Advances in Intelligent Systems and Computing 1432

Subarna Shakya
Ke-Lin Du
Klimis Ntalianis *Editors*

Sentiment Analysis and Deep Learning

Proceedings of ICSADL 2022

Springer

Using Deep Learning Models for Crop and Weed Classification at Early Stage

Akshay Dheeraj and Satish Chand

Abstract Agriculture is essential for human existence, and it plays an important role in the world economy. There is increasing demand for food to feed the ever-increasing world population. Agriculture is affected by climate changes along with weed control. Weeds are unwanted plants that compete with plants for nutrition, and sunlight and adversely affect crop quality and production. Manual weeding is a tedious and labor-intensive task because both crop and weed look the same by visual appearance. Artificial intelligence techniques like deep learning can address this problem of crop and weed classification. In this research work, a deep learning-based classification system has been proposed to classify the weed and crop based on RGB images. We investigated two popular deep learning-based transfer learning models, namely DenseNet169 and MobileNetV2, and assessed their performances for crop and weed recognition. These models perform excellently with an accuracy of 97.14 and 94.92%, respectively. The significant accuracy results make the model an important tool for farmers to identify weeds.

Keywords Plant seedlings · Deep learning · Convolutional neural network · Weed classification · MobileNetV2 · DenseNet169 · Precision agriculture

1 Introduction

Convolutional neural networks (CNNs) have become essential for researchers in computer vision tasks. Large dataset availability and evolution in computing technologies with graphics processing units (GPU) have eased these vision tasks. CNNs are popular because of their applicability to various data and excellent performances

A. Dheeraj (✉)
ICAR-Indian Institute of Soil and Water Conservation, Dehradun, Uttarakhand, India
e-mail: akshay.dheeraj@icar.gov.in

S. Chand
School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India
e-mail: schand@mail.jnu.ac.in

S. Shakya et al. (eds.), *Sentiment Analysis and Deep Learning*,
Advances in Intelligent Systems and Computing 1432,
https://doi.org/10.1007/978-981-19-5443-6_69

[1]. CNN was started with the invention of LeNet-5 which consists of three convolution layers, two sub-sampling layers, and two fully connected layers [2]. CNN achieved further progress with AlexNet in which the number of the kernel was increased and dropout regularization was introduced. After that many networks were introduced that have significant performance [3–5].

Agriculture is the major source of livelihood for world populations. With the evolution in artificial intelligence, the agriculture sector is also being transformed which led to smart farming. Agriculture is being affected by different factors like climate change along with the daunting task of weed control. One of the most important tasks in smart farming is the recognition of weeds from the field and taking necessary steps to control the weeds so that crops get enough water content, solar radiation, and other important nutrient. Manual identification of weeds in the field is a difficult task as they appear to be the same in color, and size as the crop. The objective of this research study is to develop a weed and crop classification system using a deep learning model so that early detection of weeds can be done and crop yield can be increased. Images of weeds and crops have been trained on deep learning models like DenseNet169 and MobileNetV2, and their performances have been compared. A total of 12 classes of weed and crop have been used in the research work with originally 5544 images which have been augmented to increase the size of the dataset. The proposed research study shows the accuracy of 97.14 and 94.92% on test data for DenseNet169 and MobileNetV2 models, respectively. Thus, the significant accuracy of the DenseNet169 model suggests the effectiveness of the proposed system for weed recognition at an early stage in the farmer's field.

2 Literature Survey

A good amount of work has been done in for weed detection using computer vision techniques [6–9]. In [10], the author used SVM on digital images to classify weeds and crops. The authors used a total of 224 test images of six weed species, extracted a total of 14 features that can help to distinguish crops and weeds, and achieved a classification accuracy of 97%. In [11], the authors used fuzzy decision-making and shape descriptors to develop a weed and crop classifier and achieved recognition accuracy of 92.9% on a total of 66 images.

In [12], the authors used a machine learning model specifically supporting vector machine (SVM) for weed detection in the sugar beet field. Four different weed species of sugar beet were considered for research and different shape features like moment invariant and Fourier descriptor were extracted. The research study obtained the classification accuracy of 92.92 and 95% for ANN and SVM respectively.

These days, a convolutional neural network (CNN) is being used in every field [13–18]. In [19], some of the applications of CNN-based deep learning models in agriculture have been surveyed. Research work in [20] focuses on classifying soybean and its weeds using the k-means algorithm for feature extraction and combined with CNN. The authors used 820 images having four classes as soybean and its weeds and achieved an accuracy of 92.89%. In [21], the authors proposed a CNN-based graph convolution neural network (GCN)-based model named GCN-ResNet101 to classify weeds associated with corn, lettuce, and radish crop. Authors used four different datasets having a total of 4200, 560, 280, and 5040 images for corn, lettuce, radish, and mixed weed dataset and achieved 97.80, 99.37, 98.93, and 96.51 percent classification accuracy. In [22], the authors proposed a deep learning-based transfer learning model named ResNet101 to identify nine weed species associated with maize, common wheat, and sugar beet crop and achieved a classification accuracy of 96.04%. The author used augmentation techniques, and a total of 7200 images for 12 classes having 600 images for each class were used.

3 Proposed Methodology

3.1 Data Acquisition

A publicly available dataset has been used in this research work [23]. This dataset has a total of 12 species which includes three crop species maize, common wheat, sugar beet, and their associated nine weed species. Table 1 summarizes the dataset description.

Figure 1 shows some images of the dataset. Figure 2 shows the proposed methodology. The first step of the methodology is data acquisition. Once data is collected, data augmentation techniques are used to enlarge the dataset. The third step is to

Table 1 Detail of dataset used

	Species	Number of original images	Number of augmented images	Total images
1	Black-grass	310	590	900
2	Charlock	452	473	925
3	Cleavers	335	565	900
4	Common Chickweed	713	187	900
5	Common wheat	253	647	900
6	Fat-hen	538	362	900
7	Loose silky-bent	766	134	900
8	Maize	257	643	900
9	Scentless mayweed	607	293	900
10	Shepherd's purse	274	626	900
11	Small-flowered cranesbill	576	324	900
12	Sugar beet	463	437	900
Total		5544	5281	10,825

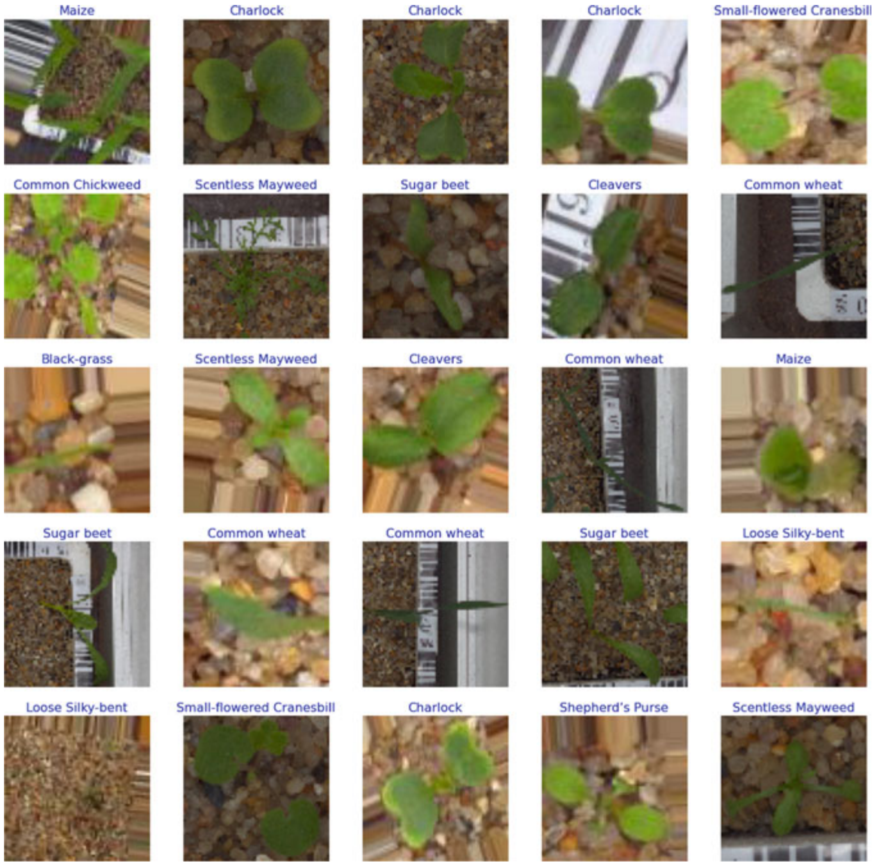

Fig. 1 Some sample images of the dataset

pre-process the data which includes resizing images according to the input size of the model used. These pre-processed data are the input to the proposed models which classify these crops and weed images.

3.2 Data Augmentation

Originally, this dataset has a total of 5544 images. To enlarge the dataset, data augmentation techniques are used. Table 2 summarizes the data augmentation techniques applied using the Keras library of Python. During data augmentation, images are flipped horizontally, rotated between the angles of 0–45°, width and height are shifted by 20%, zoomed by 20%, and shear by 20%.

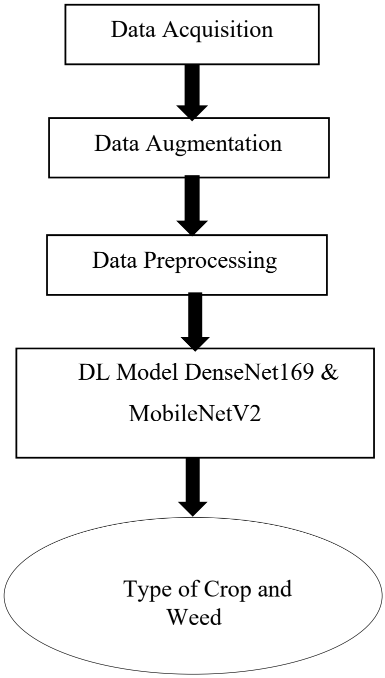

Fig. 2 Flowchart of proposed methodology

Table 2 Data augmentation techniques used

	Technique
1	Horizontal flip = true
2	Rotation = 32
3	Width shift = 0.2
4	Height shift = 0.2
5	Zoom range = 0.2
6	Shear range = 0.2
7	Fill = Nearest

3.3 CNN-based Deep Learning Model

Deep learning is being used in every field these days. It is the subfield of artificial intelligence where feature engineering is done automatically unlike machine learning where features are extracted manually. Two popular deep learning architectures have been used in the research.

- DenseNet169 [24]
- MobileNetV2 [25]

DenseNet169 Model

DenseNet consists of a number of dense blocks where within the block, the size of the feature map remains the same but the number of filters keeps on changing. DenseNet169 has four dense blocks. Each dense block defines one 1×1 convolution followed by a 3×3 convolution. The first dense block defines six 1×1 convolutions followed by a 3×3 convolution. The second dense block defines twelve 1×1 convolutions followed by 3×3 convolutions. The third dense block defines thirty-two 1×1 convolutions followed by 3×3 convolution. The fourth dense block defines thirty-two 1×1 convolution followed by 3×3 convolution There are three transition layers which does the convolution and pooling operation and are in place between the dense block. In DenseNet, all layers are connected. Feature maps from previous layers are concatenated.

MobileNetV2 Model

MobileNetV2 model was proposed by Google. This architecture uses depth-wise separable convolution through which efficiency is improved. There are 17 building blocks in MobileNetV2 after that 1×1 convolution is performed followed by the global average pooling and classification layer. One residual connection is used between the input and output layer which learns those features that are already learned and remove irrelevant features. MobileNetV2 model reduces the number of parameters which led to a reduction in computation.

DenseNet169 and MobileNetV2 have total 14.3 M and 3.5 M parameter, respectively, and top-5 accuracy of 93.2 and 90.1% on ILSVRC [26]. These smaller parameter sizes and excellent results have motivated us to choose these models for our research work.

4 Results and Discussions

Two CNN-based deep learning models DenseNet169 and MobileNetV2 have been proposed for crop and weed classification systems. We use the transfer learning approach in which these models that have been already pre-trained on the ImageNet dataset, are used for a new dataset for the research study. The proposed work has been implemented using Google Collaboratory pro which offers GPU and TPU support. These models have been trained on a batch size of 16 with 30 epochs and to optimize the training and validation loss, an Adam optimizer is used. Adam has been used to update the weight of the network iteratively. The dropout value is set to be 0.25 which means that one in four inputs are randomly dropped out during training. The momentum value is set to be 0.9, and rectified linear function (ReLu) has been used as an activation function.

Table 3 summarizes the various hyperparameters used for the proposed work.

For the above model, the same parameters have been used and their performance has been compared. Dataset was divided into 80, 10, and 10% data as training,

Table 3 Hyperparameters used for proposed work

Model	Optimizer	Epoch	Learning rate	Batch size	Dropout	Momentum	Activation function
DenseNet169	Adam	30	0.001	16	0.25	0.99	ReLU
MobileNetV2							

Table 4 Performance of proposed models on a dataset

Model	Avg. training accuracy (%)	Avg. validation accuracy (%)	Training time (sec.)	Avg. testing accuracy (%)
DenseNet169	99.82	97.32	2102	97.14
MobileNetV2	99.30	94.46	2128	94.92

validation, and test data respectively. Table 4 summarizes the performance of these two models on the dataset.

Figures 3 and 4 show the model accuracy and loss curve for DenseNet169 and MobileNetV2 models. From Table 4, it is evident that both models take the almost same time for training but the accuracy of DenseNet169 is higher than that of MobileNetV2. Although the models were trained for 30 epochs, we have used early stopping when results do not improve after three adjustments to the learning rate. Best accuracy results for both DenseNet169 and MobileNetV2 models have been obtained at epochs 22 and 23, respectively. Similarly, minimum loss for both DenseNet169 and MobileNetV2 model occurs at epochs 22 and 21, respectively. It is evident from Figs. 3 and 4 that the accuracy of the DenseNet169 and MobileNetV2 model increases with the increase in the number of epochs.

Initially, both training and validation loss were high but these losses decrease as the number of epochs increase. As shown in Figs. 3 and 4, minimum loss occurs at epochs 22 and 21 for DenseNet169 and MobileNetV2 models, respectively.

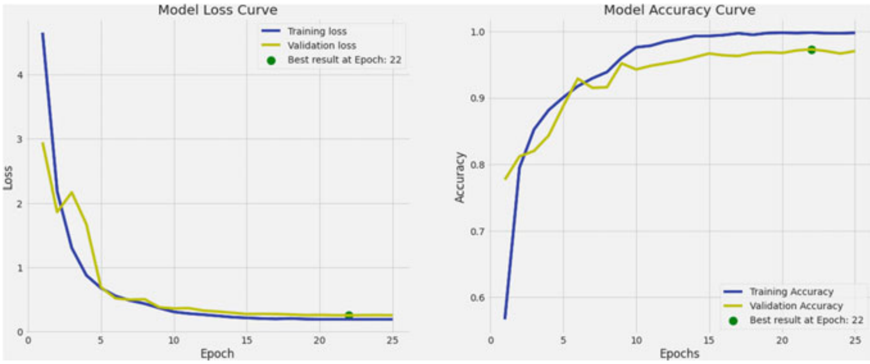

Fig. 3 Model loss and accuracy curve for DenseNet169 model

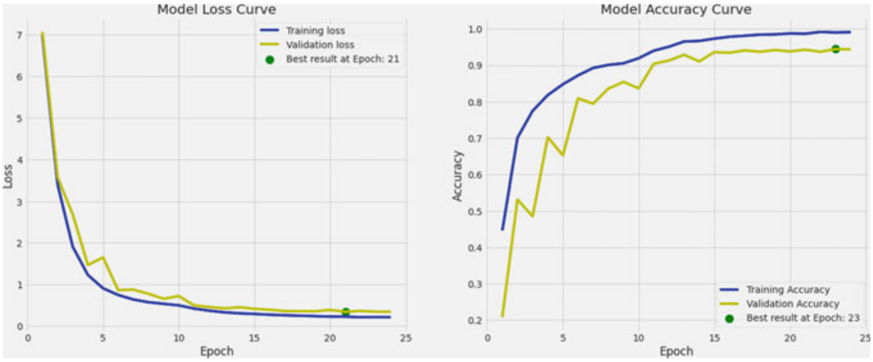

Fig. 4 Model loss and accuracy curve for MobileNetV2 model

For the object classification tasks, various performance metrics such as precision, recall, F_1-score, classification accuracy, and confusion matrix are used and calculated by using Eqs. (1), (2), (3), and (4). There are four terminologies used for calculations of these performance metrics which are true positive (TP), true negative (TN), false positive (FP), and false negative (FN). For this research study, TP means weed image is correctly classified as actual weed class and crop image is classified correctly as actual crop class.

TN means a number of correctly classified images in other classes except for the relevant class. FP means crop image is classified as weed and weed image is classified as crop image. FN means a number of incorrectly classified images in a relevant class.

$$\text{Precision}(i) = \frac{\#\text{TP}(i)}{\#\text{TP}(i) + \#\text{FP}(i)} \tag{1}$$

$$\text{Recall}(i) = \frac{\#\text{TP}(i)}{\#\text{TP}(i) + \#\text{FN}(i)} \tag{2}$$

$$\text{F}_1\text{-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Classification accuracy} = \frac{\#\text{TP}(i) + \text{TN}(i)}{\#\text{TP}(i) + \#\text{FP}(i) + \text{TN}(i) + \text{FN}(i)} \tag{4}$$

where i is the number of classes.

Precision is a measure of the ratio of correctly classified positives out of all positive instances, whereas recall is a measure of the proportion of actual positives that were identified correctly. F_1-score is the harmonic mean of precision and recall. A confusion matrix is an N $\times$ N matrix for analyzing the performance of a classification model where N is several classes. It is a visual representation of actual vs. predicted values.

The proposed model DenseNet169 and MobileNetV2 achieve the classification accuracy of 97.14 and 94.92%, respectively, on test data. Table 5 summarizes the information on precision, recall, and F_1-score of proposed models for 12 species.

The result of the proposed DenseNet169 and MobileNetV2 model for crop and weed classification has been analyzed using a confusing matrix and have been shown in Figs. 5 and 6.

From Table 5, it is evident that avg. precision for both DenseNet169 and MobileNetV2 is 97 and 95%, respectively. DenseNet169 has a precision value of less than 0.90 for two specie namely black-grass and loose silky-bent. Except for these two classes, the model has a precision value of more than 0.97 in the remaining ten classes. Similarly, MobileNetV2 has a precision value of 0.78, 0.87, and 0.89 for three species namely black-grass, loose silky-bent, and scentless mayweed. In the remaining nine classes, the model has a precision value of more than 0.96.

Table 5 Summarized information of precision, recall, and F_1-score of the proposed model for 12 species

	Species	DenseNet169			MobileNetV2		
		Precision	Recall	F_1-score	Precision	Recall	F_1 score
1	Black-grass	0.85	0.86	0.86	0.78	0.86	0.82
2	Charlock	1.00	0.98	0.99	1.00	0.98	0.99
3	Cleavers	0.99	0.99	0.99	0.99	0.99	0.99
4	Common chickweed	1.00	1.00	1.00	0.99	0.95	0.97
5	Common wheat	0.97	0.99	0.98	0.97	0.97	0.97
6	Fat-hen	0.98	0.97	0.97	0.98	0.98	0.98
7	Loose silky-bent	0.88	0.90	0.89	0.87	0.85	0.86
8	Maize	1.00	0.99	1.00	0.98	0.98	0.98
9	Scentless mayweed	0.99	0.99	0.99	0.89	0.95	0.91
10	Shepherd's purse	0.99	0.99	0.99	0.96	0.90	0.93
11	Small-flowered cranesbill	1.00	1.00	1.00	0.99	0.99	0.99
12	Sugar beet	1.00	0.99	0.99	1.00	0.98	0.99
Accuracy		0.97			0.95		
Macro Avg		0.97	0.97	0.97	0.95	0.95	0.95
Weighted Avg		0.97	0.97	0.97	0.95	0.95	0.95

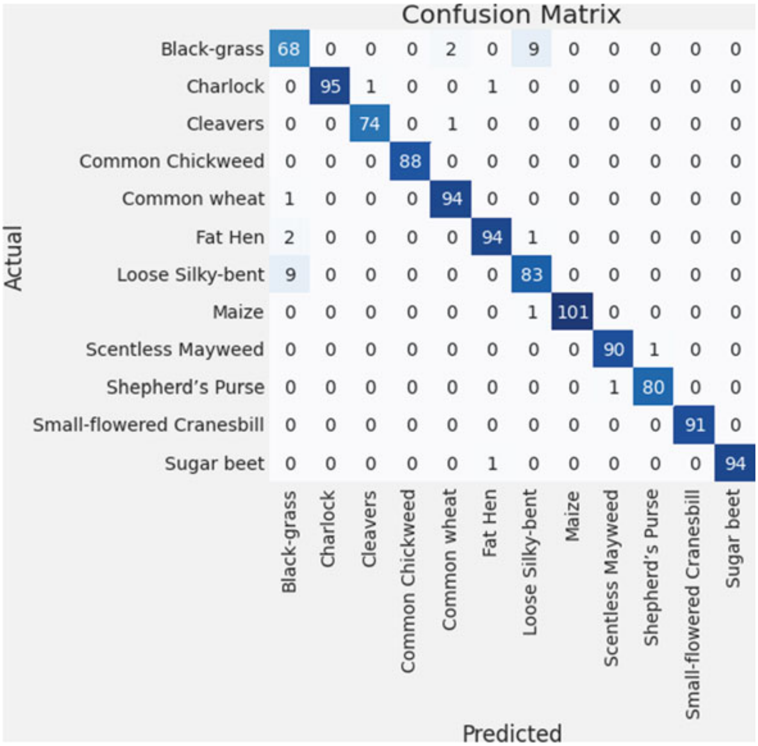

Fig. 5 Confusion matrix for DenseNet169 model

5 Conclusions

This scientific study aimed to develop a classification system for three crops and their associated nine weeds in the field. Transfer learning approaches have been used to train the proposed model. Then, we compared the performances of these two models for classifying weed and crop species into twelve classes. Both DenseNet169 and MobileNetV2 models achieved an accuracy of 97 and 95% with the dataset divided into 80, 10, and 10% as training, validation, and test data, respectively. Some data augmentation techniques were used to augment the data. Both models take almost the same time to train but the performance of DenseNet169 is better than that of MobileNetV2. Good accuracy results make the proposed system applicable to identify weeds at an early stage in the field that can assist the farmers to take action accordingly. In the future, we intend to extend the presented work by incorporating more crop and their associated weed species.

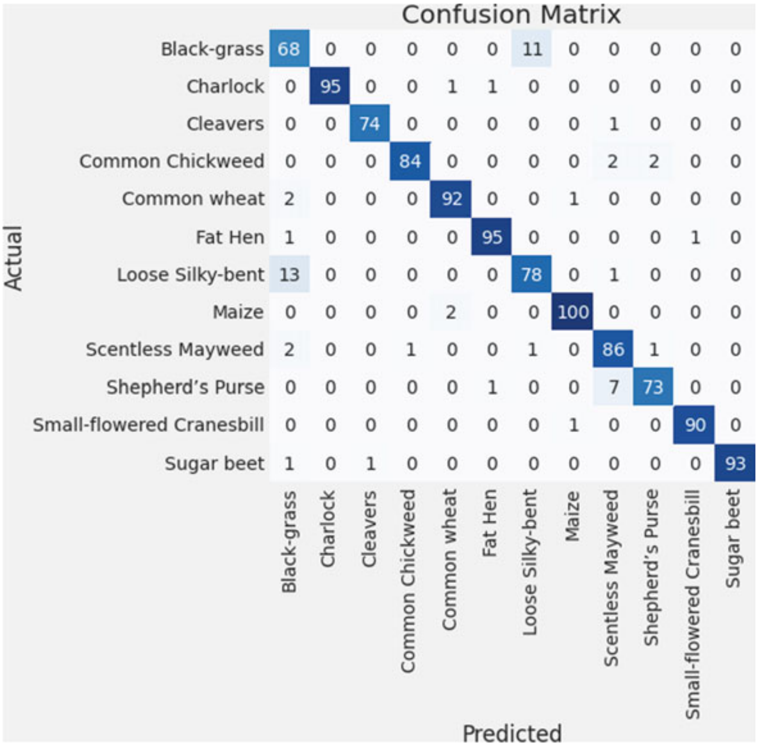

Fig. 6 Confusion matrix for MobileNetV2 model

References

1. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition, 77*, 354–377.
2. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning is applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324.
3. Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818–833). Springer.
4. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint*. arXiv:1409.1556.
5. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251–1258).
6. Hamuda, E., Glavin, M., & Jones, E. (2016). A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture, 125*, 184–199.
7. Mennan, H., Jabran, K., Zandstra, B. H., & Pala, F. (2020). Non-chemical weed management in vegetables by using cover crops: A review. *Agronomy, 10*(2), 257.
8. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors, 18*(8), 2674.

9. Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture, 158*, 226–240.
10. Ahmed, F., Al-Mamun, H. A., Bari, A. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection, 40*, 98–104.
11. Herrera, P. J., Dorado, J., & Ribeiro, Á. (2014). A novel approach for weed type classification based on shape descriptors and a fuzzy decision-making method. *Sensors, 14*(8), 15304–15324.
12. Bakhshipour, A., & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Computers and Electronics in Agriculture, 145*, 153–160.
13. Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing, 234*, 11–26.
14. Razzak, M. I., Naz, S., & Zaib, A. (2018). Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps* (pp. 323–350).
15. Kumar, T. S. (2020). Video based traffic forecasting using convolution neural network model and transfer learning techniques. *Journal of Innovative Image Processing (JIIP), 2*(03), 128–134.
16. Vijayakumar, T. (2020). Posed inverse problem rectification using novel deep convolutional neural network. *Journal of Innovative Image Processing (JIIP), 2*(03), 121–127.
17. Manoharan, J. S. (2021). Study of variants of extreme learning machine (ELM) brands and its performance measure on classification algorithm. *Journal of Soft Computing Paradigm (JSCP), 3*(02), 83–95.
18. Bashar, A. (2019). Survey on evolving deep learning neural network architectures. *Journal of Artificial Intelligence, 1*(02), 73–82.
19. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70–90.
20. Tang, J., Wang, D., Zhang, Z., He, L., Xin, J., & Xu, Y. (2017). Weed identification based on K-means feature learning combined with the convolutional neural network. *Computers and Electronics in Agriculture, 135*, 63–70.
21. Jiang, H., Zhang, C., Qiao, Y., Zhang, Z., Zhang, W., & Song, C. (2020). CNN feature-based graph convolutional network for weed and crop recognition in smart farming. *Computers and Electronics in Agriculture, 174*, 105450.
22. Sharma, P. (2019). Crops and weeds classification using convolutional neural networks via optimization of transfer learning parameters. *International Journal of Engineering and Advanced Technology (IJEAT)*, ISSN 2249-8958.
23. Giselsson, T. M., Jørgensen, R. N., Jensen, P. K., Dyrmann, M., & Midtiby, H. S. (2017). A public image database for benchmark plant seedling classification algorithms. *arXiv Preprint*. arXiv:1711.05458.
24. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700–4708).
25. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510–4520).
26. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). Imagenet large-scale visual recognition challenge. *International Journal of Computer Vision, 115*(3), 211–252.