Enhancing Forest Fire Detection: Integrated CNN And LSTM with Advanced Techniques.

MSc Research Project
Data Analytics

Sebin Chembottikkal Sunny
Student ID:X21177261

School of Computing
National College of Ireland

Supervisor: Harshani Nagahamulla

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	……………………………Sebin Chembottikkal Sunny……………………………………
Student ID:	………………………………………X21177261………………………………….……………
Programme:	……………. Data Analytics……………………… **Year:** …………2023………………
Module:	………………………..… MSc Research Project……………………………………….………
Supervisor:	……………………………… Harshani Nagahamulla …………………………….………
Submission Due Date:	……………………………… 31/01/2024…………………………………………….………
Project Title:	………… Enhancing Forest Fire Detection: Integrated CNN And LSTM with Advanced Techniques………………………………………………………
Word Count:	……………5845……………… **Page Count**……………………19……………….……..

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: …………………… Sebin Chembottikkal Sunny ……………………………………..

Date: …………………………………31/01/2024……………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Forest Fire Detection: Integrated CNN And LSTM with Advanced Techniques.

Sebin Chembottikkal Sunny
X21177261

Abstract

Forest fires pose serious threats to the environment and human safety, necessitating efficient detection methods. Traditional approaches relying on sensors or human observation are often expensive and error-prone. To address this, researchers are exploring advanced computer vision techniques, particularly leveraging convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. This study compares the performance of CNNs and CNN Bi LSTM in detecting forest fires and smoke from aerial and ground views. This research used dataset of 12,631 images from Kaggle, of forest images for training and testing. The integrated approach of CNN Bi LSTM demonstrates superior performance, exhibiting over 90% accuracy in both aerial and ground views. The results indicate that the combined spatial and temporal analysis of CNN Bi LSTM surpasses the capabilities of CNN alone. This research contributes to the development of a robust forest fire detection model, emphasizing the effectiveness of combining spatial and temporal aspects in computer vision applications.

1 Introduction

In our world, due to the most destructive natural disaster, forest fire, millions of hectares of land and lives are impacted every year (Phillips, 2022). A survey conducted by World Meteorological Organization in 2023, shows that more than ten million hectares of forest land were destroyed by wildfire. This covers around 0.7% of the global forest area, which was estimated at 1.4 billion hectares in 2020 (Phillips, 2022). A forest fire can result from humans as well as from natural elements. Factors such as arson, farming activities or power through forest can be a reason behind the forest fire. Human activities were responsible 95% of wildfire between 2005 and 2009 .Natural elements such as volcanic eruptions and lightning strikes can cause forest fires as well (Belcher, 2010). Forest fires can have serious effects on the environment in ways that they can reduce biodiversity, can increase in the emission of greenhouse gases, and that will lead to an increase in air pollution which directly impacts human health (Dillis, 2022). To avoid this, it is very crucial to detect the presence of forest fire at the beginning stage or at its origin to avoid the spread of the fire into more forest land regions.

In the domain of machine learning, deep learning holds great promise for identifying forest fires due to its ability to process vast datasets using artificial neural networks (Zuo, 2022). This technology has far-reaching applications beyond forest fire detection, including image classification, text analysis, voice recognition, and numerous other areas. For forest detection, It is used to extract useful features from images, videos, and data provided by satellites, and helps to classify them into fire and non-fire images. Deep learning does have several advantages over traditional methods such as their ability to learn from both labelled and non-labelled data and they can also adapt to different environments and scenarios (Buduma, 2022.

). Despite their impressive capabilities, deep learning models still face several challenges and drawbacks. These include the high computational requirements needed to train them, the risk of overfitting or underfitting, and potential susceptibility to adversarial attacks.

Research Question:
How effectively can the combination of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) architectures be improved for object detection, particularly in the context of comparing ground view and aerial view forest imagery with respect to accuracy?

This research outlines on the optimization of object detection models through the inclusion of Convolutional Neural Networks (CNN) and CNN Bi Long Short-Term Memory (LSTM) architectures to improve the performance of object detection. It compares the matrices of both CNN and LSTM of both ground view and aerial view to find out which is suitable for ground view and aerial view. The subsequent sections will explain the literature review, Methodology, and results with discussion, providing a comprehensive analysis that emphasizes the significance of these findings in improving the detection of the forest fire.

2 Related Work

Forest fire is one of the most disastrous natural events in the world, that need proper research on detecting the fire as well as its response method. Forest fires cause damage to human lives, the economies of the countries, and the ecosystem of the world. Burning of a forest can cause the increment of greenhouse gas emissions which consist of Carbon dioxide, methane, and nitrous oxide. This will also result in global warming (Singh, 2022). With the help of Web of Science database, using bibliometric and quantitative statistical analysis techniques, the study says that there has been a growth rate of 22.45% in forest fire over the past 20 years (Li, 2023). There are different types of forest fire detection method. These methods include satellite-based, aerial, and ground-based systems. These techniques totally rely on the sensors and cameras on towers, satellites, or vehicles. By utilizing a wide range of methodologies such as artificial intelligence, machine learning algorithms, and statistical methods, it helps in providing useful insights and information on the field of forest fire prediction and prevention (Thirumal, Forest Fire Detection and Prediction–Survey).

A pioneering deep learning method for forest fire prediction and detection has been developed by using a hybrid LSTM-CNN architecture to extract spatiotemporal information using satellite imagery. This helped in contributing the demonstration of architecture accuracy, enabling real time processing and implementing real- time alerts via Google Firebase. This paper identifies improving data preprocessing, dealing with multi-sensor data, and investigating model explainability methodologies as future research directions (Gayathri, 2022). Detection of smoke from the forest can be done using CNN. It recommends using multi-convolution kernel techniques and batch normalization which helps in improving the performance of CNN architecture. This research shows that CNN outperforms the conventional image processing techniques in smoke recognition. It also proves that CNNs can effectively extract and learn elements linked to smoke and fire from photos (Sun, 2021). A cost-effective CNN architecture could be used for forest fire detection using GoogleNet architecture. This architecture is known for its reasonable computational complexity and suitability for tasks related to image classification (Muhammad, 2018).The usage of Generative Adversarial Networks (GANs) in CNN, helps in generating high-quality forest fire samples, and improves the accuracy. Histogram of Oriented Gradients (HOG) can be used to make a primary prediction of the forest fire area in the image, which can improve the efficiency of CNNs (Liu, 2020).

Forest fire danger rating classification and forecasting model can be created by combining the advantages of two machine learning models such as the Long Short – Term Memory (LSTM) network and the Random Forest (RF) model (Chen, 2023). Some of the factors that can help in predicting forest fires are landscape, vegetation, type and weather. The LSTM method can be used to study the interconnection between factors and can help in predicting forest fires for a region or a location. Compared to conventional forest fire prediction techniques, which normally provide an accuracy of between average percent, the custom model presented in this paper demonstrated a notable improvement in accuracy in predicting the incidence of forest fires (Natekar, 2021). LSTM networks can be used to capture the temporal dynamics of fire spread, demonstrating advanced results in prediction accuracy. It creates new research opportunities in this field, such as the creation of hybrid models that combine LSTM networks with other deep learning architectures (Jindal, 2020).

Fire detection can also be done on IoT-based using LSTM and CNN. The paper claims that their system can accurately detect forest fires at an early stage, thereby helping to minimize damage and loss of life. Besides, LSTM detects forest fires from sensor data. Similarly, they have used CNN to detect forest fires from satellite images (AKSHIYA, 2020). A multistage fire detection system can be created, making use of both CNNs and LSTMs. In the initial phase, a CNN is applied to extract features from video frames, while the subsequent phase involves an LSTM to learn temporal dependencies between frames. Finally, a softmax classifier is employed to determine the presence of fire within the video. This innovative system achieves higher detection accuracy compared to state-of-the-art methods and is capable of real-time fire detection, a valuable feature for practical applications (Nguyen, 2021).

Another method combines the use of a short-term memory (LSTM) model, with unmanned aerial vehicle (UAV) images to forecast the rate at which forest fires spread. This research holds importance given the increasing risk of forest fires and the necessity, for early warning systems. The paper proposes a hybrid deep learning model that combines CNNs and LSTMs to predict the forest fire spread rate from UAV images. The CNNs extract features from the images, while the LSTM models the temporal dynamics of fire spread. Additionally, the authors' model incorporates wind speed information to account for the interaction between fire and wind (Li, 2021).There is an alternative deep learning model named SmokeyNet, which integrates multiple data sources for detecting smoke from wildland fires. This process utilize camera image data and weather sensor measurements to establish a multimodal wildland fire smoke detection system. The paper also introduces the Multimodal SmokeyNet and SmokeyNet Ensemble for multimodal wildland fire smoke detection, incorporating satellite-based fire detections, weather sensor measurements, and optical camera images (Bhamra, 2023). The below table provides a comprehensive overview of literature review papers.

Table 1 – Comparison of available techniques for forest fore detection using deep learning

No	Paper	Technique	Dataset	Accuracy %	Limitation	Significance
1	(Singh, 2022)	Remote sensing Techniques	RS datasets and Landsat datasets	NA	Data Availability and Quality	Forest fire climate mitigation.
2	(Li, 2023)	Bibliometric and Quantity Statistical Analysis	Web of Science database for international forest wildfre research	NA	Data Availability and Quality	Global wildfire research evolving
3	(Thirumal, Forest Fire Detection and Prediction–Survey).	Survey	The author collected information from old research papers	NA	Limited to detection and prediction techniques	Possibilities of implementing future technologies
4	(Gayathri, 2022)	CNN Bi LSTM	Custom Data Set	92	Limited Dataset and information	Shows advances in deeplearning, forest fire management
5	(Sun, 2021)	CNN,SVM,KNN	Custom Data Set-Through Mobile camera	94.1(CNN), 80.1(SVM), 85.2(KNN)	Single dataset and Limited accuracy	Effective CNN for smoke detection
6	(Muhammad, 2018)	CNN(GoogleNet)	Custom Dataset	86(Precision), 0.98(Recall), 0.89(F-Measure)	Lacks real-world evaluation	CNNs enhance fire surveillance
7	(Liu, 2020)	GAN + Adaboost + CNN	Custom Dataset	90	Limited training data	Advancement for forest fire detection
8	(Chen, 2023)	LSTM	Custom Data	0.32	Single region application	Improved fire danger forecasting
9	(Natekar, 2021)	LSTM	MODIS dataset(Satellite Images)	94.77	Limited Dataset and information	Improved LSTM for the forest fire detection
10	(Jindal, 2020)	LSTM	Custom Dataset	82	Limited training data	Improved LSTM for the forest fire detection
11	(AKSHIYA, 2020)	SVM, LSTM-KNN	Forest Fires in Portugal dataset from the UCI Machine Learning Repository	92.4(SVM),95.04(LSTM-KNN)	Single region application	Improved fire danger forecasting
12	(Nguyen, 2021)	CNN, Bi LSTM	Custom Dataset	96.20%	Limited evaluation	Improved forest fire image and video detection
13	(Li, 2021)	LSTM	Custom data created by other researcher Pinus sylvestris var. mongolica	NA	Limited training data	Improved forest fire prediction
14	(Bhamra, 2023)	SmokeyNet	FIgLib library	79.64	Limited evaluation	Early wildfire notification system

3 Research Methodology

In this section, it outlines how data will be collected, organized, and analysed. It includes details on data collection methods, image categorization techniques, data preprocessing steps, and data augmentation strategies. The following sections explain the works in detail.

3.1 Data Collection

The dataset, "Forest Fire," was sourced from Kaggle, a publicly available platform. Comprising 12,631 images, it encompasses classes such as fire, non-fire, and smoke. Specifically, 3,109 images were selected from the forest category for both training and testing purposes. The dataset's availability on Kaggle ensures transparency and accessibility.

Below are some sample picture that was taken from the data set.

Figure 1-Forest fire (Ground)

Figure 2 - Forest Fire(Arial)

Figure 3-Forest Smoke (Arial)

Figure 4 -Forest smoke(Ground)

3.2 Methodology

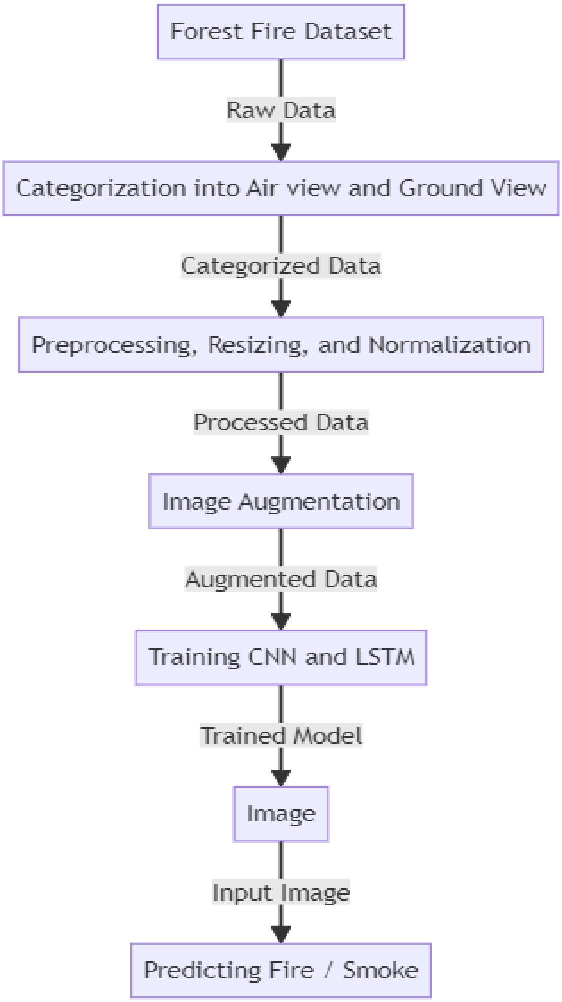

Figure 5- Overview

The Figure 5 the workflow for handling a Forest Fire Dataset to predict the presence of fire or smoke in images. Initially, the raw dataset is categorized into "Air view" and "Ground view" images to capture different perspectives of the incident. The categorized data undergoes preprocessing, including resizing and normalization, to ensure uniformity and standardize pixel values. Subsequently, the processed data is augmented through techniques like rotation and flipping, enhancing the dataset's diversity. The augmented dataset is then employed to train a model that combines Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architectures. Once trained, the model serves to predict the presence of fire or smoke in new images. The final step involves inputting an image into the trained model, which outputs predictions based on learned patterns. This systematic approach ensures the effective utilization of the dataset, comprehensive preprocessing, and the integration of advanced neural network architectures to yield a robust predictive model for identifying fire or smoke in diverse forest fire scenarios.

3.3 Image Categorization

In this research project, the focus is obtaining the optimal accuracy between Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM) in the context of image classification, specifically distinguishing between aerial and ground views.

In the implementation process, dedicated folders for fire and non-fire images were established separately for both aerial and ground views within the Convolutional Neural Network (CNN). This organizational structure aligns with CNNs, renowned for their proficiency in image-based tasks. Concurrently, a parallel organizational scheme was applied to the Long Short-Term Memory (LSTM) model, where distinct fire and non-fire folders were filled with relevant images for both aerial and ground perspectives. This deliberate approach underscores the project's commitment to systematically evaluate and compare the performance of these influential neural network architectures across different views. Furthermore, separate folders were created for smoke and non-smoke images for both aerial and ground views. This additional categorization ensures a thorough evaluation of the models' capabilities in recognizing not only fire and non-fire patterns, but also distinct characteristics associated with smoke.

This comprehensive approach ensures a robust evaluation of the models' capabilities in recognizing critical features associated with fire incidents. By separating the dataset into relevant categories and deploying both CNN and LSTM architectures, the project aims to provide valuable insights into the efficacy of these models for aerial and ground-view image classification tasks, particularly in the challenging context of identifying fire, non-fire patterns and smoke pattern.

3.4 Data Preprocessing

In image preprocessing, key steps are undertaken to ensure data preparation for subsequent machine learning tasks. The initial phase involves utilizing the cv2.cvtColor function, which is used for colour space conversion. Specifically, the transformation shifts the image from the default BGR format, commonly used by OpenCV, to the widely adopted RGB format. This conversion is crucial to align with the preferences of numerous image-processing libraries and frameworks, ensuring compatibility and consistency across diverse platforms. Another, preprocessing step is image resizing, executed through the cv2.resize function. This operation standardizes the image dimensions to a uniform 150 by 150 pixels. Beyond providing a consistent input size for machine learning models, resizing significantly contributes to reducing computational complexity and memory usage during subsequent image processing tasks. It also enhances the efficiency of model training and inference.The significance of label encoding is also emphasized in the report.

Label encoding proves essential for transforming categorical labels into numerical representations, promoting compatibility with various machine-learning algorithms that operate more effectively with numerical inputs. Array conversion, played an important role with the help of numpy.array function. This function efficiently converts lists into numpy arrays, offering advantages such as faster computation, support for vectorized operations, and versatile indexing capabilities. NumPy arrays emerge as a preferred choice for handling multidimensional data in diverse applications. Following array conversion, pixel normalization as a key preprocessing step. This process involves scaling each pixel value in the image array by 255, the maximum possible value for an 8-bit image. The scaling to the range of 0 to 1 mitigates the impact of illumination differences, enhances numerical stability, and contributes to the accelerated convergence of learning algorithms during training.

The dataset is divided into two parts: one for training and the other for testing a machine-learning model designed for a CNN. The dataset, which includes features and their corresponding target values, is randomly divided. About 20% of the data is set aside for testing, and the remaining 80% is used for training. Similar divisions of the dataset into training and testing sets are made for the LSTM procedure to facilitate LSTM model training. This division separates the combined input data and corresponding labels, with 30% allocated for testing and 70% for training. Additionally, a specific random seed is included to ensure consistent results when the code is executed multiple times. The resulting splits are stored in variables for further use in training and evaluating the LSTM model.

3.5 Data augmentation

Data augmentation refers to the process of artificially expanding a dataset by applying various transformations to the existing data. This technique is commonly employed in machine learning, particularly in computer vision tasks, to enhance model generalization and improve performance. DA refers to any method that artificially inflates the original training set with label-preserving transformations and can be represented as the mapping.

$$\phi : \mathcal{S} \mapsto \mathcal{T}$$

Where, S is the original training set and is the augmented set of S. The methods of traditional data augmentation can be divided into 2 and they are Geometric and Photometric methods. (Taylor, n.d.) The below picture shows the sample of Data augmentation used in this research project.

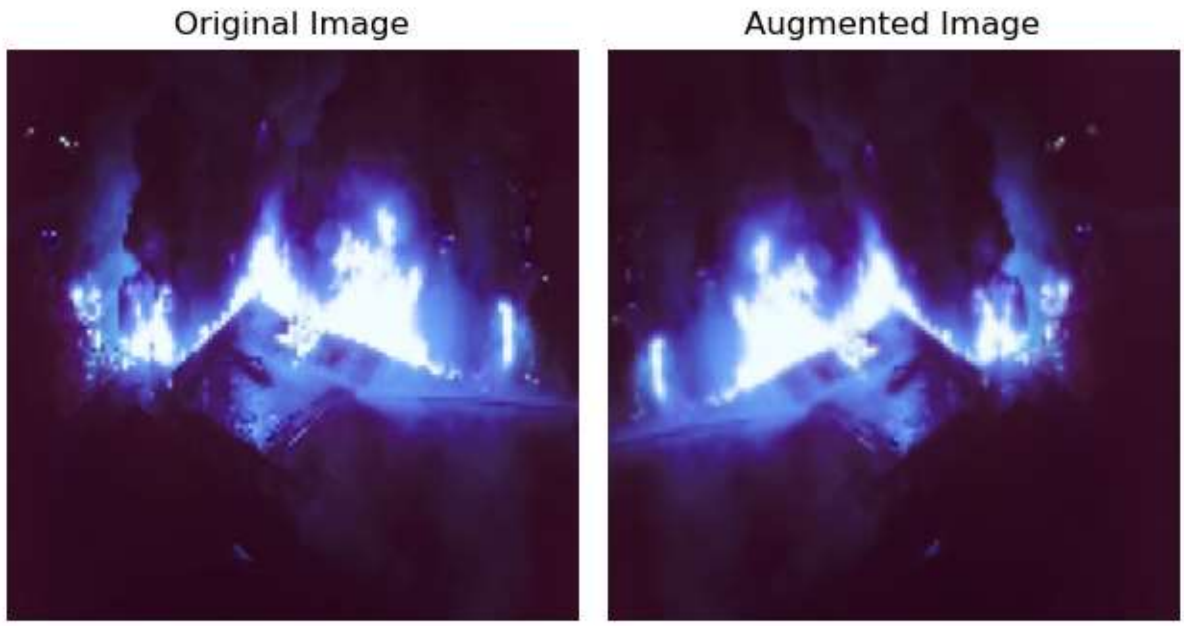

Figure 6- Data Augmentation sample

This project, various geometric transformations have been implemented to manipulate image geometry effectively. These transformations involve remapping individual pixel values to new positions while preserving the inherent shape of the represented class but introducing changes in position and orientation. Within the geometric method, techniques such as flipping, rotation, and cropping have been applied to augment the dataset, aiming to enhance its diversity and bolster the robustness of the model. A fundamental geometric augmentation method employed is horizontal flipping, chosen for its realism and computational efficiency. This technique entails the reversal of images along the horizontal axis, contributing to dataset variation. Additionally, cropping has been utilized to randomly extract portions of images, resizing them to standardized dimensions for input into machine learning models. The integration of these geometric transformations, including horizontal flipping and cropping, plays a pivotal role in enriching the dataset and improving the model's adaptability to different image orientations and positions. These methods contribute significantly to the overall performance and robustness of the model in handling diverse geometric variations in the input data. It is crucial to note that cropping may lead to samples with altered labels, requiring careful consideration to address potential label correctness issues. (Taylor, n.d.) The rotation method has been utilized during the augmentation process for both random rotations and rotations by specific angles. This has helped in reinforcing the model's robustness, facilitating effective performance across a range of scenarios. Furthermore, the research incorporates data augmentation strategies, including zoom_range for image zooming, width_shift_range for horizontal shifts, and height_shift_range for vertical shifts. The introduction of shear_range facilitates shear-based transformations. To address missing pixels, the fill_mode parameter is deliberately set to "nearest," ensuring the gaps are filled with the closest available pixel. Together, these augmentation techniques aim to cultivate a more diverse and robust dataset, with the overarching goal of enhancing the overall performance of the machine learning model within the scope of this research project.

Throughout my research project, different photometric transformations have been used to manipulate the RGB channels of images, thereby altering lighting and colour. These transformations are essential in enhancing the model's adaptability to diverse image conditions. One key technique employed is Brightness Adjustment, where pixel values are scaled to regulate image brightness. This adaptation proves valuable in preparing the model for varied lighting scenarios (Alomar, 2023). Additionally, Contrast Adjustment is implemented to modify the distinction between bright and dark regions within images (Taylor, n.d.). This helps the model effectively handle a wide range of image contrasts, contributing to its overall robustness.

4 Design Specification

4.1 Forest fire and Smoke Detection

For detecting forest fire and smoke, two approaches such as Convolutional Neural Networks (CNN) and CNN Bi Long Short-Term Memory (LSTM) architectures.

4.1.1 CNN

Convolution structures in data allow for feature extraction by a type of feedforward neural network called CNN. Manual feature extraction is not necessary for CNN. An artificial neuron is equivalent to a real neuron; CNN kernels are several receptors that can react to diverse aspects. Activation functions mimic the behaviour of neural electric signals that can only pass to the next neuron when they surpass a particular threshold. CNN possesses several advantages, making it one of the most representative algorithms in the deep learning field. Firstly, it employs local connections, where each neuron is connected to only a small number of neurons in the previous layer. This strategy effectively reduces parameters and accelerates convergence. Secondly, CNN utilizes weight sharing, allowing a group of connections to share the same weights. This not only promotes parameter reduction but also enhances efficiency. Thirdly, the down sampling dimension reduction technique in CNN, achieved through pooling layers, leverages the principle of local correlation in images. This not only reduces the amount of data, retaining crucial information, but also trims down the number of parameters by eliminating trivial features (Li, 2021). The Figure 6 shows the overall structure of a basic CNN.

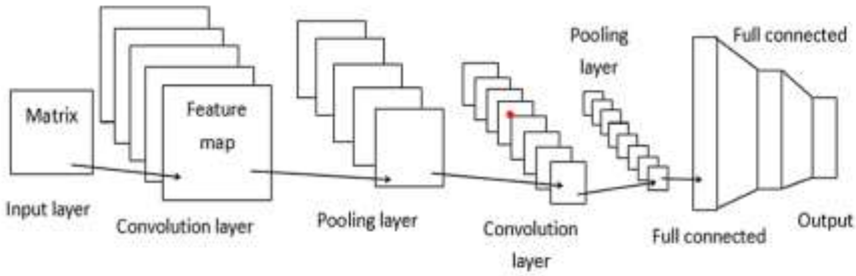

Fig 6 - The general Architecture of CNN (Sun, 2021)

The convolution layer performs feature extraction on the input image information and conducts local convolution operations on the input signal through sliding convolution kernels. The network parameters are reduced via local connections and weight-sharing mechanisms, thereby reducing the overall network complexity. The general feature extraction formula of the convolution layer is shown as follows:

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} \times k_{ij}^l + b_j^l\right)$$

where l is the number of convolution layers, $f(.)$ is the activation function, k is the convolution kernel, and b is the offset value (Sun, 2021)

Pooling layers are used to perform the reduction of dimensionality operation. The dimensionality reduction procedure is the pooling layer's primary function. In order to lower the experimental parameters, this reduces the array's size while maintaining its original characteristics. By incorporating a pooling layer, a system can become more resilient by increasing computation speed, improving the network's ability to adjust to changes in picture size, and successfully preventing overfitting issues (Li, 2021).

4.1.2 CNN

Long Short-Term Memory (LSTM) represents an enhancement over Recurrent Neural Networks (RNNs). Unlike conventional RNN units, LSTM introduces memory blocks to address the vanishing and exploding gradient problem (Islam, 2020). Additionally, it incorporates a cell state to preserve long-term states, marking a significant departure from RNNs. The distinctive feature of an LSTM network is its ability to retain and establish connections between past information and present data (Islam, 2020).

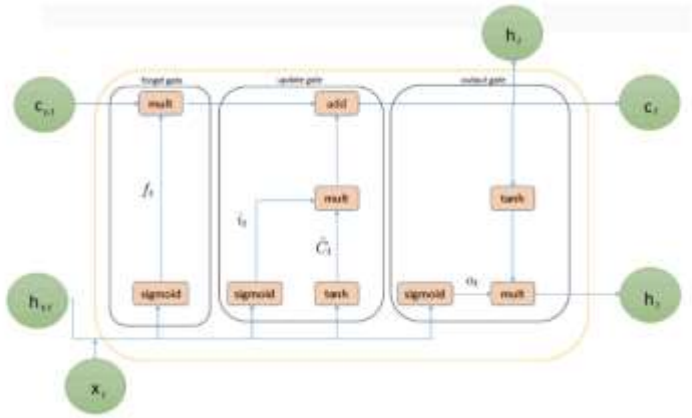

Fig 7. Basic Internal Structure of LSTM((Islam, 2020).

Basically, LSTM is combined with three gates, such as an input gate, a "forget" gate, and an output gate, where x_t refers to the current input; c_t and c_{t-1} represent the previous and new cell states, respectively; and h_t and h_{t-1} are the current and previous outputs, respectively. The internal structure of LSTM is shown in Figure 7.

The CNN BiLSTM architecture seamlessly combines Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory networks (BiLSTMs). In this hybrid structure, input data undergoes feature extraction through a CNN layer with filters and non-linear activation. The resulting feature maps are transformed into a sequence of vectors, feeding into a BiLSTM layer. Comprising forward and backward LSTM sub-layers, this bidirectional processing captures contextual information from past and future time steps. The concatenated output from the BiLSTM layer, incorporating hidden states, proves versatile for tasks like classification or regression. This architecture's adaptability makes it a potent solution for diverse projects requiring nuanced sequential data analysis.

5 Implementation

5.1 Implementation using CNN.

In this research project, a Convolutional Neural Network (CNN) architecture has been meticulously designed for the critical tasks of forest fire and smoke detection. The model unfolds sequentially, commencing with an initial convolutional layer featuring 32 filters, a (3, 3) kernel size, and a rectified linear unit (ReLU) activation function. Subsequent max-pooling layers effectively diminish spatial dimensions, paving the way for additional convolutional layers with 64 and 128 filters, progressively capturing intricate features representative of potential fire occurrences.

Following this, the architecture incorporates dense layers for processing the flattened features, culminating in a final layer with 512 units and ReLU activation. This strategic configuration ensures the extraction of high-level abstractions necessary for making informed binary classifications. The sigmoid activation function in the output layer serves to provide a probability-like output, facilitating the distinction between the presence or absence of fire in input images.

It's worth mentioning that one of the most impressive aspects of this architecture is its ability to be easily adapted for other applications, such as detecting smoke from forest fires. This extension underscores the model's versatility in addressing the nuanced challenges posed by both forest fire and smoke classification tasks. By leveraging a combination of convolutional and dense layers, the model proves to be a robust tool for image analysis, capable of discerning distinctive features associated with forest fire and smoke scenarios. This research contributes not only to the field of computer vision but also holds tangible implications for real-world applications in early fire detection systems. The below sample shows the detection of Forest fire and smoke using CNN in ground view level and in arial view level.

In CNN model, utilized the Adam optimizer, binary crossentropy loss function, and accuracy as the evaluation metric. Adam adapts learning rates during training, making it suitable for a variety of tasks. Binary crossentropy is employed as the loss function, ideal for binary classification tasks, measuring the difference between predicted probabilities and true binary labels. The accuracy metric evaluates the model's performance during training, providing insight into its ability to distinguish between images with and without fire or smoke.

5.2 Implementation using CNN Bi LSTM.

In this study, a deep learning model was employed for the classification of image sequences, combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architectures. The input data consisted of sequences with variable lengths, each containing RGB images of size 128x128 pixels. The CNN component, composed of TimeDistributed convolutional layers followed by max-pooling layers, extracted spatial features from each image within the sequence.

A TimeDistributed flatten layer was used to prepare the data for the subsequent LSTM layers. Two LSTM layers were implemented; the first with 128 units and the second with 64 units, introducing temporal dependencies and capturing sequential patterns. Dropout layers were inserted to mitigate overfitting. The model concluded with a dense layer utilizing a sigmoid activation function for binary classification.

The Adam optimizer with a learning rate of 0.001 was employed during model compilation, with binary crossentropy serving as the loss function and accuracy as the evaluation metric. The training process includes K-fold cross-validation with five folds, ensuring robust evaluation. Each fold involved training the model for one epoch on the designated training set, utilizing early stopping with a patience of 3 epochs to prevent overfitting. This approach not only leveraged the power of CNNs in spatial feature extraction but also harnessed LSTMs to effectively model temporal relationships within sequential data.

The achieved results were promising, demonstrating the efficacy of the combined CNN-LSTM architecture in the context of the classification task. The proposed model was able to predict the presence of fire, or smoke when the given image contains fire or smoke.

6 Result and Discussion

The result table provides accuracy percentages for two different types of models, CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory), in the context of predicting and classifying forest fire and forest smoke based on two perspectives: "Area View" and "Arial View."

Forest Fire		
	CNN	LSTM
Area View	92.43	87.72
Arial View	97.04	99.56

In predicting forest fires, the models exhibit varying degrees of accuracy depending on the type of view. For the "Area View," the CNN model performs slightly better than the LSTM model. However, for the "Arial View," both models show high accuracy, with the LSTM model slightly outperforming the CNN model.

Forest Smoke		
	CNN	LSTM
Area View	86.79	87.50
Arial View	99.56	100

Similarly, in the case of predicting forest smoke, the models show varying performance across different views. For the "Area View," both models have relatively similar accuracy, with the LSTM model having a slightly higher accuracy. In contrast, for the "Arial View," the CNN model achieves a accuracy of 99.56 %, while the LSTM model also performs exceptionally well with 100 % accuracy.

6.1 Confusion Matrix

6.1.1 CNN

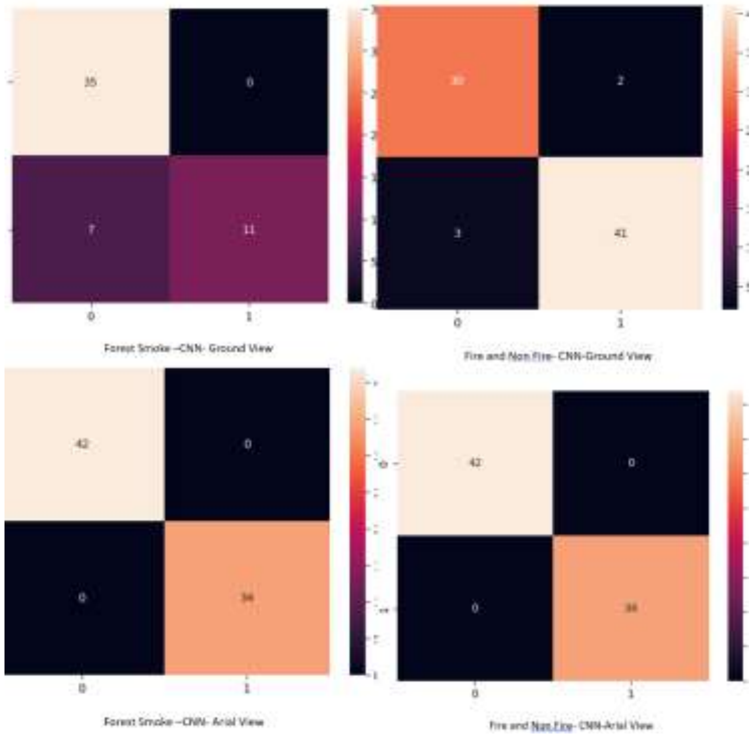

Figure 8 – Confusion Matrixes of CNN

.

The figure () The confusion matrices for ground and aerial views in forest smoke and fire detection showcase the CNN model's strong performance. Ground view smoke detection yielded 35 true positives, 0 false positives, 7 false negatives, and 11 true negatives. In ground view fire identification, there were 30 true positives, 2 false positives, 3 false negatives, and 41 true negatives. Aerial view smoke detection achieved 42 true positives, 0 false positives, 0 false negatives, and 34 true negatives. Aerial view fire detection showed 42 true positives, 0 false positives, 0 false negatives, and 34 true negatives. The model demonstrates high accuracy and reliability in both scenarios.

6.1.2 LSTM

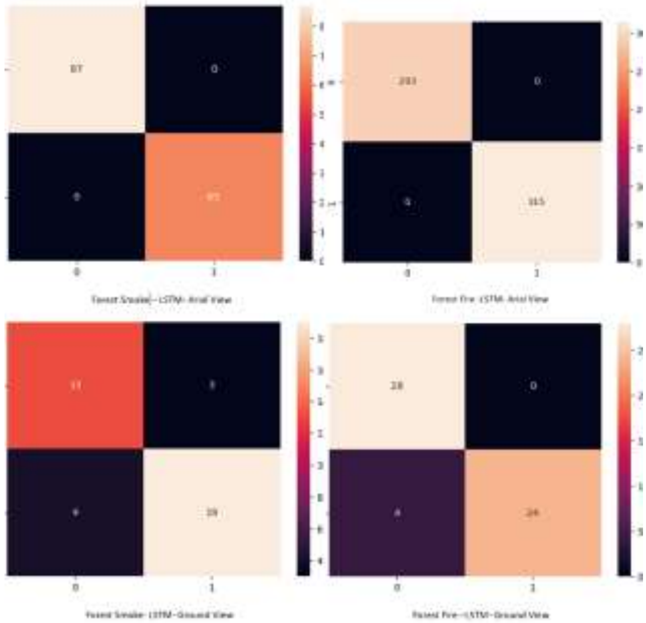

Figure 9 – Confusion matrix of LSTM

In aerial and ground view forest smoke and fire detection using LSTM, the models demonstrate impressive precision and reliability. Aerial smoke detection achieves 87 true positives with zero false positives or negatives, showcasing exceptional performance. Similarly, aerial fire detection achieves 293 true positives with zero false positives or negatives, demonstrating outstanding precision. However, ground view smoke detection exhibits 13 true positives, 3 false positives, and 4 false negatives, indicating sensitivity to smoke with room for improved precision. Ground view fire detection achieves 28 true positives, 3 false positives, and 4 false negatives, showcasing good sensitivity but with opportunities for further minimizing false positives.

7 Conclusion

The study examines the combined effect of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) architectures for forest fire detection, particularly in ground and aerial views, reveals LSTM's superiority over CNN. This superior performance across perspectives underscores LSTM's aptitude for capturing temporal dependencies crucial for recognizing evolving patterns like forest fires. The model's success in both ground and aerial scenarios signifies its versatility, making it applicable to diverse forest surveillance applications.

Future work should delve into model optimization through hyperparameter tuning and alternative architectures, dataset augmentation for improved adaptability, real-time implementation challenges, multimodal integration for a holistic understanding, and enhancing explain ability for increased trust. These avenues pave the way for a more robust and adaptable forest fire detection system.

The commercialization potential is substantial, offering opportunities for forest surveillance system enhancement, collaboration with emergency response agencies, provision of environmental monitoring services, and tailored solutions for diverse clients. This research not only advances the understanding of effective object detection in forest imagery but also lays the foundation for practical applications, contributing to the development of advanced environmental monitoring systems and emergency response solutions.

References

Phillips, C.A., Rogers, B.M., Elder, M., Cooperdock, S., Moubarak, M., Randerson, J.T. and Frumhoff, P.C., 2022. Escalating carbon emissions from North American boreal forest wildfires and the climate mitigation potential of fire management. *Science Advances*, *8*(17), p.eabl7161.

Belcher, C.M., Yearsley, J.M., Hadden, R.M., McElwain, J.C. and Rein, G., 2010. Baseline intrinsic flammability of Earth’s ecosystems estimated from paleoatmospheric oxygen over the past 350 million years. *Proceedings of the National Academy of Sciences*, *107*(52), pp.22448-22453.

Dillis, C., Butsic, V., Moanga, D., Parker-Shames, P., Wartenberg, A. and Grantham, T.E., 2022. The threat of wildfire is unique to cannabis among agricultural sectors in California. *Ecosphere*, *13*(9), p.e4205.

Zuo, C., Qian, J., Feng, S., Yin, W., Li, Y., Fan, P., Han, J., Qian, K. and Chen, Q., 2022. Deep learning in optical metrology: a review. *Light: Science & Applications*, *11*(1), p.39

Buduma, N., Buduma, N. and Papa, J., 2022. *Fundamentals of deep learning*. " O'Reilly Media, Inc.".

Singh, S., 2022. Forest fire emissions: A contribution to global climate change. *Frontiers in Forests and Global Change*, *5*, p.925480.]

Li, T., Cui, L., Liu, L., Chen, Y., Liu, H., Song, X. and Xu, Z., 2023. Advances in the study of global forest wildfires. *Journal of Soils and Sediments*, pp.1-15.

Thirumal, P.C. and Agnus, L.S.D., 2022, July. Forest Fire Detection and Prediction–Survey. In *2022 International Conference on Inventive Computation Technologies (ICICT)* (pp. 1295-1302). IEEE.

Gayathri, S., Karthi, P.A. and Sunil, S., 2022. Prediction and Detection of Forest Fires based on Deep Learning Approach. *Journal of Pharmaceutical Negative Results*, pp.429-433.

Sun, X., Sun, L. and Huang, Y., 2021. Forest fire smoke recognition based on convolutional neural network. *Journal of Forestry Research*, *32*(5), pp.1921-1927.

Muhammad, K., Ahmad, J., Mehmood, I., Rho, S. and Baik, S.W., 2018. Convolutional neural networks based fire detection in surveillance videos. *Ieee Access*, *6*, pp.18174-18183.

Liu, Z., Zhang, K., Wang, C. and Huang, S., 2020. Research on the identification method for the forest fire based on deep learning. *Optik*, *223*, p.165491.

Chen, Z., Zhang, C., Li, W., Gao, L., Liu, L., Fang, L. and Zhang, C., 2023. Fire danger forecasting using machine learning-based models and meteorological observation: a case study in Northeastern China. *Multimedia Tools and Applications*, pp.1-21.

Natekar, S., Patil, S., Nair, A. and Roychowdhury, S., 2021, May. Forest fire prediction using LSTM. In *2021 2nd International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.

Jindal, R., Kunwar, A.K., Kaur, A. and Jakhar, B.S., 2020, July. Predicting the dynamics of forest fire spread from satellite imaging using deep learning. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 344-350). IEEE.

AKSHIYA, R., DEVIBALAN, M.K., NIVETHA, J. and VINOTHINI, R., 2020. Iot Based Forest Fire Detection System. *Ilkogretim Online*, *19*(4), pp.6820-6825.

Nguyen, M.D., Vu, H.N., Pham, D.C., Choi, B. and Ro, S., 2021. Multistage real-time fire detection using convolutional neural networks and long short-term memory networks. *IEEE Access*, *9*, pp.146667-146679.

Li, X., Gao, H., Zhang, M., Zhang, S., Gao, Z., Liu, J., Sun, S., Hu, T. and Sun, L., 2021. Prediction of Forest fire spread rate using UAV images and an LSTM model considering the interaction between fire and wind. *Remote Sensing*, *13*(21), p.4325.

Bhamra, J.K., Anantha Ramaprasad, S., Baldota, S., Luna, S., Zen, E., Ramachandra, R., Kim, H., Schmidt, C., Arends, C., Block, J. and Perez, I., 2023. Multimodal Wildland Fire Smoke Detection. *Remote Sensing*, *15*(11), p.2790

Taylor, L. and Nitschke, G., 2018, November. Improving deep learning with generic data augmentation. In *2018 IEEE symposium series on computational intelligence (SSCI)* (pp. 1542-1547). IEEE.

Alomar, K., Aysel, H.I. and Cai, X., 2023. Data augmentation in classification and segmentation: A survey and new strategies. *Journal of Imaging*, *9*(2), p.46.

Li, Z., Liu, F., Yang, W., Peng, S. and Zhou, J., 2021. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*.

Sun, X., Sun, L. and Huang, Y., 2021. Forest fire smoke recognition based on convolutional neural network. *Journal of Forestry Research*, *32*(5), pp.1921-1927.

Islam, M.Z., Islam, M.M. and Asraf, A., 2020. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in medicine unlocked*, *20*, p.100412.