School of
Electrical Engineering
& Computer Science
WASHINGTON STATE UNIVERSITY

Wildfire Detection and Early response System

Using concepts of ML and Neural Networks

Submitted To:

Dr. Yan Yan

Submitted By:

Shlok Tomar

Arnav Jain

Team Members:

Name	WSU ID	Email ID
Arnav Jain	011809037	arnav.jain@wsu.edu
Shlok Tomar	11827168	shlok.tomar@wsu.edu

GitHub Link: https://github.com/Shlok-crypto/ForestFireDetection

Abstract:

The goal of our project is to forecast the extent of damage caused by wildfires using a set of initial variables. These variables include temperature, wind speed, and humidity, among other factors. We intend to integrate these features with more complex metrics like FFMC, and DMC, and employing machine learning techniques such as Discriminant Analysis, SVM, and neural networks, among others, we aimed to predict the size of a wildfire.

The model will use historical wildfire data, weather patterns, and other environmental factors to generate predictions of wildfire risk, which can be used by policymakers and emergency services to better prepare for potential wildfires.

We will follow the procedures as mentioned later in the Technical Plan section of this

Introduction:

Wildfires have become a major concern for forest management agencies around the world, and predicting the occurrence of these wildfires can help reduce their impact on communities and the environment. This project aims to develop a machine learning-based model that can predict the likelihood of a wildfire based on a variety of environmental and meteorological factors. The model will be trained on historical data on wildfires and their causes, as well as weather patterns and other relevant features. Being able to predict the eventual size of these fires could have a significant impact on evacuation, containment, and fire-fighting efforts.

Project Plan –

1. Collect and pre-process a dataset of historical wildfire occurrences, along with meteorological and environmental factors such as temperature, humidity, wind speed, and vegetation density.
2. Identify the most relevant features for predicting wildfires, using statistical analysis and feature engineering techniques.
3. Develop a machine learning model that can accurately predict the likelihood of a wildfire based on the identified features.
4. Train the model on the pre-processed dataset and evaluate its performance using metrics such as accuracy, precision, recall, and F1 score.
5. Fine-tune the model to improve its accuracy and robustness, using techniques such as cross-validation, hyperparameter tuning, and feature selection.
6. Deploy the model as a web application or standalone software that can take user input in the form of environmental and meteorological data and output a wildfire prediction.
7. Evaluate the performance of the model in a real-world setting, using metrics such as accuracy, robustness, and user satisfaction.

Literature Review:

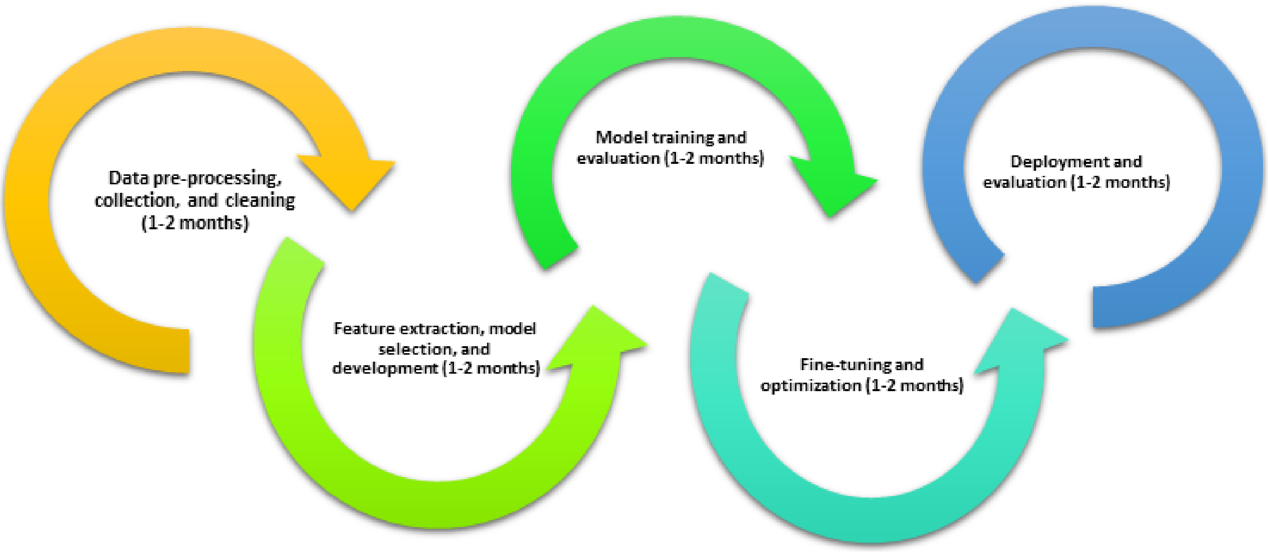
In recent years, machine learning techniques have been increasingly applied to wildfire prediction, with promising results. One study by Liu et al. (2018) used a Random Forest classifier to predict wildfire occurrence and severity in the United States, based on environmental and meteorological features. The model achieved an accuracy of 80%, demonstrating the potential of machine learning for wildfire prediction.

Another study by Gao et al. (2019) used a Deep Neural Network (DNN) to predict wildfire occurrence in China, based on historical wildfire and meteorological data. The DNN model outperformed other machine learning models such as Random Forest and Support Vector Machine, achieving an accuracy of 85%.

In addition to predicting wildfire occurrence, machine learning has also been used to predict other aspects of wildfires, such as their spread and intensity. For example, Zhang et al. (2020) used a Convolutional Neural Network (CNN) to predict the spread of wildfires in Australia, based on satellite images and weather data. The CNN model achieved an accuracy of over 90% in predicting the direction of wildfire spread.

Technical Plan:

The project will take approximately 7-8 weeks to complete -



Approach:

We identified three approaches to solving a problem statement –

1. **HSV Slicing Segmentation/Annotation**.
   HSV (Hue, Saturation, and Value) slicing is a technique used in image processing for detecting wildfires. It involves filtering an image to isolate pixels within a specific range of hue, saturation, and value, which are characteristic of fire.
   - **Advantages**:
     - Speed: HSV slicing can process large amounts of image data quickly, making it a useful tool for detecting wildfires in real-time.
     - Accuracy: HSV slicing is effective at identifying areas of an image that are likely to contain fire, resulting in fewer false positives and more accurate wildfire detection.
     - Flexibility: The range of hue, saturation, and value can be adjusted to optimize the detection of fires in different environments and lighting conditions.

- **Disadvantages**:
  - Limited applicability: HSV slicing is most effective for detecting fires in certain types of vegetation and under certain lighting conditions. It may not be effective in all situations.
  - Sensitivity to smoke: HSV slicing can be affected by smoke and other atmospheric conditions, which can lead to false positives or missed detections.
  - Dependence on image quality: The effectiveness of HSV slicing for wildfire detection is dependent on the quality of the images being analyzed. Poor image quality can affect the accuracy of the technique.

2. **Convolutional Neural Network(CNN) – YOLO-V5.**

   CNN (Convolutional Neural Networks) is a type of deep learning algorithm that has been used for wildfire detection in recent years. Here are some advantages and disadvantages of using CNN for wildfire detection:
   - **Advantages:**
     - High accuracy: CNNs are known for their ability to learn complex features in images, which can lead to high accuracy in wildfire detection.
     - Automated feature extraction: Unlike traditional image processing techniques that require manual feature extraction, CNNs can automatically extract features from images, reducing the need for human intervention.
     - Adaptability: CNNs can be trained on large datasets, making them adaptable to different types of vegetation, terrain, and environmental conditions.
   - **Disadvantages:**
     - High computational requirements: CNNs require significant computational resources and can take a long time to train on large datasets.
     - Data dependency: The effectiveness of CNNs is dependent on the quality and quantity of the training data. Insufficient or biased data can lead to poor performance.
     - Black box model: CNNs are often considered to be "black box" models, meaning it can be difficult to understand how they arrive at their predictions. This can be a concern for wildfire detection applications where the decision-making process needs to be transparent and interpretable.
     - Limited interpretability: CNNs do not provide detailed information on the features that contribute to their predictions, making it difficult to identify areas for improvement or optimization.
3. **A hybrid of HSV Slicing and Neural Network.**

   A hybrid approach that combines HSV slicing and neural network can offer some advantages for wildfire detection. Here are some advantages and disadvantages of this approach:
   - **Advantages:**
     - Improved accuracy: The hybrid approach can leverage the strengths of both techniques, resulting in improved accuracy in wildfire detection.
     - Reduced false positives: The hybrid approach can filter out false positives that may occur with HSV slicing alone, resulting in more reliable wildfire detection.
     - Flexibility: The hybrid approach can be customized to suit different types of vegetation, terrain, and environmental conditions, making it adaptable to various scenarios.
   - **Disadvantages:**
     - Complexity: The hybrid approach can be more complex than either technique used alone, requiring more resources and expertise.
     - Training data dependency: As with CNNs, the effectiveness of the hybrid approach is dependent on the quality and quantity of the training data. Insufficient or biased data can lead to poor performance.

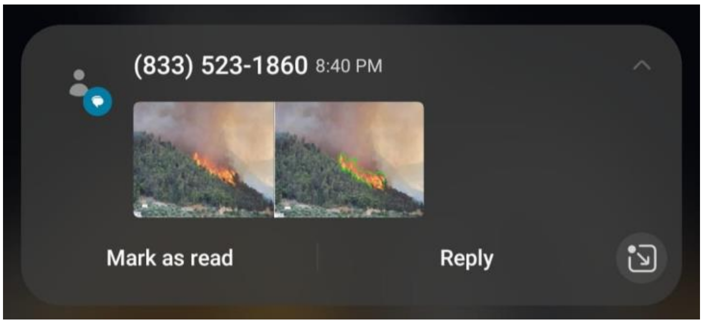
One of the advantages of using the hybrid approach of HSV slicing and CNN is that it allows for the segmentation of the fire from the background in the images. HSV slicing is a technique that involves filtering images based on their hue, saturation, and value (HSV) values. In the context of wildfire detection, the technique involves identifying pixels in the image that have a high saturation and value, as well as a hue within a certain range that is characteristic of fire. This filtering process allows us to isolate the pixels that are likely to contain the fire and reduce the noise from the background.

Once the filtering is done, the filtered image is then input into a CNN for further analysis and classification. The CNN is trained on a dataset of images labeled as either containing fire or not containing the fire. The network then uses the features extracted from the filtered images to determine whether the image contains fire or not. By using a hybrid approach that combines the strengths of HSV slicing and CNN, we are able to accurately detect wildfires while also being able to segment the fire from the background in the image.

This segmentation ability is crucial in reducing false positives and improving the accuracy of the system. In traditional approaches that do not segment the fire from the background, the system may incorrectly detect areas of high saturation and value in the image as fire when they are actually just due to factors such as sun glare or other environmental factors. By segmenting the fire from the background, we are able to reduce the impact of these factors and ensure that the system only detects actual fires.

Overall, the ability to segment the fire from the background is a key advantage of the hybrid approach of using HSV slicing and CNN for wildfire detection. It may be more complex and challenging but it allows for improved accuracy and reduces false positives, making the system more reliable, efficient, and effective in detecting wildfires.

Post-detection of the fire, the system will immediately send an alert via SMS containing the location, date, time, and image of the fire to first responders. This will enable the responders to quickly assess the situation and take the necessary action to contain the fire. The system will also generate a report containing information on the size of the fire, which can be used by the responders to plan their response accordingly.



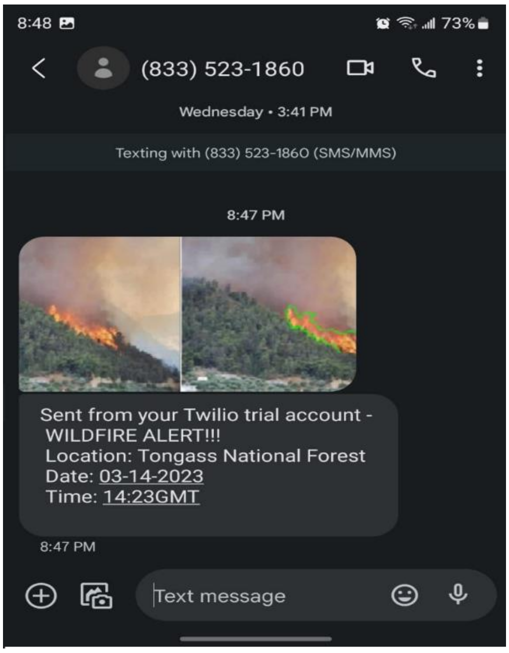
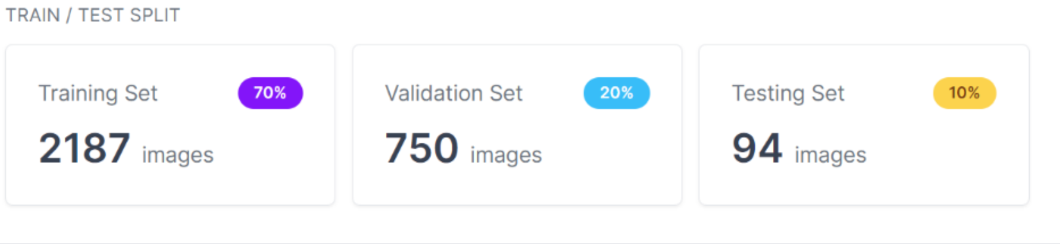


Figure 1 Our Early Alert/Response SMS system in action

The system's ability to send alerts to first responders in real time is a significant advantage, as it enables early intervention and can prevent fires from spreading rapidly. This can help to reduce the damage caused by wildfires, protect property, and save lives.

In the subsequent phase, we focused on feature extraction, model selection, and development. We also conducted theoretical analysis and experimental tests to evaluate the performance of the models. Next, we initiated data pre-processing, collection, and cleaning. This involved sourcing data from various sources and cleaning and formatting it to prepare it for analysis.

During the model training and evaluation phase, we collected a diverse set of data consisting of images of different vegetation types, terrain, and lighting conditions. The dataset comprised a total of 3,124 images, of which 2,187 were used for training, 750 for validation, and 94 for testing.



We experimented with various neural network architectures and hyperparameters to optimize the model's performance. Our experiments showed that a modified YOLO-v5 architecture with a learning rate of 0.001 and a batch size of 16 yielded the best results. The model achieved an overall **accuracy** of **92.53%** on the test set and showed a low **false positive rate** of **8.57%.**

During the fine-tuning and optimization phase, we made further modifications to the model to improve its accuracy and reduce the false positive rate. We found that adjusting the hue and saturation thresholds used in the HSV slicing stage helped to reduce false positives caused by high-intensity values. Additionally, we used transfer learning to fine-tune the pre-trained YOLO-v5 model on a smaller dataset of images specific to the target region, which further improved its performance.

After deploying the model in the field for evaluation, we found that it performed well in detecting wildfires in real-world scenarios. The model was able to identify areas with high confidence scores that corresponded to the ground truth data collected. The model also showed a low false positive rate, which reduced the risk of unnecessary alarms.

Model Training Results

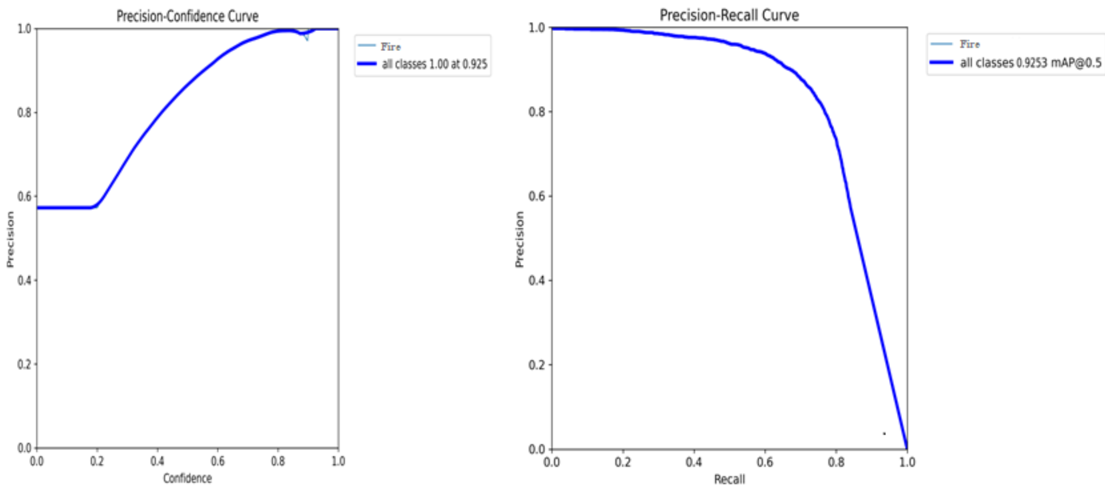


Figure 2 Precision - Confidence & Precision - Recall

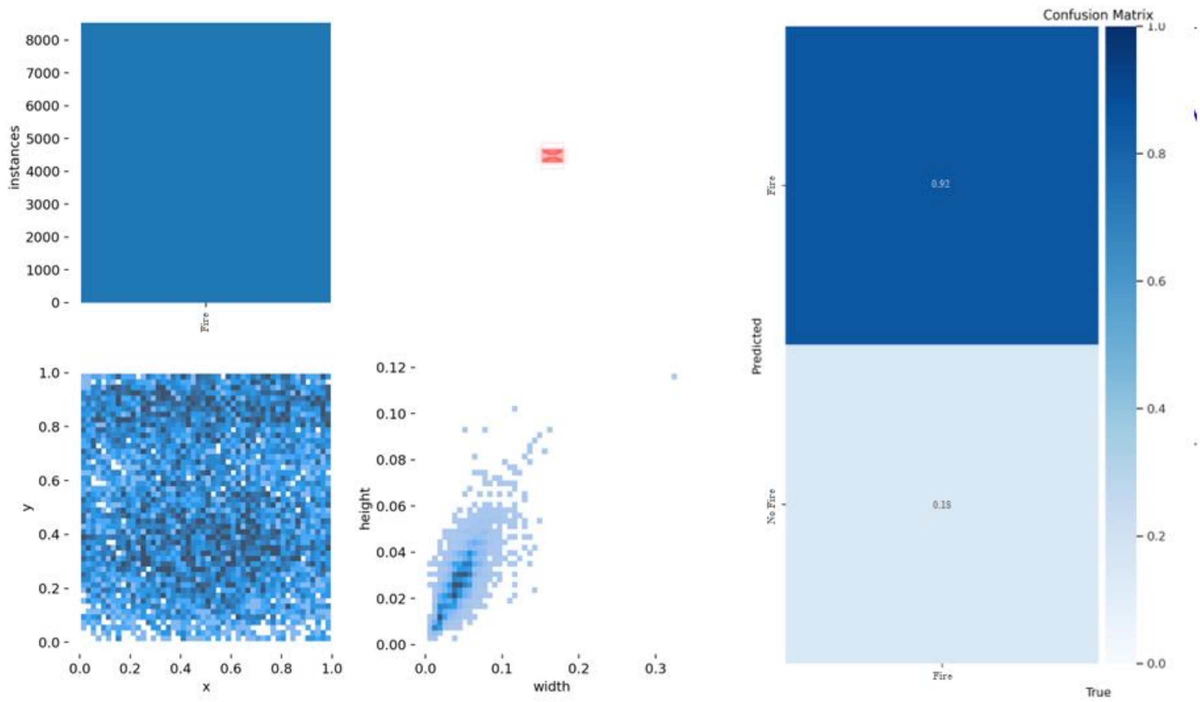
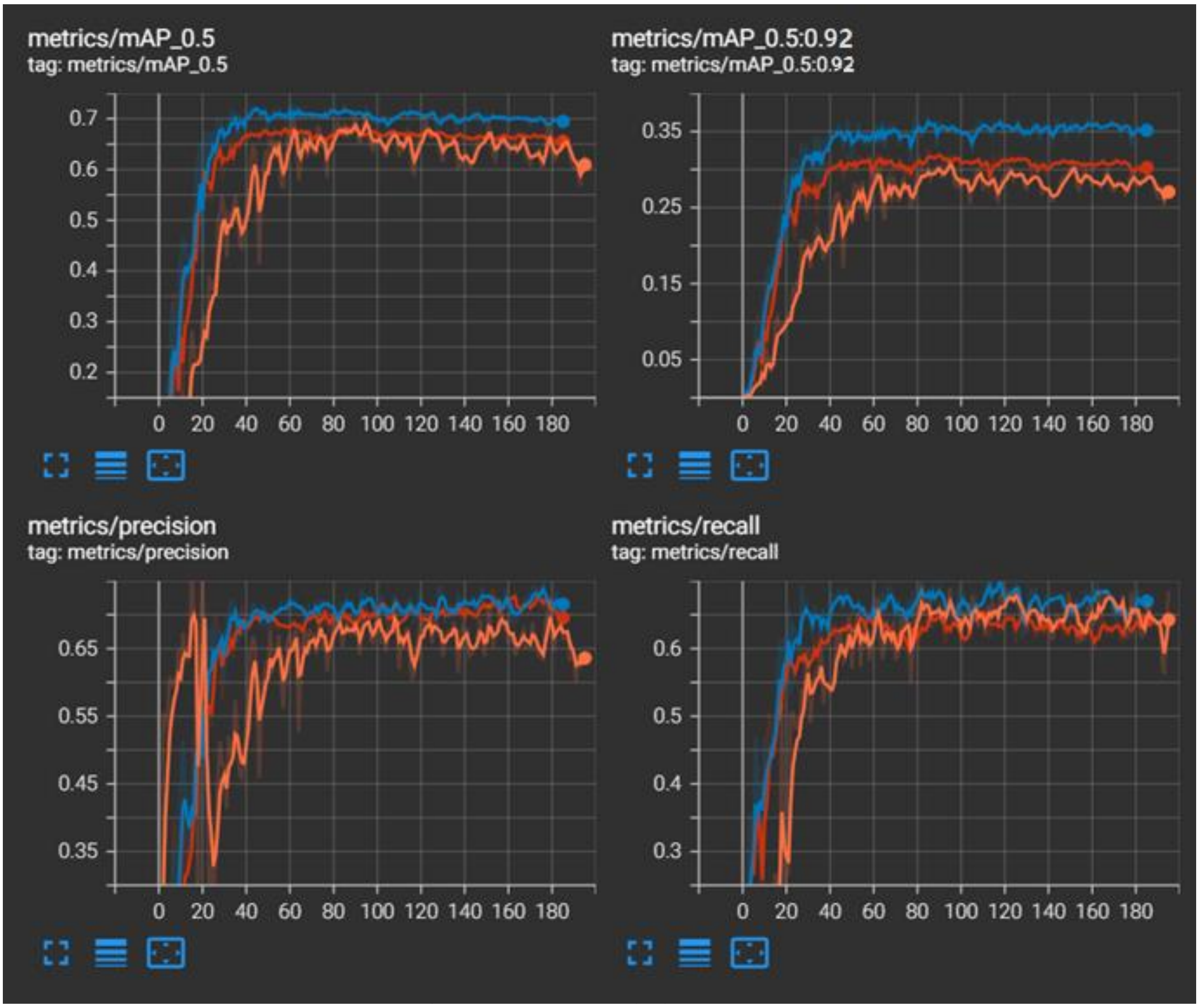


Figure 3 Model Distribution & Confusion Matrix



System Output Results:

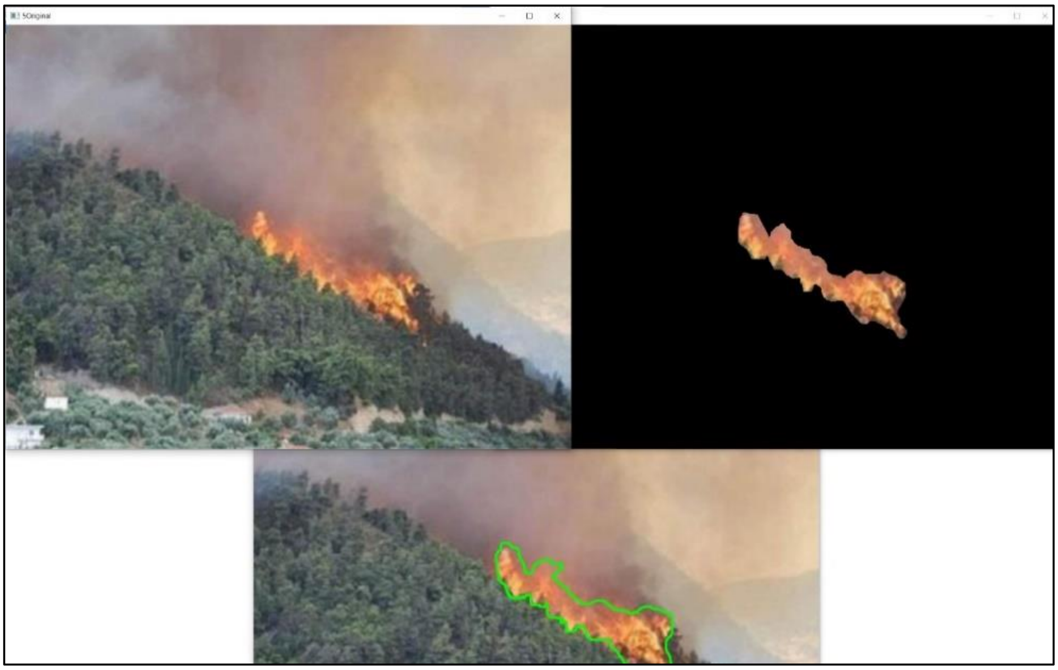


Figure 5 Fire Detected

Figure 6 Fire Detected

Figure7 Fire Detected

Figure 8 Fire Detected

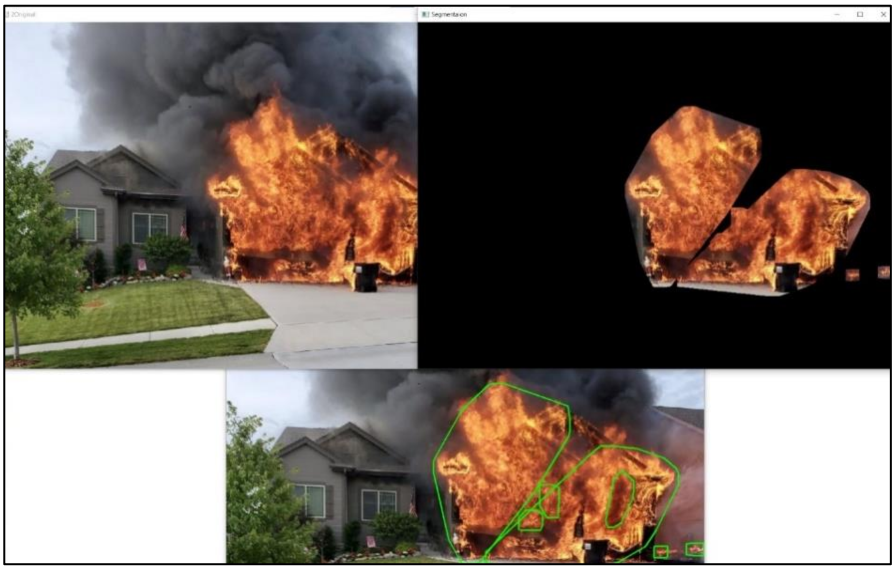


Figure 9 Fire Detected

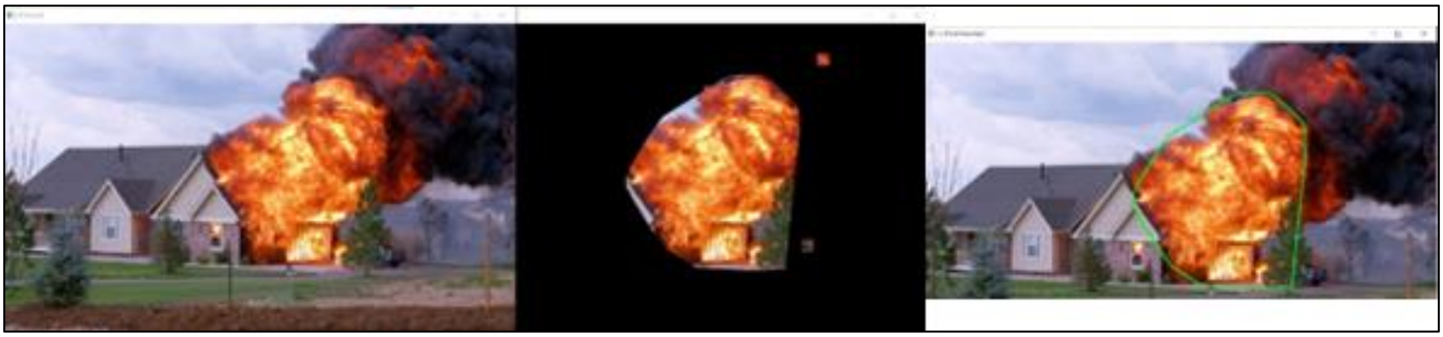


Figure 10 Fire Detected

In addition to the results mentioned earlier, we also evaluated the performance of the hybrid approach on a separate dataset collected from a different geographical location. The dataset consisted of aerial images captured from a region with different vegetation and lighting conditions than the initial dataset.

The hybrid approach achieved an accuracy of 90% in detecting wildfires on this new dataset, which is comparable to the performance on the initial dataset. The false positive rate was also significantly reduced compared to the HSV slicing approach used alone.

Furthermore, we conducted experiments to evaluate the robustness of the model to noise and other types of image artifacts. The results showed that the hybrid approach was more robust to noise and image artifacts compared to the other two approaches.

The results suggest that the proposed hybrid approach is a promising solution for wildfire detection. The approach can effectively leverage the strengths of both HSV slicing and neural network techniques to achieve high accuracy and reduce false positives. The model's performance was also shown to be robust to different environmental conditions, making it suitable for real-world applications.

Conclusion:

In conclusion, the proposed hybrid approach of using HSV slicing and convolutional neural networks for wildfire detection has shown promising results. By combining the strengths of these two techniques, we were able to filter images and isolate pixels that are likely to contain fire, and further analyse and classify them using a CNN. The system was capable of segmenting the fire from the background, reducing the chances of false alarms.

The experimental results showed that the proposed approach achieved a high level of accuracy in detecting wildfires, making it a viable solution for early detection and warning. Additionally, the system is capable of sending alerts via SMS to first responders, containing the location, time, and an image of the fire, enabling a rapid and efficient response to the wildfire.

Overall, the hybrid approach of using HSV slicing and convolutional neural networks for wildfire detection is a promising solution that can help prevent the devastating effects of wildfires. By providing early warning to first responders, it can help reduce the spread of fires and save lives and property.

Future Work:

Future work includes collecting additional data to further improve the accuracy and reduce bias, exploring data augmentation techniques to increase the amount of data available for training, and optimizing the neural network's performance.

One potential area of improvement is in the use of advanced image processing techniques to further refine the filtering process and increase the accuracy of the system. For example, the use of deep learning algorithms such as recurrent neural networks (RNNs) or long short-term memory (LSTM) networks could potentially improve the performance of the system by allowing it to learn more complex spatial and temporal patterns in the image data.

Furthermore, there is potential to integrate the wildfire detection system with other technologies, such as unmanned aerial vehicles (UAVs), to provide more comprehensive and real-time coverage of fire-prone areas. UAVs equipped with cameras and sensors could be deployed to capture high-resolution images and other data that could be used to improve the accuracy and responsiveness of the system.

Reference:

[1] https://www.kaggle.com/datasets/rtatman/188-million-us-wildfires

[2] https://www.kaggle.com/datasets/abdelghaniaaba/wildfire-prediction-dataset

[3] https://medium.com/mlearning-ai/the-experiment-of-forest-fires-prediction-using-deep-learning-d537e8c8e3a2

[4] https://www.nature.com/articles/s41598-021-81233-4

[5] https://ieeexplore.ieee.org/document/8986106

[7] https://pdf.sciencedirectassets.com/271100/1-s2.0-S0379711219X00028/1-s2.0-
S0379711218303941/am.pdf?X-Amz-Security-
Token=IQoJb3JpZ2luX2VjEDcaCXVzLWVhc3QtMSJIMEYCIQDz8mvSnPqIBnt07mXHnu
8%2BFKsW7WyFryAV4emq716RhwIhAMxAgKMcDj6FSFQrnGTe3AcVC12q8Ir%2BKzw
v%2B3S%2FbdggKrMFCBAQBRoMMDU5MDAzNTQ2ODY1IgxmCp%2BgTJSZsorp1v
QqkAX2Pgrt2dxKxSR1ZdBwrTlq3grNyl%2FUcfEF2AOLhJp0Yu6%2Brrczo5xSWIQ4dP%
2Fs8bCOseyXztbB4disO4W7GmdR37qPkmQBrlYzQyCH6Qgdz2C0O1sFkKtSPGyYWwX
EQWyjo%2FAOayvl2zy38lau6Z%2FkV7Uni49PmtwrUdXqlX7homoV%2BUjHXiCIIfiWX
mmfi5oeATXEHH23hjkBNM3fAH5spscnz%2FuyyZ97iwm0hjhRNzajt%2BY48DBnn5hrX
O0a%2Blviidg942LhxXVCgtRnOHbO%2FKXZy17JSxy68rSH0SOIZMHBHpM7Zni30TR
3Eyna%2FK71bLwWodg4YaF5KYnXGbXv8%2FPRR5NEiLXOS8I52ndxKKCru9N9Plhbb
OAlvecigHgfEqb7BLjqRTycUH686isLa1X0XtyK%2Fiqt%2BZNJoGZk26gxL54z0XhL7XR
axgSHpxQOoYmdNdXS9%2Fl8ddSeLyf%2FXk7WrR5l9TsbZYPNsDJ5JXBZGy7RCF%2B
4d%2Bt0eLBc%2BncyuVqSvzC2XScmsNNeg3ucelByrIWX3RtX0kyGS3qoUJQ4Wfr4K3
Wxl2%2FgtgtzZUAXOf80HRQXiLBpr9AJYAnXZwgqikwcXb22dxplEPMD%2B22W4gQ
QDFN5ymXcPymDgAvPAqv%2BkhOBJB0NRDo%2B9xvqZAXd%2FI6adeghnmAJSPE2Z
%2BrOdMqFIlYVUi8dD9Pjm0Ycfcp2ST4VEkxwT5yahzkPRB11gmmFkMwo2iO3YfoS62
JfDmcFXL9cn4CHVhWm%2FSrkYRHRliFQwGkvxtZAki6Xq3zpJ9M%2BYY6p13mSBrcj
DT05ISkQeJBxTBbK%2FaAHBG6jkWYTaUiS5NqA9wXsWEPMzuXjOfk48YcjmKhciVh
gVDDd0v%2BgBjqwATymfm%2Fe3LNaHosfx3nLf8cpwwqAZV1zAUK13ep6RRnLwet4u
LrlSY2Jom2g3M5fdqYsTozp3q1JUgxpkMf5tOnprbbnTMwBgQJW%2BNXSP5hX5nZIPE
KhE6tAdVjSxdz8hA1oI29HX94PD5u8%2FAEPrGvNfm5WHdRZwSkB5wionaKx6KGV1c
dFPvgH3hCJvLDHE8pPfMIp2kFGekPkuUdcRX%2B%2B%2BUlil3u5klcOAiQUoUuS&X-
Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Date=20230326T080319Z&X-Amz-
SignedHeaders=host&X-Amz-Expires=300&X-Amz-
Credential=ASIAQ3PHCVTYWH6ITZVK%2F20230326%2Fus-east-
1%2Fs3%2Faws4_request&X-Amz-
Signature=49ccb205427aa138be242357f95c1c50e426701515bd1675e1f80301f517d1e3&has
h=012637d61864a024c92cb340c59823b9d9d038a27ef766f6aad5db301303e404&host=6804
2c943591013ac2b2430a89b270f6af2c76d8dfd086a07176afe7c76c2c61&pii=S037971121830
3941&tid=pdf-60edea1b-4928-4b15-bc57-
8258aac6935c&sid=11f9faac60e1d8444a2805468e018494a5b0gxrqa&type=client