

# Research on real-time fire detection and locating for automotive firefighting robot in factories based on Convolutional Neural Network

**Haoyu Tan**

Concordia University, 1455 De Maisonneuve Blvd. W. Montreal, QC, H3G 1M8,  
CANADA

t\_haoyu@live.concordia.ca

**Abstract.** Automotive fire robots that are used in factories can carry out diverse operations in regard to patrolling, fire detection, and programmed fire rescue. An accurate detection of fire sources in factories is significantly crucial for unmanned firefighting robot in terms of building a reliable sensor system. An approach proposed by this paper to recognize the color and dynamic shape of varying flames based on HSV color algorithms and Convolutional Neural Network. As a comparison to traditional RGB image processing, this approach is more efficient in isolating colors in environment and more adaptive to a fire site that includes multiple noise factors. The research in this paper uses image processing algorithms that is trained by CNN to detect flames in simulated factory environments, followed by a HSV color locating algorithm to compute the coordinates of target fire to perform inverse kinematic analysis on an unmanned firefighting robot.

**Keywords:** robotics, firefighting robot, fire detection, image processing, deep learning.

## 1. Introduction

Fires in industrial and manufacturing properties are destructive and catastrophic for property security and human well-being. From 2011 to 2015, an annual average of 37,910 fires in U.S, including outside trash fire, non-trash outside fire, structure fire, and vehicle fire, was estimated by U.S fire departments according to their reports, associated with 6 civilian deaths, 273 civilian injuries and approximately \$1.2 billion of direct property loss [1]. Therefore, surveillance and control of fire risks of great significance to be considered in terms of protecting factory sites and their continuous productivity from those irreversible damages. Traditionally, supervisions on early fire are carried out by heat and smoke detectors, which are disadvantageous in latency, accuracy, and its maximum allowable working area [2]. Conventional fire and smoke detectors lack the capability of working in large factory sites where smoke is hard or sluggish to be captured, especially in outdoor facilities. Consequently, adapted technologies have been developed for detecting fires in various environments, such as advanced digital cameras with image processing algorithms, drones working with thermal sensor inputs and other vision-based censoring systems [3].

With the development of the industrial economy and the growing complexity of building structures, the difficulty of inspecting industrial and manufacturing sites has grown accordingly [4]. Hence, vision-

based fire detecting and locating systems are developed with smartness so that the system could detect fires happening in facility that varies in shape, dimensions, color, brightness, and all the other affective factors. However, most applications of fire detecting systems demand manual supervision or lacks flexibility. For example, CCTV fire alarming system can capture fires from video with high resolution which makes fire highly recognizable and provides precise input for image processing stage. But the installation of CCTV cameras and the configurations of monitoring angles are fixed to its specific working conditions, which loses its best efficiency when its surroundings are subjected to changes [5]. Benchmarking firefighting robots, on the other hand, are advantageous for having an ignorable number of limitations on their working environments, since their designers chose to apply thermal sensors like thermal camera on the robot for simplifying the designing process [6]. However, compared to optical cameras, thermal camera filters image information including colors, saturations etc. This application reduces the working load of the onboard computer on the firefighting robot, but also wastes the extensibility of the entire censoring system when the computer is adequately powerful and capable of advanced algorithms.

Considering the strengths and limitations of the two approaches, convolutional neural network (CNN) is introduced and proposed for applications on firefighting robots. As a branch in the field of deep learning and a revolution of artificial neural networks, it is an effective tool to perform machine learning and effectively extract the target features from images despite where the target is located in the picture [7]. In the applications of fire detection, CNN is able to automatically extract and learn the features of fire in all desired conditions, ranging from laundry factories to nuclear factories, as long as a related dataset is given [8]. It can adapt to diverse working environments since the network would be trained to be aware of how the fires visibly behave in the dataset and predict the presence of fire in the real-time video caption. Moreover, the CNN uses the raw optical camera video as its input source, where colors, saturation, and brightness are all recorded so that the video input could be extended to other usage like water leaking, hazardous exposures and other risks detections [9, 10].

In this paper, a CNN-based flame detection method is developed for an automotive firefighting differential robot project that patrols around factory facility and extinguish detected fire with equipped water gun. The fire detection system in this project is considered as a switch from patrolling mode to rescue mode, followed by fire locating algorithms to obtain the fire coordinates among relative 2D axis. It is designed for detect fire while patrolling and proceed necessary location information for kinematic analysis in lateral operations. Section 2 introduces the applied methods, including dataset collection, CNN construction for fire detecting and HSV scale conversion for fire locating. The experimental results are demonstrated in Section 3, which further validate the feasibility of the proposed method. The relevant content of this paper is analyzed and prospected in Section 4.

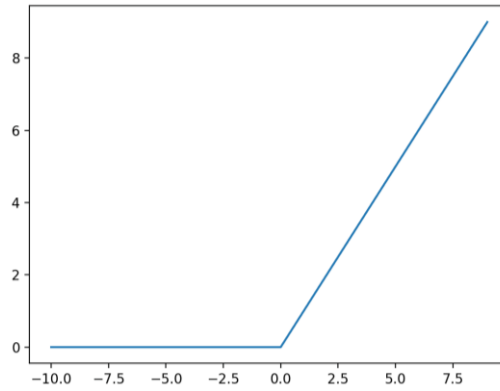
## **2. Methodology**

### *2.1. Dataset collection*

To accurately and precisely fulfill the demand for early fire detection, a dataset of 2000 images of fires in various scenarios is collected and imported from GitHub community. However, the original dataset consists of flame images in multiple scenarios that varies in city roads, buildings, and forests. Aiming to train CNN to detect fire in factories particularly, the images of facility that have similar scale, shape, and lighting conditions to those in factories are manually selected from the original dataset. This filtering process can optimize the dataset by increasing its compatibility with factory fires so that the CNN model can be trained and learnt to recognize fire in the desired environment. Some example images are shown in Figure 1.



After each convolutional operation, To add non-linearity to the CNN model, an activation function is used. Rectified Linear Unit (ReLU), which considers all minus numbers as 0 while maintaining positive values, is used in the research of this paper, allowing the network to learn complex differentiation of flame in diverse colors and surroundings. The ReLU function is shown in Figure 3.



**Figure 3.** The ReLU function.

As shown in Figure 3, the horizontal axis is the value of  $x$  in the ReLU function. And the vertical axis is the  $y$  value of the function. The blue curve represents the ReLU function.

Pooling layers are employed to decrease the spatial dimensions of the data and control overfitting. The overall pooling technique applied in this paper is max pooling, which extracts the maximum value from a subtle area of the feature map, thereby down sampling it. Average pooling is a pooling layer in CNNs that performs average pooling operations on the feature maps to down-sample and reduce spatial dimensions while preserving some spatial information from the original feature maps. Max pooling is generally considered more effective in capturing dominant features, while average pooling helps to preserve more spatial information and can be useful in certain scenarios. In addition, the dropout layer is a specific layer in a neural network where random units (neurons) are temporarily "dropped out" or "turned off" during training. This indicates that a portion of the neurons in the dropout layer are randomly set to zero during each training iteration. The dropout rate is a hyperparameter that determines the probability of dropping out a neuron. [12].

After several convolutional and pooling layers, the output is passed through fully connected layers. These layers are those of a conventional neural network, where individual neurons are linked to every other neurons in the adjacent layers. Fully connected layers help in making decisions based on the learned features.

The output layer is a single neuron or multiple neurons with an appropriate activation function. For fire detection, the output neurons will produce a high value if fire is detected and a low value otherwise. The loss function of the output computes the variance between the output of prediction and the actual ground truth label. The paper employs the categorical cross-entropy function to compute the loss, combining SoftMax activation with Cross-Entropy loss for multiclass categorization. This approach enables the training of a CNN to generate a probability distribution across  $N$  classes per image. For multiclass classification, the neural network's initial outputs undergo SoftMax activation, producing a probability vector relating to the input classes. The role of the optimizer is to adjust the network's parameters throughout training, expected to minimize the loss function. And stochastic gradient descent (SGD) and RMSprop are applied to the CNN model in this research to get the experimentally lowest value of loss.

The CNN is trained on the dataset of images classified as "fire" and "non-fire." During training, the network learns to recognize the patterns associated with fire, enabling it to make accurate predictions on unseen data. However, the convolution process is not capable of resolving flipping and zooming of images. In order to enhance both dataset diversity and model accuracy, it is essential to incorporate data

augmentation methods including horizontal flipping and zooming during the initial image pre-processing phase.

### *2.3. Fire locating with HSV*

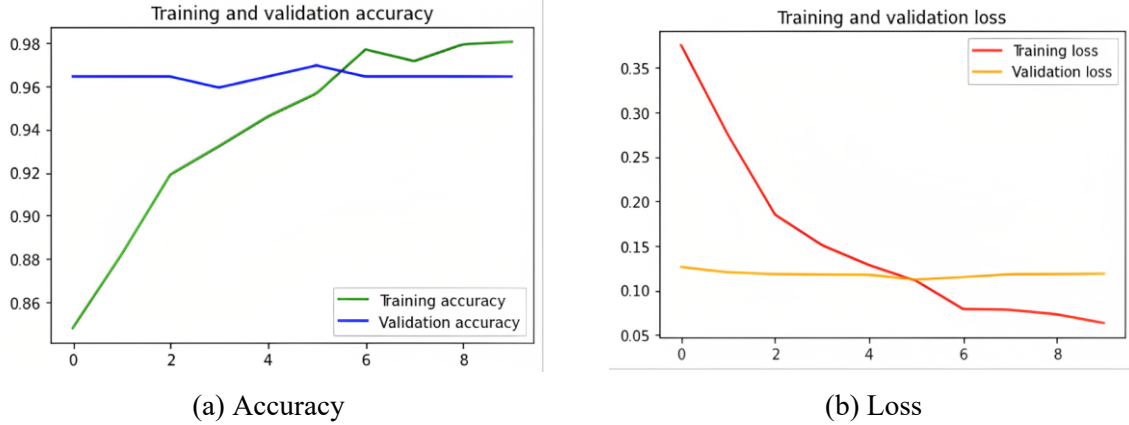
The HSV mask method is employed in fire detection to isolate specific regions within an image or video frame that are indicative of fire or flames. The Hue-Saturation-Value (HSV) color space dissects color information into three distinct components: Hue, representing the color type; Saturation, gauging color intensity or purity; and Value, denoting color brightness or intensity. The HSV mask works based on the assumption that fire typically emits colors with specific hue and saturation values, and it often appears as bright regions in the image. By defining a specific range of HSV values that correspond to the color of fire, a binary mask that highlights the potential fire regions can be created.

After the CNN model computes a valid confirmation of the presence of fire within a frame, the image is sent to a locating process. The first step is to convert the original color image (commonly in RGB color space) to the HSV color space using functions OpenCV in Python. Based on observations of fire's appearance in the HSV color space, an appropriate range of HSV values is defined that corresponds to the color of fire. For example, the hue range might cover red, orange, and yellow colors, while the saturation and value ranges might be set to capture intense and bright regions. With the defined HSV range, a binary mask is created by thresholding the HSV image. Pixels within the HSV range will have a value of 1 (white) in the mask, while pixels outside the range will have a value of 0 (black). The final HSV mask highlights the potential fire regions in the original image. The mask can be used to locate the fire by extracting contours or connected components and to draw bounding boxes around the detected regions helps visualize and identify the areas corresponding to fire or flames.

## **3. Experimental results and analysis**

### *3.1. Fire detection based on CNN*

The experimental results of the fire detection CNN algorithm demonstrate its remarkable performance and potential for real-world applications. With an impressive accuracy of 0.980 and a low loss of 0.063 on the test dataset, the CNN model has exhibited a high level of accuracy in correctly recognizing fires in images and frames. The achieved accuracy indicates that the model has effectively learned meaningful features from the input images, enabling it to make precise predictions on unseen data. Additionally, the low loss value suggests that the model's optimization process was successful in minimizing prediction errors and capturing underlying patterns in the training data. The good generalization performance observed in the test dataset further confirms that the model is not overfitting and can handle diverse fire scenarios. These results hold promising implications for fire detection systems, as the CNN algorithm has demonstrated robustness and reliability in identifying potential fire regions. However, to ensure its practical applicability, further assessment and comparison with baseline methods are necessary, along with addressing potential limitations. Overall, the experimental results highlight the CNN algorithm's significant potential in contributing to fire detection efforts and fostering safety in various environments. The experimental results are shown in Figure 4.



**Figure 4.** Experimental results.

As shown in Figure 4. (a), the green line represents the accuracy of training, and the blue line represents that of validation. As shown in Figure 4. (b) the red line represents the loss of training, and the yellow line represents that of validation. As the accuracy of the model increased while the loss decreased in training proceeding, which shows the model's growing compatibility with factory fire detection.

### 3.2. Dataset collection

After CNN validation, the HSV mask is employed to highlight fire regions and dim the background in the images. The HSV mask is created by thresholding the HSV representation of the images, isolating specific hues, saturations, and brightness values corresponding to fire. By setting the appropriate HSV range for fire color, the mask is generated, where the pixels within this range are assigned a value of 1 (white), representing potential fire regions, while pixels outside the range are set to 0 (black), representing the background. Figure 5 shows the comparison of images before and after HSV processing.



**Figure 5.** HSV mask processing.

As shown in Figure 5, the HSV mask effectively acts as a spatial filter, amplifying the regions with fire-related colors and dimming non-fire areas. This step helps to emphasize the critical visual cues indicative of fire presence, enhancing the model's fire detection capabilities. The CNN's predictions, when combined with the HSV mask, provide a more precise localization of fire regions, as the mask

acts as a fine-grained filter that highlights the subtle color variations and intensity patterns associated with fire flames.

Moreover, by isolating the fire regions and dimming the background, the HSV mask reduces visual distractions and unwanted noise in the images, leading to better discrimination between fire and non-fire elements. This approach not only improves the overall performance of the fire detection system but also contributes to a more intuitive and visually appealing output, aiding human interpretation and verification of the model's results.

By leveraging both the CNN's deep learning capabilities and the HSV mask's color-based segmentation, the fire detection system achieves a synergistic effect, resulting in a powerful and accurate fire detection solution. The model's ability to identify fire instances is enhanced by the complementary strengths of both techniques, making it highly reliable and robust for real-time fire detection in various environmental conditions and settings.

#### 4. Conclusions

In conclusion, this paper proposes a feasible approach for flame detection in factories using CNN along with HSV color mask. The results of the experimental analysis demonstrate the effectiveness and potential of the proposed method. The CNN model achieved an impressive accuracy of 98% and a low loss of 6% on the test dataset, indicating its capability to accurately distinguish between fire and non-fire images. The CNN's ability to learn and recognize fire patterns in diverse environments makes it a versatile and adaptable solution for fire detection in various factory scenarios.

Furthermore, the incorporation of the HSV mask enhances the precision of fire localization. By isolating the fire regions and dimming the background, the HSV mask acts as a spatial filter, emphasizing the distinct visual features of fire flames. The combination of CNN's predictions and the HSV mask provides a more precise localization of fire areas, facilitating the implementation of inverse kinematic analysis on an unmanned firefighting robot. The computed coordinates of the target fire locations empower the robot to respond efficiently and effectively to fire incidents, contributing to enhanced safety measures in industrial and manufacturing properties.

Overall, this research contributes to the development of a reliable and intelligent fire detection feature for unmanned firefighting robots in factories. By leveraging deep learning techniques and color-based segmentation, the proposed method addresses the limitations of conventional fire detection systems and offers a flexible and robust solution. The successful implementation of CNN and HSV color algorithms paves the way for future advancements in fire detection and response technology, contributing to the protection of factory sites, property security, and human well-being from the devastating effects of fires.

#### References

- [1] S Campbell R., Fires in U.S. Industrial or Manufacturing Properties (2018), National Fire Protection Association
- [2] Z. Zhang and Y. Hu, "Video Smoke Detection Based on Convolution Neural Network", 2017 International Conference on Computer Technology Electronics and Communication (ICCTEC)
- [3] A. Q. Nguyen, H. T. Nguyen, V. C. Tran, H. X. Pham and J. Pestana, "A Visual Real-time Fire Detection using Single Shot MultiBox Detector for UAV-based Fire Surveillance," 2020 IEEE Eighth International Conference on Communications and Electronics (ICCE), Phu Quoc Island, Vietnam, 2021, pp. 338-343, doi: 10.1109/ICCE48956.2021.9352080
- [4] Li, P., & Zhao, W. (2020). Image fire detection algorithms based on Convolutional Neural Networks. Case Studies in Thermal Engineering, 19, 100625. <https://doi.org/10.1016/j.csite.2020.100625>
- [5] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho and S. W. Baik, "Convolutional Neural Networks Based Fire Detection in Surveillance Videos," in IEEE Access, vol. 6, pp. 18174-18183, 2018, doi: 10.1109/ACCESS.2018.2812835

- [6] J. Suresh, "Fire-fighting robot," 2017 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2017, pp. 1-4, doi: 10.1109/ICCIDS.2017.8272649
- [7] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186
- [8] Muhammad, K., Ahmad, J., & Baik, S. W. (2018). Early fire detection using convolutional neural networks during surveillance for effective disaster management. *Neurocomputing*, 288, 30–42. <https://doi.org/10.1016/j.neucom.2017.04.083>
- [9] Zhang C, Alexander BJ, Stephens ML, Lambert MF, Gong J. A convolutional neural network for pipe crack and leak detection in smart water network. *Structural Health Monitoring*. 2023;22(1):232-244. doi:10.1177/14759217221080198
- [10] Seo, S.-K., Yoon, Y.-G., Lee, J., Na, J., & Lee, C.-J. (2022). Deep Neural Network-based optimization framework for safety evacuation route during toxic gas leak incidents. *Reliability Engineering & System Safety*, 218, 108102. <https://doi.org/10.1016/j.ress.2021.108102>
- [11] Google. (n.d.). Advanced guide to inception V3. Google. <https://cloud.google.com/tpu/docs/inception-v3-advanced>
- [12] Choe, J., & Shim, H. (2019). Attention-based dropout layer for weakly supervised object localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2219-2228)