Weed detection in a Rice Crop through Image Processing and Classification Using Convolutional Neural Networks

RANAIVOSON, Tojonirina and RASAMIMANANA, Hantanirina Rosiane and RASOANAIVO, Andriniaina Narindra and ANDRIANARIMANANA, Omer and ANDRIAMAMONJY, Alfred and RAZAFIMAHATRATRA, Dieudonné

University of Antananarivo, Doctoral School of Natural Resources Management and Development, Université of Antananarivo, Ecole Supérieur des Sciences Agronomiques (ESSA),Doctoral School of Natural Resources Management and Development, Université of Antananarivo, Ecole Normale Supérieur (ENS), Doctoral School of Natural Resources Management and Development

28 November 2024

Weed detection in a Rice Crop through Image Processing and Classification Using Convolutional Neural Networks

Ranaivoson Tojonirina M.Z[1], Rasamimanana Hantanirina Rosiane[2], Andrianarimanana Omer[3], Razafimahatratra Dieudonné[3], Andriamamonjy Alfred[3], Rasoanaivo Andriniaina Narindra[3]

[1]University of Antananrivo, Doctoral School of Natural Resources Management and Development
[2]University of Antananarivo, Ecole Supérieur des Sciences Agronomiques (ESSA), Doctoral School of Natural Resources Management and Development (EDGRND)
[3]University of Antananarivo, Ecole Normale Supérieur (ENS), Doctoral School of Natural Resources Management and Development (EDGRND)

Abstract

Artificial Intelligence (AI) today occupies a central ranking, especially in a context where technological progress is omnipresent. Among the most influential tools, deep learning has established itself in both professional and academic domains. This article focuses on the effectiveness of convolutional neural networks for detecting weeds competing with rice. To achieve this, an extension of the pre-trained Inception_V3 model was used for image classification, while MobileNet was employed for image processing. This innovative approach, tested on a rice field where distinguishing between rice and weeds is challenging, represents a significant advancement in the AI field. However, the training of both models revealed limitations: Inception_V3 exhibited overfitting after the 10th iteration, while MobileNet showed high volatility and overfitting from the first iteration. Despite these challenges, Inception_V3 stood out for its superior accuracy.

Keywords*: Convolutional, Neural, Pre-trained, Detection*

1 Introduction

The rise of machine learning has transformed numerous sectors, with applications ranging from spam email detection to customer segmentation for advertisements, weather forecasting, and climate change studies. Machine learning also plays a crucial role in discovering genetic sequences linked to diseases and in developing algorithms for autonomous vehicles and drones (Lantz, 2019).

Artificial neural networks, inspired by the biological functioning of the central nervous system, especially the human brain, are an integral part of the machine learning family (O'Shea & Nash, 2015). These networks, though powerful, differ from convolutional neural networks, which belong to the deep learning category. Convolutional neural networks are particularly effective in image processing and computer vision due to their complex structure based on stacking multiple convolutional layers (Tuffery, 2023).

These research advancements, accompanied by technological progress, have gradually been integrated into agriculture, where deep learning and machine learning are increasingly used (AgroTIC, 2018; Liakos et al., 2018). Computer vision, for instance, has found applications in crop monitoring (Wu et al., 2021), especially for weed detection in various plantations. Models such as CaffNet have been used to detect weeds in soybean fields (dos Santos Ferreira et al., 2017), while WeedNet-R has been applied to sugar beet plantations (Guo et al., 2023), and other models have been deployed for pepper cultivation (Subeesh et al., 2022). Remote sensing is also used to monitor weeds in rice fields (Rosle et al., 2021).

However, weed management remains a major challenge for agricultural economics. Their detection, which is complex (Khan et al., 2016), is crucial as they compete directly with crops (Buhler, 2002). Rudimentary weed detection methods may lead to the misuse of herbicides (Abbas et al., 2018), contributing to the excessive application of toxic products to crops (Griffon, 1999).

In the face of these challenges, deep learning offers promising solutions. Image classification techniques based on binary cross-entropy (Ruby & Yendapalli, 2020) and multispectral methods like WeedNet (Sa et al., 2017) have been developed to improve weed detection. The contribution of neural network models in image processing, coupled with experimentation on different loss functions, continues to push the boundaries of this field (Zhao et al., 2015).

This article proposes an innovation in deep learning by extending the use of pre-trained models for weed detection classification. This classification task is particularly complex because, unlike other classification tasks involving distinct elements, there is a strong similarity between the elements to be classified.

Indeed, the question arises: Which convolutional neural network model should be used to recognize, classify, and predict weeds in a rice field?

To address this question, the article follows the structure below: the second section presents the materials and methods, followed by the results in the third section. The fourth section is dedicated to discussion, and the article concludes with a fifth section on the conclusion.

2 Materials and Methods

2.1 Materials

The R/RStudio software was used throughout the experimentation and modeling process, with the TensorFlow and Keras modules being the most utilized. Other modules, such as tidyverse, raster, and mapview, were also employed for data preparation and visualization.

For data collection, a Phantom 4 Pro drone was used to capture aerial images. Additionally, the entire process was carried out using a computer equipped with an AMD Ryzen 7 5800X 16-core processor with 96MB cache, 32GB of RAM at 3600 MHz, and an AMD Radeon RX 6700 graphics card with 10GB cache. The quality of the work and the speed of the experimental process depend heavily on the computing technology used, particularly the components listed above.

2.2 Méthods

2.2.1 Image processing

Image processing involves training, validating, and testing a convolutional neural network model using data in the form of images. The goal is to highlight the most discriminative features within these images. To achieve this, several pre-trained models were evaluated, including: DenseNet_201, EfficientNet_B0, InceptionResNet_V2, Inception_V3, MobileNetV3_large, MobileNet, ResNet_101, ResNet_50, VGG16, VGG19, and Xception. The optimal model was selected based on the loss value, accuracy, and convergence over a given number of iterations. Before training the model, preliminary steps are necessary, such as:

- Reading the image from a specified location.
- Decoding the image (JPEG or PNG) for transformation into a tensor object.
- Resizing the image to the desired dimensions.

The pre-trained models were trained individually to assess their performance. For this purpose, each model was compiled using an optimizer algorithm to update weights (gradient descent) while minimizing a specified loss function. A precision metric was then chosen to evaluate the model's performance during training and evaluation. Finally, a loss function was selected to measure the difference between the predicted probability distribution and the actual probability distribution (labeled data). After compilation, the adjustment phase involved minimizing the loss function on the training and evaluation data over a set number of iterations and batch sizes.

For visualization, the intermediate layers of the neural network were activated, which involved performing image predictions via tensors. In other words, the outputs from the input layers or images were obtained.

2.2.2 Classification

Convolutional neural network models are renowned for their ability to perform classification tasks due to their capacity to capture the spatial features of images.

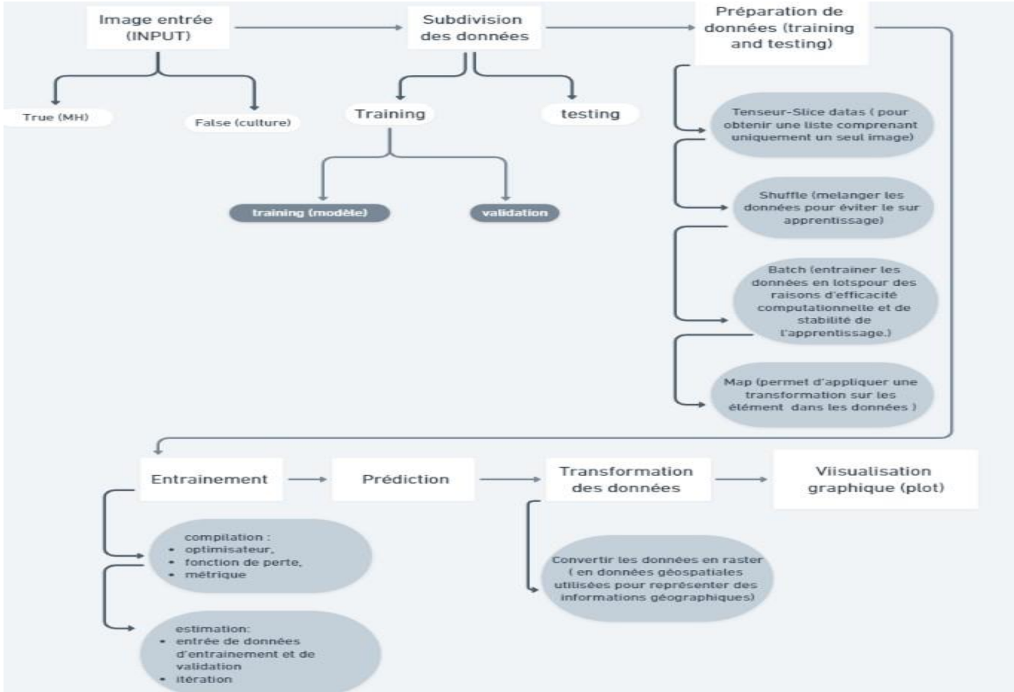


Figure 1 – Classification Procedure

The above figure outlines the sequence of steps necessary for achieving accurate classification. The first step involves structuring the input data by splitting the image into several small patches. These patches are then labeled as “1” if they represent weeds and “0” otherwise. The splitting process is carried out using a function or algorithm capable of dividing the input image into subsets. The resulting files are numbered and stored in two separate folders to ensure correct reading during the prediction phase.

The second step involves splitting the data into training and evaluation datasets. This is followed by data preparation, including converting images into tensors, randomization, batching, and ultimately transforming them into raster data. The third step is the training phase, followed by prediction in the fourth step, and ending with the transformation of training results for visualization.

3 Résults

3.1 Image processing

The figure below shows an image representing a rice field where both rice and non-rice elements can be observed. This is also a digital image (pixels) that is translated into numerical data to feed the convolutional neural network model. This serves as the input data provided to the model for making predictions or classifications.

Figure 2 – Image Captured by a Drone (Altitude: 30m)

After extracting the first activation layer from the training of pre-trained convolutional neural network models, the following results were obtained:

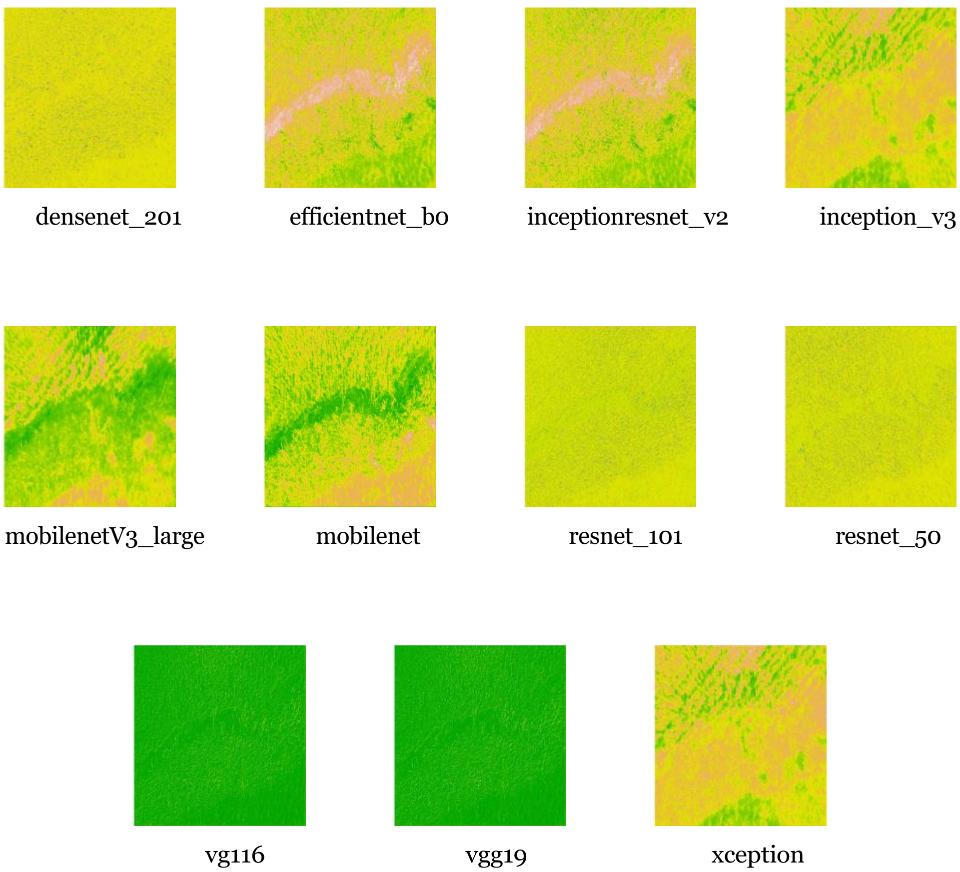


Figure 3 – Processing the Input Image with Pre-trained Neural Network Models

Three distinct colors corresponding to three different characteristics were observed in Figure 3. For the VGG16 and VGG19 models, only green was visible. For DenseNet_201, ResNet_101, and ResNet_50, only yellow was visible. For other models, the distinction of soil structure through color is apparent but not homogeneous. In other words, the color assignment to various soil elements differs across models.

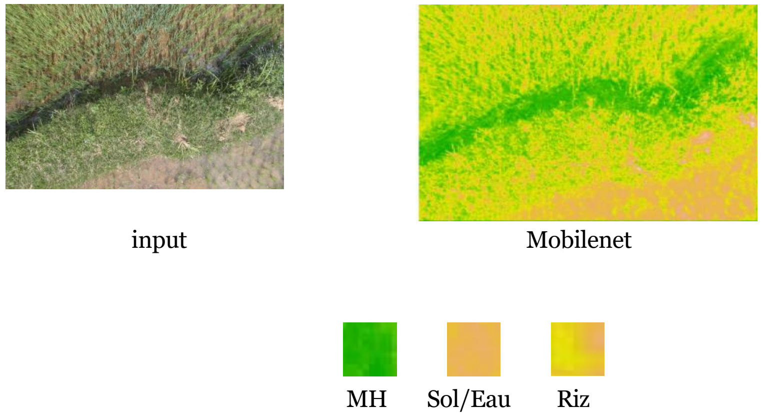


Figure 4 – Soil Occupancy Structure

Figure 4 shows that the soil occupancy structure is clearly identifiable using the pre-trained MobileNet model. Weeds, soil/water, and rice are represented by three distinct colors.

3.2 Classification

Figure 5 shows a sample of rice images obtained by subdividing the input image (Figure 2) after applying a specific function.

Figure 5 – Rice Sample from Input Image

In contrast, Figure 6 shows a sample image where no traces of rice are visible..

Figure 6 – Weed Sample from Input Image

In total, the input image was subdivided into 505 sub-images of weeds and 505 sub-images of rice, each measuring 128 × 128 pixels.

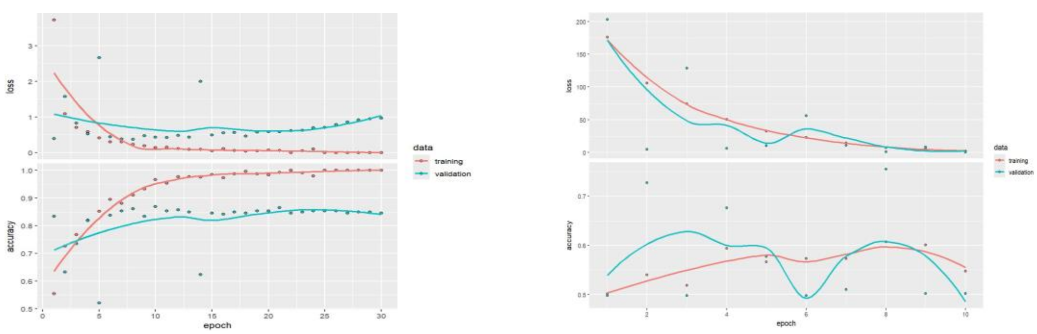
Table 1 – Comparison Between Architectures of Two Pre-Trained CNN Models

Inception_V3.1	MobileNet
One (01) input layer	One (01) input layer
Six (06) convolutional layers	Four (04) convolutional layers
Six (06) activation functions	Six (06) activation functions
Six (06) normalization operations	Six (06) normalization operations
Two (02) max-pooling layers	Two (02) zero-padding layers
One (01) flattening layer	Three (03) depthwise convolutional layers
Two (02) fully connected layers	One (01) flattening layer
Number of parameters: 2,769,152	Two (02) fully connected layers
	Number of parameters: 33,586,241

Source: Authors

The quality of image processing results from the Inception_V3 and MobileNet pre-trained models led to a comparison of their performance. An extension was implemented on the Inception_V3 model to lighten computations, prevent overfitting, and ensure the convergence of loss and accuracy values for both training and evaluation data.

The training results for the two models are shown below:



Inception_V3.1 Mobile_Net

Figure 7 – Training of a CNN Model with Extended Input Data

For the first model, a rising trajectory in accuracy and a declining trajectory in loss were observed for both training and evaluation data over 30 iterations. A slight convergence was noted, with a relatively stable and less volatile curve. Accuracy reached its maximum at the 15th iteration for training data, approaching a value of 1. For evaluation data, accuracy remained above 0.8 without exceeding 0.9.

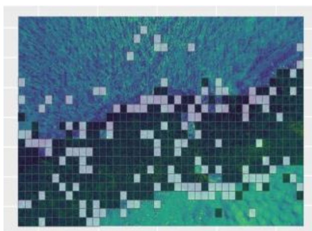
For the MobileNet model, the loss trajectory showed convergence between training and evaluation data. However, the accuracy curve exhibited high volatility and a downward trend after 30 iterations, indicating early overfitting. Moreover, accuracy on the training data never exceeded 0.6, deeming it very low.

Table 2 – Training Diagnostic

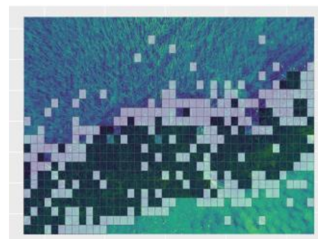
Indicator	Inception_V3		MobileNet	
	Final value	Average	Final value	Average
Loss (Training)	**0.0671**	2.0586	0.6914	9.8400
Accuracy (Training)	**0.9870**	0.9210	0.4907	0.5320
Loss (validation)	1.3500	2.7280	**0.6934**	8.1840
Accuracy (validation)	**0.7233**	0.7190	0.4980	0.5280

Source : Author's computation

The table above shows the final values and averages for each model concerning the loss and accuracy variables throughout the training and evaluation phases. Specifically, the "average" column indicates the average of each variable over 30 iterations, while the "final value" column corresponds to the value obtained at the 30th iteration.

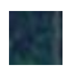


Inceptionv3



Mobilenet

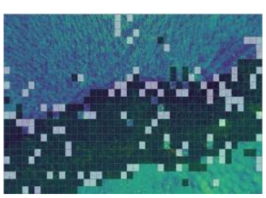
Prob ≤ 0.5



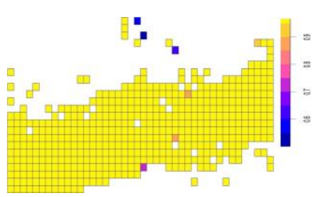
Pob > 0.5

Figure 8 – Overlay of Rice Field and Mosaic

Figure 8 illustrates the overlay of the input data (a rice field) and the output data (a distribution of tiles on the rice field). Two tiles of different colors are visible in the figure. This indicates that if the tile is gray, the probability of weed presence is less than 0.5. If it is transparent, the probability exceeds 0.5



(a : inception_V3.1)



(b)

(c)

Figure 9 – Rice fields and Mosaic

A transformation of the output data into raster format allowed us to convert grayscale data (see Figure (b)) into RGB data, on which the probability distribution was performed (see Figure (b)). Additionally, assembling the tiles into a black-and-white mosaic revealed information quite similar to that of the color mosaic. A white tile indicates the presence of weeds in the rice field with a probability approximately equal to 1 (Prob(Weeds) ≈ 1), whereas this probability is less than 0.5 if the tile is black (Prob(Weeds) < 0.5). If the tile is gray, the presence of weeds is estimated to be between 0.5 and 0.9 (0.5 ≤ Prob(Weeds) < 0.9).

4 Discussion

The image processing results using DenseNet_201, ResNet_101, ResNet_50, VGG16, and VGG19 models were inconclusive because the most specific features did not clearly emerge. Conversely, the performance of MobileNetV3_large, EfficientNet_B0, InceptionResNet_V2, Inception_V3, and Xception models was considered average. For these models, the soil occupancy data showed some confusion, and the images were not clear enough to enable precise weed distinction.

On the other hand, the MobileNet model succeeded in making weeds, soil, water, and rice visible. Regarding classification, however, MobileNet tended to underestimate the presence of weeds, assigning a probability below 0.5. Consequently, the model labeled certain areas with gray tiles where weeds were clearly present, resulting in a predominance of gray tiles in those zones (see Figure 8).

The precision rate estimated for MobileNet was 49%, compared to 83% for Inception_V3.1. Although MobileNet demonstrated satisfactory performance in image processing, its precision rate in classification disqualifies it as the optimal model.

Furthermore, the final performance of the Inception_V3.1 model surpassed MobileNet, despite MobileNet's lower loss value. On average, Inception_V3.1 provided satisfactory results.

In light of these findings, the Inception_V3.1 model meets the requirements for convergence and high precision, making it the optimal model for weed detection in a rice field, despite moderate effectiveness in image processing. The optimization of this model reduced computation time by limiting the number of parameters. The addition of fully connected layers endowed it with the ability to learn complex relationships between inputs and outputs, capturing sophisticated nonlinear dependencies within the data (Basha et al., 2020).

5 Conclusion

It is well established that deep learning is one of the most valuable fields of artificial intelligence for data science. Among revolutionary techniques, convolutional neural networks have transformed both image processing and classification. Technological advances and experimentation have led to the emergence of pre-trained models that have proven particularly effective in various competitions. To further enhance this efficiency, an extension was applied to the pre-trained Inception_V3 model to improve its precision.

Our experimentation demonstrated that while MobileNet achieved satisfactory performance in processing images of rice crop data, the modified Inception_V3.1 model outperformed others in classification tasks. It is also clear that data quality significantly influences the results obtained.

Our work produced conclusive results, leading us to conclude that the Inception_V3.1 model is optimal for weed detection in a rice field.

Références

Abbas, T., Zahir, Z. A., Naveed, M., & Kremer, R. J. (2018). Limitations of existing weed control practices necessitate development of alternative techniques based on biological approaches. Advances in Agronomy, 147, 239-280.

AgroTIC, C. (2018). Deep learning et Agriculture : comprendre le potentiel et les défis à relever.

Basha, S. S., Dubey, S. R., Pulabaigari, V., & Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. Neurocomputing, 378, 112-119.

Buhler, D. D. (2002). 50th Anniversary—Invited Article : Challenges and opportunities for integrated weed management. Weed Science, 50 (3), 273-280.

dos Santos Ferreira, A., Freitas, D. M., da Silva, G. G., Pistori, H., & Folhes, M. T. (2017). Weed detection in soybean crops using ConvNets. Computers and Electronics in Agriculture, 143, 314-324.

Griffon, M. (1999). Développement durable et agriculture : la révolution doublement verte.

Guo, Z., Goh, H. H., Li, X., Zhang, M., & Li, Y. (2023). WeedNet-R : a sugar beet field weed detection algorithm based on enhanced RetinaNet and context semantic fusion. Frontiers in Plant Science, 14.

Khan, M. A., Afridi, R. A., Hashim, S., Khattak, A. M., Ahmad, Z., Wahid, F., & Chauhan, B. S. (2016). Integrated effect of allelochemicals and herbicides on weed suppression and soil microbial activity in wheat (Triticum aestivum L.) Crop Protection, 90, 34-39.

Lantz, B. (2019). Machine Learning with R (S. Jain, Éd.). Packt publising Ltd.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture : A review. Sensors, 18 (8), 2674.

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv :1511.08458.

Rosle, R., Che'Ya, N. N., Ang, Y., Rahmat, F., Wayayok, A., Berahim, Z., Fazlil Ilahi, W. F., Ismail, M. R., & Omar, M. H. (2021). Weed detection in rice fields using remote sensing technique : A review. Applied sciences, 11 (22), 10701.

Ruby, U., & Yendapalli, V. (2020). Binary cross entropy with deep learning technique for image classification. Int. J. Adv. Trends Comput. Sci. Eng, 9 (10).

Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., & Siegwart, R. (2017). weednet : Dense semantic weed classification using multispectral images and mav for smart farming. IEEE robotics and automation letters, 3 (1), 588-595.

Subeesh, A., Bhole, S., Singh, K., Chandel, N., Rajwade, Y., Rao, K., Kumar, S., & Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. Artificial Intelligence in Agriculture, 6, 47-54. https://doi.org/https://doi.org/10.1016/j.aiia.2022.01.002

Tuffery, S. (2023). Deep learning : From big data to artificial intelligence with R : Deep learning for natural language processing.

Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. Sensors, 21 (11), 3647.

Zhao, H., Gallo, O., Frosio, I., & Kautz, J. (2015). Loss functions for neural networks for image processing. arXiv preprint arXiv :1511.08861.