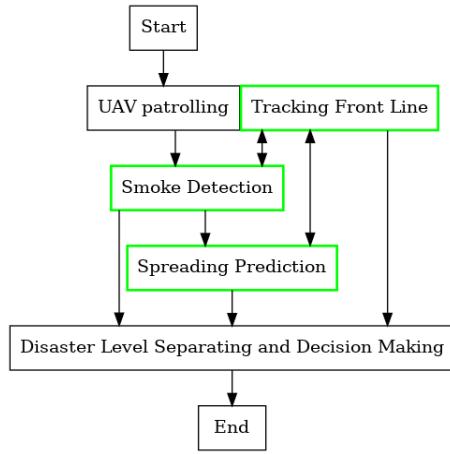


Figure 1: Main process of wildfire fighting using UAV

Encouraged by so many excellent NN-based wildfire fighting strategies in recent years, this research is trying to design and develop a NN-based system with support of UAV and ground work station for early wildfire detection, prediction, and tracking with UAV. Fig. ?? is the brief structure of the wildfire

detection and alarm system which is going to be build in this research. The main contributions could be induced in green boxed items, where they are detection, prediction and tracking. The reminder of the proposal of this research is structured as below:

In Section 2, there are statements of related works and motivation of this research. Section 3 is the description and analysis of objective of this report. Some concepts of the proposed and preparing methodologies are placed in Section 4. Conclusions and potential future works are outlined in Section 5. Timeline and some of research process are illustrated in appendix.

2 Related Works and Motivation

As analyzed in literature review before this report, traditional schemes for detecting wildfire early and analyzing wildfire actions are based on the features of wildfire smoke and flame. And, deep-neural-network-based (DNN-based) schemes focus more on the feature-learning processes in recent years. In this section, the state-of-art of NN-based wildfire detection and wildfire spreading predictions are studied.

2.1 NN-based Wildfire Detection Schemes

For wildfire detection and other recognition tasks, DNNs are higher recommended. Generally, deep networks ensured higher level features could be extracted and better generalized. Also, with the help of transfer learning, it is easier to build segmentation network models for wildfire detection. U-net, one typical fully connected network (FCN) is widely accepted and applied in wildfire detection [? ? ? ? ?]. There are also region-based convolutional network (R-CNN) [?] and recurrent neural network (RNN) [?] models could finish sequence-to-sequence analysis of wildfire data to detect smoke or flame.

However, in the most recent research, it is found that layers in these models are more computational complex. Because of the demand of early detection. The networks are often considered to be deployed on-board for UAV or UAVs. Saving the computing resource and simplifying the structure of networks are very challenging.

Under such situation, light network models are studied. In [?], depthwise seperable convolution and SqueezeNet are considered to build blocks called fire model to detect wildfire. Based on fire models, a light U-net is proposed [?] to share parameters and save computing resources. on the other hand, attention mechanism is demonstrated that they could keep more deeper information in the down sampling process and decreased false positive detection [?]. Also, it is found that attention layers is lower computational complexity by comparing with convolutional layer [?] in the research of Transformer. Vision transformer (ViT) model demonstrated that attention mechanism performed well in higher-resolution tasks [?].

However, the application of attention mechanism also has its challenges. For example, for sequence data set, attention could not capture the order information, which limited the attention to learn partial information. These researches motivated the design and implementation to optimize the encoder-decoder structured model or models combining with attention mechanism for wildfire detection to save more computational resource and perform more meticulous and accurate segmentation.

2.2 Wildfire Prediction Schemes

After detecting the wildfire, it is also important to understand the dynamics and actions of wildfire, which motivated the study of wildfire spreading models. For a long period, wildfire spreading prediction are mainly rely on two typical numerical model:

The fist one is based on physical modeling accounting for various physical and chemical phenomena. This kind of model contains computational-fluid-dynamics (CFD) [?] which could describe the spreading progress with more details. However, its performance on real fire still need more discussion because of high computational cost and the environmental complexity.

The second one is the mathematical model which based on empirical model of rate of spread (RoS). As illustrated in [?], mathematical growth prediction models is an acceptable description for wildfire spreading. Similar as fire area simulator (FARSITE) [?], a type consists of models based on continuous plane assumed the fire spreads in a continuous landscape and solving a system of partial differential equations is considered as the solution of these models[?]. A computationally faster type consists of the models based on grid [?]. Typically, there are Cellular Automata (CA) [?] and Bond-Percolation method. Compared with Bond-percolation method, CA solved its problem of describing the dynamics of the fire spread. There are also some other types of models, such as the fuzzy or neural based ones [? ?].

Generally, data-driven schemes could be applied for improving the accuracy of locating fire fronts and reduce the simulation uncertainty. In the review [?], Kalman filter, Lagrangian particle, Huygens principle and level-set method are mentioned to predict the fire front based on the real-time RoS measurement. The aforementioned schemes could be illustrated in Fig. ???. In recent years, it is found that data collected from geographic information system (GIS) and data augmentation with some NN models are also helpful in building these models. For example, the performance of FireCast of 2D-CNN [?] and LSTM [?] both encouraging.

However, these models often work for a relatively large time step (hours or days) [?]. The complex forest environment might change rapidly during the early fire period. To improve the accuracy and prediction speed need more investigation and demonstration. Such objectives motivated that the application of probabilistic models [?], attention mechanism [?], and some popular DNN models to be applied.

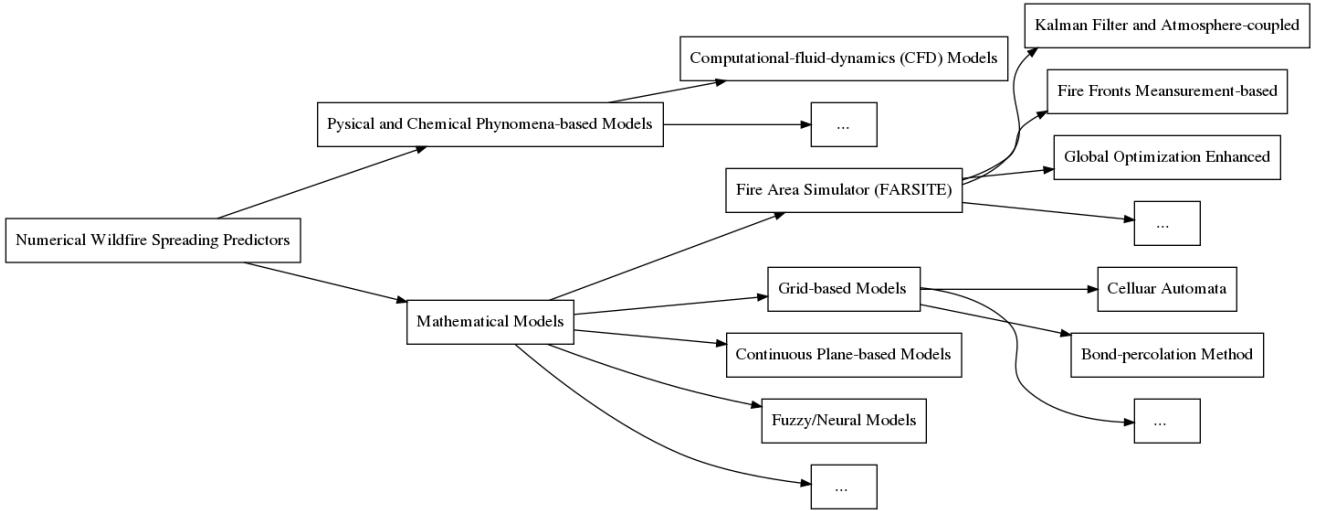


Figure 2: Main schemes for predicting the wildfire spread

2.3 Vision-based Wildfire Tracking Schemes

The prediction schemes mentioned above could be treated as a kind of auto-regressive processes. However, they are nearly to predict the wildfire scar, where the vegetation are already burnt, improving the accuracy of the fire front prediction is much more challenging for such indirect schemes by considering the wind-slope and other interacting factors. The accuracy and sensitiveness limitations are considered. Therefore, remote-sensing (equipped UAV) supported fire fronts tracking is thought about to ensure the predictions and monitoring of suspected wildfire areas.

Commonly, in forest fire management, visual object tracking (VOT) is the most efficient tracking scheme, VOT is studied in this research. VOT could be evaluated through OTB50 or OTB100 benchmarks [?]. The mainly applied VOT scheme consists of traditional generative model (GM), discriminative model (DM), and the DNN-based model getting hot in recent years. Commonly, GM build the objective model or extract the objective features at first, then it would search the similar features in the coming frames, and iterated to locate the tracking target. However, the background information and the randomness and multiple shape changing of target are not efficiently used when it is used in environment where it is lower resolution, brightness changing, which limited its accuracy in wildfire management. Compared with GM, DM performs better in mean frame per second (FPS) and mean precision. DM extracts and locates the target in local frame by comparing the different information of target and background. Related filters and some machine learning based classifiers are being applied as DM for VOT. Also, there is DNN-based model, which is different from GM and DM [?]. These main schemes could be illustrated in Fig. ??.

With the support of VOT, UAV or UAVs could track or avoid the smoke and flame area and give fire fighters more information in wildfire management tasks. The

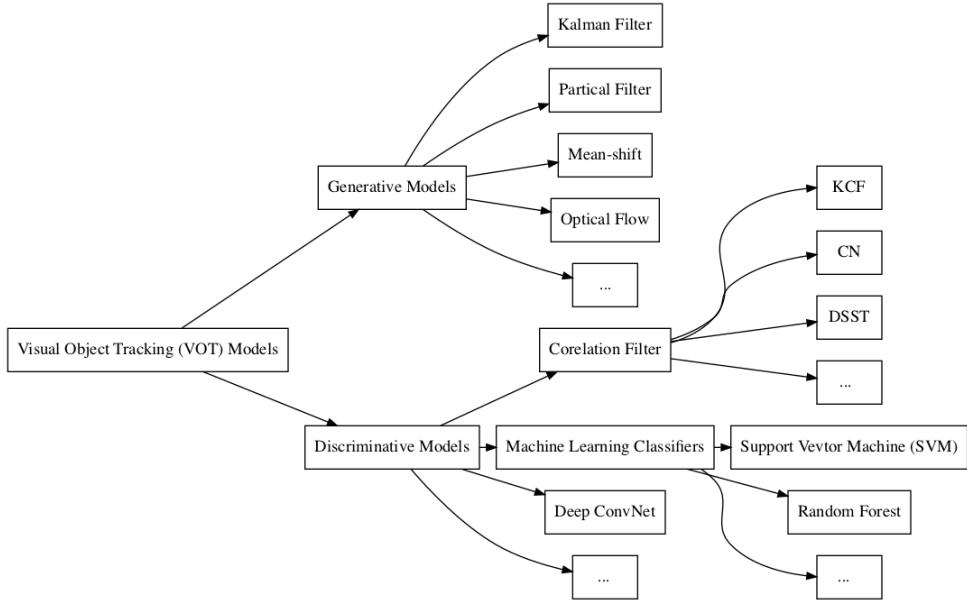


Figure 3: Main schemes could be applied for tracking fire fronts

3 Scope and Objectives

The mainly focus goals of this research contain two parts: The first part is to use DNN-based models to save computational resources when detecting and predicting the wildfire. The second part is to design the tracking scheme for UAV or UAVs to track the suspected burning area in forest. Therefore, the proposed research plan is organized around the following objectives:

- *Detection and segmentation objective:* Considering DNN model as an efficient wildfire segmentation scheme, design and develop the detection network structure so that the parameters could be used more efficient.
- *Deployment objective:* Design and develop light network structure so that it could be deployed on-board with UAV or UAVs, so that the path planning algorithm and detection scheme could work on-board with acceptable FPS at mean time.
- *Optimization objective:* Develop and test the attention mechanism in segmentation networks to share the parameters. Compared with convolutional layers, it is found that attention layers are lower computational complex.
- *Prediction objective:* Based on FARSITE and cellular model, build a wildfire spreading model which could predict the wildfire under impacts of wind and slope with acceptable accuracy.
- *Coverage and tracking objective:* Design the tracker for UAV which could track wildfire spreading in safe and efficient distance.

The proposed research in brief is expected to deploy the detection model on both UAV and ground work station. Considering the uncertain factors in forest environment, image restoration, prediction scheme, and UAV tracking should also be applied to support the detection. Because of the demand of early and accurate detection, DNNs is focused, as designed and developed for image restoration, detection, prediction and tracking. These strategies in this research and work are going to be demonstrated and verified through simulation and experiments through ground station computer and DJI M300 quad-copter equipped with H20T infrared camera.

4 Methodologies

4.1 U-net-based Wildfire Detection

Considering the rapidly shape changing and blurred edge of smoke, U-net is advocated in this research because of its fantastic performance in medical image processing fields. With the original image information or the restored images from UAV in forest, U-net could be deployed for wildfire smoke or flame segmentation. The U-net structure could be illustrated in Fig. ??.

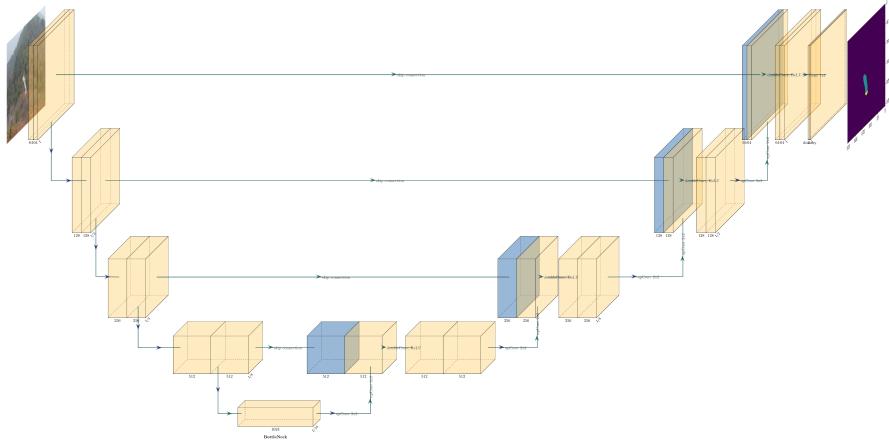


Figure 4: Brief structure of the U-net deployed on DJI M300 for the first time experiment

The structure shown in Fig. ?? is the one which based on ResNet18 and the first time experiment demonstrated that it could be deployed on-board with the support of tensor2trt to get FPS of 5 as dealing with original images. However, the segmentation accuracy still need to be improved and the size of this model could be decreased when using the concept of SqueezeNet. At ground work station, the deployment environment is not that serious as on-board, but the segmentation accuracy should be higher than on-board ones. Therefore, higher level residual networks could also be considered to build a U-net for wildfire smoke and flame segmentation. These apply-able network structures are illustrated in Table. ??.

Table 1: Residual network structures for building the U-net for wildfire detection

Layers	Output	18-layers	34-layers	50-layers	101-layers	152-layers
Conv1	112 × 112		kernel size = 7 × 7, out channel = 64, stride = 2, padding = 3			
Pool1	112 × 112		kernel size = 3 × 3, type = max pool, stride = 2			
Conv2x	56 × 56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
Conv4x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
Conv5x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$

4.2 Squeeze Model

To deploy the U-net on-board with less computational resource demands, the squeeze model is considered for developing the ResNet18-based U-net into a light model. Attention mechanism is also able to be developed to design attention gate at the skip connection parts of U-net to filter higher value features before up-sampling parts to increase the feature learning efficiency.

On the other hand, attention mechanism could not only be applied to develop the U-net for wildfire detection. It could also be helpful in building the NN-based prediction model for the fire fronts spreading. Such concept is also stated at the conclusion part of [?]. Therefore, the squeeze model and attention mechanism are going to be described as follow.

4.2.1 Squeeze Model

Fig. ?? shows the brief structure of the squeeze and unsqueeze model.

4.2.2 Attention Mechanism

4.3 GAN-based Wildfire Image Restoration

Considering the signal missing and disturbance as UAV capturing image information in wide forest environment. Generative adversarial network (GAN)-based image restoration is going to be designed. GAN could be separate into two parts: generator (G) and discriminator (D). The generating continues until the discriminator could not distinguish which is the target or which is the generated image, it is considered converged.

It should be noticed that, such strategy should only be applied in very serious forest environment and extreme situation. Because the GAN should be deployed on-board, where the computation resource distribution will be very challenging as other schemes are working to support the wildfire detection. And strictly, the restored information are not the real original information, there is risk that the restored image might misleading the segmentation of smoke edge. Commonly, the GAN structure could be illustrated in Fig. ???. It could be considered that, GAN is to balance the loss of G and D . Mean squared error (MSE)

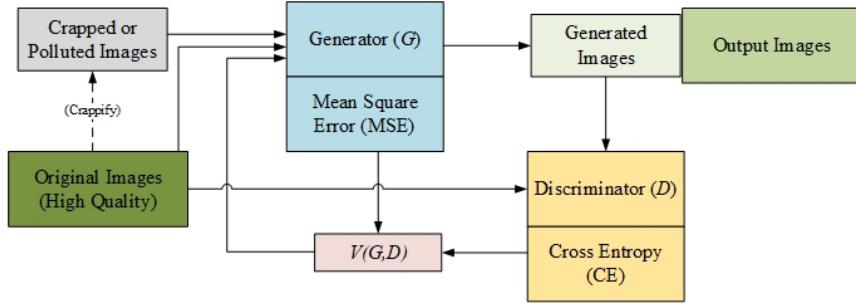


Figure 5: Brief GAN structure as training process

could be applied as the main loss function of GAN. In the training process, MSE is to compare the pixel values of generated image and high quality image, it can be written in a form of Eq. ??.

$$\text{MSE}(y, \hat{y}) = \frac{\sum_{i=0}^N (y_i - \hat{y}_i)^2}{N} \quad (1)$$

Value function $V(G, D)$ could be applied as Eq. ?? to describe the balance process of GAN.

$$\min_G \max_D V(G, D) = \mathbf{E}_{x \sim p_{data,x}}[\log D(x)] + \mathbf{E}_{z \sim p_{zz}}[\log(1 - D(G(z)))] \quad (2)$$

4.4 Cellular Model for Wildfire Prediction

4.5 VOT-based Wildfire Tracking

5 Timeline

The following Table ?? outlines the process for accomplishing the objectives of the PhD study.

6 Anticipated Significance of the Work

The merits of this proposed research can be reflected by an anticipated significant contribution to the concept of smart forestry, especially the UAV-based forest firefighting. It is an essential work to improve

Table 2: Timelines for PhD study

Main Tasks	2018-2019			2019-2020			2020-2021			2021-2022		
	W	S	F	W	S	F	W	S	F	W	S	F
Course Study	>	>	>									
Literature Review	>	>	>	>	>	>	>	>	>			
Comprehensive Exam										>		
ResNet for Wildfire Detection				>	>	>						
GAN for Image Restoration					>	>	>					
Attention Mechanism				>	>	>	>	>	>	>	>	>
Develop for On-board Deployment										>	>	
DNN Enhanced Cellular for Wildfire Prediction					>	>	>	>	>			
Vision of Tracking for Fire Fronts with UAV(UAVs)							>	>	>			
Publications	>	>	>	>			>	>	>	>	>	>

the wildfire detection and location accuracy with UAV during the early fire period. The main idea of this work is to save the limited computational resource on-board and decrease the false positive in wildfire discovering stage. The results of the proposed work would be published in journal and conference publications. Some preliminary results of the proposed research have been presented in the following conference:

7 Appendix: Process to Date

With respect to the presented timeline (Table ??) in Section. ??, this section provides a brief review of the process to date of the PhD thesis work. The works focused on the objectives of: wildfire image restoration, wildfire detection, and fire fronts prediction. They are separately stated as follow.

7.1 Wildfire Image Restoration

To develop the robustness of GAN for wildfire image restoration, a feature loss is proposed to improve the original confusion loss of U-net-based GAN. The simulation programms of the restoration are illustrated in Table ??.

Table 3: The simulation basement of proposed feature loss GAN for wildfire image restoration

Names	Illustrations
Platform	Google Cloud Platform (GCP)
Framework	Pytorch, Fastai
GPU	nvidia-tesla-t4
GAN Structure	ResNet-based U-net
Feature Loss Network	VGG16-based U-net
Dataset Resource	Google Images

The training process losses are shown in Fig. ???. There is an obvious jumping of validation loss as the entire U-net structure of GAN is unfreezed, which shows the jumping out of local minimum.

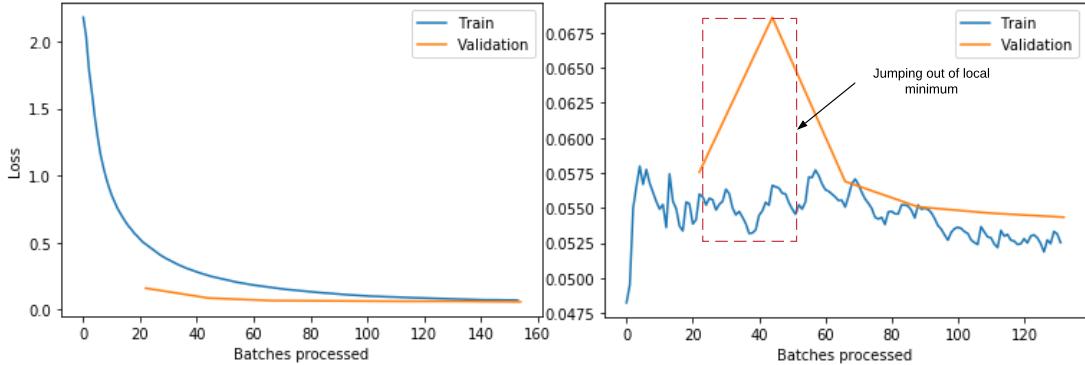


Figure 6: U-net training loss and validation loss changing. Left: Freeze encoder part of U-net; Right: Unfreezed encoder-decoder U-net. It seems over trained but jumped out of local minimum.

The final restoration performance is shown in Fig. ?? by comparing with the single VGG16-based GAN, the proposed model shows higher robustness.



Figure 7: The GAN restoration output. Left column: Crapped images, number marks are randomly added as disturbance at random location in images, higher number order means the image is randomly crapped more. Mid-left column: Restoration results of feature loss network. Mid-right column: Restoration results of proposed loss GAN. Right column: Original images (targets).

7.2 Wildfire Classification

The wildfire images classification is designed for UAV or UAVs to discriminate whether there suspected smoke or flame is contained in the image. This program is working before the segmentation on-board. Because the segmentation program might cause more computation and energy cost of on-board computer and battery. The designed classification contains ResNet34 based classifier and developed attention mechanism-based classifier. Their simulation basement are stated in Table. ??, and their classification results are shown as follow in Fig. ??.

Table 4: The simulation basement of proposed ResNet34-based and Attention mechanism-based wildfire classifiers

Names	ResNet34-based Classifier	Attention-based Classifier
Dataset Source	Google Images, Kaggle Dataset	
Platform	Personal Computer	
Framework	Pytorch, Fastai	
CPU	Intel-i7-4790K	
GPU	Nvidia-GTX-1060	Nvidia-GTX-1660
Pre-Training	False	ImageNet

7.3 Wildfire Segmentation

When the image is demonstrated as a wildfire image, it will be send into the segmentation program part so that the wildfire detection accuracy could be increased and the suspected fire point could be located more specific in the image.

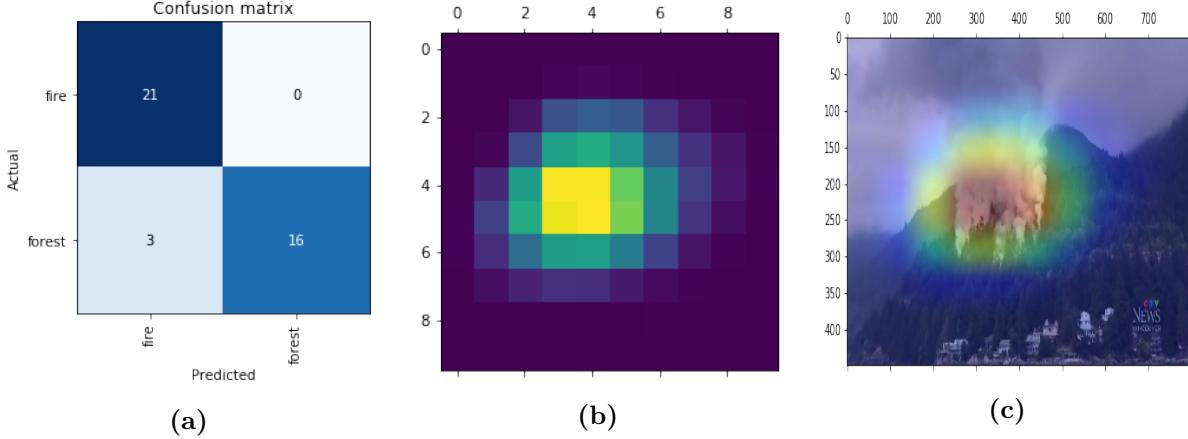


Figure 8: Classification results of designed classifiers. (a) The confusion matrix of ResNet34-based wildfire classification. (b) The result heat-map of Attention-based wildfire classification. (c) The mixed heat-map mask and original image.

The designed segmentation model on-board is based on pure U-net which has ResNet18 structure and the H20T infrared camera could be applied for directly detect the wildfire with appropriate threshold. The segmentation model on ground work station is based on ResNet34. For both the model on-board and on ground work station, attention mechanism and squeezed model should be designed to develop the model. The simulation and experience basement are separately illustrated in Table ??.

Table 5: The simulation basement of proposed on-board and ground segmentation model

Names	On-board Segmentation Model	Ground Segmentation Model
Dataset Source	Kaggle Dataset, UAV Captured Experience Dataset	
Platform	YunGuan on-board computer with DJI M300 UAV	Personal Computer
Framework	DJI OSDK, ROS, Pytorch	Pytorch, Fastai
CPU	NAN	Intel-i7-4790K
GPU	Nvidia-Jetson	Nvidia-GTX-1660
Model Structure	ResNet18	ResNet34
Pre-Training	ImageNet	ImageNet

As shown in Fig. ?? is the segmentation results of the proposed ResNet34-based U-net with attention gates in skip connection. The targets are frames randomly from a video captured by UAV, which is the dataset from Kaggle. The right column shows that the original designed model is sensitive, but with higher false positive. The developed model with attention gate sacrificed the sensitivity but decreased the false positive and increased accuracy.

As shown in Fig. ?? is the experience results of the on-board detection. The designed ResNet18 is successfully deployed on DJI M300 UAV. The experience demonstrated that the work of model size decreasing is acceptable. However, the model has not been developed yet. In the first time experience, accuracy is sacrificed to ensure the stable flight and decrease the detection as building the light model.

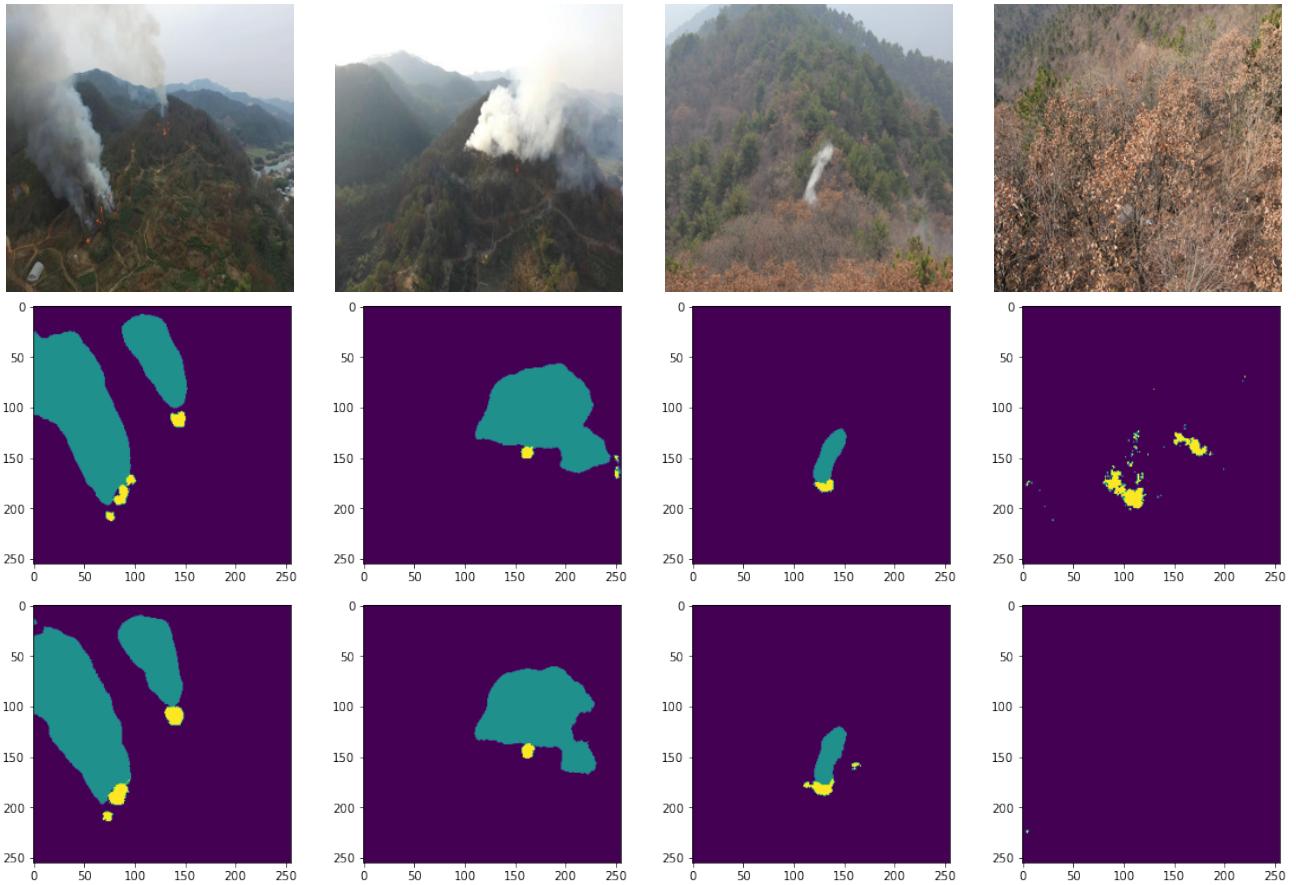


Figure 9: Segmentation results of ResNet34-based U-net on ground work station. Left two columns: Targets on Kaggle and segmentation results. Right two columns: Experience video data captured by DJI Phantom 4 UAV. Top row: Targets. Mid row: Segmentation results of the designed model. Bottom row: Segmentation results of the designed model developed by attention gates.

At the start period of experience, the cloud is detected as smoke, which means the model still need more adjustment.

7.4 Wildfire Prediction

The wildfire prediction is based on Cellular model and developed through LSTM. The results could be illustrated in Fig. ??, it could be demonstrated that the wind impacts the wildfire spreading most heavily.



Figure 10: On-board segmentation result. Left: Masked original image. Mid: Segmentation mask. Right: infrared camera image.