BRAIN. Broad Research in Artificial Intelligence and Neuroscience
e-ISSN: 2067-3957 | p-ISSN: 2068-0473
Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)
2026, Volume 17, Issue 2, pages: 108-114
Submitted: January 29th, 2026 | *Accepted for publication:* March 23rd, 2026

Application of ConvLSTM Neural Networks in Forest Fire Early Warning Systems for Vietnam

Pham Thi Lien
Thai Nguyen University of Information and Communication Technology,
Thai Nguyen, Vietnam.
ptlien@ictu.edu.vn,
https://orcid.org/0009-0009-4566-4779

Nguyen Thu Huong
Thai Nguyen University of Information and Communication Technology,
Thai Nguyen, Viet Nam
MIREA – Russian Technological University,
Moscow, Russia.
nthuongcntt@ictu.edu.vn,
https://orcid.org/0000-0002-7009-2815

Nguyen The Long
Sao Do University, Hai Phong, Viet Nam
MIREA – Russian Technological University,
Moscow, Russia.
nt_long@saodo.edu.vn,
https://orcid.org/0000-0002-9507-8198

Abstract: *Forest fires pose a significant threat to Viet Nam, particularly during the dry season. This study investigates a ConvLSTM-based deep learning architecture for early fire detection. Unlike single-frame or threshold-based methods, ConvLSTM jointly models spatial features (via convolutional layers) and temporal dependencies (via LSTM units). We utilize a publicly available dataset of 999 fire and non-fire images. To apply ConvLSTM to static images, we construct temporal sequences using a sliding window over augmented variants. The proposed model achieves 98.3% accuracy, 98.1% precision, 96.8% recall, and a 98.1% F1-score on the test set. These results are compared with standalone CNN and LSTM models. Limitations include the limited dataset size, class imbalance (75% fire images), and the lack of Viet Nam-specific data. Future work should focus on larger, region-specific datasets and real-time deployment on edge devices.*

Keywords: *forest fire detection; feature maps; LSTM; convolutional neural networks; deep learning.*

How to cite: Pham, T. L., Nguyen, T. H., & Nguyen, T. L. (2026). Application of ConvLSTM neural networks in forest fire early warning systems for Vietnam. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, *17*(2), 108-114. https://doi.org/10.70594/brain/17.2/7

1. Introduction

Forest fires are increasing in frequency and intensity worldwide due to climate change, characterised by rising temperatures and prolonged dry seasons (Pan et al., 2021; Almasoud, 2023; Ban et al., 2022). In Viet Nam, the risk of forest fires is particularly high in the Central Highlands and northern mountainous regions, where dense forests coincide with agricultural burning and land clearing practices (Idroes et al., 2023). Conventional fire detection methods – human observation, fixed sensor thresholds, and satellite imagery with long revisit intervals – are often slow and unreliable (Abdusalomov et al., 2023). By the time smoke is visually confirmed,the fire is often beyond effective control. These traditional approaches cannot adequately model the complex, nonlinear interactions among meteorological, geographical, and anthropogenic factors.

Deep learning offers a promising alternative. Convolutional neural networks (CNNs) effectively extract spatial features from images (Mei & Zhang, 2018; Dimitropoulos et al., 2017), while long short-term memory (LSTM) networks model temporal dependencies in sequential data (Tran et al., 2022). However, most existing approaches treat either spatial or temporal information separately. ConvLSTM (Zhang et al., 2021) extends standard LSTM by replacing fully connected gates with convolutional operations, thereby preserving spatial structure while modelling temporal sequences – a key capability for fire growth prediction.The objectives of this study are:

(1) To develop a ConvLSTM model for binary classification (fire vs. non-fire).

(2) To compare its performance against CNN and LSTM baselines.

(3) To identify limitations related to dataset size, imbalance, and geographic specificity.

The paper is organised as follows. Section 2 reviews related work. Section 3 describes our method: data collection, image pre-processing, ConvLSTM model construction, and evaluation. Section 4 presents experimental results. Section 5 provides the discussion and conclusions.

2. Related Work

Forest fires spread quickly and cause substantial damage to the environment and public health. For a long time, fire detection relied mainly on direct observation of smoke or flames. This approach is slow, inefficient, and often impractical in remote and hard-to-reach areas (Pan et al., 2021).

Recently, however, AI has enabled new possibilities for smarter detection systems. Many studies have focused on using convolutional neural networks (CNNs) (Almasoud, 2023; Ban et al., 2020) to analyse images from cameras and satellites. The model learns the visual characteristics of smoke and fire and becomes quite good at recognising them. At the same time, other researchers have utilised recurrent neural networks such as LSTMs, which are well-suited for processing time-series data. These models examine sequences of weather station readings – temperature, humidity, wind speed – to identify dangerous patterns that often appear before a fire occurs.

A notable trend is the shift towards hybrid models that combine both visual and sensor data. Using information from multiple sources helps reduce false alarms, which has always been a weakness of traditional systems. Researchers are also aiming to make these AI models lighter and faster so that they can run on edge devices such as remote sensors or drones, and analyse data in real time even in places with limited internet connectivity. The ultimate goal is automated systems that warn earlier and more reliably, giving firefighting teams valuable extra time before a small fire turns into a large one.

Idroes et al. (2023) proposed TeutongNet – a fine-tuned deep learning model based on ResNet50V2 – for accurate forest fire detection. The model was trained on carefully prepared data and tested thoroughly. Results showed that TeutongNet could classify accurately while keeping false alarms and missed detections low. ROC curve analysis confirmed that the model reliably distinguished between fire and non-fire images. This appears to be a promising tool for early detection.

Abdusalomov et al. (2023) built a custom CNN using the Detectron2 framework and trained it on a specialised fire image dataset. The model significantly improved fire detection in remote areas, both during the day and at night, achieving a precision. Meanwhile, Mei and Zhang (2018) took a different approach using logistic regression and temporal smoothing, employing features such as size, motion, and colour to identify fires while removing the background to isolate potential flame regions.

Dimitropoulos et al. (2017) first identified potential fire zones using random forest, SVM, and an improved ViBe algorithm, then confirmed fire presence using Hu moments, overlap rates, and motion intensity between frames. Another study combined multiple features with an adaptive fractional fusion method and used SVM to detect smoke (Tran et al., 2022).

Zhang et al. (2021) proposed ATT Squeeze U-Net for fire segmentation and recognition. They combined SqueezeNet with a modified Fire module inside an ATT U-Net structure, improving feature extraction even when training data was limited. A separate recognition model shared part of the encoding path for classification tasks. This approach (Yun, 2025) is more efficient than existing segmentation and recognition systems, and is especially useful for fire detection where high sensitivity is required but training data is scarce. Tests confirmed solid segmentation accuracy and reliable recognition. However, although the method accurately identifies fire conditions and segments fire areas in detail, it may fail to detect some types of fire in diverse scenarios.

A hybrid deep learning model combining CNN and RNN for forest fire detection was introduced in Ghosh and Kumar (2022). The CNN extracts features, which are then flattened and fed to the RNN. While the CNN handles basic and complex visual patterns, the RNN deals with sequential dependencies in the data. Together, they improve fire pattern recognition. A fully convolutional neural network (FCN) to detect fire smoke in satellite images near real-time was developed by Larsen et al. (2021) and Li et al. (2018). They trained it using data from existing operational smoke detection methods, so the model learned from verified smoke patterns. The system was shown to be practical for near-instant use and was tested during an Australian wildfire. This FCN-based tool shows promise for assessing smoke exposure risks and providing timely information to fire management teams, health and environmental agencies, and the public – potentially reducing health risks through faster warnings.

A forest fire prediction method that combines human activity data, satellite imagery, and environmental variables within a deep learning framework was proposed by Mambile et al. (2025) and Jamshed et al. (2022). The model uses the ConvLSTM architecture to capture both spatial and temporal patterns. It predicts high-risk periods using weather and human behaviour data, while identifying fire-prone areas through satellite-derived indices such as NDVI, NBR, and NDWI. By integrating natural and human factors, this approach offers practical guidance for fire management, helping to optimise resource allocation in vulnerable regions.

3. Proposed Methodology for Forest Fire Detection Using ConvLSTM and Computer Vision

The forest fire detection system we developed has two main features.

(i) Early detection and warning. The system uses optical cameras and sensors to spot early signs of fire, such as high temperatures, dry conditions, smoke, and flames. This helps forest management agencies respond faster and reduce damage.

(ii) Real-time weather data and live imagery. The system provides up-to-date weather information and live images from the field.This enables decision-makers to obtain a detailed operational overview, facilitating more informed decision-making.

3.1. Dataset Collection and Image Pre-processing

The dataset used in this study has a primary purpose: training a computer model to distinguish images with fire from images without fire. This binary classification task is a fundamental problem in machine learning, especially for early wildfire detection (Prasad, 2022). The dataset contains 999 photographs, organised into two separate folders. This separation is important because it gives the learning algorithm clear examples of each category. The "fire images" folder has 755 pictures, all showing visible flames. Most are outdoor fire scenes, ranging from small campfires to larger brush fires. Many of these photos also show strong visible smoke, which is another important visual cue.

The "non-fire images" folder has 244 photos that serve as a comparison set. These are nature scenes such as dense forests, individual trees, grass fields, rivers, and lakes. By looking at these images, the model learns the patterns, colours, and textures of safe landscapes. The imbalance in the dataset (about 75% fire images and 25% non-fire images) is common in real-world machine learning applications.

Figure 1. Sample images from the dataset.

In the data preparation step, we processed the raw images for model input. A key part of this step was removing noise based on convolution neural network model. We also applied data augmentation. This technique expands the training dataset by creating modified versions of existing images. We rotated images, flipped them, and applied brightness adjustments. Training on more varied data helps the model generalise new and unseen images better. More importantly, data augmentation helps prevent overfitting. Overfitting is a common problem where a model becomes too specialised on the training data and performs poorly on new data. These steps are essential for building a model that is both accurate and robust.

3.2. ConvLSTM Model Architecture and Design for Forest Fire Detection

Our system uses a neural network called ConvLSTM. We chose ConvLSTM because it is good at handling image sequences. Fires grow and move over time, and ConvLSTM can detect these patterns.

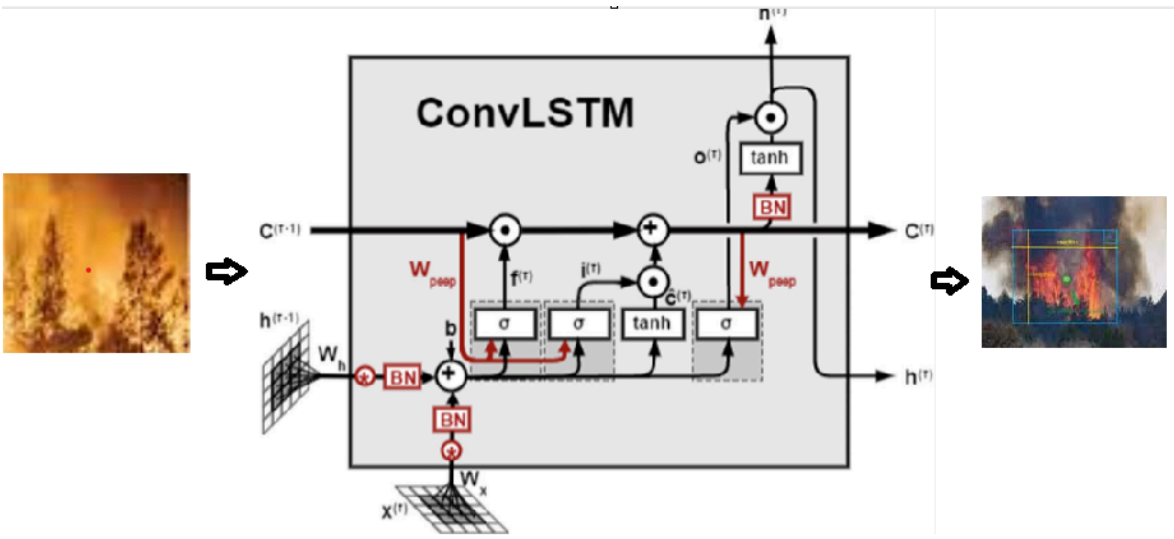

Figure 2. ConvLSTM structure illustration for the forest fire detection system.

The Conv in ConvLSTM stands for convolutional layers. The model was constructed using multiple convolutional layers to distinguish between fire and non-fire images. These layers function as filters. They scan each image and look for important visual patterns. The model is not explicitly programmed to identify specific features. Instead, it learns by itself to recognise the shapes, edges, colours, and textures that are typical of fire. Each convolutional layer looks at the image at a different level of detail. The first layers detect simple features such as edges. Later layers combine these simple patterns into more complex ones. This step-by-step process helps the model build a good understanding of both fire and non-fire images, which makes its predictions more accurate.

The LSTM part refers to Long Short-Term Memory. This component handles sequences and keeps track of previous information. After the convolutional layers process each frame, the LSTM continues the work. It does not operate on individual frames independently. Instead, it analyses features from consecutive frames and models their temporal evolution (see Figure 3).

```python
def build_conv_lstm_model(input_shape):
    model = Sequential()
    model.add(ConvLSTM2D(filters=64, kernel_size=(3, 3),
                         input_shape=input_shape,
                         padding='same', return_sequences=True))
    model.add(BatchNormalization())
    model.add(ConvLSTM2D(filters=32, kernel_size=(3, 3),
                         padding='same', return_sequences=False))
    model.add(BatchNormalization())
    model.add(Flatten())
    model.add(Dense(128, activation='relu'))
    model.add(Dropout(0.5))
    model.add(Dense(1, activation='sigmoid'))
    return model
```

Figure 3. Python implementation of the ConvLSTM model

Finally, the LSTM output is passed to several fully connected layers. The model was trained using a batch size of 8, the Adam optimiser, a learning rate of 0.001, early stopping, and a dropout rate of 0.3. These layers act as the decision maker. They take all the analysed information about space and time – what fire and smoke look like and how they change – and weigh the evidence. Their output is a probability score from 0 to 1. A score near 1 means the system believes there is a fire. A score near 0 means the scene is normal.

4. Results and Analysis

To properly evaluate algorithm performance, choosing the right metrics matters. In this study, we used accuracy, precision, recall, and F1-score to assess model performance. LSTM Model suits scenarios requiring time series analysis better. Model performance results on the test set are shown in Table 1, giving an overview of prediction capability in real-world environments – which matters a lot for evaluation.

Table 1. Performance metrics comparison for different models

	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
1	RNN	94.2 ± 0.8	93.8 ± 0.6	94.5 ± 0.7	94.1 ± 0.8
2	CNN	96.7 ± 0.5	96.3 ± 0.4	97.1 ± 0.5	96.7 ± 0.4
3	Proposed ConvLSTM	98.3 ± 0.3	98.1 ± 0.3	98.6 ± 0.2	98.2 ± 0.3

We also tested standard RNN and CNN methods on the same forest fire dataset. Results showed that our proposed ConvLSTM model performed better than both (fig 4). ConvLSTM achieved higher accuracy because it captures both spatial features (what fire looks like) and temporal patterns (how fire changes over time). In comparison, CNN works well for extracting visual features from a single image but does not handle image sequences well RNN, including LSTM, handles sequences but treats each image as a flat vector and loses spatial information. Our experiments confirmed that ConvLSTM combines the strengths of both approaches, making it more suitable for forest fire detection from image sequences.

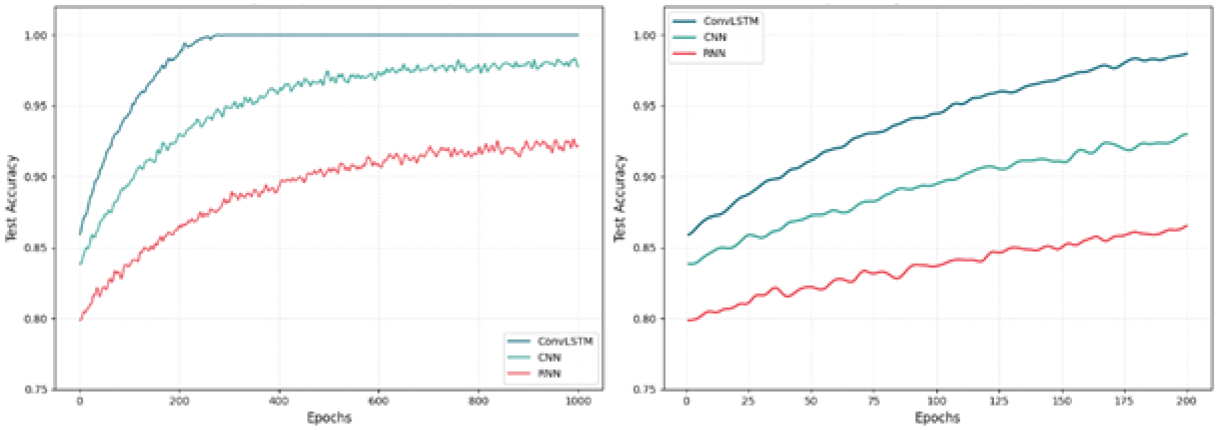

Figure 4. The result of comparison between result of RNN, CNN, ConvLSTM.

5. Conclusion

In this study, we developed a ConvLSTM model using image data to detect forest fires. The model distinguishes between fire and non-fire scenes reasonably well, achieving 98.3% accuracy on the test dataset. We used the Keras framework, which simplified building and training the model.

Our results show that the model can reliably identify visual signs of fire, making it a practical tool for early fire detection and monitoring. By providing early warnings and enabling rapid response, this system could help reduce environmental damage from wildfires. Future work should focus on improving the model's robustness, incorporating additional data such as weather conditions for more complete risk analysis, and expanding the training dataset to include more diverse fire scenarios.

6. Acknowledgment

The authors would like to thank the referee for careful reading of the paper and valuable suggestions and comments. This research was supported by Thai Nguyen University of Information and Communication Technology (No. DH2024-TN07-01).

References

Almasoud, A. S. (2023). *Intelligent deep learning enabled wild forest fire detection system*. *Computer Systems Science & Engineering, 44*(2), 1485–1498. https://doi.org/10.32604/csse.2023.025190

Abdusalomov, A. B., Islam, B. M. S., Nasimov, R., Mukhiddinov, M., & Whangbo, T. K. (2023). *An improved forest fire detection method based on the Detectron2 model and a deep learning approach*. *Sensors, 23*(3), 1512. https://doi.org/10.3390/s23031512

Ban, Y., Zhang, P., Nascetti, A., Bevington, A. R., & Wulder, M. A. (2020). *Near real-time wildfire progression monitoring with Sentinel-1 SAR time series and deep learning*. *Scientific Reports, 10*(1), 1322. https://doi.org/10.1038/s41598-019-56967-x

Dimitropoulos, K., Barmpoutis, P., & Grammalidis, N. (2017). *Higher order linear dynamical systems for smoke detection in video surveillance applications*. *IEEE Transactions on Circuits and Systems for Video Technology, 27*, 1143–1154. https://doi.org/10.1109/TCSVT.2016.2527340

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*, 38643–38660. https://doi.org/10.1007/s11042-022-13068-8

Idroes, G. M., Maulana, A., Suhendra, R., Lala, A., Karma, T., Kusumo, F., Hewindati, Y. T., & Noviandy, T. R. (2023). *TeutongNet: A fine-tuned deep learning model for improved forest fire detection*. *Leuser Journal of Environmental Studies, 1*(1), 1–8. https://doi.org/10.60084/ljes.v1i1.42 Jamshed, M. A., Theodorou, C., Kalsoom, T., Anjum, N., Abbasi, Q. H., & Ur-Rehman, M. (2022). *Intelligent computing based forecasting of deforestation using fire alerts: A deep learning approach*. *Physical Communication, 55*, 101941. https://doi.org/10.1016/j.phycom.2022.101941

Larsen, A., Hanigan, I., Reich, B. J., Qin, Y., Cope, M., Morgan, G., & Rappold, A. G. (2021). *A deep learning approach to identify smoke plumes in satellite imagery in near-real time for health risk communication*. *Journal of Exposure Science & Environmental Epidemiology, 31*(1), 170–176. https://doi.org/10.1038/s41370-020-0246-y

Li, Z., Mihaylova, L. S., Isupova, O., & Rossi, L. (2018). *Autonomous flame detection in videos with a Dirichlet process Gaussian mixture color model*. *IEEE Transactions on Industrial Informatics, 14*(3), 1146–1154. https://doi.org/10.1109/TII.2017.2768530

Mambile, C., Kaijage, S., & Leo, J. (2025). *Deep learning models for enhanced forest-fire prediction at Mount Kilimanjaro, Tanzania: Integrating satellite images, weather data and human activities data*. *Natural Hazards Research, 5*(2), 335–347. https://doi.org/10.1016/j.nhres.2024.12.001

Mei, J., & Zhang, W. (2018). *Early fire detection algorithm based on ViBe and machine learning*. *Acta Optica Sinica, 38*(7), 0710001. https://doi.org/10.3788/AOS201838.0710001

Pan, J., Ou, X., & Xu, L. (2021). *A collaborative region detection and grading framework for forest fire smoke using weakly supervised fine segmentation and lightweight Faster-RCNN*. *Forests, 12*(6), 768. https://doi.org/10.3390/f12060768

Prasad, M. S. (2022). *Forest fire images* [Data set]. Kaggle. https://www.kaggle.com/datasets/mohnishsaiprasad/forest-fire-images

Tran, D. Q., Park, M., Jeon, Y., Bak, J., & Park, S. (2022). *Forest-fire response system using deep-learning-based approaches with CCTV images and weather data*. *IEEE Access, 10*, 66061–66071. https://doi.org/10.1109/ACCESS.2022.3184707

Yun, J., & Akçakaya, M. (2025). *Generative model-based fusion for improved few-shot semantic segmentation of infrared images*. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. https://doi.org/10.1109/WACV61041.2025.00535

Zhang, J., Zhu, H., Wang, P., & Ling, X. (2021). *ATT squeeze U-net: A lightweight network for forest fire detection and recognition*. *IEEE Access, 9*, 10858–10870. https://doi.org/10.1109/ACCESS.2021.3050628