# A Deep Multi-scale Feature Fusion Approach for Early Recognition of Jute Diseases and Pests

**Rashidul Hasan Hridoy, Tanjina Yeasmin, and Md. Mahfuzullah**

**Abstract** Diseases and pests of jute hinder the quality production of fiber which is a malignant threat to the jute industry, causing severe financial losses to cultivators. Early recognition of diseases and pests of jute plant is highly vital for preventing the spread of diseases and pests which will ensure the quality improvement of the jute industry. This paper addresses a robust hybrid model, namely JuteNet, is a multi-scale feature fusion approach for early recognition of jute diseases and pests. First, a dataset of 56,108 images of jute leaves and stems is generated. Afterward, the fusion of extracted features from images by deep neural networks such as Xception, InceptionResNetV2, and InceptionV3 was used to develop JuteNet that obtained 99.47% accuracy in recognizing 2803 images of six classes of the testing set. Moreover, Xception, InceptionResNetV2, and InceptionV3 separately acquired 91.83, 96.11, and 98.86% of test accuracy, which validates the recognition efficiency of JuteNet.

**Keywords** Computer vision · Deep neural networks · Transfer learning · Feature fusion · Jute · Leaf disease recognition

## 1 Introduction

Jute (scientific name: *Corchorus capsularis*) is also called as the golden fiber which is widely cultivated in Bangladesh, India, China, Uzbekistan, and Nepal and provides affordable and eco-friendly fiber that has massive demand all over the world. In respect to production and global consumption, the fiber of jute is the second only to cotton, which is considered as a commercial, industrial, and economically crucial

R. H. Hridoy (✉) · T. Yeasmin · Md. Mahfuzullah
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
e-mail: rashidul15-8596@diu.edu.bd

T. Yeasmin
e-mail: tanjina15-758@diu.edu.bd

Md. Mahfuzullah
e-mail: mahfuzullah15-11006@diu.edu.bd



V. Suma et al. (eds.), *Inventive Systems and Control*, Lecture Notes in Networks and Systems 436, https://doi.org/10.1007/978-981-19-1012-8_37

sector of Bangladesh [1]. Jute is highly affected by several kinds of diseases and pests during cultivation, which decreases the quality of jute fiber significantly. Leaf mosaic virus, leaf curl, and stem rot are the most destructive diseases of jute. On the contrary, semilooper and hairy caterpillar are the commonly found pests of jute leaf [2]. Visual recognition technique of diseases and pests of jute through analyzing symptoms is a troublesome and time-consuming approach, but the accuracy of recognition does not meet the requirement due to lack of experience about diseases and pests. To ensure sustainable development of the jute industry, an automated approach is crucial for efficient and rapid recognition of jute diseases and pests at an initial stage.

For gaining better yields and controlling diseases and pests, several technologies were been utilized in agriculture for many years. In recent years, convolutional neural networks (CNNs) brought crucial changes in computer vision and significantly improves the efficiency of image recognition approaches [3]. Several machine learning (ML) algorithms also were utilized in recognition approaches, where preprocessing of images and feature extraction were very time-consuming tasks. Deep neural networks can learn features directly from input images that save effort and time, and also achieve higher accuracy than ML algorithms [4, 5]. Several studies were conducted using CNNs and ML algorithms in the agricultural sector for identification of plant species, diseases, and pests. However, the accurate classification of plant diseases and pests is a more challenging task than generic image recognition for the complex background of images. Fusion of extracted features provides more accurate recognition performance by merging the extracted features acquired by numerous CNNs into a hybrid recognition model [6].

A dataset of 56,108 images of six classes of jute leaves and stems, namely JL dataset, is generated in this study using twelve image augmentation approaches from 4316 collected images from several cultivations fields. The JL dataset includes 44,889 training, 8416 validation, and 2803 testing images. This research work aims to reveal an efficient recognition approach for early recognition of jute diseases and pests using the fusion of multi-scale features extracted by CNNs. In the introduced JuteNet framework, pretrained CNNs such as Xception, InceptionResNetV2, and InceptionV3 were utilized using the transfer learning technique for the extraction of features from jute leaf and stem images. All extracted features by CNNs were fed into a concatenation layer that concatenates all inputs into a single tensor. Lastly, the fusion of extracted features was fine-tuned in this study using three batch normalization (BM) layers, three dropout layers, and three fully connected (FC) layers, where the last FC layer is connected to the Softmax activation function with six neurons for classifying six classes of the JL dataset. Moreover, for comparing the recognition efficiency of the proposed hybrid model with single models, Xception, InceptionResNetV2, and InceptionV3 were also trained separately in this study. Several experimental studies were conducted for evaluating the recognition efficiency of JuteNet and single models using 2803 images of the testing set. According to experimental results, single state-of-the-art CNNs such as Xception, InceptionResNetV2, and InceptionV3 wrongly recognized 229, 109, and 32 images of the testing of the JL dataset, where the proposed JuteNet model misclassified 15 images which demonstrate that the addressed hybrid model can classify diseases and pests of jute more efficiently than single CNNs. However, JuteNet acquired 99.47% accuracy, and pretrained CNNs include Xception, InceptionResNetV2, and InceptionV3 acquired 91.83%, 96.11%, and 98.86% accuracy, respectively, on the testing set of the JL dataset, and the major contributions of this research work are as follows:

- Till now no suitable dataset of jute leaf and stem is available, and an enhanced dataset of 56,108 images is developed for the purpose of this study.
- An effective and robust hybrid model is introduced for early recognition of diseases and pests of jute using the fusion of multi-scale features.
- The recognition performance of single pretrained CNNs and the proposed hybrid model was evaluated using 2803 images of the testing set of the JL dataset.

The rest of this paper is formed as follows: Sect. 2 represents the literature review. Section 3 describes the JL dataset and the proposed framework in detail. Experimental studies are described in Sect. 4. The results obtained and their interpretations are demonstrated in Sect. 5. Finally, the conclusion and future work are given in Sect. 6.

## 2 Related Work

A remarkable number of approaches were conducted in the literature for identifying diseases and pests of several plants using CNNs and ML algorithms. Hridoy et al. addressed a computer vision method for recognizing yellow mosaic disease of the black gram using CNN which obtained 97.11% accuracy, and the performance of the addressed CNN architecture was compared with the recognition performance of three pretrained CNN models such as AlexNet, VGG16, and InceptionV3, acquired 93.78%, 95.49%, and 96.67% accuracy, respectively [3]. Hasan et al. addressed a recognition method by utilizing CNN to classify jute diseases and acquired 96.00% accuracy [4]. The classification performance of the addressed CNN architecture was compared with the performance of support vector machine (SVM), k-nearest neighbors (kNN), and random forest (RF) classifier, and these three classifiers obtained 89.00%, 86.00%, and 80.00% accuracy, respectively. Reza et al. addressed a detection method for classifying stem-based diseases of jute by utilizing the multiclass-SVM (M-SVM) classifier, obtained 86.00% accuracy with hue-based segmentation and 60.00% accuracy without hue-based segmentation [5]. For removing noise from images, the bilateral filter approach was utilized and thirteen features were extracted from images using the color co-occurrence technique for performing texture analysis. After analyzing the recognition performance of nine classifiers, Habib et al. addressed a recognition method for classifying jackfruit diseases by utilizing the RF classifier which obtained 89.59% accuracy [7]. For segmenting images, k-means clustering was utilized, and ten features were extracted from images that were fed to classifiers. Sholihati et al. addressed a classification approach using VGG16 for classifying diseases of potato leaf after analyzing the classification performance of VGG16 and VGG19 model, acquired 91.31% and 90.96% accuracy, respectively, and VGG16 consumed less training time than VGG19 [8]. For classifying diseases and pests of paddy leaf, Senan et al. addressed an automated classification method using CNN that acquired 90.30% accuracy with 50:50 training and testing image ratio and 96.60% accuracy with 70:30 training and testing image ratio [9]. However, multilayer perceptron (MLP), SVM, and artificial neural network (ANN) acquired 81.12, 81.45, and 82.60% accuracy with a 70:30 training and testing image ratio. Bhowmik et al. introduced a detection approach using CNN that contains three layers for classifying diseases of tea leaves and acquired 95.94% accuracy after 22 epochs of training [10]. Both classification performance and training time of the proposed CNN architecture were increased as the number of epochs grew during the training phase. Hridoy et al. addressed a recognition approach using CNN that was built with depthwise separable convolutions to classify diseases of betel leaf at an initial stage and acquired 96.02% accuracy [11]. The performance of three activation functions such as rectified linear unit (ReLU), scaled exponential linear unit (SELU), and Swish was analyzed, and Swish performed superior than others. Hussain et al. introduced a recognition approach for classifying diseases of cucumber leaf using the fusion of deep features with the whale optimization algorithm which was used for selecting the best features and obtained 96.50% accuracy with 45.28 s computational time [12]. VGG16 and InceptionV3 were used for extracting features from images, and the best features were classified using algorithms of supervised learning. Trang et al. addressed an identification approach for mango diseases using CNN that acquired 88.46% accuracy where three pretrained models such as InceptionV3, AlexNet, and MobileNetV2 obtained 78.48, 76.92, and 84.62% accuracy [13]. For enhancing the quality of images, rescaling and center alignment were used, and the golden section search technique was utilized to enhance the contrast. Fenu et al. introduced a multioutput learning approach for diagnosing diseases of plant and severity of stress, and five pretrained CNN models such as VGG16, VGG19, ResNet50, InceptionV3, MobileNetV2, and EfficientNetB0 were utilized [14]. Inception V3 performed better than others in diagnosing biotic stress that obtained 90.68% accuracy, where EfficientNetB0 performed better than others in diagnosing severity which acquired 78.31% accuracy. The training time of CNNs was also analyzed, and EfficientNetB0 consumed less training time than others. Zaki et al. addressed an automated approach for classifying diseases of tomato leaves by utilizing MobileNetV2 and acquired 95.94% accuracy with a batch size of 16 [15]. The performance of five optimization methods and three learning rates was analyzed where the Adagrad optimization method and the learning rate of 0.0001 performed better than others. Besides disease and pest recognition of plants, CNNs and ML algorithms were also used in several research works [16–18].

According to abovementioned research works, CNNs obtained remarkable results in recognizing plant disease and pests, and ML algorithms obtained less accuracy than CNNs. However, hybrid model is rarely utilized for identifying diseases and pests of plant. Hence, a proficient framework based on the fusion of multi-scale features is addressed in this research to recognize initial infection of jute diseases and pests.

# 3 Materials and Methods

## *3.1 JL Dataset*

A significant number of days was devoted to acquire 4316 images of jute leaves and stems, and these images were acquired with complex backgrounds after five days of initial infections. Utilizing position augmentation techniques include cropping, rotation, and flipping, and color augmentation techniques include brightness, contrast, saturation, and hue, and JL dataset of 56,108 images is generated from 4316 collected images. The JL dataset summary is given in Table 1, and Fig. 1 represents samples of six different classes.

**Table 1** The JL dataset summary

| Class | Training images | Validation images | Testing images | Total images |
|---|---|---|---|---|
| Healthy leaf | 14,508 | 2721 | 906 | 18,135 |
| Leaf mosaic virus | 7094 | 1329 | 443 | 8866 |
| Leaf curl | 5096 | 956 | 318 | 6370 |
| Stem rot | 4430 | 832 | 276 | 5538 |
| Semilooper | 7791 | 1459 | 487 | 9737 |
| Hairy caterpillar | 5970 | 1119 | 373 | 7462 |


**Fig. 1** Instances of JL dataset: (1) healthy leaf, (2) leaf mosaic virus, (3) leaf curl, (4) stem rot, (5) semilooper, and (6) hairy caterpillar


**Fig. 2** Samples of image enhancement approaches: (1) high brightness, (2) low brightness, (3) high contrast, (4) low contrast, (5) high saturation, (6) low saturation, (7) cropping, (8) rotation 90°, (9) rotation 180°, (10) rotation 270°, (11) vertical flip, and (12) horizontal flip

All images of the JL dataset were split into 3 sets randomly which include training, validation, and testing set. Twelve image enhancement approaches were utilized in this study, which are presented in Fig. 2. Image enhancement approaches reduce the overfitting issue significantly and help CNNs in enhancing the anti-interference ability during the training phase [3].

## *3.2 JuteNet*

State-of-the-art CNNs are now widely utilized in the field of computer vision for recognizing images with higher accuracy than classical approaches. Pretrained CNN models, including Xception, InceptionResNetV2, and InceptionV3, were utilized in this study for constructing the framework of JuteNet using the transfer learning approach. The approach of multi-scale feature fusion was used in JuteNet where Xception, InceptionResNetV2, and InceptionV3 were used for extracting features from images of the JL dataset.

Xception is a 71-layer deep pretrained CNN model and builds with layers of depthwise separable convolution, and $299 \times 299$ pixels images are required to train this architecture [19]. The network of Xception consists of three flows which include entry, middle, and exit, where data enters into the network via entry flow, then enters middle flow that is repeated eight times, and exits via exit flow.

InceptionResNetV2 was developed with modules of hybrid Inception-ResNet which contains 164 layers, and the input size of the image of this architecture is 299 × 299 pixels [20]. The reduction module was utilized in all hybrid Inception-ResNet modules for reducing the presentation dimension.

InceptionV3 is a 48-layer deep pretrained CNN model, and 299 × 299 pixels images are required to train this architecture, which is built with symmetric and asymmetric building blocks such as convolutions, max pooling, average pooling, dropouts, CL, and FC layers [21]. BM layers are widely applied in this architecture, and for reducing the parameters and connection numbers without decreasing the recognition ability of the model, factorization is utilized in this model.

In this study, pretrained CNN models separately extracted features from the images of jute leaf and stem, and used the global average pooling (GAP) 2D layer to flatten the output of functional CNN models into a vector that calculates the mean of input channels. Afterward, a concatenate layer (CL) was used for combining individual vectors of CNNs to generate a single vector. The framework of JuteNet is graphically represented in Fig. 3. After CL, three blocks were used in JuteNet framework for fine-tuning the output of CL to recognize six classes of the JL dataset, and each block contains one dropout, one BM, and one FC layer. For overcoming the overfitting issue during the training phase, three dropout layers were used in the JuteNet framework. BM layers were used for making the JuteNet framework faster and stable. The FC layer of the last block was connected with the Softmax activation function with six neurons.

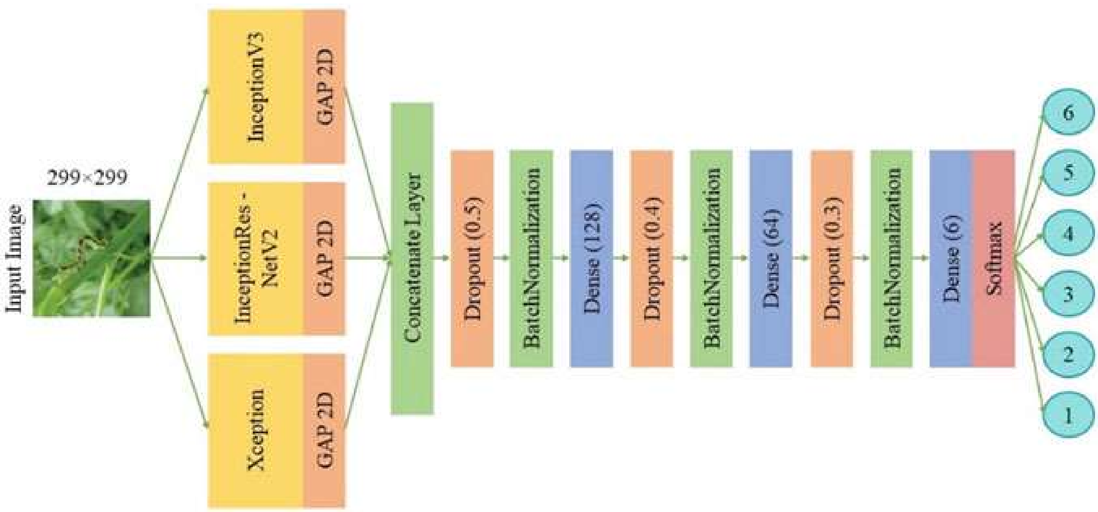


**Fig. 3** The schematic representation of the JuteNet framework

## 4 Experiments

The JuteNet framework was trained via the transfer learning strategy in this research with GPU support of Google Colab. The transfer learning approach assists networks in decreasing training time, also enhancing the baseline performance of the framework. 44,889 images of the training set and 8416 images of the validation set were used in this study to train and fit the introduced JuteNet framework. Randomly chosen 2803 images of the testing set were used for examining the efficiency of JuteNet. During the training stage of JuteNet, Adam was used as an optimization method with a learning rate of 0.0001. Categorical cross-entropy was used as a loss function, batch size was set to 32, and the epoch's number was set to 50 during the training of the JuteNet framework. During the training phase of the JuteNet network, no major fluctuations were seen in curves of training and validation accuracy and loss. Besides the JuteNet framework, single CNNs include Xception, InceptionResNetV2, and InceptionV3 were trained separately using the JL dataset in this research for comparing the efficiency of the individual CNNs with the fusion of multi-scale approach.

To examine the efficiency of the JuteNet framework and single models used in this research, statistical parameters including sensitivity (sen), specificity (spe), accuracy (acc), and precision (pre) were utilized which were computed using true positive (TP), true negative (TN), false positive (FP), and false negative (FN), and statistical parameters are given below in between Eq. 1 and 4, where *ju* represents the number of classes of the testing set of the JL dataset. Here, TP represents the summation of correctly recognized images in each class of the testing set, TN presents the summation of correctly recognized images except the relevant class, FP presents the summation of falsely predicated images in other classes except the relevant class, and FN presents the summation of falsely recognized images in the relevant class [3].

For a class *ju*,

$$\mathrm{Pre}(ju) = \frac{\mathrm{TP}(ju)}{\mathrm{TP}(ju) + \mathrm{FP}(ju)} \tag{1}$$

$$\mathrm{Rec}(ju) = \frac{\mathrm{TP}(ju)}{\mathrm{TP}(ju) + \mathrm{FN}(ju)} \tag{2}$$

$$\mathrm{Spe}(ju) = \frac{\mathrm{TN}(ju)}{\mathrm{TN}(ju) + \mathrm{FP}(ju)} \tag{3}$$

$$\mathrm{Acc}(ju) = \frac{\mathrm{TP}(ju) + \mathrm{TN}(ju)}{\mathrm{TP}(ju) + \mathrm{TN}(ju) + \mathrm{FP}(ju) + \mathrm{FN}(ju)} \tag{4}$$

## 5 Result and Discussion

This study is conducted for developing an efficient recognition approach for initial identification of diseases and pests of jute, where the fusion of multi-scale features was used to develop a hybrid robust model, namely JuteNet, that acquired 99.47% test accuracy. On the contrary, single CNNs include Xception, InceptionResNetV2, and InceptionV3 individually acquired 91.83%, 96.11%, and 98.86% accuracy, respectively, in recognizing 2803 images of the test set which contains six classes. Four normalized confusion matrices which were obtained after multiclass classification of JuteNet and three other CNNs were given in Fig. 4, where the normalized confusion matrix of JuteNet framework represented its significant recognition efficiency.

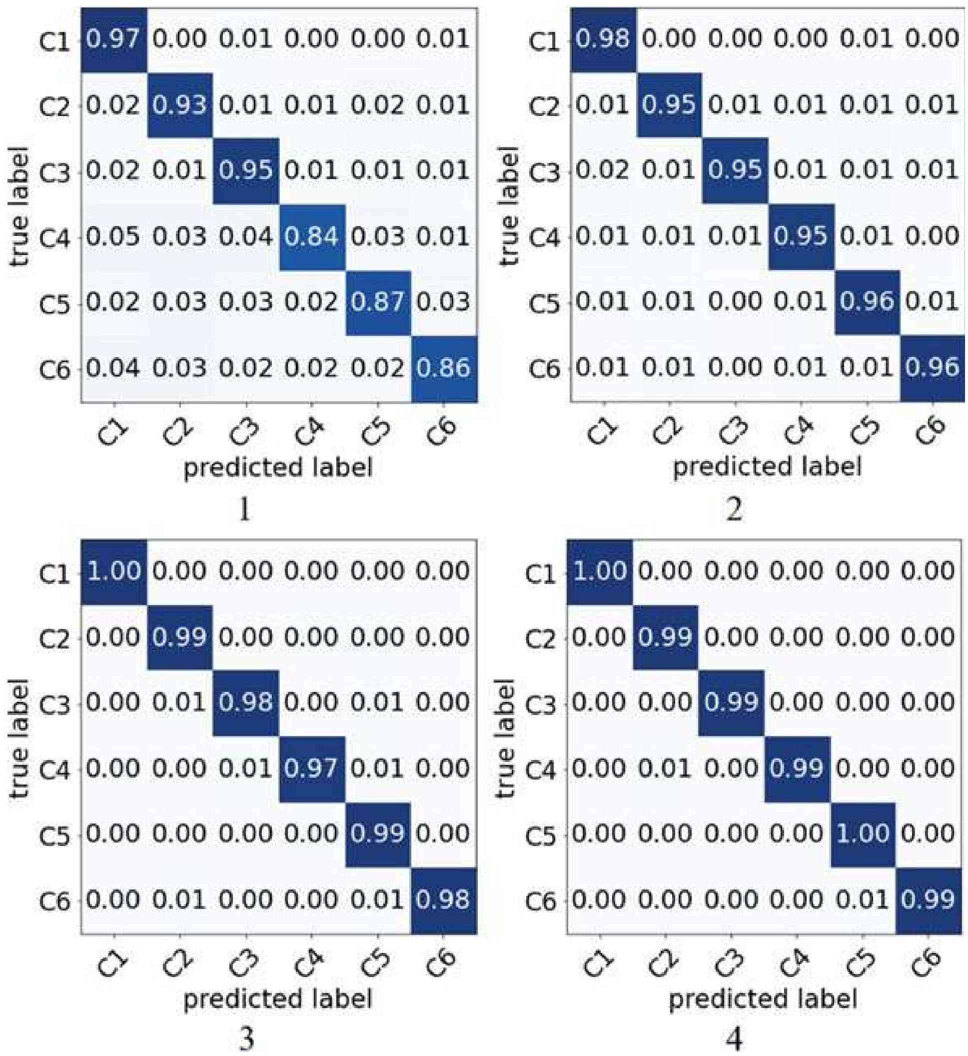


**Fig. 4** Normalized confusion matrix: (1) Xception, (2) InceptionResNetV2, (3) InceptionV3, (4) JuteNet, and (C1) healthy leaf, (C2) leaf mosaic virus, (C3) leaf curl, (C4) stem rot, (C5) semilooper, (C6) hairy caterpillar

To illustrate the classification efficiency of single CNNs and JuteNet more clearly, class-wise recognition performance was evaluated using statistical parameters. Xception acquired the highest sensitivity and precision value in the healthy leaf class, 94.84%, and 97.46%, respectively. In the leaf curl class, it obtained higher specificity and accuracy values than other classes, 99.31% and 97.72%, respectively. The class-wise recognition efficiency of Xception is presented in Table 2.

InceptionResNetV2 acquired higher sensitivity value in the healthy leaf class than other classes, 98.12%, and obtained higher specificity in the stem rot class than other classes, 99.48%. Moreover, it obtained higher accuracy in the hairy caterpillar class than other classes, 98.93%, and the precision value of the healthy leaf class is higher than others, 98.12%. Class-wise recognition performance of InceptionResNetV2, and InceptionV3 are given in Tables 3, and 4, respectively.

**Table 2** Class-wise recognition performance of Xception

| Class | Sen (%) | Spe (%) | Acc (%) | Pre (%) |
|---|---|---|---|---|
| Healthy leaf | **94.84** | 98.77 | 97.47 | **97.46** |
| Leaf mosaic virus | 90.00 | 98.76 | 97.32 | 93.45 |
| Leaf curl | 86.49 | **99.31** | **97.72** | 94.65 |
| Stem rot | 89.92 | 98.27 | 97.50 | 84.06 |
| Semilooper | 93.17 | 97.28 | 96.61 | 86.86 |
| Hairy caterpillar | 91.19 | 97.88 | 97.04 | 86.06 |

**Table 3** Class-wise recognition performance of InceptionResNetV2

| Class | Sen (%) | Spe (%) | Acc (%) | Pre (%) |
|---|---|---|---|---|
| Healthy leaf | **98.12** | 99.10 | 98.79 | **98.12** |
| Leaf mosaic virus | 95.91 | 99.11 | 98.61 | 95.26 |
| Leaf curl | 95.42 | 98.96 | 98.57 | 91.82 |
| Stem rot | 92.61 | **99.48** | 98.79 | 91.82 |
| Semilooper | 95.33 | 99.22 | 98.54 | 96.30 |
| Hairy caterpillar | 95.73 | 99.42 | **98.93** | 96.25 |

**Table 4** Class-wise recognition performance of InceptionV3

| Class | Sen (%) | Spe (%) | Acc (%) | Pre (%) |
|---|---|---|---|---|
| Healthy leaf | **99.78** | **99.84%** | **99.82%** | **99.67%** |
| Leaf mosaic virus | 98.65% | 99.79% | 99.61% | 98.87% |
| Leaf curl | 98.42% | 99.76% | 99.61% | 98.11% |
| Stem rot | 97.82% | 99.72% | 99.54% | 97.46% |
| Semilooper | 98.17% | 99.78% | 99.50% | 98.97% |
| Hairy caterpillar | 98.92% | 99.75% | 99.64% | 98.39% |

According to class-wise recognition performance, InceptionV3 and JuteNet performed better in the healthy leaf class compared with other classes. In the healthy leaf class, InceptionV3 acquired 99.78% sensitivity, 99.84% specificity, 99.82% accuracy, and 99.67% precision value. Class-wise recognition performance of JuteNet was better than three single CNNs, and it obtained 100.00% sensitivity value in the healthy class. JuteNet acquired 99.95% specificity, 99.96% accuracy, and 99.89% precision value in the healthy leaf class. However, the class-wise recognition performance of InceptinV3 and JuteNet were very close in six classes. The recognition performance of JuteNet for each class validates the significant recognition ability of this hybrid model. Among single CNNs, InceptionV3 performed better than Xception and InceptionResNetV2 in class-wise recognition evaluation. Class-wise recognition performance of JuteNet are given in Table 5.

Moreover, Xception misclassified 64 images of the semilooper class, which was the highest misclassification number of this study. On the contrary, InceptionResNetV2 wrongly predicted 26 images of the leaf curl class, which was the highest misclassification number among six classes. Inception V3 wrongly classified 3 images of the healthy leaf class, which was the lowest misclassification number among six classes. The proposed JuteNet wrongly predicted 1 image of the healthy leaf class that was the lowest false prediction number. In terms of the misclassification number of each class, JuteNet performed remarkably better than single CNNs. Class-wise misclassification of four models is given in Table 6.

The recognition time of JuteNet and single CNNs was also evaluated in this study using eighteen images of six classes. JuteNet recognized all images accurately within

**Table 5** Class-wise recognition performance of JuteNet

| Class | Sen (%) | Spe (%) | Acc (%) | Pre (%) |
|---|---|---|---|---|
| Healthy leaf | **100.00** | **99.95** | **99.96** | **99.89** |
| Leaf mosaic virus | 99.10 | 99.87 | 99.75 | 99.32 |
| Leaf curl | 99.06 | 99.92 | 99.82 | 99.37 |
| Stem rot | 99.27 | 99.84 | 99.79 | 98.55 |
| Semilooper | 99.18 | 99.91 | 99.79 | 99.59 |
| Hairy caterpillar | 99.46 | 99.88 | 99.82 | 99.20 |

**Table 6** Number of class-wise wrongly classified images of models

| Class | Xception | InceptionResNetV2 | InceptionV3 | JuteNet |
|---|---|---|---|---|
| Healthy leaf | 23 | 17 | 3 | **1** |
| Leaf mosaic virus | 29 | 21 | 5 | 3 |
| Leaf curl | 17 | 26 | 6 | 2 |
| Stem rot | 44 | 13 | 7 | 4 |
| Semilooper | 64 | 18 | 5 | 2 |
| Hairy caterpillar | 52 | 14 | 6 | 3 |

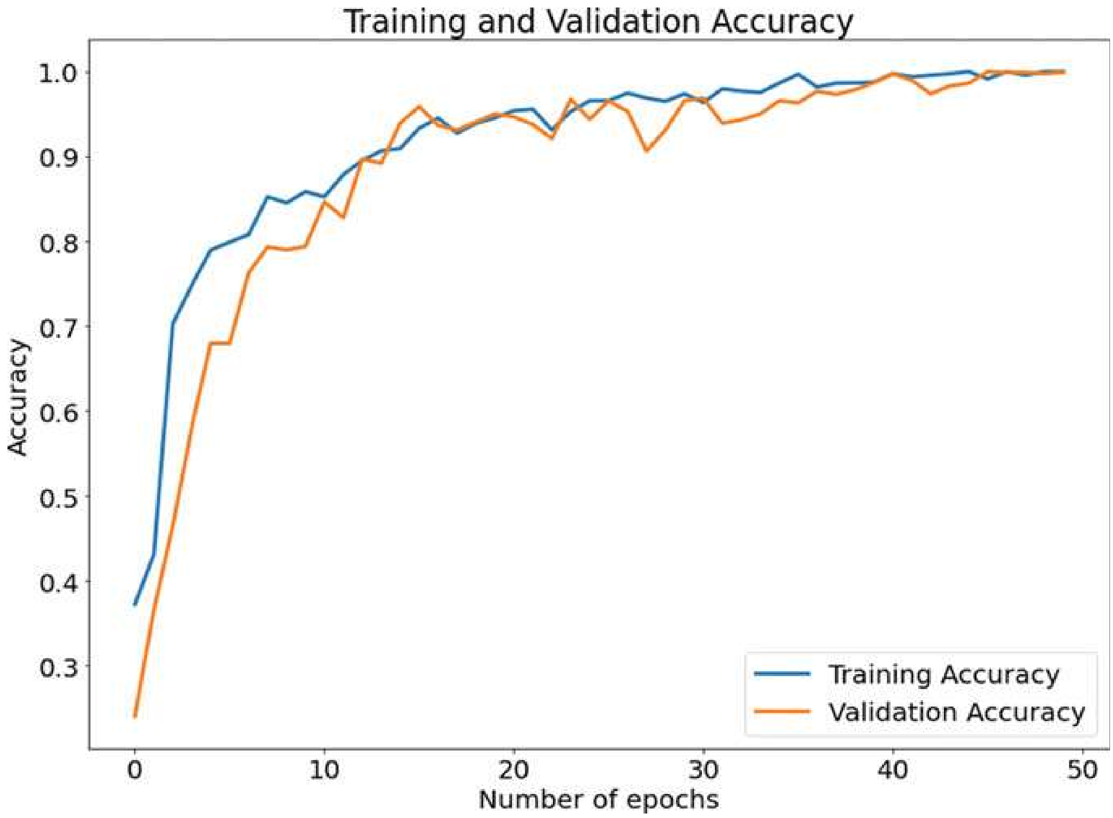


**Fig. 5** The curve of training and validation accuracy of JuteNet

11.03 s. Xception recognized sixteen images correctly within 14.68 s where InceptionResNetV2 and InceptionV3 misclassified one image and took 12.94 and 13.01 s. Several evaluation experiments were introduced using 2803 testing images, where JuteNet significantly performed better than single CNNs, which strongly ensured the recognition efficiency of the fusion of multi-scale features (FMF) approach.

The accuracy and loss curve of both training and validation of JuteNet with the training and validation set of the JL dataset are presented in Figs. 5 and 6.

The recognition efficiency of the proposed JuteNet framework was evaluated by comparing it with other approaches which had been conducted previously to identify different crop diseases and pests using several computer vision approaches and is given in Table 7.

## 6 Conclusion

Diseases and pests of jute cause enormous financial losses to cultivators and the jute industry, and are also a major threat to the sustainable development of this industry. A robust hybrid recognition model, namely JuteNet, is addressed to recognize jute diseases and pests during the initial stage in this paper. In the JuteNet framework, the fusion of multi-scale features was used, where Xception, InceptionResNetV2,

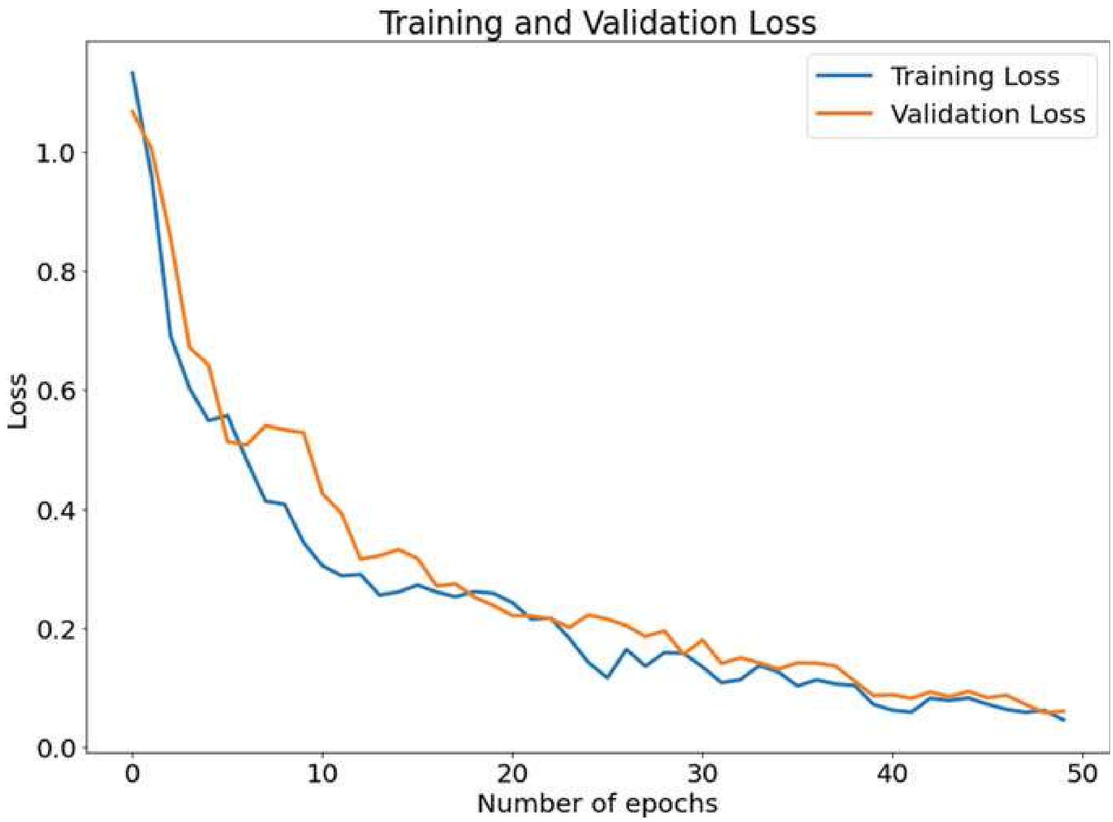


**Fig. 6** The curve of training and validation loss of JuteNet

**Table 7** Comparison of several crops disease and pest recognition approaches

| Study | Method | Number of images | Number of classes | Accuracy (%) |
|---|---|---|---|---|
| Hridoy et al. [3] | CNN | 16,980 | 3 | 97.11 |
| Hasan et al. [4] | CNN | 600 | 3 | 96.00 |
| Reza et al. [5] | M-SVM | Not Mentioned | 5 | 86.00 |
| Habib et al. [7] | RF | 480 | 5 | 89.59 |
| Sholihati et al. [8] | VGG16 | 5100 | 5 | 91.31 |
| Senan et al. [9] | CNN | 3355 | 4 | 96.60 |
| Bhowmik et al. [10] | CNN | 2341 | 3 | 95.94 |
| Hridoy et al. [11] | CNN | 10,662 | 4 | 96.02 |
| Hussain et al. [12] | FMF | 2500 | 5 | 96.50 |
| Trang et al. [13] | CNN | 394 | 4 | 88.46 |
| Fenu et al. [14] | InceptionV3 | 3057 | 4 | 90.68 |
| Zaki et al. [15] | MobileNetV2 | 4671 | 4 | 95.94 |
| Proposed | FMF | 56,108 | 6 | 99.47 |

and InceptionV3 were utilized for extracting features from images of jute leaf and stem. A dataset of 56,108 images was generated in this study which contains six classes, and several evaluation experiments were conducted using 2803 images of the test set. The proposed JuteNet framework obtained 99.47% accuracy, where Xception, InceptionResNetV2, and InceptionV3 were trained individually with the JL dataset, acquired 91.83%, 96.11%, and 98.86% accuracy, respectively. These four models were trained via the transfer learning strategy for the purpose of this study. On the contrary, the class-wise recognition performance of JuteNet and single CNNs was evaluated using statistical parameters, and JuteNet performed better than single CNNs. Moreover, JuteNet misclassified 1 image of the healthy leaf class where Xception, InceptionResNetV2, and InceptionV3 wrongly classified 23, 17, and 3 images, respectively. Among single CNNs, the class-wise recognition performance of InceptionV3 was better. The recognition time of JuteNet was also less than single CNNs, obtained 100.00% accuracy in recognizing eighteen images within 11.03 s. The outcome of several evaluation studies represented the significant recognition efficiency of the fusion of multi-scale features more perceptibly.

However, JuteNet can only recognize common diseases and pests of jute. In future work, we plan to enhance the JL dataset by including images of other crops and extend the number of images and classes to develop a more robust recognition approach for several crop diseases.

## References


1. Islam MM (2019) Advanced production technology and processing of jute. In: Hasanuzzaman M (eds) Agronomic crops. Springer, Singapore
2. Selvaraj K, Gotyal BS, Gawande SP, Satpathy S, Sarkar SK (2016) Arthropod biodiversity on jute and allied fibre crops. In: Chakravarthy A, Sridhara S (eds) Economic and ecological significance of arthropods in diversified ecosystems. Springer, Singapore
3. Hridoy RH, Rakshit A (2022) BGCNN: a computer vision approach to recognize of yellow mosaic disease for black gram. In: Smys S, Bestak R, Palanisamy R, Kotuliak I (eds) Computer networks and ınventive communication technologies. Lecture notes on data engineering and communications technologies, vol 75. Springer, Singapore
4. Hasan MZ, Ahamed MS, Rakshit A, Hasan KMZ (2019) Recognition of jute diseases by leaf ımage classification using convolutional neural network. In: 2019 10th International conference on computing, communication and networking technologies (ICCCNT), pp 1–5
5. Reza ZN, Nuzhat F, Mahsa NA, Ali MH (2016) Detecting jute plant disease using image processing and machine learning. In: 3rd International conference on electrical engineering and ınformation communication technology (ICEEICT), pp 1–6
6. Asnaoui KE (2021) Design ensemble deep learning model for pneumonia disease classification. Int J Multimed Info Retr 10:55–68
7. National Habib MT, Mia J, Uddin MS, Ahmed F (2020) An ın-depth exploration of automated jackfruit disease recognition. J King Saud Univ – Comput Inf Sci
8. Sholihati RA, Sulistijono IA, Risnumawan A, Kusumawati E (2020) Potato leaf disease classification using deep learning approach. In: 2020 International electronics symposium (IES), pp 392–397

9. Senan N, Aamir M, Ibrahim R, Taujuddin NSAM, Muda WHNW (2020) An efficient convolutional neural network for paddy leaf disease and pest classification. Int J Adv Comput Sci Appl (IJACSA) 11(7)
10. Bhowmik S, Talukdar AK, Sarma KK (2020) Detection of disease in tea leaves using convolution neural network. In: 2020 Advanced communication technologies and signal processing (ACTS), pp 1–6
11. Hridoy RH, Habib T, Jabiullah I, Rahman R, Ahmed F (2021) Early recognition of betel leaf disease using deep learning with depth-wise separable convolutions. In: 2021 IEEE region 10 symposium (TENSYMP), pp 1–7
12. Hussain N, Khan MA, Tariq U, Kadry S, Yar MAE, Mostafa AM, Alnuaim AA, Ahmad S (2022) Multiclass cucumber leaf diseases recognition using best feature selection. CMC-Comput Mater Continua 70(2):3281–3294
13. Trang K, TonThat L, Thao NGM, Thi NTT (2019) Mango diseases ıdentification by a deep residual network with contrast enhancement and transfer learning. In: IEEE Conference on sustainable utilization and development in engineering and technologies (CSUDET), pp 138–142
14. Fenu G, Malloci FM (2021) Using multioutput learning to diagnose plant disease and stress severity. In: Complexity, vol 2021
15. Zaki SZM, Zulkifley MA, Stofa MM, Kamari NAM, Mohamed NA (2020) Classification of tomato leaf diseases using MobileNet V2. IAES Int J Artif Intell (IJ-AI) 2(2):290–296
16. Vijayakumar T (2019) Comparative study of capsule neural network in various applications. J Artif Intell 1(01):19–27
17. Ranganathan G (2021) A study to find facts behind preprocessing on deep learning algorithms. J Innov Image Process (JIIP) 3(01):66–74
18. Jacob IJ (2019) Capsule network based biometric recognition system. J Artif Intell 1(02):83–94
19. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 1800–1807
20. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-ResNet and the impact of residual connections on learning. In: Proceedings of the thirty-first AAAI conference on artificial ıntelligence. AAAI Press, pp 4278–4284
21. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the ınception architecture for computer vision. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 2818–2826