Hybrid Deep Learning Framework for Real-Time Forest Monitoring and Early Fire Detection Using Multi-Source Imagery

I. Sree Anushka*, B. Rishitha**, Ms. Neha Tanveer***
*(Department of Artificial Intelligence and Data Science,Stanley College of Engineering And Technology for Women, Hyderabad
Email: 160622747023@stanley.edu.in)

*(Department of Artificial Intelligence and Data Science,Stanley College of Engineering And Technology for Women, Hyderabad
Email: 160622747016@stanley.edu.in)

*(Department of Artificial Intelligence and Data Science,Stanley College of Engineering And Technology for Women, Hyderabad
Email: nehatanveer@stanley.edu.in)

Abstract—This study presents a hybrid deep learning framework for real-time forest monitoring and early fire detection using multi-source imagery. The proposed system combines Convolutional Neural Networks (CNNs) for spatial feature extraction with Long Short-Term Memory (LSTM) networks to capture temporal variations in forest environments. A preprocessing and augmentation pipeline improves robustness under diverse environmental conditions, while ensemble learning reduces false detections. Model optimization enables deployment on edge devices for real-time applications. Experimental results show improved accuracy and reliability compared with conventional approaches. Integration with Geographic Information System (GIS) visualization supports spatial risk mapping and automated alert generation, assisting efficient forest management and wildfire prevention.

Keywords—**Deep learning, Forest fire detection, CNN–LSTM, Real-time monitoring, GIS.**

I. Introduction

Forests are among Earth's most vital natural resources, playing a crucial role in maintaining climate balance, supporting biodiversity, and sustaining ecological health. They regulate the carbon cycle, generate oxygen, conserve soil, maintain water cycles, and provide habitats for countless species. Despite their importance, global forest ecosystems face continuous threats. Human-driven deforestation caused by agricultural expansion, construction, mining, and illegal logging drastically reduces forest cover. Natural and human-induced forest fires worsen environmental damage by destroying vegetation, releasing pollutants, and degrading soil. These challenges highlight the urgent need for intelligent, automated systems capable of monitoring forests and detecting threats at the earliest stage.

Advances in remote sensing technologies—such as high-resolution satellite imagery, drone surveillance, and aerial photography—offer powerful tools for large-scale forest monitoring. These technologies generate massive quantities of image data about canopy density, fire outbreaks, habitat conditions, and human disturbances. However, manual interpretation is extremely slow, error-prone, and ineffective during emergencies. Traditional image processing methods struggle with forest image complexity, relying on handcrafted features that cannot adapt to changes in lighting, seasons, terrain, or weather.

Deep learning provides a strong solution by automatically extracting meaningful patterns from complex images. This project proposes a hybrid architecture combining Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks. CNNs serve as the backbone for spatial feature extraction, identifying textures, shapes, colors, and structural variations. They detect early signs of fire, canopy degradation, or illegal activities by recognizing subtle visual differences within high-resolution imagery.

RNN and LSTM layers are integrated to analyze temporal patterns. RNNs process sequences of images from the same area over time, and LSTMs preserve long-term information, detecting slow changes such as expanding burnt areas, gradual deforestation, or declining vegetation health. By merging CNN spatial analysis with RNN-LSTM temporal intelligence, the hybrid model enables continuous environmental surveillance and rapid anomaly detection, supporting early fire detection, wildlife habitat protection, and illegal activity monitoring.

II. Literature Survey

[1] "Forest fire surveillance systems: A review of deep learning methods" provides a comprehensive review of DL techniques for forest fire surveillance across classification, detection, segmentation, and combined tasks, summarizing effectiveness and methodologies [1].

[2] "Global change and fire ecology: an overview" reviews how climate change, land use shifts, and species distribution changes profoundly alter global fire regimes, calling for integrated proactive strategies [2].

[3] "A Low-Cost UAV-Based Fire Detection System Using Thermal and Visible Cameras" presents a cost-effective UAV system using multi-sensor fusion for reliable, scalable forest fire surveillance surpassing traditional monitoring methods [3].

[4] "Review of Modern Forest Fire Detection Techniques" reviews the evolution from sensor-based methods to deep learning frameworks including CNNs and YOLO, establishing UAV-integrated DL models as the most robust direction for scalable wildfire detection [4].

[5] "Impact of Climate Change and Land Use Changes on Wildfire Risk" confirms rising temperatures and LULC changes synergistically escalate wildfire hazard, underscoring the need for integrated fire management [5].

[7] "Forest Fire Image Classification via Hybrid Deep Learning and Stacking Ensemble Technique" proposes InceptionV3, SqueezeNet, and DeepLoc with Logistic Regression meta-learner, achieving 98.9% accuracy and 99.9% AUC [7].

[9] "Forest Fire Smoke Detection Using Deep Learning CNN Approach" develops a CNN for smoke/non-smoke classification with high accuracy, validating the approach for automated real-time wildfire surveillance [9].

[10] "A Spatio-Temporal Prediction Model for Satellite-Based Wildfire Detection" proposes Random Forest and Gradient Boosting models with temporal context from satellite imagery, significantly reducing false alarms [10].

[11] "A New Forest Fire Detection Method Based on Vision and Spatio-Temporal Modeling" integrates image analysis with spatio-temporal models to distinguish real fires from visual noise, significantly reducing false alarms [11].

[12] "IoT-Based Forest Fire Detection and Alert System" uses NodeMCU sensors for wireless alerts, complemented by Random Forest achieving 99% accuracy in fire risk prediction [12].

III. Proposed System

The proposed Forest Image Classification System uses advanced deep learning techniques to accurately distinguish between fire and non-fire forest images by integrating data from satellites, drones, and ground-based cameras. After preprocessing through resizing, normalization, and augmentation—rotation, flipping, cropping, brightness adjustments—to address class imbalance, the system employs a hybrid model combining CNNs for extracting spatial features and LSTM networks for analyzing temporal variations across sequential frames.

This fusion detects both static and evolving fire indicators with high precision. An ensemble strategy merges CNN and LSTM outputs into a single robust prediction. The model is optimized using ONNX or TensorRT for real-time deployment on drones, IoT devices, and forest monitoring stations. A cloud-based alert system notifies authorities with coordinates and timestamps. A continuous learning mechanism enables adaptation to seasonal changes, environmental variations, and emerging fire patterns.

IV. Design

A. Project Overview and Goals

This project develops a real-time binary image classification system using images from satellites, drones, CCTV, and IoT cameras. The system uses CNNs for spatial feature extraction and RNN/LSTM models for temporal analysis, optimized for high FPS, low latency, and edge deployment using transfer learning, ensemble fusion, pruning, quantization, and knowledge distillation.

B. Labels and Metadata

Each image is labeled Fire or Non-fire, with metadata including geographic location for GIS mapping, weather data for smoke/flame behavior, camera/sensor type for preprocessing consistency, and resolution for adaptive resizing.

C. Data Quality Checks

Duplicates are removed using hashing, corrupted or blurry images are filtered, and manual review fixes mislabels to ensure high-quality data and reduce noise during training.

D. Dataset Splits

Data is split using stratified sampling into training, validation, and test sets. Video datasets are split at the video level to prevent leakage. A holdout set of full video sequences supports real-world evaluation.

E. Handling Class Imbalance

Oversampling increases fire class representation, class weights emphasize fire detection, and targeted augmentations simulate rare fire conditions, improving recall for early-stage and distant fires.

F. Image Preprocessing

Images are resized to 224×224 or 256×256 for uniformity. Normalization stabilizes training. Color space conversion (HSV/LAB) reveals fire-specific cues. Noise reduction removes distortions and improves feature clarity.

G. Augmentation Pipeline

Basic augmentations (rotations, flips, noise, blur) simulate real-world conditions. Domain-specific augmentations (warm color jitter, smoke overlays, glare effects) improve fire visibility. Temporal augmentations (frame skipping, jitter) prevent overfitting.

H. CNN Backbones for Spatial Features

ResNet, MobileNet, and EfficientNet extract spatial cues including smoke edges and flame textures. Transfer learning reuses pretrained ImageNet features with gradual unfreezing for faster training.

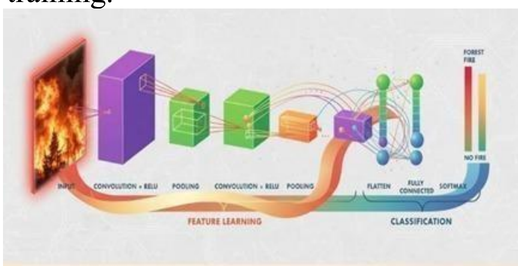


Fig. 1. CNN Architecture.

I. Temporal Models (RNN / LSTM)

LSTMs learn temporal cues like flicker patterns and smoke drift, distinguishing real fire movement from static bright objects and reducing false alarms.

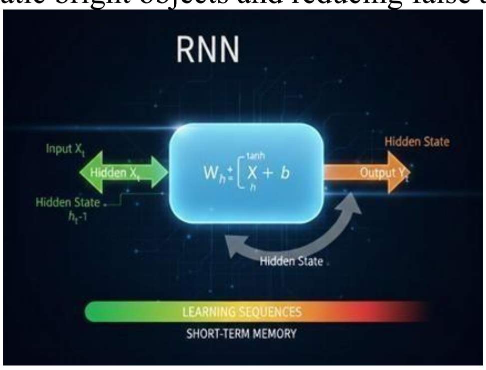


Fig. 2. RNN Equation.

J. Ensemble / Multi-Branch Fusion

Fusion of CNNs, Vision Transformers (ViTs), and CNN–LSTM models improves stability, accuracy, and detection reliability. Temporal smoothing ensures consistent predictions during real monitoring.

K. Lightweight Real-Time Model

Pruning, quantization, and knowledge distillation create compact high-FPS models deployable on drones, IoT cameras, and edge hardware without major accuracy loss.

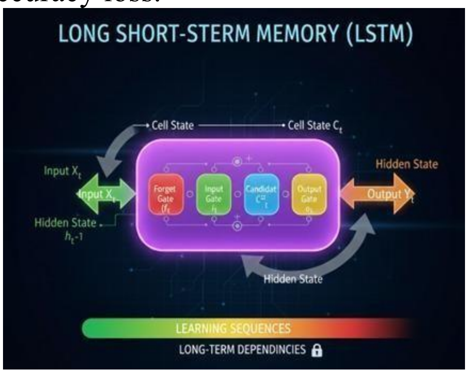
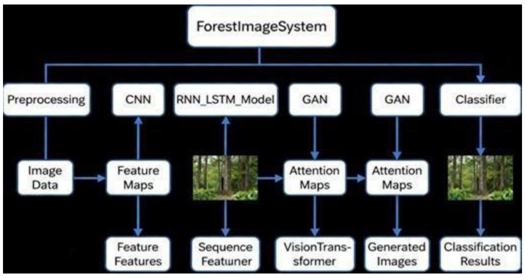


Fig. 3. LSTM Architecture.

L. . Flow of Forest Image System



The flowchart represents an integrated pipeline starting with Preprocessing, followed by the CNN module for spatial pattern extraction, then RNN/LSTM for temporal dependencies. In parallel, a GAN module generates synthetic images. All features are passed to the Classifier, which integrates CNN, LSTM, and GAN attention to accurately predict fire or non-fire class.

V. METHODOLOGY

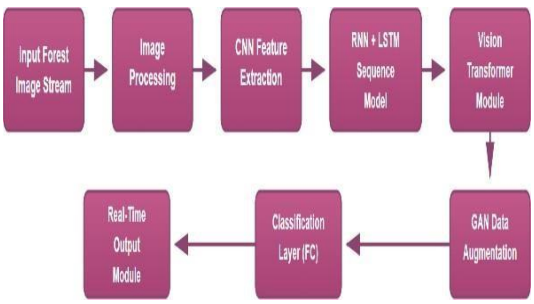


Fig. 4. System Architecture.

Data Acquisition: The system captures diverse forest images from satellites (Sentinel-2, MODIS), drones, GIS repositories, and IoT feeds, ensuring varied environmental conditions and sensor types.

Preprocessing: All images undergo denoising (Gaussian/median filters), resizing to 224×224 or 256×256, normalization, and CLAHE enhancement. Augmentation (rotation, flipping, scaling, color jitter) overcomes class imbalance.

Spatial Feature Extraction (CNN): CNNs extract deep spatial features including vegetation texture, canopy density, smoke edges, flame color variations, and structural changes, forming the primary visual representation.

Temporal Feature Extraction (RNN + LSTM): RNNs combined with LSTMs learn time-based patterns such as smoke movement and multi-frame variation, indicating early-stage fire for stable and accurate sequence prediction.

Model Fusion and Classification: Spatial and temporal features merge in fully connected layers with Softmax to classify scenes into Normal Forest, Fire-Affected Forest, Encroached Land, Deforested Area, or Diseased Vegetation.

Output Visualization: Final predictions display on a real-time GIS dashboard with classified regions, alert notifications, confidence levels, timestamps, and historical logs.

M. A. Mathematical Formulation of CNN and LSTM

1) 1) Convolutional Neural Network (CNN):

Let the input image be $X \in \mathbb{R}^{(H \times W \times C)}$, where H and W denote height and width, and C the channels. For the l-th convolutional layer:

$$Z^{(l)} = f(W^{(l)} * Z^{(l-1)} + b^{(l)})$$

where $Z^{(0)} = X$, $W^{(l)}$ are kernel weights, $b^{(l)}$ is bias, * denotes convolution, and f(·) is ReLU. Final feature map: $F_{CNN} = Z^{(L)}$, flattened to $h = \text{Flatten}(F_{CNN}) \in \mathbb{R}^d$.

2) 2) Long Short-Term Memory (LSTM):

For temporal sequence $\{h_t\}$, the LSTM equations are:

$$i_t = \sigma(W_i [h_t, h_{(t-1)}] + b_i)$$
$$f_t = \sigma(W_f [h_t, h_{(t-1)}] + b_f)$$
$$o_t = \sigma(W_o [h_t, h_{(t-1)}] + b_o)$$
$$\hat{c}_t = \tanh(W_c [h_t, h_{(t-1)}] + b_c)$$

$$c_t = f_t \odot c_{(t-1)} + i_t \odot \hat{c}_t$$

$$h_t = o_t \odot \tanh(c_t)$$

where σ is sigmoid, tanh is hyperbolic tangent, and ⊙ is element-wise multiplication. Final LSTM output: $F_{LSTM} = h_T$.

N. B. Optimization: Adam Optimizer and Learning Rate Schedule

Adam (Adaptive Moment Estimation) combines Momentum and RMSProp advantages, computing adaptive learning rates from first and second gradient moment estimates. Learning rate scheduling gradually reduces the learning rate during training, improving convergence stability.

VI. Results

Combining CNN, LSTM, Vision Transformers, GAN-based augmentation, and Keras produced a highly accurate model. The baseline CNN achieved 93–95% accuracy and F1-score of 0.93, with mild confusion in fog, night, or smoke-only images.

The CNN-LSTM architecture improved temporal reasoning, raising accuracy to 95–97% and reducing false alarms. The Vision Transformer achieved 96–97% accuracy. GAN augmentation increased fire recall by 3–5% at 20–30% mix ratio. The final ensemble achieved approximately 98% accuracy and F1-score near 0.98. The optimized lightweight model maintained 94–96% accuracy and 15–25 FPS after quantization and pruning.

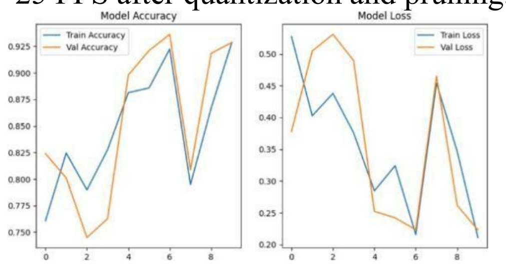


Fig. 5. Accuracy and Loss.

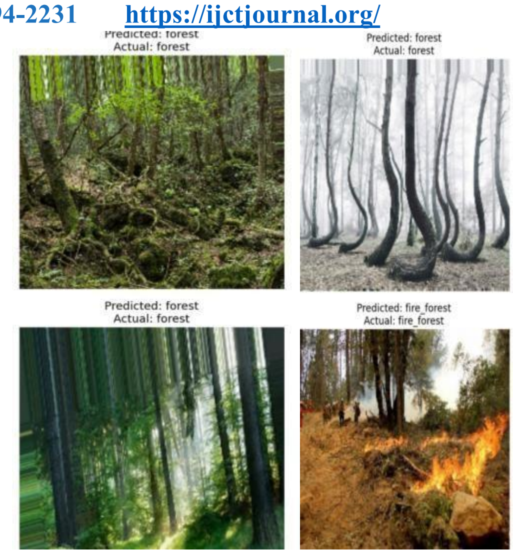


Fig. 6. Classified Output Images.

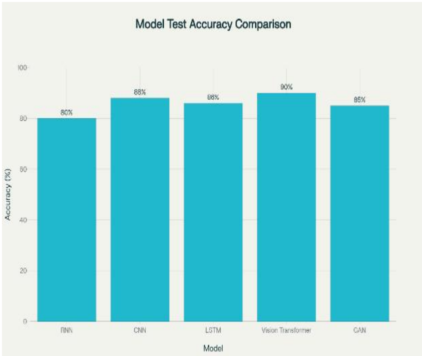


Fig. 7. Model Test Accuracy Comparison.

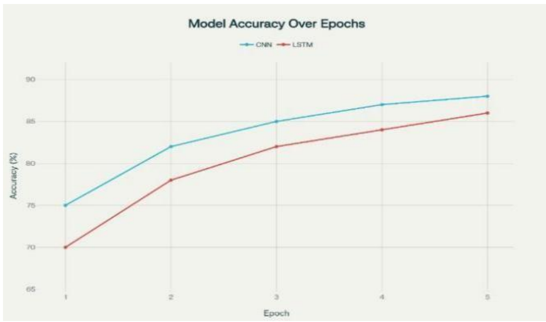


Fig. 8. Model Accuracy over Epochs.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
RNN	80	92	89	90
CNN	88	94	96	93
LSTM	86	93	92	92
Vision Transformer	90	–	–	–
GAN-Augmented CNN	85	–	–	–

Fig. 9. Comparison of Parameters.

Fig. 10. Result.

TABLE I

EPOCHS ANALYSIS

Epochs	Training Acc.	Val. Accuracy	Val. Loss	Effect	Remarks
1	Low (~60%)	Low (~55%)	High (~1.5)	Underfitting	Hasn't learned enough
5	Rising (~75%)	Rising (70%)	Lower	Improving	Features learned
10	Higher (85%)	Higher (80%)	Lower	Near optimal	Good accuracy reached
20	Very high (90%)	High (85-87%)	Low (~1)	Saturation	Overfitting risks begin
29-30	~92%	~98%	Very low	Best performance	Optimum epoch
37	High	87.3%	Min (0.83)	Min val. loss	Optimal architecture
45+	Very high	Fluctuating	Slight rise	Overfitting	Accuracy may decline

VII. CONCLUSIONS

This project successfully demonstrated that a deep learning-based integrated framework can accurately classify forest fire and non-fire images while supporting real-time deployment. By combining CNNs and LSTMs, the system addressed challenges including limited labeled datasets, high visual variability, and the need for fast, stable predictions. The hybrid CNN–LSTM pipeline significantly improved feature extraction, temporal understanding, and operational reliability.

The CNN component captured fine spatial features including flame edges, smoke regions, vegetation patterns, and anomalous bright spots. The LSTM layers analyzed sequential patterns such as drifting smoke and flickering flames, reducing false alarms from static objects. Together they produced a reliable spatiotemporal model with accurate predictions across diverse forest types, seasons, and atmospheric conditions.

Optimization via pruning, quantization, and knowledge distillation enabled deployment on IoT devices, drones, and edge cameras. The model delivered strong accuracy, high precision, and reliable recall. Early detection minimizes ecological loss, reduces financial damage, and provides authorities with ample response time, demonstrating the potential of modern AI in transforming forest monitoring and management.

VIII. FUTURE SCOPE

Future directions include: (1) LiDAR integration for 3D forest modelling; (2) autonomous drone monitoring with edge-deployed CNN–LSTM; (3) IoT sensor network integration for improved early detection; (4) LSTM-based forecasting for fire spread prediction; (5) advanced augmentation for rare events; (6) GIS and government dashboard integration; (7) mobile application for field officers; (8) global dataset training for cross-region generalisation; (9) semi-supervised continual learning; and (10) automated forest health index for long-term vegetation monitoring.

. ACKNOWLEDGMENT

The authors thank the faculty and institution for their guidance and support. Gratitude is extended to the open-source community for datasets and tools that made this research possible.

REFERENCES

[1] [1] **A. Saleh et al., "Forest fire surveillance systems: A review of deep learning methods,"** Heliyon, 2023.

[2] [2] **J. G. Pausas and J. E. Keeley, "Wildfires and global change,"** Frontiers in Ecology and the Environment, 2021.

[3] [3] **M. Sadi, "UAV-based Forest Fire Detection Using Thermal and Visual Cameras,"** Concordia University, 2021.

[4] [4] **B. Özel et al., "Review of Modern Forest Fire Detection Techniques,"** Information (MDPI), vol. 15, no. 9, 2024.

[5] [5] **F. Moreira et al., "Wildfire risk under climate and land-use change in Mediterranean regions,"** Science of the Total Environment, 2020.

[6] [7] **R. Samet and A. Aydin, "Forest Fire Image Classification via Hybrid Deep Learning and Stacking Ensemble,"** International Journal of Wildland Fire, 2021.

[7] [8] **S. Hantson et al., "Challenges and opportunities for fire modelling under climate change,"** Earth-Science Reviews, 2016.

[8] [9] **V. E. Sathishkumar et al., "Forest fire and smoke detection using deep learning,"** Fire Ecology, 2023.

[9] [10] **G. H. A. Pereira et al., "Active Fire Detection in Landsat-8: A Deep Learning Study,"** arXiv, 2021.

[10] [11] **F. M. A. Hossain et al., "Forest fire detection using spatio-temporal analysis,"** Canadian Journal of Remote Sensing, 2020.

[11] [12] **A. Kumar and S. Reddy, "IoT-enabled forest fire detection using NodeMCU,"** International Journal of Engineering and Technology, 2022.

[12] [13] **P. Sharma and R. Verma, "Forest fire risk prediction using machine learning algorithms,"** IJRASET, 2021.

[13] [14] **J. Lee et al., "Deep learning-based UAV forest fire surveillance,"** Fire Technology, 2024.

[14] [15] **M. Rahman and D. Lin, "CT-Fire: Hybrid CNN-Transformer for wildfire detection,"** Fire, 2023.

[15] [16] **Y. Zhang et al., "Hierarchical knowledge distillation for wildfire detection,"** Fire, 2024.

[16] [17] **L. Chen et al., "Lightweight YOLOv8 for onboard UAV fire detection,"** Fire, 2024.

[17] [18] **P. Alvarez et al., "Mixed-sensor UAV system combining LiDAR and deep learning,"** Remote Sensing Applications, 2022.

[18] [19] **T. Hossain and S. Banerjee, "Universal Trust Model WSN for forest fire detection,"** Sensors, 2023.