

Semi-supervised Teacher-Student Architecture for Weed Classification

Nassima Bousahba^{1,2}, Mahamed Abdelmadjidallali^{3,4}, Kawther Karnache¹, Imane Fettah¹, Faiza Boudjeltia⁵

¹ Computer Science Department, Hassiba Benbouali University of Chlef, Algeria

² EEDIS Laboratory, Djillali Liabes University, Sidi Bel Abbes, Algeria

³ SIMPA Laboratory, Mohamed Boudiaf University, Usto, Oran, Algeria

⁴ LIO Laboratory, Hassiba Benbouali University of Chlef, Algeria

⁵ The National Plant Protection Institute, Labyedh Medjadja, Chlef, Algeria

Corresponding author e-mail : n.bousaha@univ-chlef.dz

Article Info

Article history:

Received , 19/10/2024

Revised , 20/11/2024

Accepted , 08/12/2024

Keywords:

Convolutional Neural Network
Weeds Classification
Deep Learning
Mobilenet_V2
Xception
Teacher-Student Framework.

ABSTRACT

Recent advances in deep learning have demonstrated remarkable success in image classification tasks, particularly in distinguishing between weeds and crops. This research focuses on weed classification using the Weed_Ch32 dataset, which comprises 5,023 smartphone-captured images of 32 weed species collected from Algeria's National Plant Protection and Institute, Labyed Medjadja-Chlef. Our evaluation encompasses several established models including VGG16, Xception, Inception, Mobilenet, and Mobilenet v2. In our experiments, Mobilenet v2 achieved 88% accuracy in supervised learning, with performance improving to 98% when using semi-supervised techniques. Building upon these results, we explored a semi-supervised teacher-student training approach combining Mobilenet v2 and Xception architectures. This combined model achieved an accuracy of 70% on the test set.

I. Introduction

In agricultural settings, weeds emerge naturally as they are indigenous to the environment. These weeds compete with crops for essential resources including moisture, nutrients, light, and space, leading to reduced agricultural productivity. Therefore, effective weed management is crucial throughout the crop production cycle. While pre-emergence weed control can be achieved through cultivators and blanket rate herbicides, controlling weeds becomes challenging once crops are established in the field.

Without proper weed management, crop yields can suffer significantly. Studies show that weeds alone can cause 34% yield loss, while pests and diseases result in 18% and 16% losses respectively [1]. Annual crop losses due to weeds typically range between 23% and 44% in conventional agricultural fields [1]. Research indicates that timing is critical in weed treatment for improving productivity and reducing the soil weed seed bank [2][3]. Post-emergence weed control often requires precise methods, such as physical removal or targeted spraying, to eliminate weeds near crop plants while minimizing environmental impact and protecting the crops.

Weed management techniques in agricultural areas have evolved to include manual, organic, mechanical, electrical, and chemical methods. However, these conventional approaches are often time-consuming and lack precision [4]. The need for site-specific weed management has grown due to the inefficiencies of traditional methods, which are labor-intensive and costly. Conventional equipment often struggles with complete row weed

elimination. Current practices typically involve uniform herbicide application across entire fields, affecting both weeds and crops. This broad application of herbicides can have greater environmental consequences compared to targeted approaches [5]. Selective herbicide application offers improved accuracy while reducing both costs and environmental impact [6].

The foundation of successful site-specific weed control lies in accurate weed identification and classification. Modern image-based weed categorization leverages machine learning and deep learning algorithms [7][8][9]. These technologies have shown promising results across various agricultural applications, including crop identification [10], yield prediction [11], plant classification [12], and disease detection [13]. Their success extends to weed classification challenges as well [14][15][16][17].

Despite advances in autonomous weed and crop identification, challenges remain in developing effective weed categorization systems across diverse crop species. Key limitations include insufficient weed species databases and variations in mobile-based plant images due to different contributors, cameras, locations, seasons, and plant types.

To address these challenges, we developed the weed_Ch32 dataset, collecting 5,030 images of weed species from the National Plant Protection Institute, Medjadja, Chlef, and local farms using mobile phones. We implemented MobileNet_V2 in both supervised and semi-supervised learning approaches, achieving accuracy scores of 88% and 97% respectively on the weed_Ch32 dataset. Furthermore, we explored semi-supervised teacher-student training using MobileNet v2 with Xception, resulting in a combined model that achieved 70% accuracy on the test set.

The remaining sections are grouped as follows. Section 2 describes the suggested dataset and technique in depth. Section 3 illustrates the experiment and outcomes, while Section 4 depicts the conclusion and describes the scope of future study.

II. Dataset

The Weed_Ch32 dataset consists of images captured in natural environments using portable equipment. It contains 5,023 images representing 32 distinct weed species collected from the National Plant Protection Institute, Labiod Medjadja, and surrounding farms in the Chlef state. This dataset was specifically chosen for its real-world applicability, as it captures weeds in their natural growing conditions.

Each weed class contains approximately 150 images (both labeled and unlabeled) captured using a smartphone equipped with a 12 Megapixel RGB sensor capable of 3120×4208 resolution. These images reflect actual field conditions, making them particularly valuable for developing practical weed identification systems.

Table 1 provides a comprehensive overview of the dataset's characteristics, including its size, composition, and distribution of images.

Table 1. Dataset description

Dataset size	1.93 GO
Dataset type	Categorical
Number of classes	32
Labelled dataset	1524
Unlabelled dataset	3499
Total number of images	5023
Average number of images per class	~150
Dataset sources	INPV_chlef

Figure 1 presents representative samples from the dataset, showcasing the diversity of weed species captured under various natural conditions. To illustrate the distribution of different weed species in our dataset, Figure 2 displays a pie chart showing the proportion of images for each species.



Figure 1. Dataset samples

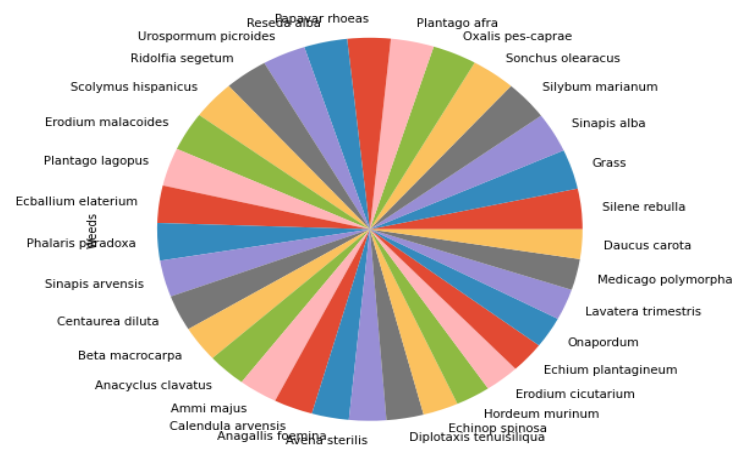


Figure 2. Pie chart of weed specifiers

For the preprocessing phase, we maintain the RGB color format of the images as input to the network. The primary preprocessing step involves resizing all images to 224×224 pixels, followed by pixel normalization through division by 255. This standardization ensures consistent input dimensions while preserving the essential visual characteristics of the weeds.

III. Problem statement

In order to identify the weed species from a (224x224) image, we will produce 32 probabilities for each of our 32 classes.

III.1. Loss function

Cross-entropy is the loss function utilised during training. The formula for this compares the similarity of two probability distributions.

$$\text{Loss} = \sum_1^K \hat{p}_i \log(p_i) \quad (1)$$

Where:

K is the number of classes

$\hat{p}_i$ Represents the actual probabilities,

p_i represents the predicted probabilities.

III.2. Evaluation metrics

Accuracy measures are utilized to evaluate the classification performance employed in this study. The accuracy is computed by dividing the number of successfully detected weed images by the total number of weed images tested.

$$Accuracy = \frac{N_{Correct}}{N_{Total}} \quad (2)$$

IV. Methodology

IV.1. Benchmark Transfer Learning Models

Transfer learning applies pre-trained weights from established models to our weed_Ch32 dataset. We selected several benchmark models based on their proven performance in image classification tasks: VGGNet[18], Xception[19], Inception_V3[20], ResNet_V2[21], Inception-ResNet_v2[22], MobileNet[23], and MobileNet_V2[24]. Our methodology progresses from these baseline models through supervised learning, advancing to semi-supervised approaches, and culminating in semi-supervised teacher-student training.

1) VGGNet

The VGGNet will be the first model we attempt. The VGGNet is essentially a collection of CNN layers with dropout and pooling. Convolutional layers in this model are 2 x 2 or 3 x 3. We employ the VGG16 particularly.

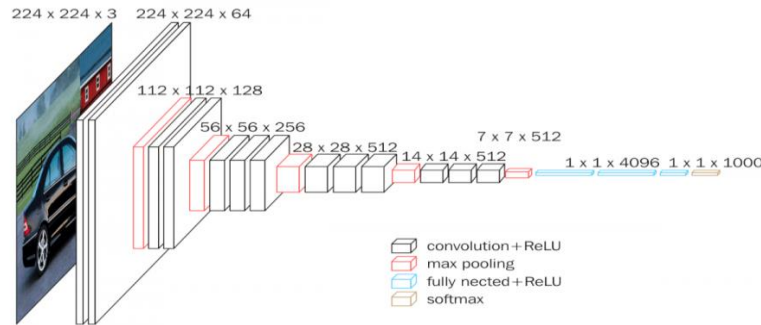


Figure 3. VGG16 Architecture

2) Inception_V3

A convolutional neural network architecture from the Inception family called Inception-v3 makes use of batch normalisation for sidehead layers, Label Smoothing, factorised 7 x 7 convolutions, an additional classifier to carry label information farther down the network.

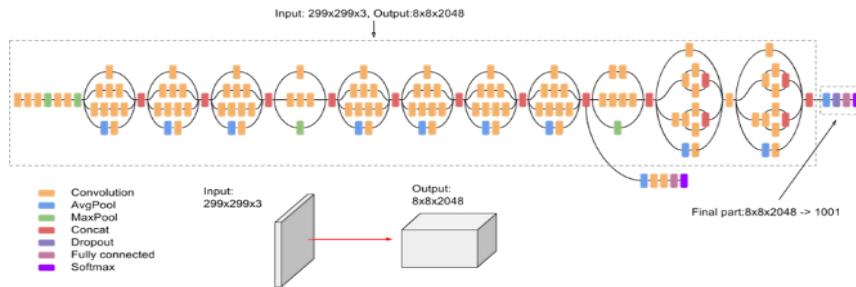


Figure 4. Inception_V3 Architecture

3) Xception

The Google Inception model served as the basis for the Xception model, a 71-layer deep CNN that takes the Inception model to its logical limit. Its design is composed of convoluted layers that are separable according to depth.

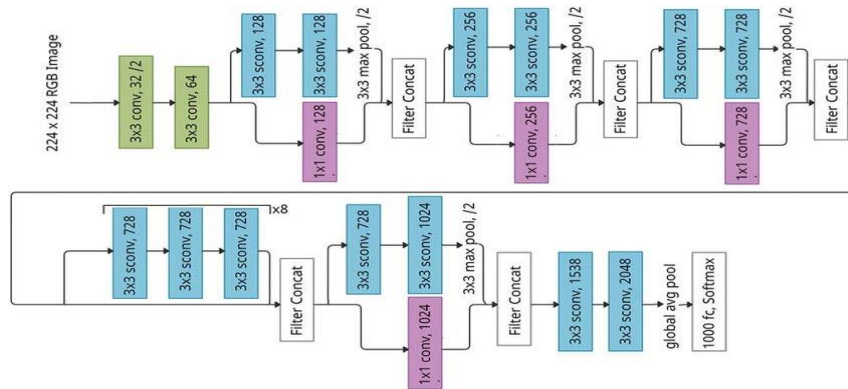


Figure5. Xception architecture

4) InceptionResNet_v2

InceptionResNet_v2 is created by combining the Inception structure and the Residual connection. In the Inception-Resnet block, convolutional filters of different sizes and residual connections are put together. The degradation issue brought on by deep structures is avoided by using residual connections, and training time is decreased. The Inception-Resnet-v2 network architecture is shown in the figure.

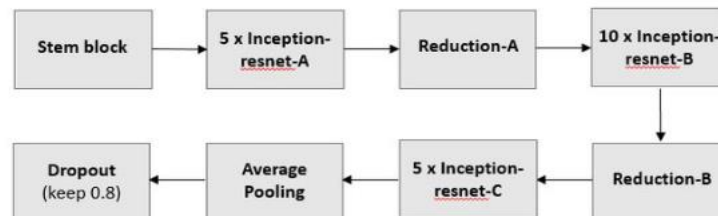


Figure 6. The architecture of InceptionResnet_v2

5) MobileNet

An effective model for embedded and mobile vision applications, MobileNet is a simplified architecture that builds shallow deep convolutional neural networks using depthwise separable convolutions [15]. As shown in Figure 7, the core of MobileNet's structure is a set of depth-wise separable filters.

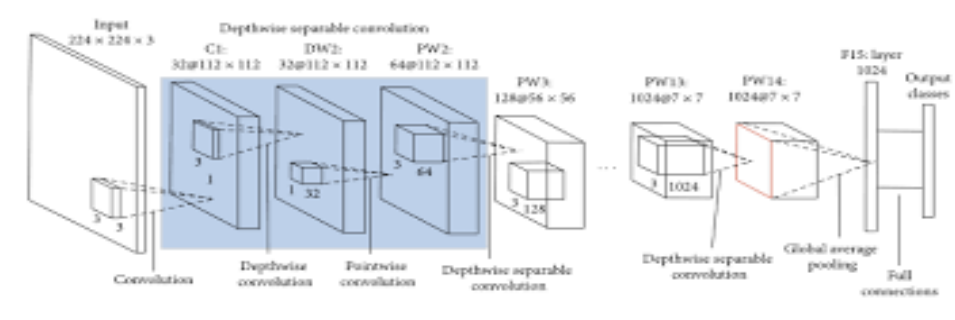


Figure 7. Architecture of MobileNet

6) MobilenetV2

In comparison to conventional residual models, which use extended abstractions in the input, the MobileNet v2 architecture is based on a reversed residual structure, in which the input and output of the residual block are constrained bottleneck layers. Features in the intermediate expansion layer are filtered using MobileNet v2's lightweight depthwise convolutions. The representational power was also maintained by eliminating non-linearities in the thin layers.

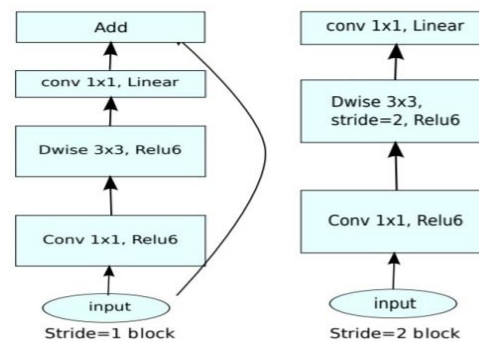


Figure 8. MobileNetV2 architecture

IV.2. Teacher-Student architecture

Building upon our findings from the benchmark models, we implement a teacher-student architecture to leverage both labeled and unlabeled data. This approach begins by training a teacher model on labeled images to generate class label distributions. These predictions are then used to create pseudo-labels for unlabeled data. A student model, typically larger in capacity, is subsequently trained on both the original labeled data and the pseudo-labeled data. This process iterates, with each trained student model becoming the teacher for the next iteration, continually refining the pseudo-labels. Figure 9 illustrates this progressive framework.

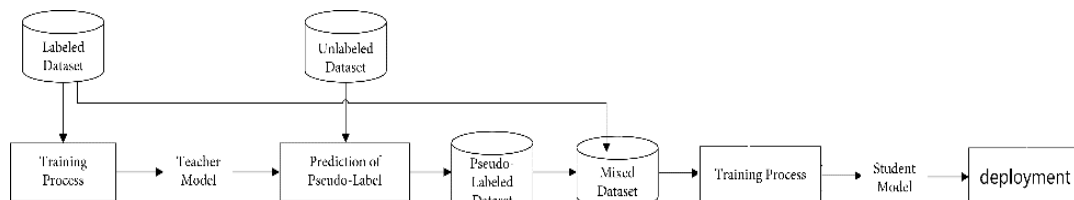


Figure 9. Teacher Student framework process

V. Experimental results

This section presents our experimental results across different learning approaches. All models were trained using consistent parameters: batch size of 32, Adam optimizer, and an initial learning rate of 1e-3 with decay over 30 epochs.

V.1. Supervised learning

We evaluated six transfer learning benchmark models: VGGNet, Inception_v3, Xception, InceptionResNet_v2, MobileNet, and MobileNet_v2. Each model was enhanced with a linear classifier for the final classification layer.

Table 2. Test accuracy for the different models

Model	Accuracy	Fit time
InceptionResNet_v2	64%	40 min
MobileNet	65%	40min
Vgg16	67%	37 min
Xception	81%	35 min
Inception_v3	84%	38 min
MobileNet_v2	88%	30 min

MobileNet_v2 achieved superior performance (88% accuracy) while requiring the least training time (30 minutes). This exceptional performance can be attributed to its efficient architecture, which uses inverted residual blocks and lightweight depth-wise convolutions, making it particularly well-suited for our dataset's characteristics. In contrast, InceptionResNet_v2, despite its deeper architecture, achieved lower accuracy (64%) and required more training time, suggesting that its complexity may have led to overfitting on our specific dataset.

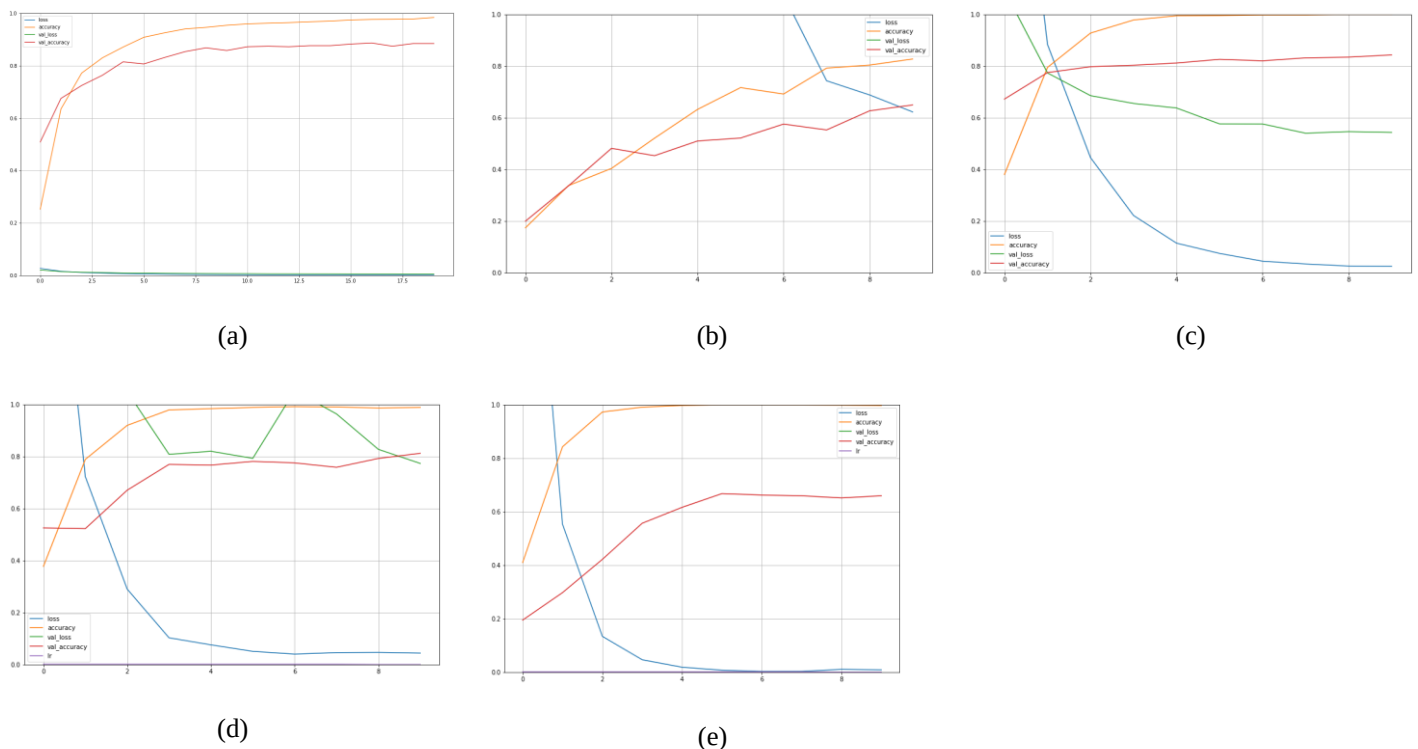


Figure 10. Accuracy and Loss graph in supervised learning (a) MobileNet_V2, (b) InceptionResNetV2, (c) InceptionV3 ,(d) Xception,(e) MobileNet.

V.2. Semi-Supervised learning

For semi-supervised experiments, we divided the dataset into 30% labeled and 70% unlabeled data.

Table 3. Test accuracy for the different models in SSL

Model	Accuracy	Fit time
Inception_v3	39 %	2h 21 min
Vgg16	77%	1h for 1 epoch
MobileNet_v2	98%	53 min

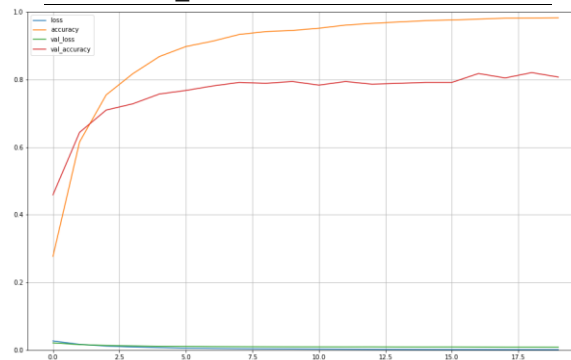


Figure11. Accuracy and Loss graph in demi- supervised learning of MobileNet_V2

MobileNet_v2 significantly outperformed other models in the semi-supervised setting, achieving 98% accuracy. This remarkable improvement over its supervised learning performance suggests that the model effectively leveraged the unlabeled data to enhance its feature representations.

V.3. Teacher-Student Framework

The teacher-student framework implementation maintained the same data distribution as our semi-supervised experiments, using a 30/70 split of labeled and unlabeled data.

Table 4. Test accuracy for the different models in Teacher-Student framework

Teacher model	Student model	Accuracy	Fit time
MobilenetV2	Inception_V3	39%	2h 30min
MobilenetV2	Mobilenet_V2	46%	2h 10min
MobilenetV2	Xception	70%	4h 20min

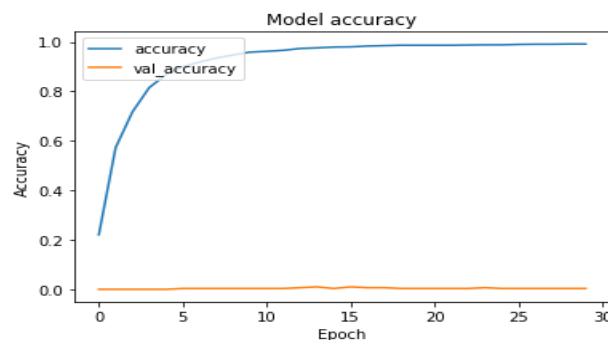


Figure 12. MobileNet_v2 as Teacher

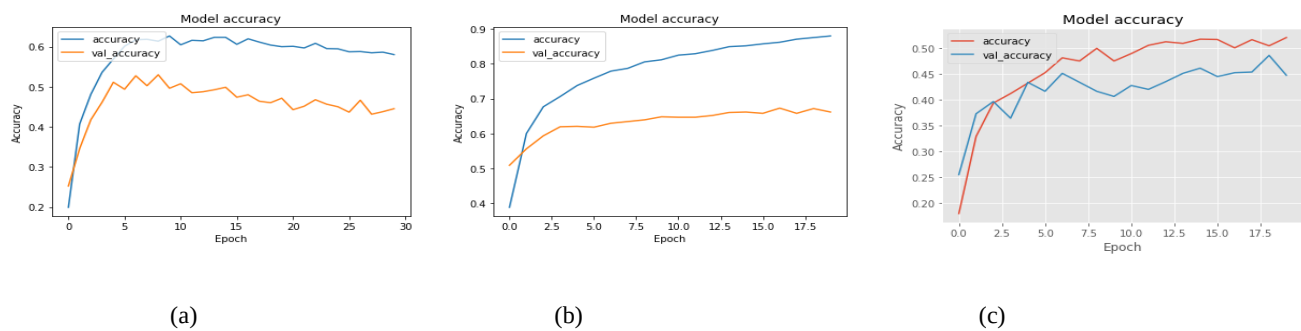


Figure 13. Student models. (a)Inception_V3, (b) Xception, (c) MobileNet_V3

The combination of MobileNet_v2 as the teacher model and Xception as the student model achieved the highest accuracy at 70%, though requiring the longest training time of 4 hours and 20 minutes. When using MobileNet_v2 as both teacher and student, the model achieved 46% accuracy with a reduced training time of 2 hours and 10 minutes. The MobileNet_v2-Inception_V3 pairing showed the lowest performance at 39% accuracy despite a significant training duration of 2 hours and 30 minutes.

We can observe that the teacher-student approach with MobileNet_v2 and Xception achieved 70% accuracy, demonstrating the potential of knowledge transfer between models, albeit at a higher computational cost than individual model training.

The results suggest that while the teacher-student framework offers promising accuracy, it comes with significantly increased computational requirements. Future work could explore optimizing this trade-off through more efficient knowledge distillation techniques or hybrid approaches.

VI. Conclusion

Precision agriculture represents an advanced management approach that aims to maximize agricultural production while maintaining sustainability and profitability. Within this domain, weed control remains a critical yet costly and environmentally impactful component of farming operations. Computer vision systems offer promising solutions for efficient weed identification and management, potentially reducing both economic and environmental costs.

This research makes significant contributions to the field of weed classification through deep learning applications in real-world agricultural settings. Our comprehensive evaluation of CNN-based transfer learning models on the Weed_Ch32 dataset yielded several important findings. Using only 30% labeled data, MobileNet V2 demonstrated superior performance in both supervised and semi-supervised learning scenarios, achieving accuracies of 88% and 98% respectively. Furthermore, our teacher-student framework, combining MobileNet V2 with Xception, achieved 70% accuracy, demonstrating the potential of knowledge transfer between models, albeit with increased computational requirements.

These results suggest that while deep learning models can effectively classify weeds in natural environments, there are important trade-offs between model complexity, computational resources, and classification accuracy. Future research directions include expanding the Weed_Ch32 dataset to encompass additional species at various growth stages, optimizing model efficiency through improved knowledge distillation techniques, and developing more robust solutions for real-time detection on mobile devices. These advancements will further contribute to the practical implementation of automated weed management systems in precision agriculture.

Acknowledgements

This research was conducted within the framework of the PRFU project ; the title is “Embaded system and IoT” and Code COOL07UN020120220002 of the Department of Computer Science of the University Hassiba Benbouali of Chlef, funded by the Algerian Ministry of Higher Education and Scientific Research.

References

- [1] E. Oerke, ‘Crop losses to animal pests, plant pathogens, and weeds’, in *Encyclopedia of Pest Management, Volume II*, CRC Press, 2007, pp. 116–120.
- [2] S. Z. Knezevic, ‘Datta The critical period for weed control: revisiting data analysis’, *Weed Sci*, vol. 63, pp. 188–202, 2015.
- [3] N. Tursun, A. Datta, M. S. Sakinmaz, Z. Kantarci, and S. Z. Knezevic, ‘Chauhan The critical period for weed control in three corn (*Zea mays* L.) types Crop Protect’, vol. 90, 2016.
- [4] A. A. Bajwa, G. Mahajan, and B. S. Chauhan, ‘Nonconventional weed management strategies for modern agriculture’, *Weed Sci.*, vol. 63, no. 4, pp. 723–747, 2015.
- [5] J. Ahmad *et al.*, ‘Visual features based boosted classification of weeds for real-time selective herbicide sprayer systems’, *Comput. Ind.*, vol. 98, pp. 23–33, 2018.
- [6] S. L. Young and D. K. Giles, *Automa-tion: The Future of Weed Control in Cropping Systems*. Dordrecht: Springer Nether-lands, 2014.
- [7] J. Behmann, A.-K. Mahlein, T. Rumpf, C. Römer, and L. Plümer, ‘A review of advanced machine learning methods for the detection of biotic stress in precision crop protection’, *Precis. Agric.*, vol. 16, no. 3, pp. 239–260, 2015.
- [8] J. Haichen, C. Qingrui, and L. Zheng Guang, ‘Weeds and crops classification using deep convolutional neural network’, in *2020 the 3rd International Conference on Control and Computer Vision*, 2020.
- [9] A. S. M. M. Hasan, F. Soheli, D. Diepeveen, H. Laga, and M. G. K. Jones, ‘A survey of deep learning techniques for weed detection from images’, *Comput. Electron. Agric.*, vol. 184, no. 106067, p. 106067, 2021.
- [10] M. Tufail, J. Iqbal, M. I. Tiwana, M. S. Alam, Z. A. Khan, and M. T. Khan, ‘Identification of tobacco crop based on machine learning for a precision agricultural sprayer’, *IEEE Access*, vol. 9, pp. 23814–23825, 2021.
- [11] P. Nevavuori, N. Narra, and T. Lipping, ‘Crop yield prediction with deep convolutional neural networks’, *Comput. Electron. Agric.*, vol. 163, no. 104859, p. 104859, 2019.
- [12] S. A. Wagle, R. Harikrishnan, S. H. M. Ali, and M. Faseehuiddin, ‘Classification of plant leaves using new compact convolutional neural network models’, *Plants*, vol. 11, no. 1, p. 24, 2021.
- [13] R. Kumar, A. Chug, A. P. Singh, and D. Singh, ‘A systematic analysis of machine learning and deep learning based approaches for plant leaf disease classification: A review’, *J. Sens.*, vol. 2022, pp. 1–13, 2022.
- [14] K. Hu, G. Coleman, S. Zeng, Z. Wang, and M. Walsh, ‘Graph weeds net: A graph-based deep learning method for weed recognition’, *Comput. Electron. Agric.*, vol. 174, no. 105520, p. 105520, 2020.
- [15] S. Khan, M. Tufail, M. T. Khan, Z. A. Khan, and S. Anwar, ‘Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer’, *Precis. Agric.*, vol. 22, no. 6, pp. 1711–1727, 2021.
- [16] Sunil, Y. Zhang, C. Koparan, M. R. Ahmed, K. Howatt, and X. Sun, ‘Weed and crop species classification using computer vision and deep learning technologies in greenhouse conditions’, *J. Agric. Food Res.*, vol. 9, no. 100325, p. 100325, 2022.
- [17] A. H. Al-Badri *et al.*, ‘Classification of weed using machine learning techniques: a review—challenges, current and future potential techniques’, *J. Plant Dis. Prot. (2006)*, vol. 129, no. 4, pp. 745–768, 2022.
- [18] K. Simonyan and A. Zisserman, ‘Very deep convolutional networks for large-scale image recognition’, *arXiv [cs.CV]*, 2014.
- [19] F. Chollet, ‘Xception: Deep learning with depthwise separable convolutions’, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, ‘Rethinking the inception architecture for computer vision’, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

- [21] K. He, X. Zhang, S. Ren, and J. Sun, 'Identity mappings in deep residual networks', in *Computer Vision – ECCV 2016*, Cham: Springer International Publishing, 2016, pp. 630–645.
- [22] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, 'Inception-v4, Inception-ResNet and the impact of residual connections on learning', *Proc. Conf. AAAI Artif. Intell.*, vol. 31, no. 1, 2017.
- [23] A. G. Howard *et al.*, 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications', *arXiv [cs.CV]*, 2017.
- [24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.