IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 June 2023, accepted 13 July 2023, date of publication 14 August 2023, date of current version 25 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3305273

RESEARCH ARTICLE

Deep Learning-Based Robust Analysis of Laser Bio-Speckle Data for Detection of Fungal-Infected Soybean Seeds

NIKHIL KALER[1], VIMAL BHATIA[1,2,3,4], (Senior Member, IEEE), AND AMIT KUMAR MISHRA[5,6], (Senior Member, IEEE)

[1]Signals and Software Group, Discipline of Electrical Engineering, Indian Institute of Technology Indore (IIT Indore), Indore 453552, India
[2]Centre for Advance Electronics, Indian Institute of Technology Indore (IIT Indore), Indore 453552, India
[3]Faculty of Informatics and Management, University of Hradec Kralove, 50003 Hradec Kralove, Czech Republic
[4]School of Electronic and Information Engineering, Soochow University, Suzhou 215006, China
[5]Division of Mathematics, Computer and Surveying Engineering, University West, 461 32 Trollhättan, Sweden
[6]Department of Electrical Engineering, University of Cape Town, Cape Town 7700, South Africa

Corresponding author: Vimal Bhatia (vimal_bhatia@ieee.org and vbhatia@iiti.ac.in)

This work was supported in part by the Government of India through Science and Engineering Research Board Project Grant CRG/2021/001215 and Grant CRG/2018/002697; in part by the Staff and Student Mobility funded by Erasmus+ Project between Indian Institute of Technology (IIT) Indore and University West; in part by the Chair Professorship from Soochow University, China; and in part by the Division of Mathematics, Computer and Surveying Engineering, University West, Sweden.

ABSTRACT Seed-borne diseases play a crucial role in affecting the overall quality of seeds, efficient disease management, and crop productivity in agriculture. Detection of seed-borne diseases using machine learning (ML) and deep learning (DL) can automate the process at large-scale industrial applications for providing healthy and high-quality seeds. ML-based methods are accurate for detecting and classifying fungal infection in seeds; however, their performance degrades in the presence of noise. In this work, we propose a laser bio-speckle based DL framework for detection and classification of disease in seeds under varying experimental parameters and noises. We develop a DL-based spatio-temporal analysis technique for bio-speckle data using DL networks, including neural networks (NN), convolutional neural networks (CNN) with long-short-term memory (LSTM), three-dimensional convolutional neural networks (3D CNN), and convolutional LSTM (ConvLSTM). The robustness of the DL models to noise is a key aspect of this spatio-temporal analysis. In this study, we find that the ConvLSTM model has an accuracy of 97.72% on the test data and is robust to different types of noises with an accuracy of 97.72%, 94.31%, 98.86%, and 96.59%. Furthermore, the robust model (ConvLSTM) is evaluated for variations in experimental data parameters such as frame rate, frame size, and number of frames used. This model is also sensitive towards detecting bio-speckle activity of different order, and it shows average test accuracy of 99% for detecting four different classes.

INDEX TERMS Agriculture, bio-speckle, convolutional neural network, deep learning, long-short term memory, neural network, noise, photonics, seed-borne fungi.

I. INTRODUCTION

Most of the food crops in the world are produced through seeds thus healthy and viable seeds are a prime requirement for high-crop yield. Seed-borne pathogens pose severe challenge that restricts availability of healthy seeds. Pathogens such as bacteria, fungi, and viruses cause biotic stress in seeds. Their early negative impact on seed germination causes weak and uneven seedling establishment, resulting in low crop yield [1]. Soybean is a globally significant crop, with over 350 million metric tons produced annually in 2020. It is a vital protein source, accounting

The associate editor coordinating the review of this manuscript and approving it for publication was Sukhdev Roy.

for 60% of total oilseed production [2]. Soybean contribute to food security, provide animal feed, support the biofuel industry, and play a crucial role in sustainable agriculture and economic development. Seed-borne diseases in soybean can spread through infected seeds, causing plant infections, decreased vigor, and lower productivity [3].

Conventional ways to determine the presence of seed-borne pathogens include visual inspection, microscopic examination, seed-soaking techniques, serological testing, and bioassay. These methods are time-consuming, inefficient, expensive, and have insufficient sensitivity, specificity, and accuracy towards different concentrations and types of seed-borne pathogens [4]. Modern molecular diagnostic techniques based on polymerase chain reaction (PCR) and deoxyribonucleic acid (DNA) analysis include conventional and several advanced PCR-based methods. The main shortcomings of molecular approaches are their complexity, failure to differentiate between vital and non-vital inoculum, and challenges in obtaining high-quality DNA templates due to PCR inhibitors in seeds [5]. Non-destructive methods for detection of contamination in a variety of agricultural products includes optical techniques like near-infrared spectroscopy (NIRS), fluorescence spectroscopy (FS), Raman spectroscopy (RS), and hyperspectral imaging (HSI) [6], [7], [8], [9]. However, optical techniques have limitations, such as in FS, where background fluorescent elements affect the fluorescence spectra, necessitating chemometric techniques for precise models. NIRS is sensitive to uneven contamination, while RS exhibits limitations in sensitivity and repeatability. High experimental cost, data dimensionality, and time-consuming image acquisition and analysis restrict the widespread adoption of HSI technique [10].

In recent years, laser bio-speckle technique has emerged as a non-destructive optical method to track biological changes inside a sample by stacking the time series speckle images, which detect morphological and physiological changes within the sample under observation [11]. A statistical interference pattern known as a 'Speckle' is formed when a specimen is illuminated with a coherent light source. In the laser bio-speckle technique, light is reflected using an optical reflecting surface onto the sample under observation. On the observation plane, interference causes the appearance of bright and dark regions [12]. The laser bio-speckle technique has several advantages over other optical-based methods, including lower memory requirements, higher spatial and temporal resolution, simple and less costly experimentation setup, and its ability to track biological and morphological changes in the sample.

The two methods commonly used in the literature to examine dynamicity of a bio-speckle pattern are numerical indexing and visual inspection. Fujii's method, generalized difference (GD), subtraction and modified subtraction average (MSA), and other visual techniques give qualitative visualization of an object's activity. Commonly used numerical indexing strategies are temporal history of speckle pattern (THSP) combined with inertia moment (IM), absolute value of difference (AVD), wavelet entropy (WE), and spatial-temporal speckle correlation technique, etc. [12], [13]. Various experimental and processing-related factors such as sampling frequency, number of frames acquired, manual region of interest (ROI) selection, single column selection from speckle frame, and illumination conditions affect performance of the above algorithms.

Frequency analysis and decomposition methods on bio-speckle data have been used in the agriculture sector for various applications such as the detection of fungi in beans [14], [15], analysis of water activity in biospecimens [16], [17], [18], and sprouting damage assessment [19]. With advancements in intelligent systems, the decision-making process is automated using machine learning (ML) for early identification of seed fungal infection and seed viability assessment [20], [21]. Extraction of handcrafted features is a lengthy and semi-automated process, requires professional expertise, and is restricted to previously known limited features for decision-making. Manually extracted features are also sensitive to noise in data, resulting in performance degradation of ML-based models when tested in noisy conditions (Table 3).

Further, feature extraction and decision-making can be automated using deep learning (DL) that allows a model to extract valuable features while learning from data and removes the disadvantage of using a specific feature extraction technique. DL-based methods have been used for seed viability assessment which utilizes a single speckle frame's spatial information [22]. Using only a single speckle frame for the bio-speckle activity assessment is prone to noise, and it does not include the dynamic component of bio-speckle activity for its assessment. A non-destructive, real-time applicable automated technique for contamination detection in seeds is of great interest in agriculture. The laser bio-speckle technique combined with DL can show great potential among emerging techniques for real-time applications in this direction.

Therefore, this investigation focuses on automatically detecting fungal-infected soybean seeds using DL-based methods. The main objectives of the study are; (i) to develop DL-based spatio-temporal analysis on bio-speckle data using different models, (ii) to compare various models for robustness against noise in data acquisition, (iii) to find the impact of experimental parameters of data such as frame size, frame rate and the number of frames used on the performance of the best robust model, and (iv) to check further model's sensitivity to detect the bio-speckle activity of different orders by extending it to multi-class data.

II. MATERIALS AND METHODS

A. EXPERIMENTAL SETUP AND DATA ACQUISITION

Majority of the anthracnose disease-causing seed-borne fungus is found in soybean seeds. *Colletotrichum truncatum*

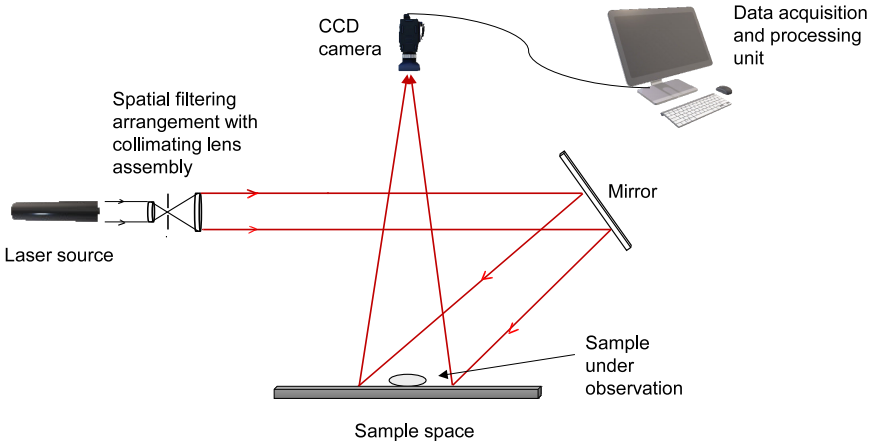


FIGURE 1. Experimental setup for data acquisition.

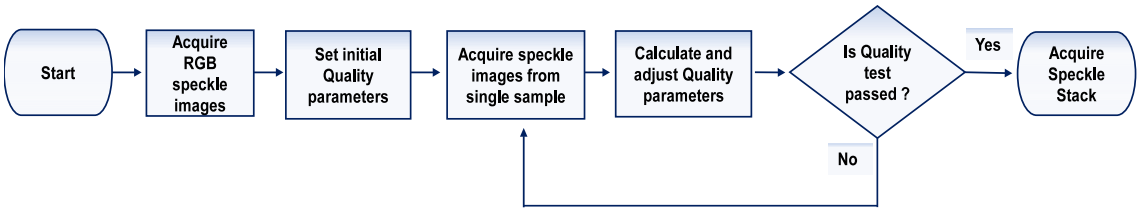


FIGURE 2. Block diagram for data acquisition process.

(*C. truncatum*) is the primary cause of anthracnose and the most common threatening seed-borne fungal infection [23]. For experiments, 1000 soybean seeds of the JS 20-29 variety were gathered from the ICAR-Indian Institute of Soybean Research (IISR) plant field, Indore (India). These seeds were disinfected with an equal proportionate solution of sodium hypochlorite and distilled water, then surface sterilized with distilled water and dried at room temperature on sterilized filter paper. Further, the seeds were separated into two different groups, namely healthy and diseased. The diseased group was created by inoculating soybean seeds with *C. truncatum* fungi using a solution containing 10^6 spores/ml. For comparison, the healthy group was treated with 20 μl of sterile distilled water.

The experimental setup arrangement is shown in Fig.1, which consists of a He-Ne laser source (wavelength = 632.8 nm and power = 15 mW), variable attenuator, spatial filtering arrangement (microscopic objects (m=40X and aperture=10 μm), collimation lens), and charged-coupled device (CCD) camera (Basler Corp., frame rate=32 frames per sec(fps); resolution=1294 $\times$ 964).

A spatial filtering arrangement creates a filtered and enlarged laser beam by beam filtering to inhibit the noise components in the laser beam. Neutral density filter (Thorlabs, USA) absorbs in visible and near-IR range to reduce the optical power of the incident beam, eliminating the risk of the sample being under or over-exposed, which could lead to erroneous results. A precision achromatic doublet lens (focal length = 250 mm, diameter = 50 mm) collates the expanded laser beam. Vibration-isolation table top was used to avoid errors due to vibrations during experimentation.

Finally, the laser was utilized to illuminate seed samples, with the backscattered light imaged onto the camera lens. A high frame rate CCD camera was used to capture sequential speckle images at varying time intervals for each sample. Quality characteristics such as contrast and saturation are highly influenced by the recording angle and size of the speckle grain, which influences bio-speckle activity [24]. As a result, these quality parameters are optimized before data acquisition. Data acquisition process is shown through block diagram in Fig. 2. Bio-speckle activity inside healthy and diseased samples are of different orders due to different physiological and biochemical processes inside the samples. Some initial frames from bio-speckle image stack of both types of sample and original sample images are shown in Fig. 3 and 4, respectively.

B. NOISE ADDITION AND NOISED TEST DATASET PREPARATION

Undesirable noise is added to the acquired data during experimentation and data acquisition stage. The quality of speckle images depends upon several factors, including atmospheric disturbances, laser energy, CCD sensitivity, and uneven

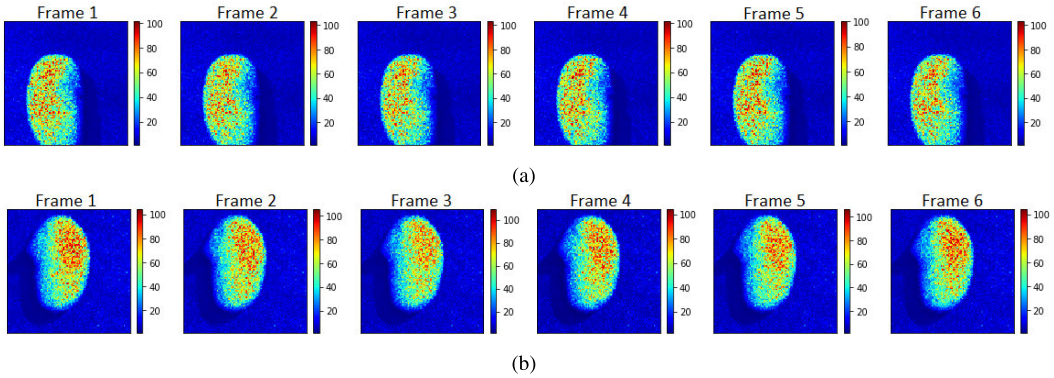


FIGURE 3. Temporal variation in recorded speckle frames for (a) healthy sample (b) diseased sample.

(a) Healthy sample (b) Diseased sample

FIGURE 4. Original images of samples.

illumination conditions, which results in various types of noises in final images. There are mainly three types of noises (impulsive noise, Gaussian noise, and speckle noise) that impact accuracy of the analysis [25]. Phase distortions, temporal coherence of the laser, and detector aperture affect the speckle images in laser speckle imaging, producing noisy speckles that contain both noise and information [26].

CCD photosensors convert the incident photons on the detector to a digital signal through multiple stages: converting photons-to-electrons, electrons-to-voltages, and voltages-to-digital signals. During this process, various noises like photon shot noise, dark current noises, sense node reset noises, and non-linearity get added at different stages, which are modeled using Poisson, Gaussian, log-normal, and inverse Gaussian distributions while designing simulation model for CCD photosensors [27].

The performance of DL-based model is affected by the presence of noise in the data. Therefore, we created four test data sets besides the original test dataset for robustness analysis by adding spatial noises in bio-speckle video frames of test data. Noise-added test datasets are described as:

1) GAUSSIAN NOISED TEST DATASET

Every pixel in the Gaussion noised image is the sum of the random Gaussian distributed noise and actual pixel values [28]. The probability density function (PDF) of Gaussian noise is mathematically represented as:

$$P(g) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(g-\mu)^2}{2\sigma^2}}, \tag{1}$$

where g is the noise pixel gray value, μ and σ are the mean and standard deviation, respectively.

2) SALT AND PEPPER NOISED TEST DATASET

Salt and pepper noise is impulsive noise that is added to an image by adding random bright (with 255-pixel value) and random dark (with 0-pixel value) values all over the image [28]. Salt and pepper noise is represented as:

$$P(g) = \begin{cases} P_a & \text{if } g = a \\ P_b & \text{if } g = b \\ 0 & \text{if } otherwise, \end{cases} \tag{2}$$

where $P(g)$ is the PDF, and g is the pixel gray value.

3) SPECKLE NOISED TEST DATASET

Speckle noise is multiplicative noise generated by multiplying random pixel values with different pixels of an image, is represented as:

$$K(i,j) = I(i,j) + N(i,j) \times I(i,j), \tag{3}$$

where $I(i,j)$ is the original image, $K(i,j)$ is speckle noised image, and $N(i,j)$ is Gaussian distributed noise.

4) MULTIPLE NOISES ADDED TEST DATASET

This test dataset is prepared using multiple noises considered for modeling of CCD photosensor. Image $I(i,j)$ is used as the λ parameter for generating Poisson random numbers to simulate shot noise in the sample frames. Furthermore, Poisson noised test samples are added with log-normal and Gaussian-distributed noises.

The PDF of log-normal distributed noise is given as:

$$P_X(x) = \frac{1}{x\sigma'\sqrt{2\pi}} \exp\left(-\frac{\left(\ln(x) - \mu'\right)^2}{2\sigma'^2}\right), \tag{4}$$

TABLE 1. Numerical quantification of noise addition using different parameters for noised test dataset samples.

Noise Used	Signal-to-Noise Ratio (SNR(dB))	Peak Signal-to-Noise Ratio (PSNR(dB))	Correlation Coefficient (CC)	Mean Square Error (MSE)	Structural Similarity Index (SSI)
Gaussian noise	14.440	30.656	0.957	55.899	0.7196
Salt and pepper noise	4.715	20.930	0.736	526.30	0.6402
Speckle noise	10.025	26.240	0.8985	154.537	0.7325
Multiple added noise	15.053	31.273	0.968	48.47	0.7931

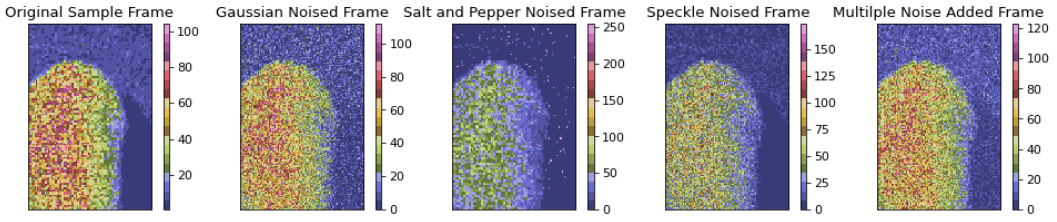


FIGURE 5. Effect of adding different noises on pixel values in a single frame of a test sample.

where μ' and σ' are mean and standard deviation of the natural logarithm of random variable X, respectively.

The PDF of Poisson distributed noise is given as:

$$P(X = k, \lambda) = \frac{e^{-\lambda}\lambda^k}{k!}, \tag{5}$$

where λ is the mean number of events, and k is the number of occurrences.

Several parameters like signal-to-noise ratio (SNR), peak signal-to-noise ratio (PSNR), mean square error (MSE), correlation coefficient (CC), and structural similarity index (SSI) can be used for the analysis of noise in the images [26]. The above parameters are calculated, and the mean value of all the sample frames is used to evaluate the extent and effect of noise addition in the test sample. Table 1 has values of the above-mentioned parameters for a single sample from different noise-added test datasets. Fig. 5 and 6 show changes in the pixel intensity values for a single frame in the sample by adding different types of noises.

C. SYNTHETIC DATA GENERATION TO EXTEND THE DATASET TO MULTI CLASSES

A model is said to be sensitive to the detection of bio-speckle activity if that can easily identify the bio-speckle activity of different orders. Bio-speckle activity in experimentally collected data is primarily of two orders, which is quantitatively evaluated in terms of the bio-speckle activity index (BSAI). The subtraction average algorithm [29] is used to find BSAI values of healthy and diseased samples. The dataset is extended to multi-class to evaluate the model's sensitivity to BSAI of different orders. Two additional synthetic classes of higher BSAI values are generated using the standard rotating diffuser model [13]. The rotating diffuser model can virtually produce bio-speckle specimens like the original samples. Firstly, we generate a speckle pattern given as:

$$I_1(a, b) = FT^{-1}\left\{H \times FT\left\{e^{j\times\varphi(a,b)}\right\}\right\}, \tag{6}$$

where a, b are the pixel coordinates, H is the coherent transfer function, $FT^{-1}\{.\}$ and $FT\{.\}$ denote inverse and direct Fourier transform respectively, and $\varphi(a, b) \in [0, 2\pi]$ is random phase matrix.

Other speckle patterns are generated by rotating each of the generated speckle patterns by an angle to get variation in the correlation between speckle images in a stack S.

$$S = [I_1, I_2, I_3, \ldots, I_n] \tag{7}$$

For H = 250, we get a synthetic speckle pattern of frame size=200 × 200 with 96 frames. Table 2 represents the experimental data and two extended classes with their BSAI values.

III. SPATIO-TEMPORAL ANALYSIS OF BIO-SPECKLE DATA USING SUPERVISED LEARNING ALGORITHMS

ML-based processing consists of data pre-processing, feature extraction, predictive model development, and its evaluation. After removing unwanted background from the speckle frames, spatial and temporal features such as temporal absolute value of difference (T-AVD), contrast, entropy, roughness, inertia moment (IM), and spatial AVD are extracted from each sample. Various ML classifiers, including k-nearest neighbour (KNN), support vector machine (SVM), decision tree (DT), random forest (RF), Gaussian Naive Bayes (GNB), and artificial neural network (ANN) are

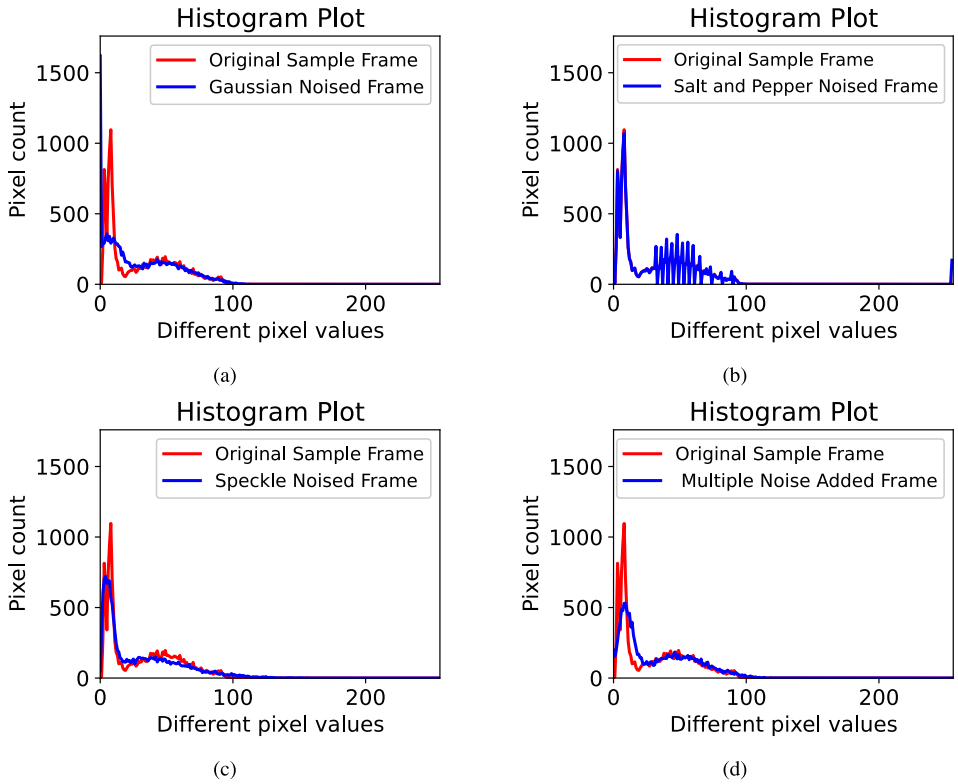


FIGURE 6. Histogram plot for different noise addition in a single frame of a test sample.

TABLE 2. Four class dataset having bio-speckle activity of different orders.

	Experimental data		Synthetic data	
	Healthy samples (Class 0)	Diseased samples (Class 1)	Class 2	Class 3
Approx. BSAI ($\times 10^3$)	0.5-0.8	1.0-1.6	2.0-2.6	3.3-3.8
Samples	300	300	240	240

trained on normalized features and then tested in normal and noisy conditions. Table 3 shows that the ML-based classifiers are not robust to noise in data.

DL-based supervised learning algorithms aim to learn a mapping function between feature vectors (inputs) and labels (output) based on input-output pairs. In forward propagation, the algorithm begins at a random point in parameter space and generates random output. The loss function computes the error based on the mismatch between the output and labels, and during backpropagation, the model parameters are updated using error [30].

Based on the main objectives mentioned in Section-I, a DL-based strategy is developed to find an accurate and robust model which considers both spatial and temporal information of bio-speckle data. The robust model is further investigated for experimental data parameters and the sensitivity of the model to bio-speckle activity of different orders. Fig. 7 represents the flow chart of the developed strategy for this study. Section III-A describes the DL networks and techniques, including neural networks, convolutional neural networks, long short-term memory networks, and transfer learning technique, which are used to develop four different models for spatio-temporal analysis of bio-speckle data. Section III-B describes the performance evaluation measures for the considered models.

A. DEEP LEARNING NETWORKS AND MODELS USED FOR SPATIO-TEMPORAL ANALYSIS

1) NEURAL NETWORK

A suitable feature extractor aims to produce low-dimensional features containing all the relevant and valuable information. Multi-layer perceptron neural networks generally have three different types of layers in the feed-forward architecture. The initial layer, called as input layer of the neural network, accepts the input vector $[x_1, \ldots, x_p]$. The input units are densely connected to the hidden layer units with the help of weights and biases as parameters. Weights (w_{ij}) are multiplied by each input neuron; values are accumulated then the bias term (b_i) is added, which results in a combined value (u_j). This weighted sum (u_j) is then given to the activation function(Θ), producing the outputs of value h_j.

TABLE 3. Performance of different machine learning models on the test and noised test datasets.

Machine learning models	Performance on test datasets						
	Original test dataset (in %)			Noised test datasets (Accuracy in %)			
	Accuracy	Precision	Recall	Gaussian Noised	Salt and Pepper Noised	Speckle Noised	Multiple Noise Added
GNB	89.77	81.25	100	51.13	89.77	51.13	51.13
SVM [1]	98.86	100	97.82	51.13	98.86	51.13	51.13
SVM [2]	98.86	97.82	100	51.13	48.86	51.13	51.13
DT	100	100	100	51.13	80.68	51.13	51.13
RF	100	100	100	51.13	93.18	51.13	51.13
KNN	98.86	100	97.82	55.68	78.40	51.13	54.54
ANN	98.86	100	97.82	48.86	48.86	48.86	48.86

[1] Radial basis function (RBF) kernel of degree = 3 is used.
[2] Polynomial kernel of degree = 3 is used.

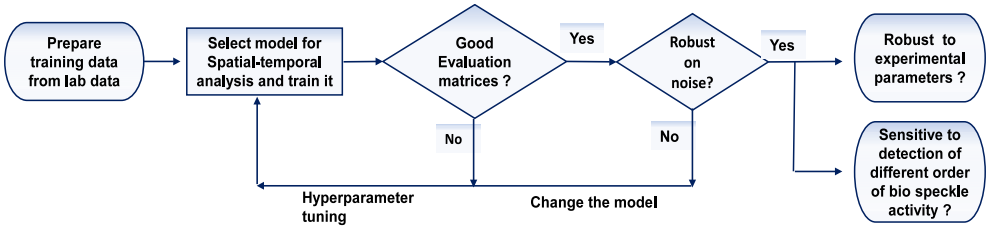


FIGURE 7. Flow chart of the strategy developed for robust spatio-temporal analysis of bio-speckle data based on a DL-based framework.

TABLE 4. Hyperparameter details for machine learning models.

ML models	Hyperparameters used	Optimal value of hyperparameters
SVM	Regularization factor C, Kernel function (linear, polynomial, and RBF), and Gamma	C = 10^6, Kernel function = RBF and polynomial, Gamma = 'scale'
DT	Maximum depth, Maximum leaf nodes	Maximum depth = 5, Maximum leaf nodes = 7
RF	Number of estimators, Criterion (gini impurity, shannon information gain)	Estimators = 100, Criterion = 'Gini impurity'
KNN	Leaf size, K-neighbours	Leaf size = 30, K-neighbours = 5
ANN	Number of neurons and hidden layers, Activation function, and Number of epochs	1 hidden layer with 3 neurons, Activation function = 'Relu', and Epochs = 200

The combined outputs obtained from the hidden layer neurons will be inputs for the output layer, which involve the computation similar to the input and hidden layer [31]. Computation at hidden layers 1 and 2 are given as:

$$U^{[1]} = W^{[1]}X + b^{[1]}, \quad H^{[1]} = \Theta\left(U^{[1]}\right),$$
$$U^{[2]} = W^{[2]}H^{[1]} + b^{[2]}, \quad H^{[2]} = \Theta\left(U^{[2]}\right) = Y, \qquad (8)$$

where $W^{[i]}$, $b^{[i]}$ are weight and bias matrices associated with the i^{th} layer, respectively. $U^{[i]}$ and $H^{[i]}$ are vectorized version of u_i and h_i, X is vectorized input matrix with p number of features for each sample, and Θ is activation function.

2) CONVOLUTIONAL NEURAL NETWORK

Convolutional neural networks (CNNs) are based on four key ideas: local connections, shared weights, pooling, and many layers. CNN uses a series of convolution, pooling, and fully connected layers to extract features for target class recognition. Convolutional layer units are organized in feature maps. By the weights of convolutional filters, each unit is connected to regional patches in the feature maps of the preceding layer. Convolutional filters of the different local receptive fields extract enriched and valuable features from the images. Pooling layers reduces spatial dimensions by combining semantically similar features into one while preserving valuable information. The most common pooling

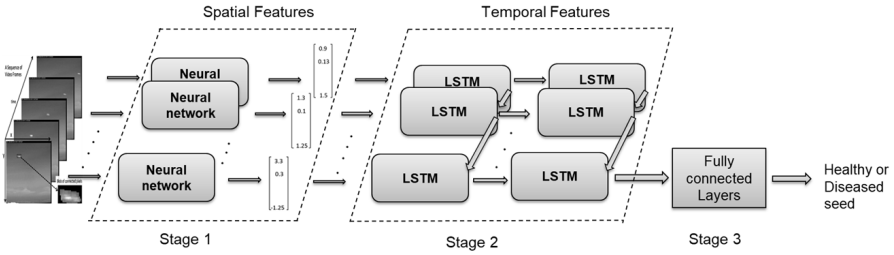


FIGURE 8. Model block diagram of neural network with LSTM network.

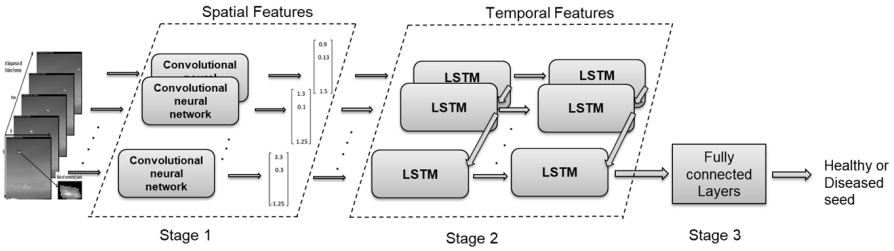


FIGURE 9. Model block diagram of CNN with LSTM network.

layers are max-pooling and average pooling. Parameters (weights and bias) of the convolutional kernel are trained using the backpropagation algorithm, just like a regular neural network [30].

During the convolution operation in convolutional layers, the value at the position (x, y) in the j^{th} feature map in the i^{th} layer, denoted as F_{ij}^{xy}, is given as:

$$\mathrm{F}_{ij}^{xy} = \theta\left(b_{ij} + \sum_{m}\sum_{h=0}^{H_i-1}\sum_{w=0}^{W_i-1} w_{ijm}^{hw}\,\mathrm{F}_{(i-1)m}^{(x+h)(y+w)}\right), \quad (9)$$

$$\mathrm{F}_{ij} = \max_{j\in m}\left(\mathrm{F}_{(i-1)(j)}\right), \quad (10)$$

where θ is the activation function used after the convolutional operation, b_{ij} is the bias for the current feature map, m is number of feature maps in the $(i-1)^{th}$layer, w_{ijk}^{hw} is the value at the position (h, w) of the kernel connected to the k^{th} feature map, $max()$ is a max-pooling function. H_i and W_i are the height and width of the kernel, respectively.

Pre-trained networks on big-labeled public datasets have powerful feature extraction capabilities that can be directly used in other tasks to save the computation required to train the models again from scratch. Transfer learning (TL) has several advantages, including reduced time for training, improved performance, better generalized, and reduced overfitting of the network with limited data [32]. In this analysis, we have used different pre-trained networks for spatial feature extraction at stage 1 of model 2 as shown in Fig. 9. Many pre-trained CNNs are available in DL-based frameworks for transfer learning applications. In this study, we used CNNs that were pre-trained on the ImageNet [33] dataset: VGG16 [34], Resnet50 [35], Densenet121 [36],

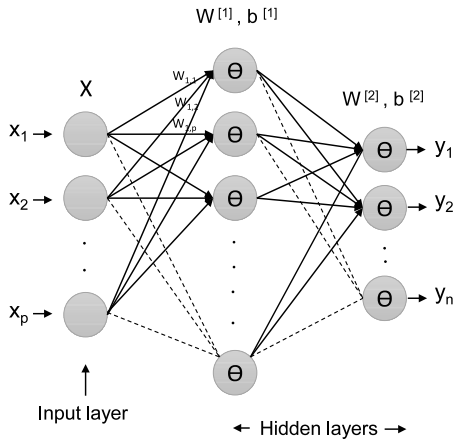


FIGURE 10. MLP neural network.

Inception_V3, Inception_Resnet_V2 [37], Efficient Net [38], Mobile Net [39], and Xception [40].

3) LONG SHORT-TERM MEMORY NETWORK

Traditional deep neural networks, like feedforward neural networks and CNNs, assume that inputs and outputs are independent. However, for sequential data, output also depends on the prior elements within the sequence. Recurrent neural networks (RNNs) are distinguished by their "memory" since they use information from previous inputs to alter the current input and output. Long short-term memory (LSTM) is a special RNN that is stable and effective for long-range dependencies in sequence. The memory cell c^t stores the state information with the help of parameterized controlling gates.

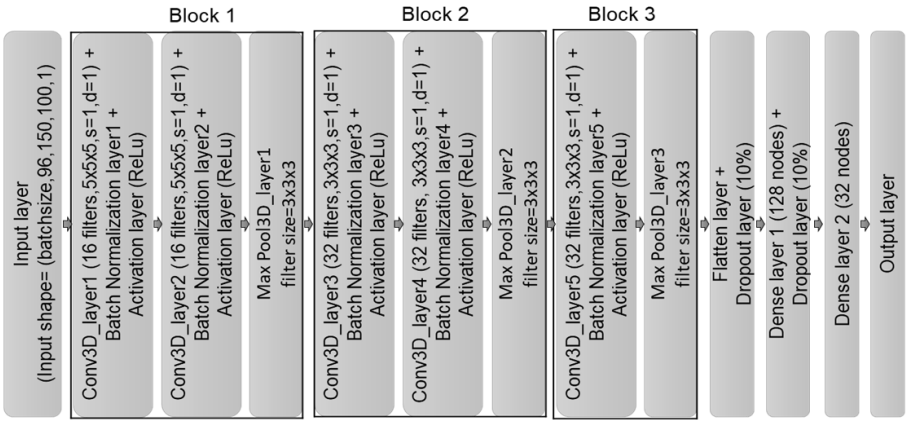


FIGURE 11. Model block diagram of 3D CNN, where s is stride and d is dilation rate.

Input gate i^t controls the accumulation of information from the new inputs in memory cell c^t. Previous cell status c^{t-1} may be "lost" if enabled by the forget gate f^t. If the output gate o^t is turned on, the most recent cell output c^t will be propagated to the final state h^t [41]. Memory cells and gated controlled information flow in the LSTM units trap the gradients and prevent them from vanishing too quickly.

The input, cell output, and states of a multivariate LSTM structure are all 1-D vectors. They are also known as fully connected LSTM (FC-LSTM) networks because of the dense connectivity between input-to-state and state-to-state transitions. The memory cell and gate computations in the LSTM cell are described by a set of operations as given as:

$$\begin{aligned} i^t &= \sigma\left(W_{xi}x^t + W_{hi}h^{t-1} + W_{ci}c^{t-1} + b_i\right) \\ f^t &= \sigma\left(W_{xf}x^t + W_{hf}h^{t-1} + W_{cf}c^{t-1} + b_f\right) \\ c^t &= f^t c^{t-1} + i^t tanh\left(W_{xc}x^t + W_{hc}h^{t-1} + b_c\right) \\ o^t &= \sigma\left(W_{xo}x^t + W_{ho}h^{t-1} + W_{co}c^{t-1} + b_o\right) \\ h^t &= o^t tanh\left(c^t\right) \end{aligned} \tag{11}$$

4) MODEL 1 (NEURAL NETWORK WITH LSTM)

In this model, spatial features are extracted from the individual frames, and extracted features are passed to the LSTM network for temporal feature processing. The output of the LSTM network is further processed through a densely connected network to classify healthy and diseased seeds.

5) MODEL 2 (CNN WITH LSTM)

In model 2, spatial features are extracted from the individual frames via different CNNs. For stage 1, pre-trained CNNs (used as a feature extractor) and CNNs trained from scratch are analyzed to see the effect on the model's performance.

6) MODEL 3 (3-D CONVOLUTIONAL NEURAL NETWORK)

In 2-D CNNs, features are extracted solely from the spatial dimension as 2-D feature maps. For bio-speckle activity analysis, capturing the variation in speckle pattern across consecutive frames is desirable. For this, we can perform 3-D convolutions in CNNs to compute features from both spatial and temporal dimensions at the same time. Convolving a 3-D kernel to the cube created by the stacked contiguous frames from the preceding layer in spatial and temporal dimensions gives 3-D spatio-temporal feature maps [42].

In 3-D convolution, the value at position (x, y, z) on the j^{th} feature map in the i^{th} layer is given by:

$$\mathrm{F}_{ij}^{xyz} = \theta\left(b_{ij} + \sum_{m}\sum_{h=0}^{H_i-1}\sum_{w=0}^{W_i-1}\sum_{d=0}^{D_i-1} w_{ijm}^{hwd}\,\mathrm{F}_{(i-1)m}^{(x+h)(y+w)(z+d)}\right), \tag{12}$$

where D_i is the size of the 3-D kernel along the temporal dimension, w_{ijm}^{hwd} is the $(h, w, d)^{th}$ value of the kernel connected to the m^{th} feature map in the previous layer.

As shown in Fig. 11, 3-D CNN model architecture has three blocks formed using 3-D convolution, batch normalization, and 3-D max pooling layers. These three blocks extract features from the spatial and temporal dimensions as 3-D feature maps. 3-D feature maps of size $(2, 4, 2)$ are obtained from the last block and converted to a 1-D vector using a flattened layer, which is further processed by classification layers to decide whether seeds are healthy or diseased.

7) MODEL 4 (CONVOLUTIONAL LSTM NETWORK)

FC-LSTM has high data redundancy for spatial data and cannot encode spatial information due to dense connections in input-to-state and state-to-state transitions. To address this, convolutional LSTM (ConvLSTM) networks with convolutional operation in both input-to-state and state-to-state transitions are proposed [43].

In ConvLSTM layer, inputs $[X_1 \ldots X_t]$, cell outputs $[C_1 \ldots C_t]$, hidden states $[H_1 \ldots H_t]$, and gates i^t, f^t, o^t

TABLE 5. Network configuration and hyperparameters details for different models.

	Stage 1	Stage 2	Stage 3
Model 1 network configuration	Temporal distributed dense layers (48 and 32 nodes), two dropout layers with 10% dropout probability	Two LSTM layers (56 and 32 nodes), dropout (10%) in input-to-state and state-to-state dense connections	Two dense layers (16 and 1 nodes)
Model 2 network configuration	Temporal distributed CNNs (pretrained and trained from scratch) followed by global average pooling2D layer	Two LSTM layers (64 and 32 nodes), dropout (15%) in input-to-state and state-to-state dense connections	Two dense layers (16 and 1 nodes)
Hyperparameters for model 1	Number of layers in stage 1 and stage 2, number of nodes in layers of stage 1 and stage 2, dropout factor, regularization factor		
Hyperparameters for model 2	Pre trained CNNs: [VGG16, Resnet50, Densenet121, Inception_V3, Efficient Net, Mobile Net, Inception_Resnet_V2, Xception],dropout factor, regularization factor		
Hyperparameters for model 3	Filters size, number of filters, dropout and regularization in 3-D convolution and dense layers		
Hyperparameters for model 4	Filters size, number of filters, dropout and regularization parameter in ConvLSTM layers, number of ConvLstm layers		

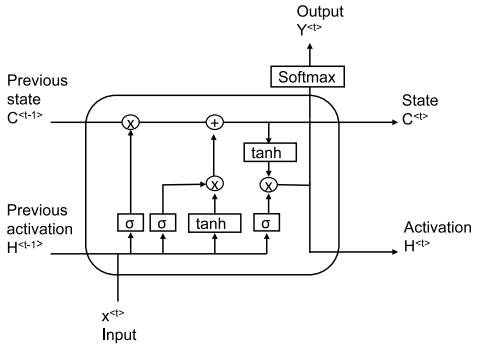


FIGURE 12. Block diagram of single LSTM cell.

are 3-D tensors with the last two dimensions being spatial (rows and columns). The key equations of ConvLSTM cell are mentioned in (13), where '$*$' denotes the convolution operator.

$$\begin{aligned} i^t &= \sigma\left(W_{xi} * x^t + W_{hi} * h^{t-1} + W_{ci} * c^{t-1} + b_i\right) \\ f^t &= \sigma\left(W_{xf} * x^t + W_{hf} * h^{t-1} + W_{cf} * c^{t-1} + b_f\right) \\ c^t &= f^t c^{t-1} + i^t tanh\left(W_{xc} * x^t + W_{hc} * h^{t-1} + b_c\right) \\ o^t &= \sigma\left(W_{xo} * x^t + W_{ho} * h^{t-1} + W_{co} * c^{t-1} + b_o\right) \\ h^t &= o^t tanh\left(c^t\right) \end{aligned} \tag{13}$$

As shown in Fig. 13, the initial two blocks in the ConvLSTM model architecture have three layers: ConvLSTM layer, batch normalization, and a 3D max pooling layer. ConvLSTM layer takes spatial data as input and returns 2-D feature maps equal to the number of filters assigned in the layer for each frame, and max pooling 3D layer reduces the spatial dimensions of feature maps.

Output from the last unit of the ConvLSTM layer in block 3, which has encoded spatial and temporal features, is utilized further after reducing the spatial dimension using the max pooling 2D layer. 2-D encoded spatial and temporal feature maps are converted to 1-D vectors via the average pooling 2D layer for the classification task.

B. MODEL PERFORMANCE EVALUATION

The model's performance is evaluated using the most common measures, including classification accuracy, precision, recall, F1-score, and area under the receiver operating characteristics curve (AUC), which are computed using a confusion matrix. Information about true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) is present in the confusion matrix.

Accuracy determines the fraction of correctly predicted classes, indicating how often the classifier is accurate. Precision shows how accurate the model is for predicted positive class and provides insight into FP, whereas recall offers a measure of FN.

Mathematical formulas for performance measures are given as:

$$Accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{14}$$

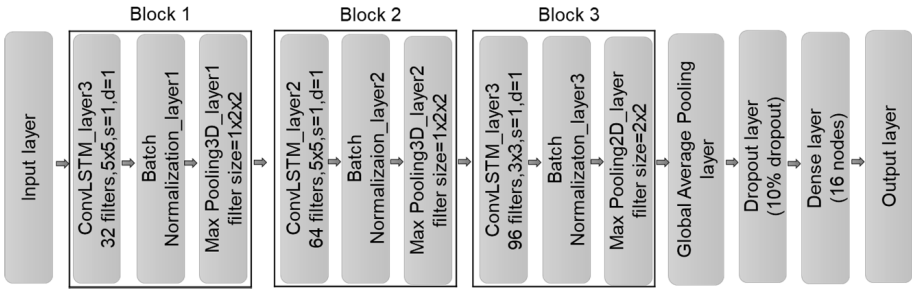


FIGURE 13. Model block diagram of Convolutional LSTM network.

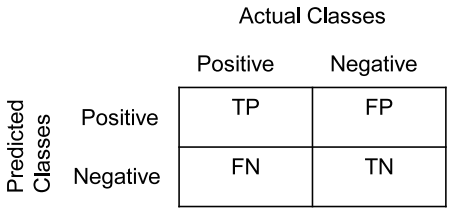


FIGURE 14. Confusion matrix for binary classification.

$$Precision(\%) = \frac{TP}{TP + FP} \times 100 \tag{15}$$

$$Recall(\%) = \frac{TP}{TP + FN} \times 100 \tag{16}$$

$$F1 - Score(\%) = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100 \tag{17}$$

$$True\ positive\ rate = \frac{TP}{TP + FN} \tag{18}$$

$$False\ positive\ rate = \frac{FP}{TN + FP} \tag{19}$$

Model predictions are the output probabilities converted to class predictions using the cut-off threshold. Performance measured using a single cut-off threshold does not provide complete insight into how well separated the classes are, and changing the cut-off thresholds changes the model performance measures. Receiver Operating Characteristics (ROC) curve is between the true positive rate (TPR) and false positive rate (FPR) at different threshold values. AUC is the area under the ROC curve from points $(0, 0)$ to $(1, 1)$. The AUC measures the classifier's ability to differentiate between classes; the higher the value of the AUC, the better the model distinguishes the positive and negative categories. AUC provides an average measurement of the model's performance over all the potential classification thresholds.

IV. RESULTS AND DISCUSSION

The experimental recorded data from setup shown in Fig. 1 comprises of 300 healthy, and 300 infected soybean seed samples. The speckle images in the recorded bio-speckle activity videos are reduced to a general frame size of 150 × 100 to remove unwanted background information and reduce computation overhead during training.

This dataset of 600 samples is divided into three splits as 70/15/15(%) (training dataset= 424 samples, validation dataset= 88 samples, test dataset = 88 samples). Each sample in the dataset has a bio-speckle activity record of 3 sec at 32 fps. For robustness analysis, four different noised test datasets are prepared using 88 samples of the test dataset as mentioned in Section II-B. TensorFlow [44], running on the Google Colab Pro platform with Tesla T4(32 GB) graphics processing unit (GPU), is used for all DL-related work.

A. PERFORMANCE AND ROBUSTNESS ANALYSIS OF DIFFERENT DEEP LEARNING MODELS

All the models mentioned in Section III-A uses the experimental dataset without any data pre-processing. DL models are trained on the training data, and their performance is evaluated after each epoch on the validation data to keep a close eye on hyperparameter tuning and overfitting of the model.

We used binary cross-entropy and Adam [45] during the training as a cost function and optimizer, respectively. For experimental data, frame size=150 × 100, frame rate=32, and frame number=96 are used. The initial learning rate for the optimizer is set to 0.001. Batch size 32, 4, 6, and 2 is used for model 1, model 2, model 3, and model 4, respectively. The test dataset performance gives the model's performance on unseen data. Table 6 summarizes the results of different models on the test and noise test datasets considered. For model 2, different pre-trained CNNs and CNNs trained from scratch are considered to explore the spatial feature extraction capabilities of different CNN architectures. The results are mentioned in Table 7 and 8.

Model 1 shows overfitting due to dense connections at all the model stages, drastically increasing the model parameters. In model 2, dense layers for spatial feature extraction are replaced by CNN, which improves the model performance, but it was not found to be robust towards noise in data. Different pre-trained CNNs and CNNs trained from scratch are considered for the first stage of the model. Results show (Table 7 and 8) that the architecture of CNN is not

TABLE 6. Performance of different models on the test and noised test datasets.

Deep learning models	Performance on test datasets						
	Original test dataset (in %)			Noised test datasets (Accuracy in %)			
	Accuracy	Precision	Recall	Gaussian Noised	Salt and Pepper Noised	Speckle Noised	Multiple Noise Added
Model 1	89.77	86.95	93.02	90.90	75.00	88.63	84.09
Model 2a[1]	94.31	91.30	97.67	77.27	56.81	57.95	86.36
Model 2b[2]	93.18	93.02	93.02	80.68	56.81	68.18	89.77
Model 3	97.72	95.55	100	96.59	51.13	57.95	93.18
Model 4	**97.72**	**100**	**95.34**	**97.72**	**94.31**	**98.86**	**96.59**

[1] Model 2a uses pre-trained Resnet50 as CNN at stage 1.
[2] Model 2b uses Resnet50 trained from scratch as CNN at stage 1.

TABLE 7. Performance of model 2 with different pretrained CNNs on test dataset and noised test datasets.

Deep learning models	Performance on test datasets						
	Original test dataset (in %)			Noised test datasets (Accuracy in %)			
	Accuracy	Precision	Recall	Gaussian Noised	Salt and Pepper Noised	Speckle Noised	Multiple Noise Added
Vgg16	93.18	91.11	95.34	87.50	52.27	69.31	90.90
Resnet50	94.31	91.30	97.67	77.27	56.81	57.95	86.36
Densenet121	90.90	87.23	95.34	51.13	47.72	52.27	52.27
Inception_V3	89.77	85.41	95.34	63.63	47.72	48.86	80.68
Efficient Net	94.31	93.18	95.34	59.09	75.00	75.00	81.81
Mobile Net	97.72	95.55	100	62.50	78.40	59.09	78.40
Inception_ Resnet_V2	47.72	0	0	44.31	50.00	36.36	43.18
Xception	93.18	91.11	95.34	79.54	47.72	48.86	85.22

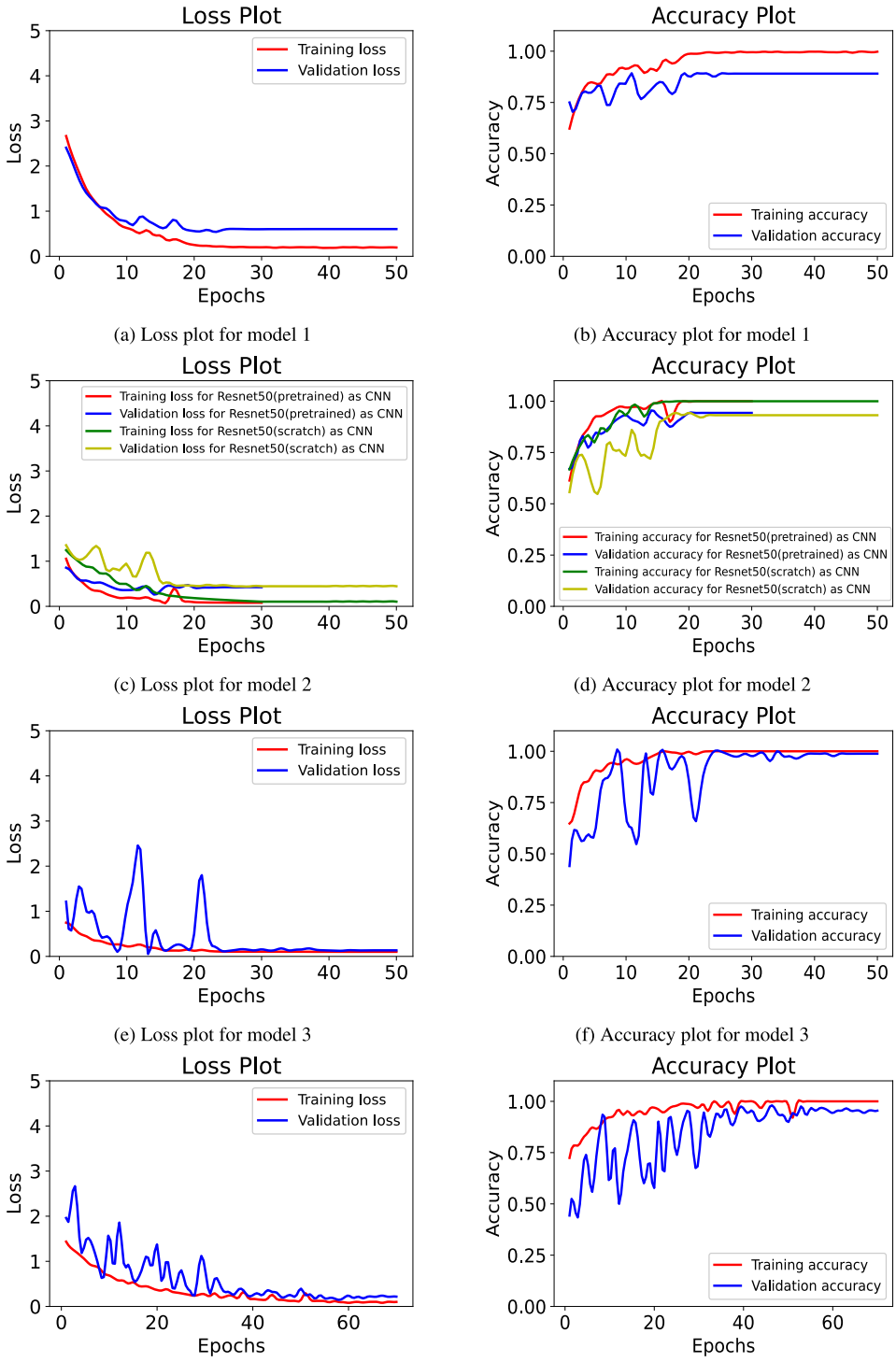
TABLE 8. Performance of model 2 with different CNNs trained from scratch on test and noised test datasets.

Deep learning models	Performance on test datasets						
	Original test dataset (in %)			Noised test datasets (Accuracy in %)			
	Accuracy	Precision	Recall	Gaussian Noised	Salt and Pepper Noised	Speckle Noised	Multiple Noise Added
Resnet50	93.18	93.02	93.02	80.68	56.81	68.18	89.77
Densenet121	93.18	93.02	93.02	78.40	48.86	55.68	95.45
Inception_V3	95.45	93.33	97.67	51.13	48.86	48.86	61.36
Mobile Net	88.63	86.66	90.69	73.86	51.13	73.86	79.54
Vgg16	51.13	0	0	51.13	51.13	51.13	51.13

responsible for the weak robustness of the model towards the noise. Spatial information loss during spatial feature extraction significantly affects the model's robustness to noise. Hence, this end-to-end model is not robust to noise. Feature extraction and their processing using convolutional operation at all stages in the ConvLSTM model prevent spatial information loss. Frame-wise spatial and temporal feature processing within a ConvLSTM layer makes the extracted features robust to noise. This model has an accuracy of 97.72%, 97.72%, 94.31%, 98.86%, and 96.59% on the test and noised test datasets, respectively (Table 6).

B. PERFORMANCE OF ConvLSTM MODEL WITH EXPERIMENTAL PARAMETER VARIATION OF DATA

Previously used bio-speckle analysis methods [12], [13] depends on the data parameters like the number of temporal frames and ROI selection from spatial dimensions for precise and accurate results. DL-based bio-speckle activity



(a) Loss plot for model 1

(b) Accuracy plot for model 1

(c) Loss plot for model 2

(d) Accuracy plot for model 2

(e) Loss plot for model 3

(f) Accuracy plot for model 3

(g) Loss plot for model 4

(h) Accuracy plot for model 4

FIGURE 15. **Loss and accuracy plots of training and validation for different models.**

analysis can overcome this dependency on experimental data parameters for accurate results. For this purpose, we have considered three significant factors: the effect of frame size, frame rate, and the number of frames. The original experimental dataset is used to create different datasets by varying these parameters, and data is normalized using z-score normalization to reduce the computational and training time of the model.

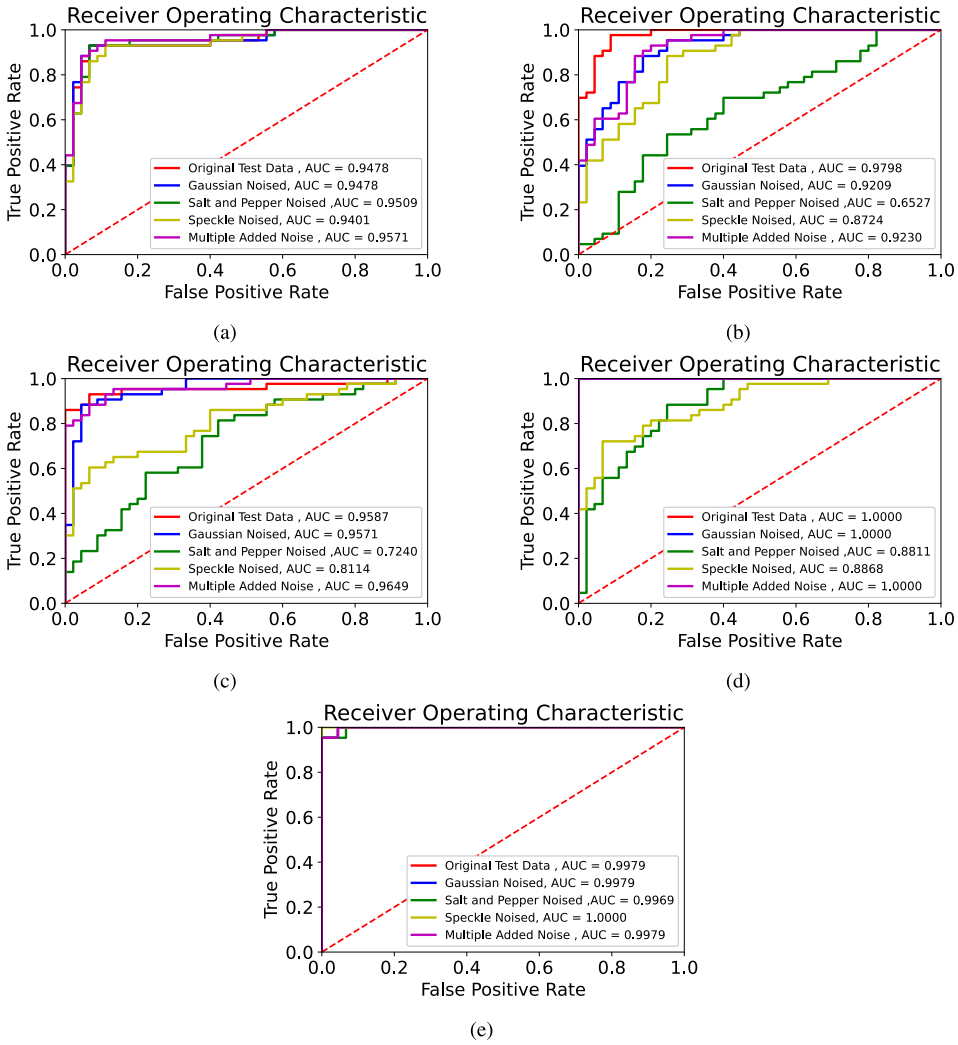


FIGURE 16. Comparison of different models using ROC plots (a) ROC plot of model 1, (b) ROC plot of model 2a, (c) ROC plot of model 2b, (d) ROC plot of model 3, (e) ROC plot of model 4.

Z-score normalization is mathematically expressed as: $Z = \frac{X - \mu}{\sigma}$, where X is original data sample, μ and σ are mean and standard deviation of data sample, respectively.

1) VARIATION IN FRAME NUMBER

The initial dataset is used to create three different datasets, keeping frame size and frame rate constant. Data parameters considered are; frame size=150 × 100, frame rate=32 fps, frames used= 96, 64, 32. Training plots and performance on test dataset are given in Fig. 19 and Table 9, respectively.

2) VARIATION IN FRAME RATE

Data parameters considered are; frame size=150 × 100, frames used = Maximum value, and frame rates= 32,16,8. Training plots and performance on test dataset are given in Fig. 18, and Table 10, respectively.

TABLE 9. Performance of ConvLSTM model on test dataset for frame number variation.

Frames used	Performance on test dataset (in %)			
	Accuracy	Precision	Recall	F1 score
96	96.59	100	93.02	96.38
64	96.59	97.61	95.34	96.46
32	97.72	97.67	97.67	97.67

TABLE 10. Performance of ConvLSTM model on test dataset for frame rate variation.

Frame rate	Performance on test dataset (in %)			
	Accuracy	Precision	Recall	F1 score
32	96.59	100	93.02	96.38
16	97.72	100	95.34	97.61
8	97.72	95.55	100	97.72

3) VARIATION IN FRAME SIZE

The parameters considered here are; frames rate= 16 fps, frames used = 48, frame size=100 × 100,150 × 100,

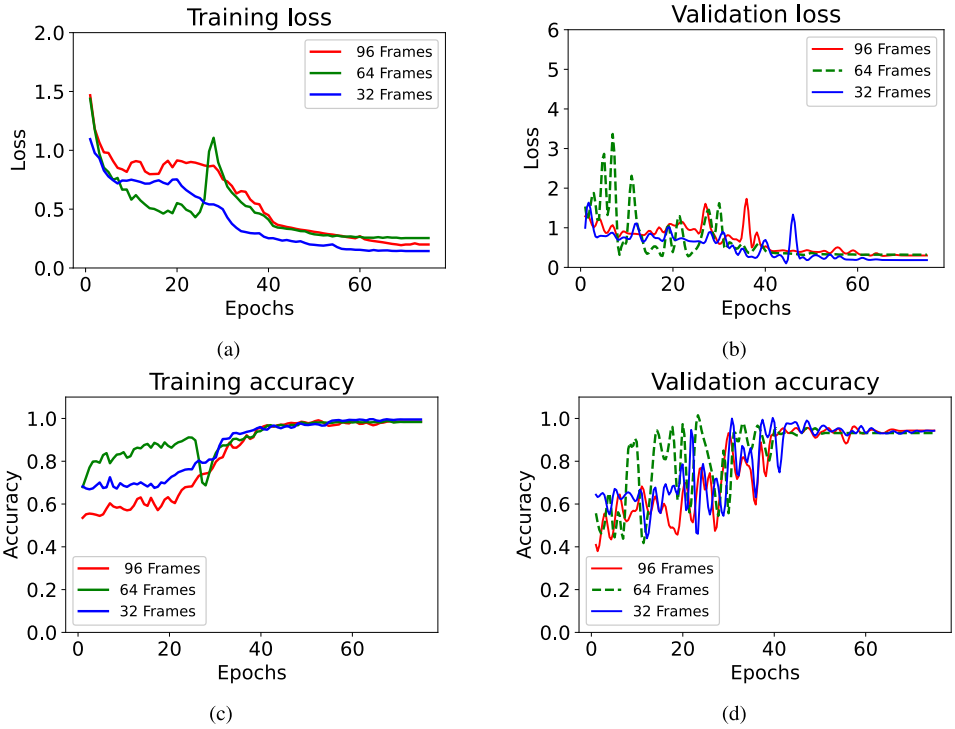


FIGURE 17. Training and validation plots for ConvLSTM model for frame number variation.

150 × 150,175 × 175. Training plots and performance on test dataset are given in Fig. 19, and Table 11, respectively.

The ConvLSTM model shows 97.72%, 97.72%, and 96.59% accuracy on the test dataset for minimum values of considered frame number, frame rate, and frame size, respectively. Results in Table 11, 9, and 10 shows that ConvLSTM model is quite robust to variation in data parameters. For all the variations considered, the model has test accuracy between 95.45% to 97.72%. The results show that only using 24 frames at 8 fps of spatial dimension 150 × 100 provides 97.72% accurate results on the test dataset. Further, this information can be used to design a computationally less expensive ConvLSTM network that takes only 24 frames at 8 fps as input, reducing the GPU computations and the time required for training.

TABLE 11. Performance of ConvLSTM on test dataset for frame size variation.

Frame size used	Performance on test dataset (in %)			
	Accuracy	Precision	Recall	F1 score
100x100	96.59	97.61	95.34	96.46
150x100	96.59	100	93.02	96.38
150x150	95.23	100	91.48	95.55
175x175	95.45	95.34	95.34	95.34

TABLE 12. Performance of ConvLSTM model on test dataset for multi-class dataset.

Classification report on test dataset				
Class name	Precision	Recall	F1-score	Test sample count
Class 0	98.00	100	99.00	46
Class 1	100	97.00	99.00	40
Class 2	100	100	100	41
Class 3	100	100	100	35
Average test accuracy = 99%				

C. SENSITIVITY OF ConvLSTM MODEL FOR BIO SPECKLE ACTIVITY OF DIFFERENT ORDERS

A sensitive model can easily identify the bio-speckle activity of different orders. We have extended the dataset to four classes with different BSAI values. Multi-class dataset has 1080 samples (Table 2) divided into three splits as 70/15/15(%). The training dataset has 756 samples, the validation dataset has 162 samples, and the test dataset has 162 samples. Each sample in the dataset has a bio-speckle activity data record of 3 sec at 32 fps. Z-score normalized data is used during the training of the model.

Categorical cross-entropy is used as a cost function with Adam optimizer (initial learning rate=0.001). The model is trained for 80 epochs with a batch size of 4 with data parameters as frame size= 150 × 100, frame number=48,

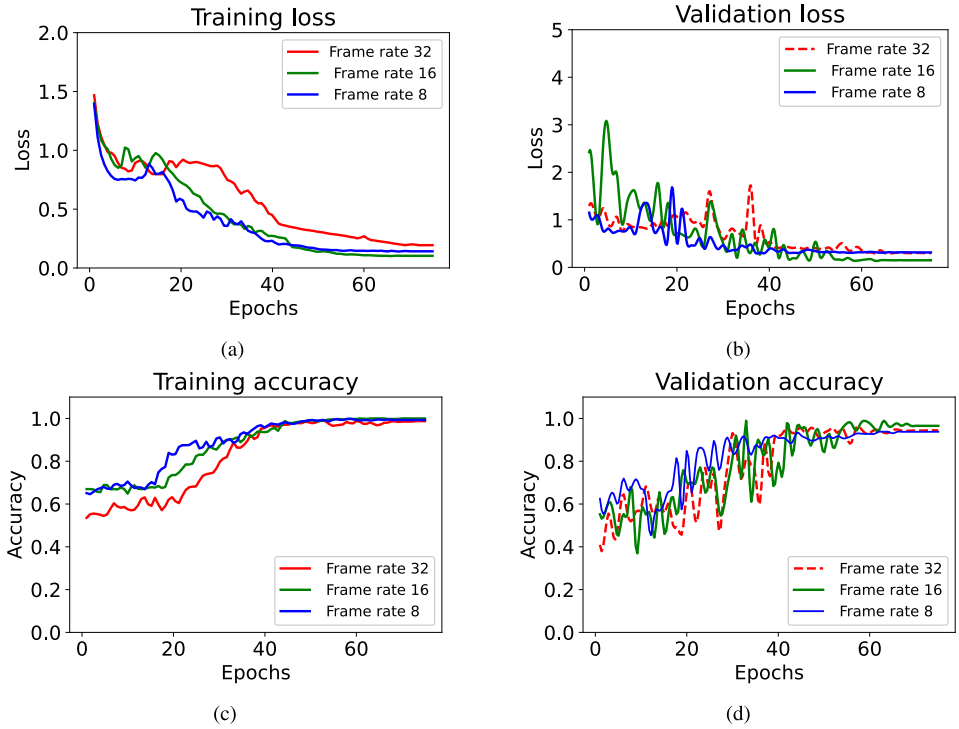


FIGURE 18. Training and validation plots for ConvLSTM model for frame rate variation.

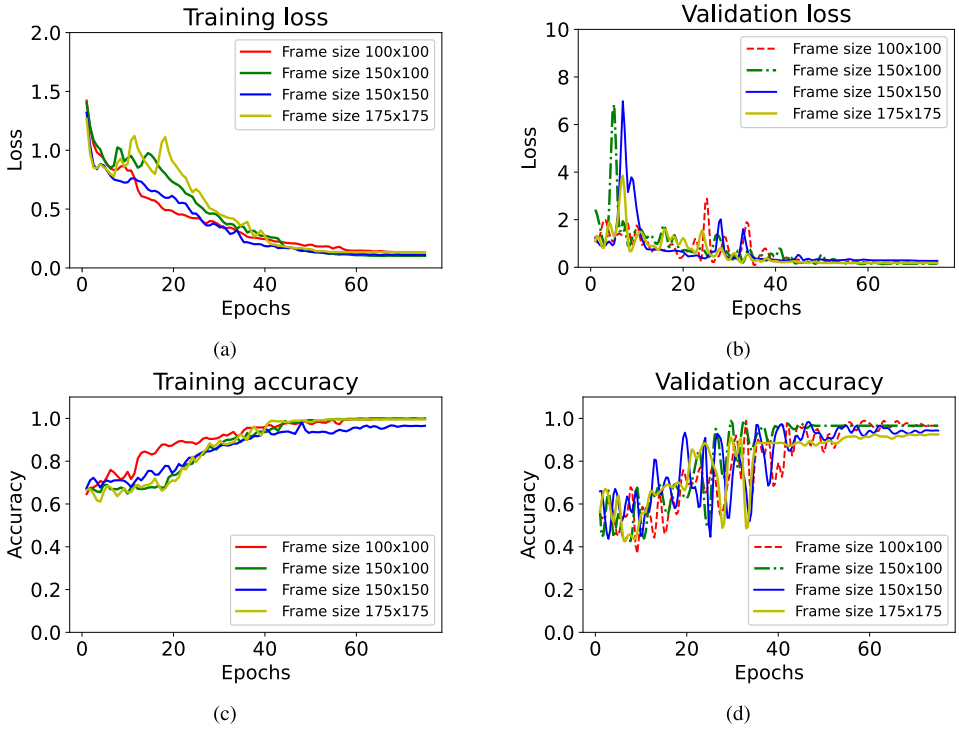
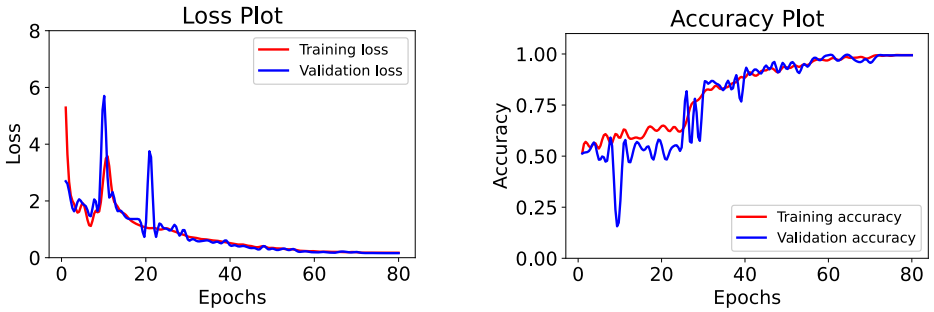


FIGURE 19. Training and validation plots for ConvLSTM model for frame size variation.

at 32fps. Training and test results are given in Fig. 20 and Table 12. ConvLSTM model gives average test accuracy of 99% for multi-class data without any hyperparameter tuning of model.



(a) (b)

FIGURE 20. Training and validation plots of ConvLSTM model on multi-class dataset.

V. CONCLUSION

Accuracy of the bio-speckle activity assessment depends upon both spatial and temporal features, as the overall activity is the cumulative effect of both these features. Among currently emerging technologies, the laser bio-speckle technique combined with deep learning can show great potential for the real-time detection of fungal infection in seeds. This study uses DL-based methods for robust spatio-temporal analysis of bio-speckle data to detect seed-borne fungal infection in soybean seeds. Among all DL models considered, the ConvLSTM model has an accuracy of 97.72% on the test dataset and is quite robust to added noises with an accuracy of 97.72%, 94.31%, 98.86%, and 96.59% on four different noised test datasets. ConvLSTM model was further tested for experimental parameter variation like frame size, frame rate, and frames used, and it is robust to the choice of experimental parameters of data used. ConvLSTM is sensitive to the detection of bio-speckle activity. When trained on multi-class data, it gives 99% test accuracy, representing that it can be potentially used to analyze the bio-speckle activity of different BSAI values. This model can be potentially used to analyze the time-varying bio-speckle activity of the sample and for seed quality assessment of different crop seeds at once in the future.

REFERENCES

[1] P. S. Thakur, V. Bhatia, L. S. Rajput, S. Rana, and S. Prakash, "Laser biospeckle technique for evaluating biotic stress on seed germination," in *Proc. Workshop Recent Adv. Photon. (WRAP)*, Mar. 2022, pp. 1–2.

[2] (2022). *International: World Oilseed Production*. [Online]. Available: http://soystats.com/international-world-oilseed-production/

[3] M. C. Pagano and M. Miransari, "The importance of soybean production worldwide," in *Abiotic and Biotic Stresses in Soybean Production*, vol. 1, 1st ed. Oxford, U.K.: Univ. Oxford, 2016, ch. 1, pp. 1–26.

[4] R. Kumar, A. Gupta, S. Srivastava, G. Devi, V. K. Singh, S. K. Goswami, M. S. Gurjar, and R. Aggarwal, "Diagnosis and detection of seed-borne fungal phytopathogens," in *Seed-Borne Diseases of Agricultural Crops: Detection, Diagnosis & Management*. Singapore: Springer, 2020, pp. 107–142.

[5] V. Mancini, S. Murolo, and G. Romanazzi, "Diagnostic methods for detecting fungal pathogens on vegetable seeds," *Plant Pathol.*, vol. 65, no. 5, pp. 691–703, Jun. 2016.

[6] T. A. Matveyeva, R. M. Sarimov, A. V. Simakin, M. E. Astashev, D. E. Burmistrov, V. N. Lednev, P. A. Sdvizhenskii, M. Y. Grishin, S. M. Pershin, N. O. Chilingaryan, N. A. Semenova, A. S. Dorokhov, and S. V. Gudkov, "Using fluorescence spectroscopy to detect rot in fruit and vegetable crops," *Appl. Sci.*, vol. 12, no. 7, p. 3391, Mar. 2022.

[7] M.-Y. Shin, C. G. Viejo, E. Tongson, T. Wiechel, P. W. J. Taylor, and S. Fuentes, "Early detection of Verticillium wilt of potatoes using near-infrared spectroscopy and machine learning modeling," *Comput. Electron. Agricult.*, vol. 204, Jan. 2023, Art. no. 107567.

[8] C. Baratto, G. Ambrosio, G. Faglia, and M. Turina, "Early detection of Esca disease in asymptomatic vines by Raman spectroscopy," *IEEE Sensors J.*, vol. 22, no. 23, pp. 23286–23292, Dec. 2022.

[9] A. Soni, Y. Dixit, M. M. Reis, and G. Brightwell, "Hyperspectral imaging and machine learning in food microbiology: Developments and challenges in detection of bacterial, fungal, and viral contaminants," *Comprehensive Rev. Food Sci. Food Saf.*, vol. 21, no. 4, pp. 3717–3745, Jul. 2022.

[10] F. Tao, H. Yao, Z. Hruska, L. W. Burger, K. Rajasekaran, and D. Bhatnagar, "Recent development of optical methods in rapid and non-destructive detection of aflatoxin and fungal contamination in agricultural products," *TrAC Trends Anal. Chem.*, vol. 100, pp. 65–81, Mar. 2018.

[11] A. Wu, J. Zhu, and T. Ren, "Detection of apple defect using laser-induced light backscattering imaging and convolutional neural network," *Comput. Electr. Eng.*, vol. 81, Jan. 2020, Art. no. 106454.

[12] A. Zdunek, A. Adamiak, P. M. Pieczywek, and A. Kurenda, "The biospeckle method for the investigation of agricultural crops: A review," *Opt. Lasers Eng.*, vol. 52, pp. 276–285, Jan. 2014.

[13] A. Chatterjee, P. Singh, V. Bhatia, and S. Prakash, "An efficient automated biospeckle indexing strategy using morphological and geo-statistical descriptors," *Opt. Lasers Eng.*, vol. 134, Nov. 2020, Art. no. 106217.

[14] R. A. Braga Jr., G. F. Rabelo, L. R. Granato, E. F. Santos, J. C. Machado, R. Arizaga, H. J. Rabal, and M. Trivi, "Detection of fungi in beans by the laser biospeckle technique," *Biosystems Eng.*, vol. 91, no. 4, pp. 465–469, Aug. 2005.

[15] G. F. Rabelo, A. M. Enes, R. A. B. Junior, and I. M. D. Fabbro, "Frequency response of biospeckle laser images of bean seeds contaminated by fungi," *Biosyst. Eng.*, vol. 110, no. 3, pp. 297–301, Nov. 2011.

[16] R. A. Braga, G. W. Horgan, A. M. Enes, D. Miron, G. F. Rabelo, and J. B. B. Filho, "Biological feature isolation by wavelets in biospeckle laser images," *Comput. Electron. Agricult.*, vol. 58, no. 2, pp. 123–132, Sep. 2007.

[17] R. R. Cardoso, A. G. Costa, C. M. B. Nobre, and R. A. Braga, "Frequency signature of water activity by biospeckle laser," *Opt. Commun.*, vol. 284, no. 8, pp. 2131–2136, Apr. 2011.

[18] J. A. Alves, R. A. Braga, and E. V. D. B. V. Boas, "Identification of respiration rate and water activity change in fresh-cut carrots using biospeckle laser and frequency approach," *Postharvest Biol. Technol.*, vol. 86, pp. 381–386, Dec. 2013.

[19] D. B. Sutton and Z. K. Punja, "Investigating biospeckle laser analysis as a diagnostic method to assess sprouting damage in wheat seeds," *Comput. Electron. Agricult.*, vol. 141, pp. 238–247, Sep. 2017.

[20] P. Singh, A. Chatterjee, L. S. Rajput, S. Rana, S. Kumar, V. Nataraj, V. Bhatia, and S. Prakash, "Development of an intelligent laser biospeckle system for early detection and classification of soybean seeds infected with seed-borne fungal pathogen (*Colletotrichum truncatum*)," *Biosyst. Eng.*, vol. 212, pp. 442–457, Dec. 2021.

[21] P. S. Thakur, A. Kumar, B. Tiwari, B. Gedam, V. Bhatia, S. Rana, and S. Prakash, "Machine learning based biospeckle technique for identification of seed viability using spatio-temporal analysis," in *Proc. Workshop Recent Adv. Photon. (WRAP)*, Mar. 2022, pp. 1–2.

[22] P. S. Thakur, B. Tiwari, A. Kumar, B. Gedam, V. Bhatia, O. Krejcar, M. Dobrovolny, J. Nebhen, and S. Prakash, "Deep transfer learning based photonics sensor for assessment of seed-quality," *Comput. Electron. Agricult.*, vol. 196, May 2022, Art. no. 106891.
[23] T. R. Boufleur, M. Ciampi-Guillardi, Ì. Tikami, F. Rogério, M. R. Thon, S. A. Sukno, N. S. M. Junior, and R. Baroncelli, "Soybean anthracnose caused by colletotrichum species: Current status and future prospects," *Mol. Plant Pathol.*, vol. 22, no. 4, pp. 393–409, Apr. 2021.
[24] P. S. Thakur, A. Chatterjee, L. S. Rajput, S. Rana, V. Bhatia, and S. Prakash, "Laser biospeckle technique for characterizing the impact of temperature and initial moisture content on seed germination," *Opt. Lasers Eng.*, vol. 153, Jun. 2022, Art. no. 106999.
[25] Y. Fan, L. Zhang, H. Guo, H. Hao, and K. Qian, "Image processing for laser imaging using adaptive homomorphic filtering and total variation," *Photonics*, vol. 7, no. 2, p. 30, Apr. 2020.
[26] K. Basak, M. Mahadevappa, and P. K. Dutta, "Multiscale noise-adaptive homomorphic filtering based speckle denoising in laser speckle imaging," in *Proc. 4th Nat. Conf. Comput. Vis., Pattern Recognit., Image Process. Graph. (NCVPRIPG)*, Dec. 2013, pp. 1–5.
[27] M. Konnik and J. Welsh, "High-level numerical simulations of noise in CCD and CMOS photosensors: Review and tutorial," 2014, *arXiv:1412.4031*.
[28] A. K. Boyat and B. K. Joshi, "A review paper: Noise models in digital image processing," 2015, *arXiv:1505.03489*.
[29] A. L. D. Pra, G. J. Meschino, M. N. Guzmán, A. G. Scandurra, M. A. González, C. Weber, M. Trivi, H. Rabal, and L. I. Passoni, "Dynamic speckle image segmentation using self-organizing maps," *J. Opt.*, vol. 18, no. 8, Aug. 2016, Art. no. 085606.
[30] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[31] I. El-Feghi, A. Tahar, and M. Ahmadi, "Efficient features extraction for fingerprint classification with multi layer perceptron neural network," in *Proc. Int. Symp. Signals, Circuits Syst.*, Jun. 2011, pp. 1–4.
[32] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Comput. Electron. Agricult.*, vol. 173, Jun. 2020, Art. no. 105393.
[33] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[36] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
[37] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. AAAI Conf. Artif. Intell.*, 2017, vol. 31, no. 1.
[38] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2019, pp. 6105–6114.
[39] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.
[40] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.
[41] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[42] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013.
[43] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 28, 2015, pp. 1–9.
[44] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, *arXiv:1603.04467*. [Online]. Available: https://www.tensorflow.org/
[45] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

NIKHIL KALER received the B.E. degree in electrical engineering from Jai Narain Vyas University, Jodhpur, India, in 2020. He is currently pursuing the M.S. degree with the Department of Electrical Engineering, Indian Institute of Technology (IIT) Indore, India. His research interests include computer vision and deep learning applications in bio-medical and photonics.

VIMAL BHATIA (Senior Member, IEEE) received the Ph.D. degree from the Institute for Digital Communications, The University of Edinburgh, Edinburgh, U.K., in 2005.

He is currently a Professor with the Indian Institute of Technology (IIT) Indore, India, and is an Adjunct Faculty Member with IIT Delhi and IIIT Delhi, India. During Ph.D., he also received the IEEE fellowship for collaborative research at the Department of Systems and Computer Engineering, Carleton University, Canada, and is a Young Faculty Research Fellow with MeitY, Government of India. He was with various IT companies for over 11 years both in India and U.K. He is a PI/a co-PI/coordinator for external projects with funding of over U.S. $20 million from MeitY, DST, UKIERI, MoE, AKA, IUSSTF, and KPMG. He is currently a DRISHTI CPS Chair Professor at IIT Indore. He is also affiliated with the School of Electronic and Information Engineering, Soochow University, China, as an Honorary Chair Professor. He has more than 350 peer-reviewed publications and has filed 13 patents (with five granted). He has supervised 18 awarded Ph.D. theses. His research interests are in the broader areas of communications, non-Gaussian non-parametric signal processing, machine/deep learning with applications to communications, and photonics. He is a reviewer for IEEE, Elsevier, Wiley, Springer, and IET. He is currently a fellow of IETE and FOSI. He is also a certified SCRUM master. He has delivered many talks, tutorials, and conducted faculty development programs for the World Bank's NPIU TEQIP-III and was invited to talk at WWRF46-Paris. He was a recipient of the Prof. SVC Aiya Memorial Award, in 2019. He is currently an Associate Editor of *IETE Technical Review*, *Frontiers in Communications and Networks*, *Frontiers in Signal Processing*, and IEEE WIRELESS COMMUNICATIONS LETTERS. He has been included in the World's Top 2% most influential Scientists by Stanford University.

AMIT KUMAR MISHRA (Senior Member, IEEE) received the Ph.D. degree from The University of Edinburgh. His current research interests include radar and sensor system design and applied machine learning. He has successfully supervised ten Ph.D. students so far and holds five patents.