Abstract

The advancement of weed detection technology has emerged as a significant development in various fields. In countries like India, where agriculture plays a crucial role in national development, weed detection has become an imperative technology. Weeds have a direct impact on crop growth, leading to reduced yields. Therefore, weed detection technology holds immense potential in addressing this issue. However, accurately classifying weed species poses a considerable challenge given the robust number of weed species. Deep learning models have emerged as a potential solution for this task. This research paper presents a baseline approach for classifying weeds growing among millet crops. The dataset used in this study comprises a total of 754 images of weeds and millets with clear and wild backgrounds, primarily grown in Chhattisgarh, India. State-of-the-art deep learning models, including MobileNet, InceptionV3, Resnet152v2, and VGG19, were employed to evaluate the dataset's accuracy. Among these models, MobileNet demonstrated superior performance, achieving a training accuracy of 92.86% when the images where captured in wild background. These results highlight the versatile applications of deep learning models in weed detection.

Dataset Collection

The initial aim of this project was to gather a size-able collection of labeled images in order to assist in the classification of various weed species using Neural network models. The use of deep learning techniques, particularly for object detection and classification, has become increasingly popular and necessary for this particular task. Consequently, several crucial factors were taken into consideration to enhance the learning process. These factors encompassed scene variability, dataset size, types of weed, weed locations, negative samples, image labelling, and clear and wild backgrounds.

In many cases, image processing frameworks fail to perform effectively in real-world applications due to unforeseen errors that arise during the initial and most critical step of the framework: image acquisition. Our objective is to employ the collected dataset to train a neural network model, specifically a Convolutional Neural Network (CNN), which will yield optimal accuracy in image classification.

Generally, the greater the number of confounding factors and variations in the dataset, the more complex and deep the model needs to be in order to achieve satisfactory performance. Therefore, a significant aspect of constructing this dataset is to capture images that encompass the full range of scene and target variations relevant to our desired application. Consequently, we have chosen to incorporate various sources of variation, such as illumination, rotation, scale, focus, dynamic backgrounds, and geographic variations in plant life.The dataset has been obtained from multiple authorised websites, such as www.gbif.org, https://portal.wiktrop.org, and several others.

GBIF | Global Biodiversity Information Facility

To ensure the dataset possesses the required variability and generality, two primary goals were established. Firstly, we aimed to collect a minimum of 200 images for each target weed species. Secondly, we aimed to achieve a balanced distribution of positive and negative class images (in a 50:50 split) for each location and species type. In this context, the negative class refers to images that do not depict any weed species. The first goal is essential when training highly complex CNNs, as they necessitate large labeled datasets. The second goal aims to prevent overfitting of the developed models to specific scene-level image features by ensuring that the targets are identified within their natural backgrounds.

Lastly, the dataset required expert analysis to label each image as either containing a target weed species or not. Additionally, any anomalous images that did not support the TensorFlow backend were removed from the obtained dataset.

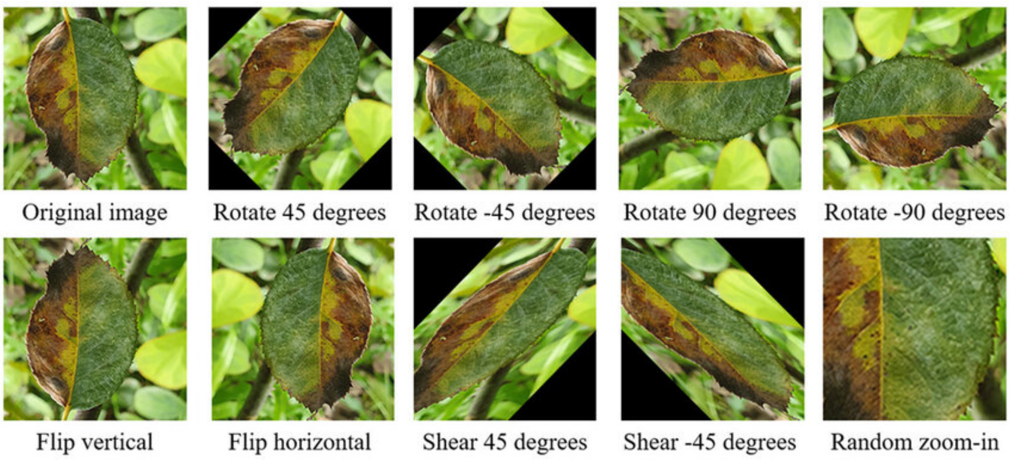
Models and their Rectification

The second objective of this study aimed to determine the expected accuracy of modern deep learning Convolutional Neural Networks (CNNs).To achieve this, the study utilized the high-level neural network Application Programming Interface (API) called Keras, along with the machine learning framework TensorFlow. Two popular CNN models, ResNet-50 (winner of ILSVRC 2015) and MobileNet (winner of ILSVRC 2017), were selected for implementation based on their strong performance on datasets with high variability and their availability in the Keras and TensorFlow backend.

Pre-Processing

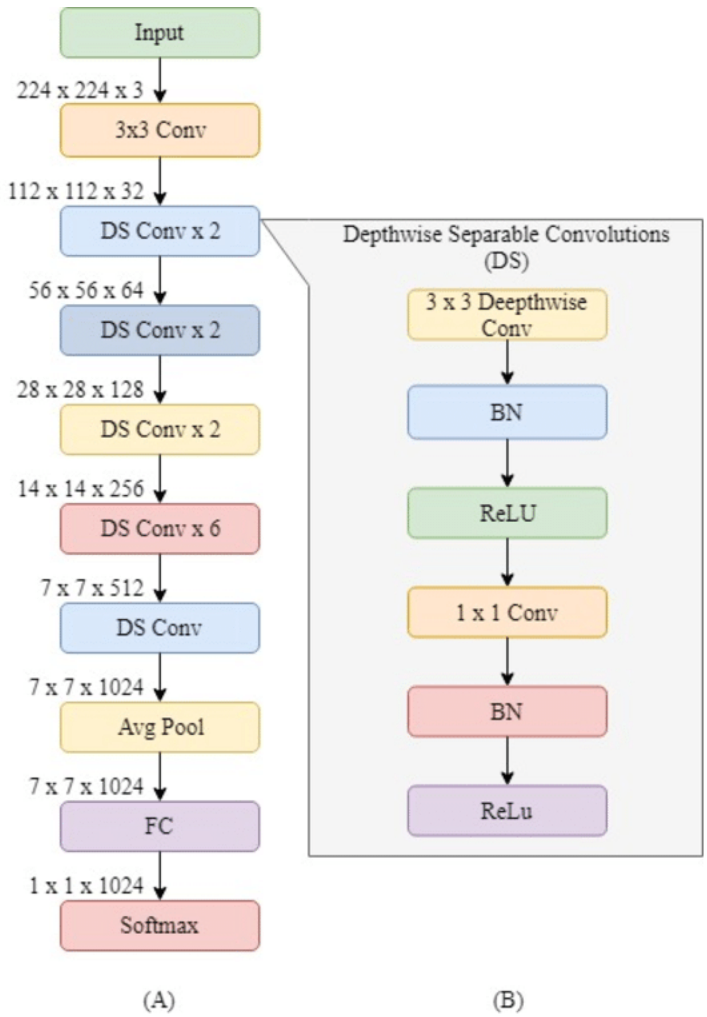
The dataset was divided into subsets for training, validation, and testing. The training subset consisted of 70% of the images, while 20% were allocated for validation to monitor the training process and prevent overfitting. The remaining 10% of images were reserved for testing and were not used during training. Before training, each model was initialised with weights pre-trained on the ImageNet dataset.

To address the high variability in the target weed classification task, various augmentations were applied to both the training and validation image subsets. These augmentations involved rotation, scaling, colour variations, illumination adjustments, and perspective transformations. OpenCV and the Keras. ImageDataGenerator module were utilized to perform these augmentations. All images were initially resized to 224 × 224 pixels and were randomly augmented for each training epoch. Each image was randomly rotated within the range of -360 to +360 degrees. Pixel intensities were shifted within the range of [0,1] for each colour channel. Additionally, images were horizontally flipped, cropped to retain the required 224 × 224 pixel size and shuffled the batches for the input layer of each architecture. Without extensive augmentation, both CNN models tended to overfit the available images by memorising the training subsets.



Model Architecture

Over the years, several image classification models have been developed, including VGG, InceptionV3, MobileNet, and ResNet. We experimented with these popular models to find the most suitable ones that could achieve the highest accuracy for our data. After thorough evaluation, we determined that both the MobileNet and ResNet-50 models, available in Keras with pre-trained weights from the Imagenet database in the TensorFlow backend, was the optimal choice. Among these two models, MobileNet emerged as the most suitable option and its original architecture, which was initially trained on the ImageNet dataset was set to be false (i.e *layer.trainable = False*), and few “Dense” layers were added at the end of the MobileNet network to classify the four species classes in the DeepWeeds dataset whose weeds are associated with the following millets:”Finger”, “Kudo", "Pearl" and “Sorghum”. This modification involved replacing the last fully connected(Flattened) layer with a Dense layer of 128 units (with ‘ReLU’ activation function and followed by a new fully connected layer containing *five neurons*(Fifth being the negative class) .



The input matrices for each network had dimensions of d × 224 × 224 × 3, where d represented the number of images per training batch and 3 denoted the number of image colour channels (RGB). The spatial average pooling layer was utilized to convert the fully-convolutional output, which was then connected to the final weed classification layer with dimensions of “d × 5” . The fifth class in this layer represented the negative class, and each batch contained 32 images with input dimensions of 224 × 224 × 3.

Since each image in the DeepWeeds dataset could contain multiple weed species, a Softmax activation function was used for each specific weed neuron in the output layer. This allowed the models to produce probabilities for each class, indicating the likelihood of an image belonging to a particular weed species. An image was classified as one of the target weeds if the Softmax-activated neuron probability for that class was high, which is determined by using the Argmax operation.

Compiling the Model

For the training of both models, the Adam optimiser, which is a gradient-based method for stochastic optimisation, was employed using the Keras implementation. The initial learning rate (lr) was set to 1 × 10^ −1. After every 5 epochs, the learning rate was halved if the validation loss did not decrease using the ***ReduceLROnPlateau*** callback. The training process was conducted in batches of 32 images, and if the validation loss did not decrease after 10 epochs, the training was Stopped using ***EarlyStopping.***

```
EarlyStopping(monitor='val_loss')
```

Problems Faced

Furthermore, apart from dealing with complex backgrounds, the presence of interfering objects, particularly neighbouring plants, can unexpectedly obstruct the view of the target foreground objects. This presents yet another unavoidable source of variation in the dataset. Additionally, the dataset needs to consider seasonal variations in the target weed species. This means that a single weed species class should include images of the weed with and without flowers and fruits, and in varying states of health. These variations can impact foliage colour, the strength of features, and other visible anomalies.

During the initial stages of collecting the dataset, we encountered issues where we couldn't train the data due to certain image formats not being supported by the TensorFlow backend. Consequently, we had to remove all the unsupported directories to proceed with our task.
Moreover, if your GPU lacks sufficient power, an alternative option is to use Google Colab as your notebook. Google Colab is a hosted Jupyter Notebook service that requires no setup and provides free access to computing resources, including GPUs and TPUs.

Maximizing the accuracy of our model was a complex undertaking that involved making several adjustments. These adjustments included adding extra layers (such as Dense, Dropout, and Batch Normalization) ensuring that the model neither overfits nor underfits the training data, as well as employing different activation functions.

Results from various models

Various models were trained and the evaluation criteria was the model accuracy and validation accuracy. Our objective was to develop a model with highest validation accuracy to get the best model possible for weed classification. Two types of datasets were evaluated, one is of images consisting of clear background and the other was image dataset of crops grown in wild. The results of all the models according to the backgrounds can be found in **Table 1** (Wild background) and **Table 2** (Clear background). The graphs in **Fig 4** (Wild background) and **Fig 5** (Clear Background) illustrate the model's performance with respect to epochs.

Model	Training accuracy	Validation accuracy
VGG-19	89.00%	76.79%
InceptionV3	89.00%	85.71%
Resnet152v2	93.22%	88.00%
MobileNet	92.86%	91.07%

Table 1 (Wild Background)

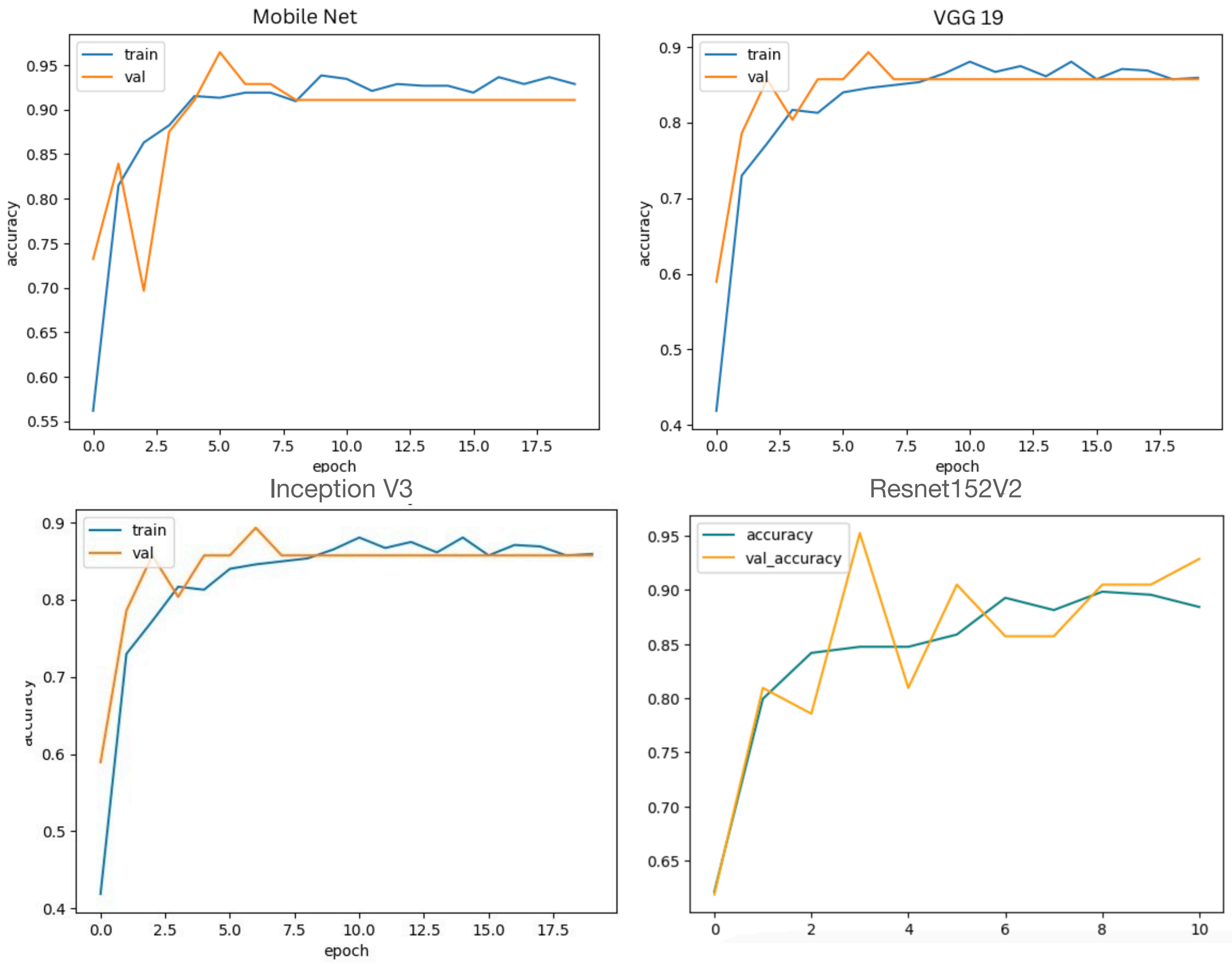


Fig 4 (Wild Background)

Model	Training accuracy	Validation accuracy
VGG-19	93.90%	75.00%
InceptionV3	93.29%	81.25%
Resnet152v2	92.07%	75.00%
MobileNet	96.34%	93.75%

Table 2 (Clear Background)

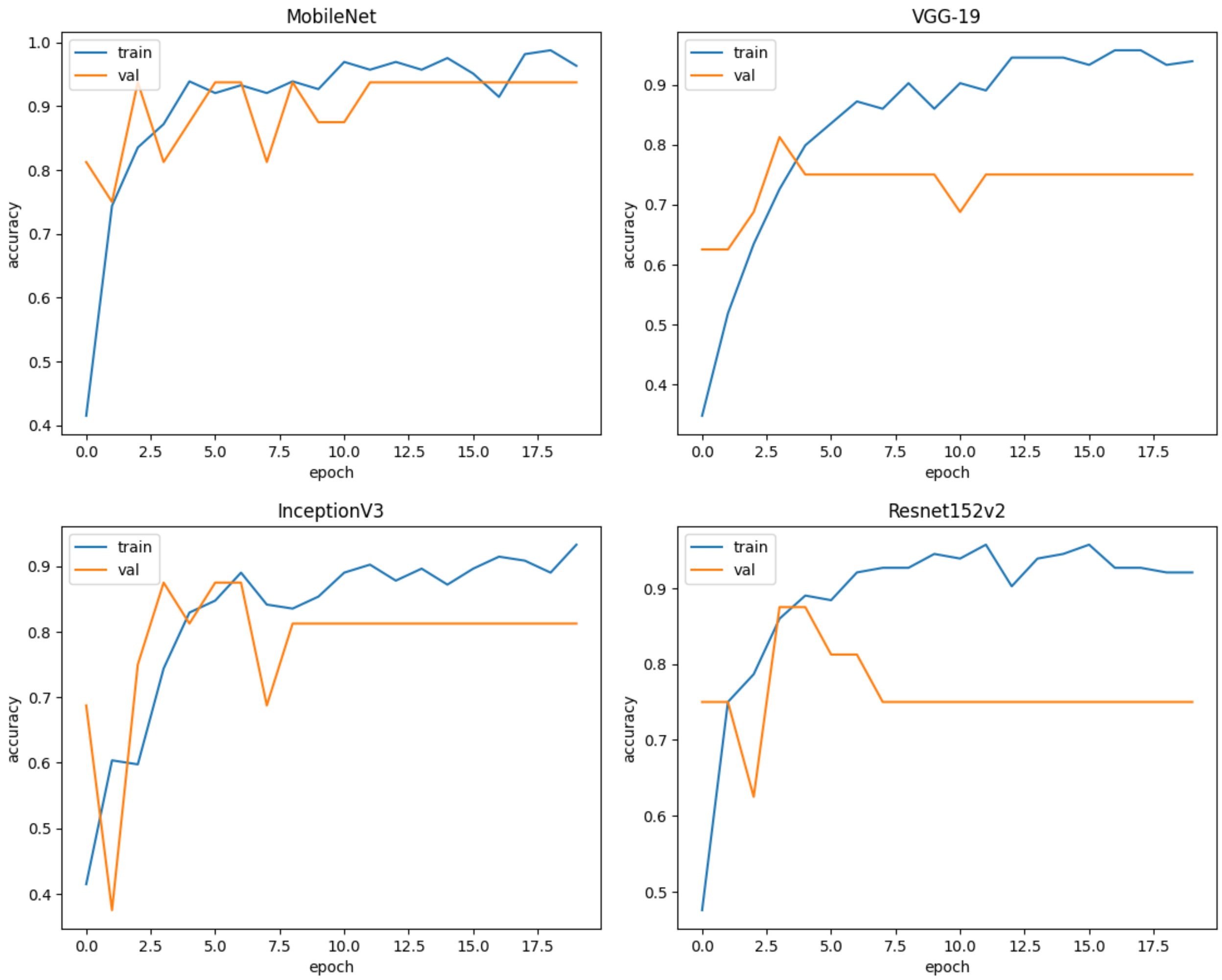


Fig 5 (Clear Background)

Conclusion

To encapsulate, this work is a multiclass weed and crop dataset evaluated over different models calibrated to find the optimal model to classify the weed species growing among crops. Speculating the results we can conclude the MobileNet model has outperformed all other state-of-the-art models with a training accuracy of 92.86% and validation accuracy of 91.07% for wild background and training accuracy of 96.34% and validation accuracy of 93.75% for clear background. Moreover, the architecture of MobileNet is relatively lightweight when compared to other models stated in this article, allowing it to be able to run on a device with smaller computational power.

Data Availability

The dataset and the source code used for this work can be found at the following GitHub repository: https://github.com/SrikarMukkamala/Weed_Classification

References

1. Olsen, Alex & Konovalov, Dmitry & Philippa, Bronson & Ridd, Peter & Wood, Jake & Johns, Jamie & Banks, Wesley & Girgenti, Benjamin & Kenny, O.P. & Whinney, James & Calvert, Brendan & Rahimi Azghadi, Mostafa & White, Ron. (2019). DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning. Scientific Reports. 9. 10.1038/s41598-018-38343-3.
2. www.gbif.org
3. https://portal.wiktrop.org
4. Bakhshipour, A. & Jafari, A. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.* **145**, 153–160 (2018).
5. Dyrmann, M., Jørgensen, R. N. & Midtiby, H. S. RoboWeedSupport - Detection of weed locations in leaf occluded cereal crops using a fully convolutional neural network. *Adv. Animal Biosci.* **8**, 842–847 (2017).
6. Lee, S. H., Chan, C. S., Mayo, S. J. & Remagnino, P. How deep learning extracts and learns leaf features for plant classification. *Pattern Recognit.* **71**, 1–13 (2017).
7. Tang, J.-L., Chen, X.-Q., Miao, R.-H. & Wang, D. Weed detection using image processing under different illumination for site-specific areas spraying. *Comput. Electron. Agric.* **122**, 103–111 (2016).
8. https://www.frontiersin.org/articles/10.3389/fpls.2022.1053329/full
9.