Efficient Wildfire Detection Framework Based on Artificial Intelligence Using Convolutional Neural Network and Multi-Color Filtering

Rabbani Nur Kumoro
Department of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
rabbani.nur.kumoro@mail.ugm.ac.id

Louis Widi Anandaputra
Department of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
louiswidianandaputra@mail.ugm.ac.id

Richardus Ferdian Dita Nugraha
Department of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
r.ferdian.dita.nugraha@mail.ugm.ac.id

Wahyono
Department of Computer Science and Electronics
Universitas Gadjah Mada
Yogyakarta, Indonesia
wahyo@mail.ugm.ac.id

***Abstract*—Wildfire prevails to be a high-risk natural disaster, posing serious damage to both human populations and the environment. As a nation with one of the largest forested regions, Indonesia faces the recurring challenge of wildfires, which annually rank among its top concerns. The significant impact of wildfires on the environment, development, and economic growth motivates this research to create a detection model based on deep learning methods. Through utilizing surveillance tools with the likes of CCTV, UAVs, and satellites, the proposed model aims to pinpoint the precise locations of fire incidents in images, achieved through an Internet of Things device-driven process. This will enable a more efficient and effective response in controlling forest fires, as well as supporting sustainable development. By employing a Convolutional Neural Network-based model using MobileNetV2 with additional fully connected layers for wildfire event classification in images, as well as multi-color filtering for segmenting fire images, the proposed model yielded impressive results. It achieved a remarkable 98.95% accuracy in classifying wildfire images and an Intersection over Union (IoU) score of 0.37 for accurately segmenting specific fire images. These outcomes surpass previous research, underscoring the proposed model's capacity to effectively detect wildfire presence in images and accurately delineate fire boundaries. The proposed model is envisioned to be seamlessly integrated into a system capable of providing real-time information on wildfire locations, serving as an effective mitigation and response solution.**

Keywords*—Artificial Intelligence, Classification, Color Rule, Deep Learning, Forest Fire, Segmentation.

I. INTRODUCTION

Wildfires have become a significant threat in recent times, causing extensive damage to human property, and vegetation, and endangering nearby communities and ecosystems [1]. Indonesia suffered a loss of 72 trillion Rupiah due to wildfires in 2019, which affected 1.6 hectares of forests [2]. This is a significant concern considering that Indonesia has one of the highest forest coverage levels globally, with 49.1% coverage in 2020 [3]. The risk of wildfires is prevalent in tropical countries due to both intentional and unintentional factors, such as forest burning for land clearing, and improper disposal of trash and cigarette butts [1]. These uncontrollable fires spread in the direction of the wind, posing significant risks to wildlife and human populations alike.

Furthermore, wildfires exacerbate climate change, contribute to global warming, and lead to increased air pollution. Therefore, it is crucial to mitigate and manage them not only to protect people and the environment but also to achieve Sustainable Development Goal (SDG) 13, which calls for urgent climate action, and SDG 15, which prioritizes the preservation and restoration of endangered forest ecosystems [4]. The Indonesian government has implemented various strategies to mitigate wildfires, especially by enhancing early detection measures, but they have not fully utilized every aspect of technology such as the use of artificial intelligence (AI) [5] despite its recent advancements. Digital image-based fire detection systems could be used by the government as a key stakeholder, in an effort to prioritize early detection and surveillance of wildfires and to effectively manage ongoing wildfire risks.

Previous research has worked with various methods to detect wildfires through images. A fast and efficient method using image processing techniques to segment and detect wildfires in digital images was proposed by introducing a color-rule model and video-based detection [6]. This method would be suitable for a live wildfire mitigation system. Another research proposed a deep learning fire detection technique that combines a quad-tree search with deep learning classification and the segmentation stage [7]. While [6] would have an advantage in terms of real-time performance as it is fast, it may need improvement in terms of the accuracy and IoU measure. In another case, [7] would have stronger detection results but it may lack fast real-time performance measures. Therefore, a strong but fast wildfire detection technique needs to be addressed.

This research aims to address the problem by proposing a strong deep learning-based classification technique followed by fast image processing methods to segment fire images. We introduce a model that uses image processing methods on a modified MobileNetV2 model. The model is enhanced with fully connected layers, color-rule-based segmentation, and morphological operations. By using an effective model architecture, this research proposes a solution to mitigate wildfire risks in Indonesia. The proposed model can process image inputs from various surveillance tools such as CCTV, UAVs, and satellites, and is projected to perform without intensive computational resources while also being able to

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00095

pinpoint the exact location of a fire in an image. This feature helps to respond more efficiently and effectively in the efforts to prevent and extinguish wildfires. In summary, this research can contribute to sustainable and advanced development of the usage of AI in Indonesia.

II. METHODOLOGY

A. Overview

Detecting fire in images begins with the classification process. The segmentation process in **Fig. 1** is initiated only if the classification results in a fire image. In the case of a no-fire image, a blank matrix is returned. The segmentation process has two distinct steps - the first step involves filtering the image using a color rule based on [8]. The second step involves performing morphological operations on the images using closing and dilation techniques. The result of this process is the segmentation masks.

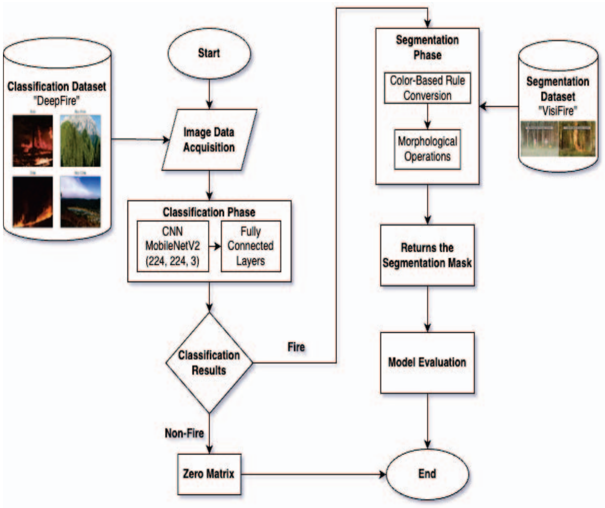

Fig. 1. Diagram of the Proposed Methodology

B. Data Acquisition

This research will utilize two datasets, namely DeepFire [9] and VisiFire [10], to develop a wildfire detection model. The DeepFire dataset comprises 1900 images, with 950 images depicting fire and non-fire situations each. These images have been divided into 931 training data, 570 testing data, and 399 validation data. This dataset is an ideal source to train and evaluate the wildfire classification model, as the images closely resemble real-life fire scenarios, as seen in **Fig. 2**. On the other hand, the VisiFire dataset consists of 2544 image data and 2544 ground truth (GT) data. This dataset will be used to evaluate the segmentation process, as shown in **Fig. 3**, which contains images of frames taken from various videos. Both datasets will be separately employed in the classification and segmentation processes.

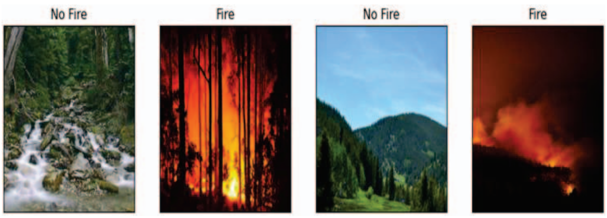

Fig. 2. Sample Dataset from DeepFire [9]

Fig. 3. Sample Dataset from VisiFire [10, 11]

C. Classification

This research proposed the MobileNetV2 model as the backbone network for classifying wildfires in forests. Many CNN models require high computational power and a significant amount of data for optimal training and accurate classification [12]. To address this issue, transfer learning, a subset of deep learning, has shown promising results [13]. The MobileNet is a lightweight deep learning model that is optimized to fulfill computation resource thresholds for various use cases, with low latency and power consumption. As presented by [12], MobileNetV2, designed for mobile devices and embedded systems, has shown better performance in different benchmarks and model sizes on the latest mobile models.

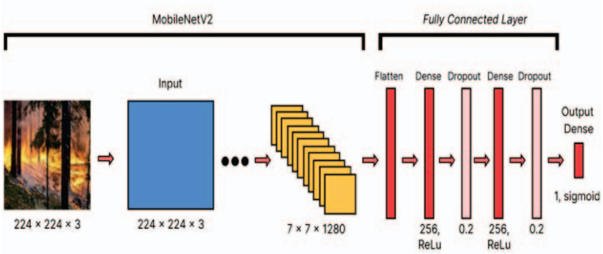

Fig. 4. Architecture of MobileNetV2 + Fully Connected Layer

In terms of transfer learning, MobileNetV2 had been trained beforehand on various datasets to classify wildfires. To help the model understand important image features, fully connected layers were added based on experimentation. The structure of the model is shown in **Fig. 4**. The model includes a dense layer with 256 neurons and a ReLu activation function, followed by a dropout layer with a dropout value of 0.2. This layer represents the previously learned features. Additionally, a second dense layer with 256 neurons and a ReLu activation function is added, followed by a dropout layer with the same value. Finally, a Dense layer with 1 neuron unit and a Sigmoid activation function is added to perform the final image classification.

D. Segmentation

The segmentation process aims to isolate areas detected as fire in images classified as fire through the classification process. The output of this process is a binary image in the form of fire masks. This stage focuses on the color dimension of the fire, which undergoes two stages. The first stage is based on the 1) color rules and the second stage involves 2) morphological operations.

1) Color-Based Rule

Fire has unique color characteristics that distinguish it from other objects. These characteristics can be evaluated through HSV, YCbCr, and RGB color channels. By utilizing this information, color-based rules can be obtained, as presented by [8], which can be used to isolate fire in images. The first rule represents the color of fire in HSV color space, as stated by (1). Then, a rule for YCbCr color space is represented by (2). Finally, (3) and (4) represent rules for RGB color space. The result of this process is the unification of (1), (2), (3), and (4), as shown in (5).

$$S1_{(i,j)} = 1, \text{ if } \quad (1)$$

$$[0 \leq H_{(i,j)} \leq 0.46] \text{ and } [0.1 \leq S_{(i,j)} \leq 0.34] \text{ and }$$

$$[0.96 \leq V_{(i,j)} \leq 1] \text{ and } [R_{(i,j)} > 180] \text{ and }$$

$$[G_{(i,j)} > 130] \text{ and } [B_{(i,j)} < 120]$$

$$S2_{(i,j)} = 1, \text{ if } \quad (2)$$

$$[Cb_{(i,j)} \leq Cr_{(i,j)} < Y_{(i,j)}] \text{ and } [180 \leq Y_{(i,j)} < 210] \text{ and }$$

$$[80 \leq Cb_{(i,j)} \leq 120] \text{ and } [80 \leq Cr_{(i,j)} < 139] \text{ and }$$

$$[R_{(i,j)} > 190] \text{ and } [G_{(i,j)} > 110] \text{ and } [B_{(i,j)} < 110]$$

$$S3_{(i,j)} = 1, \text{ if } \quad (3)$$

$$[R_{(i,j)} > 200] \text{ and } [G_{(i,j)} > 130] \text{ and } [B_{(i,j)} < 120]$$

$$S4_{(i,j)} = 1, \text{ if } \quad (4)$$

$$[B_{(i,j)} \leq G_{(i,j)} < R_{(i,j)}]$$

$$S5_{(i,j)} = (S1_{(i,j)} \cup S2_{(i,j)} \cup S3_{(i,j)}) \cap S4_{(i,j)} \quad (5)$$

2) Morphological Operations

As implementation of 1) returned a less desirable segmentation output, in which the mask looked too small compared to the original size with the addition of holes, morphological operation was needed. Closing and Dilatation were chosen as the designated operations for the mask to cover up the hollow area and increase the segmentation result area as performed by Wahyono et al [14]. This process is done using a rectangular kernel. With the same output as the previous process, this step would return a binary mask, isolating fire from other objects.

E. Evaluation Metrics

In this research, the performance and classification of the model were evaluated using accuracy and F1-Score as metrics. Accuracy measures the ability of the MobileNetV2 model to correctly classify fire images by using *True Positive* (TP), *True Negative* (TN), *False Negative* (FN), and *False Positive* (FP) values. To calculate the accuracy, the sum of TP and TN is divided by the total number of data and then multiplied by 100 to obtain a percentage. On the other hand, F1-Score combines precision and recall metrics to provide a more comprehensive understanding of the model's performance. Precision represents true positive predictions, while recall represents the model's ability to identify positive data. The F1-Score is calculated using a formula that combines precision and recall [15].

In addition, to evaluate the quality of the segmentation process, the Intersection over Union (IoU) metric is used. The IoU score calculates the extent to which the segmentation result intersects with the ground truth [16]. A higher IoU score indicates a more accurate segmentation process. The IoU is calculated by dividing the intersecting area by the union of both areas, as shown in equation (6). A good level of intersection between the prediction and ground truth indicates good segmentation quality. By using this metric, the performance of creating object masks through segmentation can be evaluated, ensuring a high-quality segmentation method.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

III. Results and Discussions

This section showcases the evaluation metrics that were used to measure the performance of the model in carrying out classification and segmentation tasks. The evaluation metrics used in measuring the model's performance were Accuracy, F1-Score, and IoU. Furthermore, a comparison of the results with previous research is presented to highlight the advantages of the proposed model. Next, the model implementation proposal will be explained as an early detection system for forest fires. This will illustrate how the model can be applied in real-life situations to improve effectiveness in preventing and detecting wildfires.

A. Classification Results

The training results with the utilization of MobileNetV2 and fully connected layers showed better results than some deep learning architectures. This is obtained through a single run with a testing set consisting of 570 images. The classification results achieved by the model used in this study and their comparisons with previous research models are displayed in **Table I**. The MobileNetV2 + FC model achieved an accuracy rate of 98.95% and an F1-Score of 0.989, demonstrating excellent performance in classifying fire and non-fire images. This model proved to be more efficient as it uses the MobileNetV2 architecture, which is designed specifically for mobile devices with fewer parameters.

TABLE I
MOBILENETV2 + FC CLASSIFICATION RESULTS

Model	*Accuracy*	*F1-Score*
MobileNetV2 + Fully Connected Layer	**98,95%**	**0,989**
InceptionV3 [17]	96,32%	0,963
ResNet152V2 [18]	96,32%	0.964

In comparison to previous research models such as InceptionV3 [17] and ResNet152V2 [18], the proposed model demonstrates better performance. The previous models achieved an accuracy of 96.32% with F1-Scores of 0.963 and 0.964, respectively. These findings suggest that the use of the proposed MobileNetV2 + FC in this study provides a significant enhancement in fire classification tasks.

TABLE II
CONFUSION MATRIX OF MOBILENETV2 + FC

	Predicted Label	
Actual Label	**Fire**	**Non-Fire**
Fire	284	4
Non-Fire	2	280

Table II shows that the MobileNetV2 model with fully connected layers accurately classified the presence of fire. The model predicted 284 TP cases and 280 TN cases, indicating a high level of accuracy. However, there were also 2 FN predictions and 4 FP predictions, where the model misclassified the image with the wrong label. Despite these errors, it can be concluded that the MobileNetV2 + FC model is successful in detecting the presence of fire, as demonstrated in **Fig. 5**.

Fig. 5. Classification Results of MobileNetV2 + FC

B. Segmentation Results

The multi-color filtering method proposed in this study demonstrated superior performance over previous research that utilized the Region-based Convolutional Neural Network (R-CNN) method and the Gaussian Mixture Models-Expectation-Maximization (GMM-EM) color probability method. Previous research reported an IoU score of 0.26 for the R-CNN method and 0.32 for the GMM-EM color probability method [14]. However, with the proposed method, the IoU score increased to 0.37, as shown in **Table III** which provides a comparison with other methods. Therefore, the use of constraints on color channel values to filter fire pixels is considered more effective than probability methods like GMM-EM or prediction using neural networks such as R-CNN.

TABLE III
MULTI-COLOR FILTERING SEGMENTATION RESULTS

Method	*IoU Score*
Multi-Color Filtering	**0,37**
Color Probability with GMM-EM [14]	0,32
R-CNN [14]	0,26

This method has an advantage in terms of efficiency. It doesn't require a large amount of computational power compared to deep learning or other prediction methods. The combination of color segmentation and morphological operations provides better results. The use of multicolor filtering can accurately determine fire pixels which is more reliable than GMM-EM and R-CNN methods. This is because the prediction process performed by GMM-EM or R-CNN has the possibility to incorrectly predict the fire image. Meanwhile, the multi-color filtering process used in this study uses fire color restrictions based on research [8]. Therefore, the multi-color filtering method followed by morphological operations can provide a reliable solution for detecting and improving segmentation accuracy in fire images as shown in **Fig. 6**.

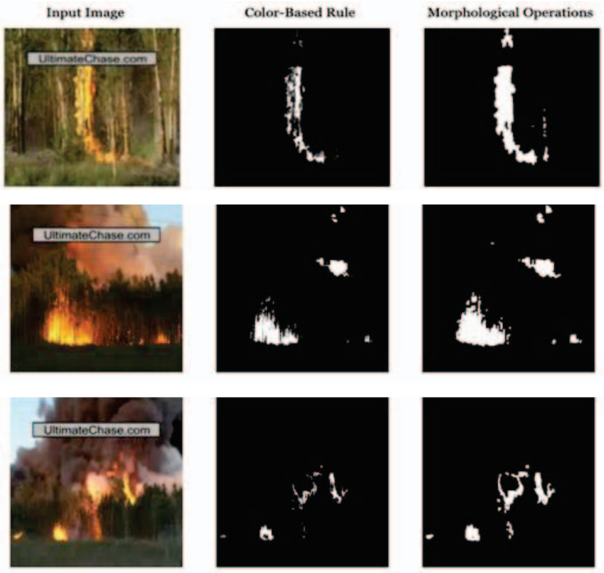

Fig. 6. Segmentation Results of Multi-Color Filtering

During testing, there are several challenges in identifying positive fire images, such as varying illumination and the presence of fumes. The issue arises due to limitations in our implemented rules, as well as the presence of non-fire objects with similar color characteristics, which can affect the accuracy of our segmentation process. Despite these challenges, the implemented color filtering rules can accurately segment fire areas in most cases.

C. Computational Cost Analysis

Efficient fire detection is critical for real-world deployment, especially when targeting low computing power devices like surveillance camera systems. In this section, a comparative analysis of the computational cost associated with three fire detection approaches is presented. The investigation was carried out using a base model MacBook Pro equipped with the Apple M1 chip.

Table IV presents a breakdown of processing time, highlighting the efficiency of the proposed method that utilizes a combination of MobileNetV2 classification and Multi-Color Filtering segmentation. This approach strikes a balance between computational efficiency and accuracy, making it ideal for low computing power devices.

TABLE IV
COMPUTATIONAL TIME COMPARISON

Method	*Average Computational Cost (CPU Time in Milliseconds)*
MobileNetV3 + DeepLabV3	299
MobileNetV2 + Multi-Color Filtering	**72,3**
YOLOv8 + Color Rule	169

The proposed method is being compared to several other wildfire detection techniques that have two steps. These methods first classify an image and then segment it to detect the wildfire. The first comparison is between the proposed method and a full deep learning detection method that uses

MobileNetV3 for classification and DeepLabV3 for segmentation. Additionally, the proposed method is also compared to another method that uses a more complex classification technique in YOLOv8.

Out of the other methods, the proposed method that uses MobileNetV2 for classification and multi-color filtering for segmentation has the lowest computational cost out of all the other methods. It takes only 72.3 milliseconds (ms) of CPU processing time, compared to the other two methods which require 299 ms and 169 ms of CPU processing time respectively. This efficiency is particularly crucial for real-time fire detection, especially on devices with limited computing power, such as surveillance camera systems.

Addressing the metric used, CPU time is used as it measures only the time during which the CPU executes the process and does not take into account the time spent waiting for network I/O or memory, nor the existence of multi-processing. Therefore, future research could focus on implementing methods that use well-optimized multi-processing techniques in order to reduce computation time even further.

D. Proposed Implementation

An effective solution to address the problem of wildfires is to implement an early detection system that uses a classification and segmentation model approach for detecting and providing warnings quickly. The system is designed to capture images of forests using Internet of Things (IoT) devices such as CCTV, UAV, and satellite. These images will be processed using the proposed MobileNetV2 model to classify whether there is a fire or not. Once classified, images indicating a fire will undergo the segmentation process to identify and delineate the fire area in more detail. The image data containing information about the fire will be sent to the central control room. The system can send online alerts to the Meteorology, Climatology, and Geophysics Agency in the control room, serving as an early warning system.

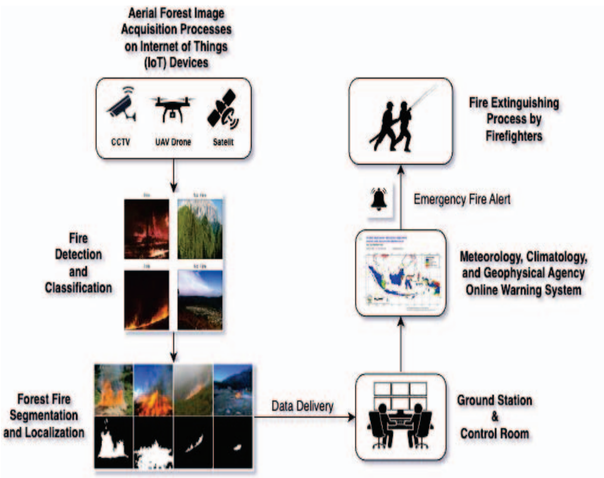

Fig. 7. Forest Fire Detection Operation Flow Diagram

A proposed system is expected to automatically detect forest fires and alert the authorities promptly. This early warning system will enable the agency or other authorities to take immediate action and provide firefighters with the necessary information to rescue and put out fires quickly. The operation flow for early detection is illustrated in **Fig. 7**, and it is hoped that this system will reduce losses caused by forest fires while ensuring environmental sustainability and community safety.

IV. Conclusion

In detecting wildfires, a fast but powerful detection method is crucial. This research aims to propose a fast but powerful wildfire detection method by introducing a method utilizing the classification of fire images followed by a segmentation step to identify the wildfire. A method utilizing MobileNetV2 for classification and multi-color filtering for segmentation is proposed. Based on this research, the proposed MobileNetV2 model has an accuracy rate of 98.95% in classifying forest fires. In addition, the use of multi-color filtering in image processing for fire segmentation also yields satisfactory results of 0.37 IoU score compared to previous research models. The proposed model is efficient and has the potential to be integrated into IoT devices, making it possible to verify wildfire events.

For future research, it is recommended to address the challenges associated with real-time fire detection, such as illumination variations and the presence of fumes, which may affect the proper identification of fire regions. This could involve exploring the integration of robust feature extraction techniques and advanced learning algorithms like YOLO or other SSD methods for real-time fire tracking. Additionally, it is suggested to collect a more structured dataset and involve other efficient deep learning algorithms to enhance the model's performance in classifying various features relevant to forest fires. One approach could involve expanding the availability of features for classification, such as smoke and fire movement. Furthermore, the proposed model could be implemented in real-life situations using low computation-powered devices to be evaluated. Hence, this will make the model more effective in detecting wildfires.

Acknowledgment

We are immensely grateful to the anonymous reviewers for generously sharing their time and expertise with us. Their invaluable feedback has undoubtedly helped us to refine our paper and take it to new heights. We recognize their contributions have been pivotal, and we want to express our sincere appreciation for their efforts.

References

[1] M. Hendri, "Perancangan Sistem Deteksi ASAP Dan Api Menggunakan Pemrosesan Citra," DSpace Repository UII, https://dspace.uii.ac.id/bitstream/handle/123456789/12602/14524051_skripsi.pdf (accessed Jun. 19, 2023).

[2] V. A. Dihni, "Luas Kebakaran Hutan dan Lahan RI Bertambah 19% pada 2021," Pusat Data Ekonomi dan Bisnis Indonesia, https://databoks.katadata.co.id/datapublish/2022/01/11/luas-kebakaran-hutan-dan-lahan-ri-bertambah-19-pada-2021 (accessed Jun. 19, 2023).

[3] D. H. C. Chen et al., "December 2019 Indonesia Economic Quarterly: Investing in people," World Bank, https://www.worldbank.org/en/country/indonesia/publication/december-2019-indonesia-economic-quarterly-investing-in-people/ (accessed Jun. 19, 2023).

[4] "Transforming our world: the 2030 Agenda for Sustainable Development," United Nations, https://sustainabledevelopment.un.org/post2015/transformingourworld/publication (accessed Jun. 19, 2023).

[5] A. Lestari, G. Rumantir, and N. Tapper, *Deteksi Dini Kebakaran Hutan dan Lahan di Kalimantan Tengah*, vol. 6, no. 2, pp. 128–134, Nov. 2015.

[6] G. Perrolas, M. Niknejad, R. Ribeiro, and A. Bernardino, "Scalable fire and smoke segmentation from aerial images using convolutional neural networks and quad-Tree Search," *Sensors*, vol. 22, no. 5, p. 1701, 2022. doi:10.3390/s22051701.

[7] T. Celik, "Fast and efficient method for fire detection using image processing," *ETRI Journal*, vol. 32, no. 6, pp. 881–890, 2010. doi:10.4218/etrij.10.0109.0695.

[8] Wahyono, A. Dharmawan, A. Harjoko, Chrystian, and F. D. Adhinata, "Region-based annotation data of fire images for Intelligent Surveillance System," *Data in Brief*, vol. 41, p. 107925, 2022. doi:10.1016/j.dib.2022.107925.

[9] Wahyono *et al.*, "Real-time forest fire detection framework based on artificial intelligence using color probability model and motion feature analysis," *Fire*, vol. 5, no. 1, p. 23, 2022. doi:10.3390/fire5010023.

[10] G. F. Shidik, F. N. Adnan, C. Supriyanto, R. A. Pramunendar, and P. N. Andono, "Multicolor feature, background subtraction and time frame selection for fire detection," *2013 International Conference on Robotics, Biomimetics, Intelligent Computational Systems*, 2013. doi:10.1109/robionetics.2013.6743589.

[11] A. Khan, B. Hassan, S. Khan, R. Ahmed, and A. Abuassba, "DeepFire: A novel dataset and deep transfer learning benchmark for forest fire detection," *Mobile Information Systems*, vol. 2022, pp. 1–14, 2022. doi:10.1155/2022/5358359.

[12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018. doi:10.1109/cvpr.2018.00474.

[13] A. E. Çetin, "Computer Vision Based Fire Detection Software," Bilkent Üniversitesi, http://signal.ee.bilkent.edu.tr/VisiFire/ (accessed Jun. 16, 2023).

[14] L. Mihalkova and R. J. Mooney, "Transfer Learning from Minimal Target Data by Mapping across Relational Domains," *International Joint Conference on Artificial Intelligence*, pp. 1163–1168, Jul. 2009.

[15] Z. Daoud, A. Ben Hamida, and C. Ben Amar, "FireClassNet: A deep convolutional neural network approach for PJF Fire Images Classification," *Neural Computing and Applications*, vol. 35, no. 26, pp. 19069–19085, 2023. doi:10.1007/s00521-023-08750-3.

[16] M. Imad, O. Doukhi, and D.-J. Lee, "Transfer learning based semantic segmentation for 3D object detection from Point Cloud," *Sensors*, vol. 21, no. 12, p. 3964, 2021. doi:10.3390/s21123964.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. doi:10.1109/cvpr.2016.308.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *Computer Vision – ECCV 2016*, pp. 630–645, 2016. doi:10.1007/978-3-319-46493-0_38.