iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79541

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Deep Learning Based Real Time Fire Detection and Alert System

Sk. Haseena Begum[1], O.Sandeep Kumar[2], M.Suresh Babu[3], M. Naga Durga Prasad[4], G. Sundar Rohith Kiran[5]
[1]*Asst. Professor,* [2,3,4,5]*Student, Department of CAI, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh*

Abstract:Fire accidents are a major threat to human lives, property, and the environment. Traditional fire detection systems that rely on smoke and heat sensors often respond slowly and may not work effectively in large or open areas. They can also generate false alarms, which reduces their reliability.To overcome these challenges, this paper presents an intelligent fire detection system based on a hybrid deep learning model that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The CNN is used to analyse individual image frames and extract important visual features such as flame colour, texture, and smoke patterns. Meanwhile, the LSTM captures the changes that occur over time by analysing consecutive frames, helping the system understand how fire develops and spreads.By combining both spatial and temporal analysis, this hybrid model improves detection accuracy and significantly reduces false alarms caused by fire-like elements such as sunlight reflections or artificial lighting. The system is trained and tested on a dataset containing both fire and non-fire images and videos, including indoor, outdoor, and forest fire situations. The performance of the model is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. The experimental results show that the proposed CNN-LSTM model performs better than traditional CNN-based methods, offering improved reliability and real-time fire detection capability.
Keywords: ***Fire Detection, Deep Learning, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Video Surveillance, Temporal Feature Extraction, Spatial Feature Extraction***

I. INTRODUCTION

Fire accidents are among the most destructive disasters, impacting homes, industries, forests, and public infrastructure. Every year, global safety reports highlight the severe economic losses and tragic loss of lives caused by fire-related incidents.

In recent years, deep learning techniquesespecially Convolutional Neural Networks (CNNs)have shown excellent performance in image classification and object detection tasks. CNNs are highly effective at extracting important visual features from images.

To address this limitation, this research introduces a hybrid deep learning framework that combines CNN with Long Short-Term Memory (LSTM) networks. In this approach, the CNN extracts spatial features from each video frame, while the LSTM analyses the sequence of frames to understand the temporal progression of fire.

By integrating both spatial and temporal learning, the proposed model enhances detection accuracy and significantly reduces false alarms caused by fire-like elements such as sunlight reflections, vehicle headlights, or artificial lighting.

II. RELATED WORK

Fire detection has traditionally depended on sensor-based technologies such as smoke detectors, heat sensors, and infrared systems. While these methods work well in closed indoor spaces, they often face challenges.

Early computer vision techniques relied on handcrafted features like color, shape, and motion patterns to identify fire. For instance, Toreyin et al. (2005) proposed a flame detection method using wavelet transforms and hidden Markov models, but their system was highly sensitive to light reflections. Similarly, Chen et al. (2004) applied RGB and YCbCr color models to detect flame-like pixels; however, their approach struggled in environments with intense lighting or direct sunlight.

The development of Convolutional Neural Networks (CNNs) significantly improved fire detection performance. Muhammad et al. (2018) introduced a CNN-based classifier for fire detection in images, demonstrating better resistance to background noise. Despite these advancements, CNN models mainly focus on spatial information, which limits their ability to distinguish real fire from static fire-like objects such as lamps or vehicle headlights.

To overcome this issue, researchers incorporated temporal learning through Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. Zhang et al. (2019) proposed a CNN-LSTM framework that captures both spatial and temporal features from video sequences, leading to higher accuracy and fewer false alarms. Likewise, Frizzi et al. (2018) combined color-based features with motion analysis to better differentiate flames from other bright light sources.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Recent studies have focused on hybrid deep learning models that integrate CNNs for spatial feature extraction and LSTMs for temporal analysis. These systems can effectively analyze flame flickering behavior, smoke movement, and time-based variations. Models from the YOLO family, introduced by Joseph Redmon et al. (2016),along with MobileNet-based architectures, have further improved real-time fire detection, especially on edge devices like surveillance cameras and drones.

The system proposed in this work builds upon previous research by implementing a CNN–LSTM hybrid architecture that captures both frame-level visual details and motion patterns over time. This design enhances performance across various scenarios, including forest fires, indoor incidents, and nighttime conditions. Moreover, it addresses key limitations found in earlier models, such as false alarms caused by lighting reflections and the absence of temporal understanding.

A. Existing System

Traditional fire detection systems mainly rely on hardware-based sensors such as smoke detectors, heat sensors, flame detectors, and infrared sensors. These systems are commonly installed in homes, industries, warehouses, and commercial buildings because they are affordable and easy to set up. However, despite their widespread use, they have several drawbacks that limit their effectiveness, particularly in terms of efficiency, response time, and real-time performance.

1) Sensor-Based Fire Detection

Most existing fire safety systems rely on devices such as smoke detectors (photoelectric or ionization), heat sensors (rate-of-rise or fixed temperature), infrared or UV flame detectors, and gas sensors that monitor gases like CO and CO_2. Although these systems are capable of detecting fire once combustion has begun, they depend on physical signs such as a rise in temperature or the accumulation of smoke. Because of this reliance on measurable changes, detection is often delayed, particularly in open spaces, high-ceiling industrial buildings, outdoor environments, or areas with strong ventilation.

2) 2.1.2 Limited Coverage and Environmental Sensitivity

Sensor-based fire detection systems are generally effective only within a limited and localized area of coverage. They are not suitable for monitoring large spaces such as warehouses, forest areas, long tunnels, or public outdoor locations. In addition, environmental conditions like humidity, dust, airflow, and exposure to sunlight can significantly affect their performance. These factors may reduce detection accuracy and increase the likelihood of false alarms or, in some cases, failure to detect an actual fire.

3) No Visual Understanding or Scene Awareness

Conventional fire detection systems are not capable of analyzing detailed visual characteristics such as flame color variations, smoke patterns, flame movement, or the dynamic spread of fire. As a result, they lack the intelligence needed to differentiate between actual fire and similar visual effects like reflections, bright lighting, lamps, or vehicle headlights. This limitation frequently leads to false alarms, which can cause unnecessary panic, operational interruptions, and reduced system reliability.

4) No Temporal or Predictive Capability

Existing sensor-based fire detection systems are unable to track the progression of fire over time, identify small flames before noticeable smoke is produced, or predict fire growth by analyzing motion patterns. These systems typically respond only after a considerable amount of heat or smoke has accumulated. As a result, early detection and timely intervention are often delayed, increasing the risk of damage and danger.

5) Limited Use in Surveillance

Most modern CCTV surveillance systems are designed primarily for recording and storing video footage. They generally do not incorporate intelligent, automated fire detection capabilities. Because of this limitation, security staff must continuously monitor camera feeds manually, which increases the likelihood of human error and can delay emergency response during critical situations.

III. METHODOLOGY

The proposed Fire Detection System is built using a hybrid deep learning architecture that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The overall methodology follows a structured approach, including stages such as data collection, data preprocessing, feature extraction, temporal modeling, and final classification.

A. Overall System Workflow

The overall workflow of the proposed system begins with capturing video or image input, which is then processed through frame extraction. Each frame undergoes image preprocessing to enhance quality and prepare it for analysis.

Next, spatial features are extracted using a Convolutional Neural Network (CNN). These features are then passed to a Long Short-Term Memory (LSTM) network to learn temporal patterns across consecutive frames. Based on the combined spatial and temporal information, the system performs binary classification to determine whether the scene contains fire or non-fire. Finally, if fire is detected, an alert is automatically generated.

B. Data Collection

The dataset used in this study includes both fire and non-fire samples collected from multiple sources, such as publicly available fire image datasets, Kaggle datasets, online surveillance footage, and custom-recorded videos. It covers a wide range of real-world scenarios, including indoor fires, forest fires, vehicle fires, nighttime fire incidents, as well as fire-like situations involving sunlight, lamps, and reflections. Incorporating diverse conditions helps improve the robustness and generalization capability of the proposed model.

C. Data Preprocessing

Before training the model, the collected data is carefully pre-processed to ensure consistency and enhance overall performance. Video inputs are first converted into individual frames at fixed intervals to effectively capture temporal changes. All images are then resized to a standard dimension, such as 224 × 224 pixels, to meet the input requirements of the CNN architecture. The pixel values are normalized to a range between 0 and 1 to improve numerical stability during training. Additionally, data augmentation techniques—including rotation, horizontal flipping, zooming, and brightness adjustment—are applied to increase dataset variability and minimize the risk of overfitting.

D. Spatial Feature Extraction

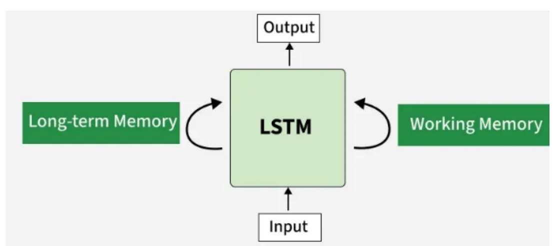

Fig: WORKING OF LONG SHORT-TERM MEMORY

Convolutional Neural Networks (CNNs) are employed to extract spatial features from each input frame. The CNN architecture is composed of convolutional layers, activation functions such as RELU, and pooling layers. Together, these components identify important visual details, including flame color distribution, smoke texture, edges, and intensity variations. After processing, the CNN transforms each frame into a meaningful feature vector that effectively represents the spatial characteristics of the scene.

E. Temporal Modelling

Since fire is a dynamic phenomenon characterized by continuous motion and changes in intensity over time, Long Short-Term Memory (LSTM) networks are utilized to model the temporal relationships between consecutive frames. The sequence of feature vectors extracted by the CNN is provided as input to the LSTM network. The LSTM learns the progression, flickering behaviour, and movement patterns of fire, enabling the system to differentiate real fire from static fire-like objects. This temporal modeling approach significantly reduces false alarms and enhances the overall reliability of fire detection.

F. Classification Layer

The output generated by the LSTM network is forwarded to a fully connected layer equipped with a Sigmoid activation function for binary classification. Based on this processing, the model determines whether the input sequence represents a fire or a non-fire event. The final output is expressed as a probability score, where values close to 1 indicate the presence of fire, and values close to 0 represent a non-fire condition.

G. Model Training

The model is trained using Binary Cross-Entropy as the loss function, while the Adam optimizer is employed to ensure efficient and adaptive weight updates during learning. The dataset is split into training, validation, and testing sets to properly assess the model's performance and generalization capability.

H. Performance Evaluation

The performance of the proposed CNN–LSTM model is assessed using standard evaluation metrics, including Accuracy, Precision, Recall, and F1-Score. In addition, a confusion matrix is utilized to provide a detailed analysis of the model's predictions by identifying true positives, true negatives, false positives, and false negatives.

IV. MODELLING AND ANALYSIS

This modelling and analysis clearly demonstrate that integrating both spatial and temporal deep learning techniques leads to a more reliable, accurate, and robust fire detection system.

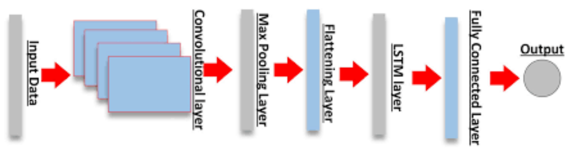

Fig: PROCESSING OF NEURAL NETWORK

A. Model Architecture Design

The proposed Fire Detection System is designed using a hybrid deep learning architecture that combines Convolutional Neural Networks and Long Short-Term Memory networks. In this framework, the CNN extracts important spatial features from individual frames, while the LSTM analyses temporal relationships across consecutive frames to understand the progression of fire over time.

.

B. Mathematical Modelling

The modeling of the proposed system can be expressed mathematically in three stages:

1) Spatial Feature Extraction (CNN)

For each input frame X_t:

$$F_t = CNN(X_t)$$

Where:

- X_t= Input frame at time t
- F_t= Extracted spatial feature vector

2) Temporal Modeling (LSTM)

The sequence of feature vectors is processed as:

$$h_t = LSTM(F_t, h_{t-1})$$

Where:

- h_t= Current hidden state
- h_{t-1}= Previous hidden state

3) Final Classification

$$Y = \sigma(W \cdot h_t + b)$$

Where:

- Y= Output probability
- W= Weight matrix
- b= Bias
- σ= Sigmoid activation function

If $Y > 0.5$, the system classifies the input as Fire; otherwise, Non-Fire.

4) Training Analysis

The The model is trained using Binary Cross-Entropy loss, defined as:

$$Loss = -[y\log(p) + (1-y)\log(1-p)]$$

Where:

- y= True label
- p= Predicted probability

The Adam optimizer is used for adaptive learning rate optimization. The dataset is divided into training, validation, and testing sets to prevent overfitting and ensure generalization.

C. Performance Metrics Analysis

The To evaluate model effectiveness, the following metrics are used:
Accuracy – Overall correctness of predictions
Precision – Correct fire detections among predicted fire cases
Recall – Ability to detect actual fire cases
F1-Score – Harmonic mean of precision and recall

D. Comparative Analysis

To evaluate the effectiveness of the hybrid approach, it is compared with a standalone CNN model. Although a CNN can efficiently extract spatial features, it does not capture temporal relationships between frames. By incorporating LSTM into the architecture, the model gains the ability to understand dynamic patterns over time, which significantly enhances detection accuracy in real-world scenarios and reduces false alarms caused by static fire-like objects.

V. RESULTS AND DISCUSSIONS

We The results confirm that the hybrid CNN–LSTM model offers a reliable and efficient solution for real-time fire detection in surveillance systems and safety-critical environments. By effectively combining spatial feature extraction, temporal modelling, and accurate classification, the system demonstrates improved performance and robustness compared to traditional approaches.

A. Experiment Result

The The proposed Fire Detection System based on the CNN + LSTM architecture was evaluated using a dataset containing both fire and non-fire image sequences. To ensure an unbiased assessment of performance, the dataset was divided into training, validation, and testing sets. The model was trained over multiple epochs using the Adam optimizer for efficient learning and the Binary Cross-Entropy loss function to optimize binary classification performance.

B. Confusion Matrix Analysis

	Actual Fire Positive	Actual No Fire Negative
Predicted Fire	TP = 48%	FP = 3.5%
Predicted No Fire	FN = 2.0%	TN = 46.5%

The number of True Positives (TP) is considerably high, demonstrating the model's strong ability to correctly detect fire events. Similarly, the True Negatives (TN) indicate accurate identification of non-fire situations. Compared to standalone CNN models, the False Positives (FP) are significantly reduced, reflecting improved system reliability. The False Negatives (FN) are minimal, which is crucial for ensuring timely fire detection. Reducing false positives is especially important in real-world applications, as unnecessary alarms can cause operational disruptions and reduce trust in the system.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

C. Discussion

The The experimental results clearly show that the proposed CNN–LSTM architecture successfully captures both the spatial and temporal characteristics of fire. The CNN component extracts important visual details such as flame texture, colour distribution, and smoke patterns from individual frames.

One of the major advantages observed during the analysis is the model's ability to distinguish real fire from static bright objects. Traditional CNN-based methods may incorrectly classify stationary light sources as fire because of similar colour features. However, by incorporating LSTM to analyse motion patterns over time, the system effectively reduces such false detections and improves overall reliability.

.

VI. OUTPUT SCREENS

The final output screen of the FireWatch AI system shows how the model detects fire directly from an input image and displays the result on the screen. In this stage, the system analyses the uploaded frame using the proposed Hybrid CNN + LSTM model and highlights the detected fire region along with the prediction information.

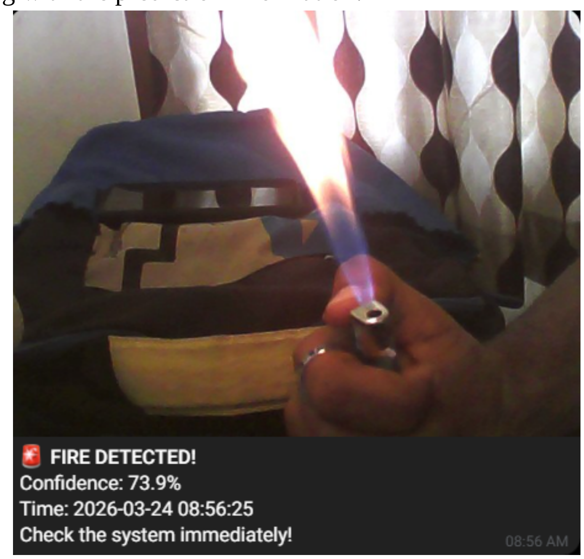

Fig: OUTPUT SCREEN

This label appearson the frame once the model detects fire. It provides three key pieces of information: the predicted class, the confidence level of the prediction, and the approximate location of the detected region in the image.

First, the label "FIRE" indicates that the system has classified the visual input as containing fire. The convolutional neural network (CNN) analyses the image and extracts important visual features such as colour intensity, flame texture, and brightness patterns.

The value "1.00" represents the confidence score of the prediction. This means the model is almost completely certain that the detected object in the frame is fire. Such a high confidence value usually occurs when the flames are clearly visible and the visual characteristics strongly match the fire patterns learned during training.

The final part of the label "[bottom_left_x2.00]" provides positional information about where the fire is detected within the frame. This helps the system identify the approximate location of the fire, allowing it to highlight the area for monitoring or further action.

The model detects fire by identifying several visual characteristics, including bright red, orange, and yellow colour patterns, irregular flame shapes, and small spark-like particles around the burning region. These patterns are captured by the CNN layers, which specialize in detecting spatial features from images.

After feature extraction, the LSTM component helps confirm whether the detected pattern behaves like real fire by analysing sequential information between frames. This combination improves the reliability of the system and reduces the chances of false alarms caused by bright lights or reflections.

When the confidence level exceeds the predefined threshold, the system generates a fire alert and displays the detection result directly on the image. This real-time visual feedback allows users to quickly understand the situation and take appropriate action if necessary.

VII. CONCLUSIONS

This paper presented an intelligent Fire Detection System built on a hybrid deep learning architecture that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The proposed method overcomes the limitations of traditional sensor-based fire detection systems as well as standalone image-based models by integrating both spatial and temporal feature learning into a unified framework.

In this architecture, the CNN effectively extracts spatial features from individual frames, including flame colour distribution, texture patterns, and smoke characteristics. Meanwhile, the LSTM network captures temporal relationships and motion patterns across consecutive frames. The integration of these two components significantly enhances detection accuracy and minimizes false positives caused by fire-like elements such as sunlight reflections, artificial lighting, and other bright surfaces.

VIII. FUTURE SCOPE

The proposed CNN–LSTM-based fire detection system demonstrates strong potential for real-time detection across diverse environments. However, there remain several opportunities for further improvement and large-scale deployment. Future enhancements can focus on increasing model efficiency, expanding dataset diversity, optimizing performance for edge devices, and strengthening integration with IoT and smart monitoring systems. By addressing these areas, the system's overall accuracy, scalability, and practical usability can be significantly improved.

A. Integration with IoT and Smart Devices

The proposed system can be further enhanced by integrating it with IoT-based smoke and temperature sensors, smart alarms, automatic sprinkler systems, and edge computing devices such as Raspberry Pi and NVIDIA Jetson Nano. Such integration would enable a fully automated fire response mechanism, significantly improve reaction time while minimize reliance on manual monitoring and intervention.

B. Creation of a Large Real-World Fire Dataset

Current fire detection datasets are limited and do not fully represent the wide range of real-world fire scenarios. Future research can focus on developing a large-scale and diverse fire and smoke dataset that includes extreme lighting conditions, varying weather environments, reflective surfaces, and multi-angle video recordings of both indoor and outdoor fires. Expanding dataset diversity will enhance model generalization and robustness.

C. Multi-Class Detection (Fire, Smoke, Explosion, Hazard Events)

The system can be extended beyond binary classification to support multi-class detection. This may include identifying different smoke levels, detecting gas leak indicators, recognizing explosions, and estimating fire intensity and spread. A multi-class framework would provide more comprehensive monitoring capabilities, especially for industrial safety systems and public infrastructure protection.

D. 3D Fire Localization and Spread Prediction

Future enhancements may incorporate depth estimation techniques to determine the precise 3D location of fire within a scene. Additionally, advanced temporal models can be developed to predict the direction of fire spread and estimate fire size and associated risk levels. These improvements would support proactive decision-making and disaster management planning.

E. Integration with Drone Surveillance

The system can also be deployed on drones equipped with AI-based vision systems to monitor forest regions, industrial facilities, and high-risk construction zones. Integrating the CNN–LSTM architecture into drone platforms would enable autonomous fire detection, even in remote or difficult-to-access terrains, significantly improving early warning capabilities.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NIPS), 2012.

[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in International Conference on Learning Representations (ICLR), 2015.
[4] C. Szegedy et al., "Going Deeper with Convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[6] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[7] A. Graves, "Long Short-Term Memory," in Supervised Sequence Labelling with Recurrent Neural Networks, Springer, 2012.
[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015.
[9] G. Muhammad, M. H. Al-Mutib, M. Alsulaiman, and Z. Ali, "Automatic Fire Detection Using Deep Learning Approaches," IEEE Access, vol. 6, pp. 18105–18112, 2018.
[10] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Wavelet Based Real-Time Smoke Detection in Video," in IEEE International Conference on Image Processing (ICIP), 2005.
[11] T. Celik and H. Demirel, "Fire Detection in Video Sequences Using a Generic Color Model," Fire Safety Journal, vol. 44, no. 2, pp. 147–158, 2009.
[12] P. Barmpoutis, P. Papaioannou, K. Dimitropoulos, and N. Grammalidis, "A Review on Early Forest Fire Detection Systems Using Optical Remote Sensing," Sensors, vol. 20, no. 22, 2020.
[13] J. Chen, Y. He, and J. Wang, "Fire Smoke Detection Based on Deep Convolutional Neural Networks," in IEEE International Conference on Computer and Information Technology, 2017.
[14] S. Zhang, Y. Wu, T. Menon, and Y. Zhao, "Video-Based Smoke Detection Using Spatio-Temporal Deep Neural Networks," Fire Technology, 2019.
[15] D. Jain, A. Singh, and S. Kumar, "Deep Learning Based Fire and Smoke Detection Using CNN," in International Conference on Advances in Computing and Communication Engineering, 2019.