Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 2, pp. 501–513
DOI : 10.47974/JSMS-1291

Exploring agricultural image data through deep learning for visual information analysis : A study on Cabbage (Brassica Oleracea var. Capitata) farming

Dibyanshu *
R. K. S. Rajput
Department of Mathematics, Statistics and Computer Science
G.B. Pant University of Agriculture and Technology
Pantnagar 263145
Uttarakhand
India

Abstract

This article outlines a methodology for developing and evaluating efficient image classification models using Convolutional Neural Networks (CNNs). It begins with meticulous network architecture design and training on a dataset comprising 9000 images of Cabbage, Weeds, and empty areas in cabbage plants, with the aim of achieving accurate image classification based on features and patterns. The paper conducts a detailed comparative analysis with established models such as AlexNet and ResNet, as well as modified versions of AlexNet and ResNet, employing performance metrics such as accuracy, precision, recall, and the F1 score. This comprehensive evaluation highlights the proficiency of the developed models and their relative effectiveness. Real-world datasets featuring fragmented agricultural plot images validate the models in practical scenarios, affirming their accuracy and reliability in agricultural data analysis. This article is helpful for the implementation of CNN in the automation of weeding in cabbage crops.

Subject Classification: *68T45 Machine vision and scene understanding.*

Keywords: *Convolutional neural network, ResNet, AlexNet, Weeds.*

Introduction

Agriculture is the backbone of Indian economy and trade, and rural communities rely on agriculture for their survival (Musale et al., 2020). Weeds are unwanted plants that grow alongside cultivated crops,

E-mail: **dibyanshubhatt2000@gmail.com**

disrupting agricultural ecosystems (Harlan and deWet, 1965). They compete with crops for essential resources such as water, nutrients, and sunlight, thereby reducing crop yield and quality, making weed detection and removal essential (Monteiro and Santos, 2022). Traditionally, farmers have employed various methods, such as manual weeding and chemical herbicides, to control weed populations (Gianessi, 2013).

In recent years, technological advancements, particularly in the fields of computer vision, machine learning, and robotics, have opened up new avenues for weed detection and control (Esteva et al., 2021).

Therefore, we propose a novel method for weed identification using AlexNet and ResNet Convolutional Neural Networks. In the ever-evolving landscape of agricultural technology, the need for efficient and sustainable weed management strategies has grown increasingly (Zhan et al., 2023). Weeds, which continually threaten crop yields and necessitate necessary control measures, have become a focal point of agricultural innovation (Neve et al., 2018). Our study addresses this challenge by harnessing the power of deep learning, specifically the formidable capabilities of AlexNet and ResNet, to revolutionize weed identification.

Weeds, often indistinguishable to the human eye amidst a sea of crops, pose a formidable obstacle to modern farming (Scott, 2017). As image data is quite large, and training a good AI network requires a large number of images (Shah and Mishra, 2020).

Data augmentation introduces variations to the input images, such as rotations, flips, and alterations in brightness and contrast, effectively exposing the neural networks to a broader spectrum of visual data (Wang, 2020).

Cabbage (Brassica olearaceae var. capitata) and Cauliflower (Brassica oleracea var. botrytis) are prominent vegetables cultivated and widely enjoyed in India.

Application of Deep Neural Network Techniques for Identification of Weeds: The application of Deep Neural Networks (DNNs) in agriculture holds significant potential to transform weed management practices. DNNs can play a crucial role in weed detection and classification in images and videos captured in the field, thereby facilitating precision agriculture by enabling farmers to differentiate between crops and weeds (Yadav et al., 2022). DNNs contribute to optimizing herbicide usage by accurately identifying weed species and their distribution, ensuring that herbicides are applied only where necessary (Patel et al., 2019).

The adoption of data-driven approaches, alongside the reduction in labor costs and minimized environmental impact, are among the manifold benefits that DNNs can bring to weed management in agriculture (Tantalaki et al., 2019).

Weeds compete with cultivated crops for essential resources such as sunlight, water, and nutrients, which can hinder crop growth and yield. Furthermore, weeds can act as hosts for pests and diseases (Kumar et al., 2021).

Weed control aids in disease and pest management, thereby reducing the risk of diseases affecting the cabbage crop (Cooper and Dobson, 2007).

Common Weeds Found in Cabbage Crops: Below are some weed species commonly found alongside cabbage plants:

1. Blackjack (Bidens pilosa): This annual weed produces small, sticky seeds that can cling to clothing and farm equipment, facilitating its spread (Image of plant shown in the Figure 1).
2. Love Grass (Setaria spp.): Love grass is a warm-season grass weed that can rapidly establish itself in fields and gardens, posing a threat to crops (Image of plant shown in the Figure 2).
3. Star Grass (Cynodon spp.): Star grass includes species like Bermuda grass, which can be invasive and challenging to eradicate due to its extensive root system (Image of plant shown in the Figure 3).
4. Wandering Jew (Commelina benghalensis): Wandering Jew, also known as wandering jewweed, is a low-growing, sprawling weed that can be invasive in gardens and agricultural areas (Image of plant shown in the Figure 4).
5. Oxalis (Oxalis spp.): Oxalis, commonly known as wood sorrel, is a weed with heart-shaped leaves that can be invasive in lawns and garden beds (Image of plant shown in the Figure 5).
6. Couch Grass (Digitaria spp.): Couch grass, also known as quackgrass, is a persistent and hard-to-control perennial grass weed (Image of plant shown in the Figure 6).
7. Mexican Marigold (Tagetes minuta): This weed, often called wild marigold, can invade crop fields and gardens and compete with desirable plants (Image of plant shown in the Figure 7).
8. Goose Grass (Eleusine indica): Goose grass is a common weed in lawns and agricultural areas, known for its fine, bristly seed heads (Image of plant shown in the Figure 8).
9. Pigweed (Amaranthus spp.): As mentioned previously, pigweed is a fast-growing weed that can outcompete crops for resources (Image of plant shown in the Figure 9).
10. Nutsedge (Cyperus spp.): Nutsedge is a persistent and challenging-to-control weed with triangular stems and tuberous roots (Image of plant shown in the Figure 10).

Figure 1
Blackjack Weed

Figure 2
Love Grass

Figure 3
Star Grass

Figure 4
Wandering Jew

Figure 5
Oxalis

Figure 6
Couch Grass

Figure 7
Mexican Marigold

Figure 8
Goose Grass

Figure 9
Pigweed

Figure 10
Nutsedge

Digital Images: Digital images are visual representations of objects captured and stored in a digital format, typically referred to as a grid of pixels. Unlike traditional analog images, digital images are encoded as numerical data, enabling precise manipulation, storage, and transmission (Doermann, 1998). Digital images play a pivotal role in fields like computer vision, where algorithms analyze and interpret visual data, and in medical imaging, aiding in diagnoses and treatments (Patil and Deore, 2013).

AUC-ROC Curves: As shown in Figure 11, the AUC-ROC curve serves as a performance metric in classification tasks across various threshold configurations (Sonego et al., 2008). It measures the model's capacity to correctly classify instances as either class 0 or class 1. (Elder et al., 2010). The ROC curve is generated by plotting the True Positive Rate (TPR) on the vertical axis and the False Positive Rate (FPR) on the horizontal axis.

Confusion Matrix: As shown in the Figure 12, a confusion matrix serves as a summary tool for evaluating the performance of a machine learning model when tested against a dataset (Reddy et al., 2020). True Positives (TP) correspond to instances where both the model's prediction and the actual value are 'Cabbage,' while True Negatives (TN) represent cases where both the prediction and actual value are 'Not Cabbage.' False Positives (FP) encompass instances where the prediction is 'Cabbage,' but the actual value is 'Not Cabbage,' and False Negatives (FN) entail cases where the prediction is 'Not Cabbage,' but the actual value is 'Cabbage.

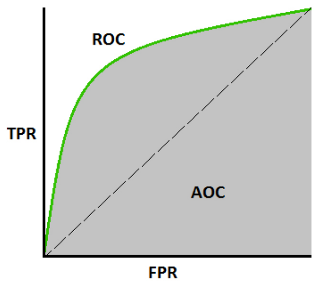

Figure 11

AUC- ROC Curve

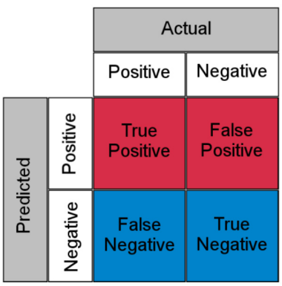

Figure 12

Confusion Matrix for classification task

Related Work

Yu et. al. (2019b) introduced the VCGNet model, a deep convolutional neural network (DCNN) that outperformed GoogleNet in detecting various weed species within Bermuda grass. Their study highlighted the potential of DCNNs for accurate weed detection, both in actively growing and dormant bermudagrass. Asad and Bais (2020) presented a methodology for accelerating the labeling of pixels for weed density estimation and variable-rate herbicide prescriptions. Their use of Semantic Segmentation and the SegNet model demonstrated promising results, with high intersection over the union values.

Ofori et. al. (2020) explored the performance of DCNNs in classifying crops and weeds at different growth stages, showing that training from scratch and fine-tuning pre-trained models can yield excellent results. Osorio et. al. (2020) introduced multiple deep learning methods for weed estimation in lettuce crops, with YOLOv3 and Mask R-CNN achieving high F1 scores, highlighting the effectiveness of deep learning in crop detection. Grace (2020) utilized deep learning-based computer vision techniques such as ConvNet or CNN to analyze images and identify weeds and crops in farm fields. They used the publicly available Plant Seedling Classification dataset from Kaggle for evaluation, achieving an accuracy of 89% with their proposed CNN algorithm.

Ruigork et. al. (2020) emphasized the importance of application-specific evaluation, conducting assessments at different levels to analyze weed detection and spraying decisions made by a robot. Wu et. al. (2020) introduced a multicamera system for weed control, incorporating advanced techniques for precise weed removal even with delays in detection, demonstrating the system's potential in real-field situations. Priyanka (2021) employed a ResNet50 pre-trained model for the prediction of cancerous cells, achieving accurate results with a high accuracy of 74.04%.

Methodology

This entails meticulously designing the network architecture, carefully selecting appropriate parameters, and conducting training using the collected dataset comprising 9000 images (sample images are shown in Figure 13, 14 and 15). The objective is to create models that effectively classify images based on their features and patterns.

Figure 13

Sample Image of Cabbage Dataset

Figure 14

Sample Image of Weed Dataset

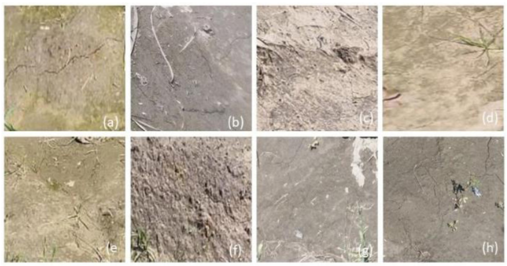

Figure 15

Sample Image of Empty Field Dataset

To analyze the performance of the developed models, a comparative analysis is carried out against existing models such as AlexNet and ResNet. The trained models undergo rigorous evaluation using various performance metrics, including accuracy, precision, recall, and F1 score. This comprehensive evaluation allows for an assessment of the models' effectiveness in accurately classifying images. Comparisons with established models like AlexNet and ResNet provide insights into the relative performance of the developed models.

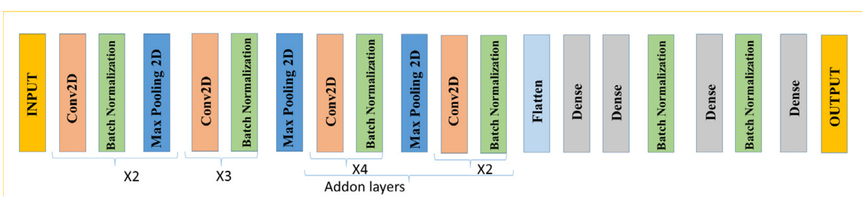

Figure 16

Architecture of Improved AlexNet

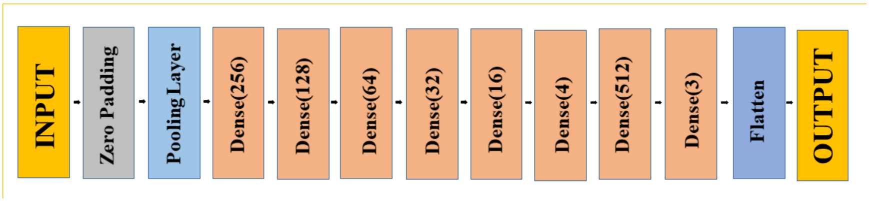

Figure 17

Architecture of Improved ResNet

Real datasets consisting of fragmented images of an agricultural plot are employed to validate the models' performance in real-world scenarios. These fragmented images, captured from the agricultural plot, serve as representative samples for validating the accuracy and reliability of the models.

Results

An experiment was conducted on Google Colab with GPU support, and the data were stored in Google Drive. The study involved two pre-trained architectures, AlexNet and ResNet, as well as two improved architectures that underwent serial adjustments of nodes (as shown in Figure 16 and 17). The performance of all models was evaluated based on classification accuracy, precision, recall, and F1-score. Table 1 displays the classification report containing all the metric performances computed by both pre-trained and improved models. It is observed from Table 1 that the accuracy of the improved models is significantly better.

Table 1

Results for Pretrained and Improved models

Model	Accuracy	Precision	Recall	F1-Score
ResNet Pretrained	0.78	0.78	0.80	0.76
ResNet Improved	0.87	0.87	0.87	0.86
AlexNet Pretrained	0.91	0.92	0.91	0.91
AlexNet Improved	0.91	0.92	0.91	0.91

The Figures 18-21 depict an upward trend in accuracy and a downward trend in loss. The confusion matrices of the models are displayed in Figures 22-25, illustrating the relationships among cabbage, weed, and empty fields. All models were executed for 50 epochs.

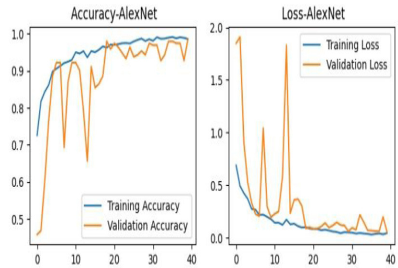

Figure 18

Training and validation curve for accuracy and loss for AlexNet

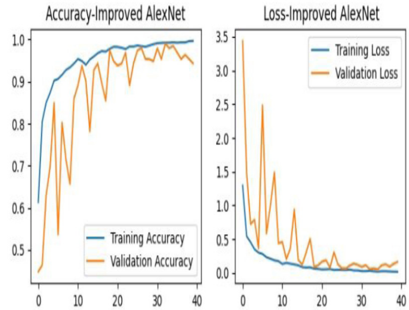

Figure 19

Training and validation curve for accuracy and loss for improved AlexNet

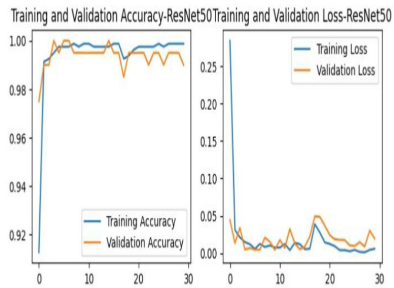

Figure 20

Training and validation curve for accuracy and loss for ResNet

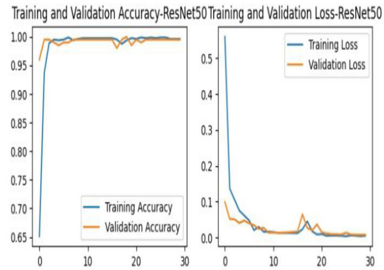

Figure 21

Training and validation curve for accuracy and loss for improved ResNet

Conclusion

The utilization of deep learning models, particularly AlexNet and ResNet, for weed identification in image detection has demonstrated promising outcomes. These models were trained on a substantial dataset comprising weed images, enabling them to assimilate relevant features and effectively classify the presence of weeds. Our evaluation metrics, including accuracy, precision, recall, and F1-score, indicate commendable performance from both models in weed detection.

Despite AlexNet's shallow architecture, it exhibited satisfactory performance in weed detection, as similar results were observed in both the improved and pretrained models. Conversely, the ResNet model

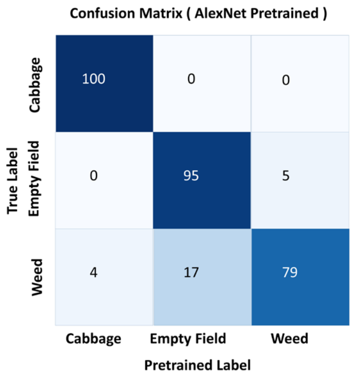

Figure 22

Confusion Matrix pretrained AlexNet

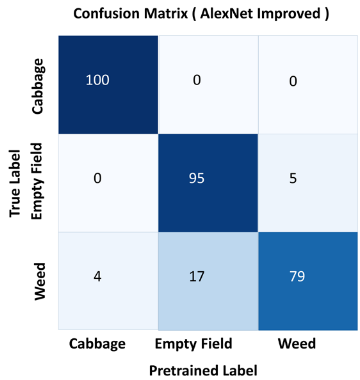

Figure 23

Confusion Matrix Improved AlexNet

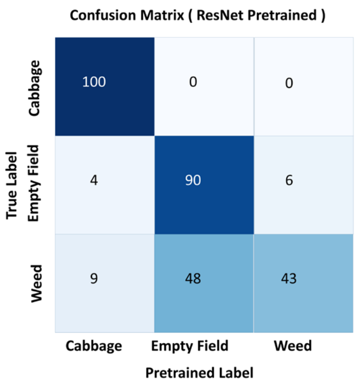

Figure 24

Confusion Matrix pretrained ResNet

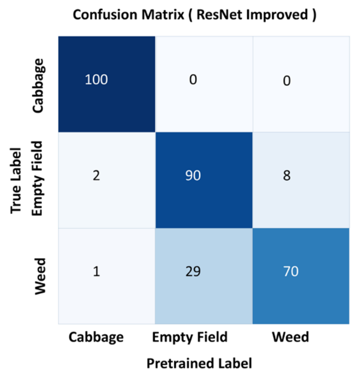

Figure 25

Confusion Matrix Improved ResNet

demonstrated lower accuracy compared to AlexNet, but the improved version exhibited superior accuracy over the pretrained model.

Future Work

The study in this article suggests some ideas for future research. First, we could look into transfer learning to make weed detection better. This means using models that already learned from lots of pictures to help spot weeds. We might not need to train them from scratch, which can save time.

Tweaking these models a bit to fit with specific types of weeds could also help them work better.

References

[1] Asad, M. H. and Bais, A. Weed Detection in Canola Fields using maximum likelihood Classification and deep convolutional neural network. Inf. Process. Agric., 7(4): 535–545 (2020).

[2] Cooper, J., & Dobson, H. The benefits of pesticides to mankind and the environment. Crop Protection, 26(9), 1337-1348 (2007).

[3] Doermann, D. The indexing and retrieval of document images: A survey. Computer Vision and Image Understanding, 70(3), 287-298 (1998).

[4] Eder, L., Chandran, V., Shen, H., Cook, R. J., & Gladman, D. D. Is ASDAS better than BASDAI as a measure of disease activity in axial psoriatic arthritis?. Annals of the rheumatic diseases, 69(12), 2160-2164 (2010).

[5] Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., & Socher, R. Deep learning-enabled medical computer vision. NPJ digital medicine, 4(1): 5 (2021).

[6] Gianessi, L. P. The increasing importance of herbicides in worldwide crop production. Pest management science, 69(10), 1099-1105 (2013).

[7] Grace, R. K. Crop and weed classification using deep learning. *Turkish Journal of Computer and Mathematics Education*, 12(7): 935-938 (2021).

[8] Harlan, J. R., & deWet, J. M. Some thoughts about weeds. *Economic botany*, 19(1), 16-24 (1965).

[9] Monteiro, A., & Santos, S. Sustainable approach to weed management: The role of precision weed management. Agronomy, 12(1): 118 (2022).

[10] Musale, S., Khedkar, A., Shriniwar, M., & Gadekar, T. Evaluating performance of different neural network models on crop and weed detection and classification. *Turkish Journal of Computer and Mathematics Education*, 11(2), 1199-1204 (2020).

[11] Neve, P., Barney, J. N., Buckley, Y., Cousens, R. D., Graham, S., Jordan, N. R. & Williams, M. Reviewing research priorities in weed ecology, evolution and management: a horizon scan. Weed Research, 58(4): 250-258 (2018).

[12] Ofori, M. and El-Gayar, O. F., Towards Deep Learning for weed detection: Deep Convolutional Neural Network Architectures for plant seedling Classification. 'In Americas Conference on Information Systems (AMCIS-2020)' at Salt Lake City, Utahduring, during. August 2020, 10-14. pp. 1-10 (2020).

[13] Osorio, K., Puerto, A., Pedraza, C., Jamaica, D. and Rodríguez, L. A Deep Learning approach for Weed Detection in lettuce crops using Multispectral Images. Agri. Engineering, 2(3): 471-488 (2020).

[14] Patel, M., Jernigan, S., Richardson, R., Ferguson, S., & Buckner, G. Autonomous robotics for identification and management of invasive aquatic plant species. *Applied Sciences,* 9(12), 2410 (2019).

[15] Patil, D. D., & Deore, S. G. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing*, 2(1), 22-27 (2013).

[16] Priyanka, B. J. Machine learning approach for prediction of cervical cancer. *Turkish Journal of Computer and Mathematics Education,* 12(8): 3050-3058 (2021).

[17] Reddy, G. T., Bhattacharya, S., Ramakrishnan, S. S., Chowdhary, C. L., Hakak, S., Kaluri, R., & Reddy, M. P. K. An ensemble-based machine learning model for diabetic retinopathy classification. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (IC-ETITE) (pp. 1-6). IEEE (2020, February).

[18] Ruigrok, T., Van Henten, E., Booij, R., Van Boheemen, K. and Kootstra, G. Application-Specific Evaluation of a Weed-Detection Algorithm for Plant-Specific Spraying. Sensors, 20(24): 1-20 (2020).

[19] Scott, J. C. Against the grain: A deep history of the earliest states. Yale University Press (2017).

[20] Shah, S. K., Mishra, R., Mishra, B. S. P., & Pandey, O. Prediction of abnormal hepatic region using ROI thresholding based segmentation and deep learning based classification. *International Journal of Computer Applications in Technology*, 64(4): 382-392 (2020).

[21] Sonego, P., Kocsor, A., & Pongor, S. ROC analysis: applications to the classification of biological sequences and 3D structures. *Briefings in bioinformatics*, 9(3), 198-209 (2008).

[22] Tantalaki, N., Souravlas, S., & Roumeliotis, M. Data-driven decision making in precision agriculture: The rise of big data in agricultural systems. *Journal of Agricultural & Food Information*, 20(4), 344-380 (2019).

[23] Wu, X., Aravecchia, S., Lottes, P., Stachniss, C. and Pradalier, C. Robotic Weed Control using Automated Weed and Crop Classification. *J. Field Robot*, 37(2): 322–340 (2020).

[24] Yadav, S., Kaushik, A., Sharma, M., & Sharma, S. Disruptive technologies in smart farming: an expanded view with sentiment analysis. Agri Engineering, 4(2), 424-460 (2022).

[25] Yu, J., Schumann, A. W., Cao, Z., Sharpe, S. M. and Boyd, N. S. B. Weed detection in perennial ryegrass with Deep Learning Convolutional Neural Network. Front. Plant Sci., 10: 1422 (2019).

[26] Zhan, J., Thrall, P. H., & Burdon, J. J. Achieving sustainable plant disease management through evolutionary principles. *Trends in Plant Science*, 19(9), 570-575 (2014).