# FIRE ACCIDENT DETECTION USING PANOPTIC SEGMENTATION

**Hanif Mohammed A[*1], Aswin C[2], Abinesh B S[3], Abdul Kalam Azad A[4], Dhanish Chander P[5]**

[1,2,3,4,5] Students, Artificial Intelligence and Data Science

[1,2,3,4,5] Panimalar Engineering College

**Abstract**

In India, there are more than 27,000 fire-related fatalities and injury cases per year. In order for the fire engine to reach the incident spot as soon as possible, a person must inform the Fire Rescue Service Department and provide an accurate location. The fire engine will be late in arriving if the informant is uncertain of the precise location. The aim of this study is to use artificial intelligence to spot fires on CCTV cameras while concurrently using GPS to pinpoint the precise location and transferring that information to the Fire Service Department and Traffic police. We will be using a machine learning algorithm to detect fire and send the popup notification to the control center.

**Keywords:** Fire detection, Fire accident detection, Panoptic Algorithm, CCTV Surveillance Camera.

## 1. Introduction

India has experienced a considerable increase in fire incidents, which pose risks to property, human life, and the environment. Homes, businesses, industrial sites, and public areas have all experienced terrible effects from fires, including fatalities, injuries, and monetary losses. India still struggles to avoid and manage fire accidents despite attempts to reduce fire dangers; in order to fully comprehend their impact and find practical mitigation solutions, a thorough investigation is required. Around all over India, around 35 people die over a day due to fire accident and 14000 above die due to fire accident annually.

Machine learning technology has significantly advanced with the usage of panoptic segmentation. With great accuracy, it enables the identification of things in a picture, such fire and smoke. In this instance, the device makes use of CCTV cameras to record images and quickly identify fire accidents. An alarm is automatically sent from the fire detection point to the control room, notifying the workforce to take prompt action.

To confirm that the information provided by the technology is accurate and that it is not a false alarm, the staff members in the control center will manually verify it. This action is essential to avoiding unneeded interruption and lowering false alarms. The relevant authorities, including the police and fire stations, are informed of the fire accident as soon as it is confirmed, along with the location of the accident site. This could shorten the response time and guarantee that the right resources are sent to the scene of the fire accident right away.

This technology can help India's emergency response system operate more effectively overall by reducing response times. This is crucial in a nation like India where there are many fire accidents and the response time is frequently slowed down by traffic jams and other logistical issues. This technology can assist in preventing the spread of the fire, saving lives, and reducing property damage by automatically detecting fire accidents and transmitting information to the control center.

## 2. Existing Methodology

In the existing system, a person needs to call and inform the location of the fire accident spot and the Fire Engine will arrive at the spot and they put off the fire. There is frequently a risk of misinterpretation or inaccurate information being delivered when relying only on humans. The risk of property damage and fatalities rises, emergency responders become disoriented, and response times are prolonged. The existing methodology contains the following drawbacks,

**Human dependence**

The current methods for reporting a fire accident rely heavily on people. As a result, someone must see the accident and notify the authorities. This can cause the incident to be reported later than expected or, in some situations, not at all.

**Lack of reporting**

Even if someone sees an accident involving a fire, they might not always report it to the appropriate authorities. It's possible that they opt to ignore the accident because they don't realise how serious the problem is, they don't know the right steps to take, or all three. This may cause response times to take longer, which may cause more harm and casualties.

**Inaccurate location reporting**

The site of the fire accident may not be precise, which is another drawback of the current techniques. The person reporting the accident might not be able to give the police correct information if they are unfamiliar with the area. As a result, the authorities can take longer to respond and may have trouble locating the scene of the accident.

**Lack of knowledge of reporting procedures**

In some instances, the individual reporting the fire accident might not be aware of the proper steps to take while filing the report. The authorities might not be able to get to the accident scene fast as a result of response time delays. To lessen damage and fatalities, it is crucial that people are informed about the right steps to take when reporting a fire accident.

**Time-consuming**

It can take longer to report a fire accident because every step of the procedure depends on people. This is due to the fact that notifying the authorities of the accident and waiting for their response to the occurrence both take time. In some circumstances, this delay can cause more harm and fatalities.

**No human presence**

Before the authorities are alerted, the entire area can catch fire at the accident scene if no one is there. This could result in catastrophic damage and casualties, as well as make it challenging for the authorities to put out the fire. Our methodology addresses and resolves the aforementioned issues because it is totally automated.

## 3. Related Works

Alexander Kirillov et al [1]. proposed a novel approach to object detection in the given field of view known as Panoptic Segmentation. He points out that the panoptic segmentation's operation is based on the combination of the instance and semantic segmentation principles. For object labeling, he used 4 datasets in his research. They are Cityscapes [2], which includes street view datasets, COCO [3] which includes big datasets on objects and things, Mapillary Vistas [4] which includes road level dataset, and ADE20k [5] includes databases for both indoor and outdoor objects.

Dahun Kim et al. [6] expanded the work of Alexander Kirillov's Panoptic Segmentation [1] by anticipating the objects and giving instance IDs to each and every pixel on the video. Because it separates the video pixel by pixel, Dahun Kim termed this task "Video panoptic Segmentation"[6]. The datasets used for Panoptic Segmentation are incompatible with Video Panoptic Segmentation. Therefore, he employed the modified VIPeR (Viewpoint Invariant Pedestrian Recognition) [7] dataset and Cityscapes [3] dataset for Video Panoptic Segmentation (VPS).

Rohit Mohan et al. [8] developed Efficient PS (EPS), which is more effective and precise method of detection in panoptic segmentation. There are two steps to it. The first stage of semantic segmentation [9] uses a convolutional neural network (CNN) to provide a class label to each pixel in the image. In the second stage, instance segmentation [10], which recognizes and classifies each individual object in the image, is carried out using a separate CNN. Because it provides a unified framework for semantic and instance segmentation, uses efficient network architecture, and optimization techniques, and streamlines the image processing workflow while requiring fewer additional post-processing steps, efficient panoptic segmentation has higher accuracy levels than panoptic segmentation.

Waqar S. Qureshi et al. [12] suggested a technique for detecting fire in a scene using a combined video approach, which tries to determine the existence of fire in a scene by analysing dynamic changes in the look and behaviour of the flames. The suggested approach consists of three phases. To begin, the input video is split into individual frames. To identify any changes in the scene, motion detection techniques are used. While this algorithm [13] detects fire in the video sections, it may be less accurate. To increase the method's accuracy, the next step is to use a colour analysis algorithm [14] to detect places with fire by utilising the distinct characteristic colours of fire, such as red, orange, and yellow. Nonetheless, Because the flames are not consistent in shape, the ML model may still fail to appropriately detect the fire. To remedy this, the third phase is detecting the flames' form features using edge detection [15] and other shape analysis techniques. This can assist in identifying the fire from other items in the scene and increase the fire detection system's accuracy. It attempts to develop a more efficient and accurate early warning system for detecting fires by utilising computer vision and image processing techniques.

## 4. Development of the model

### A. Datasets

A dataset is a collection of data which is used for training the data and testing ML models as they provide the model with a set of examples to learn from. The quality of the data and the size of the dataset determine the performance and success rate of the model.

Based on several criteria, including the data source, the data format, and the application domain, datasets can be divided into numerous kinds. A few examples of frequent dataset kinds are:

**Image datasets** These datasets are collections of images which are utilized for computer vision applications like object detection, semantic segmentation, and image classification.

**Text datasets** These datasets are collections of text documents used for language translation, sentiment analysis, and other natural language processing activities.

**Time series datasets** These datasets are collections of data points collected over time at regular intervals, such as stock prices or weather information.

**Audio datasets** These datasets are collections of audio signals which are utilized for applications like music classification, speech recognition, and other audio-related activities.

**Structured datasets** These datasets, which are utilised for functions like regression and classification, are made up of structured data in the form of tables.

In this approach, the ml algorithm is trained using picture datasets. We will train the ML model for detecting fire accidents using the image datasets. In accordance with the sensitivity of the data and the planned use of the model, datasets may be privately produced or publicly available predefined datasets. It is crucial to make sure that the+ datasets are gathered and used in an ethical and responsible manner.

**B. Training the model**

Training a machine learning segmentation model involves several steps, including data preparation, model architecture selection, training, and evaluation. Here is an overview of the process:

**Data preparation** It is the initial step in the segmentation model training process. In order to do this, a dataset of images with labelled semantic and instance segmentation masks must be gathered. To make sure that the data is consistent, the photos must be pre-processed to normalize the size and colour balance.

**Model architecture selection** Selecting a suitable segmentation model architecture that can learn to divide images into semantic classes and instances is the next stage. The size, complexity, and availability of the computational resources should all be taken into consideration while choosing the architecture.

**Training** The model can be trained after the data and model architecture are ready. A pre-trained model may be adjusted during the training process, or a new model may be created, employing methods like data augmentation and transfer learning to boost the model's performance.

**Evaluation** Following training, the model's effectiveness needs to be evaluated on a different validation set. The model architecture and training parameters can be adjusted based on the evaluation results to enhance the model's performance.

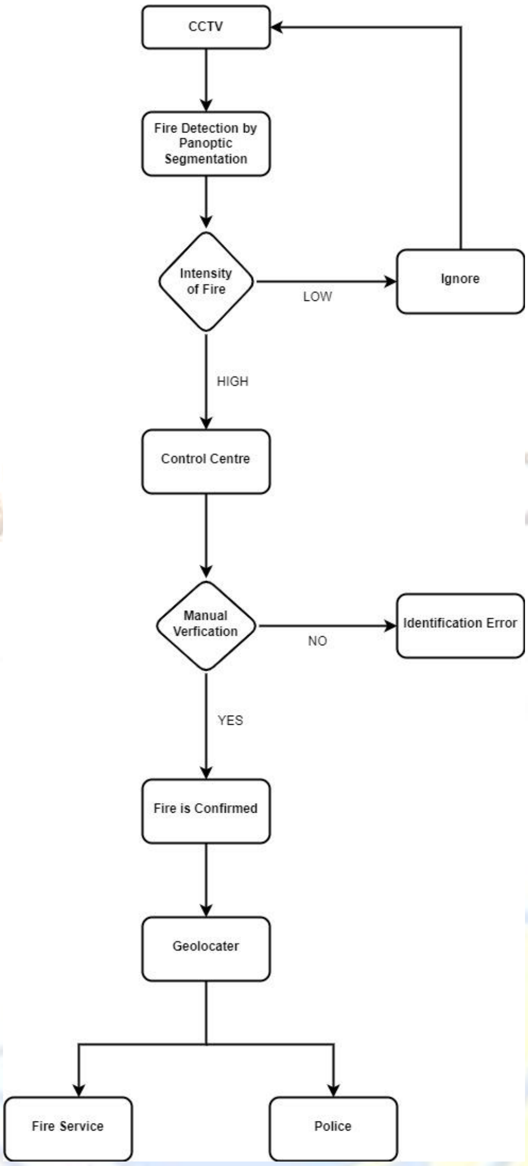
**Deployment** After the success of the model, it can be deployed in a real w0orld environment to carry out segmentation on fresh images or video streams.

**5. Proposed Methodology**

This approach includes a number of elements, such as GPS, a control centre, CCTV surveillance cameras, and a machine learning system for fire accident detection. CCTV cameras record streams of the scene's activity and send them to the control centre for analysis. Real-time fire accident detection is performed by the machine learning algorithm after processing the video feed. The GPS component is utilized to locate the accident and launch the necessary actions, such as sending out emergency services.

In this method, the CCTV security camera transmits the raw video to the machine learning algorithm. Using the trained dataset, the panoptic segmentation machine learning algorithm divides the raw video material into various classes and locates the fire accident. The alert notification is delivered to the control centre after the fire accident has been identified. The fire accident is manually verified. The fire station and the police station are notified of the fire accident's geolocation after verification of the fire accident.

**6. Architectural Diagram**




**7. CCTV Camera**

CCTV cameras are video surveillance cameras that record and transmit video footage of a specified area to a monitoring station or recording device. They are frequently utilised in both public and private settings for security and surveillance purposes. The CCTV cameras in this instance send raw video to the machine learning system for fire accident detection. The algorithm analyses the video and spots any intense fire threats, such as smoke or flames.

**8. Detection of Fire**

Real-time video panoptic segmentation is a method for breaking down a scene into useful objects and regions by analysing live video streams. In order to offer a thorough comprehension of the video stream, this method combines two computer vision techniques, semantic and instance segmentation.

The Semantic Segmentation involves breaking the scene up into areas according to the semantic significance of the things that are present. For instance, different parts of a street scene can be distinguished as being cars, people, and buildings.

Instance segmentation is used for identification and differentiation among specific instances of objects in a video or picture stream. In contrast with semantic segmentation, which assigns the pixel in the image to a certain class ands instance segmentation is concerned with locating each instance of an object in the image. For instance, instance segmentation may recognise and differentiate between each individual car in an image with several cars.

Real-time video panoptic segmentation can give a thorough comprehension of the scene, including the names and locations of objects present as well as the layout and structure of the environment, by integrating these two methodologies.

In this proposed methodology, fire accidents are found using real-time video panoptic segmentation. To recognise the visual cues connected with fire, such as colour shifts or motion patterns, the system uses pre-trained datasets. Whenever a fire is detected by the system, the dataset is compared to the fire's intensity to decide whether the accident qualifies as significant. When a fire is minimal in intensity, it is not taken into account and keep an eye on that; however, when a fire is strong in intensity, the system sends a notification of an alert to the control centre, where the CCTV cameras are continuously watched.

*Fig. In the above image, a typical fire is identified using panoptic segmentation.*

**9. Control Centre**

The control centre is in charge of keeping an eye on the CCTV security cameras that the government has installed. The control centre receives an alert notification when the real-time video panoptic segmentation system spots a fire with high intensity. To validate the true nature of a fire accident, the control centre employees will then manually check the camera footage. The crew will proceed in accordance with established norms, such as calling the fire department or emergency services, if the existence of a fire accident is confirmed. The control centre is essential to the fire accident detection system because it makes sure that the accidents are correctly identified and dealt with as soon as possible.

**10. Geolocator**

The process of detecting a fire accident moves to the next level after the control centre staff has manually confirmed the existence of a fire accident. The closest fire department and police station receive the geographical position of the fire accident and other relevant details via automatic transmission from the system. This is made achievable by the placement of the government-installed CCTV cameras, which enables the precise site of the accident to be determined. The accident can then be controlled before it escalates out of control with the help of the fire and rescue department and police department. This automated procedure assists in ensuring an immediate reaction to fire incidents, possibly minimizing the harm and severity of the incident.

**11. Working Principle**

The algorithm used in our research paper is Panoptic Feature Pyramid Network (PFPN). An extension of an FPN that utilises FPN to produce both instance and semantic segmentations is referred to as a Panoptic FPN. The process begins with an FPN backbone and subsequently adds a branch to carry out semantic segmentation concurrently with the already-existing region-based branch for instance segmentation. The dense-prediction branch is added without disrupting the FPN backbone, making it compatible with current instance segmentation techniques. Following is how the new semantic segmentation branch succeeds in its objective. We execute 3 up sampling stages starting from the deepest FPN level (at 1/32 scale) to produce a feature map at 1/4 scale. Every up sampling stage consists of 33 convolution, group norm, ReLU, and 2 bilinear up sampling. For FPN scales 1/16, 1/8, and 1/4, the same approach is used again with fewer up sampling phases. The outcome is a collection of feature maps with the same 1/4 scale, which are then added up elementally. The per-pixel class labels are produced using a SoftMax final 11 convolution, 4 bilinear up sampling, and final 1 x 1 convolution at the initial image resolution. Besides adding to stuff classes, it also generates a distinct 'other' class for all the pixels in the image that are associated along with objects.

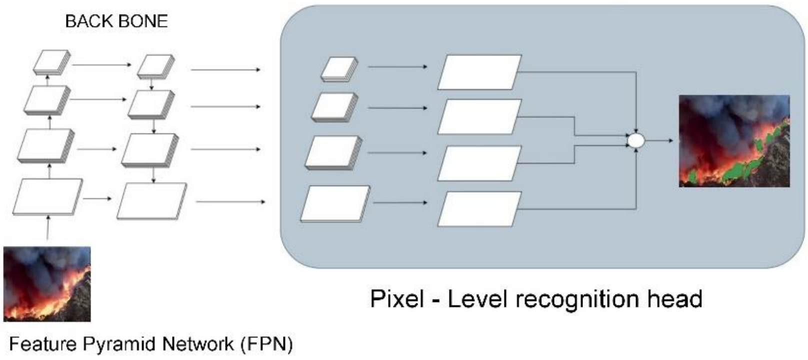



*Fig. The working of the FPFN is described in the above diagram.*

PQ is calculated individually for each class and averaged across courses. As a result, PQ is indifferent to class inequality. The unique matching separates the data for each class. The anticipated and ground truth segments are divided into three categories: true positives (TP), false positives (FP), and false negatives (FN). (FN), corresponding to matched pairs of segments, unmatched predicted segments, and unmatched ground truth segments. Figure 2 illustrates one example. PQ is defined using these three sets.

$$PQ = \frac{\sum_{(p,q)\in TP} IoU(p,q)}{|TP| + \frac{1}{2}|FP| + \frac{1}{2}|FN|}$$

## 12. Algorithm

The technique we proposed uses a panoptic segmentation model in CCTV security cameras to find high-intensity flames. The algorithm is explained in detail below:

**Step 1:** Start

**Step 2:** Prepare or gather the datasets for the segmentation model's training

You need to gather or prepare a dataset of photos or videos that feature fire accidents before training the panoptic segmentation model. This dataset will be used to accurately train the machine to identify fire accidents.

**Step 3:** Use the training datasets to calibrate the panoptic segmentation model.

In this stage, you will use the gathered or produced dataset of photos or videos that involve fire accidents to train the panoptic segmentation model. To recognise the patterns that point to the occurrence of a fire accident, the model will be trained.

**Step 4:** Implement the panoptic segmentation model to the CCTV security cameras.

You will use the panoptic segmentation model in the CCTV security cameras after training it. This will make it possible for the cameras to precisely detect fire accidents.

**Step 5:** Enter the RAW CCTV camera video as input into the panoptic segmentation model.

The panoptic segmentation model will receive the raw video data from the CCTV security camera as input. The model will examine the footage and determine if a fire accident occurred or not.

**Step 6:** If (High Intense Fire Detected)

The model will send a notification to the control room if it discovers a high-intensity fire. The message will inform the staff in the control room to review the CCTV footage and manually confirm the fire accident.

**Step 7:** An employee at the control centre manually verifies the fire accident after notifying the notification.

The control room staff will manually confirm the fire accident in this phase by reviewing the CCTV footage. If the fire accident is verified, they will proceed to step 8 after that.

**Step 8:** If (Fire Accident is Confirmed)

The geolocation of the fire accident site will be given to the police and fire stations once the fire accident is confirmed. They will be able to respond to the fire mishap more swiftly and limit further damage as a result.

**Step 9:** Else

If the control room employee determines that the notification is a false alarm, then the notification message is ignored

End of Algorithm.

This algorithm describes the procedures for creating a system that can identify high-intensity flames in video footage from security cameras. Creating datasets of photos and videos that include examples of both fires and non-fires must come first. Following that, the datasets are utilised to train a panoptic segmentation model, a deep learning technique that can identify and classify various objects and regions in an image or video. Once trained, the model is placed into use on the CCTV cameras, and the raw video footage that is recorded by these cameras is fed into the model as input. The model sends a notification to the control room when it discovers a high-intensity fire. The human operator can respond appropriately after receiving this alert, which informs them that a probable fire has been discovered in the monitored region. A person working in the control room then manually confirms the alert. They will look at the video of the potential fire to determine if one is there. If the fire is confirmed, the system notifies the appropriate authorities, including the fire department and the police department, of the fire's geolocation so they can take prompt action. The alert is ignored if it turns out to be a false alarm. This guarantees that the system is not overloaded with false alarms and helps prevent pointless emergency replies. Overall, this technology can assist in early fire detection and response, thus reducing property damage and potentially saving lives. The technology can alert the right authorities and automate the fire detection process to send out early warnings, allowing them more time to react and take the necessary action.

## 13. Conclusion

In our study, a methodology is proposed with the goal of identifying fire accidents before they become serious. By using this method, the control centre can be informed of the fire accident without requiring human intervention. In comparison to current approaches, our technology is more effective because it requires less time to transmit the information and relies less on people. Compared to existing approaches, this methodology's location detection is more accurate than the location provided by the informant. The authorities can act quickly to control the situation and stop the fire from spreading by locating the fire accident zone. Overall, our approach provides a quick, accurate, and efficient way to identify fire accidents, thereby saving lives and minimizing property damage.

**14. Reference**


[1] Kirillov, Alexander, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollár. "Panoptic segmentation." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9404-9413. 2019.

[2] Cordts, Marius, Mohamed Omran, Sebastian Ramos, Timo Scharwächter, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. "The cityscapes dataset." In CVPR Workshop on the Future of Datasets in Vision, vol. 2. sn, 2015.

[3] Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. "Microsoft coco: Common objects in context." In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, pp. 740-755. Springer International Publishing, 2014.

[4] Neuhold, Gerhard, Tobias Ollmann, Samuel Rota Bulo, and Peter Kontschieder. "The mapillary vistas dataset for semantic understanding of street scenes." In Proceedings of the IEEE international conference on computer vision, pp. 4990-4999. 2017.

[5] Zhou, Bolei, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. "Scene parsing through ade20k dataset." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 633-641. 2017.

[6] Kim, Dahun, Sanghyun Woo, Joon-Young Lee, and In So Kweon. "Video panoptic segmentation." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9859-9868. 2020.

[7] Richter, Stephan R., Zeeshan Hayder, and Vladlen Koltun. "Playing for benchmarks." In Proceedings of the IEEE International Conference on Computer Vision, pp. 2213-2222. 2017.

[8] Mohan, Rohit, and Abhinav Valada. "Efficientps: Efficient panoptic segmentation." International Journal of Computer Vision 129, no. 5 (2021): 1551-1579.

[9] O'Shea, Keiron, and Ryan Nash. "An introduction to convolutional neural networks." arXiv preprint arXiv:1511.08458 (2015).

[10] Guo, Yanming, Yu Liu, Theodoros Georgiou, and Michael S. Lew. "A review of semantic segmentation using deep neural networks." International journal of multimedia information retrieval 7 (2018): 87-93.

[11] Hafiz, Abdul Mueed, and Ghulam Mohiuddin Bhat. "A survey on instance segmentation: state of the art." International journal of multimedia information retrieval 9, no. 3 (2020): 171-189.

[12] Qureshi, Waqar S., Mongkol Ekpanyapong, Matthew N. Dailey, Suchet Rinsurongkawong, Anton Malenichev,and Olga Krasotkina. "QuickBlaze: early fire detection using a combined video processing approach." Fire technology 52 (2016): 1293-1317.

[13] Manzanera, Antoine, and Julien C. Richefeu. "A new motion detection algorithm based on Σ–Δ background estimation." Pattern Recognition Letters 28, no. 3 (2007): 320-328.

[14] Shao, Jing, Guanxiang Wang, and Wei Guo. "An image-based fire detection method using color analysis." In 2012 International Conference on Computer Science and Information Processing (CSIP), pp. 1008-1011. IEEE, 2012.

[15] Qiu, Tian, Yong Yan, and Gang Lu. "An autoadaptive edge-detection algorithm for flame and fire image processing. " IEEE Transactions on instrumentation and measurement 61, no. 5 (2011): 1486-1493.