Fire and Smoke Detection using CNNs trained with Fully Supervised methods and Search by Quad-Tree

Gonçalo Perrolas
goncalo.perrolas@tecnico.ulisboa.pt
Alexandre Bernardino
alex@isr.tecnico.ulisboa.pt
Ricardo Ribeiro
ribeiro@isr.tecnico.ulisboa.pt

Instituto de Sistemas e Robótica
Instituto Superior Técnico
Lisboa, Portugal

Abstract

Wildfires prevail as one of the most destructive and uncontrollable natural disasters for man-kind. The fire-fighting combat team can greatly benefit from reliable information about the current position of active burning areas. It is also important to detect as early as possible the fire ignition sources through the smoke columns produced.

In this work we use aerial images taken from drones of the wildfires to detect fire and smoke using deep neural networks. A set of tests is presented to evaluate the performance of the network used in the fire and smoke detection. To solve the multi-scale detection problem we use a Quad-Tree method in the search task.

The proposed system shows an adequate performance in real drone aerial images.

1 Introduction

This article presents initial results of the Firefront Project (`www.firefront.pt`). Its main objective is to create a support system to the fire combat teams. The system will transmit valuable information about the wildfire in real-time to the ground teams using aerial vehicles to capture images of the scene. Next, the images taken, are used to detect fire and smoke using a convolution neural network able to segment the respective areas.

It comes as a challenge the detection of variable size portions of both fire and smoke, and due to the small input size image that can be feed to the common CNNs an extra algorithm was used for doing detection on smaller image sections dynamically using a Quad-Tree search method.

1.1 State-of-the-Art

The problem of fire and smoke detection using RGB images has been highly studied on the last few years, using different approaches and techniques. Existing methods can be divided into two big groups: classic methods and methods that use deep learning.

The classic methods tend to use the RGB components of images via histogram analysis, to evaluate what colours are more related to fire areas. One example is the method used by Cruz *et al.* [1] that relies on indexes based on RGB components to boost fire detection and then applies a threshold to binarize the image into the classes. This process is very time and computing power efficient but results in poor detection performances and this method is very constrained to the environment conditions of the images taken and the camera characteristics. The active system CICLOPE that is operating in Portugal, developed by Batista *et al.* [2] uses a background subtraction method using fixed cameras to do smoke detection. This technique makes use of the advantages of comparing consecutive frames with the same background, showing a good performance but, as we intend to use aerial vehicles to gather the images, a different method must be used.

At the other side of the spectre, deep neural networks can be used in this detection problem, as this area of investigation is evolving increasingly year after year. The most common type of neural network used in these types of problems is convolutional neural networks. These are able to extract image information on an abstract level allowing to choose the characteristics that better represent fire or smoke portions of the image. Frizzi et al. [3] used a type of CNN using weak supervision training to do fire and smoke segmentation. Training is a crucial stage while working with CNNs and requires a large set of images paired with the corresponding labels for fire and smoke. The results shown proved that the trained network was over-fitted to the data because of the small dataset used, which is an important challenge due to the lack of fire and smoke datasets available. There are a lot of CNN extensions, one of them being the R-CNN. Barmpoutis *et al.* [4] made use of this type of CNN and got good results, it showed an efficiency way of detecting one instance of fire with a bounding box. On that work it was used an interesting and complete fire dataset called Corsican Fire Dataset [5] which was pretty useful for the training of our networks.

2 Methodology

Firstly the system requires an aerial vehicle to capture the images from a wildfire or the beginning of one. Then, two independent systems are used, one in charge of fire segmentation and other for smoke segmentation. As input to these systems we are going to fed small portions of the images (patches) to be able to detect small or big areas of fire/smoke on those images. The process of slicing the image into smaller patches it is going to be done using a Quad-Tree methodology. An overall diagram of the complete system and the corresponding results are shown in Figure 1.

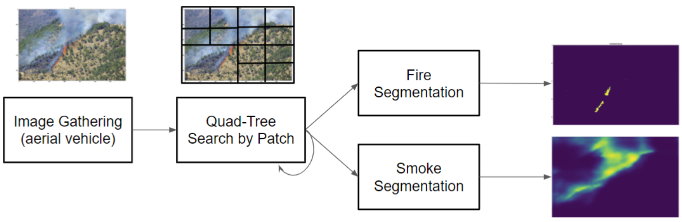


Figure 1: Global Diagram of the system.

2.1 Search by Quad-Tree

Quad-Tree methods are often used to partition a two-dimensional space by recursively subdividing it into four quadrants or regions. This is useful for doing a dynamic search for fire or smoke occurrences on the images, starting first with larger scale patches and, if nothing is detected, then moving on to smaller patches by slicing the previous ones in four segments. By using this method we solve the issues related to the multi-scale nature of fire/smoke, areas, in large patches are difficult to be detected by the neural networks (small input size) and patches that are almost filled completely with fire or smoke can have dubious detection results due to the lack of more external features that define the phenomenon.

2.2 Patch Processing (Classification + Segmentation)

The detection system is composed of two different neural networks, one for classification and other for segmentation. The classification network chosen was a SqueezeNet model [6] due to its level of accuracy compared to state-of-art models like Alex-Net while having a lot fewer parameters making the model more suitable for smaller datasets. For segmentation we used U-net [7], one of the most used state-of-the-art models for semantic segmentation.

By using a classification stage before doing segmentation the overall performance is increased due to the fact that it is easier to create a more complete dataset with image level labels then with pixel level. This reduced the number of false detection on the segmentation.

The overall logic of the system can be explained as following: each patch of the image produced by the Quad-Tree stage is given as input to the SqueezetNet [6] to determine if the patch contains fire or smoke. If a phenomenon is detected then the patch moves along to the segmentation stage where the areas of fire/smoke are segmented by the U-Net [7]. If nothing is detected in the classification stage the patch is not segmented and the process moves to the next patch in the sequence.

2.3 Datasets

Both networks need a set of images with the corresponding labels identifying the respective classes. For the classification network the labels needed are on image level (the images does/doesn't contain fire/smoke) for the segmentation model we need pixel level labels. The images gathered for the fire dataset mainly came from three different sources: Corsican Dataset [5] (RGB images with pixel wise labelling), smaller datasets found online and a batch of images gathered online that were manually labelled to extend as much as possible the size of the dataset. The smoke dataset consists also of datasets with pixel wise labels found online and some more images segmented manually.

The datasets for classification include the ones used for the segmentation training together with some more images, as this one doesn't need extensive labelling. In table 1 we present a small overview of the datasets used for the training phase where is identified the corresponding number of images.

Classification	Fire	Positive	800 imgs
		Negative	520 imgs
	Smoke	Positive	500 imgs
		Negative	300 imgs
Segmentation	Fire	Containing Fire	700 imgs
		Negative	450 imgs
	Smoke	Containg Smoke	300 imgs
		Negative	60 imgs

Table 1: Overview of the Dataset

Good negative cases for fire are sunrises, sky with red tonalities, red objects and rooftops. For smoke the main concern was related to clouds due to the very challenging similarities.

3 Results

The dataset was divided as three different subsets randomly, one for training, one for validation and one for testing purposes (70%,20%,10%). After training all the four networks, we evaluated the performance of the methods here proposed. On table 2, the full system segmentation results are shown. The results show a good general performance, although it is rather important to refer that the classes fire/negative and smoke/negative are unbalanced within each image, typically the images have a smaller area of fire or smoke. We can also conclude that the smoke detection performance is a bit worse than the fire. In terms of processing time, for an average size image (800 x 500), the processing takes roughly 4 secs, running on a Tesla K80 GPU and Intel Xeon CPU.

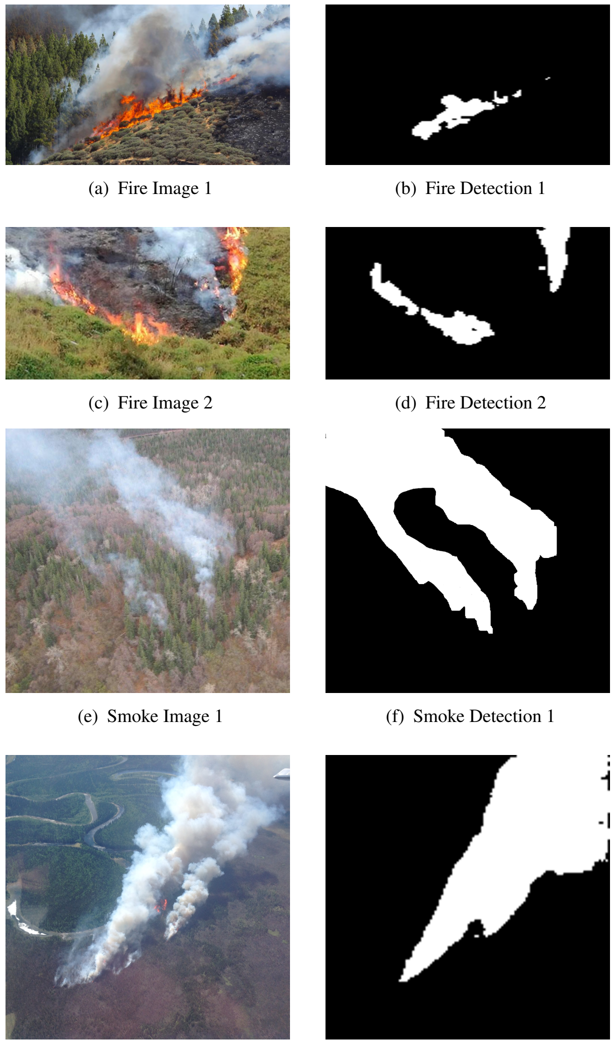
	Avg. IoU	Pixel Accu.
Fire	0.8692	0.8348
Smoke	0.8404	0.7519

Table 2: Performance Metrics on test set

On figure 2, are shown some examples of the results of the detections produced by the developed system.

4 Conclusions

The datasets are the core asset of the neural networks used in this approach. As we improved our dataset we observed a boost on the robustness of the performance of the detection for both fire and smoke. It would be a huge progress in the development of this area of study if it was created a centralised and accessible database with more diverse set of images. The results shown here prove that the system is reliable and can be applied to real life monitoring of fire situations and it has potential to represent a big improvement in the way we deal with fire.



(a) Fire Image 1 (b) Fire Detection 1
(c) Fire Image 2 (d) Fire Detection 2
(e) Smoke Image 1 (f) Smoke Detection 1
(g) Smoke Image 2 (h) Smoke Detection 2

Figure 2: Examples of results produced for both fire and smoke detection

Acknowledgements

This work was supported by FCT with the LARSyS - FCT Project UIDB/50009/2020 and project FIREFRONT (PCIF/SSI/0096/2017)

References

[1] H. Cruz M. Eckert J. Meneses J. Martínez. Efficient forest fire detection index for application in unmanned aerial systems (uass). *Sensors (Basel)*, June 2016.

[2] M. Batista B.Oliveira P.Chaves J. Ferreira T. Brandão. Improved real-time wildfire detection using a surveillance system. In *Proc. of the World Congress on Engineering 2019*, 2019.

[3] S. Frizzi R. Kaabi M. Bouchouicha J. Ginoux E. Moreau F. Fnaiech. Convolutional neural network for video fire and smoke detection. *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence*, 2016.

[4] P. Barmpoutis K. Dimitropoulos K. Kaza and N. Grammalidis. Fire detection from images using faster r-cnn and multidimensional texture analysis. *ICASSP - IEEE*, 2019.

[5] T. Toulouse L. Rossi A. Campana T. Celik M. Akhloufi. Computer vision for wildfire research: an evolving image dataset for processing and analysis. *Fire Safety Journal*, 2017.

[6] Forrest N. Iandola Song Han Matthew W. Moskewicz Khalid Ashraf William J. Dally Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size. 2016.

[7] O. Ronneberger P. Fischer T. Brox. U-net: Convolutional networks for biomedical image segmentation. 2015.