

# **Significance of Image Pre-processing in Computer Vision Based Fire Detection, A Review**

## **Task 2: Literature Review**

Ashwin Rajesh - 14259321

### **Contents**

1 Abstract .....	2
2 Introduction .....	2
3 Literature Review .....	2
4 Image Pre-Processing Algorithms .....	4
4.a HSV filter .....	4
4.b Edge & Contour Detection .....	4
4.b.a Sobel Operator .....	5
4.c Corner Detection .....	5
4.d Dark Channel Prior .....	6
4.e Histogram Equalization .....	7
Bibliography .....	7

# 1 Abstract

## 2 Introduction

Wildfires pose a significant threat to humans, wildlife and the environment alike. Left unchecked, A wildfire can spread rapidly causing large scale destruction to forests & infrastructure, as well as releasing large amounts of pollutants into the atmosphere. Therefore, it has become increasingly crucial to detect them as early as possible. This report aims to explore different methods employed in improving the performance and efficiency of wildfire detection systems, with a focus on computer vision based fire detection using Convolutional Neural Networks (CNNs).

## 3 Literature Review

UAVs have seen extensive research for this application, as they provide valuable visual data which can be employed to carry out search and rescue operations, save forest resources, help firefighters navigate efficiently among numerous other enhancements. Rather than sending ground crews to monitor hazardous environments, UAVs can drastically reduce the risk to firefighting crews by remotely scanning large amounts of dangerous forest area. However UAVs tend to perform poorly in harsh weather conditions, have limited flight time and require a human operator to have a visual line of sight (Afghah, 2024). These factors combined with expensive operation costs severely limit the effectiveness of UAVs in many circumstances.

Edge computing sensor nodes can be placed at high risk locations that can monitor temperature, humidity and other characteristics of air in the surrounding area. These systems generally use a microcontroller to operate sensors, relying on solar cells for long-term power. By setting up multiple such nodes in different areas forming a Wireless Sensor Network (WSN), the collected sensor data may then be analyzed to determine the locations of fire events (Mohapatra & Trinh, 2022). More advance methods of data classification with the use of artificial neural networks boast a high accuracy of >82% with multiple sensors (Yan et al., 2016). The low cost of WSN systems has made it an increasingly popular choice for real-time monitoring of forest fires. However, sensor nodes fail to be viable in some situations. Establishing wireless communications for a WSN can prove challenging in rural or untamed areas such as forests (Talaat & ZainEldin, 2023). The sensor nodes may also be prone to damage from the wildfire, needing to be replaced in order to continue using the system.

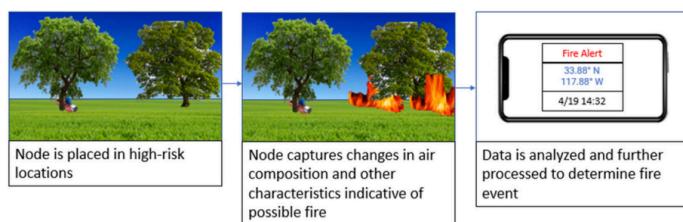


Figure 1: Basic operation of WSNs. Images sourced from Creative Commons, following the guidelines on Attribution 3.0 Unported, CC By 3.0

Satellite-based fire detection can potentially offer significant advantages over traditional methods due to the vast areas they can monitor. The benefits and limitations of these systems largely depend on the satellite's orbit. Satellites in Sun-Synchronous Orbit (SSO) provide high spatial resolution but revisit the same location only after several days, resulting in low temporal resolution. This delay makes SSO satellites less effective for real-time wildfire detection (Thangavel et al., 2023). In contrast, Geostationary Earth Orbit (GEO) satellites remain fixed over the same region, as their orbital period matches Earth's rotation. Equipped with multispectral imaging sensors, GEO satellites provide continuous monitoring, making them ideal for detecting and tracking fires in real time.

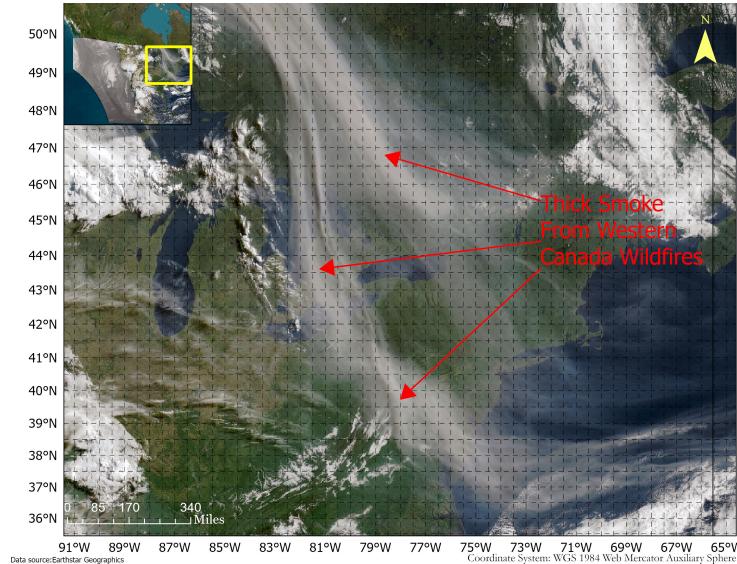


Figure 2: Satellite Based Wildfire Detection

Despite the various classifications of technologies within wildfire detection, advancements in computer vision object detection has spurred use in terrestrial, aerial as well as satellite systems.

[Thangavel et al. \(2023\)](#) explores the feasibility of space-borne fire detection using onboard computer vision to rapidly generate alerts. The traditional method of downloading image data from satellites has too much latency for time-sensitive applications such as fire detection. By examining performance of hardware accelerators for edge computing such as Nvidia Jetson and Movidius Myriad 2, the report found that AI inference on-board the satellite is possible in terms of performance as well as power consumption, allowing for low latency detection of wildfires. [Ghali et al. \(2022\)](#) details methods to optimise Deep Learning to deal with UAV limitations such as image degradation, small fire size & background complexity. By combining EfficientNet-B5, vision transformers and EfficientSeg convolutional model, an accuracy of 85.12% was obtained, outperforming many state-of-the-art models. [Talaat & ZainEldin \(2023\)](#) proposes a low cost WSN fire detection approach based on YOLO object detection models. The system, named SDFS was trained on a dataset of 26,520 images and achieved a high precision rate of 97.1% for smoke and fire classification.

Due to the importance of computer vision based object detection in wildfire detection, it is imperative that they are as accurate as possible. There are many different techniques to improve the performance of neural networks. Transfer learning is the process of using a model that has been pre-trained on a large dataset and further training and fine tuning it to a specific task. [Asif et al. \(2023\)](#) studies the impact of transfer learning on classification by comparing six different pre-trained architectures. The models were trained using an open access dataset of 3886 images and found that all six architectures were able to achieve a testing accuracy of above 90% despite a relatively small training dataset. Augmenting data with synthetically generated samples investigated by [Wong et al. \(2016\)](#) found that inserting data-space transformations such as warped images provided improved performance and reduced overfitting. [Ryu & Kwak \(2022\)](#) proposed an image pre-processing pipeline using advanced techniques such as HSV filters and corner detection to assist models with classification by eliminating unwanted noise in images. This method has observed to improve fire detection accuracy by 5.5% and smoke detection by 6% in object detection models.

This paper contributes a benchmark of different pre-processing filters and algorithms to uncover insights on an ideal pipeline for wildfire detection, using metrics such as speed, power consumption and model accuracy. The benchmarks are carried out on a Raspberry Pi 4B, to simulate low-powered

edge computing hardware that is consistently used in terrestrial, UAV and satellite systems. Large deep-learning networks are unviable on these systems due to the high computational intensity, while models with reduced parameters are more efficient but suffer from less accurate inference. The preprocessing pipeline aims to improve the accuracy of lightweight models by highlighting important features in fire and smoke, while reducing unwanted noise in the image. For classification,

## 4 Image Pre-Processing Algorithms

### 4.a HSV filter

Hue, Saturation, Value (HSV) is a cylindrical-coordinate representation of points in an RGB colour model. It is an alternative representation of the RGB color model that intends to describe colors in a way that is more aligned with human perception. Colour masks can be created by setting upper and lower bounds within the HSV channels. This can then be applied to an image to filter specific tones of colour. A study by [Hassan & Saud \(2023\)](#) utilized HSV thresholding to detect human skin in an image. The process achieved an impressive 99.587% accuracy on natural images under varying light conditions. [Ryu & Kwak \(2022\)](#) used a HSV filter in addition to other preprocessing stages to isolate colours of fire in an image to assist object detection models with inference, which resulted in a 5.5% increase in fire detection accuracy. The following bounds were used to preprocess images in the paper:

$$\text{RoI}_{\text{HSV}(x,y)} := \begin{cases} 1 & \text{if } 20 < H(x,y) < 40 \text{ and } 50 < S(x,y) < 255 \text{ and } 50 < V(x,y) < 255 \\ 0 & \text{otherwise} \end{cases}$$



Figure 3: Resulting Image after HSV mask

### 4.b Edge & Contour Detection

Edge detection involves identifying regions of an image where the contrast between pixels changes dramatically. Capturing edges can be used to reveal important properties in the image as these edges often correspond to changes in depth, illumination, orientation or material. Edge detection can be particularly useful in detecting fire, as it uncovers data about positions of contours and contrasts where fire and smoke could potentially exist. More specifically, method to use low edge responses in an image region may be useful in differentiating between smoke and sky. This method is explored further in this paper

Edge detection has seen use in many fields to extract valuable information from visual data. [Sharifrazi et al. \(2021\)](#) aims to accurately detect COVID-19 patients by using edge detection to improve detection accuracy of a CNN on CT images of the lungs. The addition of sobel edge detection to a CNN proved to be an effective approach, achieving an accuracy of 99.02% on a custom dataset. [Ghandorh et al. \(2022\)](#) proposes an approach to detect roads in satellite images using edge detection and semantic-segmentation. The results showed accurate segmentation & edge detection even in complicated backgrounds. This shows the potential of edge detection in satellite and UAV based systems, where small details in an image are of significant importance.

#### 4.b.a Sobel Operator

A popular method of edge detection thanks to its computational simplicity is the Sobel operator. The sobel filter involves convolving the image with a specific kernel which calculates the gradient of the image in x and y directions (Sobel, 2014). For a given image, let us consider a pixel region such as:

50	50	100	100
50	<b>50</b>	100	100
50	50	100	100
50	50	100	100

Where the value in each cell is the brightness of the pixel in that position within the image. In the region shown above, we can see that the pixel brightness rapidly changes between the 2nd and 3rd columns, which would be perceived as an edge by humans. We can then convolve the gradient filter shown below over each pixel of the image.

1	0	1
-2	0	-2
1	0	1

Let us consider the pixel in the second column and second row as the pixel currently being processed. The gradient filter is multiplied with the neighbouring 3x3 area centered by our pixel of interest, and repeated with every other pixel to produce the partial derivative of the image in the x direction. By obtaining the y partial derivative in a similar manner, we may combine the images to produce a resultant image containing high absolute values near edges and a value close to 0 everywhere else.

#### 4.c Corner Detection

Corner detection is a common technique used in computer vision to infer features from an image. Ryu & Kwak (2022) shows that a corner detection algorithm is a vital preprocessing stage, as it separates fire from other objects of similar color, which would fall within the thresholds of a HSV filter. The popular Harris corner detector uses the autocorrelation function of the image to determine intensity differences within patches of an image. Using a Taylor expansion, the autocorrelation function can be approximated as (Ryu & Kwak, 2022; Sánchez et al., 2018):

$$I(x_i, y_i) + [I_{x(x_i, y_i)} I_{y(x_i, y_i)}] \begin{bmatrix} u \\ v \end{bmatrix}$$

Where  $I$  represents intensity and  $u$  and  $v$  represent the shift in the region from the reference pixel  $(x_i, y_i)$ . This shows that the change in intensity depends on the partial derivatives  $I_x$  and  $I_y$  of the image. When written in matrix form, the expression is as follows:

$$M = \begin{pmatrix} \sum(I_{x(x_i, y_i)}^2) & \sum(I_{x(x_i, y_i)} I_{y(x_i, y_i)}) \\ \sum(I_{x(x_i, y_i)} I_{y(x_i, y_i)}) & \sum(I_{y(x_i, y_i)}^2) \end{pmatrix}$$

The eigenvalues of the matrix can be found using the determinant and trace:

$$\det(M) = AB - C^2 = \lambda_1 \lambda_2$$

$$\text{trace}(M) = A + B = \lambda_1 + \lambda_2$$

There are three possible situations based on their values:

- Both eigenvalues are small: this happens when the pixels are in a flat region
- One eigenvalue is bigger than the other eigenvalue: The region likely is an edge
- Both eigenvalues are large: the region is a corner

Therefore, corner regions within an image will output a high corner strength. These regions can be used as a candidate region that can be inferred through a CNN.

#### 4.d Dark Channel Prior

Dark Channel Prior has been commonly used to measure the degree of haziness as well as haze-removal in images. The technique is based on the observation that in most outdoor images, pixels tend to have low intensities in atleast one colour channel (dark channel). This property can be used to estimate the transmission map of an image, representing the amount of haze affecting the scene. Atmospheric haze can be modelled as follows (He et al., 2011):

$$I(x) = J(x)t(x) + A(1 - t(x))$$

$I(x)$  represents a pixel that reached the camera.  $J(x)$  represents the undistorted pixel.  $t(x)$  is the transmission map, representing how much scene radiance is retained, where a value of 1 means no haze and 0 means maximum haze. Due to the scattering of light from haze, low intensity channels in hazy patches of an image have an inherently higher value. As a result, DCP can be used to estimate  $t(x)$  providing the areas of the image affected by haze. Ryu & Kwak (2022) and Liu et al. (2019) note that dark channel prior methods can apply to smoke due to the similar nature, having relatively higher values in their dark channels. This causes smoke to be picked up as an area of high haze in the transmission map. Ryu & Kwak (2022) apply a threshold to the transmission map, extracting areas with high dark channel values and suppressing the rest. As a result, a 6% increase in detection accuracy was achieved compared to detection without extra preprocessing stages.



Figure 4: Image processed with Dark channel prior & thresholded to high intensity values

As shown in image Figure 4, DCP reveals areas with high intensity values in all three channels, causing haze, smoke, and sky regions to appear very bright compared to the rest in the second image. In the third image, regions below a certain intensity threshold are suppressed to zero. This gives us the regions where there is a high chance of smoke or fog.

Despite the notable improvements in smoke detection using dark channel prior, there are some drawbacks to be considered. Dark channel prior tends to be unreliable when the image consists of a large portion of sky, since the sky tends to have a high dark channel value, causing the algorithm to mistake it as hazy or smoky area.

#### 4.e Histogram Equalization

Histogram Equalization increases the global contrast in images, which can enhance the visibility of finer details within an image. The algorithm spreads the intensity values out in an image so that it utilizes the full range of values more efficiently (Patel & Goswami, 2014). As a result, HE is most effective on images with a narrow range of intensity values.

HE's effectiveness at enhancing image quality makes it a practical technique in many scenarios. Ergasheva et al. (2024) presents a system to detect early fires inside a ship using a YOLO computer vision model. Images were preprocessed with Histogram Equalization to reduce the impact of water vapour on the quality of the images, contributing to the remarkable 99% accuracy of the model. Nguyen et al. (2021) investigates using CLAHE, an advanced form of histogram equalization along with a YOLOv4 model to detect bone fracture features in xray images, obtaining a result of 81.91% when trained on a small dataset.

Let us consider  $n_i$  as the number of occurrences of the gray level  $i$  in a greyscale image. The probability of a pixel with level  $i$  is as follows:

$$p(i) = \frac{n_i}{n}, 0 \leq i < L$$

Where  $L$  is the total number of gray levels in the image. The cumulative distribution function can then be defined:

$$\text{cdf}(i) = \sum_{j=0}^L p(j)$$

In order to achieve a flat histogram of values, the CDF must be linearized. This can be carried out by the following equation:

$$h(v) = \text{round} \left( \frac{\text{cdf}(v) - \text{cdf}_{\min}}{N - \text{cdf}_{\min}} \right)$$

Where  $N$  is the number of pixels in the image.

### Bibliography

- Afghah, F. (2024). Autonomous Unmanned Aerial Vehicle Systems in Wildfire Detection and Management-Challenges and Opportunities. In E. Blasch, F. Darema, & A. Aved (Eds.), *Dynamic Data Driven Applications Systems: Dynamic Data Driven Applications Systems*.
- Asif, S., Wenhui, Y., Amjad, K., Jin, H., Tao, Y., & Jinhai, S. (2023). Detection of COVID-19 from chest X-ray images: Boosting the performance with convolutional neural network and transfer learning. *Expert Systems*, 40(1).
- Ergasheva, A., Akhmedov, F., Abdusalomov, A., & Kim, W. (2024). Advancing Maritime Safety: Early Detection of Ship Fires through Computer Vision, Deep Learning Approaches, and Histogram Equalization Techniques. *Fire*, 7(3), 84. <http://ezproxy.lib.uts.edu.au/login?url=https://www.proquest.com/scholarly-journals/advancing-maritime-safety-early-detection-ship/docview/3001523245/se-2>
- Ghali, R., Akhloufi, M. A., & Mseddi, W. S. (2022). Deep Learning and Transformer Approaches for UAV-Based Wildfire Detection and Segmentation. *Sensors (Basel, Switzerland)*, 22(5), 1977–.

- Ghandorh, H., Boulila, W., Masood, S., Koubaa, A., Ahmed, F., & Ahmad, J. (2022). Semantic Segmentation and Edge Detection—Approach to Road Detection in Very High Resolution Satellite Images. *Remote Sensing (Basel, Switzerland)*, 14(3), 613–.
- Hassan, E. K., & Saud, J. H. (2023). HSV color model and logical filter for human skin detection. *AIP Conference Proceedings*, 2457(1).
- He, K., Sun, J., & Tang, X. (2011). Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12), 2341–2353.
- Liu, Y., Qin, W., Liu, K., Zhang, F., & Xiao, Z. (2019). A Dual Convolution Network Using Dark Channel Prior for Image Smoke Classification. *IEEE Access*, 7, 60697–60706. <https://doi.org/10.1109/ACCESS.2019.2915599>
- Mohapatra, A., & Trinh, T. (2022). Early Wildfire Detection Technologies in Practice—A Review. *Sustainability*, 14(19), 12270–.
- Nguyen, H. P., Hoang, T. P., & Nguyen, H. H. (2021). A deep learning based fracture detection in arm bone X-ray images. *2021 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, 0, 1–6. <https://doi.org/10.1109/MAPR53640.2021.9585292>
- Patel, S., & Goswami, M. (2014). Comparative analysis of Histogram Equalization techniques. *2014 International Conference on Contemporary Computing and Informatics (Ic3i)*, 0, 167–168. <https://doi.org/10.1109/IC3I.2014.7019808>
- Ryu, J., & Kwak, D. (2022). A Study on a Complex Flame and Smoke Detection Method Using Computer Vision Detection and Convolutional Neural Network. *Fire*, 5(4). <https://doi.org/10.3390/fire5040108>
- Sharifrazi, D., Alizadehsani, R., Roshanzamir, M., Joloudari, J. H., Shoeibi, A., Jafari, M., Hussain, S., Sani, Z. A., Hasanzadeh, F., Khozeimeh, F., Khosravi, A., Nahavandi, S., Panahiazar, M., Zare, A., Islam, S. M. S., & Acharya, U. R. (2021). Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using X-ray images. *Biomedical Signal Processing and Control*, 68, 102622. <https://doi.org/https://doi.org/10.1016/j.bspc.2021.102622>
- Sobel, I. (2014). An Isotropic 3x3 Image Gradient Operator. *Presentation at Stanford A.I. Project 1968*, .
- Sánchez, J., Monzón, N., & Salgado, A. (2018). An Analysis and Implementation of the Harris Corner Detector. *Image Processing On Line*, 8, 305–328.
- Talaat, F. M., & ZainEldin, H. (2023). An improved fire detection approach based on YOLO-v8 for smart cities. *Neural Computing and Applications*, 35(28), 20939–20954. <https://doi.org/10.1007/s00521-023-08809-1>
- Thangavel, K., Spiller, D., Sabatini, R., Amici, S., Sasidharan, S. T., Fayek, H., & Marzocca, P. (2023). Autonomous Satellite Wildfire Detection Using Hyperspectral Imagery and Neural Networks: A Case Study on Australian Wildfire. *Remote Sensing*, 15(3). <https://doi.org/10.3390/rs15030720>
- Wong, S. C., Gatt, A., Stamatescu, V., & McDonnell, M. D. (2016). Understanding Data Augmentation for Classification: When to Warp?. *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 1–6.
- Yan, X., Cheng, H., Zhao, Y., Yu, W., Huang, H., & Zheng, X. (2016). Real-Time Identification of Smoldering and Flaming Combustion Phases in Forest Using a Wireless Sensor Network-Based Multi-Sensor System and Artificial Neural Network. *Sensors*, 16(8). <https://doi.org/10.3390/s16081228>