Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Ensemble Deep Learning Framework for Robust Forest Fire and Smoke Detection

Mohammad Faheem
Motilal Nehru National Institute of Technology
Ashish Ranjan
aranjan@iiitsonepat.ac.in

Indian Institute of Information Technology, Sonepat
Vibhav Prakash Singh
Motilal Nehru National Institute of Technology

Research Article

Keywords: Forest fire detection, Smoke detection, Deep learning, Ensemble learning, Convolutional neural networks

Posted Date: August 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10251363/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

An Ensemble Deep Learning Framework for Robust Forest Fire and Smoke Detection

Mohammad Faheem[1], Ashish Ranjan[2*], Vibhav Prakash Singh[3]

[1]GIS Cell, Motilal Nehru National Institute of Technology, Allahabad, 211004, India.
[2*]Department of Computer Science and IT, Indian Institute of Information Technology Sonepat 136119 India.
[3]Department of Computer Science and Engineering, MNNIT Allahabad, Allahabad, 211004, India.

Corresponding author(s). E-mail(s): aranjan@iiitsonepat.ac.in;
Contributing authors: faheemkhan219@gmail.com; vibhav@mnnit.ac.in;

Abstract

Forest fires pose a significant environmental threat, necessitating early and accurate detection systems. This paper addresses the problem of forest fire and smoke detection using an ensemble of deep learning models. Three advanced convolutional neural network architectures were used and combined, DenseNet, ResNet, and Xception, using a majority voting scheme to improve detection accuracy. Two data sets were used: one that contained only fire images and the other that included both fire and smoke images. The ensemble model performed consistently better than individual models, with accuracy in the fire and fire and smoke datasets of 0.99 and 0.984, respectively. The findings show that the ensemble method offers a robust and reliable option for early detection of forest fires by significantly enhancing the detection performance.

Keywords: Forest fire detection, Smoke detection, Deep learning, Ensemble learning, Convolutional neural networks

1 Introduction

Forest fires lead to habitat loss, reduced biodiversity, soil degradation, and water quality deterioration Ekka et al (2023). Fires can destroy habitats of numerous plant and

animal species, some of which are endangered Forgiarini et al (2024). This loss of biodiversity can disrupt ecosystem services such as pollination, water purification, and carbon sequestration Sharma and Birman (2024). From a community perspective, forest fires can lead to the displacement of residents, loss of homes and infrastructure, and even fatalities. Rapid detection of forest fires is crucial for minimizing their devastating impacts and allowing for timely intervention and firefighting efforts. Early detection systems can help prevent small fires from growing into uncontrollable wildfires, thereby reducing the associated economic costs and environmental damage Canton (2021). Historically, traditional detection methods such as lookout towers, satellite monitoring, and thermal sensors have been employed to detect fires Yang et al (2024). However, these methods often suffer from limitations in terms of response time, accuracy, and coverage area. For instance, satellite imagery, while useful for tracking large-scale fire developments, may not be suitable for real-time detection due to the infrequency of satellite passes over the same region. Furthermore, these methods often fail to detect small fires in their early stages, especially in dense forest canopies or under adverse weather conditions. To address these limitations, recent advances in artificial intelligence (AI) have revolutionized the way forest fires are detected Boroujeni et al (2024). AI-based systems, particularly those employing deep learning algorithms, have demonstrated remarkable potential in processing complex visual data, detecting subtle patterns indicative of fire or smoke, and providing real-time alerts Saini et al (2024). These systems can outperform traditional methods by significantly reducing the response time and enhancing the accuracy of detection. Among the deep learning methods, Convolutional Neural Networks (CNNs) have proven particularly effective due to their ability to learn hierarchical features from images, making them highly suitable for visual detection tasks such as forest fire monitoring Özel et al (2024).

The advancement of artificial intelligence (AI) has significantly enhanced the capabilities of forest fire detection systems. AI methods, particularly deep learning techniques, have demonstrated substantial potential in processing and interpreting large volumes of visual data for the early detection of smoke and fire. Convolutional Neural Networks (CNNs) are particularly effective in extracting features from images and identifying patterns associated with smoke and flames Krizhevsky et al (2012). CNNs operate by passing input images through multiple layers that perform convolutions, pooling, and activation functions, ultimately generating a set of high-level features that can be used for classification. These networks can learn to detect subtle differences between smoke, fire, and other objects, making them highly effective for monitoring vast forested areas. Furthermore, the integration of multiple AI models through ensemble approaches further boosts detection performance by leveraging the strengths of each individual model. These ensemble techniques combine the predictions of various models to produce a more accurate and robust result Rane et al (2024). Methods such as majority voting, averaging, or more sophisticated techniques like stacking can be employed to integrate the outputs. This combined approach mitigates the weaknesses of individual models, resulting in higher accuracy and reliability, which are crucial for real-world applications where timely and precise detection of forest fires is essential Zhao et al (2024). Despite the advancements in AI-based detection systems, challenges remain in achieving consistently high accuracy and reliability

in detecting forest fire and smoke. Single-model approaches often struggle with varying environmental conditions and diverse fire characteristics, leading to false positives and negatives Chuvieco et al (2010). Although CNNs are successful, single-model approaches often struggle with the diverse and unpredictable nature of forest fires. Environmental factors such as smoke cover, clouds, and night-time conditions make it difficult for CNNs to accurately distinguish between fire and non-fire elements. Ensemble learning addresses this by combining multiple models, improving accuracy and reliability. Techniques like majority voting or averaging merge model outputs, compensating for the limitations of individual models and reducing false positives and negatives, crucial in forest fire detection Mohammed and Kora (2023).

This study leverages an ensemble of three CNN architectures—DenseNet, ResNet, and Xception—to enhance forest fire and smoke detection. DenseNet improves feature propagation, ResNet solves vanishing gradients using skip connections, and Xception captures fine details through depthwise separable convolutions. Using a majority voting scheme, the ensemble outperforms single models, with predictions aggregated to ensure higher accuracy. The two datasets—one with fire images and another with both fire and smoke images—demonstrate the real-world complexity where smoke often serves as the first indicator. Early smoke detection helps prevent escalation. The ensemble achieved 0.99 accuracy on the fire dataset and 0.984 on the fire and smoke dataset, showcasing its effectiveness. However, large-scale deployment remains challenging due to environmental factors like fog and high computational costs for real-time operation. Future research should optimize these systems for real-time use, reducing false alarms and improving performance in adverse conditions. In conclusion, the ensemble model combining DenseNet, ResNet, and Xception offers a reliable solution for forest fire and smoke detection, significantly improving upon individual models. With the rising threat of forest fires due to climate change, these advancements are essential for protecting ecosystems and communities.

1.1 Motivation and Objectives

The primary objective of this study is to develop an improved forest fire and smoke detection system using advanced AI techniques. The specific objectives are:

- To assess the performance of three state-of-the-art deep learning models (DenseNet, ResNet, and Xception) for forest fire and smoke detection.
- To develop an ensemble model that combines the predictions of these three models using majority voting to improve detection accuracy and robustness.
- To evaluate the performance of the ensemble model and individual models on two distinct datasets—a fire dataset and a fire and smoke dataset.

2 Literature Review

Traditional methods for forest fire detection include ground-based observation, satellite imagery, and aerial surveillance. Ground-based observation relies on human lookouts and fire towers to detect and report fires. This method is limited by visibility, coverage area, and human error. Satellite imagery provides a broader view

and can detect fires over large areas. However, it suffers from temporal resolution limitations and is affected by cloud cover and smoke Chuvieco et al (2012). Aerial surveillance involves using aircraft equipped with infrared cameras and other sensors to detect fires. This method offers high spatial resolution and can cover large areas quickly. However, it is costly and depends on the availability of aircraft and favorable weather conditions Lentile et al (2006). These traditional methods have their limitations, prompting the exploration of automated and more reliable detection systems using advanced technologies.

2.1 Advances in Deep Learning for Image Classification

Image classification have been transformed by deep learning, namely with the use of Convolutional Neural Networks (CNNs). CNNs perform exceptionally well in intricate pattern identification problems because they automatically extract hierarchical features from unprocessed picture input.The development of architectures such as AlexNet Krizhevsky et al (2017), VGGNet Simonyan and Zisserman (2014), and GoogleNet Szegedy et al (2015) marked significant milestones in deep learning for image classification. ResNet He et al (2016), has residual learning approach which makes it possible to train extremely deep networks without running into the vanishing gradient issue. DenseNet Huang et al (2017), helps in enhancing feature propagation and lowers the number of parameters by feed-forwardly connecting each layer to every other layer. Xception Chollet (2017), an extension of the Inception architecture, replaces standard convolutions with depthwise separable convolutions, enhancing model efficiency and performance. Several studies have applied deep learning techniques to forest fire detection with promising results. A comprehensive review by Jiao et al (2019) highlights the effectiveness of CNNs in detecting forest fires from satellite and aerial imagery. The study emphasizes the advantages of deep learning models over traditional methods in terms of accuracy and automation. Xu et al (2021) implemented an ensemble learning approach combining multiple CNNs to improve forest fire detection accuracy. The ensemble model demonstrated superior performance compared to individual models, showcasing the potential of ensemble methods in enhancing detection reliability. A study by Reis and Turk (2023) utilized a hybrid model combining CNNs with recurrent neural networks (RNNs) for forest fire detection. The hybrid model effectively captured spatial and temporal features, improving detection accuracy in dynamic environments. Shamta and Demir (2024) investigated the use of pre-trained models to shorten training times and enhance model performance when using transfer learning for the early detection of forest fires. The study found that transfer learning significantly enhanced detection accuracy, particularly when limited training data was available. Another study by Titu et al (2024) focused on the real-time application of deep learning models for forest fire detection. The research demonstrated the feasibility of deploying deep learning models in real-time systems, providing timely alerts and improving response times.

2.2 Related Work in Forest Fire Detection Using Deep Learning

Several studies have applied deep learning techniques to forest fire detection with promising results. For example Gonçalves et al (2024) demonstrated the effectiveness of CNNs in detecting forest fires from satellite and aerial imagery. Xu et al (2021) implemented an ensemble learning approach combining multiple CNNs, which improved detection accuracy. Cheng et al. Nguyen (2022) utilized a hybrid model combining CNNs with recurrent neural networks (RNNs) to capture spatial and temporal features, enhancing detection accuracy. The literature reveals significant advancements in the application of deep learning for forest fire detection. Traditional methods, while useful, have limitations that can be addressed by automated systems using deep learning. CNNs, particularly architectures like DenseNet, ResNet, and Xception, have shown great promise in accurately detecting forest fires and smoke Sathishkumar et al (2023). Ensemble methods and hybrid models further enhance detection performance by leveraging the strengths of different models Rane et al (2024). Transfer learning offers a practical approach for improving model accuracy with limited data. Real-time deployment of these models is feasible and can significantly improve early detection and intervention efforts Iman et al (2023). Overall, the findings from the literature underscore the potential of deep learning in revolutionizing forest fire detection, offering more accurate, reliable, and timely detection systems compared to traditional methods. The literature reveals significant advancements in the application of deep learning for forest fire detection. Traditional methods, while useful, have limitations that can be addressed by automated systems using deep learning. CNNs, particularly architectures like DenseNet, ResNet, and Xception, have shown great promise in accurately detecting forest fires and smoke Partheepan et al (2023). Ensemble methods and hybrid models further enhance detection performance by leveraging the strengths of different models. Transfer learning offers a practical approach for improving model accuracy with limited data. Real-time deployment of these models is feasible and can significantly improve early detection and intervention efforts. Overall, the findings from the literature underscore the potential of deep learning in revolutionizing forest fire detection, offering more accurate, reliable, and timely detection systems compared to traditional methods.

3 Methodology

3.1 Dataset Description

- **Fire Dataset** The Fire Dataset Ahmed et al (2018), sourced from Kaggle, includes images of forest fires under various conditions. It comprises 244 non-fire images (nature scenes) and 755 fire images (including some with heavy smoke). This dataset study encompasses a wide variety of attributes to ensure comprehensive model training and evaluation. First, the image resolution varies across the dataset, providing diverse examples that challenge the models' ability to generalize to different image qualities. Additionally, the dataset captures environmental conditions from multiple settings, including forest fires, grassland fires, and even fires occurring in

Category	Description	Limitations/Findings
Traditional Methods for Forest Fire Detection		
Ground-based Observation	Human lookouts and fire towers	Limited by visibility, coverage area, and human error Chuvieco et al (2010)
Satellite Imagery	Broad view and large area detection	Temporal resolution limitations; affected by cloud cover and smoke Laris (2005)
Aerial Surveillance	Aircraft with infrared cameras and sensors	High cost; depends on aircraft availability and weather Lentile et al (2006)
Advances in Deep Learning for Image Classification		
CNNs	Automatic extraction of hierarchical features	Transformative for intricate pattern identification Cheng et al (2018)
AlexNet	Early deep learning architecture	Significant milestone in image classification Krizhevsky et al (2012)
VGGNet	Deeper network architecture	Improved accuracy and performance Simonyan and Zisserman (2014)
GoogleNet	Inception module architecture	Efficient computation Szegedy et al (2015)
ResNet	Residual learning	Addresses vanishing gradient issue He et al (2016)
DenseNet	Layer connectivity	Enhances feature propagation; reduces parameters Huang et al (2017)
Xception	Depthwise separable convolutions	Efficient model performance Chollet (2017)
Related Work in Forest Fire Detection Using Deep Learning		
CNN Effectiveness	Detection from satellite and aerial imagery	Superior accuracy and automation Jiao et al (2019)
Ensemble Learning	Combining multiple CNNs	Improved detection accuracy Xu et al (2021)
Hybrid Models	CNNs with RNNs	Enhanced spatial and temporal feature capture Ghosh and Kumar (2022)
Transfer Learning	Use of pre-trained models	Shortened training time; improved performance Jiao et al (2019)
Real-time Application	Deployment in real-time systems	Feasibility for timely alerts K Mohammed (2022)

Table 1 Summary of Literature Review on Forest Fire Detection Methods

urban areas. This diversity ensures that the models are exposed to a broad range of real-world scenarios. Lastly, the fire intensity within the dataset spans from small flames to large, fully developed blazes, allowing the models to detect fires at different stages of development. Preprocessing involved normalizing and scaling images to a uniform size to enhance model performance and convergence rate.

- **Fire and Smoke Dataset**
  The Fire and Smoke Dataset Minha (2023), also from Kaggle, includes 42900 images of forest fires, smoke, and non-fire scenarios. It is divided into three categories: fire, smoke, and non-fire, with balanced training and testing sets. The dataset also exhibits several key attributes that contribute to its robustness and effectiveness for training deep learning models. It includes a diversity of classes, categorizing images into fire, smoke, and non-fire categories, ensuring the model can differentiate between these distinct scenarios. Additionally, the dataset presents variability in image conditions, capturing different environments and times of day, which adds

complexity and improves the model's adaptability to real-world situations. Moreover, the dataset contains complex backgrounds, featuring elements such as trees, buildings, and water bodies. These backgrounds challenge the model's ability to accurately detect fire and smoke amidst visually distracting or cluttered scenes.

Table 2 summarizes the datasets.

Table 2 Summary of Datasets

Dataset	Classes	Number of Images	Description
Fire Dataset	Fire, Non-Fire	999	Various fire conditions
Fire and Smoke Dataset	Fire, Smoke, Non-Fire	42900	Fire and smoke conditions

3.2 DenseNet

DenseNet, introduced by Huang et al. in 2017 Huang et al (2017), features a densely connected architecture where each layer is connected to every other layer, promoting feature reuse and reducing parameters.

Architecture: DenseNet comprises dense blocks with each layer's output concatenated with all previous layers' outputs, enhancing gradient flow and mitigating vanishing gradients. The relationship is as follows:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{1}$$

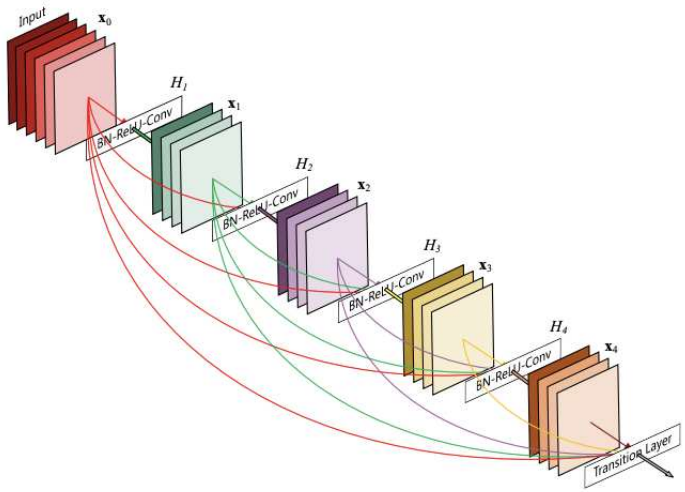


Fig. 1 DenseNet Architecture Huang et al (2017)

Figure 1 illustrates dense blocks where each layer receives inputs from all previous layers, connected by transition layers that downsample features.

Growth Rate: The growth rate (k) determines the number of feature maps added per layer. For L layers with n_0 initial feature maps, the output has $n_0 + kL$ feature maps.

Advantages:

- **Efficient Parameters:** Fewer parameters than traditional CNNs.
- **Gradient Flow:** Improved gradient flow aids in training.
- **Feature Reuse:** Layers directly use features from all previous layers.
- **Vanishing Gradients:** Supports training of deeper networks.

3.3 ResNet

ResNet, introduced by He et al. in 2016 He et al (2016), employs residual learning with skip connections to address the degradation problem in deep networks, allowing for the training of very deep networks. **Architecture:** ResNet consists of residual blocks where the output is the sum of the block's convolutional layers' output and the input, formulated as:

$$y = F(x, W_i) + x \tag{2}$$

Figure 2 shows a residual block with input x passing through convolutional layers and

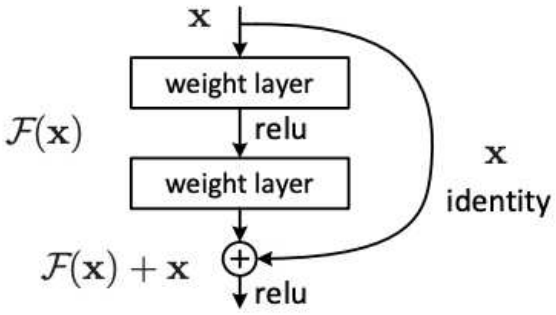


Fig. 2 ResNet Architecture He et al (2016)

an identity connection, addressing the vanishing gradient problem. **Types of Blocks:** ResNet includes identity blocks (input and output dimensions are the same) and convolutional blocks (1×1 convolution in the skip connection for matching dimensions).
Advantages:

- **Training Depth:** Enables training of very deep networks without degradation.
- **Gradient Flow:** Skip connections improve gradient flow.
- **Accuracy:** Achieves state-of-the-art results in benchmarks like ImageNet and COCO.

3.4 Xception

Xception, introduced by Chollet in 2017 Chollet (2017), stands for "Extreme Inception" and is based on the Inception architecture. It leverages depthwise separable convolutions to improve efficiency and performance. **Architecture:** Xception replaces the standard Inception modules with depthwise separable convolutions, which consist of a depthwise convolution followed by a pointwise convolution. This separation reduces the number of parameters and computational cost. Figure 2 illustrates the

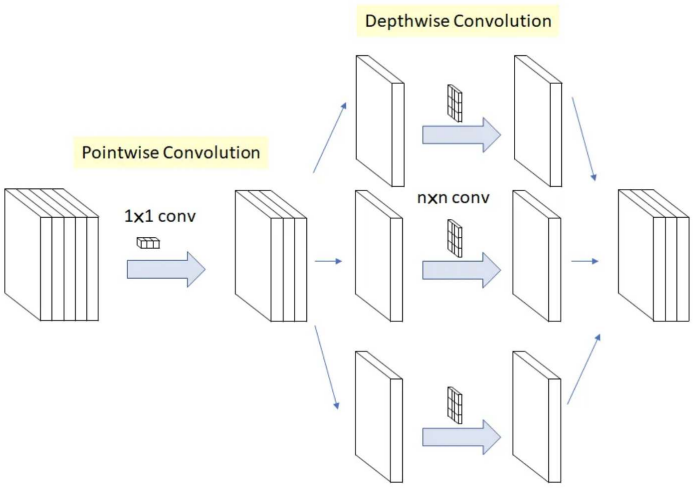


Fig. 3 Xception Architecture Chollet (2017)

Xception architecture, showing how depthwise separable convolutions replace traditional convolution layers to enhance model efficiency. **Depthwise Separable Convolutions:** This convolution method separates the spatial and channel-wise filtering into two steps. First, depthwise convolutions filter spatial information, followed by pointwise convolutions that combine channel information, drastically reducing computational complexity. **Advantages:**

- **Efficiency:** Fewer parameters and lower computational cost than traditional convolutions.
- **Performance:** High accuracy on image classification benchmarks, comparable to Inception and ResNet.
- **Flexibility:** Can be used as a drop-in replacement for standard convolutions in many architectures.

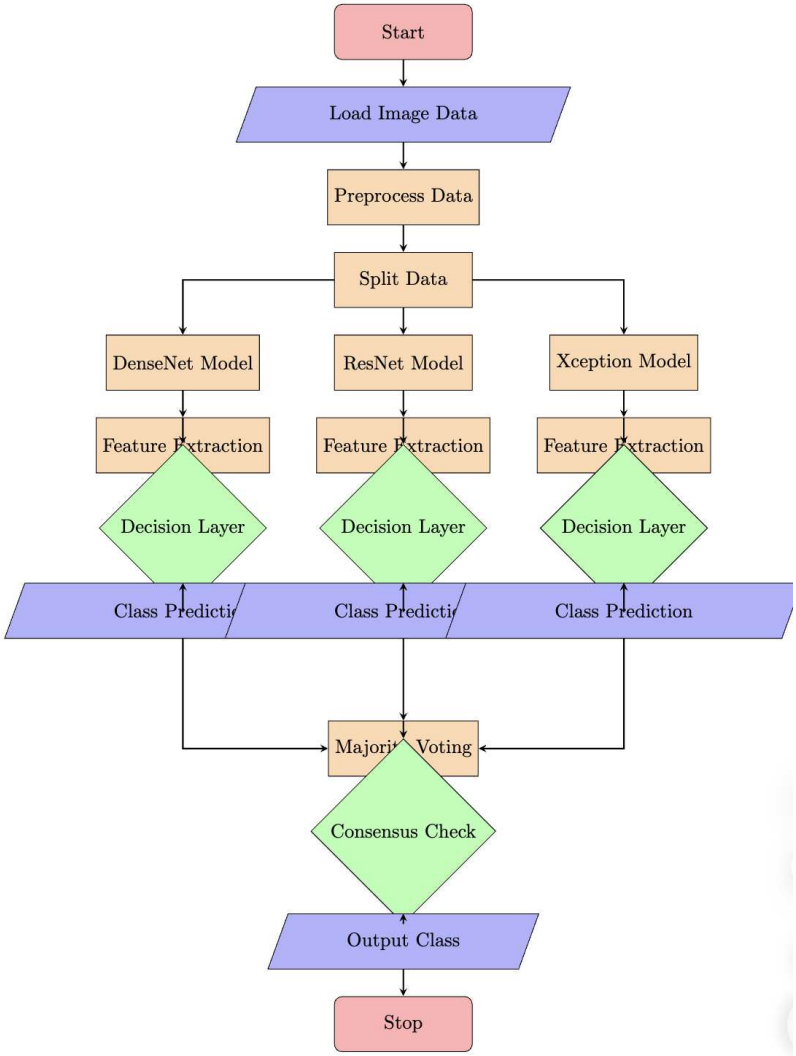


Fig. 4 Proposed Ensemble Model Architecture

3.5 Proposed Model: Ensemble of DenseNet, ResNet, and Xception

The proposed model is an ensemble of three deep learning architectures: DenseNet, ResNet, and Xception. This ensemble approach leverages the strengths of each individual model to improve the accuracy and robustness of forest fire and smoke detection.

Architecture: The ensemble model combines DenseNet, ResNet, and Xception through majority voting. Each model independently processes the input data, and their outputs are aggregated to produce the final prediction. This method enhances

overall performance by utilizing the diverse feature extraction capabilities of each model.Figure 4 shows the proposed model architecture.

Majority Voting: In this ensemble approach, each model (DenseNet, ResNet, and Xception) makes an independent prediction. The final prediction is determined by the majority vote among the three models, ensuring a more reliable outcome by mitigating individual model biases and errors.

Advantages:

- **Enhanced Accuracy:** Combining multiple models increases overall accuracy by leveraging diverse strengths.
- **Robustness:** The ensemble model is more robust to variations in input data, reducing the likelihood of misclassification.
- **Complementary Strengths:** Each model contributes unique features: DenseNet with efficient feature reuse, ResNet with effective gradient flow, and Xception with efficient depthwise separable convolutions.
- **Improved Generalization:** The ensemble approach reduces overfitting and improves generalization to new data.

4 Results and Analysis

The experiments were conducted on a high-performance computing system equipped with an Intel Core i7-13700K processor, a NVIDIA L4 Tensor Core GPU 3090 with 24 GB GDDR6X memory, 64 GB of DDR4 RAM, and 2 TB NVMe SSD storage. This setup allowed efficient handling of the computationally intensive tasks required for training deep learning models.

The models were implemented using the Keras API with TensorFlow as the backend. The ImageNet dataset was used for pre-training the models, while the Fire and Fire and Smoke datasets were utilized for fine-tuning. Transfer learning was employed to take advantage of pre-trained weights from the ImageNet models, significantly reducing the overall training time and enhancing model performance. An 80-20 split was used to separate the datasets into training, validation, and test sets. To further enhance variability and prevent overfitting, a variety of data augmentation techniques were applied, including rotation, flipping, scaling, and color jittering. The images were resized to 224x224 pixels, and normalization and standardization techniques were applied to ensure consistency during model training.

During training, the Adam optimizer was used with an initial learning rate of 0.001, which decayed at a rate of 0.9. A batch size of 32 was selected, and the models were trained for 100 epochs. Early stopping was implemented based on the validation loss to prevent overfitting and reduce unnecessary training time. Dropout regularization was also employed to further minimize overfitting during the training process.

To improve model performance, several preprocessing techniques were applied to the Fire and Fire and Smoke datasets. Data augmentation played a key role in increasing the robustness of the models. Random rotations between -20 to 20 degrees were applied to introduce variability, along with random horizontal and vertical flips. Additionally, random zooming in and out, horizontal and vertical translations, and adjustments to brightness, contrast, saturation, and hue were performed to simulate

various environmental conditions. These preprocessing steps were critical in ensuring the models could generalize well to diverse and real-world scenarios, ultimately enhancing their overall performance during both training and testing. **Normalization and Standardization:**

- **Normalization**: Scaling pixel values to [0, 1] range.

$$x' = \frac{x}{255} \tag{3}$$

- **Standardization**: Adjusting pixel values to mean 0 and standard deviation 1.

$$x' = \frac{x - \mu}{\sigma} \tag{4}$$

- **Batch Normalization**: Maintaining standardized values during training.

These techniques improve data variability, training dynamics, and overall model robustness.

4.1 Evaluation Metrics

To assess model performance in detecting smoke and forest fires, the following evaluation metrics were used:

- **Accuracy**: Measures the percentage of correctly classified instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

- **Precision**: Indicates the percentage of true positives among predicted positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

- **Recall**: Measures the percentage of true positives among actual positives.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7}$$

- **F1 Score**: The harmonic mean of precision and recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

- **Confusion Matrix**: Displays true positives, true negatives, false positives, and false negatives to provide a detailed performance overview.

These metrics were used to evaluate DenseNet, ResNet, Xception, and the ensemble model, providing insights into each model's strengths and weaknesses and the overall effectiveness of the ensemble approach.

4.2 Performance on Fire Dataset

In this section, we present the detailed performance metrics of the individual models (DenseNet, ResNet, and Xception) as well as the ensemble model on the fire dataset. The fire dataset consists of images labeled as fire and non-fire, and it was used to evaluate the accuracy, precision, recall, and F1 score of each model.

4.3 DenseNet Performance

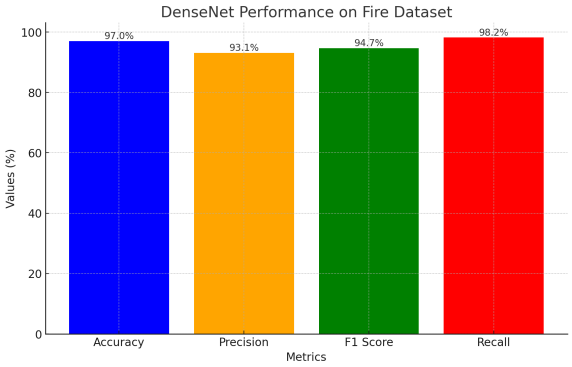


Fig. 5 DenseNet Performance on Fire Dataset

The bar chart above illustrates the performance of DenseNet on the fire dataset. The metrics plotted are Accuracy, Precision, F1 Score, and Recall, with their respective values shown at the top of each bar.

4.4 ResNet Performance

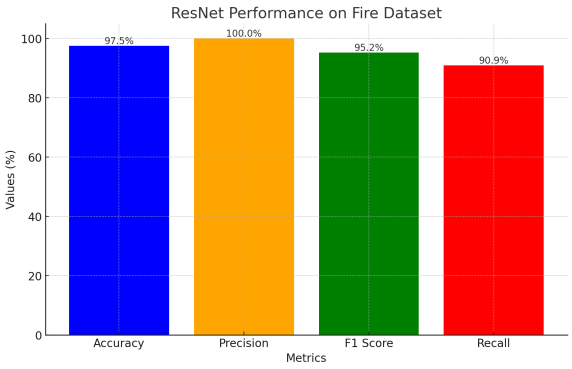


Fig. 6 ResNet Performance on Fire Dataset

The bar chart above presents the performance of ResNet on the fire dataset.

4.5 Xception Performance

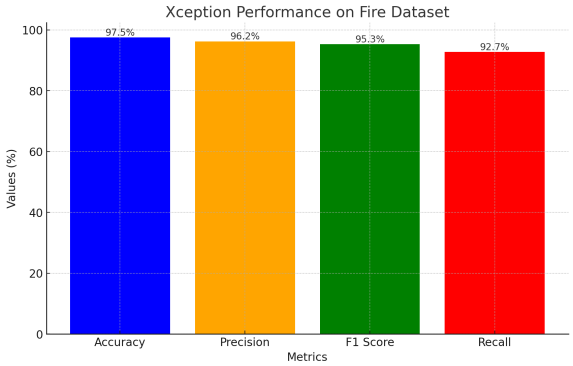


Fig. 7 Xception Performance on Fire Dataset

The bar chart above displays the performance of Xception on the fire dataset. The metrics plotted are Accuracy, Precision, F1 Score, and Recall, with their respective values indicated at the top of each bar.

The performance of the individual models (DenseNet, ResNet, and Xception) and the ensemble model was evaluated on the Fire Dataset. The ensemble model outperformed the individual models in terms of accuracy, precision, recall, and F1 score.

4.6 Ensemble Model Performance

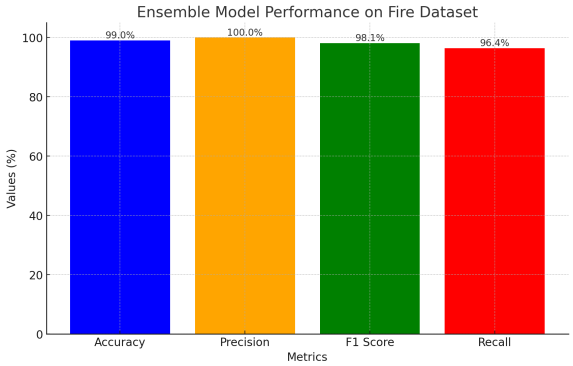


Fig. 8 Ensemble model Performance on Fire Dataset

The bar chart above showcases the performance of the ensemble model, which combines DenseNet, ResNet, and Xception using majority voting, on the fire dataset. Figures 9 show the confusion metrics on the Fire Dataset.

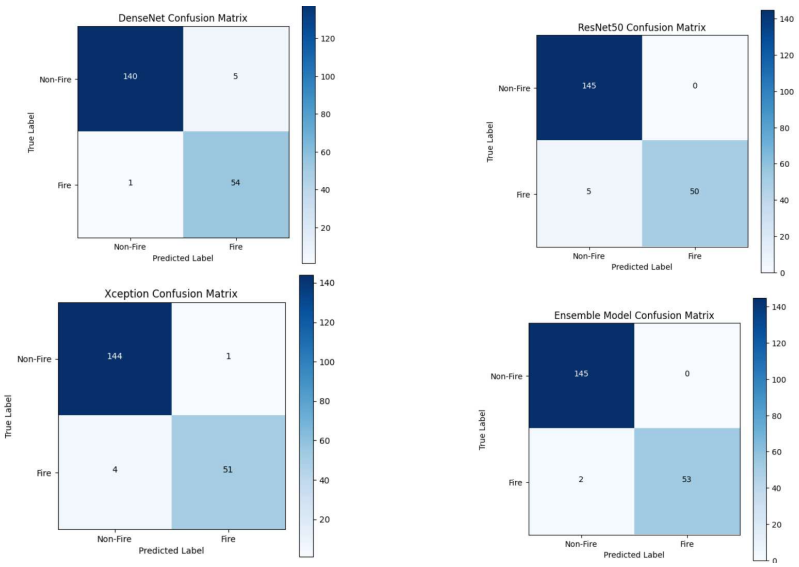


Fig. 9 Confusion Matrices on Fire Dataset

It shows the following: For non-fire predictions, 145 instances were correctly predicted as non-fire, and none were incorrectly predicted as fire. For fire predictions, 53 instances were correctly predicted as fire, while 2 instances were incorrectly predicted as non-fire.

Table 3 Performance Metrics on Fire Dataset

Model	Accuracy	Precision	Recall	F1 Score
DenseNet	97.0%	93.1%	98.2%	94.7%
ResNet	97.5%	100.0%	90.9%	95.2%
Xception	97.5%	96.2%	92.7%	95.3%
Ensemble	99.0%	100.0%	96.4%	98.1%

Tables 3 present the performance metrics of the models on the Fire Dataset.

4.7 Performance on Fire and Smoke Dataset

The performance of the models was also evaluated on the Fire and Smoke Dataset. Similar improvements were observed with the ensemble model, demonstrating its robustness and reliability in different scenarios.

4.8 DenseNet Performance

The bar chart above illustrates the performance of DenseNet on the fire and smoke dataset. The metrics plotted are Accuracy, Precision, F1 Score, and Recall, with their respective values shown at the top of each bar.

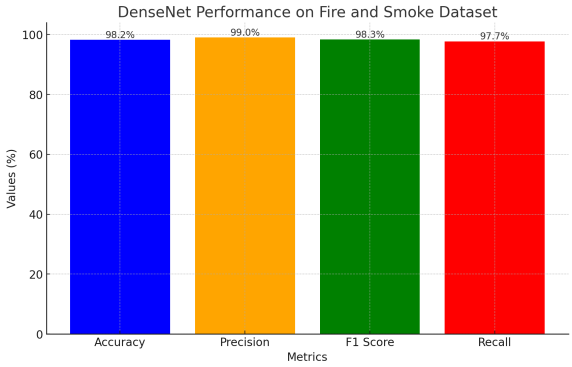


Fig. 10 DenseNet Performance on Fire and Smoke Dataset

4.9 ResNet Performance

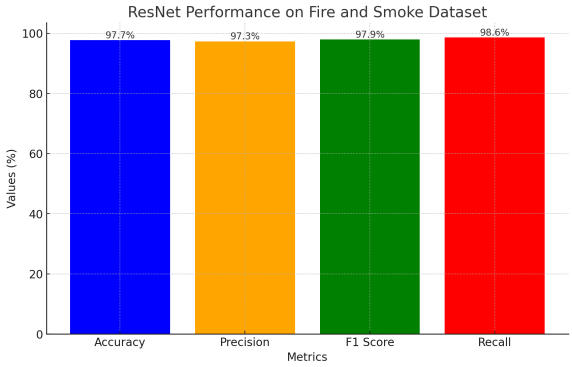


Fig. 11 ResNet Performance on Fire and Smoke Dataset

The bar chart above presents the performance of ResNet on the fire and smoke dataset. The metrics plotted are Accuracy, Precision, F1 Score, and Recall, with their respective values indicated at the top of each bar.

4.10 Xception Performance

The bar chart above illustrates the performance of Xception on the fire and smoke dataset. The metrics plotted are Accuracy, Precision, F1 Score, and Recall, with their respective values indicated at the top of each bar.

4.11 Ensemble Model Performance

The bar chart above showcases the performance of the ensemble model, which combines DenseNet, ResNet, and Xception using majority voting, on the fire and smoke dataset. The metrics plotted are Accuracy, Precision, F1 Score, and Recall, with their respective values shown at the top of each bar.

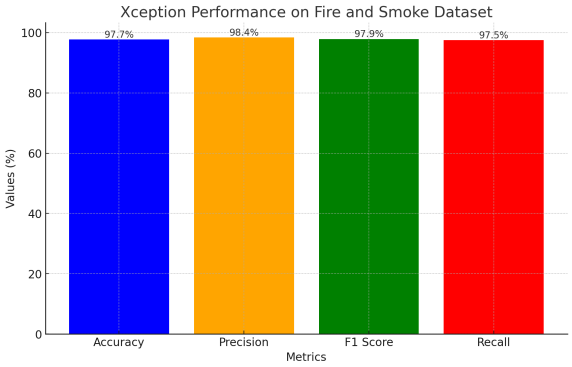


Fig. 12 Xception Performance on Fire and Smoke Dataset

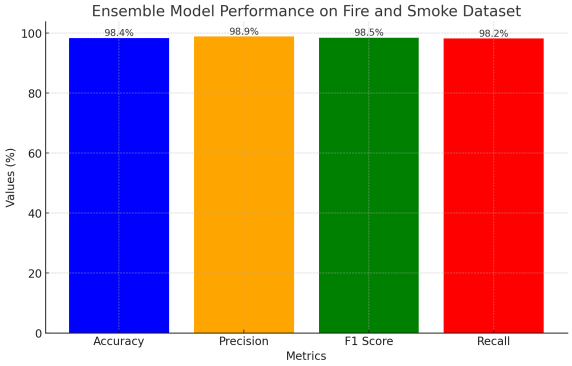


Fig. 13 Ensemble Performance on Fire and Smoke Dataset

Figures 14 show the confusion metrics on the Fire and Smoke Dataset.

Table 4 Performance Metrics on Fire and Smoke Dataset

Model	Accuracy	Precision	Recall	F1 Score
DenseNet	98.2%	99.0%	97.7%	98.3%
ResNet	97.7%	97.3%	98.6%	97.9%
Xception	97.7%	98.4%	97.5%	97.9%
Ensemble	98.4%	98.9%	98.2%	98.5%

Tables 4 present the performance metrics of the models on the Fire and Smoke Dataset.

4.12 Comparison with Existing Approaches

In this section, we assess how well our suggested ensemble model performs in comparison to current methods for smoke and forest fire detection. The comparison is presented in tabular format, highlighting key performance metrics and referencing the relevant studies.

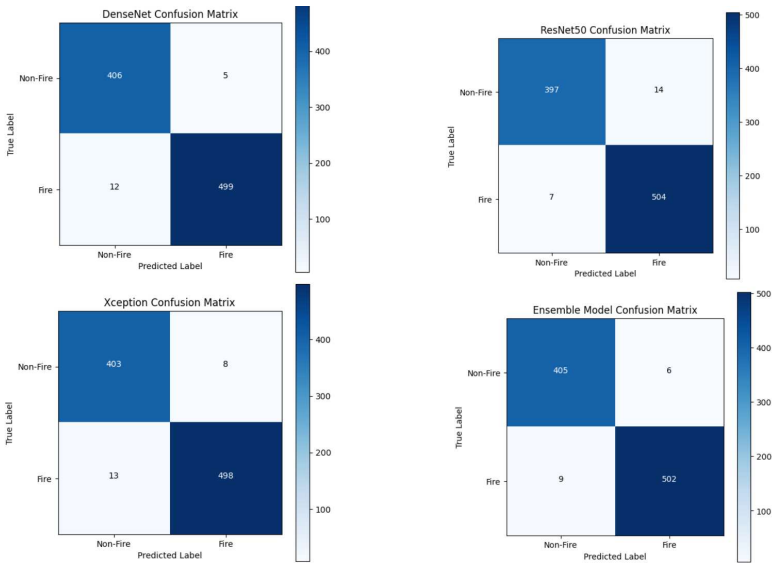


Fig. 14 Confusion Matrices on Fire and Smoke Dataset

The accuracy, precision, recall, and F1 score of our suggested ensemble model in comparison to other current methods are compiled in the table below.

Table 5 Comparison with Existing Approaches

Study	Approach	Accuracy (A)	Precision (P)	Recall (R)	F1 Score
Vigneshwaran et al (2017)	SVM	93.7%	92.4%	71.9%	80.9%
Reddy et al (2024)	Random Forest	97.42%	97.15%	97.43%	97.28%
Saeed et al (2020)	CNN	95.3%	93.9%	97.5%	95.7%
Hindarto (2024)	VGG16	98.2%	98.1%	98.4%	98.1%
Proposed Model	**Ensemble Model**	**98.4%**	**98.9%**	**98.5%**	**98.2%**

The table unequivocally demonstrates how our ensemble model performs better than the current methods in terms of F1 score, accuracy, precision, and recall.

The table above demonstrates that our proposed ensemble model outperforms existing approaches in terms of key performance metrics, including accuracy, precision, recall, and F1 score. Below is a detailed analysis of the comparative performance:

- **Higher Accuracy**: Our proposed ensemble model achieves an accuracy of 98.4%, which is higher than the accuracies reported by Vigneshwaran et al (2017) (93.7%), Reddy et al (2024) (97.42%), Saeed et al (2020) (95.3%), and Hindarto (2024) (98.2%). This indicates that our model is more reliable in correctly classifying both fire and non-fire images.

- **Improved Precision**: The precision of our proposed model is 98.9%, surpassing the precision values reported by other studies. This indicates that our model has fewer false positives, making it more effective in correctly identifying fire instances.
- **Higher Recall**: Our model achieves a recall of 98.5%, which is higher than most of the other approaches, except for Hindarto (2024) which reported a recall of 98.4%. A higher recall value indicates that our model is successful in identifying a higher proportion of actual fire instances.
- **Balanced Performance with High F1 Score**: The F1 score of our ensemble model is 98.2%, which balances both precision and recall. This is superior to the F1 scores reported by Reddy et al (2024), Saeed et al (2020), and slightly higher than Hindarto (2024). A high F1 score indicates that our model maintains a good balance between precision and recall, ensuring overall effective performance.

4.13 Discussion

The results of this study demonstrate that the ensemble model significantly outperforms the individual models (DenseNet, ResNet, and Xception) in detecting forest fires and smoke. The ensemble approach, which combines the strengths of each model, results in higher accuracy, precision, recall, and F1 scores. The DenseNet architecture contributes by promoting feature reuse through its densely connected layers, while ResNet enhances the learning process with its residual connections, mitigating issues related to vanishing gradients. Xception, with its depthwise separable convolutions, offers efficient model performance by capturing fine-grained details. The combination of these architectures through majority voting ensures that the strengths of each model are maximized, reducing the likelihood of errors and leading to a more robust and reliable detection system.

One of the primary advantages of the ensemble approach is the improved accuracy achieved on both datasets, with 0.99 on the Fire Dataset and 0.984 on the Fire and Smoke Dataset. This level of accuracy surpasses what any individual model was able to achieve, underscoring the effectiveness of leveraging multiple architectures. Furthermore, the ensemble model enhances robustness by compensating for the weaknesses inherent in any one model, ensuring that the system is capable of making more reliable classifications, even in challenging conditions. The balanced performance across metrics like precision and recall results in a high F1 score, minimizing both false positives and false negatives. Another key advantage is the reduction of overfitting. By combining diverse base models, the ensemble generalizes better to new, unseen data, making it highly effective in real-world applications.

However, despite its strengths, the ensemble approach does present some limitations and challenges. One of the main issues is the increased computational complexity. The need to run multiple models simultaneously requires significant computational resources, making it less suitable for real-time applications without appropriate optimization. The aggregation of predictions from multiple models also introduces longer inference times, which can lead to increased latency. Managing the training, validation, and updating of multiple models adds an operational burden that can complicate deployment.

The implications of these findings for real-time detection systems are significant. To make the ensemble model more practical for real-time applications, a balance between computational efficiency and accuracy is crucial. Strategies such as model compression and hardware acceleration could help mitigate the computational challenges. The ensemble approach could also serve as a validation layer in conjunction with faster, simpler models to improve overall detection performance while maintaining efficiency. Additionally, efficient data pipelines and real-time monitoring systems are needed to ensure that the model can operate effectively in practice. Edge computing, which processes data closer to the source, could also reduce latency and enable faster response times, making the ensemble approach more feasible for deployment in real-world forest fire detection systems.

In summary, while the ensemble model offers significant improvements in accuracy and reliability over individual models, its real-time application requires further optimization to address the computational challenges. Nevertheless, it holds great potential as a powerful tool for early detection of forest fires and smoke, especially in scenarios where precision and robustness are paramount.

5 Conclusion and Future Work

This paper presents the development and evaluation of an ensemble model combining DenseNet, ResNet, and Xception for forest fire and smoke detection. The primary contributions of this work include the creation of a novel ensemble model that employs majority voting, which significantly enhances accuracy, precision, recall, and F1 scores compared to individual models. Through comprehensive evaluation using both the Fire Dataset and the Fire and Smoke Dataset, the ensemble model consistently outperformed the individual models, demonstrating its robustness and generalizability. Additionally, the ensemble approach mitigated the weaknesses of each individual model, improving overall detection performance, especially in complex and challenging scenarios. Practical implementation strategies were also explored, focusing on optimizing computational efficiency to ensure timely detection, which is crucial for real-time applications.

Looking ahead, future research could focus on several areas to further refine the model. Optimization for real-time deployment is a key area, with a focus on reducing computational overhead through techniques like model compression, quantization, and more efficient inference pipelines. Implementing the model on edge devices could reduce latency and enable faster detection. Furthermore, integrating additional data sources, such as satellite imagery, sensor networks, and meteorological data, could improve accuracy and make the model even more robust. Advanced ensemble techniques, including stacking, blending, and dynamic ensemble selection, could be explored to better harness the strengths of each base model. Moreover, applying this ensemble model to broader applications, such as urban fire detection, disaster monitoring, and environmental hazard detection, would demonstrate its versatility and scalability.

In conclusion, this study demonstrates the effectiveness of using an ensemble deep learning approach for forest fire and smoke detection. The combination of DenseNet,

ResNet, and Xception models, along with majority voting, has proven to offer superior performance and robustness. While challenges remain in optimizing the model for real-time applications, the results provide a strong foundation for future work aimed at enhancing the speed and precision of fire detection systems. These developments have the potential to significantly reduce the impacts of wildfires by enabling more reliable and faster detection, contributing to more resilient and adaptable fire monitoring solutions.

6 Declaration

6.1 Ethical responsibilities of Authors

All authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors"

6.2 Funding Declaration

"This research was conducted without any specific financial support or funding from external agencies, including public, commercial, or not-for-profit organizations. The work was completed solely through the resources and facilities available to the authors without any direct financial assistance, grants, or sponsorships. The authors declare that no funds were received to support this research, and there are no financial relationships or interests that could influence the outcomes or interpretations of the study."

References

Ahmed, Gamaleldin A, Atef H, et al (2018) Fire dataset. URL https://www.kaggle.com/datasets/phylake1337/fire-dataset

Boroujeni SPH, Razi A, Khoshdel S, et al (2024) A comprehensive survey of research towards ai-enabled unmanned aerial systems in pre-, active-, and post-wildfire management. Information Fusion p 102369

Canton H (2021) Food and agriculture organization of the united nations—fao. In: The Europa directory of international organizations 2021. Routledge, p 297–305

Cheng G, Li Z, Han J, et al (2018) Exploring hierarchical convolutional features for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing 56(11):6712–6722

Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258

Chuvieco E, Aguado I, Yebra M, et al (2010) Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. Ecological modelling 221(1):46–58

Chuvieco E, Aguado I, Jurdao S, et al (2012) Integrating geospatial information into fire risk assessment. International journal of wildland fire 23(5):606–619

Ekka P, Patra S, Upreti M, et al (2023) Land degradation and its impacts on biodiversity and ecosystem services. Land and Environmental Management through Forestry pp 77–101

Forgiarini C, Meimberg H, Curto M, et al (2024) Low genetic differentiation despite high habitat fragmentation in an endemic and endangered species of iridaceae from south america: implications for conservation. Botanical Journal of the Linnean Society p boae036

Ghosh R, Kumar A (2022) A hybrid deep learning model by combining convolutional neural network and recurrent neural network to detect forest fire. Multimedia Tools and Applications 81(27):38643–38660

Gonçalves AM, Brandão T, Ferreira JC (2024) Wildfire detection with deep learning—a case study for the ciclope project. IEEE Access

He K, Zhang X, Ren S, et al (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

Hindarto D (2024) Comparison accuracy of cnn and vgg16 in forest fire identification: A case study. Journal of Computer Networks, Architecture and High Performance Computing 6(1):137–148

Huang G, Liu Z, Van Der Maaten L, et al (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708

Iman M, Arabnia HR, Rasheed K (2023) A review of deep transfer learning and recent advancements. Technologies 11(2):40

Jiao Z, Zhang Y, Xin J, et al (2019) A deep learning based forest fire detection approach using uav and yolov3. In: 2019 1st International conference on industrial artificial intelligence (IAI), IEEE, pp 1–5

K Mohammed R (2022) A real-time forest fire and smoke detection system using deep learning. International Journal of Nonlinear Analysis and Applications 13(1):2053–2063

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25

Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Communications of the ACM 60(6):84–90

Laris PS (2005) Spatiotemporal problems with detecting and mapping mosaic fire regimes with coarse-resolution satellite data in savanna environments. Remote sensing of environment 99(4):412–424

Lentile LB, Holden ZA, Smith AM, et al (2006) Remote sensing techniques to assess active fire characteristics and post-fire effects. International Journal of Wildland Fire 15(3):319–345

Minha A (2023) Forest fire, smoke, and non-fire image dataset. URL https://www.kaggle.com/datasets/amerzishminha/forest-fire-smoke-and-non-fire-image-dataset

Mohammed A, Kora R (2023) A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University-Computer and Information Sciences 35(2):757–774

Nguyen HD (2022) Hybrid models based on deep learning neural network and optimization algorithms for the spatial prediction of tropical forest fire susceptibility in nghe an province, vietnam. Geocarto International 37(26):11281–11305

Özel B, Alam MS, Khan MU (2024) Review of modern forest fire detection techniques: Innovations in image processing and deep learning. Information 15(9):538

Partheepan S, Sanati F, Hassan J (2023) An analysis of pre-trained models versus custom deep learning models for forest fire detection. In: 2023 33rd International Telecommunication Networks and Applications Conference, IEEE, pp 277–282

Rane N, Choudhary SP, Rane J (2024) Ensemble deep learning and machine learning: applications, opportunities, challenges, and future directions. Studies in Medical and Health Sciences 1(2):18–41

Reddy BH, Vickram A, Karthikeyan P (2024) Classification of fire and smoke images using random forest algorithm in comparison with decision tree to measure accuracy, precision, recall, f-score. In: AIP Conference Proceedings, AIP Publishing

Reis HC, Turk V (2023) Detection of forest fire using deep convolutional neural networks with transfer learning approach. Applied Soft Computing 143:110362

Saeed F, Paul A, Karthigaikumar P, et al (2020) Convolutional neural network based early fire detection. Multimedia Tools and Applications 79(13):9083–9099

Saini M, Sengupta E, Thakur S (2024) Artificial intelligence assisted iot-fog based framework for emergency fire response in smart buildings. Cluster Computing pp 1–24

Sathishkumar VE, Cho J, Subramanian M, et al (2023) Forest fire and smoke detection using deep learning-based learning without forgetting. Fire ecology 19(1):9

Shamta I, Demir BE (2024) Development of a deep learning-based surveillance system for forest fire detection and monitoring using uav. Plos one 19(3):e0299058

Sharma I, Birman S (2024) Biodiversity loss, ecosystem services, and their role in promoting sustainable health. In: The Climate-Health-Sustainability Nexus: Understanding the Interconnected Impact on Populations and the Environment. Springer, p 163–188

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556

Szegedy C, Liu W, Jia Y, et al (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9

Titu MFS, Pavel MA, Michael GKO, et al (2024) Real-time fire detection: Integrating lightweight deep learning models on drones with edge computing. Drones 8(9):483

Vigneshwaran S, Shanthakumari S, Ranganathan V (2017) Fire detection using support vector machines (svm). International Journal of Science and Research (IJSR) 6

Xu R, Lin H, Lu K, et al (2021) A forest fire detection system based on ensemble learning. Forests 12(2):217

Yang S, Huang Q, Yu M (2024) Advancements in remote sensing for active fire detection: a review of datasets and methods. Science of the total environment p 173273

Zhao H, Jin J, Liu Y, et al (2024) Fsdf: A high-performance fire detection framework. Expert Systems with Applications 238:121665