International Journal of Advanced Technology and Engineering Exploration, Vol 10(102)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2022.10100289

Research Article

Detection of whitefly pests in crops employing image enhancement and machine learning

Lal Chand[*], Amardeep Singh Dhiman and Sikander Singh
Department of Computer Science and Engineering, Punjabi University, Patiala, Punjab, India

Received: 10-October-2022; Revised: 27-May-2023; Accepted: 29-May-2023

Abstract
Agricultural research is currently undergoing a transformation with the emergence of precision agriculture, which utilizes automated monitoring, data collection, and analysis technologies. This new paradigm is expected to have a profound impact on agricultural practices, aiming to significantly improve both the quantity and quality of crop yields. One crucial challenge in precision agriculture is the automated detection of pests, as they can cause substantial damage to agricultural produce. However, the diverse nature of pests and the variety of crops they attack pose significant challenges for automated pest detection. A deep neural network-based approach has been proposed for the automated detection of whitefly pests in common plants. Before the actual training process, the captured images are subjected to contrast enhancement to ensure uniformity, as they are typically taken under varying lighting and partial shading conditions. The preprocessing step has been shown to enhance the accuracy of the proposed method by making the system more resilient to image degradations. The techniques utilized in this research employ decision tree (DT), convolutional neural networks (CNN), residual networks (ResNet), and attention-based CNN. The experimental results indicate that the proposed technique achieves accuracy rates of 81%, 96%, 97.5%, and 98% for the four models, namely DT, CNN, ResNet, and attention-based CNN, respectively. By comparing the results with those of baseline contemporary techniques, it is evident that the proposed model outperforms other deep learning models in terms of classification accuracy. Consequently, the method presented in this study can be considered an effective automated technique for accurately detecting whitefly pests and identifying pest infestations in crops.

Keywords
Precision agriculture, Whitefly pest detection, Machine learning, Histogram normalization, Feature extraction, Classification accuracy.

1.Introduction
The domain of agriculture is seeing a paradigm shift with the increasing use of technology. Employing robotics, automation and communication technology has opened a completely new field termed as precision agriculture [1]. One of the major challenges which agriculturists face is the attack of pests on the crops which can severely damage the crops and subsequent yield. Different crops are subjected to infestations by variety of pests. Due to the rapid multiplication of pests, it becomes necessary to devise mechanisms for quick and accurate detection. Manual detection is often a tedious and time consuming job, which becomes even more difficult if the farm size is large. Thus, accurate automated systems are necessary for pest infestation detection. Use of high-end drone technology combined with image recognition methods based on machine learning (ML), automated detection of pest attacks has gained prominence [2].

Precision agriculture entails an approach for farming management based precise observation and measurement methods of variety of crops depending on their variability. One of the major catalysts for precision agriculture is use of unmanned aerial vehicles (UAVs) for capturing data and sending it for observation and analysis [3]. The UAVs are generally less expensive and equipped with image capturing technologies. Vegetative images capturing methods are supported by these machines that can facilitate the detection and classification of a large variety of pests. The use of artificial intelligence (AI) and ML based techniques can aid in the process of pest detection and control. The use of automated tools greatly facilitates the process of technology driven

*Author for correspondence

agricultural systems optimizing the use of resources and increasing productivity [4]. Precision agriculture has thus emerged as a promising and much sought after technique for automated and quick detection of pests in agricultural farms. The development of such automated integrated pest management (IPM) techniques allows higher productivity, lesser use of pesticides and insecticides and reduced losses. (Bemisia tabaci) commonly known as the Whitefly pest, happens to be one of the most common yet menacing pests to destroy several crop categories such as cotton, soyabean along with a wide variety of fruit and vegetable crops.

Figure 1 depicts a typical cotton plant infested by Whitefly pests. The detection of white flies is particularly challenging due to the extremely small size of the white flies which often resemble spots on the leaves. The whitefly undergoes four major stages prior to development into an adult fly [5]. The stages are depicted in *Figure 2.* The most effective way to tackle the white fly attack is to detect it in the nymph stage. Adult flies multiply extremely rapidly which can infest more than two hundred species of crops. Whitefly infestation has resulted in damage to around 60% of total cotton crops during peaks of whitefly infestations. The whiteflies reproduce rapidly resulting in quick damage to crops if not cited and neutralized quickly. Being small, these flies may often go unnoticed during the initial stages of infestation.

The lifecycle of the white fly consists of four major stages or instars. The whitefly bears a close resemblance with aphids and develops into a full grown fly after the completion of the fourth instar. The flies can multiply in abundance and damage crops in both farmlands and greenhouses. They may even remain dormant in cold weather for long durations and start to multiply suddenly as soon as they get warm, dry and favourable conditions [6]. *Figure 2* depicts the lifecycle of the whitefly pest.

One of the most challenging aspects of controlling whitefly infestations is the fact that the flies often become resistant to chemical pesticides over time. Hence indiscriminate use of pesticides to alleviate the infestation has not proven to be an effective technique. Rather, detecting them early and using isolation and biological methods are recommended for controlling infestation. Hence, to minimize the losses incurred due to the whitefly pest, early and possible automated detection of the pests is of great significance. The major challenge in this aspect though is the quality of the images available for automated detection which often happen to be degraded as they are captured through (UAVs).

Figure 1 Whitefly pest infestation

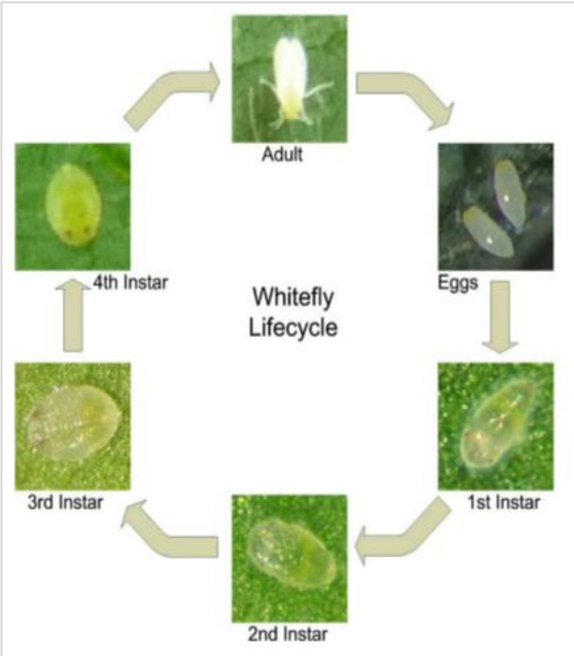

Figure 2 The whitefly life cycle

Thus, the motivation behind the proposed work is enhancement of the images prior to training any automated ML algorithm and subsequently develop an algorithm which serves as an effective automated classifier mitigating the problem of severe crop damage by whitefly pests.

The objectives of this research paper are preparing a comprehensive dataset for the whitefly pest, developing and effective image enhancement approach and finally designing an effective classifier which attains high classification accuracy.

The major contributions of this work are preparing a comprehensive and exhaustive dataset for analysis, designing an image enhancement technique which can reduce the noise effects while capturing and finally designing an automated classifier which is robust to image capturing noise effects that yields significantly high classification accuracy with respect to baseline techniques.

This paper is structured into the following sections. Section 1 provides an overview of precision agriculture, including the background, the motivation behind automated detection of whitefly pests, the associated challenges, and the objectives and contributions of the proposed work. Section 2 presents a comprehensive literature review of the contemporary research in this field. Section 3 discusses the ML models that have been developed for the automated detection of whiteflies. Section 4 presents the experimental results obtained from the study. Section 5 offers a detailed discussion of the experimental results and highlights the limitations of the proposed method. Finally, Section 6 concludes the paper and outlines future directions for enhancement.

2.Literature review

A brief review of the latest research in the domain is cited in this section so as to render insight into the contemporary techniques being used for the detection of whitefly and similar pests in precision agriculture applications.

With the advent of ML and computer vision techniques for agricultural applications, different approaches to detect pests and aphids are being developed. The broad categories of techniques have employed predominantly for the detection of pets entail the following methods:

1) Feature extraction and subsequent application of ML applications for identification/classification of pests and aphids [7].
2) Deep learning based approaches for detection and classification of pests/aphids.

Some of the major noteworthy contribution in the field of study is summarized briefly in this section. De Castro et al., (2022) [8] proposed a deep learning approach based on you only look once (YOLO) version-4 (YOLOv4) with data augmentation and image mosaicking which achieved a F-1 score of 0.87. Parab et al., (2022) [9] compared two deep learning approaches which happen to be YOLOv4, and a single-shot detector faster-RCNN. The faster-RCNN yielded precision of 95.08%, F-1 Score of 0.96 and recall of 98.69%.

Approaches based on identifying patches of abnormality in crops using Faster RCNN, EfficientDet, RetinaNet and YOLOv5 have been explored to detection possible pest infestation, with a highest precision score of 75% for Cassava plants [10].

Chou et al. (2023) [11], proposed an artificial intelligence of things (AIoT) system comprising of internet of things (IoT) sensors to detect environmental sensors and a convolutional neural network (CNN) based method to estimate whitefly pest infestation on Asparagus crops. The proposed work attained highest accuracy of 95.8%. Huddar et al. (2012) [12] employed image augmentation and feature extraction for identification of whitefly pests. The average accuracy obtained was 96%. Legaspi et al. (2021) [13] proposed the YOLO algorithm for the combined multi class detection of white flies and fruit flies in plants. The average accuracy measure obtained as 83% for both the classes. Pattnaik and Parvathi (2021) [14] proposed a ML approach based on hand selected feature extraction followed by support vector machine (SVM) based classification for aphids. Both the histogram of oriented gradient (HOG) and local binary pattern (LBP) features were used for training the SVM. The average accuracy of the proposed model was 97% for the used dataset. One major limitation of their work was the lack of analysis of the model for large and varied datasets as the SVM tends exhibit saturation in performance with the addition of more training data.

This is particularly relevant in the agricultural domain with large divergences in the crop and fly image attributes. Pattnaik et al. (2020) [15] used the CNN model with transfer learning for automated detection of pests in tomato crops. The approach explored the used of transfer learning so as to find out the effect of training of a CNN model with a particular dataset and then using the pre-trained model for classification of other datasets. The approach is useful in the sense that exhaustive labelled training datasets may not be available for multiple crops or may be extremely time consuming and tedious. The average accuracy obtained as 88.3%. Nagar and Sharma (2020) [16] used a multi targeted CNN, RCNN and YOLO approach for detection of very small pests with reduced resolution images. The approach tried to circumvent the challenge of low resolution images typically captured using moving drones under poor lighting conditions. The approach selected the YOLO and CNN due to the speed of computation often required for precision agriculture applications with exhaustive scrutiny of images. The accuracy obtained was 62%, 73.7% and 62.5%. Rustia et al. (2020) [17] developed an embedded systems application with sensors and microcontroller for pest counting in farms with an accuracy of 93%. Liu et al. (2019) [18] proposed a fused channel-spatial attention (CSA) module and CNN approach for multi-class pest detection. The approach is termed as regional proposal network (RPN) which attains an accuracy of 75.46% for the multi-class pest dataset. Chen et al. (2020) [19] proposed the use of the AlexNet and CNN for the detection and estimation of damages by pests. The average accuracy obtained as 82%. Deng et al. (2018) [20] proposed an SVM model trained by handpicked scale invariant feature transform (SIFT) features. The ML approach attained an accuracy of 85.5%. Giakoumoglou et al., 2022 [21] employed YOLO, RCNN and RetinaNet deep learning algorithms for the identification of whitefly pests obtaining an average precision of 75%. Rajan et al., 2016 [22] used a colour features trained SVM model for the detection of pets and aphids. The classification accuracy obtained was 95%. Gašparović et al. (2020) [23] proposed a semi-supervised approach for detection of weeds in sunflower plants. The semi-supervised approach was a step towards exploring ML approaches which could be trained with limited or small datasets, typically to reduce the effort and time in labelling large datasets. The semi-supervised SVM approach was proposed in this work. The work was relevant considering the fact that SVM models regularly exhibit saturation and well as exacerbation in performance for large datasets. Potena et al. (2017) [24] proposed a CNN approach with relatively lesser convolution and pooling layers for the detection of weeds in crops targeting fast training convergence. The approach achieved an accuracy of 96%. Omrani et al. (2014) [25] proposed a K-Means clustering and SVM approach for crop disease detection. The K-Means algorithm was effectively used to primarily group common attributes to facilitate training for the SVM model. A classification accuracy of 93% was obtained using this approach. Cho et al. (2007) [26] applied size, shape, and color features for regression analysis among multiple classes of pests, whiteflies, and aphids. The image analysis techniques based on the features attained an accuracy of 55%. Qiao et al. (2008) [27] proposed a segmentation-based approach for the detection of white flied in traps which attained an accuracy of 81%. Wang et al. (2013) [28] developed an image processing assisted smart devise for the detection of plant and vegetable pests. An average classification accuracy of 82% was achieved. Wang et al. (2020) [29] proposed a deep learning model base on multi-projection model and convolutional CNN for generating high resolved features used in the classification of three categories of plant pests which were the sticky worm, rice planthopper and wheat mice. The average accuracy obtained was 73.9% for the model.

A summary of different contemporary approaches for the detection of pests and aphids is summarized in *Table 1*.

Table 1 Noteworthy contribution in the field

Authors	Approach used	Performance
Joochim et al., 2023 [10]	Faster RCNN, EfficientDet, RetinaNet and YOLOv5 employed to identify abnormal patches in plants.	Precision of 75% attained.
Chou et al., 2023 [11]	Artificial intelligence of things (AIoT) based method comprising data collected by IoT sensors and deep learning for identification of Whitefly Pest infestation.	Highest Accuracy of 95.8%
De Castro et al. 2022 [8]	Deep learning object detection algorithm (YOLOv4) with data augmentation and image mosaicking	F-1 Score of 0.87.
Parab et al., 2022 [9]	Deep learning-based approach for whitefly detection on a yellow-sticky tape (YST) at an IoT remote whitefly monitoring station.	Precision of 95.08%, F-1 Score of 0.96 and recall of 98.69%, beating YOLOv4 which obtained Precision of 71.77%, F-1 Score of 0.83 and recall of 73.31%,

Authors	Approach used	Performance
Huddar et al., 2012 [12]	Image Segmentation and feature extraction-based classification.	Accuracy of 96%
Legaspi et al., 2021 [13]	Classification of white fly and fruit fly infestation using YOLO algorithm.	Accuracy of 83% achieved.
Pattnaik and Parvathi, 2021 [14]	HOG and LBP features extracted along with classification using SVM.	Accuracy of 97% achieved for used dataset.
Pattnaik et al., 2020 [15]	CNN with transfer learning	Accuracy of 88.83%
Wang et al., 2020 [29]	CNN, R-CNN and Yolo algorithms applied for classification.	Accuracy of 62%, 73.7% and 62.5% achieved.
Rustia et al., 2020 [17]	A pest counting technique based on embedded sensing module in wireless sensor networks (WSN)	Average accuracy achieved for automated pest counting was 93%
Liu et al., 2019 [18]	RPN that is adopted by fusing the CSA module and CNN	Classification Accuracy of 75.46% achieved.
Chen et al., 2018 [19]	Alexnet and CNN used to evaluate crop damage	Mean Accuracy of 82% achieved.
Deng et al., 2018 [20]	bio-inspired hierarchical model and SIFT were used for feature extraction and classification was done using the SVM	Accuracy of 88.5% achieved.
Giakoumoglou et al., 2022 [21]	YOLO, RCNN and RetinaNet deep learning algorithms have been used for identification of whitefly pests.	Average precision of 75% achieved.
Rajan et al., 2016 [22]	Segmentation employed, color features used to train the SVM to classify the pest pixels and leaf pixels.	The classification accuracy was 95%.
Gašparović et al., 2020 [23]	Semi-supervised learning approach for weed mapping in sunflower crops.	The classification accuracy was 95%.
Potena et al., 2017 [24]	Lightweight CNN used for crop weed classification.	Highest Accuracy of 96% achieved
Omrani et al., 2014 [25]	A combination of K-Means clustering and SVM used for crop disease detection.	Accuracy of 93% achieved
Cho et al., 2007 [26]	Size, shape, and color features were extracted and feature similarity was used for classification pf aphids and pests.	Achieved accuracy of 55%
Qiao et al., 2008 [27]	Binary counting of whiteflies based on object detection	Detection accuracy of 81% achieved
Wang et al., 2013 [28]	A thresholding-based counting approach was used.	Accuracy of 82% achieved.

The research work also compares multiple state of the art techniques in terms of the classification accuracy to render insight into the choice of method for real time precision agriculture applications.

The salient points of the existing contemporary literature can be summarized as:

1. Feature extraction of LBP, SIFT or HOG features followed by classification using ML models such as the SVM has the advantage of reduced annotated data and hardware

requirements compared to deep learning models.
2. The challenge of the aforesaid method lies in the selection of features and feature combinations to train the ML model.
3. Deep Learning models such as the CNN, RCNN, Faster RCNN, AlexNet, YOLO and its variants completely bypass the necessity to meticulously identify features and feature combinations thereby making the system more readily usable.
4. The limitation of such an approach however lies in the fact that it requires copious amounts of annotated data and significantly higher hardware for attaining high classification accuracy.

Thus, both feature extraction followed by a ML classifier and deep learning models have their own advantages and limitations. The choice of the method to be employed should be based on the system constraints and fault tolerance. As the divergences in crop texture and pets is substantial, deep learning-based models are better suited for the application.

3.Methods

This section presents a detailed account of the different baseline approaches along with the developed attention-based model adopted for automated classification. The analysis of previous work indicates the fact that two major approaches have been used for the detection of pests and weeds in plants. One of them is the ML based approach with hand-picked features and the other is the deep learning-based approach. The most common type of Deep Learning approaches used are the CNN, recurrent networks, and YOLO [30]. Limited work seems to be done on image pre-processing and enhancement, which may substantially improve the performance of any classifier as plant images with pests or diseases or weeds are subject to noise and degradation depending upon the capturing technique. The proposed approach presents both ML and deep learning-based techniques aiming to address the limitations of existing work in the domain. The ML and deep learning-based methods employed in this work can be pictorially understood using the system block diagram depicted in *Figure 3.*

While the ML based model needs a separate feature selection stage, the deep learning model bypasses the step, and both approaches have their own advantages and limitations as discussed earlier. The steps to be incorporated for the implementation of the proposed system are discussed in detail subsequently.

3.1Image enhancement

As images captured by UAVs are often affected by noise and degradation effects, hence to facilitate the classification process, image enhancement techniques have proven to be extremely useful. The discrete wavelet transform (DWT) was employed for effective de-noising in existing literature [17]. Unlike the Fourier Transform and its derivatives, the DWT is based on a non-smooth kernel function. This allows for the use of the wavelet transform for the analysis and denoising of image data using wavelet transform as image data exhibits sudden changes and spikes in image statistical parameters. The wavelet transform can be thought of as a combination of high pass and low pass filtering techniques explained in Equation 1.

$$F(n) \xrightarrow[DWT]{} Z_{LPF}, Z_{HPF} \qquad (1)$$

Here,
DWT represents the discrete wavelet transform operator.
Z_{LPF} are the low pass filtered coefficient values.
Z_{HPF} are the high pass filtered coefficient values [31].

Typically, the high pass coefficient values contain the fluctuations and the low pass components contain the original information of the image. The decomposition of the images using wavelet transform can be done as a decomposition tree in which each decomposition level would yield the approximate co-efficient values, the detailed co-efficient values, the horizontal coefficient values and the vertical co-efficient values. Thus, the image in the spatial domain would be converted to the wavelet domain co-efficient as [32], explained in Equation 2.

$$F(x, y) \xrightarrow[DWT2]{} C_A, C_D, C_H, C_V \qquad (2)$$

Here,
C_A represents the approximate co-efficient values.
C_D represents the detailed co-efficient values.
C_V represents the vertical co-efficient values.
C_H represents the horizontal co-efficient values.

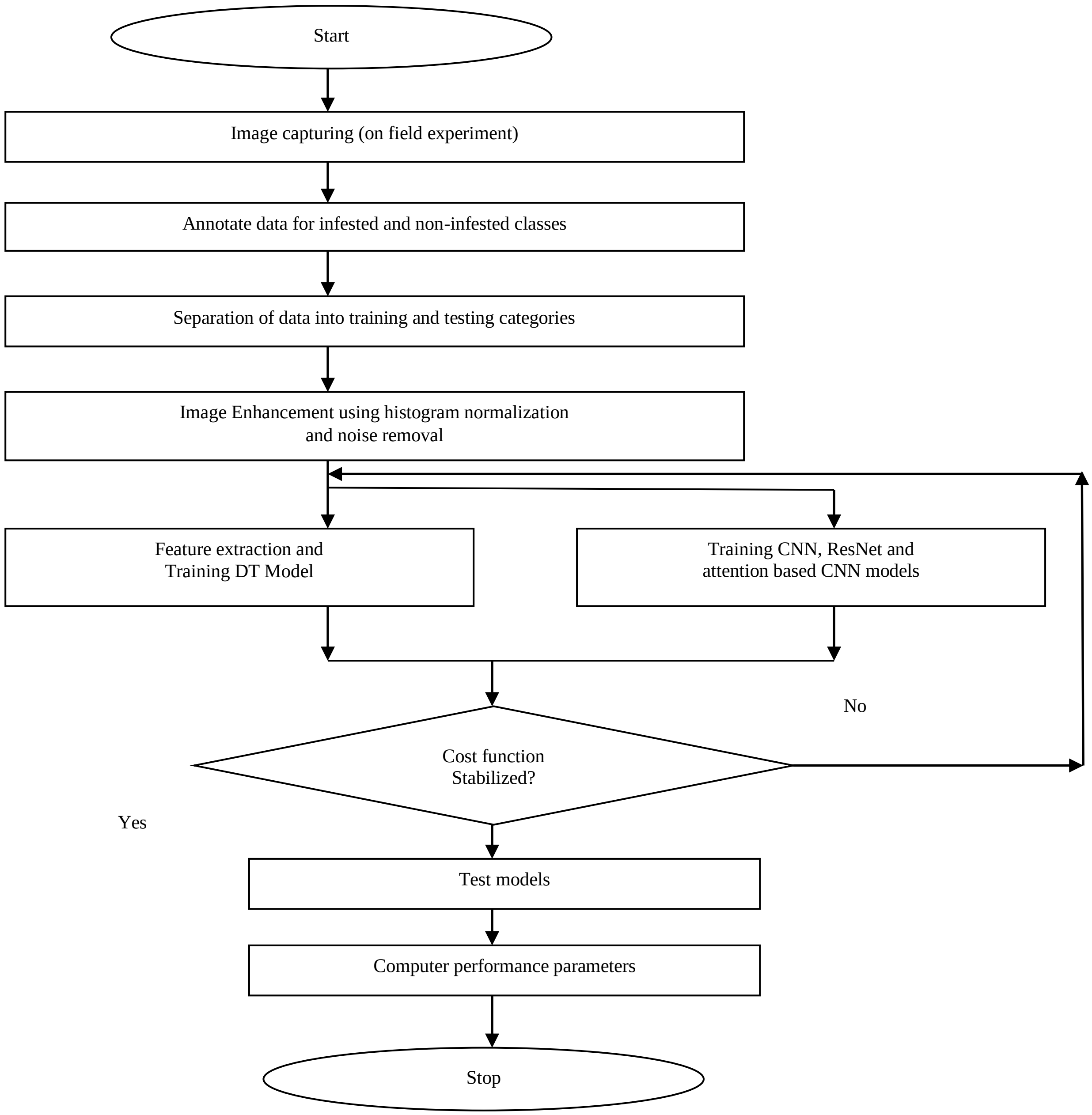

Figure 3 System block diagram

$DWT2$ represents the discrete wavelet transform on two-dimensional image data. The approximate co-efficient values are retained while discarding the noise detailed coefficient values in a recursive manner as [33]:

$for\ (i = 1, i \leq n, i++)$

{

$Image_i \xrightarrow{DWT2} CA_{i,1}, CD_{i,1}$

$CA_{i,1} \xrightarrow{DWT2} CA_{i,2},, CD_{i,2}$ ...

$CA_{i,k-1} \xrightarrow{DWT2} CA_{i,k},, CD_{i,k}$

}

Here,

i represents the index of the i^{th} image.

The running index value 1, 2,..k represents the decomposition value or level

CA represents the approximate coefficient of a particular level

CD represents the detailed coefficient of a particular level

DWT2 represents the wavelet transform function applied to 2-D images.

The histogram analysis allows us to infer the image enhancement results of the 2-D DWT applied. The normal histogram and be computed using Equation 3.

$$h = \sum_{k=1}^{l} s_l \quad (3)$$

Here,
s represents function for histogram decomposition.
h represents the normal histogram
k represents the histogram index
l represents distributed bin count of the DWT-histogram analysis.

The computation of the bits is done using Equation 4:

$$l = \frac{\max(X) - \min(X)}{l_w} \quad (4)$$

Here,
$\max(X)$ denotes the maximum value of the random variable $'X'$
$\min(X)$ denotes the minimum value of the random variable $'X'$
l_w converges with distinct values for the variable in the distribution.

As the UAVs encounter moderate to severe image degradation effects, filtering out the same is mandatory to ensure accurate feature extraction.

Moreover, it is necessary to separate or segment out the region 'A' with a central reference 'C_0', based on the computation of the maximum value of the changing gradient computed using Equation 5 [34].

$$g_r = max(r, C_0,) |G_\sigma(r) \frac{\partial}{\partial r} \oint_{r,C_0,}^{R} \frac{I(x,y)}{2\pi r} dA| \quad (5)$$

Here,
g_r is the gradient along 'r'
$I(x,y)$ is a 2-D image.
C_0 is the reference where the gradient computation is initiated.
G_σ represents a Gaussian kernel
r is the radial increment.
R is the maximum value of the radial increment computed from central reference.
max is the maximum value operation.
dA is the differential increase in area.
Segmentation allows clear demarcation of the affected area within the area of interest.

3.2Feature extraction

Subsequent to image enhancement and segmentation to highlight the areas of interest, the next step is feature extraction coupled with automated classification. Image classification can be done employing different benchmark techniques [35]. Feature extraction from the labelled data is extremely critical to accurate classification since the features help the classifier to find pattern in different groups of data. One of the ways to classify is by treating the image pixels as the features. This may be effective in case of object recognition or image classification where the categories have distinct boundaries [36]. In case of detection of pest attacks, especially extremely small pests such as white flies, this method may succumb to overlapping or fuzzy boundaries among the pixel values of images. Moreover, for RGB images, with 3 distinct R, G and B channels, the dimension of the feature dataset may be exceedingly large for lesser sophisticated hardware for agricultural systems. Hence statistical feature extraction may be adopted as the baseline for the classifier. The feature selection should be made with the following points of consideration [37]:

1) The features should be pervasive for different crop types.
2) The computation complexity must not be high keeping in mind the system constraints.
3) Accurate feature extraction is possible for the dataset to be used.

The stochastic features computed in this paper are mean, median, variance, standard deviation, skewness and kurtosis. The image-based features such as energy, entropy, homogeneity, root mean square (RMS) value and smoothness are also computed to identify both statistical as well as image-based artefacts [38].

3.3Feature combination and classification

After the feature extraction part, the design of the ML based classifier is done. In this paper, four different feature vectors have been experimented based on different feature combinations. The feature combinations include minimalistic stochastic, stochastic, image texture based and combination of stochastic and image texture-based features. These feature combinations are termed as feature vectors 1, 2, 3 and 4 respectively and have been discussed in the results section. The feature combination analysis allows to analyse the effects of addition/deletion of features on the performance of the system. This helps in deciding the less impactful features and hence allows streamlining the feature extraction process.

This is especially useful for handpicked feature extraction for ML models which are to be implemented on constrained hardware platforms.

As ML has kept gaining more prominence in the state of the practice over a wide range of applications, one of the most impactful ML techniques to reach the forefront has been the artificial neural network (ANN) The different classification paradigms used in this work are explained subsequently. Neural Networks: ANNs have gained a lot of prominence recently due to the emergence of deep neural networks and deep learning. *Figure 4* depicts the structure of a single neuron.

The neuron depicted in *Figure 4* is a multi-input neuron which can process parallel inputs simultaneously. The output of the single neuron unit can be computed using Equation 6.

$$y = \varphi(\sum_{i=1}^{n} x_i w_i + \beta) \quad (6)$$

Here,
x denotes the inputs to the neuron.
y denotes the output of the neuron.
w denotes the weights.
β denotes the bias.
φ denotes the activation function.

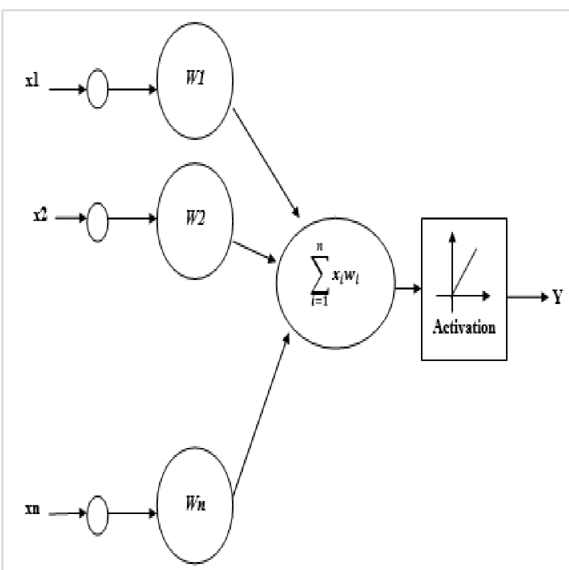

Figure 4 The multi-input neural network model

A dense interconnection of such neurons is often termed as a neural network. Increasing the data processing or hidden layers of the network allows to make complex computations.

The most commonly used classification approaches involving neural networks are:
1) Neural network based ML approach.
2) Deep neural network based deep learning approach.

Designing deep neural networks (termed as Deep Learning) with cascaded neural layers is found to be extremely effective in automated classification problems. Rather, the features are computed by the network directly. In general, the deep learning approach often has a much deeper networks with stacked hidden layers. Some of the common neural network configurations which have been used for classification problems pertaining to pests are the decision tree (DT) and different variants of the CNN [39]. The DT algorithm works on the principle of Bayes theorem of conditional probability to compute the maximal probability of an unknown data sample in each class [40]. Another popular and effective neural network structure presented in the CNN which has evolved as one of the most effective deep neural network strictures which works on the principle of convolutions [41]. The outer layers of the networks are used to compute low level features while the deeper layers are used to compute the higher level features. Adjusting the parameters (often termed as Hyperparameters) of the network allows to tune the network to a given data set.

The ML model is depicted in *Figure 4.* Recurrent networks often use a cascading structure of applying the output of one layer to the input of another layer in loops. CNNs sometime encounter the problem of a vanishing gradient and overfitting [42]. To overcome these limitations, residual networks (called ResNets) are sometimes employed which have skip connections among the hidden layers which do not directly connect the layers in cascade thereby decreasing the chances of overfitting [43].

The hypothesis of any designed algorithm lies in the classification of partially overlapping datasets based on the deep probabilistic classifiers. The major challenge in this aspect is the similarity among infested and non-infested plants, hence accurate decision making is challenging for overlapping data samples. The intersection of sets overlapping sets makes the decision making challenging and hence a probabilistic classification is needed. One of the most common deep learning approaches used for image classification happens to be the CNN.

The CNN is an extremely effective deep learning based classifier which performs pattern recognition in each of its layers based on stochastic computing. The fundamental operation in the CNN hidden layers is the convolution operation mathematically computed using Equation 7.

$$x(t) \times h(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau \qquad (7)$$

Here,
x(t) is the input
h(t) is the system under consideration.
y is the output
×is the convolution operation in continuous domain
For a discrete or digital counterpart of the data sequence, the convolution is computed using Equation 8.

$$y(n) = \sum_{-\infty}^{\infty} x(k)h(n-k) \qquad (8)$$

Here
x(n) is the input
h(n) is the system under consideration.
y is the output
×is the convolution operation in discrete domain

The architecture of the CNN primarily comprises of the input, output, and convolution, pooling and fully connected layers. The input layer is fed with annotated or grouped classes of the data. For instance, in case of image classification, the grouped images are fed to the input layer. The convolution layers are stacked one after the other computing the image features. In general, higher lower order features are computed in the output layers while the higher order features are computed in the inner layers. Typically, strided convolutions are calculated as they help in covering all the data samples rather than just the internal samples of the data matrix. The pooling layers are used typically for reducing the computational complexity of the network. The fully connected layer connects the features to the hidden layer which is subsequently connected to the output layer which renders the classification result.

In this paper, DT are used for handpicked feature extraction and classification while the CNN and its developed variants have been employed for automated classification of the whitefly pests. A derivative of the CNN termed as the residual network or ResNet has also been designed and used in this paper which tries to mitigate the limitations of the conventional CNN. Interleaved or skip connecting the hidden layer neurons avoid the changes of vanishing gradient of errors with respect to weights and additionally avoids overfitting. While ML and deep learning-based approaches assign equally weighted connections to the hidden layer, this work presents an attention-based model for automated classification of Whitefly pests. The essence of the approach lies in the fact that all the statistical features in the feature vector computed by any deep learning network does not have the same impact on the classification process. Hence, to make the classifier more robust and immune to overlapping feature values, high impacting features are given more scaling weights termed as attention. The contextual attention vector α_cis computed in the proposed approach using Equation 9 [44].

$$\alpha_c = \sum_{i=1}^{T} \alpha_{,i} v_i \qquad (9)$$

The composite contextual vector $v_{c,i}$, for attention weights $\alpha_{t,i}$ is calculated using Equation 10.

$$attention\left(v_{c,i}, \alpha_{c,i}\right) = \sum \alpha_i \times a_{c,i} \qquad (10)$$

The attributes of the data set are discussed in the results section in detail. The detailed description of different methods employed in this work can be summarised as using the proposed algorithm.:

Start
Step. 1: Prepare dataset.
Step.2: Divide dataset in the ratio of 75:25 for training: testing splitting ratio.
Step.3: Employ DWT for removal of inherent noise and degradation effects in the image, defined in equation (2).
Step.4: Apply radial gradient $g_r \xrightarrow[Segment]{} A_{i9X,Y)}$ for segmentation as defined in equation 5.
Step.5: Compute stochastic feature for the DT model.
Step.6: Train the DT model with different feature combinations.
Step.7: Directly apply training dataset for the CNN, ResNet and attention-based models.
Step.8: Terminate training if:
Cost Function $mse = \frac{1}{n}\sum_{i=1}^{n} e_i^2$ stabilizes for multiple iterations
Maximum pre-defined iterations (1000) are reached, whichever happens earlier.
Step.9: Compute classification error and accuracy.
Stop

The accuracy of the proposed system is computed using Equation 11.

$$Ac = \frac{TP+TN}{TP+TN_{FP}+FN} \quad (11)$$

Here,
TP, TN, FP and FN denote the true positive, true negative, false positive and false negative respectively.

The attention based model tries to add another degree of freedom to the conventional loss function of the CNN or ResNet in the sense that the minimization function now would have conventional root mean squared error (RMSE) as one part of the loss and the attention vector based cross entropy as the second loss. The RMSE is computed using Equation 12.

$$rmse = \sqrt{\frac{\sum_1^n (p_i - a_i)^2}{n}} \quad (12)$$

The attention vector cross entropy is computed using Equation 13.

$$E(L,a) = -\sum_1^n a_i log_2 L_i + (1-a_i) log_2 a_i \quad (13)$$

Here,
L stands for the Labelled attention vector.
a_i Stands for the contextual attention vector.
n is the total number of samples.

Thus the loss function incorporates both the rmse and the cross entropy to add the effects of both the convolution based learning as well as the feature vector sequences which happen to find maximum match for consecutive correct classifications. The Hyperparameters of the proposed model are: The convolution layers in the Net have been designed as 50, with the rectified linear (ReLu) activation function and batch normalization incorporated. A stride of 2 and max pool of 3×3 has been used. The input layer size of 256×256 is used and multiple learning rates of 0.01, 0.02 and 0.04 have been experimented with. The optimal learning rate has been found to be 0.02. A drop out of 30% or 0.3 has also been incorporated to fasten up the convergence.
A batch size of 128 has been employed with a weight decay of 10^{-4}. The system has been limited to a depth of 50, to limit the hardware complexity of the system as it is intended for agricultural applications, typically constrained on hardware resources.

The softmax f_c1000 connected layer has been used for the study. The proposed attention-based model is depicted in *Figure 5*. The experimental setup for the proposed approach along with the results obtained and their significance is discussed in detail in the subsequent section.

4.Results

The first step for designing the automated classification model is the preparation of an exhaustive dataset. 18,000 high resolutions .jpg images from the regions of Mansa, Bathinda, Abohar and Fazilka, belonging to Punjab, India have been captured. The sites chosen for data collection were done under the aegis of experts from regional research station, Punjab University of Agricultural (PAU), Bhatinda. These specific locations were chosen after discussions with agricultural scientists working in the domain, who confirmed that these sites are prone to whitefly infestation.

The size of each of the images varies between 5MB to 7MB. The image format is .jpg. The images captured are 3 colour channel images with RGB channels.

Manual labelling of the dataset has been done as:
1. Whitefly infested.
2. Whitefly non-infested

Care has been taken while capturing the images, to maintain almost similar capturing conditions pertaining to lighting, angle of capture, distance of object from capturing device along with identical background.

As all the images captured with or without whitefly infestation are from cotton crops of the same geographical location, homogeneity in capturing attributes has been observed. The horizontal and vertical resolution of the images are 350dpi. The bit depth for the images is 24 with a compression ratio of 3 bits/pixel. The exposure time for image capture is 1/400 s with a maximum aperture of 3.44531.

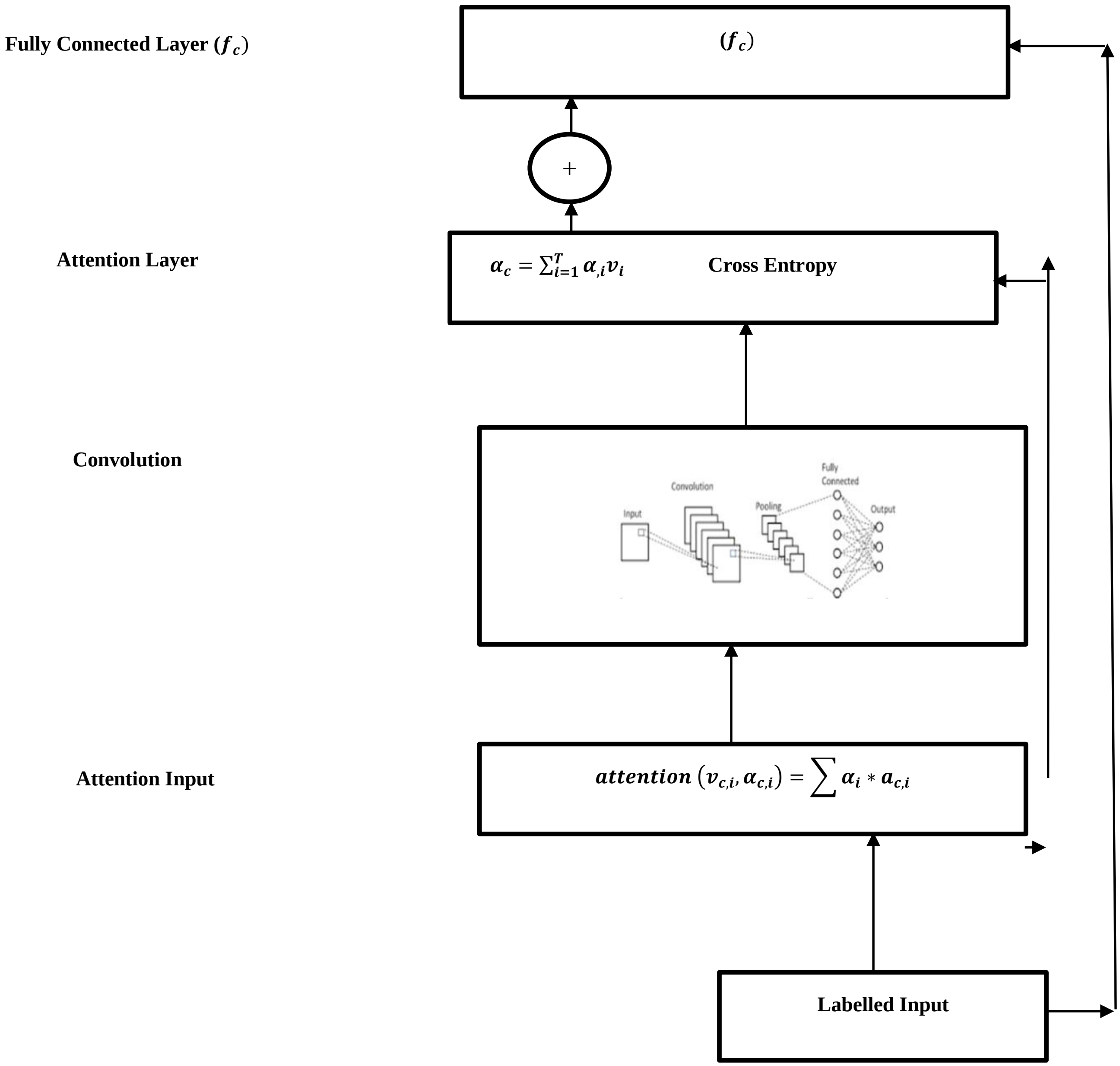

Figure 5 Proposed attention-based model

The data set is subsequently divided into 2 categories for training and testing samples respectively in the ratio of 75:25. The experimental set up has been designed with a 1:1 split ratio of positive and negative annotated images so as to avoid imbalanced instances. The advantage of data preparation by authors themselves has also aided the purpose. Thus data resampling was not required in the study due to the available exhaustive dataset collected and annotated by the authors. Prior to GLCM feature extraction, images are enhanced and segmented based on the DWT filtering and threshold based segmentation approach explained in section 2 of this paper. White the DT based model is trained with 12 statistical features, the Deep Learning model uses 0.66M parameters to train. The experimental tool for implementing the proposed approach has been chosen as Matrix Laboratory (MATLAB 2021a) due to the data visualization and ML tools available. The functions in the image processing, bio-informatics and deep learning toolboxes have been used for the data processing, feature extraction and classification purposes. A detailed step by step process has been presented in this section which highlights each step which is performed in the experiment. *Figure 6* depicts a sample jpg image from the dataset. *Figure 7* depicts the histogram analysis of the original image. It can be observed that the histogram distribution is not uniform for the image.

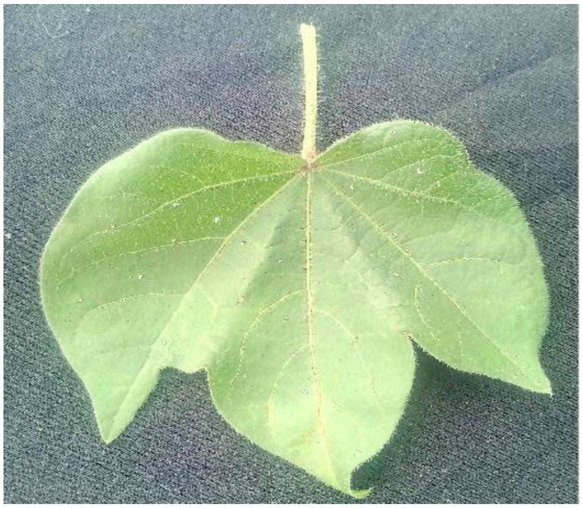

Figure 6 Original image

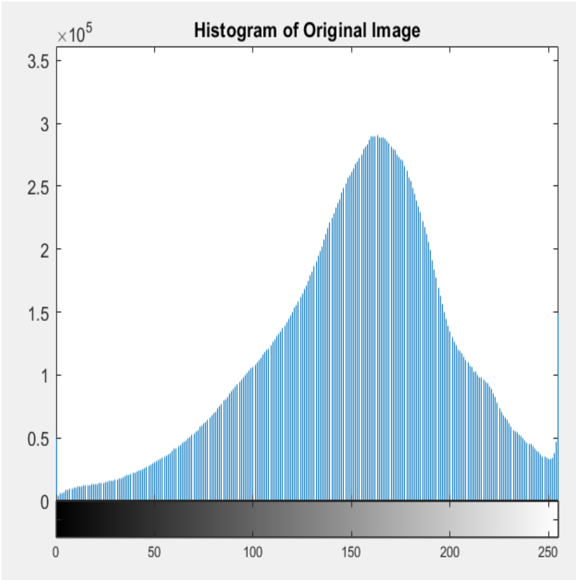

Figure 7 Histogram of original image

Figure 8 depicts the contrast enhanced image whose histogram is depicted in *Figure 9*.

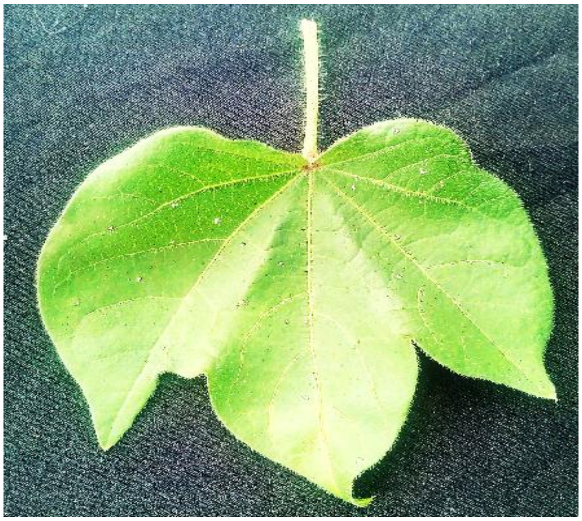

Figure 8 Contrast enhanced image

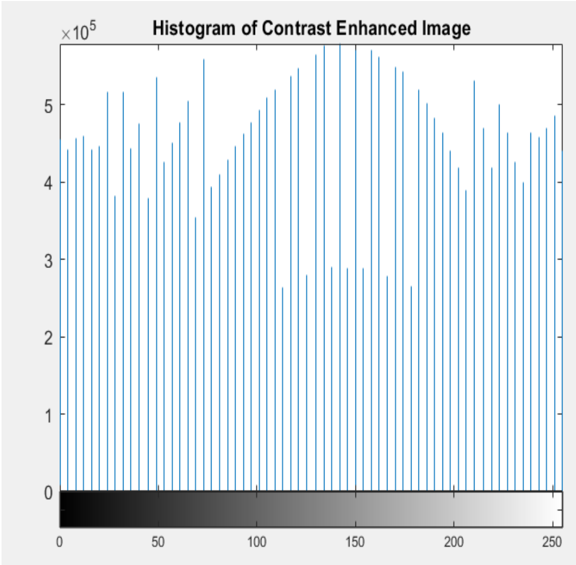

Figure 9 Histogram of contrast enhanced image

Figure 9 depicts the histogram of the contrast enhanced image. A more uniform histogram distribution of the image can be observed compared to the original image. It can be observed from the histogram of the contrast enhanced image that the histogram has a more uniform distribution subsequent to contrast enhancement as compared to the raw or unprocessed image. This is typically useful in the case of crop images captured by UAVs under varying lighting and partial shading conditions. The contrast enhancement helps to normalize the feature values computed from the images.

Different feature combinations have been explored in this paper with 4 different cases of the input vectors:
1) Feature Vector 1: mean, median, variance and standard deviation.
2) Feature Vector 2: mean, median, variance, standard deviation, skewness, and kurtosis.
3) Feature Vector 3: energy, entropy, homogeneity, RMS value and smoothness.
4) Feature Vector 4: mean, median, variance, standard deviation, skewness, kurtosis, energy, entropy, homogeneity, RMS value and smoothness.

A comparative analysis of the classification performance for stochastic features alone, image texture features alone and combination of both stochastic and image texture features is presented in *Table 2.*

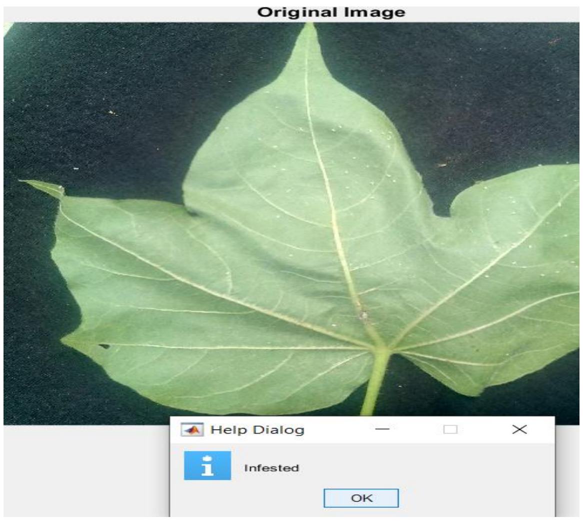

Figure 10 Classification as infested sample

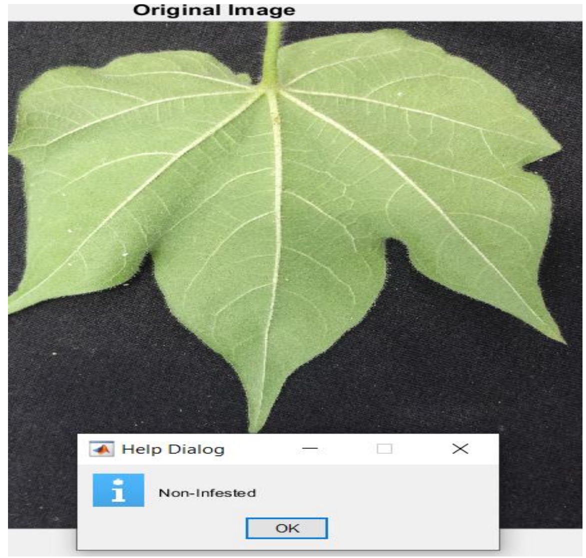

Figure 11 Classification as non-infested sample

As simple graphical user interface (GUI) has been designed for the final classification of any new sample as infested or non-infested and is depicted through *Figures 10* and *11*. The final classification is based on the decision of the classifier as infested or non-infested. A comparative analysis among the various classifiers used in this work has been summarized in *Table 2*. The parameters computed are accuracy (%), precision, recall, specificity and the F-1 score. *Figure 12* depicts the confusion matrix for the proposed attention-based CNN model. Similar confusion matrices can be generated for the DT, CNN and ResNet Models. The confusion matrix clearly depicts the TP, TN, FP, FN values along with the class distribution for the testing phase which can be seen to be 1:1 (same as the training phase).

Table 2 Comparative accuracy of classification for different approaches

S. No.	Technique	Classification accuracy	Precision	Recall	Specificity	F-1 Score
1.	DT	69% (feature vector 1) 73% (feature vector 2) 53% (feature vector 3) 81% (feature vector 4, best case)	0.8163 (feature vector 4)	0.8	0.82	0.8129
2.	CNN with contrast enhancement.	96%	0.9662	0.9533	0.966	0.9597
3.	ResNet	97.5%	0.9702	0.98	0.97	0.9749
4.	Attention-based CNN	98%	0.9730	0.9870	0.97266	0.9795

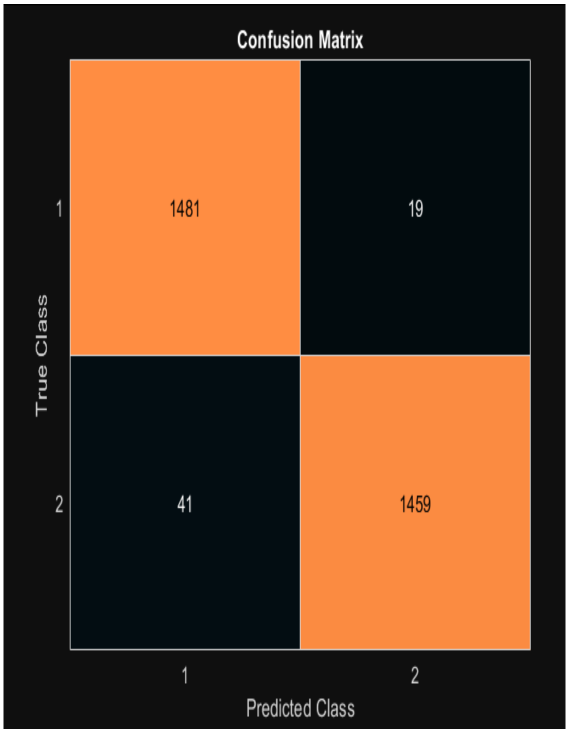

Figure 12 Confusion matrix for the attention-based CNN model

Figure 13 presents a comparative analysis of the different ML models employed for automated classification in this work.

It can be observed from *Figure 13* that the attention-based CNN obtains the highest accuracy followed by ResNet, CNN and the DT algorithms.

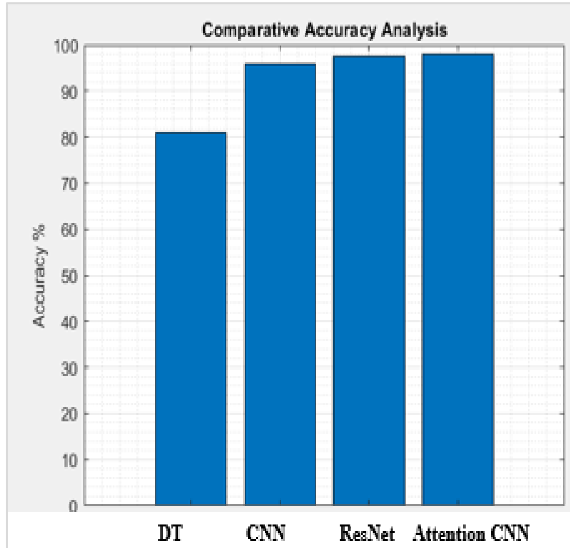

Figure 13 Comparative accuracy analysis

To evaluate the proposed work, the results have been pitted against the results of classification models cited in recent literature. A comparative analysis in terms of classification accuracy has also been made existing contemporary work is also cited in *Table 3.*

Table 3 Comparative analysis with previous work

S. No.	Author	Classification Accuracy	Approach	Dataset	Attributes
1.	Legaspi et al. [13]	83.07%	YOLO.	Image dataset of total 470 images prepared by authors.	The YOLO algorithm typically outperforms conventional CNN or R-CNN as it uses end to end training.
2.	Cho et al. [26]	59%	SVM	Image dataset collected at laboratory of Entomology Division, National Institute of Agricultural Science and Technology, Rural Development Administration, Korea.	Simpler to implement but saturation of accuracy occurs relatively quickly compared to deep learning models.
3.	Qiao et al. [27]	81%	boundary tracking and size estimation based density estimation	Dataset collected at the laboratory at the Entomology Division of the National Institute of Agricultural Science and Technology (NIAST), Rural Development Administration, Republic of	Boundary based approaches may not yield accuracy results under image variability.

S. No.	Author	Classification Accuracy	Approach	Dataset	Attributes
				Korea. Yellow sticky traps (18 cm × 13 cm, Green Agro Tech Co., Ltd. ®, Gyeongsan-si, Korea)	
4.	Gondal and Khan [45]	97%	SVM	Authors own dataset	Simpler implementation, need to lesser data samples but relatively lesser accuracy compared to deep learning models.
5.	De Castro et al. [8]	87%	YOLOv4	Real and annotated data set prepared by authors with annotations under the supervision of an entomologist.	Faster and more accurate compared to conventional CNN. R-CNN as image is analysed as a whole instead of blocks.
6.	Parab et al. [9]	95.08%	Single-Shot and Two-Shot Deep Neural Network	Data set prepared by authors at Web-RAID (Web-Based Remote Autonomous Insect Detector) Station	Higher accuracy compared to conventional CNN.
7.	Joochim et al. [10]	75%	Mask R-CNN, RetinaNet and YOLOv5	Dataset prepared by authors. Images captured through drone designed, equipped with a camera to be used for surveying.	Standard deep learning algorithms used. However, relatively lesser accuracy is achieved.
8.	Chou et al. [11]	85.8%	Sensor based data collection followed by CNN based analysis.	Dataset prepared by authors using sticky paper traps placed in a greenhouse, for Asparagus crops.	CNN is one of the most commonly used deep learning classifiers but needs enormous datasets to train and render high accuracy.
9.	**Proposed Work**	81% DT (best case, feature vector 4) 96% CNN with contrast enhancement 97.5% ResNet 98% (Attention-based CNN)	DT, CNN, ResNet, Attention-based CNN.	Dataset prepared by authors from areas of Mansa, Bathinda, Abohar and Fazilka, belonging to Punjab, India), under the supervision of experts from regional research station, Punjab University of Agricultural (PAU), Bhatinda, Punjab, India.	Both feature extraction subsequent to image enhancement as well as Deep Learning models explored. ResNet avoids chances of overfitting and vanishing gradient. Attention-based CNN is more robust and immune to overlapping feature values.

A comparison with existing work clearly established the fact that the proposed work especially the contrast enhancement coupled with ResNet and attention-based CNN clearly beats baseline contemporary techniques. Out of the four models designed, it can be observed that the attention-based CNN outperforms the DT, CNN with contrast enhancement and the ResNet.

5.Discussion

The proposed approach presents a comparative analysis of four different approaches wherein the handpicked feature-based classifier happens to be the DT while the CNN, RCNN and attention-based CNN happen to be the deep learning-based approaches. One of the reasons for picking the DT approach is to control the features to be computed for further classification at relatively lesser necessity of hardware complexity. As the data set prepared from actual data is exhaustive (18,000 images), the training: testing split is chosen as 75:25. Similar results can be obtained through a 70:30 or 80:20 split. Due to the large dataset prepared (for both infested and non-infested) cases, imbalanced instances in each of the categories is avoided. The DT model achieves convergence in 55 iterations, with a training time of approximately 3 minutes. The attention-based CNN model train in 200 iterations with a training time of 65 minutes. Both the CNN and ResNet reach convergence in 60 iterations and with training times of 38 minutes and 34 minutes respectively. The accuracy comparison is presented in *Figure 13*.

It can be observed that the DT attains converges in a much quicker time compared to the deep learning models. Out of the deep learning models, the attention-based CNN takes the maximum time to reach convergence, while the CNN and ResNet exhibit similar times of convergence.

The DT based classification model attains a classification accuracy of 69%, 73%, 53% and 81% for the 4 categories of feature vectors generated through feature combination. It can be observed that the combination of the stochastic as well as image texture based features outperforms the other feature combinations. This can be attributed to the fact that more information is captured through both stochastic and image texture based features combined compared to the stochastic to image texture based features alone. Due to the necessity of feature selection, combination and evaluation of feature impact on the classification accuracy, deep learning models become handy for large datasets. While deep learning models may be more convenient, yet they may require higher hardware resources compared to ML models. Thus the choice of the model would depend on the hardware constraints of the system. It should be noted though, that although the DT model needs lesser time of converge and lesser hardware requirements, its classification accuracy is significantly less compared to the deep learning models. This too is true for the best case scenario for the DT model trained with feature vector 4 (combination of stochastic and image texture features, both).

The confusion matrix yields the accuracy and error rate estimation. The error rates achieved by the DT, CNN, ResNet and attention-based CNN models are 19%, 4%, 2.5% and 2% respectively. It is evident from the results that the deep learning approaches clearly beat the handpicked feature selection followed by ML based classification approach. The conventional CNN yields slightly lesser accuracy of 96% compared to the ResNet which attains an accuracy of 97.5%. The highest accuracy among all the approaches is attained by the attention-based CNN model which attains an accuracy of 98%. The major advantage of the proposed work is the fact that the models designed in this approach have been tested on real images captured from the Malwa region on Punjab, India, which are not prior-enhanced or processed such as datasets available in the public domain. Moreover, the attention-based CNN approach fuses the cross entropy at the attention layer to the fully connected layer thereby incorporating the changes in the deeper layer features for the various images in the training dataset. As mentioned earlier, it incorporates additional degrees of freedom to the loss function computation fusing the conventional RMSE as one part of the loss and the attention vector based cross entropy as the second loss. This allows for a more robust feature computation and selection criteria which is not considered in both the conventional CNN or ResNet models.

This is significantly important to make the model more robust as real images without prior-processing exhibit significantly larger divergences in the contrast, capturing angles, intensity probability distribution along with stochastic features. Hence, higher entropy features (implying more information) needs to be separately processed for such datasets, thereby closely emulating the practical scenario where a UAV would capture images on a large field travelling at varying velocities rendering varying capturing angles, contrast and noise effects. As the model attains high accuracy for the contrast enhanced CNN, ResNet and attention-based CNN, the model can be tested on other pest classes as well. The background being cotton crops, makes the model a generic useful model, however other background with weeds etc. can be tested as well. The feature enhancement and attention network would imply that the proposed model can be effectively used for various lighting and capturing conditions for UAV applications.

Based on the comparative evaluation of the designed models, it can be observed that the feature extraction and feature combination model employing the DT approach attains much lesser accuracy compared to the deep learning models. The major reason for the same can be thought of the fact that while homogenous datasets may render better classification accuracy, datasets with high levels of divergences would lead to randomized information capture through handpicked features or feature combinations, therefore resulting in lesser accuracy of classification. On the other hand, deep learning models would attain much higher accuracy owing to the exhaustive features computed through convolutions in the deep hidden layers of the network. This would however come at the cost of computational complexity and execution time. The ResNet clearly performs better compared to the CNN owing to lesser redundancy in feature computation and better weight and bias updates. The proposed attention-based CNN model clearly outperforms the CNN and ResNet owing to the additional degree of freedom rendered by the entropy cost function for the features to align with the annotated class. This can be visualized as:

Consider the original training vector to be X and the attention-based training vector to be X_A. The contextual attention parameter $\sum_{i=1}^{T} \alpha_{,i} v_i$ can be considered as F_A where F_A renders the additional degree of freedom by training the attention-based CNN model in each iteration with annotated X samples as well as a random counterpart of the same F_A thereby creating the vector as:

$$X_A = [X, F_A] \quad (14)$$

Clearly, as F_A is chosen randomly in each iteration for the training, it would render higher entropy to the features thereby at some instances more information that the vector X alone. This would ultimately lead to higher classification accuracy (which is validated through experimental results), at the cost of higher computational complexity. It is worth noting that increasing the size of F_A would intuitively mean that a duplicate copy of the vector X is being used. This would result in degradation of the performance of the system. Thus, size of α_c is considered 0.1 (10%) of the original vector X to render two major benefits:

1) Preserving the randomness in the entropy function to impact the training process significantly.
2) Limiting the computational complexity of the system in terms of both number of iterations and training time.

This approach may however become less impactful as the data samples become more homogenous such as in cases of medical datasets such as chest X-Rays, Cancer cases etc. Apart from the above mentioned of the proposed attention-based CNN model, the identified limitations of the proposed work can be viewed as the dependence of the results of the proposed work on the data set used for training. With wide variations in the crop size and texture, along with variations in the whitefly pests, the effectiveness of the approach needs to be validated through rigorous tests on geographical regions across India as well as outside India where whitefly pest attacks are prevalent. Some other countries bearing geographical and topographical resemblances which may be considered are Pakistan, Bangladesh, and Indonesia.

A complete list of abbreviations is shown in *Appendix I.*

6.Conclusion and future work

This paper presents both handpicked feature extraction followed by classification using the DT approach and deep learning approaches such as the CNN, ResNet and attention-based CNN. While deep learning approaches bypass the need for handpicked feature selection and feature extraction, they typically need much larger labelled datasets in order to evaluate the traits of infested and non-infested data samples. Classification accuracy of the classifier is chosen as the major performance evaluation parameter. It has been shown that the proposed attention-based CNN outperforms the other models in terms of accuracy of classification. The contrast enhancement model has shown to increase the proposed systems performance by making it more immune and robust towards partial shading, uneven lighting and noise effects typically encountered during image capturing through UAVs, used for precision agriculture applications. A valuable future enhancement of the proposed method can be testing the proposed model on actual images captured by UAVs in the Punjab region of India along with other geographical locations being inflicted by the whiteflies such as real data collected from regions of Pakistan, Bangladesh, and Indonesia etc. Transfer learning models can also be explored to use pre-trained models on other datasets to test for variabilities in data pertaining to texture, size, capturing and orientation.

Acknowledgment

The authors acknowledge the valuable suggestion, critical comments and suggestions of the members of the department of computer science and Engineering, Punjabi University, Patiala, India which have immensely helped in shaping the paper. The authors also acknowledge the valuable suggestions form Dr. Pathania and other experts form regional research station, Punjab University of Agricultural (PAU), Bhatinda, Punjab, India.

Conflicts of interest

The authors have no conflicts of interest to declare.

Author's contribution statement

Dr. Amardeep Singh Dhiman and Dr. Sikander Singh: Developing the problem statement, methodology, and investigation. **Mr. Lal Chand:** Data collection, annotation, implementation, and compilation of results.

References

[1] Pérez-Ortiz M, Peña JM, Gutiérrez PA, Torres-Sánchez J, Hervás-Martínez C, López-Granados F. A semi-supervised system for weed mapping in sunflower crops using unmanned aerial vehicles and a crop row detection method. Applied Soft Computing. 2015; 37:533-44.

[2] Dimitriadis S, Goumopoulos C. Applying machine learning to extract new knowledge in precision agriculture applications. In panhellenic conference on informatics 2008 (pp. 100-4). IEEE.

[3] Brunelli D, Albanese A, d'Acunto D, Nardello M. Energy neutral machine learning based iot device for pest detection in precision agriculture. IEEE Internet of Things Magazine. 2019; 2(4):10-3.

[4] Pedersen SM, Lind KM. Precision agriculture–from mapping to site-specific application. Precision agriculture: Technology and Economic Perspectives. 2017:1-20.

[5] Byrne DN, Bellows Jr TS. Whitefly biology. Annual Review of Entomology. 1991; 36(1):431-57.

[6] Li Y, Qu C, Yan X, Sun X, Yin Z, Zhao H. Effect of feeding stage and density of whiteflies on subsequent aphid performance on tobacco plants. Agronomy. 2022; 12(5):1-12.

[7] Moranduzzo T, Melgani F. A SIFT-SVM method for detecting cars in UAV images. In IEEE international geoscience and remote sensing symposium 2012 (pp. 6868-71). IEEE.

[8] De Castro Pereira R, Hirose E, De Carvalho OL, Da Costa RM, Borges DL. Detection and classification of whiteflies and development stages on soybean leaves images using an improved deep learning strategy. Computers and Electronics in Agriculture. 2022; 199: 107132.

[9] Parab CU, Mwitta C, Hayes M, Schmidt JM, Riley D, Fue K, et al. Comparison of single-shot and two-shot deep neural network models for whitefly detection in IoT Web application. AgriEngineering. 2022; 4(2):507-22.

[10] Joochim O, Satharanond K, Kumkun W. Development of intelligent drone for cassava farming. in recent advances in manufacturing engineering and processes: proceedings of ICMEP 2023 (pp. 37-45). Singapore: Springer Nature Singapore.

[11] Chou CY, Chang SC, Zhong ZP, Guo MC, Hsieh MH, Peng JC, et al. Development of AIoT System for facility asparagus cultivation. Computers and Electronics in Agriculture. 2023; 126:107665.

[12] Huddar SR, Gowri S, Keerthana K, Vasanthi S, Rupanagudi SR. Novel algorithm for segmentation and automatic identification of pests on plants using image processing. In third international conference on computing, communication and networking technologies 2012 (pp. 1-5). IEEE.

[13] Legaspi KR, Sison NW, Villaverde JF. Detection and classification of whiteflies and fruit flies using YOLO. In 13th international conference on computer and automation engineering 2021 (pp. 1-4). IEEE.

[14] Pattnaik G, Parvathi K. Automatic detection and classification of tomato pests using support vector machine based on hog and lbp feature extraction technique. In progress in advanced computing and intelligent engineering: proceedings of ICACIE 2021 (pp. 49-55). Springer Singapore.

[15] Pattnaik G, Shrivastava VK, Parvathi K. Transfer learning-based framework for classification of pest in tomato plants. Applied Artificial Intelligence. 2020; 34(13):981-93.

[16] Nagar H, Sharma RS. A comprehensive survey on pest detection techniques using image processing. In 4th international conference on intelligent computing and control systems 2020 (pp. 43-8). IEEE.

[17] Rustia DJ, Lin CE, Chung JY, Zhuang YJ, Hsu JC, Lin TT. Application of an image and environmental sensor network for automated greenhouse insect pest monitoring. Journal of Asia-Pacific Entomology. 2020; 23(1):17-28.

[18] Liu L, Wang R, Xie C, Yang P, Wang F, Sudirman S, Liu W. PestNet: an end-to-end deep learning approach for large-scale multi-class pest detection and classification. IEEE Access. 2019; 7:45301-12.

[19] Chen CJ, Wu JS, Chang CY, Huang YM. Agricultural pests damage detection using deep learning. In advances in networked-based information systems: the 22nd international conference on network-based information systems 2020 (pp. 545-54). Springer International Publishing.

[20] Deng L, Wang Y, Han Z, Yu R. Research on insect pest image detection and recognition based on bio-inspired methods. Biosystems Engineering. 2018; 169:139-48.

[21] Giakoumoglou N, Pechlivani EM, Katsoulas N, Tzovaras D. White flies and black aphids detection in field vegetable crops using deep learning. In 5th international conference on image processing applications and systems 2022 (pp. 1-6). IEEE.

[22] Rajan P, Radhakrishnan B, Suresh LP. Detection and classification of pests from crop images using support vector machine. In international conference on emerging technological trends 2016 (pp. 1-6). IEEE.

[23] Gašparović M, Zrinjski M, Barković Đ, Radočaj D. An automatic method for weed mapping in oat fields based on UAV imagery. Computers and Electronics in Agriculture. 2020; 173:1-12.

[24] Potena C, Nardi D, Pretto A. Fast and accurate crop and weed identification with summarized train sets for precision agriculture. In intelligent autonomous systems 14: proceedings of the 14th international conference IAS-14 2017 (pp. 105-21). Springer International Publishing.

[25] Omrani E, Khoshnevisan B, Shamshirband S, Saboohi H, Anuar NB, Nasir MH. Potential of radial basis function-based support vector regression for apple disease detection. Measurement. 2014; 55:512-9.
[26] Cho J, Choi J, Qiao M, Ji CW, Kim HY, Uhm KB, Chon TS. Automatic identification of whiteflies, aphids and thrips in greenhouse based on image analysis. Red. 2007; 346(246):1-8.
[27] Qiao M, Lim J, Ji CW, Chung BK, Kim HY, Uhm KB, et al. Density estimation of Bemisia tabaci (Hemiptera: Aleyrodidae) in a greenhouse using sticky traps in conjunction with an image processing system. Journal of Asia-Pacific Entomology. 2008; 11(1):25-9.
[28] Wang K, Zhang S, Wang Z, Liu Z, Yang F. Mobile smart device-based vegetable disease and insect pest recognition method. Intelligent Automation & Soft Computing. 2013; 19(3):263-73.
[29] Wang QJ, Zhang SY, Dong SF, Zhang GC, Yang J, Li R, Wang HQ. Pest24: a large-scale very small object data set of agricultural pests for multi-target detection. Computers and Electronics in Agriculture. 2020; 175:1-9.
[30] Saleem MH, Potgieter J, Arif KM. Automation in agriculture by machine and deep learning techniques: a review of recent developments. Precision Agriculture. 2021; 22:2053-91.
[31] Zebari DA, Haron H, Zeebaree SR, Zeebaree DQ. Enhance the mammogram images for both segmentation and feature extraction using wavelet transform. In international conference on advanced science and engineering 2019 (pp. 100-5). IEEE.
[32] Gunawan D. Denoising images using wavelet transform. In pacific RIM conference on communications, computers and signal processing 1999 (pp. 83-5). IEEE.
[33] Abdulrahman AA, Rasheed M, Shihab S. The analytic of image processing smoothing spaces using wavelet. In journal of physics: conference series 2021 (p. 022118). IOP Publishing.
[34] Mardia KV, Hainsworth TJ. A spatial thresholding method for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1988; 10(6):919-27.
[35] Huang FC, Huang SY, Ker JW, Chen YC. High-performance SIFT hardware accelerator for real-time image feature extraction. IEEE Transactions on Circuits and Systems for Video Technology. 2011; 22(3):340-51.
[36] Wong WK, Lai Z, Wen J, Fang X, Lu Y. Low-rank embedding for robust image feature extraction. IEEE Transactions on Image Processing. 2017; 26(6):2905-17.
[37] Wang X, Bai X, Liu W, Latecki LJ. Feature context for image classification and object detection. In CVPR 2011 (pp. 961-8). IEEE.
[38] Rehman TU, Mahmud MS, Chang YK, Jin J, Shin J. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Computers and Electronics in Agriculture. 2019; 156:585-605.
[39] Burhan SA, Minhas S, Tariq A, Hassan MN. Comparative study of deep learning algorithms for disease and pest detection in rice crops. In 12th international conference on electronics, computers and artificial intelligence 2020 (pp. 1-5). IEEE.
[40] Rajesh B, Vardhan MV, Sujihelen L. Leaf disease detection and classification by decision tree. In 4th international conference on trends in electronics and informatics 2020 (pp. 705-8). IEEE.
[41] Nam NT, Hung PD. Pest detection on traps using deep convolutional neural networks. In proceedings of the 1st international conference on control and computer vision 2018 (pp. 33-8).
[42] Thanapol P, Lavangnananda K, Bouvry P, Pinel F, Leprévost F. Reducing overfitting and improving generalization in training convolutional neural network (CNN) under limited sample sizes in image recognition. In 5th international conference on information technology 2020 (pp. 300-5). IEEE.
[43] Reddy AS, Juliet DS. Transfer learning with ResNet-50 for malaria cell-image classification. In international conference on communication and signal processing 2019 (pp. 945-49). IEEE.
[44] Majid S, Alenezi F, Masood S, Ahmad M, Gündüz ES, Polat K. Attention based CNN model for fire detection and localization in real-world images. Expert Systems with Applications. 2022; 189: 116114.
[45] Gondal MD, Khan YN. Early pest detection from crop using image processing and computational intelligence. FAST-NU Research Journal. 2015; 1(1):59-68.

Lal Chand is presently serving as an Assistant Professor in the Department of Computer Science and Engineering at Punjabi University, Patiala. With over 14 years of experience in academia and research, his areas of expertise encompass Machine Learning, Computer Vision, and Computer Networks. Currently, his research focuses on the applications of Machine Learning in the field of Agriculture.
Email: Lc.panwar5876@gmail.com

Amardeep Singh Dhimanis a Professor in the Department of Computer Science and Engineering at Punjabi University, Patiala. With over 20 years of teaching experience, he has supervised 40 research scholars at the Master's level and twelve research scholars at the Doctorate level. He has authored over 40 research papers published in reputable international and national journals, and has written two books on Bioinformatics Computing and Soft Computing. Additionally, he has successfully managed several state-funded projects.
Email: Amardeep_dhiman@yahoo.com

Dr. Sikander Singh is currently working as an Assistant Professor in the Department of Computer Science and Engineering at Punjabi University, Patiala. His areas of interest encompass Computer Vision, Pattern Recognition, Deep Learning, and Data Science. He has numerous research articles published in top national and international journals.
Email: Sikander@pbi.ac.in

Appendix I

S. No.	Abbreviation	Description
1	AIoT	Artificial Intelligence of Things
2	ANN	Artificial Neural Network
3	CNN	Convolutional Neural Network
4	CSA	Channel-Spatial Attention
5	DPI	Dots Per Inch
6	DT	Decision Tree
7	DWT	Discrete Wavelet Transform
8	FN	False Negative
9	FP	False Positive
10	GUI	Graphical User Interface
11	HOG	Histogram of Oriented Gradients
12	IPM	Integrated Pest Management
13	IoT	Internet of Things
14	LBP	Local Binary Pattern
15	MAPE	Mean Absolute Percentage Error
16	ML	Machine Learning
17	RCNN	Regions with Convolutional Neural Networks
18	RPN	Regional Proposal Network
19	ResNet	Residual Network
20	RMS	Root Mean Square
21	RMSE	Root Mean Squared Error
22	SIFT	Scale Invariant Feature Transform
23	SVM	Support Vector Machine
24	UAV	Unmanned Aerial Vehicles
25	WSN	Wireless Sensor Networks
26	YOLO	You Only Look Once
27	YST	Yellow-Sticky Tape