A Convolutional Neural Network Approach to Fire Classification and Segmentation

Computer Vision and Image Analysis - Final Project Report

Aji Kautsar Yusuf
(20/454525/PA/19556)
Department of Computer Science and Electronics, Universitas Gadjah Mada
Building C, 4th Floor North
Yogyakarta, Indonesia
ajikautsar04@mail.ugm.ac.id

M. Rafi Syafrinaldi
(21/472772/PA/20335)
Department of Computer Science and Electronics, Universitas Gadjah Mada
Building C, 4th Floor North
Yogyakarta, Indonesia
muhammad.rafi.syafrinaldi@mail.ugm.ac.id

Mohammad Abdan Syakura
(21/475044/PA/20519)
Department of Computer Science and Electronics, Universitas Gadjah Mada
Building C, 4th Floor North
Yogyakarta, Indonesia
mohammad.abdan.syakura@mail.ugm.ac.id

Rabbani Nur Kumoro
(21/472599/PA/20310)
Department of Computer Science and Electronics, Universitas Gadjah Mada
Building C, 4th Floor North
Yogyakarta, Indonesia
rabbani.nur.kumoro@mail.ugm.ac.id

Vania Clementine Lumbanbatu
(21/475023/PA/20509)
Department of Computer Science and Electronics, Universitas Gadjah Mada
Building C, 4th Floor North
Yogyakarta, Indonesia
vania.cle2003@mail.ugm.ac.id

***Abstract* — Forest fires pose significant threats to ecosystems, necessitating accurate detection and control methods. Traditional approaches suffer from limitations such as restricted field of view, delayed responses, and high costs. To address these challenges, this study proposes a solution that combines Image Processing and Convolutional Neural Network (CNN) architecture for precise fire classification and segmentation. The proposed methodology employs MobileNetV2 for fire classification and DeepLabv3+ for semantic segmentation. The approach achieves impressive results, with an accuracy exceeding 97% and an IoU score of 63%, showcasing its effectiveness in accurately classifying and segmenting fire regions. By leveraging advanced CNN architectures and semantic segmentation techniques, this solution aims to overcome limitations, provide a reliable method for forest fire identification, and enable early detection of wildfires, thus supporting effective firefighting strategies.**

Keywords — Classification, Convolutional Neural Network (CNN), DeepLabv3+, Fire, MobileNetV2, Segmentation

I. Introduction

For humans, forests provide a variety of wealth and play an essential role in living. Forests are also considered our planet's lungs, they filter the air by adding oxygen (O2) and lowering carbon dioxide (CO2) levels. Additionally, they purify the water from the majority of pollution-causing agents. Over the last few years, forest fires are becoming a kind of natural tragedy for forest lands, and wild habitats as it severely affects the ecosystem **[2]**. Forest fires have caused substantial economic losses, air pollution, environmental degradation, and other challenges all over the world, wreaking havoc on human life, animals, and plants **[5]**. According to incomplete statistics, over 200,000 forest fires occur worldwide, destroying approximately 10 million hectares of forest land **[12]**.

In recent years, forest-fire monitoring methods represented by deep learning have been developed rapidly. Using drone technology and optimizing existing models to improve forest-fire recognition accuracy and segmentation quality is of great significance for understanding the spatial distribution of forest fires and protecting forest resources. Due to the spreading and irregular nature of fire, it is extremely tough to detect fire accurately in a complex environment **[12]**. Fire classification and segmentation are helpful for firefighters to understand the fire scale and formulate a reasonable fire-fighting plan.

The proposed study involves the use of a classification and semantic segmentation approach using the combination of image preprocessing, MobileNetV2, and DeepLabv3+ CNN architecture for segmenting wildfire and detecting fire regions. Specifically, our used model, DeepLabv3+, extends DeepLabv3 by adding a simple yet effective decoder module to refine the segmentation results, especially along object boundaries **[12]**. DeepLabv3+ is a semantic segmentation architecture that improves upon DeepLabv3 with several improvements, such as adding a simple yet effective decoder module to refine the segmentation results **[4]**. This study aims to develop precise classification and segmentation methods that can effectively identify forest fires and enable early detection of wildfires.

II. Problem Statements

Fire is a natural process in several ecosystems, but in others, it negatively affects biodiversity, particularly when natural fire regimes change abruptly. Wildfires are one of the critical factors affecting global ecosystems and societies, impacting atmospheric composition, vegetation succession, soil erosion and runoff, and societal values (resources and assets). Wildfires also have important socio-economic implications and can affect lives, houses, and other human values. Recent studies estimate that every year more than 4 million km^2, which is approximately the size of India and Pakistan combined, are burned globally. Notably, those estimations are likely to be conservative, as they are based on coarse-resolution satellite images, which tend to miss small fires that are quite frequent in tropical regions **[6]**.

Numerous conventional approaches have been used to detect and control forest fires, such as human-operated watchtowers, patrolling helicopters, and remote-sensing satellites for forest surveillance. Watchtowers are equipped with multiple wireless sensors, such as temperature, gas, smoke, and optical, in the forest to identify fires. These detectors have limitations, such as a restricted field of view, delayed responses, and high manufacturing costs. Human-powered helicopters can be costly if detection activities are carried out often and can be easily affected by unfavorable situations. Remote sensing satellite methods provide a broad picture but demand enormous resources due to low spatial and temporal image resolution, making it hard to ensure quality monitoring in the early stages **[2]**. These methods are time-consuming, laborious, and generally inefficient in recognition.

As a kind of forest "fault", fire is highly destructive and difficult to rescue. Deep learning has drawn interesting results for pixel-level classification for smoke detection, but few systems are proposed for fire flame detection. Therefore, this study proposes a fire classification and segmentation method based on deep learning.

III. Literature Study

The literature study presented in our study focused on the use of Convolutional Neural Networks (CNNs) for fire classification and semantic segmentation. Several real-life studies have demonstrated the effectiveness of CNNs in detecting wildfires from satellite imagery. For example, a recent study by Ghali et al. (2021) **[3]** used deep vision transformers to manually segment wildfire regions from satellite images with high accuracy.

Another study by Li et al. (2019) **[7]** proposed a CNN-based method for the early detection of forest fires using thermal infrared images. These studies have shown that CNNs can accurately identify fire regions by learning distinctive features such as color, texture, and shape. In addition to CNNs, we also explored the use of MobileNetV2, a state-of-the-art algorithm that has been successfully applied in various computer vision tasks.

By combining these two techniques, we aimed to develop an effective model for fire classification that can classify fires in real-time with high accuracy and precision. Overall, the literature study provided valuable insights into the current state-of-the-art in fire classification and semantic segmentation and helped guide our approach to developing an effective model for detecting and preventing forest fires.

IV. Methodology

The methodology depicts the process for classifying and segmenting fire from an image using a series of key steps. Initially, the image is loaded. Then, a Convolutional Neural Network (CNN) called MobileNetV2 is employed to analyze the image, identify, and classify the presence of fire. If no fire is detected, the process concludes, and an empty matrix is returned.

However, if a fire is detected by MobileNetV2, the image undergoes pre-processing to enhance the accuracy of the subsequent classification steps. This pre-processing includes RGB Thresholding, where the image is analyzed based on pixel values in the RGB color space. Additionally, YCbCr Thresholding is applied to examine pixel values in the YCbCr color space. Furthermore, segmentation using DeepLabv3+ is performed to segment the image and identify regions of interest

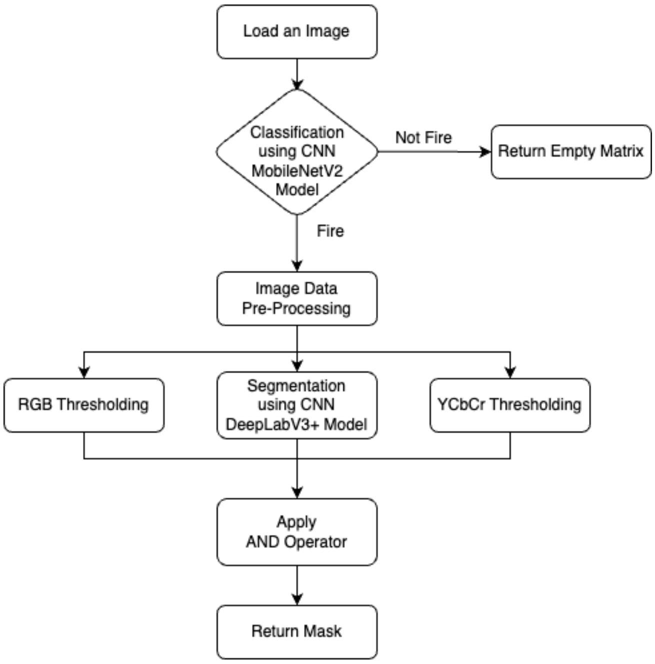

Figure 1. Flowchart Diagram

The outputs of the pre-processing steps are then combined using logical AND operators. For a pixel to be considered part of the fire, it must be identified as fire by at least one thresholding technique and also align with the segmentation model's results. Finally, a mask is generated to highlight the areas of the image that contain the fire, providing a visual representation of the detected fire regions. This overview from **Figure 1.** illustrates the systematic approach used to detect fire in an image and highlights the various stages involved in the process.

A. *Classification Data Acquisition*

The dataset we use for our classification model is the "FIRE Dataset" obtained from Kaggle. The dataset was specifically curated to develop a model capable of recognizing images containing a fire, with the objective of binary classification.

Figure 2. Frame Samples from FIRE Dataset

The dataset from **Figure 2** consists of two classes: fire_images and non-fire_images. The fire_images folder contains a total of 755 outdoor-fire images, some of which may include heavy smoke. These images were collected from various sources, encompassing different fire-related scenarios. The dataset captures the diversity and complexity of fire incidents, enabling the development of models that can discern fire occurrences from other environmental factors. The non-fire_images folder encompasses 244 nature images that serve as the reference for non-fire instances. These images depict various natural scenes such as forests, trees, grasslands, rivers, people, foggy forests, lakes, animals, roads, and waterfalls. The inclusion of non-fire images helps establish a clear distinction between fire and non-fire classes during the binary classification process.

B. Segmentation Data Acquisition

The FLAME (Fire Luminosity Airborne-based ML Evaluation) Dataset, the primary dataset used in this study, is a collection of high-resolution aerial images of pile burn detection using drones or UAVs **[10]**. It comprises a vast and diverse collection of aerial imagery captured during a prescribed pile burn in Northern Arizona, USA, using drones. The dataset includes raw aerial videos and raw heatmap footage obtained from an infrared thermal camera. What sets the FLAME dataset apart is its accessibility. It is publicly available through IEEE Dataport and is leveraged in this study due to its applicability in the fire segmentation task.

Figure 3. Frame Samples from FLAME Dataset

The dataset comprises several repositories, each holding diverse and valuable data as shown in **Figure 3**. We use the ninth and tenth repositories from the FLAME Dataset. The ninth repository is dedicated to fire segmentation and contains 2,003 fire frames with a resolution of 3480x2160, with a total size of 5.3 GB. While the tenth repository holds the ground truth data for the fire segmentation problem and consists of 2,003 mask frames, each with a 4K resolution of 3480x2160, amounting to 23.4 MB in total size.

Together, these repositories form an inclusive and diverse dataset, enabling the exploration of various facets of fire segmentation using convolutional neural networks. The images, videos, and thermal footage offer a rich foundation for training the model to distinguish between fire and no-fire scenarios and accurately segment fire in various conditions.

C. Classification Modelling

In our approach, we utilized fire classification modelling as a preliminary step to enhance the accuracy of image segmentation for fire detection. By employing a dedicated fire classification model, we aimed to filter out images that do not contain a fire, thereby reducing the number of false positives in the subsequent segmentation process. In the development of our fire classification system, we employed transfer learning by utilizing the MobileNetV2 model pre-trained on the widely recognized ImageNet dataset as our foundational architecture.

The architecture of MobileNetV2 is built upon an inverted residual structure, which differs from conventional residual models **[8]**. In this structure shown in **Figure 4**, the input and output of the residual block are represented by thin bottleneck layers, as opposed to expanded representations. This design choice helps to optimize the model's performance while minimizing computational complexity.

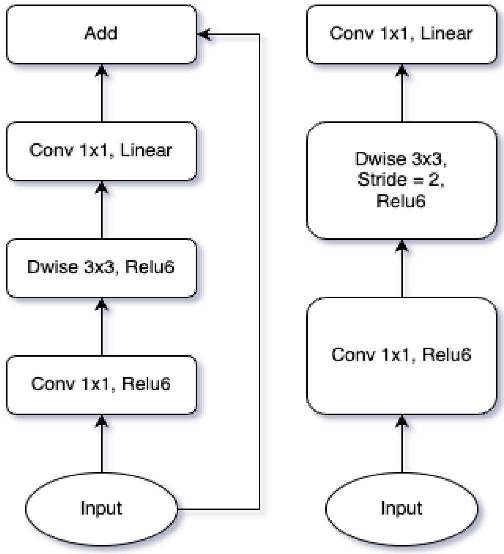

Figure 4. MobileNetV2 Architecture

To adapt MobileNetV2 for fire classification, we replaced the final classification layer with a custom-designed layer tailored specifically for our task. This new layer was developed to accurately classify images into either fire or non-fire categories.

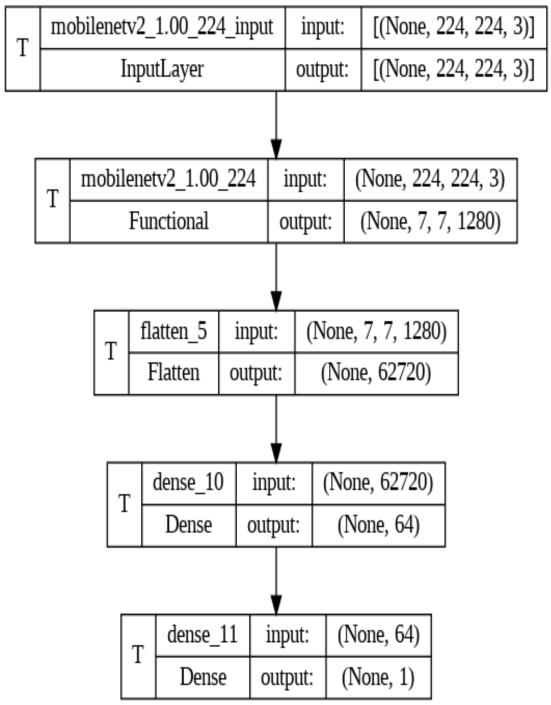

Figure 5. Proposed Model Architecture

By training the modified MobileNetV2 model from **Figure 5** on a curated dataset of fire images, we fine-tuned the model's parameters and optimized its weights to better suit the nuances of the fire classification task.

Figure 6. Result of the Classification Model

The output of the classification model depicted in **Figure 6** is straightforward and intuitive. When a fire is detected, the result is classified as "Fire." On the other hand, if no fire is detected, the output is classified as "Non Fire." This classification approach simplifies the interpretation of the model's predictions, providing a clear indication of whether a fire is present or not in the given input data and helps reduce false positive rates for the segmentation model.

D. Image Data Pre-Processing

In this study, we propose a pre-processing method based on color space conversion and image processing techniques to enhance the visibility and distinguishability of fire regions in images. The method involves converting the RGB image to the HSV color space, applying histogram equalization to the Value (V) channel, performing median filtering on the V channel, and finally converting the modified HSV image back to RGB color space. The experimental results display the effectiveness of the proposed method in reducing noise interference.

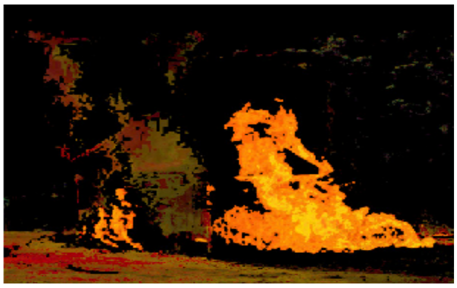

Figure 7. HSV Color Space Image

In our approach, we begin by converting the RGB images from **Figure 7** into the HSV color space. This conversion is beneficial as it enables us to take advantage of the perceptual characteristics of color and isolate specific components of the image, namely hue, saturation, and value. By operating in the HSV color space, we can effectively identify and extract fire regions based on their distinct hue and saturation values **[11]**. This initial step of converting the image from the RGB color space to the HSV color space allows us to separate the image into different components, making it more convenient to manipulate and analyze specific aspects of the image.

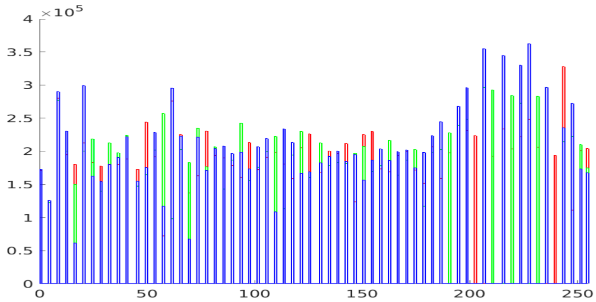

Figure 8. Histogram Equalization

Following the conversion to the HSV color space, we proceed to apply histogram equalization specifically to the value (V) channel of the HSV image in **Figure 8**. This technique is employed to enhance the contrast and distribution of pixel intensities within the V channel, resulting in improved visibility of fire regions **[11]**. Histogram equalization works by redistributing the intensity values across the entire range, effectively making both brighter and darker areas more distinguishable. By performing histogram equalization on the V channel, we aim to enhance the image's contrast, bring out more details, and overall improve the quality of the image.

123	125	126	130	140
122	124	126	127	135
118	120	150	125	134
119	115	119	123	133
111	116	110	120	130

Figure 9. Median Filtering

To further improve the quality of the fire regions, we apply a median filtering operation to the equalized V channel as shown in **Figure 9**. The purpose of median filtering is to reduce noise in the image while preserving the edges and details of the fire regions. This technique effectively eliminates unwanted artifacts and enhances the overall clarity of the fire boundaries. By replacing each pixel value with the median value of its neighboring pixels, median filtering helps to smooth out the image and reduce the impact of noise while retaining important features. This step plays a crucial role in ensuring that the fire regions are accurately represented and enhances the visual quality of the final output.

Figure 10. Pre-Processed RGB Image

Finally, the pre-processed HSV image undergoes a conversion back to the RGB color space, resulting in an enhanced and visually appealing representation of the fire regions, as depicted in **Figure 10.** By converting the image from the HSV color space back to the RGB color space, we ensure that the final processed image is in the widely used RGB format, making it compatible with various applications and visualization tools. This processed image, along with the filtered V channel, serves as valuable input for fire segmentation algorithms. By incorporating these enhancements, we can expect improved accuracy and performance in the task of accurately identifying and segmenting fire regions within the image.

E. Image Thresholding

Thresholding techniques are vital in the fire segmentation task for separating flame pixels from the background. Two approaches are used in this study which are RGB and YCbCr thresholding. RGB thresholding analyzes pixel values in the RGB color space, exploiting color characteristics to isolate flames. YCbCr thresholding involves converting the image to YCbCr color space, separating it into components for statistical analysis, and applying threshold values to identify flames based on their color distribution **[11]**. These techniques enable accurate flame extraction, contributing to precise fire segmentation.

In the RGB thresholding step, we focus on the pixel values in the RGB color space to extract flames. Flames typically exhibit distinctive colors, with higher intensity in the red and green channels compared to the blue channel. By isolating the red and green channels, we can concentrate on these relevant color components. We define specific criteria to identify flames, such as a higher red value compared to green, and apply a threshold with a value of 177 in our case to separate flame pixels from the rest of the image.

In the YCbCr thresholding step, we convert the RGB image to the YCbCr color space, which separates the image into its luminance (Y) and chrominance (Cb and Cr) components. This conversion enhances our ability to capture color information effectively. Flames in the YCbCr color space exhibit specific color distribution characteristics. To differentiate flame pixels from the background, we determine threshold values based on the statistical properties of the Cb and Cr channels. By analyzing the distribution of these channels, we can identify flames based on their distinct color characteristics. Furthermore, we can refine the flame extraction process by considering additional criteria, such as setting a threshold value for the Y channel, which takes into account the luminance characteristics of flames.

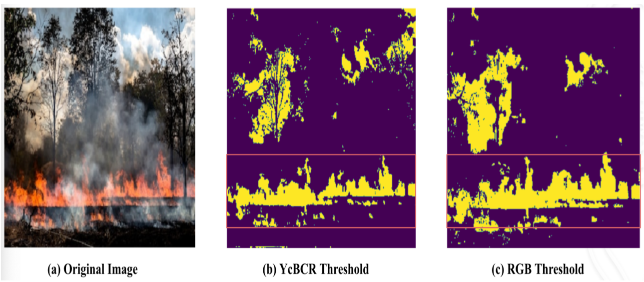

Figure 11. (a) (b) (c) Thresholding Process Flow

The thresholding techniques, RGB and YCbCr, have been applied to the image in **Figure 11**, resulting in a thresholded image where the flames are prominently highlighted. By setting appropriate threshold values, the flames are separated from the background and other objects, effectively suppressing a large portion of the image. This process aids in isolating the flames and enhancing their visibility, facilitating accurate detection and differentiation from the surrounding environment.

F. Segmentation Modelling

In this study, we employ the state-of-the-art DeepLabv3+ model, specifically designed for semantic image segmentation, to accomplish this task effectively, the model architecture is shown in **Figure 12**. The model was trained on the FLAME Dataset, employing an 80/20 train-validation split and over a course of 30 epochs. To prevent overfitting and ensure generalization, an early stopping mechanism was employed.

The architecture of DeepLabv3+ consists of three essential components: the backbone network, the atrous spatial pyramid pooling (ASPP) module, and the decoder **[4]**. The backbone network is based on a pre-trained EfficientNet convolutional neural network (CNN) and serves as a feature extractor. It processes the input image and extracts high-level features **[9]**.

The ASPP module employs parallel atrous convolutions with different dilation rates to capture multi-scale contextual information **[5]**. By considering multiple dilation rates, the ASPP module effectively analyzes objects at various scales and preserves fine-grained details **[4]**. The outputs of the parallel convolutions are combined to create a comprehensive representation of the image.

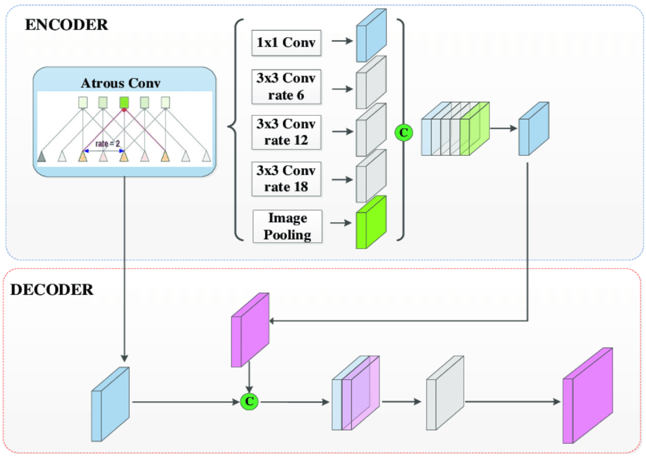

Figure 12. DeepLabv3+ Architecture

The decoder module refines the output from the ASPP module and generates the final segmentation map. It incorporates upsampling layers to increase the spatial resolution of the feature maps gradually. Additionally, skip connections are utilized to merge low-level and high-level features, preserving fine details during the upsampling process. The decoder module enhances the segmentation map by incorporating more detailed information, leading to improved localization accuracy.

G. Applying AND Operator

The segmentation mask obtained through the DeepLabv3+ model provides a binary representation of the regions of interest within the image with values of 1 representing fire areas and 0 representing the non-fire areas.

The mask is then combined with the thresholded RGB and YcBCR image using the AND operator which performs a pixel-wise comparison between the corresponding pixels of the segmentation mask and the thresholded image. The resulting image assigns a value of 1 only to pixels that satisfy the conditions of both the segmentation mask and the thresholded image.

H. Results

Figure 13. Segmented Fire from Test Dataset

Figure 14. Segmented Fire from VISIFIRE Dataset

The segmentation results obtained from the model showcase its effectiveness in accurately identifying and segmenting fire regions in both the Test and the VISIFIRE Dataset in **Figure 12** and **Figure 13**. The model successfully minimizes false positive segmentations, ensuring that the identified fire areas align closely with the ground truth labels. This demonstrates the model's capability to precisely capture fire regions, highlighting its accuracy and reliability in fire segmentation tasks.

V. Performance Evaluation

The metrics used in the classification report from **Table 1**, including accuracy, recall, precision, and F1-score, are commonly employed to evaluate the performance of classification models. Accuracy measures the overall correctness of the model's predictions, while recall quantifies the model's ability to identify all instances of the positive class. Precision assesses the proportion of correctly predicted positive instances among all predicted positive instances. F1-score provides a balanced measure that considers both precision and recall, making it suitable for evaluating classification models **[12]**.

In the segmentation report shown in **Table 2**, the metrics used, namely accuracy, IoU score, dice loss, and focal loss, are specific to the task of semantic segmentation. Accuracy measures the correctness of the model's pixel-wise predictions. IoU score, or Intersection over Union, calculates the overlap between the predicted segmentation masks and the ground truth masks, providing a measure of segmentation accuracy. Dice loss assesses the similarity between predicted and ground truth masks, while focal loss focuses on challenging regions for better training performance **[9]**.

These metrics are chosen based on their relevance to the specific tasks of classification and semantic segmentation, providing insights into different aspects of the model's performance and allowing for an in-depth evaluation of their effectiveness.

Table 1. Classification Report

Model	*Accuracy*	*Recall*	*Precision*	*F1-Score*
MobileNetv2	0.985	1.0	0.945	0.972

MobileNetV2 achieves an accuracy of 0.985, indicating a high level of correctness in fire classification predictions. With a recall of 1.0, the model effectively captures all true positive fire regions, resulting in a low rate of false negatives. It demonstrates a precision of 0.945, meaning that when it identifies a pixel as a fire region, it is correct approximately 94.5% of the time, demonstrating a relatively low rate of false positives. The F1 score of 0.972, which combines precision and recall, showcases the model's strong performance in fire classification tasks, effectively balancing the trade-off between minimizing false negatives and false positives. These metrics reflect the performance of MobileNetV2, highlighting its accuracy, recall, precision, and the overall balance between precision and recall, as measured by the F1 score.

Table 2. Segmentation Report

Model	*Accuracy*	*IoU Score*	*Dice Loss*	*Focal Loss*
DeepLabv3+	0.976	0.630	0.985	0.074

The accuracy of the DeepLabv3+ model is reported as 0.976, indicating a high level of overall correctness in its segmentation predictions. However, the IoU score is 0.630, which suggests a lower level of overlap between the predicted and ground truth masks compared to other metrics. This indicates that the model may struggle in accurately segmenting fire regions and capturing their boundaries. There could be several speculations as to why the IoU score is not as good as the other metrics. One possibility is that the model may have difficulty in capturing fine details and precise boundaries of fire regions due to the complexity and variability of their shapes and appearances. Another factor could be the presence of image artifacts, such as noise or low contrast, which might hinder the model's ability to accurately segment fire regions.

To address the lower IoU score, potential solutions can be considered. One approach is to explore more advanced architectural modifications or techniques specific to semantic segmentation tasks, such as using dilated convolutions or incorporating contextual information through the use of attention mechanisms. Additionally, incorporating more diverse and representative training data, including a wide range of fire scenarios and environmental conditions, could help improve the model's ability to handle different variations in fire appearances.

Another solution could involve refining the post-processing steps, such as applying morphological operations or boundary refinement techniques, to enhance the accuracy of the segmented fire regions. Furthermore, exploring data augmentation techniques specifically designed to address the challenges associated with fire segmentation, such as rotation, scaling, or introducing variations in lighting conditions, may also contribute to improving the IoU score.

VI. Conclusions

Our study demonstrates the successful utilization of MobileNetV2 and DeepLabv3+ models for fire classification and semantic segmentation tasks. These models exhibit exceptional accuracy, surpassing 97%, and achieve an IoU score of 63% in accurately classifying and segmenting fire regions within images. This level of performance ensures precise identification of fire areas, essential for effective fire detection and monitoring.

Moreover, MobileNetV2 and DeepLabv3+ showcase their resilience in handling challenging scenarios, including low-quality images such as CCTV footage, while maintaining high accuracy. This resilience highlights the effectiveness of deep learning techniques in addressing complex computer vision tasks like fire detection. By leveraging these advanced techniques, we can explore new possibilities for improving fire safety and emergency response systems through image analysis and understanding.

In conclusion, the successful implementation of MobileNetV2 and DeepLabv3+ models, their high accuracy, improved segmentation quality, and ability to handle low-quality images, demonstrate the effectiveness of deep learning in fire classification and semantic segmentation. These findings pave the way for further advancements in fire safety by harnessing the capabilities of advanced image analysis and understanding offered by deep learning techniques.

To further enhance the performance of the models, it is recommended to explore data augmentation techniques that can improve their ability to handle diverse environments and lighting conditions **[10]**. Incorporating temporal information from video sequences can help improve fire detection accuracy and enable real-time monitoring of fire incidents. Evaluating the models' performance on large-scale datasets and comparing them with other state-of-the-art methods will provide valuable insights and validation of their effectiveness.

Additionally, exploring the integration of additional sensors and data sources, such as thermal imaging and multi-modal fusion, can contribute to enhancing the accuracy and reliability of fire detection systems by incorporating complementary information from different sources **[11]**. These future directions can advance the capabilities of the models and further improve their applicability in real-world scenarios.

VII. Acknowledgment

We very much appreciate the guidance and advice that Mr. Wahyono, S. Kom., Ph.D., throughout the Computer Vision and Image Analysis course in this semester. Thank you for all the knowledge that you have given to us.

We also are aware of the drawbacks that it isn't comparable to other research out there that uses complex neural networks and many more. But throughout the final project completion, we are proud of what we can create and how it could be used and applied in real world problems.

VIII. References

[1] Chen, L.-C. et al. (2018) 'Encoder-decoder with atrous separable convolution for Semantic Image segmentation', Computer Vision – ECCV 2018, pp. 833–851. Available at: https://doi.org/10.1007/978-3-030-01234-2_49.

[2] Chuvieco, E. et al. (2023) 'Towards an integrated approach to wildfire risk assessment: When, where, what and how may the landscapes burn', Fire, 6(5), p. 215. Available at: https://doi.org/10.3390/fire6050215.

[3] Ghali R, Akhloufi MA, Jmal M, Souidene Mseddi W, & Attia R. (2021). Wildfire Segmentation Using Deep Vision Transformers. Remote Sensing. 13(17):3527. Available at: https://doi.org/10.3390/rs13173527

[4] Harkat, H., Nascimento, J.M. and Bernardino, A. (2020) 'Fire segmentation using a deeplabv3+ architecture', Image and Signal Processing for Remote Sensing XXVI. Available at https://doi.org/10.1117/12.2573902

[5] Lang, Y., & Moeini-Meybod, H. (2021). Wildfires – a growing concern for sustainable development (2021) UN Department of Economic and Social Affairs (DESA) Policy Briefs. Available at: https://doi.org/10.18356/27081990-111.

[6] Mseddi, W.S. et al. (2021) 'Fire detection and segmentation using Yolov5 and U-NET', 2021 29th European Signal Processing Conference (EUSIPCO). Available at: https://doi.org/10.23919/eusipco54536.2021.9616026.

[7] P. Li and W. Zhao, "Image Fire Detection Algorithms Based on Convolutional Neural Networks", Case Studies in Thermal Engineering, Vol. 19, pp. 100625 (2020). Available at: https://doi.org/10.1016/j.csite.2020.100625.

[8] Sandler, M. et al. (2018) 'MobileNetV2: Inverted residuals and linear bottlenecks', 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Available at: https://doi.org/10.1109/cvpr.2018.00474.

[9] Shahid, M. et al. (2023) 'Forest fire segmentation via temporal transformer from aerial images', Forests, 14(3), p. 563. Available at: https://doi.org/10.3390/f14030563.

[10] Shamsoshoara, A. et al. (2021) 'Aerial imagery pile burn detection using Deep learning: The flame dataset', Computer Networks, 193, p. 108001. https://doi.org/10.1016/j.comnet.2021.108001.

[11] Wahyono, Harjoko A, Dharmawan A, Adhinata FD, Kosala G, & Jo K-H. (2022). Real-Time Forest Fire Detection Framework Based on Artificial Intelligence Using Color Probability Model and Motion Feature Analysis. Fire. 5(1):23. Available at: https://doi.org/10.3390/fire5010023

[12] Guan, Z. et al. (2022) 'Forest fire segmentation from aerial imagery data using an improved instance segmentation model', Remote Sensing, 14(13), p. 3159. https://doi.org/10.3390/rs14133159.