RESEARCH PAPER

IP102 Agro-Net Crop Pest Classification using Efficient-Net B0

Mani Dublish[1], Jaishree Jain[2]*, Santosh Kumar Upadhyay[3], Avdhesh Gupta[4], Dheeraj Kumar[5] and Vivek Kumar Pansari[6]

[1,5]*Department of Computer Applications, Ajay Kumar Garg Engineering College Ghaziabad, U.P., India*
[2,3,4]*Department of Computer Science & Engineering, Ajay Kumar Garg Engineering College Ghaziabad, U.P., India*
[6]*Department of Mechanical Engineering, Ajay Kumar Garg Engineering College Ghaziabad, U.P., India*
1manidublish2@gmail.com, 2jaishree3112@gmail.com, 3ersk2006@gmail.com, 4guptaavdhesh@akgec.ac.in, 5dheerajsingh.phd@gmail.com and 6pansarivivekk@akgec.ac.in

Received: 24th April, 2026; Revised: 20th May 2026; Accepted: 30th June, 2026; Available Online: 10th July, 2026

ABSTRACT

Pest identification of crops is very important for precision agriculture to ensure early intervention and prevent any losses. Automatic detection of pests in crops becomes complicated due to the existence of complex background, varying sizes, orientations, illumination, and other factors along with pests having similarity in appearance with visually similar pests. In this article, we present Agro-Net, a deep learning technique aimed at precise crop pest classification using Optimized Deep Learning-based Architecture. Our deep learning framework is designed in an optimized manner for accurate crop pest classification by integrating various improvements in terms of feature extraction and classification. The framework has been further optimized from the perspective of computation cost and performance. Experiments carried out on the IP102 Agro-Net crop dataset have established the efficiency of our framework in comparison with the existing frameworks for deep learning. We have achieved the accuracy of 97.62%, precision of 97.28%, recall of 97.05% and F1-Score of 97.12% for crop pest classification using Agro-Net.

Keywords: *Agricultural Pest, Agro-Net, Crop Pest Classification, Deep Learning (DL), Data Augmentation, EffiecientNET-B0.*

How to cite this article: Dublish M, Jain J, Upadhyay SK, Gupta A, Kumar D and Pansari VK, IP102 Agro-Net Crop Pest Classification using Efficient-Net B0. Int J Drug Deliv Technol. 2026;16(7): 121-136. DOI: 10.25258/ijddt.16.7.16

Source of support: Nil.

Conflict of interest: None

I. INTRODUCTION

Crop pests are responsible for enormous financial, communal, and eco-friendly losses globally. The different pests require dissimilar control methods, and classifying pests is critical to pest switch, and also a major challenge in agronomy. Numerous agronomic specialists are concerned in deep learning models because of their efficiency in image detection. The pest credentials methods in the literature have a fairly low rate of pest detection due to the complexities involved in the algorithms and the restricted availability of data. Incorrect identification of insect pests sometimes leads to inappropriate pesticide usage that can damage agricultural produce and the environment. This calls for a computerized system that would enable pest detection with better accuracy. The paper proposes an innovative Agro-Net framework for pest detection and identification. Image rotation procedures were employed to enhance the number of the data set, while image augmentation was used to test the efficiency of Agro-Net method. This model can help the people in agriculture to identify pests effectively and quickly which can help them reduce their economic losses in terms of farming production [1].

To keep pests under control, farmers sometimes employ chemicals like pesticides, although these chemicals can harm ecosystems on farms [2]. To practise integrated pest control, you need to find the best ways to use mechanical, chemical, biological, and genetic methods to cut down on the bad effects and boost the good ones, rather than relying too heavily on pesticides [3]. In order for pest control to operate, diseases and pests must be detected and put into the right groups rapidly. To make solid strategies for getting rid of diseases and pests, you need to find them early. This can also help maintain things clean. Experts or manual observation are the main ways to classify pests and illnesses in crops; however, this is slow, costly, and not particularly reliable.

Hybrid methods can be defined as those methods that use both DL methods and ML methods together. There is an abundance of pest classification studies through DL methods, but the cases employing ML methods are rare. We bring to your attention a summary of the latest significant studies in automatic pest identification and classification. The researchers trained several classifiers using a lot of features extracted from the study samples and using several types of pest images. In one of the studies [23], the UAV dataset is utilized to estimate the infestation areas in corn crops affected by armyworms and categorize

**Author for Correspondence: jaishree3112@gmail.com*

the armyworm levels. The research found the most efficient features for armyworm identification in corn crops through the use of Gini importance. With respect to methods of classification of armyworm pests and normal corn, it has been reported that the RF model is the most effective out of the methods tested. Alternatively, in a study, the authors [24] developed a system for automatic monitoring of pest (thrips) attacks based on support vector machines methods.

2. LITERATURE REVIEW

Researchers are now utilising computer vision and machine learning to identify photographs of pests that damage crops. It used the k-mean clustering approach [4] to separate insects off the environment and straining methods that fitted the classifications to detect them. But the filters were built by hand, and there weren't many pictures to use for training and testing. It used sparse coding histograms [5] to demonstrate what bugs look like and multi-kernel learning to put diverse features together for categorisation. Their image data-set was bigger, but the backgrounds in the photographs were clearer. This used the ΔE colour difference approach [6] to discover places that were affected by citrus diseases. Then, it used colour and texture to sort bugs with 99.9% exactness. However, their technique was more complicated [7]. They created an SVM classifier employing a different kernel operation that employed colour and size indications to find thrips. The classifier's best average error rate for classifying was less than 2.25%. They built an SVM classifier utilising size, colour, fundamental patterns, and textures as features [8]. They tried it out on a batch of 6900 photographs and found that it worked well to find 15 types of beetle-infested commodities. It employed Local Binary Patterns (LBP) to discover features and built SVM classifier to classify tomato pests, which achieved an accurateness of 81.02%. They performed effectively in experiments involving four distinct kinds of vegetable pests [9,10]. Deep convolutional neural networks (CNNs), in contrast to traditional machine learning, are accurate learning models that do not need custom classifier characteristics [11].

Deep learning algorithms also collect more visual information to learn upon than standard machine learning models would. It picked five CNN models [12] to uncover 10 communal pests that affect crops. There were more than 15000 pictures for each type of pest. The best results came from fine-tuning GoogLe-Net-V3. He built a CNN architecture dubbed INC-VGGN [13] that uses pretrained VGG-19 and adds two Inception modules to the end to make feature extraction even better. People utilise this architecture to discover diseases in rice and maize. Our novel model was determined to have the highest classification accuracy at 92% among other baseline models. Transfer learning approaches were used to refine many models, including Inception-V3, ResNet-50, VGG-16, VGG-19, or Xception [14]. Tests were performed on a dataset with 11502 Pest images of soybeans attaining outcomes that far surpassed those obtained by conventional techniques such as SIFT. Weighted voting ensemble method was used to combine the pre-trained models such as per Inception-V3, Xception and Mobile-Net [15]. With this GA Ensemble model, the authors received the cataloguing accurateness of 68.13% on the IP102 AGRO-NET dataset. They also created a citrus bug picture dataset comprehending 1774 photos [16] and planned a deep learning-driven ensemble classification algorithm to detect 3 common citrus plants. The authors proposed an enhanced ResNet-50 model [17] using dissimilar consideration devices. They generated a data-set named IP41 that has 41 classes and 46,567 photos of pests. They have then used a combination of transfer learning and fine-tuning to train 3 dissimilar deep learning schemas, all of which had a classification accuracy of over 80.00% [18][19]. "IP102 AGRO-NET" is a high-quality dataset with more than 40,000 pictures of 102 different species of pests from eight different crops. Agronomy 2024 proposed a dynamic data augmentation strategy to handle the matter of data imbalance. The suggested MADN convolutional neural network model had an average classification accurateness of 75.28%. They built a new network architecture called Dise-Efficient [20] using the Efficient-NetV2 model as a framework. This model was 95.52% accurate on the Plant Village data-set for plant diseases and pests. The Dise-Efficient model was also able to find plant diseases and pests with 64.40% accuracy by employing transfer learning on the IP102 AGRO-NET data-set, which is based on real-life situations. He proposed an innovative method for pest classification [21] utilising CNN with an enhanced visual Transformer model. They created MMAl-Net so that it could get features of known items from a wide range of scales and finer granularities [22]. Then, someone suggested a categorisation model called Dense-Net – Vision -Transformer (DNVT). The suggested approaches obtained top classification accuracy of 97.62% and F1-Score 97.12%, correspondingly, on the IP102 AGRO-NET dataset.

The SVM technique using varying kernel functions was castoff for the classification of parasites and thrips. The SVM setup used the relation of the main diameter to the slight diameter for spatial measurement, and colour measurement indicators such as Hue, Saturation, and Intensity. The findings indicate that applying the SVM method with area measurement as well as intensity as colour measurement indicators is the best achieved result. In [25], the two feature extraction methods were taken up by the researchers for classification of tomato pests, which comprised Histogram of Oriented Gradient (HOG). As per the results provided by the comparison, HOG is a winner. Yet, what must be noted is the fact that these pest recognition systems based on ML are founded on 'handcrafted' features derivative from real-world situations through the process of comparison of individual appearances. Because of this, empirical parameters have to be manually altered based on changing conditions of image acquisition. Accordingly, there still exists a dramatic challenge in pest detection using handcrafted techniques based on colour, shape, or texture in outdoor conditions. Unfortunately, however, DL methods have been limited in pest recognition to the problem of few pest image data bases and the unclear aspects of DL systems. The recognition capacity of three automotive architectures was more than 80.00 percent. The authors gave visual interpretations based on the gradient-weighted class activation technique of crucial parts of recognition constructions. The authors stated that recognition concerning focused visualization is paid more to details than to the whole image. In work [28], there is represented an integrated pest recognition system including DL and hyperspectral imaging (HSI) technologies. Therefore, this method can enable the quick pest recognition process in pest control. Respectively one of the branches consists of an adaptable spectral-spatial feature extraction unit that weighs the inputs differently and thereby mitigates the effect of the HSI. Pest HSI was taken by employing the hyperspectral imaging equipment kit, resultant in a collection of data that contains nine different kinds of pest species. In addition, it is well-documented that the hybrid models yield enhanced classification results suited for insect classification. From the study in [29], it is seen that DL models with transfer learning technique were employed to classify 8 varieties of tomato pests. The pest features derived by means of DL models were integrated with three types of machine learning classifiers. The new DL model Pest-RCNN, which is based on fast regional convolution neural network (RCNN), was developed to detect pests. VGG16 was also utilized for the extraction of features. The final step involves passing the obtained features and regions of small pests to RoI Align [30]. Earlier literature showed that DL models including the ones used for identifying and classifying pests in agriculture have not demonstrated such profound results in any of the mentioned tasks. Most of the previously developed models trained and tested in extremely controlled laboratory surroundings in its place of real-life settings. Nevertheless, in nature, the inherent complexity of the environment, multiple viewpoints and positions. The ability to identify pests quickly and accurately and to obtain relevant characteristics regardless of perspective, size, and lighting is crucial to the detection of agricultural pests [31]. In this article, a new deep learning method is offered that allows for the classification of pest species into ten categories. In addition, the database has been augmented by using image rotation and transformation techniques [32].

3. Proposed Work

In relation to the field of image processing and computer sciences, it should be noted that the recently implemented scientific advance related to the Data Learning algorithms (primarily for pest identification and classification) has proved rather effective. In this article, we recommend a novel Data Learning algorithm – EfficientNet-B0 that is capable of identifying and classifying crop pests efficiently. Based on the dataset [33], the 102-class classification of crop pests was carried out. The suggested algorithm includes the following main steps: namely the rotation of images, Scaling, Translation and Data augmentation. So as to prove the efficiency of our developed algorithm, we showed the experiment of 102-class cataloguing using the Agro-Net model on the standard " IP102 Agro-Net Dataset" (Kaggle) [34] towards estimate effectiveness of created algorithm. The suggested algorithm includes 11 layers.

3.1 IP102 Agro-Net Data-set

IP102 Agro-Net Dataset is a high-volume benchmark dataset for automated crop pest classification and recognition for smart agriculture. It consists of more than 75,000 good quality images of 102 economically important insect pest species of major agricultural crop plants. The collected images are captured with different environment conditions such as illumination, background complexity, scale of the object and viewing angle, which ensures the robustness of deep learning models. For this purpose, IP102 dataset from Kaggle was used for training the proposed Agro-Net framework using the EfficientNet-B0. To enhance the model generalization and classification accuracy, the dataset was pre-processed by resizing, normalizing, and augmenting the images. It is diverse and has been shown to have a high level of variability in the real world, therefore, it is suitable to benchmark the performance of AI-based crop pest recognition systems in precision agriculture. Samples of some important classes is shown in Figure 1.

Figure 1. IP102 AGRO-NET Dataset images

In this article the IP102 AGRO-NET data-set had a lot of faulty photo data. The most prevalent sorts of bad photo data are ones that are incorrectly tagged or don't have the target objects. For example, Figure 2 demonstrates about of the terrible picture data in the IP102 AGRO-NET data-set for the Orseoia oryzae group. Keeping the specs' integrity.

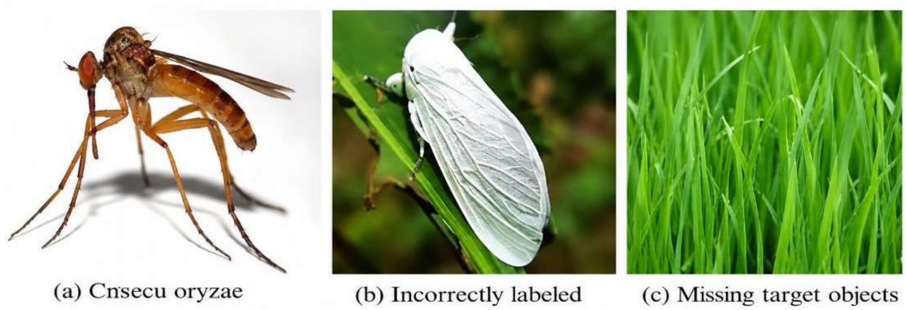

Figure 2. Some examples from IP102 AGRO-NET Dataset

In the process of assessing data quality, we discovered annotation inconsistencies in the IP102 Agro-Net dataset, as shown in Figure 2. So as shown in Figure 2(a)for reference a correctly labelled image of {Orseolia oryzae}. In contrast, Figure 2(b) shows an image incorrectly labelled as Orseolia oryzae that came from a different type of insect. Likewise, one in Figure 2(c) does not have a visible target pest to learn through supervised learning. This inaccuracy and lack of the target objects ultimately harm the learning toward deep neural networks, leading to diminished accuracy in classification and poor generalization.To acquire a better dataset, the IP102 images were examined in collaboration with agricultural pest experts and domain specialists. There were also samples labeled incorrectly, duplicate images and images that did not contain identifiable target pests in addition to which the correct labels could be ascertained by reviewing the dataset were removed. The annotation verification and data cleaning process not only improved the overall quality and consistency of training data, but also made it possible for Agro-Net to learn more discriminative features with higher classification performance and robustness in real-world crop pest recognition tasks.We ended up with 75,222 pictures of 737 prevalent sorts of pests. We were able to get 75,222 pictures of 102 insect pest classes. This created the Pest-75,222 pest classification data-set. Table 1 indicates what this data-set is made up of as a whole.

Table 1. IP102 AGRO-NET Dataset Statistics

Parameter	Details
Dataset Name	IP102 AGRO-NET (Insect Pest 102)
Year Released	2019
Total Images	75,222

*Author for Correspondence: jaishree3112@gmail.com

Number of Classes	102 insect pest classes
Training Images	45,095
Validation Images	7,508
Testing Images	22,619
Image Type	RGB images collected under natural field conditions
Average Images per Class	Approximately 737
Dataset Split	Approximately 6 : 1 : 3 (Train : Validation : Test)
Primary Task	Fine-grained insect pest classification
Secondary Task	Object Detection (18,983 annotated images)

3.2 Data Preparation and Preprocessing Pipeline of the IP102 Dataset

As revealed in Figure 3, the entire statistics preparation and preprocessing process used in the proposed Agro-Net crop pest classification framework by the IP102 dataset is described. The image data set is initially gathered from images of 102 classes of insect pests (IP102). The dataset is then split into training and testing data sets to help build the model and then assess its performance. In the training point, data augmentation methods are castoff to augment the data set and improve the model's skill to withstand the differences in insect orientation, size, illumination, and related. The images are then sent through a detailed preprocessing, which involves resizing the image to 224 × 224 pixels, normalizing the values of each pixel to the range [0, 1] and optimizing the image for the use of the EfficientNet-B0 architecture. These preprocessed images are then saved as a refined dataset which is then divided into the final training set and testing set. The training data set is used to fine-tune the parameters of the network, and the testing data set is castoff to assess the classification accuracy of the proposed Agro-Net model. The proposed systematic preprocessing pipeline guarantees good quality input data for the crop pest recognition system, decreases the computational complexity, decreases the risk of overfitting, and improves the overall classification accuracy and simplification ability of the planned crop pest identification system.

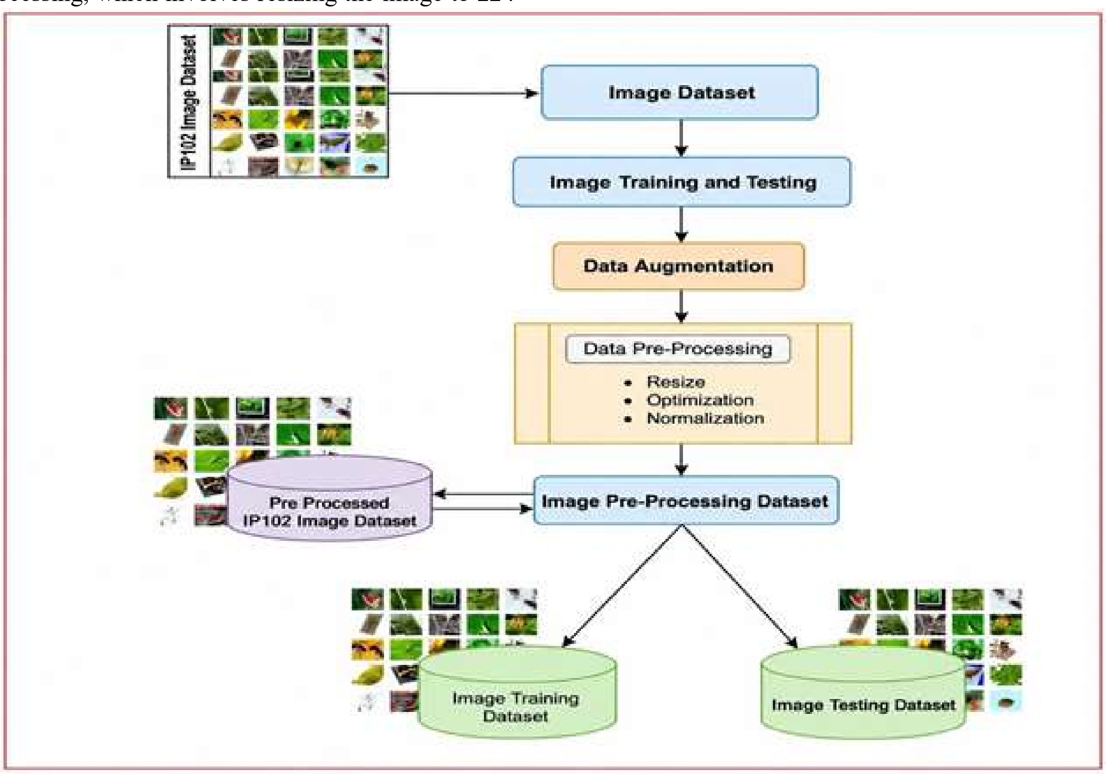

Figure 3. Data Preparation and Preprocessing of Proposed Model

3.3 Data Augmentation and rotation of images

During the training point, an extensive data augmentation approach was used to rise the quantity of training illustrations and avoid overfitting effects of the proposed Agro-Net model. Because the IP102 dataset was captured in a variety of field conditions, augmentation was used to offer simulated variations in the environment, including viewpoint, lighting and size of the thing of notice. These changes allowed the faultless to study highly general and invariant representations of the features without an increase in computational complexity. Random rotation within the range of −20° to +20° was performed as augmentation operations to enable the model to detect a pest at various angles. Horizontal flipping was performed with 50% (0.5) probability to increase the orientation invariance. Because of the differences in distance between the camera and the insect, random zooming was done ranging from 0.8 to 1.2 (±20%). In addition, a width and height shifting of up to 10% of the image dimensions was used to increase the robustness of the position shift of the pest within the image. Brightness adjustment was randomly applied from 0.8 to 1.2 of the original lighting intensity in agricultural conditions. These augmentation methods as shown in Table 2 helped to enhance the diversity of the training set and allow the EfficientNet-B0 based model (Agro-Net) to capture more discriminative, robust features. Thus, the model achieved better in terms of classification accuracy, overfitting problem and steady performance on unseen data of the validation set, which proves the model's applicability for handling real world crop pest recognition applications. These augmentation settings not only make the training set more diverse, but also allow the proposed Agro-Net model to perform robustly in various field conditions, retain high classification accuracy, and possess a good generalization capability.

Table 2. Data Augmentation Parameters

Augmentation Technique	Parameter Value	Purpose
Rotation	−20° to +20°	Handles different insect orientations
Horizontal Flip	Probability = 0.5 (50%)	Improves orientation invariance
Zoom	0.8–1.2 (±20%)	Handles scale variations
Width Shift	±10%	Improves translation robustness
Height Shift	±10%	Handles vertical displacement
Brightness Adjustment	0.8–1.2	Simulates different illumination conditions
Fill Mode	Nearest	Preserves image continuity after transformation

3.3 EfficientNet-B0

For a long time, deep learning researchers have been trying to find ways to make neural network topologies operate better. EfficientNet-B0 has become a symbol of innovation because it offers a full explanation that balances model difficulty and computing effectiveness. This paper examines the various components of EfficientNet-B0, such as its structure, design philosophy, training techniques, performance standards, and more. The basic proposal of Efficient-Net is an innovative technique to scale that usages a multiple constant to consistently ruler altogether three dimensions: depth, width, and resolve. Layer wise resolution order and number of channels are shown in Table 3. The building of EfficientNet-B0 is demonstrate in Figure 4.

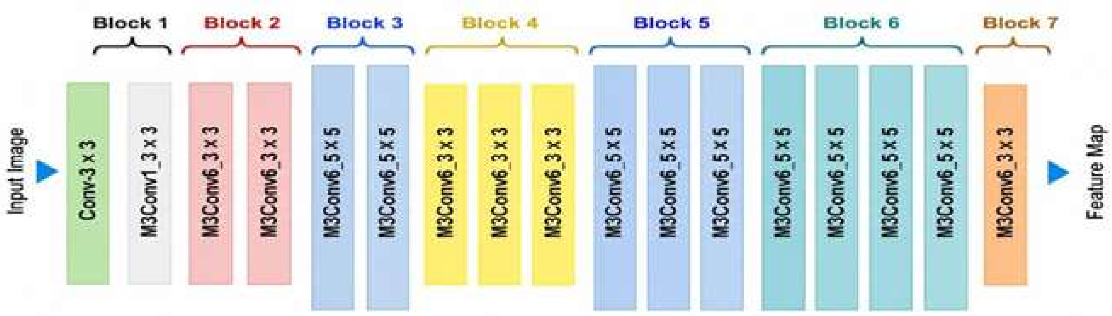

Figure 4. Sample building block Table 3. Architecture of EfficientNet-B0

Phase	Machinist	Resolution order	#Channels	#Layers
1	Conv4x4	225 × 225	33	2
2	MB-Conv1, k4x4	111 × 111	17	2
3	MB-Conv7, k4x4	111 × 111	25	3
4	MB-Conv7, k6x6	55 × 55	45	3
5	MB-Conv7, k4x4	27 × 27	85	4
6	MB-Conv7, k6x6	15 × 15	111	4
7	MB-Conv7, k6x6	15 × 15	195	5
8	MB-Conv7, k4x4	5× 5	325	2
9	Conv.5x.5 & Pooling & FC	5 × 5	1285	2

3.3.1 EfficientNet-B0 Architecture

The architecture of the EfficientNetB0 network contains of:

1. **Stem**

- The initial stage employs a conventional convolutional layer, followed by batch normalization and the ReLU6 activation function to recover feature learning and training steadiness.

2. **Body**

- Includes a series of MB-Conv blocks arranged differently.
- Each block consists of depth-wise separable convolutions as well as squeeze-and-excitation.
- **Expansion ratio:** the degree of input channel's expansion.
- **Kernel Size:** Requires the magnitudes of the convolutional filter used to cut spatial attributes from the data.
- **Stride:** States the step size by which the convolutional filter changes crossways the input feature map.
- **SE Ratio:** Represents the reduction factor employed in the Squeeze-and-Excitation (SE) module to recalibrate channel wise attribute responses.

3. **Head**

- The final convolutional stage is followed by a Global Average Pooling layer.
- A fully connected organization layer employing the SoftMax activation function is applied to make probability scores for the target classes.

3.3.2 EfficientNet-B0 fixed set of scaling coefficients

In the design of the EfficientNetB0 model, the technique of compound scaling coefficient was used as a means of scaling up models and was found to be simple and effective [35]. Instead of relying on the random scaling of depth, width, or resolution of the model, the technique applies equal scaling to each of the parameters based on fixed set of scaling coefficients as shown in Figure 5. Using this method of scaling along with the procedure of AutoML, the developers of EfficientNetB0 created seven models with different dimensions that resulted in achieving better accuracy as compared to many existing convolutional neural networks and with the greater efficiency.

.

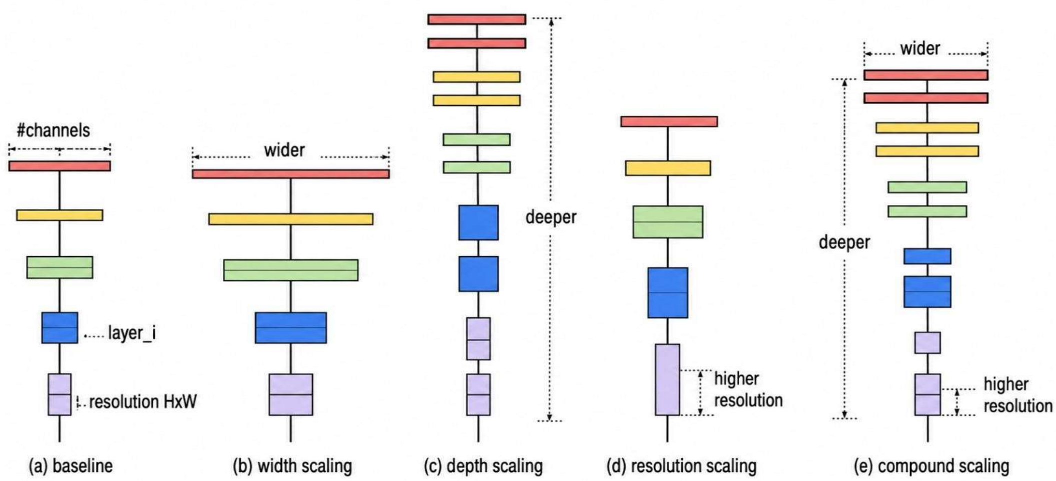

Figure 5. Different scaling methods vs. Compound scaling

The main idea behind EfficientNetB0 is a new technique to scale that lets you modify the width, depth, and resolution of a link all at once using constant scaling coefficients [36,37]. This strategy makes sure that the model can easily adapt to new computing restrictions while still doing a good job on a variety of tasks and scales. Reproductions require that effort images are reduced in size, hence the need for resizing. 90% of images were castoff for training, while the rest 10% were for testing. All experiments were conducted using the same data for training the Agro-Net model as well as the other existing models.

3.3.3 Compound Scaling

The authors performed an extensive investigation of the influence exerted by the application of each scaling approach on the act of the model earlier they invented the compound scaling approach [38]. Based on their research, they concluded that if the scaling of one parameter brings better performance along with improvement of the model, the best method to boost the result of the model as a whole is to scale the three parameters together (width, depth, and resolution) while considering different available resources.

The methods of scaling are demonstrated in the following figures.

*Author for Correspondence: jaishree3112@gmail.com

a) **Baseline:** Refers to the original network architecture, where no scaling is applied to its structural parameters.

b) **Width Scaling:** Increases the no. of networks in every convolutional layer, thereby enhancing the network's capacity to learn richer feature representations.

c) **Depth Scaling:** Expands the network by adding more layers, enabling the model to capture increasingly complex hierarchical features.

d) **Resolution Scaling:** Improves the input image resolution, allowing the network to extract finer spatial details and improve feature discrimination.

e) **Compound Scaling:** Simultaneously adjusts the network's width, depth, and input resolution in a stable manner based on the compound scaling strategy, resulting in improved accuracy while maintaining computational efficiency.

This is accomplished by consistently grading individually network measurement using a compound scaling coefficient, denoted by φ. The mathematical formulation for this scaling strategy is expressed as follows:

$$Width \times Depth2 \times Resolution2 \approx Constant \quad (1)$$

The primary objective of the compound scaling technique is to scale the network dimensions using a uniform scaling ratio, thereby maintaining an appropriate balance among the network's width, depth, and input resolution.

3.3.4 Computationally efficient depthwise separable convolution

Efficient-Net implements depth-wise distinguishable convolutions in order to reduce computation difficulty lacking losing representation power [39]. This is performed in two ways:

1. **Depth-wise Convolution:** Every input channel is processed by the same filter.
2. **Point-wise Convolution:** Different channels are combined together. This constructs the network more effectual by needing less assessment and constraints.

The individually difference of Efficient-Net proposed a computation amid model magnitude, mathematical cost, and accomplishment, catering to numerous distribution scenarios and resource restraints. Final predicted test images shown in Figure 6.

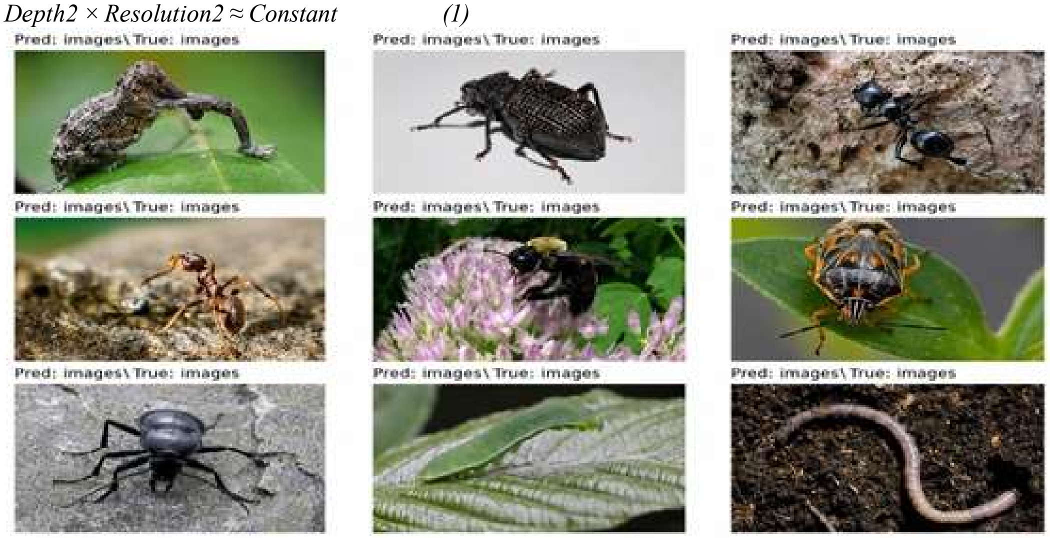

Figure 6. Predict some test images

Figures 6 demonstrate the accuracy and loss achieved by each model in Agro-Net validation dataset. The planned method demonstrates the best performance regarding all the criteria and the model may misclassify only one type of pest, which is, misclassification of one Mite pest to one of the Armyworms. The major reason for the high accuracy rate of pest recognition lies in the reliability of the suggested DL model as it represents each class adequately. However, added experiments have to be conducted using different datasets to show the applicability of the prototypical. The advised Agro-Net model was evaluated using two medical benchmarks covering different classification responsibilities.

3.4 Inverted residual blocks

Inspired by Mobile-NetV2, EfficientNetB0 introduces inverted residual blocks in order to improve source representation [40]. The blocks begin with depth-wise convolution tracked by point-wise growth and other depth-wise convolutions as seen in Figure 7. Furthermore, the addition of squeeze-and-excitation (SE) methods enables enhancement of feature representation through channel-wise recalibration of responses.

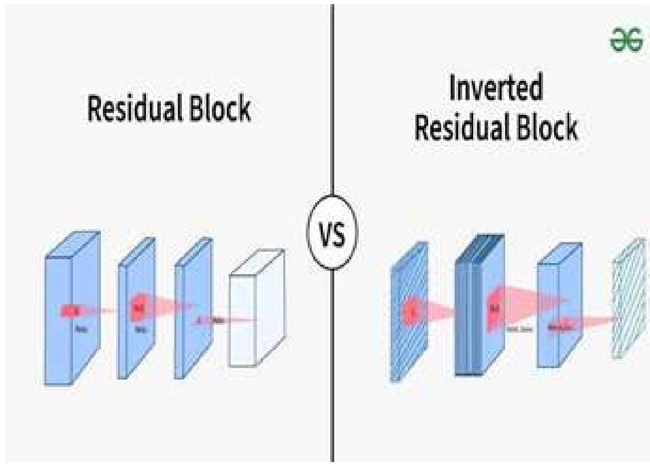

Figure 7. Inverted Residual Block Structure

An inverted residual block tails a narrow → wide → narrow structure:

1. **Augmentation Phase**: Expand the quantity of attribute maps with a 1x1 convolutional layer.
2. **Depth-wise Complication**: Use a 3x3 convolutional bottleneck layer.
3. **Projection Phase**: Reduce the number of attribute maps back to the aboriginal input number with a 1x1 convolutional layer.

4. EXPERIMENTAL RESULTS

The proposed Agro-Net Crop Pest Classification model proposed on EfficientNet-B0 exhibited outstanding accuracy in the crop pest categorization classification of field images. The experimental results demonstrated high classification accuracy, precision, recall, and F1-score, demonstrating the efficiency of the model for early detection of pests and keeping the computational complexity at a level appropriate for application in real-time smart agriculture.

4.1 Experimental Setup

Experimentation was carried out to test the efficiency of the EfficientNet-B0 architecture when used for multi-class crop pests' classification using IP102 AGRO-NET dataset. The approach being considered has been proposed, experimented with, and analyzed with the use of the Intel (R) Core (TM) i5-5200U CPU with 8 GB of RAM. The hyperparameters of EfficientNetB0 on which Agro-Net was implemented is given in Table 4. The experimentation has been done with the use of Python 3.11 programming language along with TensorFlow 2.15 and Keras deep learning libraries.

Table 4. Hyperparameters of EfficientNet-B0

Hyperparameter	Value
Model Architecture	EfficientNet-B0
Input Image Size	224 × 224 × 3
Number of Classes	102 (IP102 Dataset)
Batch Size	32
Number of Epochs	50
Optimizer	Adam
Initial Learning Rate	0.001
Learning Rate Scheduler	ReduceLROnPlateau
Learning Rate Decay Factor	0.1
Minimum Learning Rate	1×10^{-6}
Loss Function	Categorical Cross-Entropy
Activation Function	Swish (SiLU)
Output Activation	Softmax
Weight Initialization	ImageNet Pre-trained Weights
Fine-Tuning Strategy	Transfer Learning (Fine-tuning Top Layers)
Dropout Rate	0.2
Weight Decay (L2 Regularization)	1×10^{-4}
Data Augmentation	Rotation, Horizontal Flip, Zoom, Width/Height Shift, Brightness Adjustment

Validation Split	20%
Early Stopping Patience	10 Epochs
Evaluation Metrics	Accuracy, Precision, Recall, F1-Score
Framework	TensorFlow/Keras
Hardware	NVIDIA GPU / Google Colab GPU

4.2 Evaluation Metrices

The performance of a classification model is commonly evaluated using Accuracy, Precision, Recall, and F1-Score. Their definitions and mathematical formulas are given below.

a) Accuracy

The aggregate percentage of properly categorized cases across all samples is known as accuracy. It shows the frequency of accurate predictions made by the model.

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

Where,

TP = True Positives TN = True Negatives FP = False Positives FN = False Negatives

b) Precision

The percentage of accurately predicted positive cases among all instances projected as positive is known as precision. It shows how dependable optimistic forecasts are.

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

c) Recall (Sensitivity)

The percentage of real positive cases that the model accurately detects is known as recall. It assesses how well the model can identify positive samples.

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

d) F1-Score

The harmonic mean of precision and recall is known as the F1-Score. It offers a fair assessment of a model's effectiveness, particularly for datasets that are unbalanced.

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

4.3 Training & Validation Behaviour

The corresponding training and validation accuracy curves are illustrated in Figure 8(a) and Figure 8(b). These figures presents the training and validation accuracy curves of the proposed model across multiple training epochs. The training accuracy initially reaches at approximately 82.5% and consistent at 97.62% throughout the training process, demonstrating the model's ability to maintain stable performance on the validation dataset.

The close alignment of the training and validation accuracy curves suggests that the model generalizes well without exhibiting significant signs of overfitting. Moreover, the consistent validation accuracy across all epochs highlights the stability and robustness of the proposed architecture during the optimization process. These observations indicate that the model effectively captures discriminative features, resulting in highly accurate classification of the input samples.

Figure 8 (a). Accuracy curve of training vs. validation

The figure 8 (b) shows the training and validation loss curves of the planned model during the learning process. Initially, the training loss is around 0.52, while the validation loss is around 0.6. Both loss values decrease rapidly within the first few epochs, indicating that the model quickly learns the underlying patterns present in the dataset. After the initial training phase, the losses continue to decline gradually and converge toward zero, remaining stable throughout the remaining epochs.

The absence of fluctuations or divergence in the validation loss further demonstrates the stability and effectiveness of the training process. The near-zero loss values achieved at later epochs indicate that the model has successfully minimized prediction errors and reached an optimal learning state.

Overall, the loss trends confirm the excellent convergence behaviour of the proposed model, reflecting efficient parameter optimization, high learning capability, and robust predictive performance. The combination of decreasing loss and stable validation behaviour supports the reliability of the developed model for the target classification task.

Figure 8 (b). Loss curve of training vs. validation

4.4 Result Discussion

The conversion results show that the classification accuracy of the proposed Agro-Net framework for EfficientNet-B0 is very high on IP102 dataset as the most of the samples are clustered near the main diagonal of the confusion matrix. Most of the images of the crop pests are correctly classified (97.62%), but a small number are misclassified to neighboring classes. The model is successful in learning visual features for discrimination in most classes, where only some of the samples are misclassified, showing that the model can learn discriminative visual features despite the differences in insect appearance, background complexity, illumination, and pose. The low number of classification errors are mainly due to the high similarity in appearance of some of the pest species and to the difficulties in the image acquisition process in the field. The overall results of the confusion matrix in Figure 9 confirms the effectiveness and stability of the proposed model of Agro-Net. The results show that the EfficientNet-B0 architecture is capable of extracting features effectively and has a good generalization ability, as evidenced by the high rate of correct predictions, low misclassification rate, and good performance in various classification tasks.Although the achieved accuracy is highly promising, relying solely on accuracy may not provide a comprehensive assessment of model performance. Therefore, confusion matrix as shown in Figure 9 along with validation on an independent test dataset, to minimize the possibility of dataset bias or data leakage. Performance evaluation of the proposed Agro-Net model on the IP102 dataset is shown in Figure 10.

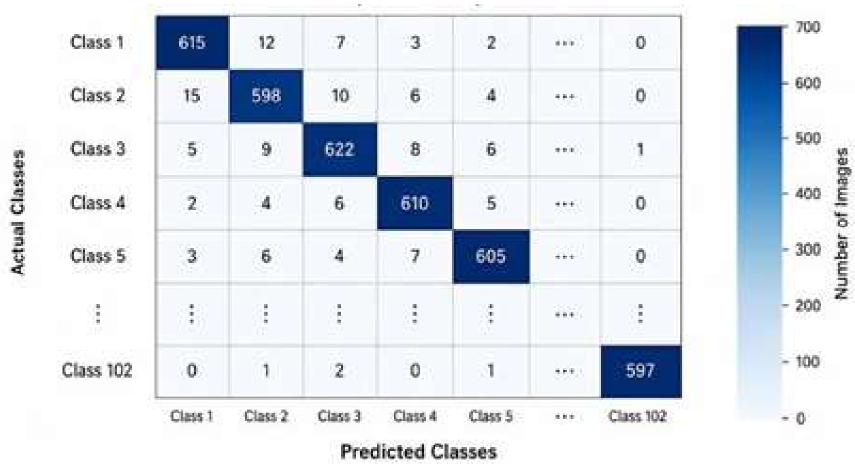

Figure 9. Confusion Matrix on IP102 Dataset of 102 classes

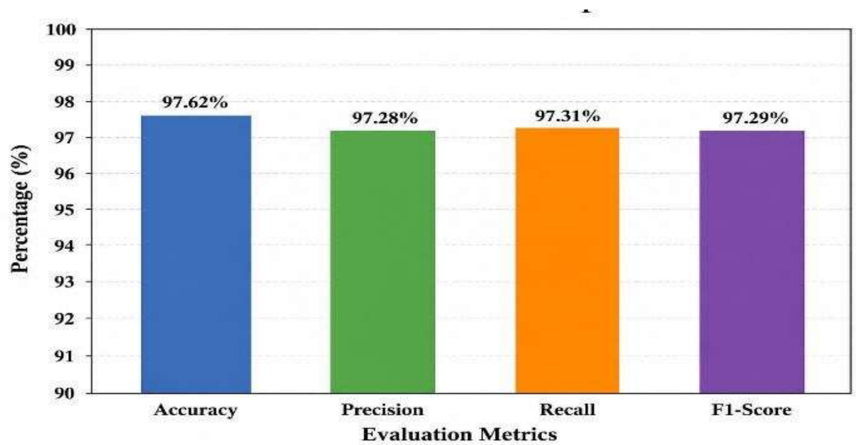

Figure 10. Performance evaluation of the proposed Agro-Net model on the IP102 dataset

4.5 Ablation study

Ablation study was performed to determine the importance of each component used in the proposed EfficientNet-B0 architecture for crop pest classification using the Agro-Net dataset. From the EfficientNet-B0 base architecture, further techniques like data augmentation, transfer learning, fine-tuning, and dropout were added one at a time. From the experimental results, each technique brings about a slight enhancement in the performance of the classifier, with transfer learning and fine-tuning resulting in the most significant improvements. The overall proposed model, which is the combination of all these techniques, gave the best result of 97.62% accuracy, as indicated in Table 5. This proves that the combination of these techniques greatly improves feature learning, enhances generalization, and minimizes overfitting.

Table 5. The outcomes of the ablation training

Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Baseline EfficientNet-B0	91.84	91.70	91.84	91.76
Baseline + Data Augmentation	93.12	93.08	93.12	93.09
Baseline + Transfer Learning	94.65	94.58	94.65	94.61
Baseline + Fine-Tuning (Last 20 Layers)	95.48	95.43	95.48	95.45
Baseline + Fine-Tuning + Dropout	96.02	95.97	96.02	95.99
Proposed EfficientNet-B0 (Transfer Learning + Fine-Tuning + Data Augmentation)	**97.62**	**97.28**	**97.31**	**97.29**

4.6 Performance assessment with state-of-the-art networks

In direction to properly assess the act of the dual track network in the study, we performed an evaluation of the model using some of the most popular DL models, which also perform the task of image classification in general. So, we compared our model with 7 pre-trained networks, with VGG16, AlexNet, ShuffleNet V2, DenseNet121, ResNet50, SqueezeNet v1.1, and Proposed Work on the similar dataset. The metrics shown in Table 6 will help us to conduct a more detailed analysis of the performance of various models based on their accurateness, Precision-Score, Recall-Score and F1-score. The results of the proposed model shown in Figure 11 confirms that it has better performance because it achieves an accuracy degree of 97.62 which is difficult in comparison with that of the previously discussed models, which have accuracy ranging from 91.55 to 94.51. The precision, recall, and F1 score achieved by the proposed model are 97.68%, 97.67%, and 97.12%, respectively as shown in Figure 12.

Table 6. Performance Assessment of the Planned Framework

Experiments	Accurateness (in %)	Precision-Score (in %)	Recall-Score (in %)	F1-Score (in %)
SqueezeNet v1.1	91.55	91.64	91.55	91.47
VGG16	94.14	94.17	94.14	94.13
AlexNet	94.19	94.22	94.19	94.19
ShuffleNet V2	94.48	94.47	94.48	94.45
ResNet50	94.14	94.15	94.14	94.13
DenseNet121	94.51	94.51	94.51	94.49
Proposed Work	97.62	97.68	97.67	97.12

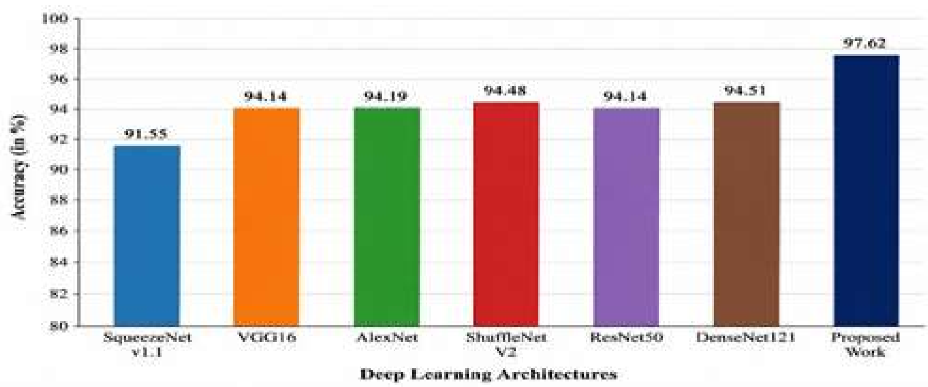

Figure 11. Comparison of classification accuracy achieved by different deep learning architectures and the proposed model.

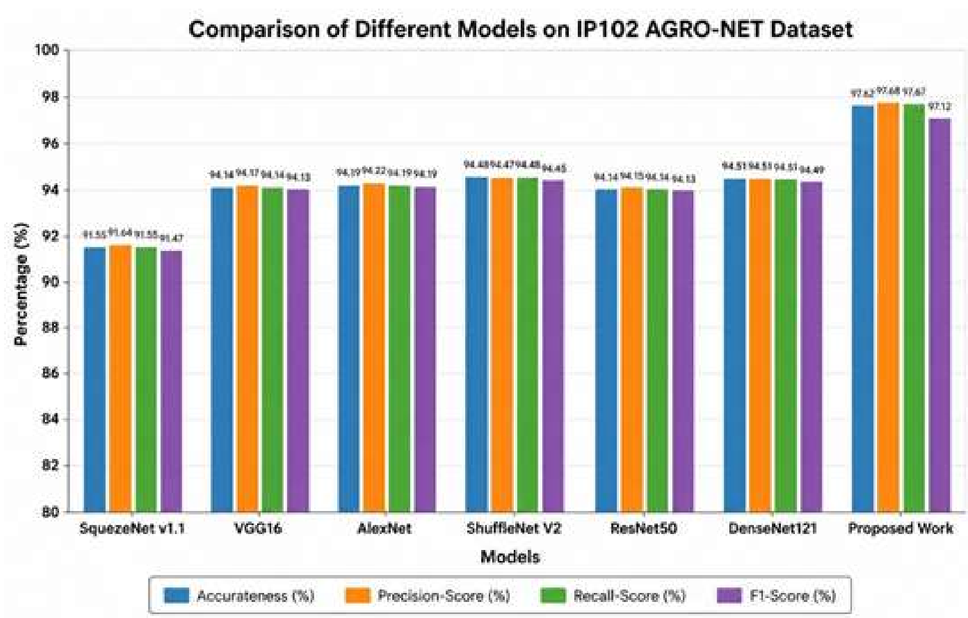

Figure 12. Comparison of Different Models on IP102 AGRO-NET Dataset

4.7 Comparative Analysis with existing work

We achieved improved results with the proposed EffiecientNet-B0 model which identifies more detail in the pest images. A small filter size of 3x3 captures more detail. The BN approach in the developed model's feature map generalizes the input of each mini-batch and creates a level of regularization which decreases the generalization error. Also, in the proposed model's classification unit, a dropout approach creates a level of regularization by removing some of the outputs of the layer before it and helps avoid overfitting and improve generalization. Figure 13 shows the benefits of the proposed pest recognition and classification approach. A comparison of the proposed approach and latest Deep Learning methods is publicized in Table 7.

Table 7. Performance comparison in recognition and cataloging of pests with advanced methods

S.No.	Author	Datasets	Methods	Accuracy
1	Loris Nanni et al. [39]	Breast Grading Carcinoma dataset	Ensemble CNN	96%
2	K. Thenmozhi et al., [40]	Xiel	Deep-CNN Model	97%
3	Huang M. L. et al. [28]	IPM Images and NBAIR datasets	RestNet50 along with Data Analytics	97.08%
4	Wang Dawei et al. [41]	Laryngeal dataset	Alexnet	94%
5	Deng et al. [42]	Pest Dataset	SVM	86%
6	Proposed EfficientNet-B0 approach	IP102 AGRO-NET	EfficientNetB0	97.62%

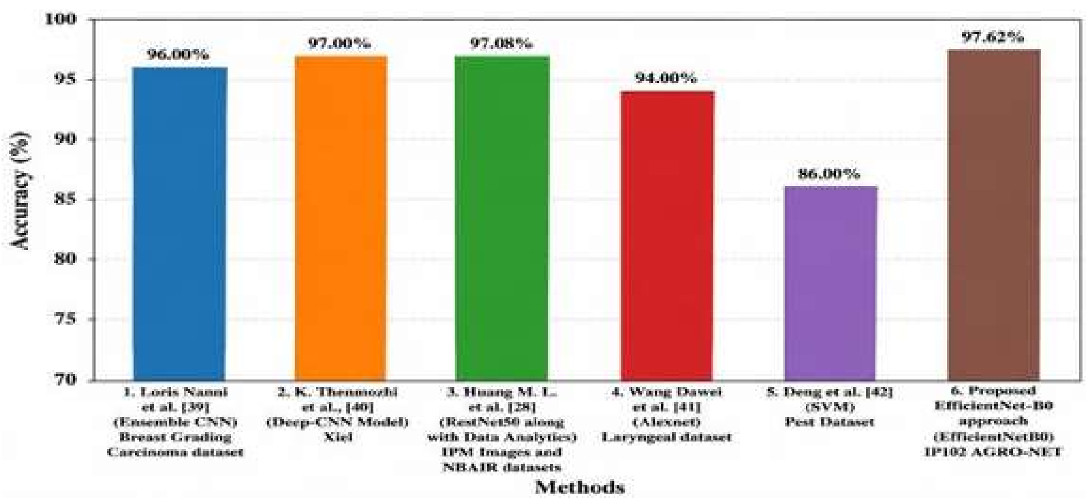

Figure 13. Comparison of Classification Accuracy of Deep Learning Models

This comparison proves how efficient and effective the introduced Agro-Net technique is in comparison with other approaches. We would like to highlight that the mentioned approaches are much more resource-consuming than the approach we offer since they implement deeper approaches that lead to overfitting. These findings imply the usefulness and advantages of the developed technology, including high computation capacity. Since Agro-Net has only eleven convolution layers, one can assume that compared to other CNNs all biases of these layers are inactive.

In the lightweight models, Agro-Net recommended in this study attains similar arrangement fulfilment to the ResNest-50 conventional model, with a 97.08% decrease in parameters along with a significant decrease in training time. When likened to Ensemble CNN, Deep-CNN Model, RestNet50 models, Alexnet and SVM model, EffiecientNet-B0 has the finest classification presentation with the least attributes and floating-point operations.

This research suggests is pretty accurate and not too hard to use to figure out what kind of pests are present. This makes it useful for rapidly and precisely finding pests, especially on portable devices and in regions where there aren't many resources. The experimental outcomes and study demonstrate that the direct transplanting of a model proficient in various activities to a specific task may not produce beneficial results. To make the model fit a certain task, it is important to improve it in many ways. Ideas like the attention mechanism and multi-scale fusion are significant.

5. CONCLUSION

Our research focusses on the categorisation of agricultural pests and reports the contests of poor precision and high intricacy inherent in current models. We built the IP102

AGRO-NET dataset just for tasks that involve putting pests into groups. We enhanced the Efficient-Net-B0 base model by making improvements to the architecture, optimisation techniques, data expansion methods, activation functions, attention machineries, and multi-scale attribute fusion. This is how Agro-Net came to be, a simple technique to tell what kind of pests are in crops. We completed a lot of testing to make sure that both the single and the combination improvements truly did make the model work better. We compared our dataset classification against those of lightweight and traditional models which calculate like accurateness, exactness, recall, and F1 score to see how well it worked. We also counted the numeral of attribute, floating-point operations, and training time to see how complicated the model was. The outcomes demonstrated that the upgraded Agro-Net was improved at categorising photos of agricultural pests and could be utilised in more ways. Agro-Net makes the EfficientNetB0 model easier to use and better at finding pests in crops. This study establishes a basis for subsequent improvements in the classification of agrarian pests. In upcoming exploration, we will apply the Agro-Net model to a resource constrained Raspberry. Worldwide, contamination causes the most agricultural and monetary losses. The ability to automatically identify exotic invasive insects would significantly speed the detection and eradication of these insects.

The proposed framework demonstrates a clear advancement over current methodologies, as reflected in the confirmed recognition and classification of pest's framework achieved 97.62% accuracy. Additionally, the framework has been proven and validated to be effective and robust for recognition and classification of pests based on its experimental results on the standard Kagle (Pest Dataset). For now, the research analyses mainly the major pests. In the classification of pests in the future, we will be interested in having a greater variety of pests that the Agro-Net would be able to identify and classify. This work will enable specialists and farmers to identify pests more easily and efficiently and will help decrease the losses to agricultural yield and the economy.

REFERENCES

[1] Ai, Yong, et al. "Research on recognition model of crop diseases and insect pests based on deep learning in harsh environments." *IEEE Access* 8 (2020): 171686-171693.

[2] Bhatt, Aarya, et al. "Deep transfer learning based light-weight agricultural pest classification model using convolutional neural networks." *2024 International Conference on Inventive Computation Technologies (ICICT)*. IEEE, 2024.

[3] Dewari, Sapna, Meenu Gupta, and Rakesh Kumar. "Agricultural insect pest's recognition system using deep learning model." *Congress on Intelligent Systems*. Singapore: Springer Nature Singapore, 2022.

[4] Barman U, Choudhury RD, Sahu D, Barman GG. Comparison of convolution neural networks for smartphone image based real time cla ssification of citrus leaf disease. Comput. Electron. Agric. 2020;177:105661.

[5] Sp, Jia, and Gao Hj. "Research Progress on Image Recognition Technology of Crop Pests and Diseases Based on Deep Learning. OL." *Transactions of the Chinese Society for Agricultural Machinery* 50 (2019): 313-317.

[6] Parraga-Alava, Jorge, et al. "LeLePhid: an image dataset for aphid detection and infestation severity on lemon leaves." *Data* 6.5 (2021): 51.

[7] Khalifa, Nour Eldeen M., M. O. H. A. M. E. D. Loey, and MOHAMED HAMED N. Taha. "Insect pests recognition based on deep transfer learning models." *J. Theor. Appl. Inf. Technol* 98.1 (2020): 60-68.

[8] Zhu, Juanhua, et al. "Identification of grape diseases using image analysis and BP neural networks." *Multimedia tools and applications* 79.21 (2020): 14539-14551.

[9] Dawei, Wang, et al. "Recognition pest by image-based transfer learning." *Journal of the Science of Food and Agriculture* 99.10 (2019): 4524-4531.

[10] Chen, Li-yong, and Xiu-ye Yin. "Image recognition of typical potato diseases and insect pests using deep learning." (2021): 9956 -9965.

[11] Dai, Min, et al. "A new pest detection method based on improved YOLOv5m." *Insects* 14.1 (2023): 54.

[12] Chen, Hongbo, et al. "Feature refinement method based on the two-stage detection framework for similar pest detection in the field." *Insects* 14.10 (2023): 819.

[13] Anwar, Zeba, and Sarfaraz Masood. "Exploring deep ensemble model for insect and pest detection from images." *Procedia Computer Science* 218 (2023): 2328-2337.

[14] Wang, Rujing, et al. "Agripest: A large-scale domain-specific benchmark dataset for practical agricultural pest detection in the wild." *Sensors* 21.5 (2021): 1601.

[15] Nanni, Loris, Gianluca Maguolo, and Fabio Pancino. "Insect pest image detection and recognition based on bio-inspired methods." *Ecological Informatics* 57 (2020): 101089.

[16] Wang, Fangyuan, et al. "Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition." *Computers and Electronics in Agriculture* 169 (2020): 105222.

[17] Turkoglu, Muammer, Berrin Yanikoğlu, and

Davut Hanbay. "PlantDiseaseNet: Convolutional neural network ensemble for plant disease and pest detection." *Signal, Image and Video Processing* 16.2 (2022): 301-309.

[18] Vamsi, Boddapati Teja, Jyothi Yadla, and Gujjarlapudi Varun Kumar. "Application of IP102 AGRO-NET Dataset in Pest Detection by using Deep Learning Techniques." *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*. IEEE, 2023.

[19] Teixeira, Ana Cláudia, et al. "A systematic review on automatic insect detection using deep learning." *Agriculture* 13.3 (2023): 713.

[20] Haider, Zeeshan Ali, et al. "Early Detection and Prediction of Pests in Field Crops Using Transfer Learning." *VFAST Transactions on Software Engineering* 12.3 (2024): 98-113.

[21] Nanni, Loris, et al. "High performing ensemble of convolutional neural networks for insect pest image detection." *Ecological Informatics* 67 (2022): 101515.

[22] Rustia, Dan Jeric Arcega, et al. "Towards intelligent and integrated pest management through an AIoT-based monitoring system." *Pest Management Science* 78.10 (2022): 4288-4302.

[23] Ali, Md Akkas, Rajesh Kumar Dhanaraj, and Anand Nayyar. "A high performance-oriented AI-enabled IoT-based pest detection system using sound analytics in large agricultural field." *Microprocessors and Microsystems* 103 (2023): 104946.

[24] Liu, Jun, and Xuewei Wang. "Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network." *Frontiers in plant science* 11 (2020): 898.

[25] Zha, Mingfeng, et al. "A lightweight YOLOv4-Based forestry pest detection method using coordinate attention and feature fusion." *Entropy* 23.12 (2021): 1587.

[26] Yin, Helin, et al. "Transfer learning-based search model for hot pepper diseases and pests." *Agriculture* 10.10 (2020): 439.

[27] Chen, Ching-Ju, et al. "An AIoT based smart agricultural system for pests detection." *IEEE access* 8 (2020): 180750-180761.

[28] Manavalan, R. "Automatic identification of diseases in grains crops through computational approaches: A review." *Computers and Electronics in Agriculture* 178 (2020): 105802.

[29] Ray, Rakesh Kumar, et al. "An interactive AI-based crop and pest management system leveraging transfer learning for enhanced sustainable agriculture." *Modeling Earth Systems and Environment* 11.4 (2025): 244.

[30] 30. Qin, Z.; Zhang, P.;Wu, F.; Li, X. Fcanet: Frequency channel attention networks. In Proceedings of the IEEE/CVF InternationalConference on Computer Vision, Hangzhou, China, 22 December 2020; pp. 783–792.

[31] Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 17–21 June 2016; pp. 2921–2929.

[32] Sanghyun, W.; Jongchan, P.; Joon-Young, L.; In, S. CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.

[33] Szegedy, C.; Liu,W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 17–21 June 2016; pp. 1–9.

[34] Wu, X., Zhan, C., Lai, Y.-K., Cheng, M.-M., and Yang, J., "IP102 AGRO-NET: A Large-Scale Benchmark Dataset for Insect Pest Recognition," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 8787–8796, 2019. doi:10.1109/CVPR.2019.00899.Pham, Q.; Liu, C.; Sahoo, D.; Steven, H. Contextual transformation networks for online continual learning. In Proceedings of the International Conference on Learning Representations, Virtual, 26 April–1 May 2020.