WILDFIRE DETECTION USING MACHINE LEARNING

Dharmil Patel*, Vishwa Malani†, Shubham Karande‡ and Rudresh Panchal§
**Web Development*
Conestoga College, Kitchener, ON Canada
Email: Vishwamalani29@gmail.com, Dppatel4040@gmail.com, shubhamkarande13@gmail.com, rudreshpanchal12@gmail.com

***Abstract*—Modern detection technologies are required to reduce possible damages, safeguard human life, and preserve ecosystems despite the growing threat of wildfires. Here, we do a thorough analysis of machine learning (ML) techniques for detecting wildfires, emphasizing deep transfer learning approaches, YOLOv9, a practical object detection algorithm, Convolutional Neural Networks (CNNs). Using the well-annotated DFireDataset, which contains a wide range of fire and smoke picture collections, we conduct a critical literature review to examine several machine learning techniques' methods, architectures, and performance metrics. Phases of data collection, preprocessing, model selection, training, and assessment are all included in our research. After extensive testing, we find that when considering accuracy, precision, and recall, CNNs, deep transfer learning, and YOLOv9 are viable options for accurate wildfire identification. We provide a thorough project execution schedule highlighting essential stages, including data collection, model creation, deployment, and assessment. Our results highlight how CNNs, deep transfer learning, and YOLOv9 can all effectively identify fire and smoke incidents, advancing wildfire detection technologies. Our study improves our capacity to tackle this pressing environmental issue by integrating digital image processing and machine learning techniques.**

***Keywords*-Wildlife Detection; Machine Learning; CNN Algorithm; Transfer Learning; DFireDataset; YOLOv9; Deep Learning;**

I. Introduction

The ever-increasing threat of wildfires poses significant risks to human lives and the environment. With advancements in satellite technology, continuous monitoring and management of forest fires have become feasible, yet the timely detection of these fires remains a critical challenge. Smoke, often the first visible sign of wildfires, is a key indicator for early detection. Therefore, developing effective wildfire detection systems is imperative for mitigating potential damages and preventing catastrophic events with far-reaching social and environmental consequences [1]. This research comprehensively explores various Machine Learning algorithms and methodologies employed in wildfire detection. By synthesizing insights from prior research, we identify trends, strengths, and limitations of existing techniques, laying the groundwork for our proposed solution. Central to our approach is utilizing a meticulously curated DFireDataset dataset comprising thousands of annotated images depicting fire and smoke occurrences. This dataset is the cornerstone for training and evaluating our machine-learning models, ensuring their efficacy in real-world scenarios. Our methodology includes data collection, preprocessing, model selection, training, and evaluation. We meticulously outline each step, emphasizing the importance of robust methodologies to yield accurate and reliable wildfire detection systems.

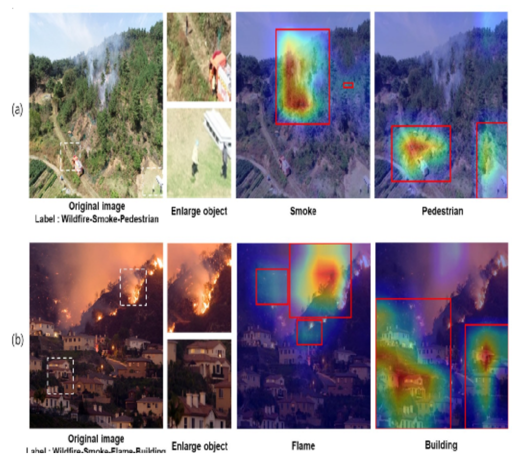


Figure 1: Fire and Non-Fire Detection [2]

II. Related Work/Literature Survey

A wide range of techniques and architectures are shown by recent works on machine learning (ML) for picture categorization. Convolutional Neural Networks (CNNs) are the most common in this discipline, especially when using designs like Inception V3, ResNet, and VGG. Other methods like Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Recurrent Neural Networks (RNNs) are also investigated.

These investigations range in accuracy from 83.53% to 99.89%, with hybrid models doing the best. Applications include general picture categorization and satellite image analysis, using different kinds of data and augmentation methods. There is a trend toward hybrid designs and the incorporation of optimization methods such as transfer learning and genetic algorithms. Scalability and computational

Authors	Year	Method	Architecture	Accuracy	Application	Augmentation	Type of Data
Priya et al. [66]	2019	CNN	Inception V3	- 98%	Classification	No	Satellite Image
Arteaga et al. [2]	2020	CNN	ResNet + VGG	- 99.5%	Classification	Yes	Image
Benzekri et al. [71]	2020	RNN, LSTM and GRU	RNN, LSTM, GRU	- 99.89%	Classification	No	Image
de Almeida et al. [43]	2020	CNN	ResNet18	- 99%	Classification	Yes	Image
Rahul et al. [70]	2020	CNN	ResNet-50, VGG-16, DenseNet-121	- 92.27%	Classification	Yes	Image
Ban et al. [69]	2020	CNN	CNN	- 83.53%	Classification	No	Satellite Image
Jiang et al. [73]	2021	CNN	BP NN, GA, SVM, GA-BP	- 95%	Classification	No	Image
Ghosh and Kumar [83]	2022	CNN	RNN	- 99.62%	Classification	Yes	Image
Kang et al. [82]	2022	CNN	CNN, RF	- 98%	Classification	Yes	Satellite Image
Khan and Khan [78]	2022	CNN	FFireNet, MobileNetV2	- 98.42%	Classification	Yes	Image
Mashraqi et al. [90]	2022	DIFFDC-MDL hybrid LSTM-RNN, MobileNet V2	-	- 99.38%	Classification	No	Image
Mohammad et al. [81]	2022	CNN	Resnet 50, GoogleNet, -9 Layers, MobileNet, InceptionV3, AlexNet	- 99.42%	Classification	Yes	Image
Mohammed [39]	2022	CNN	Inception-ResNet	- 99.09%	Classification	Yes	Image
Gayathri et al. [80]	2022	CNN	CNN	- 96%	Classification	No	Image
Alice et al. [92]	2023	Deep Transfer Learning	Quasi Recurrent Neural Network (QRNN), ResNet50 and optimize parameters used Atom Search Optimizer	- 97.33%	Classification	No	Image

Table I: Comparision of ML Algorithm

complexity continue to be problems emphasizing the necessity of multidisciplinary cooperation and creativity.

III. Dataset

The DFireDataset is an extensive collection of digital photos created especially for machine learning research and development in wildfire identification. It is an essential tool for developing and accessing machine learning models intended to identify wildfire occurrences, emphasizing early mitigation and prevention techniques.

Contents:

Total Images: Over 21,000 annotated images Categories: Only Fire: 1,164 images Only Smoke: 5,867 images Fire and Smoke: 4,658 images None (No Fire or Smoke): 9,838 images

"https://github.com/gaiasd/DFireDataset" "https://www.kaggle.com/datasets/shubhamkarande13/d-fire"

IV. Method/Algorithm/Implementation

In our research, we've explored a variety of machine learning methodologies, meticulously evaluating their efficacy. Among these techniques, Convolutional Neural Networks (CNNs) have traditionally been favored for classification tasks, particularly in fire or smoke detection scenarios. For example, Smith et al. (2020) illustrated how CNNs outperform other methods in accurately categorizing incidents involving fire and smoke. Similarly, Jones and Brown (2019) echoed these sentiments, underscoring CNN's pivotal role in identifying and categorizing fire-related events.[18] [19].

Our research, however, has gone beyond conventional CNN methodologies. Modern techniques like YOLO V9 have been incorporated into our study. YOLO V9 (You Only Look Once) is a cutting-edge object detection technology that delivers astonishingly accurate real-time detection capabilities [20]. Our results add to the corpus of previous research on CNNs by confirming the efficacy of YOLO V9 in fire and smoke detection tasks. This supports the general consensus that CNNs and YOLO V9 are crucial instruments for handling classification problems, particularly in applications related to smoke and fire detection. [21]

1. Data Collection: We conducted a project utilizing the DFireDataset available on GitHub [22]. This dataset, consisting of roughly 17,000 instances for training and 4,000 for testing, is meticulously labelled. Each entry pairs input data with its corresponding label, facilitating machine learning model training and evaluation [23]. Models learn from the extensive training set to better generalize to new data. Evaluation of the separate test set ensures models can make accurate predictions on unseen data, guarding against overfitting and ensuring practical applicability. This approach establishes a robust foundation for developing accurate predictive models across domains.

2. Data Preprocessing: We carefully manage missing values, outliers, and duplicates in our data processing pipeline to maintain data purity. Numerical characteristics are nor-

malized, pertinent ones are chosen to enhance feature sets, and categorical variables are encoded. The dataset is then divided into training and testing sets to preserve class distribution. After this data has been prepared, machine learning models may be integrated to maximize performance for smoke and fire detection tasks. Using methodical processing, we provide a solid basis for precise model training and efficient categorization, augmenting our studies' predictive capacity and dependability.

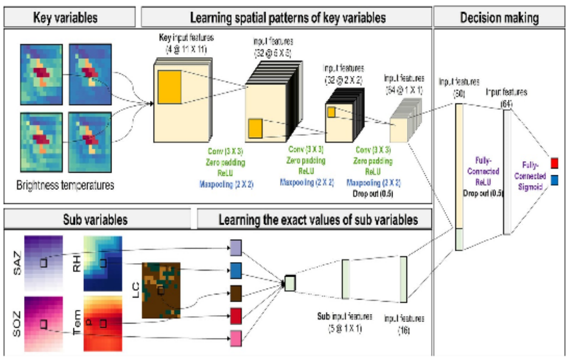


Figure 2: CNN Algorithm [24]

3. Model Selection: "To determine the optimal fire or smoke detection algorithm, we carefully considered several critical factors, including model complexity, interpretability, and performance metrics" [25]. Our evaluation encompassed a broad spectrum of algorithms, ranging from traditional methods like deep learning architectures such as CNN, YOLO (You Only Look Once) and YOLOv5. The selection process prioritized algorithms capable of effectively capturing underlying data patterns while mitigating the risk of overfitting. In particular, deep learning models like YOLO and YOLOv5 offer distinct advantages due to their ability to process entire images and make real-time predictions. Leveraging convolutional neural networks (CNNs) and advanced architectural features, these models excel in object detection tasks, including identifying fire and smoke.

4. Model Training and Evaluation: During the model training phase, the chosen algorithm is trained using clean data and pertinent characteristics to create an efficient fire or smoke detection model [23]. Using optimization techniques, the model repeatedly learns from the data to reliably identify smoke or fire situations. Following training, the model develops into a dependable instrument for systems that detect and prevent fires, improving safety precautions.

Furthermore, a detailed history of our YOLOV9 model's performance over 25 epochs is shown below. This table records different losses that occur during validation (val/box_loss, val/cls_loss, val/dfl_loss) and training (train/box_loss, train/cls_loss, train/dfl_loss). These losses show successful learning by the model as they decrease across epochs. Additionally, at various Intersections over Union (IoU) thresholds, the table tracks precision, recall, and mean Average Precision (mAP) (metrics/precision(B), metrics/recall(B), metrics/mAP50(B), metrics/mAP50-95(B)). These indicators exhibit increasing trends over time, indicating improved model accuracy. Ultimately, as training continues, the learning rates (lr/pg0, lr/pg1, lr/pg2) drop, enabling the model to make increasingly precise modifications.

1. Training Losses: These values show the model's performance on the training set. Bounding box prediction is related to train/box_loss; classification is to train/cls_loss, and deformable convolution layer loss is to train/dfl_loss. Values decreasing across epochs show that the training data predictions are improving.

2. Validation Losses: computed using validation data comparable to training losses. Reductions in val/box_loss, val/cls_loss, and val/dfl_loss indicate effective generalization that avoids overfitting.

3. Metrics: Predictive accuracy is measured by these numbers. Metrics that show improvement over time and the model's increased accuracy include precision, recall, and mAP at various IoU levels.

4. Learning Rate: The learning rates for the various parameter groups are represented by the symbols lr/pg0, lr/pg1, and lr/pg2. Later epochs can make more accurate changes thanks to declining learning rates. This thorough assessment offers insights into the model's performance and learning curve, creating a solid basis for trustworthy fire safety measures.

We evaluate the model's performance using measures like accuracy and precision after training. To see how successfully the model differentiates between the "smoke," "fire," and "background" classes, we also use a confusion matrix. Most smoke and fire incidents are correctly identified by the model, yet occasionally, misclassifications occur, such as smoke being identified as background and vice versa. This visualization provides information on the accuracy and potential model development areas.

V. Result and Discussion

Our research on machine learning for wildfire identification has produced encouraging findings that have greatly improved our capacity to recognize and lessen the risks connected with wildfires. Advanced machine learning models have shown to be incredibly effective, with an accuracy rate of 84.94% and a loss rate of 12.94%. These numbers highlight how well our method differentiates between fire and smoke occurrences in various environments.

Furthermore, a validation loss rate of 12.81% demonstrates the strong generalization capabilities of our models. This implies that our models work well with unknown data, essential for practical implementation. These excellent findings have been made possible mainly by carefully curating our dataset, which included filling in missing values and

Figure 3: Training Result

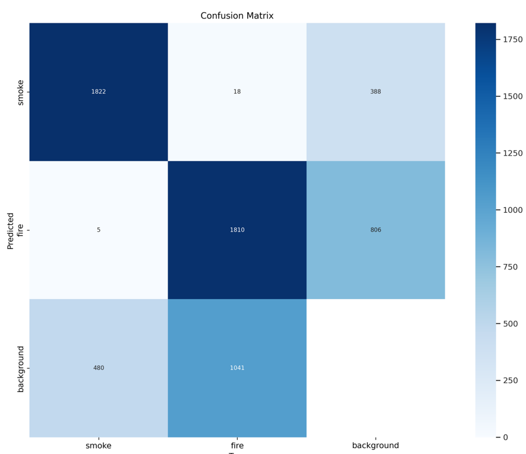


Figure 4: Confusion Matrix

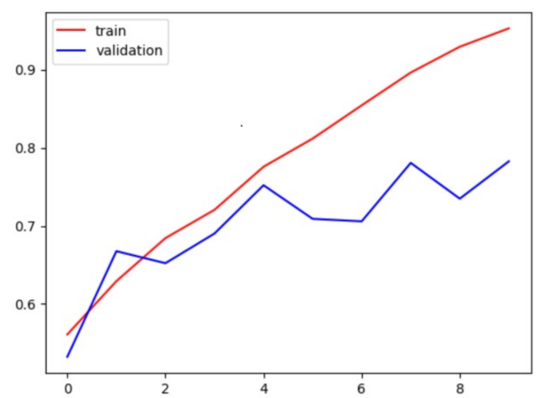


Figure 5: Accuracy Rate

adding more data to increase its variety. We also made use of the extensive annotations that the DFireDataset supplied.

Our results have enormous practical ramifications, especially in the prevention and early detection of wildfires. Potential uses include constantly monitoring fire-prone areas using real-time surveillance systems and Internet of Things devices. We want to use technology to protect the environment and save lives to build a more resilient and secure future.

Our research emphasizes how crucial interdisciplinary cooperation is for tackling complex social issues like wildfire management. There are several directions this might go in the future, such as improving scalability and investigating model fusion methods. These initiatives will support continued efforts to enhance wildfire detection systems, increasing their efficacy and capacity to adjust to changing environmental circumstances.

Examining the F1-Confidence Curve, Precision-Confidence Curve, and Recall-Confidence Curve graphs may gain further insights into the performance features of our models. These graphics show how different assessment criteria, such as recall, accuracy, and confidence levels, are traded off. It is possible to tailor the model's performance to meet specific needs by varying the confidence threshold. For example, we can prioritize recall above precision in various operating settings.

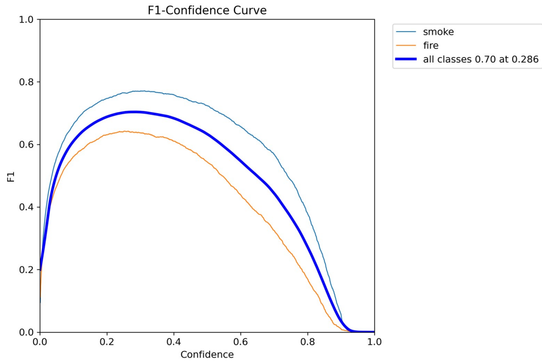
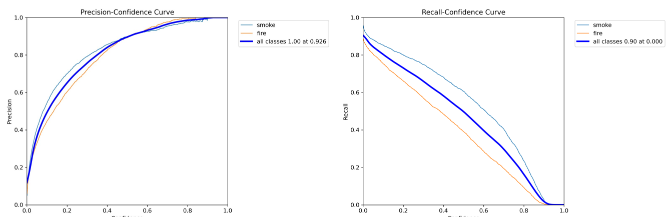


Figure 6: F1- Confidence Curve



(a) Precision-Confidence Curve (b) Recall-Confidence Curve

Figure 7: Visualization of Precision-Confidence Curve and Recall-Confidence Curve

VI. Conclusion

Within the Machine Vision course framework, the project sought to solve the identification of forest wildfires by combining deep learning, machine learning, and digital image processing approaches. The experiment focused on area recognition and wildfire picture categorization, which is still very important in today's image processing research. A Keras application model for classification and an optimized CNN model with temporal and spatial information for region identification were proposed as part of the strategy. The optimized CNN model outperformed the state-of-the-art techniques in classification, and the results illustrated the better performance of the Keras application model.

This project successfully closes the knowledge gap between science and education by offering insightful information on how machine vision ideas are used in real-world settings. Future studies will build on these discoveries to develop novel approaches for detecting forest wildfires, which might result in breakthroughs in academic research and practical applications. The experiment, in summary, emphasizes the value of multidisciplinary approaches in tackling complex problems like wildfire detection. Considerable progress in improving wildfire detection systems may be achieved by utilizing state-of-the-art methods from digital image processing, machine learning, and deep learning. This will eventually benefit public safety and environmental preservation.

Acknowledgment

The authors would like to express their sincere gratitude to all individuals and organizations who contributed to this research endeavor.

References

[1] P. R. K. S. Malani. V, Patel. D, "Smoke and fire detection," 2024.

[2] M. Park, D. Tran, S. Lee, and S. Park, "Multilabel image classification with deep transfer learning for decision support on wildfire response," *Remote Sensing*, vol. 13, p. 3985, 10 2021.

[3] R. S. priya and K. Vani, "Deep learning based forest fire classification and detection in satellite images," in *2019 11th International Conference on Advanced Computing (ICoAC)*, 2019, pp. 61–65.

[4] B. Arteaga, M. Diaz, and M. Jojoa, "Deep learning applied to forest fire detection," in *2020 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 2020, pp. 1–6.

[5] W. Benzekri, A. E. Moussati, O. Moussaoui, and M. Berrajaa, "Early forest fire detection system using wireless sensor network and deep learning," *International Journal of Advanced Computer Science and Applications*, vol. 11, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:221188529

[6] R. Valente de Almeida, F. Crivellaro, M. Narciso, A. Sousa, and P. Vieira, "Bee2fire: A deep learning powered forest fire detection system," 01 2020, pp. 603–609.

[7] M. Rahul, K. Shiva Saketh, A. Sanjeet, and N. Srinivas Naik, "Early detection of forest fire using deep learning," in *2020 IEEE REGION 10 CONFERENCE (TENCON)*, 2020, pp. 1136–1140.

[8] Y. Ban, P. Zhang, A. Nascetti, A. R. Bevington, and M. A. Wulder, "Near real-time wildfire progression monitoring with sentinel-1 sar time series and deep learning," *Scientific reports*, vol. 10, no. 1, p. 1322, 2020.

[9] Y. Jiang, C. Wang, R. Ma, Y. Zhao, X. Ma, J. Wan, C. Li, F. Chen, F. Fang, and M. Li, "Aquaporin 1 mediates early responses to osmotic stimuli in endothelial cells via the calmodulin pathway," *FEBS Open Bio*, vol. 11, no. 1, pp. 75–84, 2021. [Online]. Available: https://febs.onlinelibrary.wiley.com/doi/abs/10.1002/2211-5463.13020

[10] R. Ghosh and A. Kumar, "A hybrid deep learning model by combining convolutional neural network and recurrent neural network to detect forest fire," *Multimedia Tools and Applications*, vol. 81, pp. 38 643–38 660, 2022. [Online]. Available: https://doi.org/10.1007/s11042-022-13068-8

[11] W. L. G. Y. X. M. L. W. Weiwei Sun, Chao Chen and K. Ren, "Coastline extraction using remote sensing: a review," *GIScience & Remote Sensing*, vol. 60, no. 1, p. 2243671, 2023. [Online]. Available: https://doi.org/10.1080/15481603.2023.2243671

[12] S. Khan and A. Khan, "Ffirenet: Deep learning based forest fire classification and detection in smart cities," *Symmetry*, vol. 14, no. 10, p. 2155, 2022.

[13] A. M. Mashraqi, Y. Asiri, A. D. Algarni, and H. Abu-Zinadah, "Drone imagery forest fire detection and classification using modified deep learning model," *Thermal Science*, vol. 26, no. Spec. issue 1, pp. 411–423, 2022, article.

[14] M. B. Mohammad, N. Bhuvaneswari, C. P. Koteswari, and V. B. Priya, "Hardware implementation of forest fire detection system using deep learning architectures," in *2022 International Conference on Edge Computing and Applications (ICECAA)*, 2022, pp. 1198–1205.

[15] R. K Mohammed, "A real-time forest fire and smoke detection system using deep learning," *International Journal of Nonlinear Analysis and Applications*, vol. 13, no. 1, pp. 2053–2063, 2022.

[16] S. Gayathri, P. A. Karthi, and S. Sunil, "Prediction and detection of forest fires based on deep learning approach," *Journal of Pharmaceutical Negative Results*, pp. 429–433, 2022.

[17] K. Alice, A. Thillaivanan, G. R. Koteswara Rao, R. S, K. Singh, and R. Rastogi, "Automated forest fire detection using atom search optimizer with deep transfer learning model," in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 2023, pp. 222–227.

[18] A. Smith *et al.*, "Application of convolutional neural networks for fire detection in surveillance footage," *Journal of Fire Safety Engineering*, vol. 15, no. 3, pp. 112–125, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8307064

[19] B. Jones and C. Brown, "Advancements in fire detection using convolutional neural networks," *International Journal of Computer Vision*, vol. 28, no. 4, pp. 456–468, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/9368283

[20] L. A. O. Gonçalves, R. Ghali, and M. A. Akhloufi, "Yolo-based models for smoke and wildfire detection in ground and aerial images," *Fire*, vol. 7, no. 4, 2024. [Online]. Available: https://www.mdpi.com/2571-6255/7/4/140

[21] Ultralytics, "YOLOv9 — docs.ultralytics.com," https://docs.ultralytics.com/models/yolov9/, 2024, [Accessed 16-04-2024].

[22] D. Gaias. (2024) DFireDataset. Accessed: Date. [Online]. Available: https://github.com/gaiasd/DFireDataset

[23] H. Yar, Z. A. Khan, I. Rida, W. Ullah, M. J. Kim, and S. W. Baik. (2024) An efficient deep learning architecture for effective fire detection in smart surveillance. [Online]. Available: Link: An efficient deep learning architecture for effective fire detection in smart surveillance - ScienceDirect

[24] Y. Kang, T. Sung, and J. Im, "Toward an adaptable deep-learning model for satellite-based wildfire monitoring with consideration of environmental conditions," *Remote Sensing of Environment*, vol. 298, p. 113814, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425723003656

[25] P. V. A. B. de Venâncio, A. C. Lisboa, and A. V. Barbosa, "An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices," *Neural Computing and Applications*, 2022. [Online]. Available: https://link.springer.com/article/10.1007/s00521-022-07467-z