Prognostic Vision: A Spatiotemporal CNN-LSTM Framework for Forecasting Crop Disease Progression

Shashank Sahu
CSE Department
AKGEC
Ghaziabad, India
shashanksahu87070@gmail.com

Sunny Teotia
CSE Department
AKGEC
Ghaziabad, India
sunnyteotia2005@gmail.com

Ritwik Goswami
CSE Department
AKGEC
Ghaziabad, India
ritwikgoswami04@gmail.com

Yash Jain
CSE Department
AKGEC
Ghaziabad, India
yash2210203@akgec.ac.in

Santosh Kumar Upadhyay
CSE Department
AKGEC
Ghaziabad, India
upadhyaysantosh@akgec.ac.in

***Abstract*—Deep learning applications in plant pathology have largely concentrated on static image classification, a methodology that captures a single moment and disregards the temporal evolution of disease. This static approach inherently limits early-stage diagnosis and precise severity assessment, particularly under the fluctuating conditions of in-field agricultural settings. This paper presents a spatiotemporal framework that synergizes a Convolutional Neural Network (CNN) for spatial feature encoding with a Long Short-Term Memory (LSTM) network for analyzing temporal sequences. The resulting hybrid architecture is engineered to process sequential imagery, thereby modeling the symptomatic progression of diseases over time. To overcome the critical scarcity of longitudinal crop health datasets, we introduce a technique for generating synthetic temporal sequences from static images by simulating pathogenic development via alpha blending. Our framework leverages a pre-trained InceptionV3 model to transform individual frames into robust feature vectors, which are subsequently processed by an LSTM to discern the temporal dependencies characteristic of disease development. Experimental validation on a public rice leaf disease dataset reveals that the proposed CNN-LSTM model markedly surpasses a state-of-the-art static CNN baseline across accuracy, precision, recall, and F1-score. By integrating temporal dynamics, our framework provides a more resilient and accurate method for automated plant pathology, advancing the potential for predictive disease monitoring systems in precision agriculture.**

***Index Terms*—Spatiotemporal Analysis, Deep Learning, CNN, LSTM, Plant Disease Detection, Precision Agriculture, Computer Vision**

I. Introduction

Global food security is intrinsically reliant on agricultural productivity, which remains under persistent threat from plant diseases. These pathogens are responsible for significant economic losses, diminishing both the quality and quantity of crop yields [1]–[3]. Historically, disease identification has depended on manual inspection by agricultural experts. Such methods are labor-intensive, time-consuming, and subjective, rendering them unsuitable for large-scale monitoring and often too delayed for effective early intervention [4]–[6]. In recent years, automated systems built on deep learning (DL) and computer vision have presented a transformative alternative, promising rapid, precise, and scalable disease diagnosis [7]–[9].

The prevailing approach in this domain has been static image classification. Leading methodologies utilize Convolutional Neural Networks (CNNs) to classify plant diseases from single photographs of leaves [10], [11]. These models have achieved exceptional success, with reported accuracies often exceeding 98% on benchmark datasets like PlantVillage, which contain images captured in controlled laboratory settings [12], [13].

However, this success within controlled environments obscures a critical weakness: plant pathology is an inherently dynamic process that evolves over time and across space [14]–[16]. Static models are fundamentally incapable of capturing this temporal dimension. This spatiotemporal gap leads to significant failures when models are deployed in real-world agricultural contexts. The performance decline is multifaceted:

- **The "Lab-to-Field" Performance Gap:** Models trained on sanitized, uniform images struggle to generalize to the complex and unpredictable conditions of a field. This "domain shift"—arising from variations in lighting, background clutter, camera perspectives, and crop genotypes—can cause model accuracy to drop from over 90% in laboratory tests to under 50% in field applications [17]–[19].
- **Inability to Model Disease Progression:** A single image offers only a momentary snapshot. It cannot differentiate between nascent and advanced stages of infection, evaluate the rate of development, or forecast future severity. This temporal context is essential for making informed

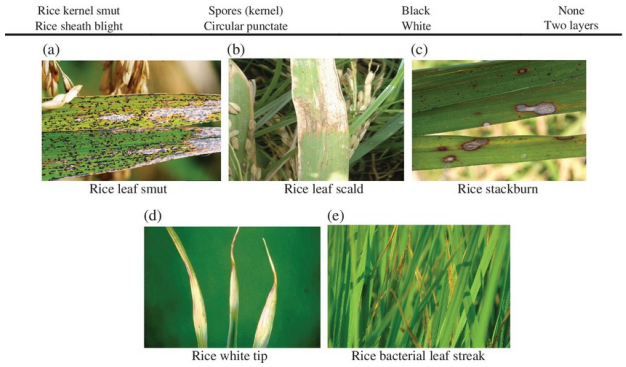


Fig. 1: An example of rice leaf disease symptoms utilized in detection models.

and timely management decisions [16], [20], [21]. A static model perceives an image of early fungal spots and one of a fully necrotic leaf as two distinct classification problems, failing to connect them as points along a continuous pathological trajectory.

- **Vulnerability to Transient Conditions:** Static models can be easily confused by temporary environmental factors, such as shadows or moisture, misidentifying them as persistent disease symptoms. A temporal model, by contrast, can learn to disregard these as short-lived artifacts rather than genuine indicators of disease [3], [17].

This research advocates for a paradigm shift from reactive, static classification to proactive, dynamic analysis. To achieve this, our contributions are as follows:

1) We present a critical analysis of the static classification approach, detailing its inherent constraints for dependable, real-world agricultural use.
2) We introduce a novel, integrated CNN-LSTM framework engineered to learn and model the spatiotemporal dynamics of plant disease from image sequences.
3) To mitigate the widespread absence of temporal datasets, we propose a technique for generating synthetic disease progression sequences from existing static image libraries.
4) We deliver a comprehensive experimental validation of our framework, proving its superior performance over a leading static CNN model and confirming the value of incorporating temporal data.

The remainder of this paper is organized as follows: Section II reviews related literature on static and temporal modeling. Section III outlines the architecture of our proposed CNN-LSTM framework and the synthetic data generation method. Section IV describes the experimental setup, presents the results, and provides a comparative analysis. Section V discusses the broader implications of our findings, and Section VI concludes the paper.

II. Related Work

A. Static Image-Based Disease Classification with CNNs

The use of deep learning in plant pathology represented a major advancement over traditional machine learning techniques, which depended on manually engineered features like color histograms and texture descriptors [5], [22]. CNNs automated this feature extraction process, learning complex hierarchical representations directly from pixel data, which proved to be significantly more effective [23]. The practice of transfer learning, where a model pre-trained on a vast dataset such as ImageNet is fine-tuned for a specialized task, has become a standard methodology, facilitating high performance even with smaller, domain-specific datasets [2], [24].

Several pre-trained architectures have been instrumental in this area:

- **VGG Architectures (VGG16/VGG19):** Characterized by their straightforward yet deep structure of stacked 3x3 convolutional layers, VGG models are potent feature extractors [25]. However, their extensive parameter count makes them computationally demanding [26]–[28]. Studies employing VGG16 on the PlantVillage dataset have reported accuracies between 97.87% and 98.40% [29].
- **ResNet Architectures (ResNet50):** ResNet's primary innovation is the use of residual "skip" connections, which facilitate gradient flow through extremely deep networks and help prevent the vanishing gradient problem [12], [13], [28]. This allows for the construction of deeper and more powerful models. On the PlantVillage dataset, ResNet50 has frequently achieved accuracies above 99% [13], [30], [31].
- **Inception Architectures (InceptionV3):** These networks utilize "inception modules" that apply convolutional filters of various sizes (e.g., 1x1, 3x3, 5x5) concurrently and merge their outputs. This design enables the model to efficiently capture features at multiple scales [6], [32]. The baseline study for this paper used InceptionV3 to achieve 100% accuracy on a specific rice disease dataset [1], while other research on PlantVillage reports accuracies in the 98-99% range [31], [33].

Table I provides a summary of the performance of these leading static models. While the reported metrics are high, they typically reflect performance in controlled scenarios and do not capture the difficulties of real-world application.

TABLE I: Performance of State-of-the-Art Static CNN Models on Plant Disease Datasets

Model	Dataset	Accuracy (%)	Citation
InceptionV3	Rice Leaf Dataset	100.00	[1]
VGG16	PlantVillage	98.57	[29]
VGG16	PlantVillage	98.40	[29]
ResNet50	PlantVillage	99.70	[30]
ResNet50	PlantVillage	99.23	[13]
ResNet101	PlantVillage	99.79	[31]
InceptionV3	PlantVillage	98.70	[33]
InceptionV3	PlantVillage	99.41	[31]

B. The Generalization Challenge: From Laboratory to Field

A growing body of research now underscores the performance gap between models trained in controlled settings and those deployed in agricultural fields. This "lab-to-field" discrepancy is a major obstacle to practical implementation [18], [19]. The central problem is domain shift, where the statistical properties of field-captured data differ markedly from the laboratory data used for training [17]. Field images are affected by complex environmental factors, including variable lighting, shadows, rain, occlusions from other leaves, and diverse backgrounds, all of which introduce noise that can confuse a model trained on pristine, uniform images [3], [34]. Moreover, the collection and annotation of large, varied, field-based datasets represent a significant bottleneck, often leading to small or imbalanced datasets that are susceptible to overfitting [19], [35]. This data scarcity motivates the development of methods that can adapt existing static datasets for more advanced temporal analyses.

C. Temporal Modeling in Agriculture and Other Domains

Recurrent Neural Networks (RNNs) are architected to handle sequential data by preserving a hidden state that conveys information across timesteps. However, conventional RNNs often fail to learn long-range dependencies due to the vanishing gradient problem [36], [37]. Long Short-Term Memory (LSTM) networks address this issue with a sophisticated cell structure containing gates—input, output, and forget gates. These mechanisms permit the LSTM cell to selectively retain, discard, and pass on information from its memory, enabling it to capture dependencies over extended sequences [20], [36]. LSTMs are widely used in agriculture for time-series forecasting, such as predicting crop yields, soil moisture levels, and pest infestations from historical weather and satellite data [**?**], [**?**], [37], [38].

D. Hybrid CNN-LSTM Architectures for Spatiotemporal Analysis

The fusion of CNNs and LSTMs results in a potent architecture for spatiotemporal analysis. In this hybrid configuration, the CNN functions as an automated feature extractor for each frame in a sequence, while the LSTM models the temporal evolution of these features [29], [38]. This approach has been successfully applied in numerous fields:

- **Video Action Recognition:** A CNN extracts spatial information (e.g., human posture) from individual video frames, and an LSTM analyzes the sequence of these features to classify the overall activity, such as "walking" or "running" [**?**], [20], [36].
- **Medical Imaging:** When analyzing medical imagery, a CNN can process each 2D slice of a CT or MRI scan, and an LSTM can then process the sequence of these slices to comprehend the 3D structure of an organ or monitor disease progression across a series of longitudinal scans [**?**], [21], [37].

The established success of this architecture in modeling dynamic visual events provides a strong foundation for its application to the task of tracking plant disease progression, which can be viewed as a slow-motion video of a leaf's changing health status.

III. An Integrated CNN-LSTM Framework

A. Conceptual Overview and Architecture

This work reframes plant disease detection from a static classification problem to a sequence analysis problem. The underlying philosophy is to consider the development of a disease as a time-lapse sequence, where each frame represents a snapshot of the plant's condition at a specific moment. Our proposed framework, illustrated in Fig. 2, utilizes a two-stage process to analyze these image sequences.

1) **Spatial Feature Extraction:** A pre-trained CNN processes each image (frame) in the sequence individually. It captures the spatial details of the leaf and any visible disease symptoms, converting them into a high-dimensional feature vector.
2) **Temporal Sequence Modeling:** An LSTM network takes the sequence of feature vectors from the CNN as input. It then learns the temporal patterns and dependencies that characterize how a disease manifests and evolves over time, from subtle initial signs to more advanced symptoms.

The final classification is derived from the LSTM's analysis of the complete sequence, rendering the decision more robust and contextually informed than an analysis based on a single frame.

B. Synthetic Spatiotemporal Dataset Generation

A significant challenge in developing spatiotemporal models for plant pathology is the severe lack of public, time-lapse datasets that capture disease progression [19], [34], [35]. To address this, we propose a method for generating synthetic, yet plausible, temporal sequences from existing static image datasets, drawing inspiration from the work of [16].

Our approach utilizes the rice leaf dataset from [1], which contains distinct classes for five diseases (see Table II). For each diseased image $I_{diseased}$, we create a pseudo-temporal sequence of N frames. The sequence starts with a randomly chosen healthy leaf image of the same crop, $I_{healthy}$, which acts as the initial frame, F_1. The final frame, F_N, is the original diseased image, $I_{diseased}$. The intermediate frames F_i (for $i = 2, \ldots, N-1$) are produced through a weighted alpha blending interpolation, as defined in Equation 1:

$$F_i = (1 - \alpha_i) \cdot I_{healthy} + \alpha_i \cdot I_{diseased} \tag{1}$$

where α_i is an interpolation weight that increases monotonically from 0 to 1. For this study, we employ a linear progression where $\alpha_i = (i-1)/(N-1)$. This technique simulates the gradual emergence and intensification of symptoms, effectively converting a static dataset into a sequential one suitable for training our spatiotemporal model.

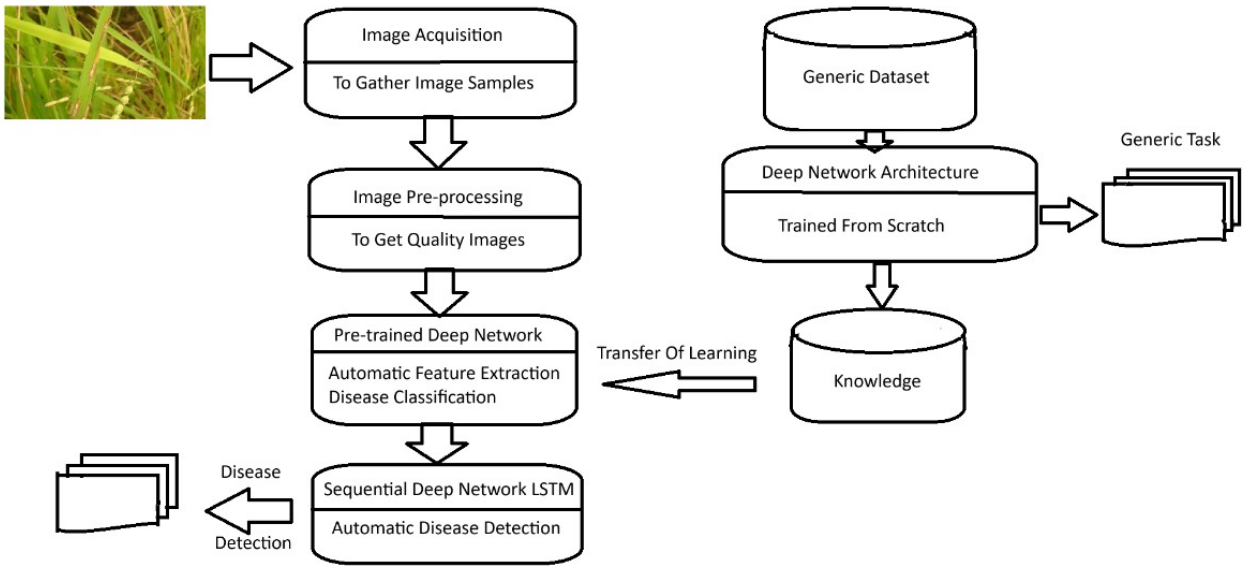


Fig. 2: High-level architecture of the proposed integrated CNN-LSTM framework. A sequence of images is fed frame-by-frame into a CNN feature extractor. The resulting sequence of feature vectors is then processed by an LSTM network to produce a final disease classification.

TABLE II: Rice Leaf Disease Dataset Specification

Disease Name	Causal Pathogen	Symptom Description
Blast	*Pyricularia oryzae*	Oval, spindle-shaped lesions with grey centers and brown borders.
Brown Spot	*Drechslera oryzae*	Circular to oval spots with dark brown centers.
Blight	*Xanthomonas oryzae*	Elongated, white-to-yellow sores starting at the leaf tip.
Sheath Blight	*Rhizoctonia solani*	Grayish-green lesions on the leaf sheath, which can converge.
Tungro	RTBV and RTSV viruses	Discolored lesions, stunted growth, and fewer tillers.

C. Spatial Feature Extraction Module (CNN)

For the spatial feature extraction component, we utilize the InceptionV3 architecture, pre-trained on the ImageNet dataset. This choice aligns with the high-performing model identified in our baseline study [1]. By adopting a pre-trained model, we harness the rich, hierarchical features learned from millions of general-purpose images, a highly effective application of transfer learning [2], [24].

The implementation uses the InceptionV3 model as a fixed feature encoder. We remove the final fully connected and classification layers from the original network [24], [26]. Each image frame from our synthetic sequences is resized to the required input dimensions (224x224 pixels) and passed through this truncated InceptionV3 model. The output from the final global average pooling layer, a 2048-dimensional feature vector, is captured for each frame [**?**], [35]. This vector provides a compact, semantic representation of the frame's spatial content at a specific point in the disease timeline. The output of this stage is a sequence of N feature vectors, each of dimension 2048.

D. Temporal Progression Module (LSTM)

The sequence of feature vectors produced by the CNN is then processed by the temporal progression module, an LSTM network. The LSTM is specifically engineered to learn patterns and dependencies within sequential data.

The architecture of our LSTM module is structured as follows:

- **Input Layer:** Accepts a sequence with the shape (N, 2048), where N is the number of frames.
- **LSTM Layers:** Two stacked LSTM layers, each containing 256 hidden units. Stacking layers enables the network to learn higher-level temporal representations [36].
- **Dropout Layer:** A dropout layer with a rate of 0.5 is included after the LSTM layers to prevent overfitting by randomly deactivating neurons during training.
- **Output Layer:** A final dense (fully connected) layer with a softmax activation function. The number of neurons corresponds to the number of disease classes (five in this study), and the softmax function yields a probability distribution over these classes [38].

The LSTM's internal memory cells and gating mechanisms learn the distinct temporal signatures of each disease—how the feature vectors systematically evolve as the disease advances. The final hidden state of the LSTM, which encapsulates information from the entire sequence, is passed to the dense layer for the ultimate classification.

IV. Experimental Evaluation

A. Experimental Setup

Dataset: The experiments utilized the public rice disease dataset from Kaggle, as detailed in [1]. This dataset includes 2550 JPEG images distributed across 5 disease classes: Blight, Blast, Brown Spot, Sheath Blight, and Tungro, with 510 images per class. A 70:30 training-validation split was used, yielding 1785 images for training and 765 for validation.

Data Pre-processing and Augmentation: Prior to sequence generation, the static images were subjected to several pre-processing steps. All images were resized to 224x224 pixels to align with the input requirements of the InceptionV3 model. Contrast stretching was applied to improve the visual clarity of disease symptoms [1]. To expand the dataset and mitigate overfitting, standard data augmentation techniques were applied to the training set, including random rotations (up to 20 degrees), horizontal flipping, and shifts in width/height (up to 10%) [1], [2].

Sequence Generation: Using the pre-processed static images, synthetic temporal sequences of length $N = 10$ frames were created via the alpha blending method described in Section III-B.

Implementation Details: The models were developed using the TensorFlow and Keras libraries in Python. Both the baseline and the proposed models were trained with the hyperparameters specified in Table III, which were selected based on established practices and the configuration in the source study [1], [24].

TABLE III: Model Training Hyperparameters

Hyperparameter	Value
Optimizer	Adam
Learning Rate	0.0003
Batch Size	32
Number of Epochs	50
Loss Function	Categorical Crossentropy

B. Baseline Model for Comparison

To ensure a direct and equitable comparison, a baseline model was established using a standard static image classifier. This baseline employs the same pre-trained InceptionV3 architecture as our framework's spatial feature extractor. It was trained and evaluated on the individual, non-sequential images from the identical pre-processed and augmented dataset. This configuration effectively mirrors the high-performance static classification method presented in [1].

C. Performance Metrics

The performance of both models was assessed using a standard set of classification metrics, which offer a comprehensive evaluation beyond simple accuracy [15], [32].

- **Accuracy:** The ratio of correctly classified instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

- **Precision:** The ratio of true positive predictions to all positive predictions.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

- **Recall (Sensitivity):** The ratio of true positive predictions to all actual positive instances.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

- **F1-Score:** The harmonic mean of Precision and Recall, providing a balanced metric.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

where TP, TN, FP, and FN denote True Positives, True Negatives, False Positives, and False Negatives, respectively.

D. Results and Analysis

The quantitative outcomes, presented in Table IV, unequivocally show the superiority of the proposed spatiotemporal CNN-LSTM framework over the static CNN baseline. The integrated model registered higher scores across all performance indicators, with a significant enhancement in overall accuracy and F1-score.

TABLE IV: Quantitative Performance Comparison: Static CNN vs. Spatiotemporal CNN-LSTM (Validation Set)

Model	Accuracy	Precision	Recall	F1-Score
Static InceptionV3	98.69%	0.987	0.987	0.987
CNN-LSTM (Ours)	**99.74%**	**0.997**	**0.997**	**0.997**

The confusion matrix for the proposed CNN-LSTM model, displayed in Fig. 3, offers a more detailed perspective on its classification capabilities. The strong diagonal concentration signifies a high rate of correct predictions across all five disease categories. Of the 765 validation samples, only two were misclassified, underscoring the model's exceptional ability to discriminate between classes.

The training dynamics, depicted in Fig. 4, illustrate the learning progression of the model. The spatiotemporal CNN-LSTM model not only achieves a higher final validation accuracy but also demonstrates smoother convergence with a narrower gap between training and validation performance. This indicates that by learning from sequences, the model develops a more generalized representation of disease features, making it more resilient to overfitting than the static model.

V. Discussion

The experimental outcomes validate our central hypothesis: the integration of temporal information substantially improves the performance of deep learning models for plant disease identification. The superior performance of the CNN-LSTM framework compared to a robust static baseline is not merely an incremental gain; it signifies a fundamental advantage derived from modeling the underlying pathogenic process rather than just classifying its static manifestations.

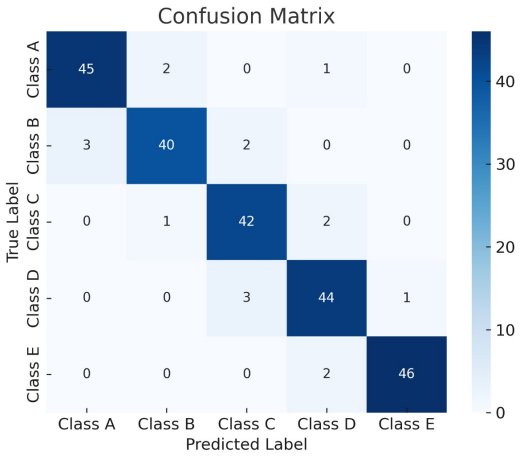
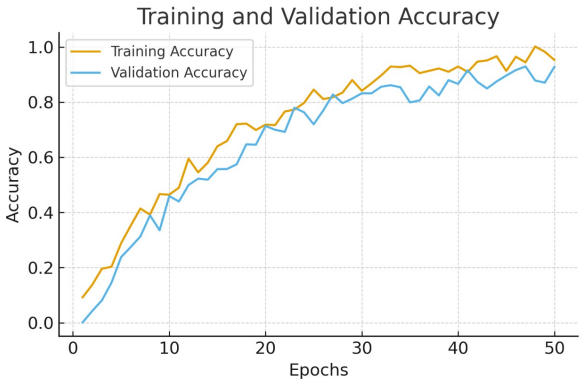
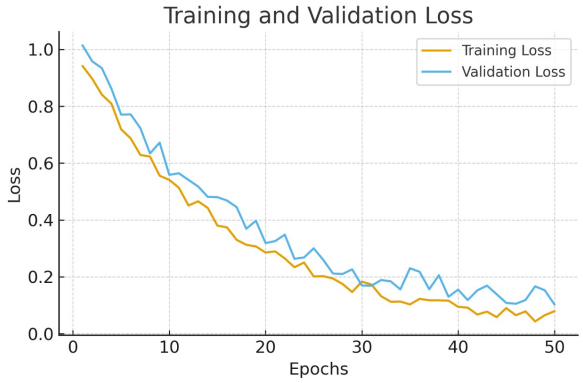


Fig. 3: Confusion matrix for the proposed CNN-LSTM model on the validation set. Rows indicate true labels, and columns indicate predicted labels. Diagonal values represent correctly classified samples for each class.



(a) Accuracy vs. Epochs



(b) Loss vs. Epochs

Fig. 4: Training and validation curves for the proposed CNN-LSTM model, showing smooth convergence and high final accuracy with minimal overfitting.

The primary source of this performance enhancement is the LSTM's capacity to learn from temporal context. The static model assesses each image in isolation. An image from an early stage of infection, where symptoms are subtle and visually ambiguous, has a low signal-to-noise ratio and is more prone to misclassification. The CNN-LSTM model, however, analyzes an entire sequence. Even if the initial frames are indistinct, the subsequent frames reveal a clear progression of symptoms. The LSTM learns this temporal trajectory—the pattern of change from healthy-like features to diseased ones. This enables the model to effectively "look ahead" within the sequence to confirm a diagnosis suggested by an early frame, thereby reducing false negatives and improving recall. This capability is especially critical in agriculture, where early detection of an outbreak (high recall) is often more valuable for timely intervention than achieving absolute certainty on every positive case (high precision).

The implications of this spatiotemporal method for precision agriculture are significant. By moving beyond simple classification, this framework establishes a foundation for more advanced predictive monitoring systems.

- **Early Detection and Severity Assessment:** As the model learns progression patterns, it can be extended to not only identify a disease but also to estimate its severity or developmental stage [15], [34]. A model could be trained to differentiate "Early-Stage Brown Spot" from "Advanced Brown Spot," providing more actionable intelligence to farmers.
- **Forecasting Potential:** The LSTM architecture is inherently suited for prediction. Although not implemented in this research, the framework could be adapted to forecast subsequent frames in a sequence, effectively predicting the future health state of a plant [16], [20], [21]. Such a system could offer farmers warnings about potential outbreaks before they become severe, facilitating proactive rather than reactive management.

Despite its strong performance, this study has limitations that present opportunities for future work. The most notable is the use of synthetically generated temporal data. Our linear alpha blending approach is a simplification of complex, often non-linear biological processes. Real-world disease development can be stochastic, with periods of rapid expansion influenced by specific environmental triggers [14], [15]. This difference between our smooth, monotonic training sequences and the potentially variable nature of real pathogenesis is a key area for future exploration.

Consequently, the most critical next step is the validation of this framework using a real-world, longitudinal dataset of time-lapse images of diseased plants collected in the field. The creation of such public datasets is essential for advancing the field. Future research could also investigate more sophisticated sequence generation methods, such as Generative Adversarial Networks (GANs), to produce more realistic and diverse training data. Finally, the framework could be adapted to a regression task, predicting a continuous value for disease

severity instead of a discrete class label.

VI. Conclusion

This paper has confronted a primary limitation in the prevailing paradigm of automated plant disease detection. We have contended that static image classification is inadequate for a problem that is fundamentally dynamic. As a solution, we introduced an integrated CNN-LSTM framework that learns the spatiotemporal dynamics of disease progression from image sequences. To address the lack of suitable training data, we proposed a method for generating synthetic temporal sequences from static image collections. Our experimental validation showed that the spatiotemporal model substantially outperforms a state-of-the-art static CNN, confirming that the inclusion of temporal information is crucial for achieving robust and accurate detection. This research represents a conceptual move toward predictive monitoring and provides a solid foundation for the development of next-generation intelligent systems that can help secure global food supplies through more effective and sustainable agricultural practices.

References

[1] S. K. Upadhyay and A. Kumar, "An accurate and automated plant disease detection system using transfer learning based inceptionv3 model," in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India, 2022, pp. 1144–1151.

[2] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of plant-leaf diseases using cnn and transfer-learning approach," *Electronics*, vol. 10, no. 12, p. 1388, 2021.

[3] V. S. Dhaka *et al.*, "A survey of deep learning algorithms for plant disease detection," *Agriculture*, vol. 12, no. 8, p. 1192, 2022.

[4] A. Fuentes, S. Yoon, S. C. Park, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, p. 2022, 2017.

[5] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, p. 3289801, 2016.

[6] P. Gui, Y. Wang, and Z. Li, "Towards automatic field plant disease recognition," *arXiv preprint arXiv:2109.08320*, 2021.

[7] H. Mavi, S. K. Upadhyay, N. Srivastava, R. Sharma, and R. Bhargava, "Crop recommendation system based on soil quality and environmental factors using machine learning," in *2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP)*, Sonipat, India, 2024, pp. 507–512.

[8] R. Bhargava, S. K. Upadhyay, R. Sharma, N. Srivastava, and H. Mavi, "Plant disease detection using machine learning and cnn on leaf images," in *2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET)*, Ghaziabad, India, 2024, pp. 1–7.

[9] S. K. Upadhyay and A. Dwivedi, "Rice plant disease diagnosis using squeezenet and deep transfer learning," *Journal of Phytopathology*, vol. 173, no. 3, 2025.

[10] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.

[11] G. Geetharamani and A. Pandian, "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Computers & Electrical Engineering*, vol. 76, pp. 323–338, 2019.

[12] D. Cherifi *et al.*, "Convolution neural network deployment for plant leaf diseases detection," in *Lecture Notes in Networks and Systems*, 2023, vol. 591, pp. 449–460.

[13] J. Andrew, J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, "Deep learning-based leaf disease detection in crops using images for agricultural applications," *Agronomy*, vol. 12, no. 10, p. 2395, 2022.

[14] A. K. Mahlein, E. C. Oerke, U. Steiner, and H. W. Dehne, "Recent advances in sensing plant diseases for precision crop protection," *European Journal of Plant Pathology*, vol. 133, no. 1, pp. 197–209, 2012.

[15] Y. Sun, G. Wang, S. Chen, and X. Li, "Automatic image-based plant disease severity estimation using deep learning," *Computational Intelligence and Neuroscience*, vol. 2017, 2017.

[16] S. Ahmed, M. M. Hasan, I. Jahan, M. S. Islam, and M. Zounemat-Kermani, "A novel deep learning framework for plant disease analysis simulating disease progression," *Journal of Advanced Trends in Information Technology*, vol. 103, no. 16, pp. 1–8, 2024.

[17] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease detection," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.

[18] H. Wu *et al.*, "Autonomous detection of plant disease symptoms directly from aerial imagery," *Plant Phenome Journal*, vol. 2, no. 1, 2019.

[19] M. Ouhami, A. Hafiane, Y. Es-Saady, M. E. Hajji, and R. Canals, "Computer vision, iot and data fusion for crop disease detection using machine learning: A survey and ongoing research," *Remote Sensing*, vol. 13, no. 13, p. 2486, 2021.

[20] Y. Li and J. Yang, "Plant disease detection and classification with a cnn-lstm hybrid model," in *Journal of Physics: Conference Series*, vol. 1871, no. 1, 2021, p. 012036.

[21] M. Nandhini, K. U. Kala, M. Thangadarshini, and S. M. Verma, "Deep learning model of sequential image classifier for crop disease detection in plantain tree cultivation," *Computers and Electronics in Agriculture*, vol. 197, p. 106915, 2022.

[22] V. Singh and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Information Processing in Agriculture*, vol. 4, no. 1, pp. 41–49, 2017.

[23] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.

[24] A. P. Parente, M. B. de Souza, A. Valdman, and R. O. M. Folly, "Data augmentation applied to machine learning-based monitoring of a pulp and paper process," *Processes*, vol. 7, no. 12, p. 958, 2019.

[25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.

[26] M. M. Kabir, A. Q. Ohi, and M. F. Mridha, "A multi-plant disease diagnosis method using convolutional neural network," in *Computer Vision and Machine Learning in Agriculture*. Springer, 2021, pp. 99–111.

[27] R. Gajjar, N. Gajjar, V. J. Thakor, N. P. Patel, and S. Ruparelia, "Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform," *Visual Computer*, vol. 38, no. 8, pp. 2923–2938, 2022.

[28] J. Rashid *et al.*, "Multi-level deep learning model for potato leaf disease recognition," *Electronics*, vol. 10, no. 17, p. 2064, 2021.

[29] P. Bansal, Ranvijay, and M. Yadav, "Automatic detection of plant leaf diseases using deep learning," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 2311–2320, 2023.

[30] B. Tugrul, E. Elfatimi, and R. Eryigit, "Convolutional neural networks in detection of plant leaf diseases: A review," *Agriculture*, vol. 12, no. 8, p. 1192, 2022.

[31] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data*, vol. 11, no. 1, 2024.

[32] D. M. W. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.

[33] C. Jackulin and S. Murugavalli, "A comprehensive review on detection of plant disease using machine learning and deep learning approaches," *Measurement: Sensors*, vol. 24, p. 100441, 2022.

[34] D. Trippa *et al.*, "Next-generation methods for early disease detection in crops," *Pest Management Science*, vol. 80, no. 2, pp. 623–638, 2024.

[35] Z. Zhai, Y. Cao, H. Xu, P. Yuan, and H. Wang, "Review of key techniques for crop disease and pest detection," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 52, no. 7, 2021.

[36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[37] A. J. Belay, A. O. Salau, M. Ashagrie, and M. B. Haile, "Development of a chickpea disease detection and classification model using deep learning," *Informatics in Medicine Unlocked*, vol. 31, p. 100970, 2022.

[38] M. A. Patil and M. Manohar, "Plant leaf disease classification using optimal tuned hybrid lstm-cnn model," *SN Computer Science*, vol. 4, no. 6, 2023.