Enhancing agricultural pest identification using COnvolutional neural networks and transfer learning

Oluwafemi Olobashola

University of Hull

1. **Introduction**

Agriculture is the process through which land is cultivated for the purpose of plant and animal production. From these cultivated plants and animals, food, wood, and fabrics can be derived. Food is an absolute necessity for human existence, so is fabrics, and wood for construction, production of furniture, paper, etc. (Mizushina, 2021). According to the World Food Program (2023), an estimated 345million people are facing high levels of food insecurity globally. This, combined with the new realities from COVID-19 and global conflicts; especially the Russian-Ukraine war has necessitated the need for every possible means of mitigating food insecurity. Amongst the major risk factors for food insecurity is agricultural pests and diseases, which accounted for up to 40% post-harvest losses (Savary et al, 2019). These pests and diseases are usually transboundary in nature, and as such they can move rapidly from one region to the other along with the movement of food and animal products. With their introduction into new environments, they have the potential of inflicting losses. In addition to exacerbating food crises, these pests are also capable of causing diseases in humans, and sometimes lead to epidemics, examples of which are swine flu and avian influenza (FAO, 2017). The threat poised by agricultural pests therefore requires adequate attention, which in turn requires a lot of knowledge and experience. The farmer needs to be able to detect that something has gone wrong on their farm, or detect the presence of alien organisms, relay same to the plant pathologist who with knowledge and experience is able to carefully identify the causative agent and then make recommendations for remedy. This process could usually take quite some time, and when plant pests with very rapid progression are involved, it would amount to massive losses before the process can be completed. These, coupled with the agricultural sector already facing workforce shortages of up to 500,000 out of the estimated 4.1 million required (House of Commons, 2022) calls for the deployment of intelligent systems capable of detecting these disease-causing agents and pests in real time, identifying them, and making cost effective recommendations for prompt remedial actions very necessary. Convolutional Neural Network (CNN) algorithms have been used for the identification of diseases, plants, and weeds to high levels of accuracy (Kasinathat, T., Singaraju, D. and Uyyala, S., 2021).

In this project work, the focus is therefore the use of CNN algorithms to achieve high levels of accuracy in the classification and subsequent identification of agricultural pests.

1. **Background**

Kasinathat, T., Singaraju, D. and Uyyala, S. (2021), in their study: “Insect classification and detection in field crops using modern machine learning techniques” made use of two datasets containing 9 and 24 classes of insects. They proposed a CNN model and compared with Artificial Neural Networks (ANN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Naive Beyes (NB) classifiers. Their proposed CNN model had 5 convolutional layers, 3 max pooling layers, 1 flatten layer, a fully connected layer and a softmax output layer. The models were trained using a learning rate of 0.0001 and a batch size of 64. Their proposed CNN model performed best with an accuracy of 91.5% on the 9-class dataset and 90% on the 24-class dataset. In their work: “Image classification of pests with residual neural network based on transfer learning” Li, C., et al (2022) used a residual CNN model based on the following algorithms: DenseNet121, EfficientNet-B0, VGG19, ResNet-50, ResNeSt-50 and ResNeXt-50(32 x 4d), using the IP102 agricultural pest image dataset containing 75, 222 images of 102 pest species, and were able to achieve the highest accuracy of 86.95%, precision of 84.6% and recall of 85.6% with the ResNeXt-50(32 x 4d) algorithm. A learning rate of 0.0001 and CutMix data augmentation was used. So also, Kundur, N.C., and Mallikarjuns, P.B., (2022), in their work: “Insect pest image detection and classification using deep learning” used the 5, 10 and 15 class insect pest images of the IP102 dataset. The datasets had 14,490, 29,210 and 43,210 images respectively. They used two CNN algorithms: EfficientNetB4 and EfficientNetB7, images resizing, normalization and augmentation (flipping, rescaling, and zooming) were implemented on the datasets to help increase the data and avoid overfitting. The EfficientNetB7 performed best with an accuracy of 99.0%, 96.0% and 93.0% of the 5, 10 and 15 class insect pest datasets respectively. Doan, T.N. (2023), in their work: “Large-scale insect pest image classification” used 3 datasets: Xie24, D0 and IP102 containing 24, 40 and 102 insect classes respectively. They combined EfficientNet-B5 and Power mean SVM to attain higher level of sophistication and accuracy. In training their algorithm, they used a learning rate of 0.0001, images size of 600x600 was used, batch size of 32, and data augmentation was applied. The best performance recorded by their algorithm stood at 72.3% on the IP102 dataset. Xia,D., et al (2018) in their own work: “Insect detection and classification based on an improved Convolutional Neural Network” used the Xie dataset containing 24 classes of crop insect. Data augmentation was implemented by downloading additional images from the internet and applying the conventional image augmentation techniques like bipolar interpolation, rotation, as well as salt and pepper noise. The data was trained on a VGG19 model with a learning rate of 0.001 to obtain a precision of 89.22% at its best.

This project work proposes to use a computationally efficient base CNN model trained with transfer learning algorithms (VGG16, EfficientNet-B0 and MobileNetV2) on the ImageNet dataset to successfully identify and classify agricultural pests to an accuracy level of 98% or more to help mitigate the havoc being wrecked on global food security by these pests in real-time.

1. **Objectives**

With about a million named insect species globally, and about 70,000 pest and diseases that have been known to damage agricultural produce (McKirdy, S.J, Sharma, S.B, and Bayliss, K.L., 2014), classification and identification of disease-causing agricultural pests becomes a difficult task, especially for those without the knowledge and experience. To address this problem, computer vision with machine learning algorithms (Base-CNN algorithm using pretrained models) is explored to attain higher levels of accuracy, speed, and efficiency. The model is aimed to achieve an accuracy of 98% using the least number of computational resources possible to aid the ease of deployment for real-time pest identification on the field.

1. **Methodology**

**4.1 Base-CNN Model**

In this study, a Base-CNN model was built and fine-tuned to improve on its performance. The model consists of an input layer, four hidden layers, a flattening layer that converts the outputs from the hidden layer into 1D, a fully connected layer, and an output layer. Figure 1.

* The input layers use 64 filters with a 3x3 kernel size, activation function of ReLU, and input image size of 128x128 and 3 color channels. The choice of 128x128 is informed by the desire to have a computationally agile model.
* The hidden layer consists of 4 convolutional blocks with one 2D convolutional layer and one 2D maxpooling layer each, and all have ReLU applied as their activation function. Each of the convolutional layers has a kernel size of 3x3, while the maxpooling layers each has a size of 2x2. The 1st, 2nd, 3rd and 4th convolutional layers however have filter sizes of 64, 128, 256 and 512 respectively.
* The flattening layer serves to convert the output from the convolutional layers into 1D outputs to be fed into the fully connected layer.
* The output layer has 12 neurons, which stands for the number of classes in the dataset and uses softmax as activation function.

**Transfer Learning Models**

Transfer learning is a branch of Machine learning that adapts already existing algorithms that have been trained on a different dataset but for a similar task. It leverages on the knowledge it has gained from the previous task, transfers this knowledge into the current task, thereby saving the computational resources required to start the task from scratch (Hosna et al, 2022). Transfer learning achieves this by transferring the weights from previous models to the new task, significantly speeding up the training process and the performance on the new task, especially when both tasks are related, e.g. image classification. For this study, the transfer learning algorithms used were VGG16, EfficientNetB0 and MobileNetV2.

A yellow square with colorful squares

Description automatically generated with medium confidence

**Figure 1 Base-CNN Model**

**4.2 VGG-16**

VGG-16 is one of the types of the deep convolutional neural networks developed by Simonya and Zisserman in 2014. It consists of 13 convolutional layers, using 3x3 filters, the maxpooling layers of the VGG-16 architecture are also made up of 2x2 pooling layers, followed by 2 fully connected layers and a final layer as the softmax output (Thenmozhi, k., and Reddy, U.S., 2019).

* 1. **EfficientNet**

EfficientNet is a deep convolutional neural network that uses a unique scaling approach to improve its accuracy and efficiency. The scaling approach uniformly adjusts the depth, width, and resolution of the network using a compound coefficient. This coefficient is predefined and systematically enhances the network dimensions. EfficientNet was developed using a multi-objective neural architecture search that maximizes precision. EfficientNet has been shown to outperform other CNNs in terms of both accuracy and efficiency. For example, EfficientNet-B7 achieved a top-1 accuracy of 84.3% on the ImageNet dataset, while having a significantly smaller network size and faster processing time. EfficientNet hs also been shown to be effective on other datasets, such as CIFAR-100 (91.7% accuracy) and Flowers (98.8% accuracy). This study makes use of EfficientNet-B0 for the agricultural pest classification task (Doan, T.N. (2023).

A diagram of a diagram

Description automatically generated with medium confidence

**Figure 2 EfficientNet-B0 General Architecture**

* 1. **MobileNet**

MobileNet is a simple convolutional neural network (CNN) that was introduced in 2017 and designed to be efficient in terms of both size and computation. Despite its compact nature, it was still able to achieve high accuracy on image classification tasks. MobileNetV1 uses a technique called depth wise separable convolutions, which allows it to reduce the number of parameters and computations required. It also uses a set of hyper-parameters that can be tuned to optimize the model for a different task. This model has been used for a variety of tasks including augmented reality and image classification. MobileNetV2 is an improved version of MobileNetV1 and was first introduced in 2018. It addresses some of the limitations of MobileNetV1, such as information degradation in non-linear convolution layers, and uses a novel structure called inverted residuals to help preserve information integrity (Dong et al, 2020).

A diagram of a computer program

Description automatically generated

**Figure 3 MobileNetV2 Architecture**

1. **Experiment**

**5.1 Dataset**

The dataset for this classification task is the Agricultural Pests Image Dataset downloaded from Kaggle, and consists of 12 pest classes: Ants, Bees, Beetles, Caterpillars, Earthworms, Earwigs, Grasshoppers, Moths, Slugs, Snails, Wasps, and Weevils. There was a total of 5,504 images as shown in Table 1.

**Table 1 Dataset**

|  |  |
| --- | --- |
| Pest Class | Total |
| Ants | 499 |
| Bees | 500 |
| Beetle | 416 |
| Caterpillar | 434 |
| Earthworms | 323 |
| Earwig | 466 |
| Grasshopper | 485 |
| Moth | 497 |
| Slug | 391 |
| Snail | 510 |
| Wasp | 498 |
| Weevil | 485 |

**5.2 Exploratory Data Analysis and Preprocessing**

The images as contained in their corresponding folders were extracted and had their containing folder names assigned to them individually as their labels. The images and labels were subsequently organized into numpy arrays. Because cv2 reads images in the BlueGreenRed (BGR), the images were converted into and saved into RedGreenBlue (RGB). The classes value count visualization was done to check for balance amongst the data classes and they were found to be relatively balanced (Figure 4).

Twelve random images were selected for visualization to get a feel of some of the classes in the dataset (Figure 5).

The dataset was subsequently split into training set (80%), validation set (10%) and test set (10%).

A graph of different colored bars

Description automatically generated

**Figure 4 Data Classes Value Counts**

A collage of insects

Description automatically generated

**Figure 5 Sample Crop Pests**

**5.3 Data Normalization**

This was achieved by converting the pixel values of the images I to floating points between 0 and 1 helping the model to generalize well and avoid overfitting, optimize the gradient descent, and mitigate the effects of images having wide pixel ranges (Sun et al, 20212).

* 1. **Experimental Setup**

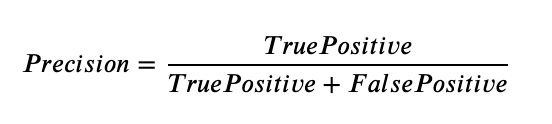
Six CNN algorithms were trained on the dataset; 3 were Base-CNN models, while 3 were models that had been pretrained on ImageNet dataset.

To evaluate the performance of our classification algorithms, a few metrices are calculated. These metrices include.

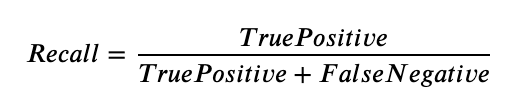
**5.4.1 Accuracy:** This gives the amount of correct image classifications of all the classification carried out by the model.



**5.4.2 Precision:** This measures how well the model classifies the classes by calculating the fraction of correctly classified pests’ class amongst all the images classified as being in that class.



**5.4.3 Recall:** this is an indication of how effectively the model identifies the images in each class. This is achieved by taking a count of the accurate positive pest predictions for a particular class of pests out of all the possible pests in that class.



**5.4.4 Log Loss:** This is another evaluation method especially useful in checking the performance of multi-class classification tasks. It assigns a probability to each predicted pest class and gets penalized for incorrect classification.

For example, if there are N samples in M classes, the Log Loss is derived as follows:

A number and mathematical equation

Description automatically generated with medium confidence

(Mishra, A., 2018).

* 1. **Base-CNN model**

As earlier stated, the Base-CNN model consists of 4 convolutional blocks which are very crucial in the extraction of important features from the inputs. Non-linearity is introduced into these inputs, and ReLU serves as the activation function. 512 neurons were applied to the dense layer of the architecture, with a dropout of 0.5, helping to prevent overfitting. A batch size of 32 was also applied to the algorithm. The input image size was set at 128x128 due to computational resource considerations.

* 1. **Base-CNN 2**

This had padding introduced to maintain the dimension of the input in the output of the convolutional layers. Batch normalization was also introduced to help improve the stability of the model.

* 1. **Base-CNN 3**

For this architecture, data augmentation was used to increase the volume of the data. Rotation, height shift, width shift, horizontal flip, vertical flip, and shear were the parameters used in augmenting the data to generate more images. The learning rate was also varied from the default 0.001 to 0.0001 and the training batch size increased to 64.

* 1. **VGG-16 Transfer Learning Model**

This model has previously been trained on the ImageNet dataset and performed well in extracting key features in that data set. The feature extraction capabilities as contained in the convolutional and maxpooling layers are therefore adopted for the agricultural pest image dataset, using the fully connected layer of the Base-CNN model above, maintaining the learning rate of 0.0001, batch size of 64 and a dropout of 0.5.

* 1. **EfficientNetB0 Transfer Learning Model**

This model was also previously trained on the efficient net dataset and is applied just like the VGG-16 model. GlobalAveragePooling 2D was introduced to compute the average value of the feature map extracted by the pretrained model, replacing the fully connected layer, and further saving computational time and reducing the number of parameters. This overall helps to maintain important feature information (Akhtar, N., and Ragavendran, U., 2020).

* 1. **MobileNetV2 Transfer Learning Model**

This model is also applied using GlobalAveragePooling 2D to replace the fully connected layer, learning rate of 0.0001, dropout of 0.5, batch size of 64 and data augmentation were all maintained.

1. **Results**

Table 2 shows the classification results of the 6 models used in the identification and classification of pests in the agricultural pest image dataset used for this project. The accuracies achieved on the training, validation and testing datasets are indicated.

**Table 2 Performance of proposed agricultural pest classification models**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Training Accuracy | Validation Accuracy | Testing Accuracy |
| Base-CNN Model | 97.4% | 92.7% | 92.4% |
| Base-CNN Model 2 | 91.8% | 91.7% | 91.8% |
| Base-CNN Model 3 | 93% | 93% | 93% |
| VGG16 Model | 94.5% | 94.8% | 95% |
| EfficientNetB0 Model | 97.6% | 90.3% | 90.7 % |
| MobileNetV2 | 96.4% | 97.5% | 96.9% |

* The first base model showed good performance only on the training data set, having an accuracy of 97.4%, The loss function however remained high at 2.26 while its accuracy on the testing data came to 92.4%.
* Introduction of padding and batch normalization into the base model for the Base-CNN model 2 helped the model to generalize better, but it did not achieve a good enough accuracy or loss function at 91.7% and 2.43 respectively.
* Base-CNN model 3 with data augmentation, increased batch size and new learning rate of 0.0001 gave an improved and generalized performance on all data splits at accuracy level of 93% on all, and a loss function of 1.72, which is an improvement of the first two models. All three initial base models had computational times of between 245 ms/epoch to 297 ms/epoch.
* The VGG16 model showed an improvement with an accuracy of 95% but had a low recall of 49%. Predictions with this model showed a misclassification is 3 out of the selected 12 pests, where a beetle was labelled as a snail, slug as beetle, and slug as a grasshopper. This model however had the highest computational time of 630 ms/epoch.
* The EfficientNet surprisingly performed poorly on the validation and test splits when compared to the training. This showed overfitting and was made evident in its wrong classification of 8 out of the 12 pest images selected for identification.
* MobileNetV2 showed the best performance on all metrices, and on all dataset splits. It had an accuracy of 96.9% on the test split, a loss of 0.69, a recall of 85.9% and the best computational time of 243 ms/epoch. The MobileNetV2 model identified 11 out of the selected images correctly.

Table 3 shows the performance of all six proposed models using the selected metrices. During the training of all models, the validation and training accuracies and loss were monitored, and the MobileNetV2 model showed the best performance this is shown of figures 6 – 11, and this also performed better than the compared works, except for the EfficientNetB7 performance on the 5-class dataset as shown in Table 4.

**Table 3 Model performance metrices on testing data splits**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | Loss | Accuracy | Precision | Recall | AUC | Time/Epoch |
| Base-CNN Model | 2.25 | 92.7% | 58.1% | 44.3% | 0.85 | 460ms |
| Base-CNN Model 2 | 2.33 | 91.7% | 100% | 0.4% | 0.66 | 549ms |
| Base-CNN Model 3 | 1.63 | 93.0% | 67.1% | 31.8% | 0.87 | 361ms |
| VGG16 Model | 1.15 | 94.5% | 84.0% | 48.5% | 0.93 | 2000ms |
| EfficientNetB0 Model | 2.62 | 90.7% | 36.6% | 16.3% | 0.71 | 758ms |
| MobileNetV2 Model | 0.69 | 96.9% | 85.9% | 75.1% | 0.98 | 29 |

**Table 4 Results from related works**

A green and white table with black text

Description automatically generated

A graph of a line

Description automatically generated with medium confidence

**Figure 6. Base-CNN Model Accuracy and Loss Performance**

A graph of a training and training accuracy

Description automatically generated with medium confidence

**Figure 7. Base-CNN Model2 Accuracy and Loss Performance**

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

**Figure 8. Base-CNN Model3 Accuracy and Loss Performance**

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

**Figure 9. VGG16 Model Accuracy and Loss Performance**

A graph of training and training accuracy

Description automatically generated

**Figure 10. EfficientNetB0 Model Accuracy and Loss Performance**

A graph of a training and training accuracy

Description automatically generated with medium confidence

**Figure 11. MobileNetV2 Mode2l Accuracy and Loss Performance**

A collage of insects

Description automatically generated

**Figure 12. VGG16 Predictions**

A collage of insects

Description automatically generated

**Figure 13. MobileNetV2 Prediction**

1. **Conclusion**

In this study, the challenging task of classifying and identifying disease-causing agricultural pests using CNN and transfer learning was addressed. The aim was to achieve accuracy of up to of 98% while minimizing computational resources to make its real-time field deployment more realistic, and this has yielded valuable insights.

Among the models tested, the VGG16 model achieved 95% accuracy on testing data, though its recall rate was comparatively lower. The EfficientNetB0 model exhibited surprising overfitting issues, incorrectly classifying 8 out of 12 pest images during testing. MobileNetV2 model however stood out as the top performer, boasting a remarkable 96.9% accuracy on the test split, 85.9% recall, and a rapid computational time of 243 ms/epoch. In future works further data augmentation techniques could be applied to get better results from the agricultural pest dataset. Such techniques as; brightness and contrast adjustment, colour jittering, random cropping, gaussian noise, and scaling/zooming Takahashi, R., Matsubara, T. and Uehara, K. (2019).

The study therefore highlights the potential of employing MobileNetV2 for the accurate and efficient identification of agricultural pests, thus aiding timely intervention and resource optimization, which would potentially lead to higher agricultural yields and incomes as well as improved global food security. It makes ready the development of a user-friendly application that integrates the MobileNetV2 model could facilitate on-field pest identification, aiding farmers in making informed decisions promptly.

**Reference:**

Agrawal, S., 2021. Metrics to Evaluate your Classification Model to take the right decisions. Retrieved from https://www.analyticsvidhya.com/blog/2021/07/metrics-to-evaluate-your-classification-model-to-take-the-right-decisions/

Alhichri, H., Alswayed, A.S., Bazi, Y., Ammour, N. and Alajlan, N.A., 2021. Classification of remote sensing images using EfficientNet-B3 CNN model with attention. *IEEE access*, *9*, pp.14078-14094.

Akhtar, N. and Ragavendran, U., 2020. Interpretation of intelligence in CNN-pooling processes: a methodological survey. *Neural computing and applications*, *32*(3), pp.879-898.

Doan, T.N., 2023. Large-Scale Insect Pest Image Classification. *Journal of Advances in Information Technology*, *14*(2).

Dong, K., Zhou, C., Ruan, Y. and Li, Y., 2020. MobileNetV2 model for image classification. In *2020 2nd International Conference on Information Technology and Computer Application (ITCA)* (pp. 476-480). IEEE.

FAO. 2017. *The future of food and agriculture – Trends and challenges.* Rome.

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z. and Azim, M.A., 2022. Transfer learning: a friendly introduction. *Journal of Big Data*, *9*(1), p.102.

House of Commons, 2022. Labour shortages in the food and farming sector. Environment, Food and Rural Affairs Committee Fourth Report of Session 2021 – 2022.

Kasinathan, T., Singaraju, D. and Uyyala, S.R., 2021. Insect classification and detection in field crops using modern machine learning techniques. *Information Processing in Agriculture*, *8*(3), pp.446-457.

Kundur, N. and Mallikarjuna, P., 2022. Insect pest image detection and classification using deep learning. *Int. J. Adv. Comput. Scie. Appl. https://doi. org/10.14569/IJACSA*.

Li, C., Zhen, T. and Li, Z., 2022. Image classification of pests with residual neural network based on transfer learning. *Applied Sciences*, *12*(9), p.4356.

McKirdy, S.J., Sharma, S.B. and Bayliss, K.L., 2014. Quarantine and biosecurity. In *Encyclopedia of Agriculture and Food Systems* (pp. 11-20). Academic Press.

Mishra, A., 2018. Metrics to evaluate your machine learning algorithm. *Towards data science*, pp.1-8.

Mizushina, Y., 2021. Importance of Agriculture and Research. *Open Access Journal of Science and Technology*, *9*(2).

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N. and Nelson, A., 2019. The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, *3*(3), pp.430-439.

Stork, N.E., 2018. How many species of insects and other terrestrial arthropods are there on Earth?. *Annual review of entomology*, *63*, pp.31-45.

Sun, J., Cao, X., Liang, H., Huang, W., Chen, Z. and Li, Z., 2020, April. New interpretations of normalization methods in deep learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 04, pp. 5875-5882).

Takahashi, R., Matsubara, T. and Uehara, K., 2019. Data augmentation using random image cropping and patching for deep CNNs. *IEEE Transactions on Circuits and Systems for Video Technology*, *30*(9), pp.2917-2931.

Thenmozhi, K. and Reddy, U.S., 2019. Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, *164*, p.104906.

Torrey, L. and Shavlik, J., 2010. Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI global.

Xia, D., Chen, P., Wang, B., Zhang, J. and Xie, C., 2018. Insect detection and classification based on an improved convolutional neural network. *Sensors*, *18*(12), p.4169.