International Journal of Agricultural and Environmental Information Systems
Volume 11 • Issue 4 • October-December 2020

Plant-Seedling Classification Using Transfer Learning-Based Deep Convolutional Neural Networks

Keshav Gupta, Deloitte Consulting India Pvt. Ltd., Hyderabad, India

Rajneesh Rani, Department of Computer Science and Engineering, NIT Jalandhar Punjab, India

https://orcid.org/0000-0003-2104-227X

Nimratveer Kaur Bahia, Department of Computer Science and Engineering, NIT Jalandhar Punjab, India

ABSTRACT

The ever-growing population of this world needs more food production every year. The loss caused in crops due to weeds is a major issue for the upcoming years. This issue has attracted the attention of many researchers working in the field of agriculture. There have been many attempts to solve the problem by using image classification techniques. These techniques are attracting researchers because they can prevent the use of herbicides in the fields for controlling weed invasion, reducing the amount of time required for weed control methods. This article presents use of images and deep learning-based approach for classifying weeds and crops into their respective classes. In this paper, five pre-trained convolution neural networks (CNN), namely ResNet50, VGG16, VGG19, Xception, and MobileNetV2, have been used to classify weed and crop into their respective classes. The experiments have been done on V2 plant seedling classification dataset. Amongst these five models, ResNet50 gave the best results with 95.23% testing accuracy.

KEYWORDS

MobileNetV2, Plant SeedlingResNet50, VGG16, VGG19, Xception

INTRODUCTION

Agriculture is the backbone of Indian economy. It is a source of livelihood for more than half of the population of India. Along with its allied sectors, agriculture accounted for 23% of Indian GDP, and employed 59% of the total workforce in India(1). The industrial sector also depends on agriculture for raw products. Food production must be increased by 70% to feed the world population by 2050(3). The population of the world is likely to increase by two billion to three billion by 2050 which will double the demand of food, according to several studies. Even today when there is plenty of food, the world population almost one billion people suffer from chronic hunger. . The loss in crop yield has to be taken into consideration for improving the return on investment for people depending on agriculture as well as for strengthening the economy of the country. The yield losses are caused by three major factors weed invasion, pests and pathogens, from which weeds are a major issue. Weeds are the unwanted plants that grow along with the crops e.g Chickweed, Scentless Mayweed etc. Weed control is the process of reducing or completely eliminating the loss in crop yield caused due to weed

DOI: 10.4018/IJAEIS.2020100102

invasion. The broad categorization of weed control methods is done in three of the classes that are Cultural weed control, Mechanical weed control, and Chemical weed control.

One of the easiest ways to control weeds is through prevention or cultural control. Close planting in the garden can reduce weed growth by eliminating open space. Cover crops are good for this as well. Cover crops are the crops grown in the space left between the rows of in crop fields. These crops are used to eliminate the space to prevent weed growth. Adding mulch will prevent light from getting to weed seeds and prevents growth. The problem with this method is that it requires high knowledge of spacing of crops and the type of cover crop that can be used. Chemical control is done by using herbicides which kill certain targets while leaving the desired crop relatively unharmed. Some of these act by interfering with the growth of the weed and are often based on plant hormones. Herbicides can have several side effects on biotic and abiotic environment and bear a risk to harm human health. Therefore a reduction in the amount of herbicides used in the modern agriculture is a relevant step towards sustainable agriculture. Mechanical weeding means machines equipped with classifiers, classifying the plant as crop or weed. The classification task is related to computer vision which in turn is based on Artificial Intelligence.

The main objective of this article is to classify an image of the plant as crop or weed using deep learning and to evaluate the performance of state of the art models on a benchmark dataset(V2 Plant Seedling dataset) using various performance parameters.

The rest of the article is planned as follows: Section 2 explains the background as a literature survey of the proposed work. Section 3 describes the experimental study design and methodology used in the work. Section 4 presents the experimental results and analysis. Section 5 concludes the work with some possible future directions.

BACKGROUND

In recent years, various projects have dealt with automated recognition of weeds using cameras with the aim of developing new farming machinery that can control the weeds more intelligently. It is then of major technical and economic importance to implement computer based methods for reliable and fast identification and classification of weed/crop for empowering machinery that can control weed invasion. The first step in both of these is to identify weeds using their images. This is accomplished using computer vision. Automatic systems can be based on images of weed and crop plants from which classification features associated to size, shape, texture and colour can be readily obtained. For this task numerous image processing algorithms are available which complement with classification methods to make the field of machine vision suitable for weed identification. In this section, the literature related to this area is presented. Deep learning is the current favorite choice of researches working around the field of image processing. It can provide great deal of help in the problem of classifying images of plants as weed or crop. Application of deep learning in agriculture-related literature is presented to observe how deep learning has helped to resolve problems in agriculture.

A method based on color information for discriminating between vegetation and background and shape analysis to distinguish between crop and weed is proposed by (Perez, Lopez, Benlloch, & Christensen, 2000). They have compared the algorithm results with human classification. They have claimed that their study shows that despite the difficulties in correctly determining the number of seedlings; it is feasible to use image classification techniques to estimate the relative leaf area of weed while moving across the field and use this data in a stratified manual survey of the field. They have validated the proposed method using 32 color images taken in natural lighting conditions. Color analysis technique was applied to the dataset and thus 1950 objects were obtained. Among these 1950 objects, 1433 were crop pieces and 517 were weeds. They have reported an accuracy of 89.7% for crop and 74.5% for weed using Bayes rule and an accuracy of 89.0% for crop and 79.2% for weed using k-NN(k=5).

In 2003, (Aitkenhead, Dalgetty, Mullins, McDonald, & Strachan, 2003) evaluated the performance of morphological characteristic measurements and neural networks for discriminating carrot seedlings from that of fat hen and ryegrass using digital imaging. In the first method, morphological characteristic measurements were used to discriminate the plants exploiting the variations depending on the plant size. The method was to use a neural network for the task. The first method had varying effectiveness of 52-74% while the neural network achieved an accuracy exceeding 75% on their own dataset collected in a commercial carrot field.

In the year 2012, (Ahmed, Al-Mamun, Bari, Hossain, & Kwan, 2012) evaluated the performance of support vector machine (SVM) for the classification of weed and crop using images. They conducted experiments to test the combinations of 14 features to find the optimal combination of features to classify the crop and weed plants with a high classification rate. They have achieved an accuracy of 97% on 224 test images. The images used in this study are taken in a chili field. They have included five weed species found in the chili field in Bangladesh. The data was pre-processed before experimentation using binarization and morphological opening.

The performance of support vector machine (SVM) for the classification of weed and crop using images is evaluated in the year 2012 by (Ahmed et al., 2012) . They conducted experiments to test the combinations of 14 features to find the optimal combination of features to classify the crop and weed plants with high classification rate. They have attained the maximum accuracy of 97% on 224 test images. The images used in this study are taken in a chili field. Five weed species were included which is found in the chili field in Bangladesh. The data was pre-processed before experimentation using binarization and morphological opening.

Random Forest-based approach has been proposed to classify weed and crop without segmentation. (Haug, Michaels, Biber, & Ostermann, 2014) have claimed that their proposed system can discriminate crop and weed growing close to each other, handling the overlap between the two of them. The essence of their proposed model is estimating crop or weed sparse pixel positions using Random Forest classifier and then spatially smoothing these pixels using Markov Random Field and inferring continuous crop or weed regions through interpolation. They have tested the system on their own dataset collected in an organic carrot farm. The proposed system achieves an accuracy of 91.5% before smoothing is applied which increases to an accuracy of 93.8% after smoothing is applied.

A variation of PCANet (Principal Component Analyses Network) for the classification of weed and crop is proposed by Xinshao in 2015 (Xinshao & Cheng, 2015) . The features extracted by the PCANet and LMC(Large Margin Classifier) is used to construct linear classifiers. Use of LMC classifier is different from SVM. The dataset is preprocessed before applying the proposed method to it by applying image filter extraction by using PCA filter banks, binarization and histogram counting. The dataset used contains 3,980 images of 91 types of weed seeds. They have achieved an average accuracy of 90.96%.

In the year 2016 Dyrmann (Dyrmann, Karstoft, & Midtiby, 2016) proposed a convolution neural network, build from scratch for plant and weed classification. They combined six datasets for their experiments to make the dataset of 22 species and 10,413 images of crop and weed. Image mirroring and rotation based data augmentation were applied before training the proposed model on the dataset. The experiments were performed using the Theano framework. They have reported a classification accuracy of 86.2% for the 22 species of plants.

A system based on a combination of vegetation detection and deep learning to classify vegetation in a field into useful vegetable and weed is proposed by (A. Milioto, Lottes, & Stachniss, 2017). They have evaluated the system on their own datasets designated as dataset A and dataset B. They have reported an accuracy of 97.3% on test set A and 89.2% on test set B when the model is trained on train set A. They have given another result showing the accuracy of the model on test set B before and after retraining the model on a small subset of train set B as 89.2% before retraining and 96.1% after retraining on train set B.

To recognize the sugar beet plants and weeds in the field based solely on image data is addressed by (A. Milioto et al., 2017). They execute a vegetation detection to eliminate irrelevant soil information from the images using multispectral images containing RGB and near infra-red information and then performed blob segmentation to extract patches containing singular crops or weeds. Classification of each patch with a convolutional neural network is done which is trained in an early growth stage to avoid overlapping crops and weeds in the data. This approach will accurately identifying the weeds on the field by taking images from different sugar beet fields.

An InceptionV3 plus lightweight deep convolutional neural network plus a set of K lightweight models based mix-model for the classification of carrot and weed plants is proposed by (McCool, Perez, & Upcroft, 2017). The approach consists of three stages (i) adopt a pre-trained model(InceptionV3) to the task at hand,(ii) apply model compression techniques to learn DCNN that is less accurate but has two orders of magnitude fewer parameters and (iii) combine K lightweight models as a mixture model to further enhance the performance of the lightweight model. They have used the Crop/Weed Field Image dataset consisting of 60 images. They achieved an accuracy of 93.9% by the complicated pre-trained model but this system can process only 0.12 frames per second. To make the approach fast lightweight deep CNNs are learned, which when combined give an accuracy of greater than 90% and processing between 1.07 to 1.83 frames per second.

The performance of traditional algorithms, Support Vector Machine(SVM) and K- Nearest Neighbor (KNN) with background segmentation, and a Convolutional Neural Network(CNN) on the plant seedling dataset is compared which is proposed by (Nkemelu, Omeiza, & Lubalo, 2018). The architecture of their CNN consisted of six convolution layers, each followed by a rectified linear unit (ReLU). The first two convolution layers had 64 filters, the next had 128 and the last one 256. Each convolution layer had zero padding. A max-pooling layer was embedded after each convolution layer. The successors of convolution layers were three fully connected layers with the last one of them having the softmax activation function, giving the probability of an input image belonging to each of the 12 classes. The accuracy achieved by KNN and SVM was 56.84% and 61.47% respectively. The maximum accuracy achieved by them is 92.6% on the dataset with CNN.

Fully convolution neural networks with encoder-decoder structure incorporated with spatial information by considering image sequence is used by (Lottes, Behley, Milioto, & Stachniss, 2018). They have claimed that their system generalizes well to previously unseen fields under varying environmental conditions. The basic idea of their work is to exploit geometric patterns that result from the fact that some crops are sowed in rows. The comparisons to other state-of-the-art approaches are presented and show that system substantially improves the accuracy of crop-weed classification without requiring a retraining of the model.

Problem of CNN-based semantic segmentation of crop fields separating sugar beet plants, weeds, and background exclusively based on RGB data is addressed by (Lottes et al., 2018; Andres Milioto, Lottes, & Stachniss, 2018). A deep encoder-decoder CNN for semantic segmentation is maintained with a 14-channel image storing vegetation indexes and has been used to solve crop-weed classification tasks. This approach can accurately perform pixel wise semantic segmentation of crops, weeds, and soil, properly dealing with heavy plant flap in all growth phases.

From the study of literature, it has been concluded that a plenty of research is done in the field of agriculture. But there is requirement to explore more state-of-art feature extraction and classification techniques to classify weed and crop. As majority of the methods used in literature is done by machine learning techniques which are applied to various datasets available for weed and crop classification with a lot of variations. These have achieved high accuracy in this field. Deep learning techniques have not been applied as much as machine learning techniques. The literature review shows that CNN used on the Plant seedling dataset gives better results than the machine learning techniques. But, CNN used was a shallow network with only 6 convolutional layers. We intend to fill this gap and apply some deep convolutional neural networks on the dataset using transfer learning. The deep learning application in agriculture-related literature shows that transfer learning has given good results in the problems of classification from images in agriculture. Applying these techniques may help in improving the accuracy achieved so far.

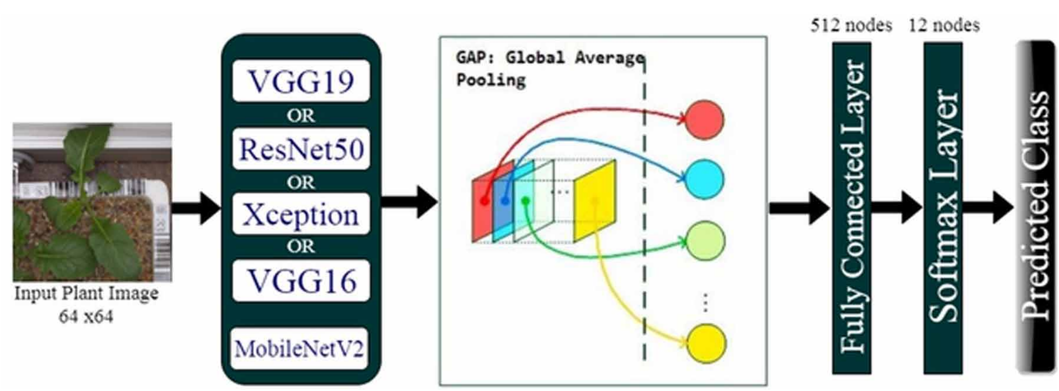
METHODOLOGY

This section presents the architecture that is proposing for the problem of weed and crop classification. Figure.1 shows the CNN models that we have used along with the layers that are substituted for the last layer of VGG19. As shown in Figure, it has used VGG19 or ResNet50 or Xception or VGG16 or MobileNetV2 as the starting architecture which takes an image of the plant as an input. In this study, there is the removal of the final layer of the architecture and replaced it with a Global Average Pooling (GAP) layer followed by a fully connected (FC) layer with 512 nodes and finally a softmax layer with 12 nodes, which gives one of the 12 classes in the dataset as output. We settled on 512 nodes in the FC layer after experimenting with the data with various numbers of nodes ranging from 264-4086 nodes for getting the best results.

Global Average pooling (GAP) layer is being used by experts to minimize over-fitting by reducing the total number of parameters. It is a variant of its counterpart Max Pooling layer. The pooling operation doesn't just reduce the dimensionality of the input but it also makes the input robust against rotation and orientation. The max-pooling layer reduces the dimensionality of the given input by taking the max value in the max-pooling filter used. But after experimentation, we have seen that global average pooling gave better results than max-pooling layer.

The last step of CNNs is to combine high-level features, which come from convolutional and pooling layers in order to classify images or, more in general, learn non-linear relations among features. The layer that performs this combination is called a fully-connected layer. As the name suggested, all neurons of that layer are connected to all neurons of the previous one to collect all information and send them to another fully connected layer or produce output values. The most common use of the FC layer is an image classification task, where the number of neurons corresponds to a number of classes that we want to classify. The fully connected layer in the proposed neural network consists of 512 nodes and ReLU as its activation function. The soft-max layer is just a fully connected layer with a soft-max activation function. The Soft-max activation function is mostly used in the output layer of problems consisting of more than two classes of output. The number of nodes in the soft-max layer corresponds to the number of classes in the dataset. In the proposed model the soft-max layer has 12 nodes each for each of the 12 species of plants. We have used Stochastic gradient descent(SGD) optimizer with a momentum of 0.9 and a learning rate of e-3and with a batch size of 16.After,

Figure 1. Proposed architecture



experimenting with the tuneable parameters once we reached an accuracy which didn't improve any further we finalized those parameters and configurations.

EXPERIMENTAL SETUP

The experiments have been performed using the Anaconda 2019.03 for windows with Python 3.7. The open-source Anaconda Distribution is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 11 million users worldwide, it is the industry standard for developing, testing, and training on a single machine. We have used the Spyder integrated development environment for running the algorithm. The whole algorithm for our experiment was implemented using Python 3.7. Tensor flow was used at the backend as the machine learning library with Keras as the wrapper library.

We have divided this section into three parts which include dataset discussion, data augmentation and performance metrics used to evaluate the results.

Dataset

The dataset, which is made publically available by Aarhus University (Giselsson, Jørgensen, Jensen, Dyrmann, & Midtiby, 2017) has been used in proposed work. The dataset that used is collected by the Aarhus University Signal Processing group in collaboration with the University of Southern Denmark. In this study, the second version of the dataset is used. The first version contained 4275 images of approximately 960 unique plants belonging to 12 species. In the second version number of images has been increased to 5539.These 12 species are common plant species found in Danish agriculture. The images in the dataset were recorded multiple times over a 20 day period at an interval of 2 to 3 days, starting a few days after emergence. The dataset is primarily targeted to research that tries to identify plant species at a dear growth stage, so as to allow farmers to conduct weeding before weed start competing with the crop for nutrients. This is why it is known as Plant Seedling Dataset. Table 1 shows the distribution of training and testing images in each of the 12 classes. It shows that Loose Silky-bent is the largest class in the dataset with 762 images and Common Wheat is the smallest class in the dataset with 253 images.

Each class contains colored images that show plants at different growth stages. The images are in various sizes and are in png format. The 12 classes in the dataset are shown in Figure 2 with an example image of each class. In experiments, this dataset is divided into 80:20 ratios with 80 percent data(4431 images) used for training the network and 20 percent of data(1108 images) used for testing of the network and image augmentation is used before training the network.

Data Augmentation

Data augmentation is the process of artificially creating new data for training and testing of a machine learning model. This helps in improving the performing of the model because in supervised learning the more your model know the better it can perform. It also helps to prevent overfitting on the training set. In this study, data augmentation is done for improving the performance of the model. Figure 3 shows an example of how data augmentation can help in improving the performance of a machine learning model.

In this images are classified into their respective classes and although there is large dataset containing 5,539 images but it is not as large as required to train the deep convolutional neural networks that we are using. Image augmentations are used in different modes: Widthshift range, Height shift range, Rotation range, Rotation range, Zoom range, Horizontal flip, Fill mode, Shear range. The value of the parameters used for image augmentation is set as shown below:

- width_shift_range = 0.2

Table 1. Distribution of images in V2 Plant Seedling Dataset

Class	Species	Training Images	Testing Images
1	Sugar Beet	370	93
2	Black grass	247	62
3	Charlock	361	91
4	Cleavers	268	67
5	Common Chickweed	572	143
6	Common Wheat	203	50
7	Fat Hen	430	108
8	Loosy Silky-bent	609	153
9	Maize	206	51
10	Scentless Mayweed	485	122
11	Shepherd's purse	219	55
12	Small-flowered cranesbill	461	115

- height_shift_range = 0.2
- rotation_range = 20
- zoom_range = 0.2
- horizontal_flip = True
- fill_mode = "nearest"
- shear_range = 0.2

In Figure 3 the images of a plant with (a) being the original image and (b) and (c) being the horizontally and vertically flipped versions of the original image.

Performance Metrics

The results for each of the 5 pre-trained based models are shown separately. For each of the model, we have presented the following (i) accuracy plot for training and testing, (ii) classification report,(iii) confusion matrix and (iv) the accuracy of the model on V2 Plant Seedling dataset for classification of the input image into its respective class.

Accuracy Plot

The accuracy plot is the curve that shows how the accuracy (training and testing) of the model is changing with each epoch. The X-axis of the plot represents the epochs(1-12) and the Y-axis represents the accuracy values(0-100) in percentage. The training and testing accuracy for each of the models is presented in the accuracy plot.

Classification Report

The classification report displays the precision, recall, F1 scores for the model. The parameters in this report are useful for comparing the performance of models on a dataset. The report shows the value of these parameters for each of the classes. This helps to analyze the performance of the model in each of the classes. The parameters in the classification report are defined as:

Figure 2. The 12 classes in the dataset with an example image

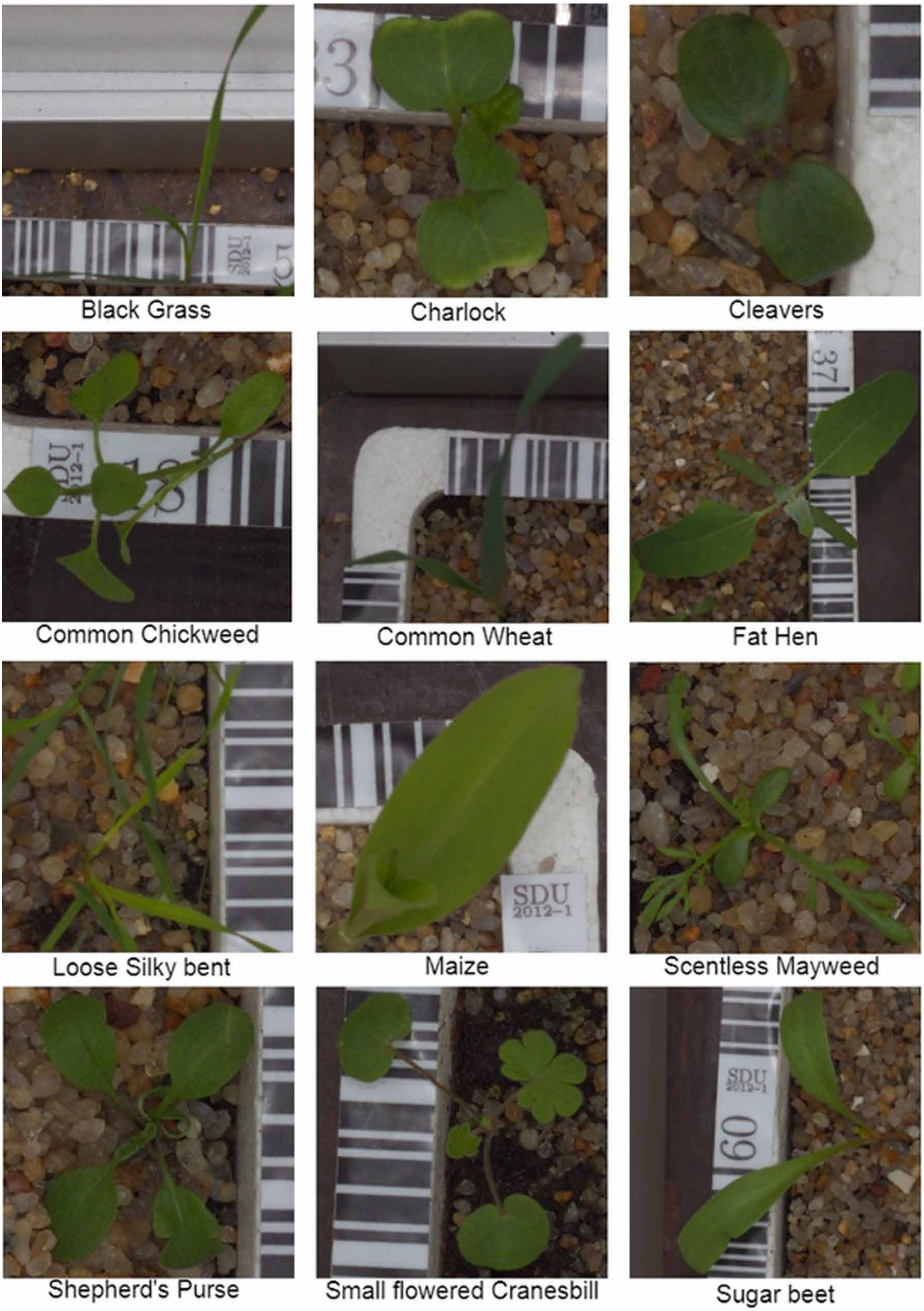
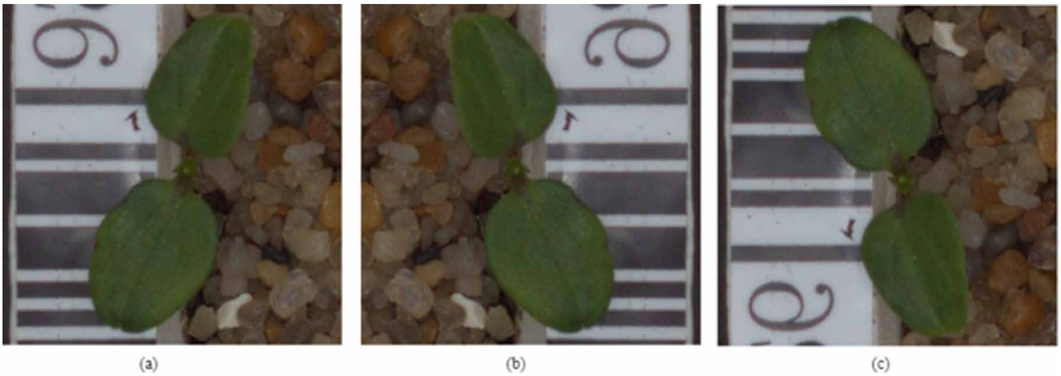


Figure 3. (a)the original image,(b)horizontally flipped, and(c)vertical flipped



Precision

Precision is the ability of a classifier not to label an instance positive that is actually negative. In simple words, precision is the accuracy of positive predictions. It is given by the following equation.

$$Precision = \left(\frac{TP}{TP + FP}\right)$$

In the above equation, TP stands for true positive and FP stands for false negative. In the context of the problem true positive is when a weed image is classified in the correct weed class and a crop image is classified in the correct crop class. Similarly, false positive is when a crop image is classified as a weed image and a weed image is classified as a crop image.

Recall

A recall is the ability of a classifier to find all positive instances. In other words, it is the fraction of correctly identified positive entries.

$$Recall = \left(\frac{TP}{TP + FN}\right)$$

In the above equation, FN stands for false negative. In the context of problem false negative is when a weed/crop image belonging to a class is rejected as belonging to that class.

F-1 Score

The F1 score is a weighted harmonic mean of precision and recalls such that the best score is 1.0 and the worst is 0.0. Generally speaking, F1 scores are lower than accuracy measures as they embed precision and recall into their computation. As a rule of thumb, the weighted average of F1 should be used to compare classifier models. It is given by the following equation.

$$F1Score = 2 * \left(\frac{P * R}{P + R}\right)$$

In the above equation, P stands for precision and R stands for recall.

Confusion Matrix

The confusion matrix is a table where the rows and columns represent the classes in the dataset, giving the fraction of misclassification. The ordering of the classes in rows and columns is similar. The rows represent the true labels and the columns correspond to the predicted label. This way the diagonal elements of the confusion matrix give the number of images that are correctly classified into their respective classes and the rest of the elements give the number of images that are misclassified in the test set. The non-diagonal entries can be analyzed to see which classes are mostly misclassified or confused with each other.

Accuracy

Accuracy is a fraction of correctly classified images. It is given by the following equation.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

In the context of the problem, accuracy specifies the percentage of plant images correctly classified into their respective class.

RESULTS AND DISCUSSIONS

In this section, we have presented the results obtained by training these models on the V2 Plant Seedling dataset. We have shown the results obtained by the trained networks on the testing set. These five different models are taken VGG16, VGG19, ResNet50, Xception, MobileNetV2 and are discussed in the next section.

Accuracy Precision Scale

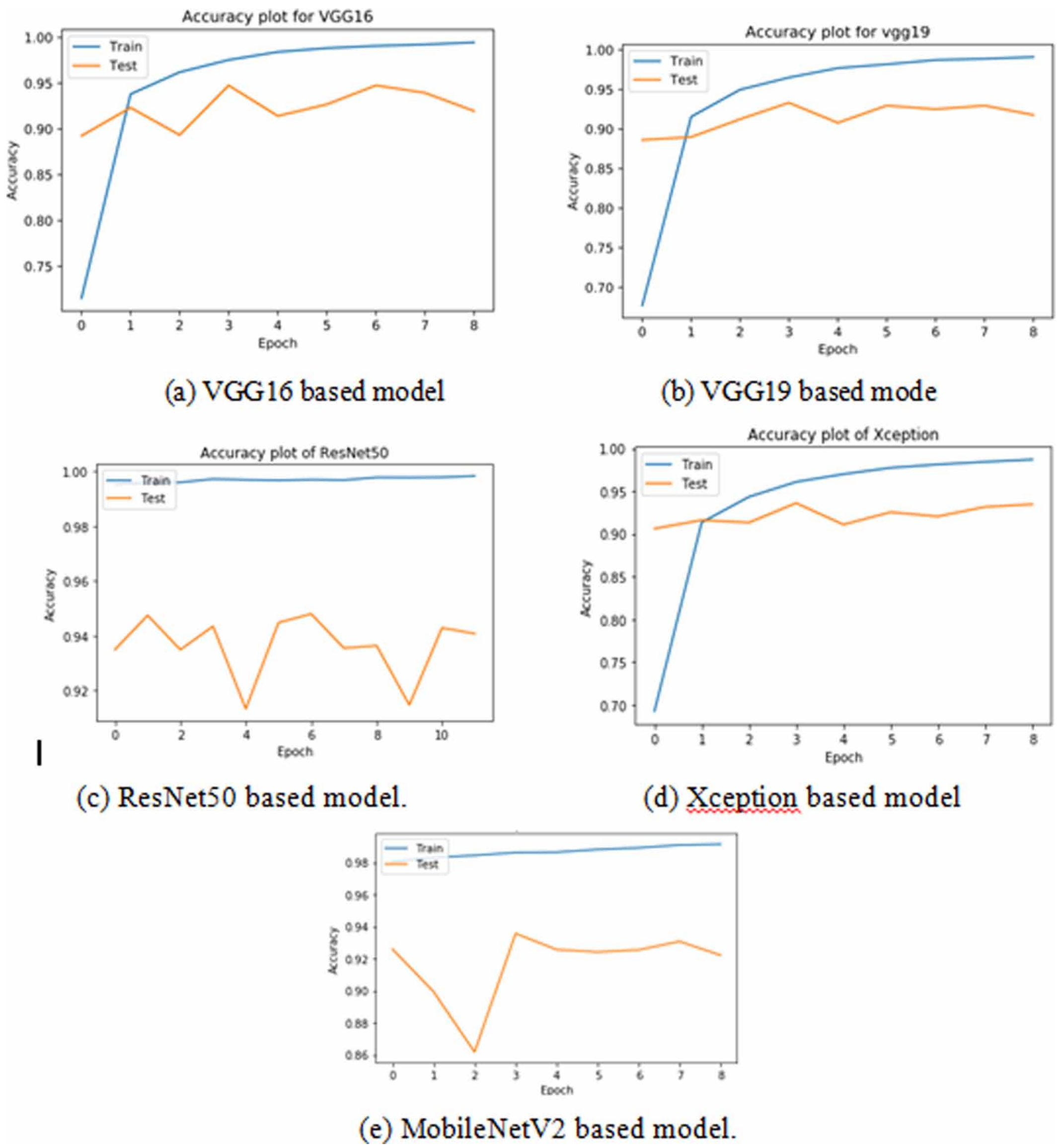
Graphically representation of the accuracy of the differences based model for training and testing on the dataset are shown in figure 5. It is proposed that VGG16 shows a training accuracy of 99.46% and a testing accuracy of 94.76% on the V2 Plant Seedling Dataset. The training accuracy plot shows that even in this case the previous knowledge has not proved useful but again, the second measure of transfer learning has helped to improved learning on the targeted dataset. The testing accuracy varies with each epoch with a peak value of 94.76%.

VGG19 has achieved a training accuracy of 99.10% and a testing accuracy of 93.32%., The accuracy achieved by the model is better than the previous accuracy on the dataset in the literature.

ResNet50 was the first pre-trained that was trained in our experiments. It has performed the best in terms of accuracy with 99.84% training accuracy and 95.23% testing accuracy. The accuracy plot shows that the training accuracy has been steady throughout the training period. The initial value of training accuracy shows that the measure of transfer learning that is working here is, transferred knowledge because the network has shown an initial accuracy of more than 98% in the first epoch. The testing accuracy has varied in a small range continuously with a peak value of 95.23%.

Xception has performed the best on the ImageNet validation dataset. The accuracy plot shows that the training accuracy has started from a very low value. The value of training accuracy has drastically improved after the training of the Xception based model on the V2 Plant Seedling dataset. The training accuracy plot shows that in this case, the previous knowledge has not proved useful. We achieved a training accuracy of 98.70% and a testing accuracy of 93.59% using Xception as the base model.

Figure 4. Accuracy plot of different models



The number of parameters in MobileNetV2 is least amongst all the pre-trained models used. The pre-trained architecture that is second in the list of a number of parameters is ResNet50, which when compared to MobileNetV2 has approximately 8 times as many parameters as ResNet50. This huge gap makes the running time of MobileNetV2 less as compared to others with a huge decrease in the number of calculations required per epoch. We achieved a training accuracy of 99.13% and a testing accuracy of 93.50% by using the model based on MobileNetV2. Accuracy plot of all the models are shown in figure 4.

The precision, recall, and F1-score of all species are calculated individually and then averages of all are taken and are shown in table 2.

Precision is the ability of a classifier not to label an instance positive that is actually negative. Recall is the ability of a classifier to find all positive instances The F1 score is a weighted harmonic

Table 2. Average classification report of all models

Models	Precision	Recall	F1-Score
VGG16	**0.95**	**0.95**	**0.94**
VGG19	0.93	0.93	0.93
ResNet 50	**0.95**	**0.95**	**0.95**
Xception	0.94	0.94	0.94
MobileNetV2	0.94	0.94	0.93

mean of precision and recall such that the best score is 1.0 and the worst is 0.0. F1 scores are lower than accuracy measures as they embed precision and recall into their computation. It is observed that VGG16 and ResNet50 give the best results for precision, recall, and F1-score.

Confusion Matrix

In this out of five models, the confusion matrix of the best two models is discussed. The confusion matrix for the VGG16 based model shown in Figure 5 suggests the same conclusion. The confusion matrix shows that 11 of the images belonging to Black grass class are classified as Loose Silky-bent and 19 of the images belonging to Loose Silky-bent class have been classified as Black-grass. Although, the total number of misclassifications between the two classes has been reduced to 30, which is the best amongst all the presented models.

Figure 6 shows the confusion matrix for ResNet50 based model. The confusion matrix shows that 28 Black-grass class images are classified as Loose-Silky bent. Therefore the model has highly misclassified these two classes but if compare with other models it gives the best results.

The reason for this high misclassification is that these two classes are highly similar to each other in their appearance. Figure 7 shows one instance each of images from each of these classes. In the figure you can see that the images in these two classes are highly similar to each other. This makes feature selection for correctly classifying these classes into their respective class difficult.

Comparison With Previous Work

Results are also compared with other existing work in Table 3 and concluded that the proposed method has achieved better accuracy. We have tuned five pre-trained CNN models used for image classification on the V2 Plant Seedling dataset to classify images of crop and weed seedling into their respective classes.

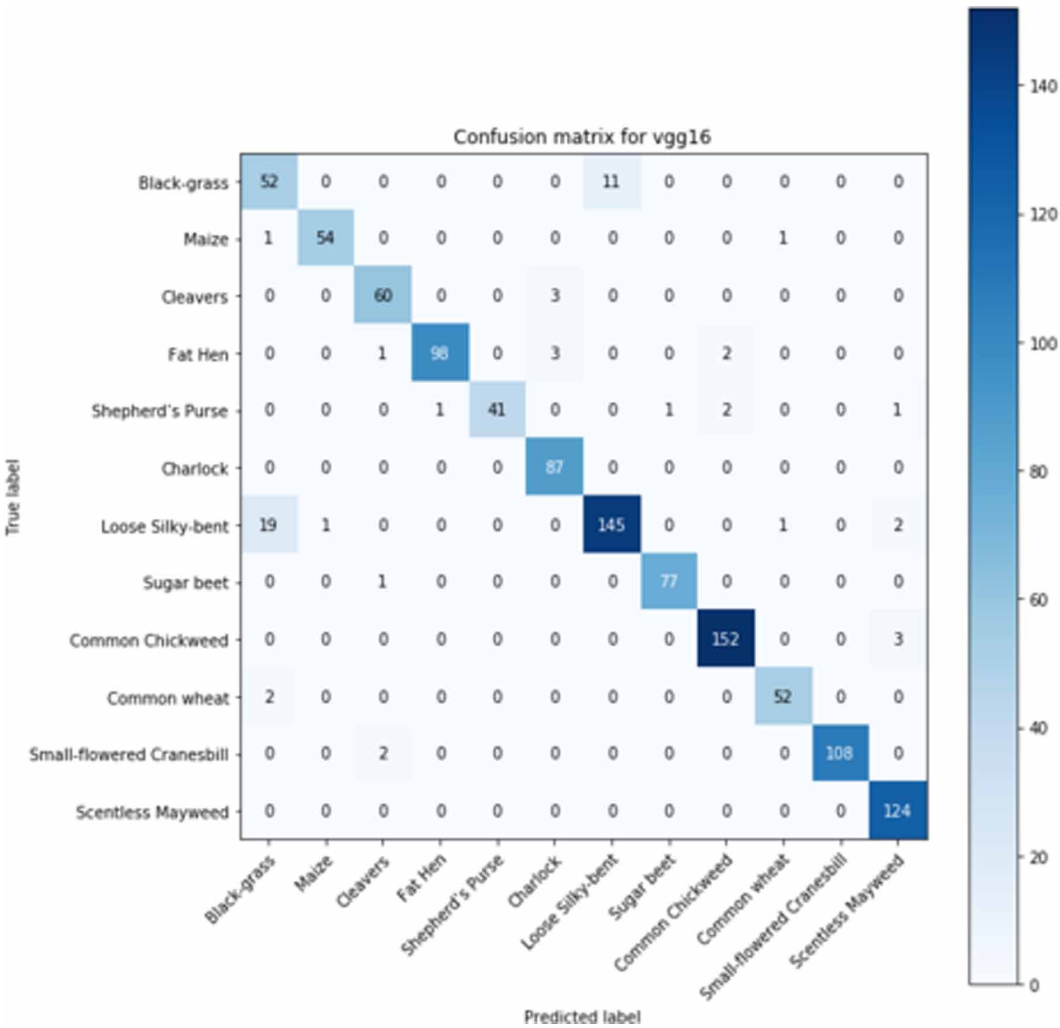
It shows that the deep learning techniques have performed much better than machine learning techniques on the dataset. We can see that amongst the deep learning techniques deep neural networks have performed better than the shallow neural network previously used on the dataset. Amongst the deep neural networks, ResNet50 has given the best results in our experiments.

CONCLUSION

In the proposed work five pre-trained CNN models are used for image classification on Plant seedling dataset to classify images of crop and weed seedling into their respective classes and have achieved improved results on the dataset. The five models used in this work are ResNet50, Xception, MobileNetV2, VGG16, and VGG19. Amongst these five models, ResNet50 gave the best results with 95.23% testing accuracy.

There are many models pre-trained on ImageNet dataset like GoogleNet, NASNet, etc. There could be an effort in the future to extend this work to evaluate all the pre-trained networks for the

Figure 5. Confusion matrix for VGG16 based model



task of plant seedling classification. There could be an attempt in the future to collect and publish a dataset that has all species of weeds in it with a sufficient number of images of each species to train a neural network.

Figure 6. Confusion matrix for ResNet50 based model

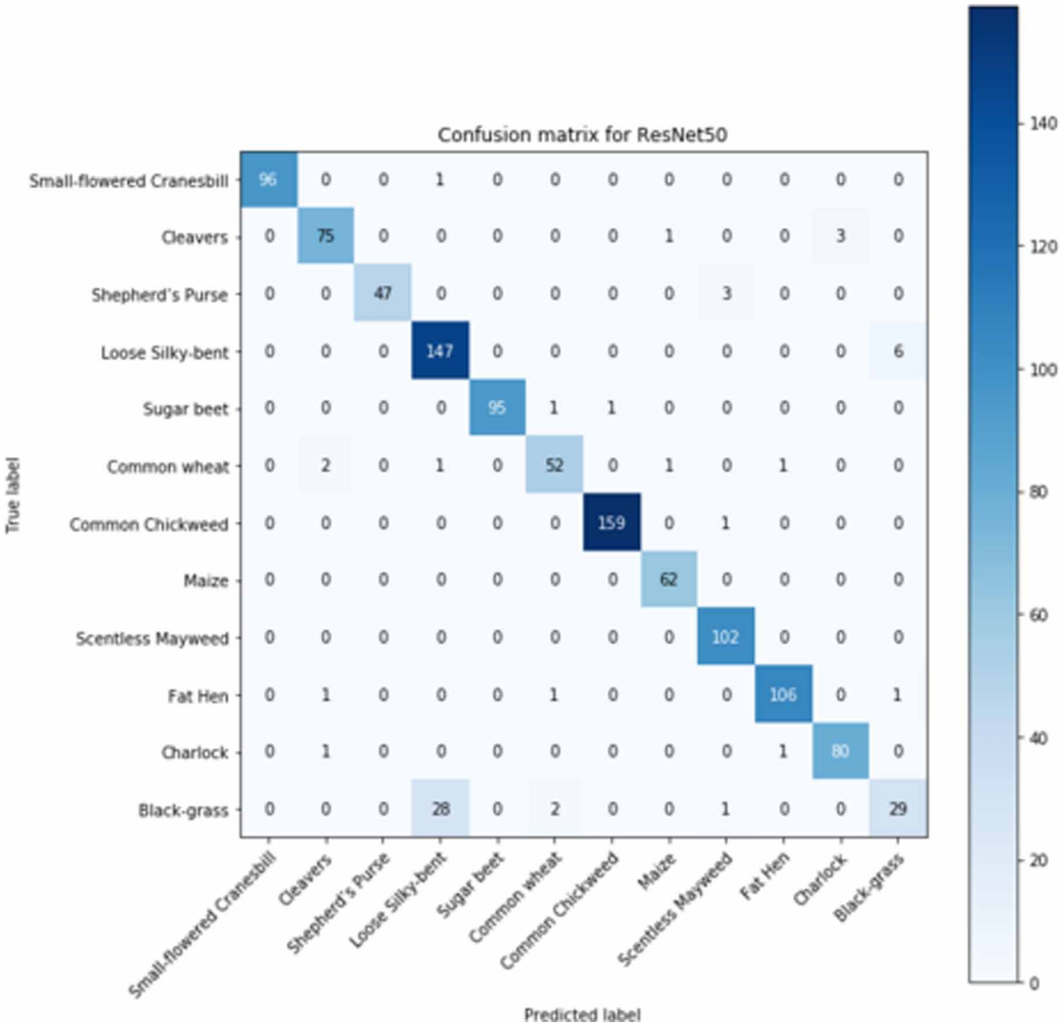


Figure 7. Black grass and Loose Silky-bent

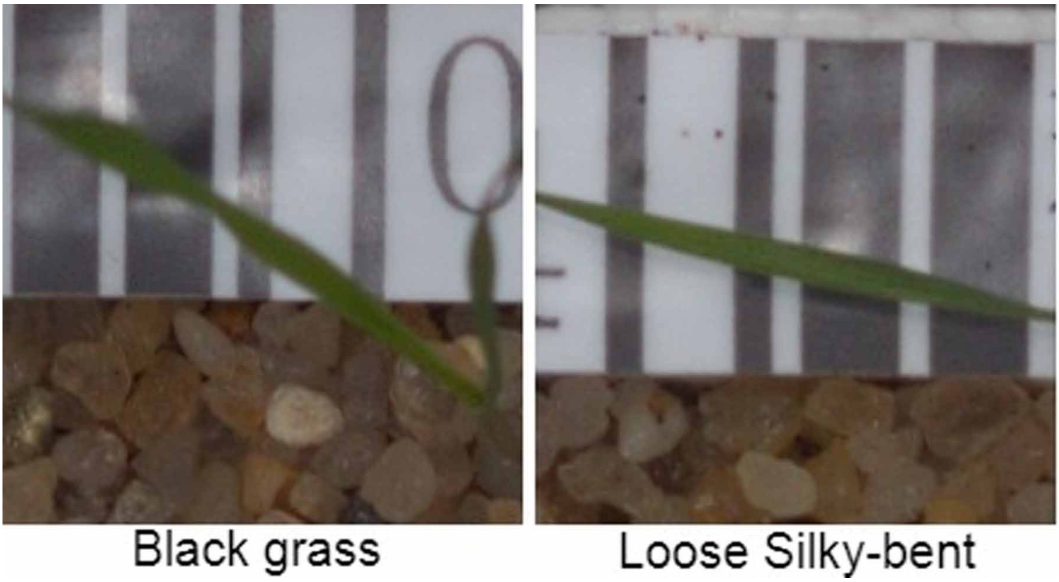


Table 3. Accuracy comparison with existing approaches

Author	Model	Accuracy(%)
Nkemelu, D. K. et.al.	KNN	56.84
	SVM	61.87
	CNN	92.60
Proposed work	ResNet50	**95.23**
	Xception	93.59
	VGG16	94.26
	VGG19	93.32
	MobileNetV2	93.50

REFERENCES

Ahmed, F., Al-Mamun, H. A., Bari, A. S. M. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection (Guildford, Surrey)*, *40*, 98–104. doi:10.1016/j.cropro.2012.04.024

Aitkenhead, M. J., Dalgetty, I. A., Mullins, C. E., McDonald, A. J. S., & Strachan, N. J. C. (2003). Weed and crop discrimination using image analysis and artificial intelligence methods. *Computers and Electronics in Agriculture*, *39*(3), 157–171. doi:10.1016/S0168-1699(03)00076-0

Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural network. *Biosystems Engineering*, *151*, 72–80. doi:10.1016/j.biosystemseng.2016.08.024

Giselsson, T. M., Jørgensen, R. N., Jensen, P. K., Dyrmann, M., & Midtiby, H. S. (2017). *A Public Image Database for Benchmark of Plant Seedling Classification Algorithms*. Retrieved from https://arxiv.org/abs/1711.05458

Haug, S., Michaels, A., Biber, P., & Ostermann, J. (2014). Plant classification system for crop /weed discrimination without segmentation. *2014 IEEE Winter Conference on Applications of Computer Vision, WACV 2014*, 1142–1149. doi:10.1109/WACV.2014.6835733

Lottes, P., Behley, J., Milioto, A., & Stachniss, C. (2018). Fully convolutional networks with sequential information for robust crop and weed detection in precision farming. *IEEE Robotics and Automation Letters*, *3*(4), 2870–2877. doi:10.1109/LRA.2018.2846289

McCool, C., Perez, T., & Upcroft, B. (2017). Mixtures of Lightweight Deep Convolutional Neural Networks: Applied to Agricultural Robotics. *IEEE Robotics and Automation Letters*, *2*(3), 1344–1351. doi:10.1109/LRA.2017.2667039

Milioto, A., Lottes, P., & Stachniss, C. (2018). Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs. In *2018 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2229–2235). IEEE. doi:10.1109/ICRA.2018.8460962

Milioto, A., Lottes, P., & Stachniss, C. (2017). Real-Time Blob-Wise Sugar Beets Vs Weeds Classification for Monitoring Fields using Convolutional Neural Networks. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4*(2W3), 41–48. 10.5194/isprs-annals-IV-2-W3-41-2017

Nkemelu, D. K., Omeiza, D., & Lubalo, N. (2018). *Deep Convolutional Neural Network for Plant Seedlings Classification*. Retrieved from https://arxiv.org/abs/1811.08404

Oerke, E. C. (2006). Crop losses to pests. *The Journal of Agricultural Science*, *144*(1), 31–43. doi:10.1017/S0021859605005708

Perez, A. J., Lopez, F., Benlloch, J. V., & Christensen, S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture*, *25*(3), 197–212. doi:10.1016/S0168-1699(99)00068-X

Xinshao, W., & Cheng, C. (2015). Weed seeds classification based on PCANet deep learning baseline. In *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)* (pp. 408–415). IEEE. doi:10.1109/APSIPA.2015.7415304