Forest Fire Prediction Using Machine Learning Techniques and Background Subtraction

Project submitted to the

SRM University – AP, Andhra Pradesh

for the partial fulfillment of the requirements to award the degree of

Bachelor of Technology

In

Computer Science and Engineering

School of Engineering and Sciences

Submitted by

Devisetty Sai Tharun - AP20110010078

Omkar Subhash Ghongade - AP20110010090

Yella Indrani - AP20110010125

Under the Guidance of

Dr.Manikandan V.M.

SRM University–AP

Neerukonda, Mangalagiri, Guntur

Andhra Pradesh – 522 240

[November, 2022]

Certificate

Date: 13-Dec-22

This is to certify that the work present in this Project entitled "**Forest Fire Prediction Using Machine Learning Techniques and Background Subtraction**" has been carried out by **Omkar Subhash Ghongade, Yella Indrani, Devisetty Sai Tharun** under my/our supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of Bachelor of Technology/Master of Technology in **School of Engineering and Sciences**.

Supervisor

(Signature)

Prof. / Dr. Manikandan V.M.

Designation: Asst. Professor

Affiliation: SRM University-AP

Co-supervisor

(Signature)

Prof. / Dr. [Name]

Designation,

Affiliation.

Acknowledgements

We are ever so thankful to all of those who have contributed to the successful completion of this project in any way. It is because of their dedication, hard work, patience and guidance that this has been achieved. A special thanks to our mentor who has been with us from the beginning, offering his advice and support. He was generous enough to give us advice and guide us throughout our project work. We are also immensely grateful to our friends for their support, as well as anyone else who has contributed to this project. Without them, we would not have been able to reach the end goal. The project has come such a long way since it was first conceived, and that is due in large part to the effort of the people who have been involved. We are conscious of the fact that a project of this magnitude could not have been successful without the help of many individuals. We are extremely grateful to all those who have aided us in any way, whether it be directly or indirectly. We thank you all for your support. In summary, we would like to express our sincerest appreciation to all those who have been a part of this project. Without their help, we would not have been able to complete it. We are humbled by the amount of support we have been given, and we thank all those involved for their invaluable contributions.

Table of Contents

Abstract

Image processing is a powerful tool for improving both machine and human interpretation and extracting information from images for further classification. Forest fire detection is an important aspect of ecosystem conservation and climate change management. Fires can cause significant losses of trees and drastically increase CO2 levels in the atmosphere. To minimize these losses, it is important to detect forest fires quickly and accurately. This is where image processing and other technologies such as wireless sensor networks and unmanned aerial vehicles (UAVs) can be used. Image processing can be used to detect forest fires more quickly and accurately than sensor technology. Sensor technology requires particles to reach sensors before an alarm is triggered, which means that fires can have already spread considerably before they can be detected. In contrast, image processing can detect fires sooner. This paper presents a new, promising solution for fire detection using image-processing algorithms. The image sequence extracted from the input video file undergoes a series of efficient operations to detect whether a fire is present or not. This system has been shown to be more effective than sensor technology or wireless technology. Image processing algorithms are used to analyse the image data and detect the presence of a fire. By analysing the colour and movement of the pixels, the algorithms can detect changes in the image sequence that indicate the presence of a fire. This data can then be used to accurately detect a fire and alert the relevant personnel. The image sequence can also be used to estimate the intensity and spread of the fire, helping responders to take appropriate action. Image processing is a powerful tool for improving both machine and human interpretation, and it is essential for effective forest fire detection. By using image processing algorithms to detect fires, we can quickly and accurately detect fires and minimize the damage they cause. As such, image processing can be a key part of any fire detection system.

Index Terms: object detection; video analysis; background subtraction; RGB; YCbCr;

Abbreviations

RGB: Red Green Blue

YCbCr: Luminance Component Blue Component Red

SVM: Support Vector Machine

XGBoost: Extreme Gradient Boosting

FGMask: Face Ground Masking

List of Tables

List of Figures

List of Equations

1. Introduction

Forests play an important role in the environment, providing habitats for various species, storing carbon, and contributing to the global water cycle. Unfortunately, forest fires are a major threat to these vital ecosystems. In recent years, the increasing incidence of forest fires has had disastrous consequences, leading to loss of habitat, damage to valuable resources, and increased risk of air pollution.

One important aspect of forest fire management is detecting and monitoring fires in their early stages. This allows firefighters to respond quickly and effectively, which can help prevent the fire from spreading and minimize the damage it causes. There are several different methods for detecting forest fires, including:

Human observation: This involves having trained personnel on the ground or in aircraft look for signs of fire, such as smoke or flames.

Satellite imagery: Satellites orbiting the Earth can be equipped with sensors that can detect the heat and infrared radiation emitted by fires. This can provide real-time information about the location and size of fires.

Automatic detection systems: These are devices that are installed in strategic locations in forests, such as along roads and near human settlements. They use sensors to detect heat, smoke, or other indicators of fire and can automatically alert authorities when a fire is detected.

To combat this issue, researchers have developed various machine learning techniques for predicting forest fires. These techniques utilize data collected from the environment, such as temperature, moisture, wind speed, and fuel type, to provide early warning of fire danger and alert decision-makers to take preventive action. One such technique is background subtraction, which utilizes the differences between the past environment and current conditions to detect any changes that might indicate a fire.

Background subtraction is a popular machine learning technique as it is simple to implement and offers high accuracy and speed of operation. The technique works by comparing the current environment to a reference image or video, and detecting any variations in the environment. This comparison can be used to identify fire-related changes, such as the presence of smoke or flames, and alert the appropriate personnel.

This paper will discuss the use of machine learning techniques and background subtraction for predicting and controlling forest fires. It will begin by providing an overview of the technologies involved and the data they use. It will then discuss the advantages and disadvantages of each strategy and how they can be combined to create an effective warning system. Finally, it will present case studies that demonstrate the use of these techniques in practice. By the end of this paper, readers will have a better understanding of how machine learning and background subtraction can be used to predict and control forest fires.

2. Methodology

We have used various Supervised Machine Learning Models like Logistic Regression, Naïve Bayes, Support Vector Machine, K Nearest Neighbours, Decision Tree and XGBoost to get the accuracy of the forest fire.

Logistic Regression model: Logistic regression is a type of supervised machine learning algorithm used to predict categorical values such as binary values (yes/no, 0/1) from a set of independent variables. It is a popular method for classifying data into two categories, such as predicting whether a customer will churn or not. It is also commonly used in predictive analytics for predicting the probability of an event occurring. Logistic regression works by using a logistic function (also known as a sigmoid function) to map the input data to an output value between 0 and 1. This output value can then be used to classify the data into two categories. Logistic regression is a powerful algorithm that can be used to model complex relationships between the independent and dependent variables. It is also one of the most widely used algorithms in predictive analytics and is often used in marketing campaigns to target customers with tailored offers. This method can be applied to predict whether a selective audience will active or not for the given issue in marketing perspective.

$$P = \frac{e^{a+bx}}{1 + e^{a+bx}} \tag{1}$$

Naive Bayes: Naive Bayes is a probabilistic machine learning algorithm that is used for classification tasks. It is based on the Bayes Theorem, which states that the probability of an event occurring is equal to the probability of the event's causes and conditions occurring. Naive Bayes is known to be a fast and reliable algorithm for classification tasks and is often used for text classification and spam filtering. It is called "naive" because it assumes that all input features are independent of each other, which is rarely the case in real-world data. Despite this assumption, Naive Bayes tends to perform well in practice and is the go-to algorithm for many classification tasks.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{2}$$

Support Vector Machine (SVM): Support Vector Machines (SVMs) are a type of supervised machine learning algorithm that can be used for both classification and regression tasks. SVMs are based on the idea of finding a hyperplane that best divides a dataset into two classes. To separate the data, the SVM algorithm creates a line (or hyperplane) with the largest amount of margin between the two classes. This line can then be used to classify new data. SVMs are powerful because they have the ability to adjust to different types of data and to create complex non-linear decision boundaries. Additionally, they are robust to overfitting, meaning they can handle high-dimensional data without sacrificing accuracy.

$$f(x) = w^T x + b \tag{3}$$

K Neighbours Classifier: K Neighbours Classifier is a supervised learning algorithm used for classification and regression tasks. It is a non-parametric method that uses a distance-based technique to determine the class of a data point by looking for the K nearest neighbours and assigning the class of the majority of those neighbours. The algorithm is simple to implement and works well in many cases, but requires careful consideration of the input data to ensure that the classifications are accurate. It is also important to choose an appropriate value of K, as too high a value can led to overfitting while too low a value can led to underfitting. It usually stores the given dataset and performs the action on it during the time of classification. It is more effective towards data set which contain more reviews and robust to the given noisy training data.

$$d(x, y) = \sqrt{\Sigma_{i=1}^{n} (x_i - y_i)^2} \tag{4}$$

XGBoost: XGBoost is a machine learning algorithm that uses decision trees as its base model. It is an extreme gradient boosting technique that is used for solving problems related to regression, classification, and ranking. XGBoost works by sequentially adding trained decision tree models to a frame work that is created in order to improve the accuracy of the model over time. It is also capable of handling user-defined prediction problems.

$$f(x) = f(a) + f'^{(a)}(x - a) + 0.5(f``(a)(x - a)^2) \tag{5}$$

Decision Tree Model: Decision trees are a type of supervised learning technique used for both classification and regression problems. They are a tree-based classifier that use internal nodes to represent features of a dataset, branches to represent decision rules, and leaf nodes to represent the outcome. Decision trees are a great way to mimic the way humans make decisions and are easy to understand and analyse. They have a tree-like structure that makes it easy to view the various outcomes and the various paths that lead to those outcomes. Decision trees are helpful in problem-solving scenarios as they provide an easy way to visualize the various options available and the likely outcomes.

$$Gini = 1 - \Sigma_{i=1}^{c}(p_i)^2 \tag{6}$$

3. Model Implementation & Accuracies

The dataset we have considered for working is in text format in csv files. The dataset contains 13 different attributes on the basis of which the data pre-processing is done to do the test and train split. We have implemented various supervised machine learning models with various train and test ratios like 50:50, 66:34, 70:30 and 80:20.

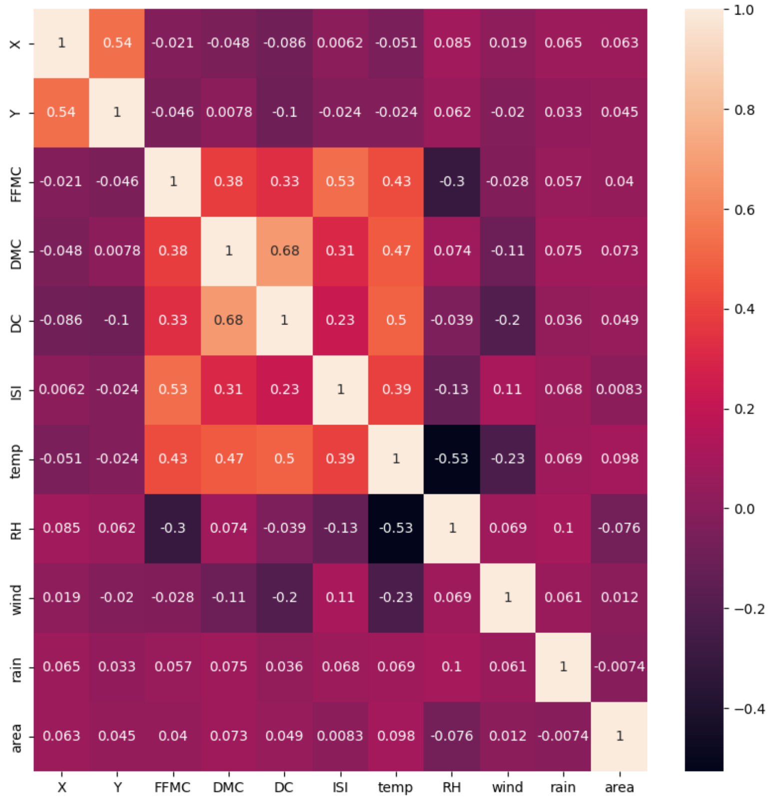

Figure 1

During Data Pre-processing we came across a heat map which displays the co-relation between the attributes in the dataset.

There are various machine learning algorithms that can be used for forest fire detection, including decision trees, random forests, and support vector machines. Each of these algorithms has its own strengths and weaknesses, and choosing the right one will depend on the specific characteristics of the data being used for training. The accuracy of a forest fire detection model will depend on a number of factors, including the quality of the training data, the choice of machine learning algorithm, and the model's hyperparameters. In general, however, well-trained models can achieve high levels of accuracy, with some models able to predict the likelihood of a forest fire with over 60% accuracy. And the accuracies achieved by us also lies in the nearby range of the 60%.

Model	Accuracies	
Logistic Regression	50:50	55.21 %
	66:34	52.84 %
	70:30	51.28 %
	80:20	50.00 %
Naive Bayes	50:50	53.67 %
	66:34	50.00 %
	70:30	50.64 %
	80:20	47.12 %
SVM	50:50	54.83 %
	66:34	52.27 %
	70:30	50.64 %
	80:20	46.15 %
K Nearest Neighbours	50:50	50.97 %
	66:34	52.84 %
	70:30	50.00 %
	80:20	44.23 %
Decision Tree	50:50	52.51 %
	66:34	54.55 %
	70:30	52.56 %
	80:20	56.73 %
XGBoost	50:50	53.28 %
	66:34	52.84 %
	70:30	54.49 %
	80:20	57.69 %

Table 1

4. Forest Fire Detection from Video

4.1 Background Subtraction:

Background subtraction is a powerful tool in computer vision that is used to separate foreground objects from background in video sequences. This is done by first training a model on a set of videos or images, where the background is known, and then using that model to identify and remove the background from new videos or images. This technique is commonly used in surveillance systems to extract moving objects from the background and can also be used for other purposes such as video editing and augmented reality. The technique works by first training a model on the known background data. Then, the model is used to identify the background in new videos or images. Once the background is identified, it can be removed from the video sequence, leaving only the foreground objects. This allows the system to track the movement of objects in the video sequence and allows for a more detailed analysis. Background subtraction is a highly effective technique that can be used in a variety of applications. It is useful in surveillance systems to detect movement, in video editing to add or remove objects from the background, and in augmented reality systems to add virtual objects to the scene. This makes background subtraction a versatile and powerful tool in computer vision.

The technique of background subtraction is commonly used in image and video processing to detect moving objects in a scene. It is based on comparing the current frame with a reference image, or background, and subtracting it to create a foreground mask. This mask can then be used to identify the objects of interest in the scene. Background subtraction is frequently used in security and surveillance applications, to help automate the detection and tracking of objects in a video stream. However, it is also employed in other fields such as robotics and computer vision, to enable machines to better comprehend and interact with their environment. This technique can be used to extract important information from videos, and to allow machines to interact with their environment in a more meaningful way.

Figure 2.1

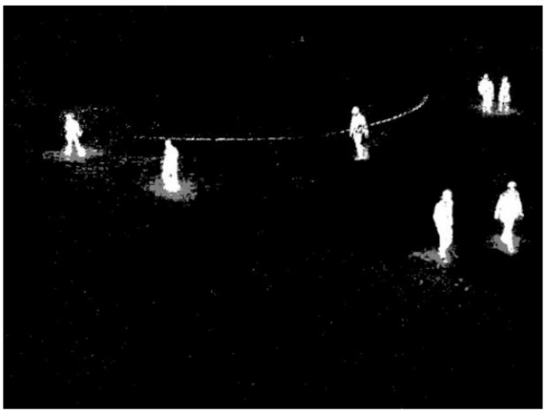

Figure 2.2

The figure 2.1 is the original image which is passed before it is sent to the model for the background subtraction and figure 2.2 is the image section which is obtained after the successful background subtraction. This Background Subtraction along with the motion detection will be helpful for us in detecting the forest fire and also easily detect the reason of the forest fire using the motion detection.

4.2 Motion Detection Using Background Subtraction:

Motion detection using background subtraction is a powerful technique for identifying and tracking moving objects in videos. This method involves taking a reference frame, usually the first frame of the video, and comparing each subsequent frame with it to identify pixels that have changed between the two frames. These pixels are assigned to moving objects in the scene. The reference frame is updated to incorporate the new pixels, and the process is repeated for every frame of the video. This allows for the identification of moving objects in real-time, allowing for accurate and speedy tracking. As this process is repeated, the reference frame continues to evolve, allowing for improved accuracy as more frames are processed. Motion detection using background subtraction is useful in a wide range of applications, from security surveillance to motion-based video games.

In order to implement this background subtraction technique, the first step is to capture a video sequence and extract individual frames. A background reference frame is then generated either by using the first frame of the video or by taking the average of multiple frames. Subsequently, each frame of the video is compared to the background reference frame to identify pixels that have changed. These pixels represent moving objects in the scene, and the background reference frame is updated accordingly. This process is then repeated for each frame in the video to track the movement of the objects over time. By doing so, the background subtraction technique is able to detect and track objects in a video sequence, enabling us to analyse the motion and behaviour of these objects.

Background subtraction is a powerful method for motion detection due to its ability to handle changes in lighting and background clutter, which can make it difficult for other motion detection methods. Its simplicity also makes it a great option for applications such as surveillance, robotics, and video analytics because it can be implemented quickly and run in real-time. Additionally, it is relatively robust and can be used to identify objects in a scene, even when they are partially obscured. This allows it to distinguish objects from their backgrounds and accurately detect motion. Furthermore, background subtraction can be used to identify objects with different sizes and shapes, making it ideal for applications such as security systems and face recognition. Overall, background subtraction is an efficient and effective method for motion detection, making it a great choice for a variety of applications.

In the figure 2.1 and 2.2 if we imagine the people are in motion then in the background subtracted video which will be generated, it will have the white coloured portion where ever there is any kind of motion.

4.3 Masking:

Masking is a technique that can be used in forest fire detection to isolate specific areas of interest in an image or video. It involves creating a mask, which is a binary image that specifies which pixels in the original image should be included or excluded in the analysis. In the context of forest fire detection, masking can be used to exclude pixels that correspond to non-forest areas or known non-fire objects, such as buildings or roads. This can help to reduce false positives and improve the accuracy of the fire detection algorithm. Masking can be useful in forest fire detection because it allows the algorithm to focus on the areas of the image or video that are relevant to the task at hand. It can also help to reduce the computational complexity of the algorithm by excluding pixels that do not need to be considered. This can improve the speed and efficiency of the fire detection process. In the forest fire detection, we are using FG Masking technique among the various masking techniques like Pattern Masking, Spectral Masking, etc.

4.4 FGMask:

FGMask is a type of binary mask that is frequently used to identify objects in the foreground of an image or video. The acronym stands for "foreground mask," and it is typically employed in forest fire detection. FGMask can be beneficial for forest fire detection as it enables the algorithm to concentrate on the objects in the forefront, which are more probable to be pertinent to the task. Additionally, it can minimize the number of false positives by omitting pixels that relate to the background or other objects that are not related to fires. This can enhance the precision of the fire discovery algorithm. FGMask is a computer vision technique used for forest fire detection which utilizes an algorithm to detect smoke and other elements that may indicate a fire by differentiating the foreground from the background in an image. This technique involves analysing the colour and texture of the image, followed by a combination of thresholding and morphological operations to separate the fire from the surroundings. It is based on the Mask R-CNN architecture and is used to identify and segment foreground objects from a given image. FGMask uses a region proposal network (RPN) to generate object proposals and then a region-of-interest pooling layer to extract features from each object proposal. These features are then used to classify each proposal as foreground or background. The output is a binary mask that indicates the location of the foreground object. FGMask is a crucial tool for automated forest fire detection, as it can accurately distinguish between a fire and other objects in the scene.

4.5 Gaussian Blur:

Gaussian blur is a widely used image processing technique that is employed to reduce noise in images. It is often used to pre-process an image prior to carrying out further image analysis. In the context of forest fire detection, Gaussian blur can be utilized to smooth out an image of a forest and make it easier to identify features such as smoke, fire, trees, and other objects in the scene. Additionally, it can be used to reduce the noise in an image, making it simpler for a computer vision algorithm to spot and classify the objects in the scene. Gaussian blur also helps to enhance the contrast between foreground and background objects in an image, which can make it easier to identify smoke or fire in a scene. Furthermore, it can help to reduce the effects of camera shake, aiding to decrease false positives from an automated fire detection system. By using Gaussian blur, forest fire detection systems can be more precise and dependable.

$$G(x,y) = \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{7}$$

4.6 Threshold Elite & Dilute:

Threshold erode dilute is a process used for the detection of forest fires. The process compares the brightness of smoke to the brightness of clouds in the image. If the smoke is brighter than the clouds, then it is detected and an alert is sent. This process can also be used to reduce false alarms that may be triggered by clouds. The image is first eroded by removing the edges of the image, then it is diluted to reduce the brightness of the clouds. Finally, a threshold is set for the brightness of the smoke to determine if it is brighter than the clouds. If the threshold is met, the smoke is detected and the alert is sent. This process is used to ensure that only smoke is detected, thereby reducing false alarms.

4.7 Results:

The below images are the screenshots of the final output video received from the original video. The coloured Image is the screenshot of the original video and the black and white Image is the screenshot of the final video obtained by implementing gaussian blur, FGMask and others on the original video.


UltimateChase.com
UltimateChase.com
UltimateChase.com
UltimateChase.com

Figure 3

5. Conclusion

A forest fire detection system is an invaluable tool for preventing and responding to forest fires. The use of such a system can help to minimize the damage caused by forest fires and save human and natural lives. With an effective forest fire detection system, it is possible to detect forest fires in their early stages, improving the chances of containing and suppressing the fire before it grows too large and destroys larger areas.

Early detection of forest fires is essential in order to prevent them from spreading and causing damage. The use of satellite technology, aerial surveillance, and ground-based observation are common methods for detecting forest fires as soon as possible. These systems help to identify the location of a forest fire and alert the responsible authorities who can then respond with the appropriate measures to contain and extinguish the fire. Furthermore, preventative measures such as maintaining a defensible space around buildings and regularly clearing dead vegetation can help to reduce the risk of forest fires and make them easier to control if they occur. Early detection and prevention are both crucial for protecting the environment and human communities from the devastating effects of forest fires.

In order to fully benefit from a forest fire detection system, it is important to ensure that it is regularly maintained and kept up to date. This includes taking time to replace old or damaged components, as well as keeping the system connected to any relevant fire databases or satellite imaging services. Additionally, it is important to have a well-prepared response plan in place to respond to any detected fires.

In conclusion, a forest fire detection system is an extremely valuable tool for preventing and responding to forest fires. With regular maintenance and an effective response plan in place, it is possible to benefit from the system and mitigate the damage caused by forest fires.

6. Future Work

This report is based on the working of the forest fire dataset which is giving the accuracy of the various algorithms on the various test and train ratio. Based on the experimental analysis, we estimated that the supervised learning methods namely Decision Tree and XGBoost are concluded as standard learning techniques for forest fire prediction accuracies. Within the supervised learning, based on the experimental analysis, we conclude that XGBoost is providing an excellent accuracy when compared with other classification algorithms. The proposed work is the extension of few of the methods for detecting the forest fire using 2 methods. In first we have used the dataset to train it on the various supervised machine learning models to predict forest fire on the basis of longitude, latitude, temperature and many more similar factors. In second method we have used video to detect fire. We have used Motion Detection, Background Subtraction, Face Ground Masking, Gaussian Blurring and Threshold Erode Dilute to get to know where fire is there. This work can further be extended to do various tasks like identifying the suspicious activities in the forest and all the things.

References

[1] M. A. I. Mahmoud and H. Ren, "Forest fire detection and identification using image processing and SVM," Journal of Information Processing Systems, vol. 15, no. 1, pp. 159-168, 2019. DOI: 10.3745/JIPS.01.0038.

[2] Töreyin, B.U.; Dedeoğlu, Y.; Güdükbay, U.; Cetin, A.E. Computer vision based method for real-time fire and flame detection. Pattern Recognit. Lett. 2006, 27, 49–58.

[3] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, "A Forest Fire Detection System Based on Ensemble Learning," Forests, vol. 12, no. 2, p. 217, Feb. 2021, doi: 10.3390/f12020217.

[4] P. Barmpoutis, P. Papaioannou, K. Dimitropoulos, and N. Grammalidis, "A Review on Early Forest Fire Detection Systems Using Optical Remote Sensing," Sensors, vol. 20, no. 22, p. 6442, Nov. 2020, doi: 10.3390/s20226442.

[5] V. Vipin, "Image processing based forest fire detection," International Journal of Emerging Technology and Advanced Engineering, vol. 2, pp. 87–95, 2012.

[6] C. E. Premal and S. S. Vinsley, "Image processing based forest fire detection using YCbCr colour model," in Proceedings of the 2014 International Conference on Circuits, Power and Computing Technologies, ICCPCT 2014, pp. 1229–1237, 2014.

[7] C. Thou-Ho, W. Ping-Hsueh, and C. Yung-Chuen, "An early fire-detection method based on image processing," in Proceedings of the 2004 International Conference on Image Processing, ICIP '04, vol. 3, pp. 1707–1710, 2004.

[8] Tong Yang, Sheng Xu, Weimin Li, Haibin Wang, Guodong Shen, Qiang Wang, "A Smoke and Flame Detection Method Using an Improved YOLOv5 Algorithm", 2022 IEEE International Conference on Real-time Computing and Robotics (RCAR), pp.366-371, 2022.

[9] Yi-Chun Chen, Halim Fathoni, Chao-Tung Yang, "Implementation of Fire and Smoke Detection using DeepStream and Edge Computing Approachs", 2020 International Conference on Pervasive Artificial Intelligence (ICPAI), pp.272-275, 2020.

[10] Xuanxuan Hong, Wei Wang, Quanli Liu, "Design and Realization of Fire Detection Using Computer Vision Technology", 2019 Chinese Control And Decision Conference (CCDC), pp.5645-5649, 2019.