Assessing the Impact of Including Diseased Crop Images in Training Datasets on the Performance of Convolutional Neural Networks for Crop Classification

QUINCY-ALEXANDRE LELASSEUX, University of Twente, The Netherlands

This study investigates the impact of including diseased crop images in training datasets on the performance of ResNet18, a convolutional neural network (CNN), in crop classification tasks. The focus lies on assessing model performance in scenarios where both healthy and diseased crop images are present in the dataset. Under these conditions, the model can adapt and classify crops more effectively by fine-tuning the CNN using transfer learning. Datasets containing varying proportions of diseased crop images (e.g., 0%, 10%, 30%, and 50%) were systematically analyzed to evaluate the effects of these variations on classification accuracy, precision, and F1-score. Experiments included training with both mixed and separated datasets as baselines for discussion. To ensure robust and reliable results, the training and evaluation process incorporated k-fold cross-validation. In the results of this study it was found that the absence of diseased samples during training significantly reduces the model's ability to generalize to real-world conditions, whereas incorporating such images enhances robustness and accuracy.

Additional Key Words and Phrases: ResNet18, Convolutional neural network (CNN), crop classification, k-fold cross validation, transfer learning

1 INTRODUCTION

Computer vision and machine learning (CVML) techniques are playing an increasingly vital role in addressing global challenges in agriculture, particularly in meeting the rising demand for food production [5, 11]. Precision agriculture is one way of addressing such challenges, by ensuring plants or animals get precisely the treatment they need, which can be achieved with great accuracy thanks to the latest technology. The need for precision agriculture and efficient crop production has driven the application of machine learning (ML) and deep learning (DL) methods in crop classification. These techniques aim to optimize agricultural output by maximizing crop yield while minimizing resource utilization [13].

Crop classification involves categorizing images into predefined groups based on attributes such as shape, texture, and color. Convolutional neural networks (CNNs), such as ResNet18, have been widely used in this domain to automate tasks including production prediction, resource allocation, and disease management, thereby improving agricultural efficiency. However, the reliability of these models is heavily influenced by the quality of the training data.

A challenge with real-world agricultural datasets is often the presence of irrelevant features such as soil, weed and non crop-objects. Also, diseased crop leaves frequently appear in such datasets, introducing another layer of complexity by obscuring features critical for accurate classification. The presence of sick leaves may cause models trained on healthy datasets to misclassify crops, thereby affecting their performance in real-world scenarios.

This research focuses specifically on the inclusion of diseased crop leaves, which adds complexity to classification tasks. Understanding how diseased crop images impact model performance is crucial for designing robust crop classification systems. By evaluating the effect of varying proportions of diseased crop images on crop classification tasks, this study aims to improve the efficiency of deep learning models for the agricultural industry.

This study investigates the classification of a variety of crops, encompassing multiple types of crop leaves available in public datasets. Representative samples of the datasets used in this research are illustrated in Figure 1. By introducing varying proportions of diseased crop images into the training and testing datasets, this research seeks to:

- Quantify the effect on a model's performance when the datasets include varying proportions of diseased crop images and are tested in different conditions.
- Understand the relationship between training datasets with diseased crop images and model robustness.

Unlike prior studies that focus primarily on the detection of diseases in crops [17], this work instead shifts the emphasis to evaluating how the inclusion of diseased crop images affects the broader task of crop classification. The findings of this research will provide valuable guidelines for the development and deployment of deep learning models in agricultural environments.

The remainder of this paper is organized as follows: the related works section reviews existing research on crop classification and the effects of diseased crop images on model performance. The problem statement and research questions define the research objectives, while the datasets section outlines the datasets used and the proportions of diseased crop images introduced. The methodology section details the experimental setup, including ResNet18, dataset preparation, and k-fold cross-validation. It is then followed by the results, which analyze the impact of diseased crop images on model performance through metrics and visualizations.

2 RELATED WORKS

Image processing technology has become increasingly used in agricultural sensing applications such as disease detection and crop growth monitoring [1].

The impact that the inclusion of diseased crop images has on machine learning models performing these tasks has been studied

TScIT 42, January 15, 2025, Enschede, The Netherlands

across various domains. [16] is an example of one such study, where the research focused on the effect of mislabeled data in the training data set. Another example of similar research is [18], which explores how such factors affect regression tasks, focusing on linear and polynomial regression models. These studies highlight the detrimental effects of including irrelevant or mislabeled data in training datasets and emphasize the need for strategies to improve robustness. While this work provides foundational insights, it primarily addresses synthetic issues in numerical data and regression tasks, leaving open questions about how this transfers into the context of image-based classification tasks.

The study [7] categorizes issues like these into attribute and class challenges, highlighting their adverse effects and strategies for mitigation, such as data cleaning and developing robust algorithms. While this review offers broad insights into how to identify and deal with data quality problems in datasets, this research specifically examines the impact of diseased crop images on the performance of ResNet18 in crop classification.

Similarly, study [21] dives more deeply into the differences between attribute issues and class problems and their impact on classification performance .

By systematically introducing varying proportions of diseased crop images (e.g., 0%, 10%, 30%, and 50%) into training and testing datasets, this research in this paper examines how convolutional neural networks (CNNs), specifically ResNet18, respond to datasets with differing qualities of training data.

Furthermore, while studies such as Mohammad et al. [14] and Majumdar et al. [12] have addressed related challenges in disease recognition and specific crop classifications, they lack focus on the broader implications of including diseased crop images in datasets.

In the aspect of datasets, the dataset [3] from Harvard Dataverse offers valuable crop classification images. However, since its crops were already well-represented in this study's dataset, it was excluded to avoid redundancy and ensure the class inbalances were not too high.

This study complements the existing literature by expanding the understanding of how diseased crop images affect model performance, extending from simpler tasks such as regression and segmentation to complex image-based classification tasks.

3 PROBLEM STATEMENT AND RESEARCH QUESTION

While CNNs like ResNet18 have proven to be effective in general image classification tasks in an agricultural context [10], their performance can be hindered by challenges that are found in real-world agricultural implementations. One of the biggest problems is the inclusion of irrelevant data in the real world, such as soil, weeds, or non-crop objects, which can make it harder to detect relevant crop features. The inclusion of diseased crop images introduces irrelevant or challenging features that reduce the accuracy of the classifier.

In this study, the focus is on the inclusion of diseased crop images which may distort the visual features of healthy leaves, making classification more challenging.

3.1 Research Question

To address the identified problem, this study focuses on answering the following research question:

"How does the performance of crop recognition models vary when trained on datasets with and without diseased crop samples?"

4 TOOLS AND DATASETS

4.1 Tools

This research made use of a variety of tools and frameworks to facilitate data preprocessing, model training, evaluation, and visualization. The tools used in this study include:

4.1.1 Development Environment and Hardware. The local Jupyter server of the University of Twente was used to create and save multiple python notebooks that were used in the experimenting and evaluating phases of this research. Additionally, GPU-accelerated computing was utilized to deal with the computationally intensive tasks such as the fine-tuning of the CNN.

4.1.2 Programming Language and Libraries. Python was the primary programming language used for all experiments and analyses. This choice was justified by the extensive amount of machine learning libraries and frameworks that are available and easily accessible. The deep learning framework PyTorch was used for implementing and fine-tuning the ResNet18 convolutional neural network (CNN). Smaller libraries such as Scikit-learn and Matplotlib were utilized for the k-fold cross-validation, data splitting, evaluation calculation, and visualization.

4.2 Multi-Crop Dataset Preparation

4.2.1 Introduction to Dataset Preparation. The dataset for this study was created by merging several publicly available datasets containing images of crop leaves. These datasets included both healthy and diseased samples to simulate real-world agricultural scenarios.

4.2.2 Dataset Details and Class Distribution. Table 1 lists the datasets used and merged along with their characteristics. The class distribution and image samples are provided in Figure 1. The dataset is considered unbalanced, as certain classes contain significantly more samples than others. This imbalance was taken into account during the interpretation of the results.

4.2.3 Images Presenting Diseases. Diseased images were relabeled as healthy in certain experiments to analyze performance under different training conditions. The separation of healthy and diseased images was maintained during initial dataset preparation to facilitate such analyses.

4.2.4 Dataset Splitting and Setup. The dataset was split into training and testing subsets using a 90/10 ratio to ensure reliable evaluation. This split was consistent across all experimental conditions, enabling controlled comparisons between different setups. Diseased images were strategically introduced into training and testing datasets to simulate realistic conditions to be evaluated.

Dataset	Healthy	Diseased	Total
Cassava [6]	336	203	537
Multiclass [2]	2529	1580	4109
Plantdoc [19]	755	1118	1873
Plantvillage [8]	15084	25028	40112

Table 1. Table of the datasets that have been used and merged.

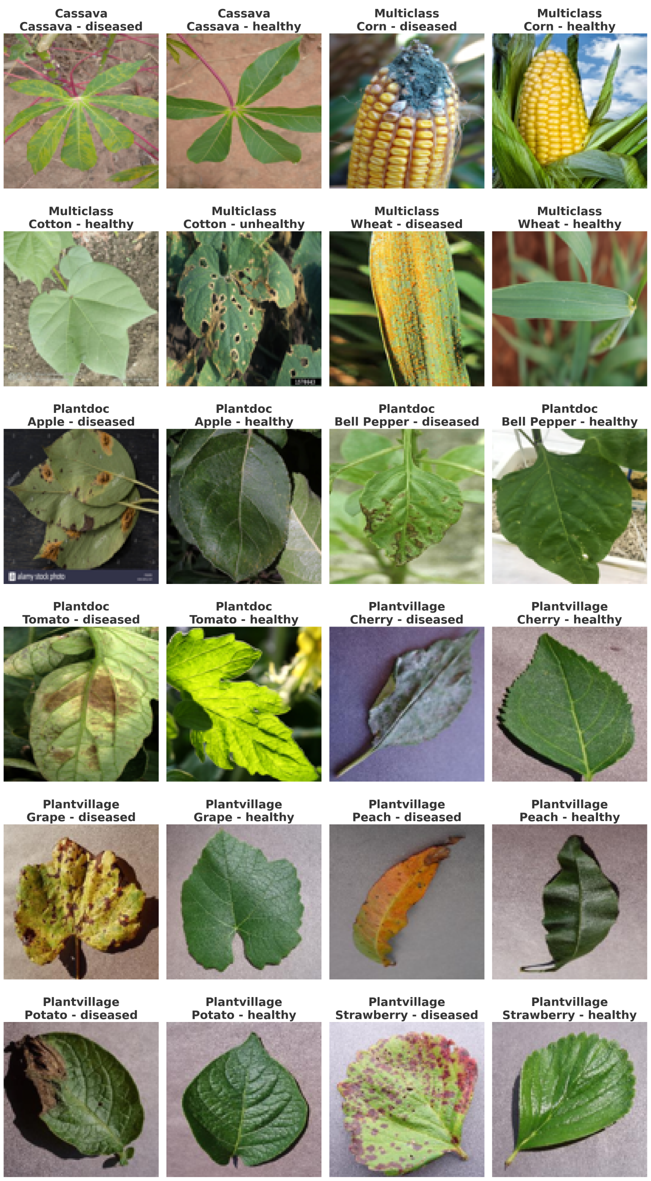


Fig. 1. Class distribution and examples of all classes in the datasets.

5 METHODOLOGY

This section reveals the systematic approach that was followed to investigate the impact of diseased images on the performance of ResNet18 in crop classification tasks. The methodology consists of several stages including model selection and adaptation, introduction of diseased crop images, training and evaluation, and performance analysis. Each stage is described in detail below.

5.1 Model Selection and Adaptation

The ResNet18 convolutional neural network (CNN) was chosen as the model for this research due to its proven effectiveness in image classification tasks and its relative smaller size [9]. To adapt the model for the specific task of crop classification, transfer learning was employed by fine-tuning the pre-trained ResNet18 model on the custom dataset.

5.2 Training and Evaluation

The training phase employed transfer learning to fine-tune the pre-trained ResNet18 model for the specific crop classification task. During this phase, only the final fully connected layer of the network was modified to accommodate the number of crop classes in the dataset, while the earlier layers, responsible for feature extraction, retained the knowledge from ImageNet pretraining. The following steps were performed and repeated for each of the desired proportions of diseased crop images in the training data (0%, 10%, 30%, 50%).

5.2.1 k-fold cross-validation. Training was conducted using k-fold cross-validation, specifically with $k = 5$, to ensure robust evaluation and mitigate the impact of data variability. Each fold involved splitting the dataset into training and validation subsets, allowing the model to be trained on $k - 1$ folds and to be validated on the remaining fold. This approach ensured that all samples of the classes contributed to both training and validation, providing a comprehensive evaluation of the model's performance. Each fold resulted in a model that was trained on a certain distribution of the training dataset. The highest performing model from each fold was saved and stored in the notebook for future evaluation. This process resulted in 5 models per proportions of diseased crop image level.

5.2.2 Epoch and Early stopping. An epoch refers to a complete pass through the entire training dataset during the training process [4]. In deep learning, multiple epochs are typically required for a model to learn the patterns in the data effectively. Each epoch involves forward and backward passes, where the model's weights are adjusted based on the computed gradients to minimize the loss function.

In this study, the training process was designed with a maximum of 50 epochs. However, early stopping was implemented as a regularization technique to prevent overfitting and optimize training time. Early stopping monitors the validation accuracy during training and halts the process if no improvement is observed for three consecutive epochs. This approach ensures that the model does not overfit to the training data and remains generalizable to unseen samples. Early stopping also enhances computational efficiency by avoiding unnecessary epochs when the model has already reached its optimal performance.

By combining the use of epochs with early stopping, the training process balanced sufficient learning with computational efficiency and model generalizability.

5.2.3 Evaluation Phase. After training, each model was tested on datasets containing proportions of diseased crop images of 0%, 10%, 30%, and 50%, covering all possible combinations of training and testing conditions. This comprehensive evaluation strategy allowed for the analysis of how the inclusion of diseased crop images in both training and testing datasets influenced model performance.

5.3 Performance Analysis

Performance in this study is evaluated using the metrics accuracy and F1-score, to provide a comprehensive analysis of the model's ability to classify crops under varying proportions of diseased crop images. Since the dataset is unbalanced, the F1-score serves as the primary metric for interpreting the results.

In addition to the standard metrics, confusion matrices were generated for the model that was closest to the average performance among the five models trained for each proportion of diseased crop images. These matrices highlighted which classes were frequently misclassified, providing valuable insights into the impact of diseased crop images on the model's ability to distinguish between crop classes.

The analysis was conducted at both individual model levels and aggregated levels. The performance of the five models trained on each proportion of diseased crop images was averaged, and standard deviations were calculated to assess the stability and robustness of the training process. Testing on datasets with varying proportions of diseased crop images further allowed for a detailed examination on how well models generalized to different conditions, revealing thresholds at which performance significantly deteriorated.

6 RESULTS AND INTERPRETATION

The results presented in this section aim to evaluate the impact of varying proportions of diseased crop images on the performance of ResNet18 in the task of crop classification.

The analysis is structured to first present the outcomes of the training and validation phases, followed by an evaluation of test performance. Trends in the metrics across proportions of diseased crop images are explored to identify critical thresholds and provide insights into the robustness of ResNet18 under these conditions. These results are then interpreted and used to answer the research question.

6.1 Training and validation results

Table 2 summarizes the average accuracy of the models across all folds on the validation set. Figure 3 summarizes the trends and validation accuracy for the models across different proportions of diseased leaves.

6.1.1 Interpretation of training and validation results. From Table 2 we can conclude that accuracy decreases very slightly as noise increases, from 94.44% at 0% diseased images to 92.85% at 30%, but rebounds to 93.78% at 50%, demonstrating robustness to noise. The F1-score remains stable (0.91–0.93), indicating consistent balance between precision and recall despite the increasing noise. Overall, the model generalizes well across varying noise levels.

% of Diseased Images	Accuracy (%)	F1-score
0%	94.44 ± 1.03	0.91 ± 0.06
10%	93.30 ± 0.86	0.93 ± 0.01
30%	92.85 ± 1.11	0.92 ± 0.01
50%	93.78 ± 0.35	0.93 ± 0.01

Table 2. Average accuracy and F1-score of the models across all folds with varying diseased images in validation datasets.

6.2 Testing results

The F1-score and accuracy graphs illustrate model performance across datasets with varying proportions of diseased images. The F1-score graph highlights the balance between precision and recall, while the accuracy graphs in the appendix show overall correctness. Together, they reveal how training data composition impacts model robustness and generalization.

6.2.1 Interpretation of testing results. Figure 2 shows that as the percentage of diseased images increases, F1-scores generally decrease, indicating reduced precision-recall balance. Models trained on higher proportions of diseased images (e.g., 50%, 30%) exhibit more stable performance with narrower error bars, while those trained on lower proportions (e.g., 0%) show greater variability. This suggests that training on higher diseased proportions improves robustness to noisy datasets.

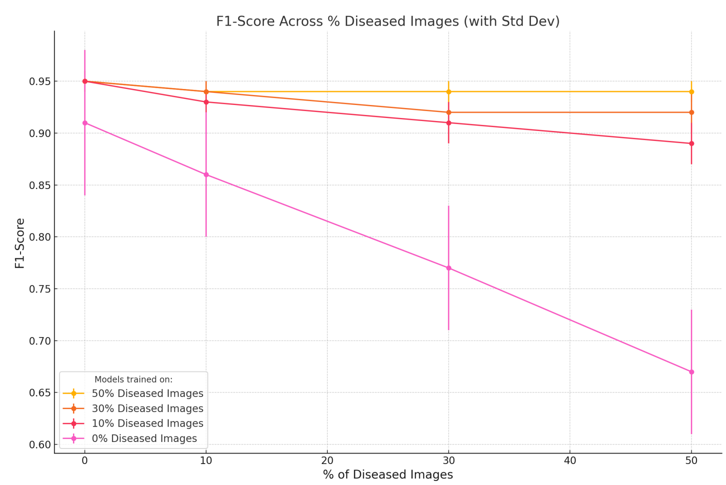
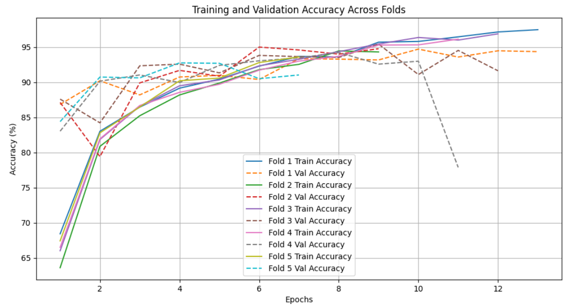
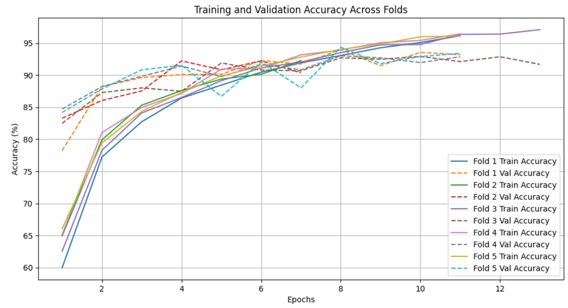


Fig. 2. Average F1-score of the models that have been trained on different ratios of diseased images.



(a) 0% Diseased images

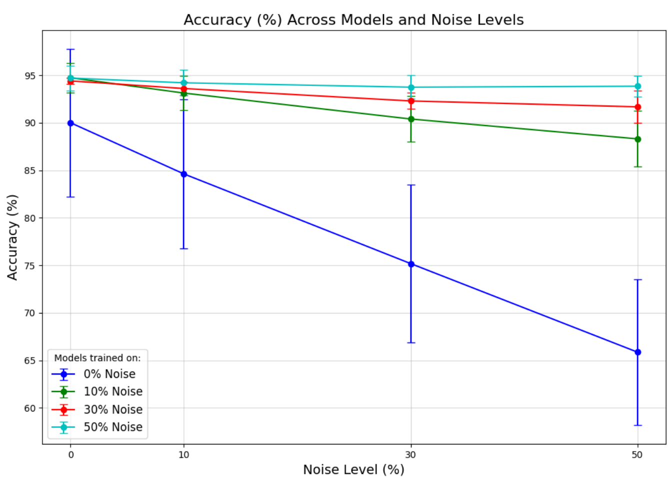


(b) 10% Diseased images

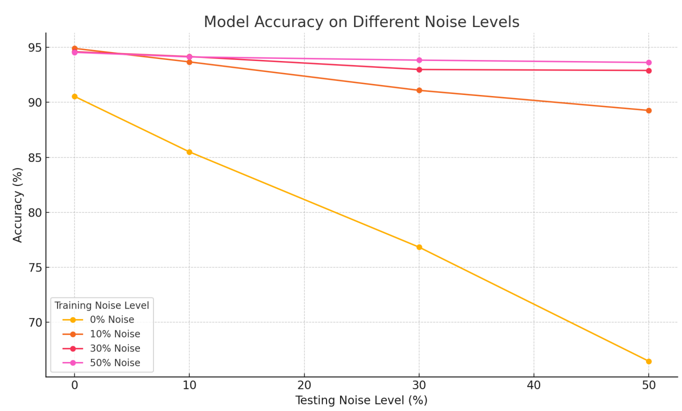
(c) 30% Diseased images

(d) 50% Diseased images

Fig. 3. Training and validation accuracy for each fold and proportion of diseased images. The same trends can be found in every graph; training accuracy, represented by solid lines, steadily increases and plateaus near 95%, indicating that the model effectively learns from the training data. Validation accuracy, shown as dashed lines, also improves but exhibits more variability across folds, likely due to differences in the validation subsets. While there is a small gap between training and validation accuracy in later epochs, suggesting mild overfitting, the overall trend shows that the model generalizes well. The performance plateaus after several epochs, signaling convergence and highlighting the model's robustness despite potential diseased images in the dataset.



(a) Average testing accuracy of the models.



(b) Testing accuracy of the model closest to the average of the 5 models per noise level.

Fig. 4. Testing accuracy of individual and combined models.

6.2.2 Correlation between accuracy and F1-score. Figures 4a and 4b depict the accuracy of the models across different levels of diseased images. Despite the imbalance in the dataset, the accuracy and F1-score graphs show similar trends. Both metrics generally decrease as the proportion of diseased images in the dataset increases. Models trained on datasets with higher levels of diseased images (e.g., 30% and 50%) are more stable, maintaining consistent performance even as the proportion of diseased images increases. On the other hand, models trained on datasets with lower proportions of diseased images (e.g., 0%) show a sharper drop in both accuracy and F1-score, with more variation as indicated by wider error bars.

These similar trends suggest that accuracy and F1-score are both affected by the training data's composition and noise levels, even though they measure performance differently. While accuracy focuses on overall correctness, F1-score highlights the balance between precision and recall. Considering the dataset's imbalance, one might expect F1-score to behave differently, but the similarities show that both metrics respond to the same factors. This supports the idea that training on datasets with higher ratio of diseased proportions improves the model's ability to handle noisy or imbalanced data.

6.2.3 Analysis of classification errors. The confusion matrix in appendix B 5d demonstrates the performance of the ResNet18 model tested on a dataset with 50% diseased images, despite being trained only on healthy samples (0% diseased images), making this its lowest-performing mode. The model struggles to generalize to diseased samples, resulting in poor accuracy across multiple classes. For instance, cherry achieves only 46.2% accuracy, with frequent misclassifications as apple (14.8%) and peach (11.0%), likely due to diseased features resembling those of other crops. Similarly, corn is often misclassified as cotton (16.5%), while peach shows severe errors, with 25% of its samples labeled as apple. These misclassifications highlight the model's inability to handle the altered features of diseased crops, underscoring the importance of including diseased images during training to improve resilience to the presence of diseased images.

7 DISCUSSION

7.1 Conclusion of the research question

To answer the proposed research question, the experiment was conducted and the results were interpreted. The inclusion of diseased crop samples in the training dataset significantly improves the performance of crop recognition models. Models trained with diseased samples demonstrate greater robustness and accuracy, effectively handling the variability introduced by diseased features, whereas those trained without such samples struggle to generalize, leading to reduced classification accuracy and increased misclassifications.

7.2 Relation to prior research

The results align with prior research indicating that the dataset the model is trained on has a significant impact on the testing results [20]. Similarly, another study [15] found that models trained on controlled datasets showed significant accuracy drops when tested on images from different conditions, emphasizing the importance of diverse training data. These studies align with our findings, showing that including diseased crop images in the training dataset has big influences on a model's performance in classification.

7.3 Contextualization

These results are important for real-world agricultural challenges, especially in precision agriculture. Models that do not account for the distortions caused by diseased crops may perform poorly, leading to misclassifications in identifying crops. This can make it harder to monitor crops and manage diseases effectively. Including diseased images in training datasets helps create more reliable systems that work well in different conditions.

7.4 Limitations of this research

Despite these contributions, this study has limitations. The analysis was restricted to the ResNet18 architecture, and it remains unclear whether more complex models, such as ResNet50 or DenseNet, could achieve better generalization to noisy data. Furthermore, the datasets used, while representative, may not fully capture the diversity of diseased crop features present in real-world scenarios, potentially limiting the generalizability of the results.

7.5 Future research

Future research could address these limitations by exploring additional architectures and employing data augmentation strategies to simulate varying levels of disease. Including more samples from real-life scenarios can also improve the effectiveness of a model's training phase. Techniques such as segmentation to isolate diseased regions may help models focus on relevant features, further enhancing classification performance. Additionally, evaluating the impact of pre-training on diverse datasets or employing transfer learning from related agricultural tasks could provide valuable insights into improving model robustness. By addressing these areas, future work can build on the findings of this study to develop more effective and reliable crop recognition systems for the agricultural industry.

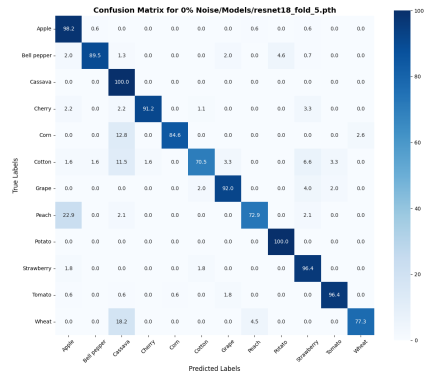
8 CONCLUSION

In conclusion, this study investigated the effect of including diseased crop samples in the training dataset on the performance of crop recognition models. It was found that the absence of diseased samples during training significantly reduces the model's ability to generalize to real-world conditions, whereas incorporating such images enhances robustness and accuracy. These findings highlight the importance of curating diverse and representative datasets for agricultural applications, particularly in precision farming and automated crop monitoring. Future improvements, such as using advanced architectures, data augmentation, and segmentation techniques, could further enhance model performance and support the development of reliable, real-world-ready crop recognition systems.

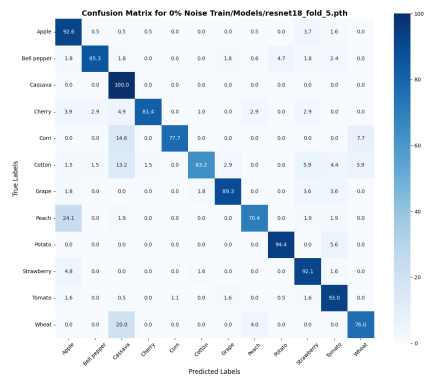
A APPENDIX A: USE OF AI TOOLS

This research utilized ChatGPT to generate example code and paraphrase certain text segments. AI assistance was employed throughout the paper to enhance coherence and readability while preserving the originality and credibility of the ideas. All content was subsequently reviewed and edited by Quincy Lelasseux, who assumes full responsibility for the final work.

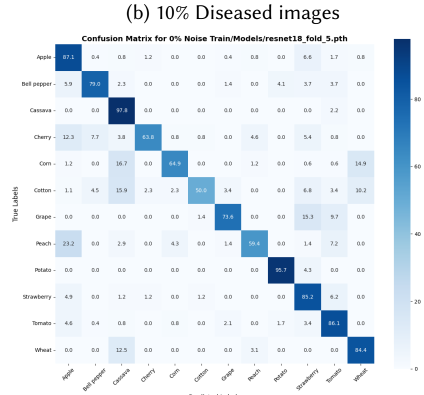
B APPENDIX B: CONFUSION MATRICES



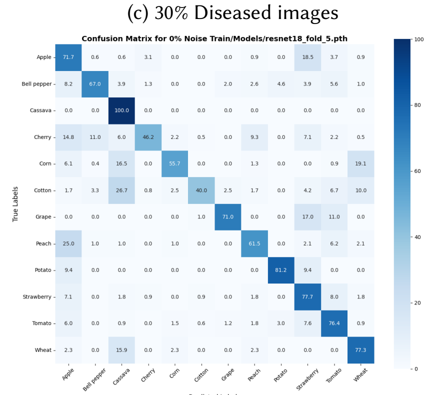
(a) 0% Diseased images



(b) 10% Diseased images



(c) 30% Diseased images



(d) 50% Diseased images

Fig. 5. Confusion matrices for the model closest to the average trained on 0% diseased images, tested on several ratios of diseased images.

REFERENCES

- [1] Alwaseela Abdalla, Haiyan Cen, Liang Wan, Reem Rashid, Haiyong Weng, Weijun Zhou, and Yong He. 2019. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. *Computers and electronics in agriculture* 167 (2019), 105091.
- [2] Jawad Ali. 2020. 20k+ Multi-Class Crop Disease Images. https://www.kaggle.com/datasets/abdallahalidev/20k-crop-disease-dataset Accessed: 2025-01-15.
- [3] Mutebi Chodrine Nakatumba-Nabende Joyce Namanya Gloria Bright Hodasi Babirye, Claire. 2022. Crop Image and Spectral Datasets for Machine Learning. https://dataverse.harvard.edu/dataverse/airlabug.
- [4] Jason Brownlee. 2018. What is the Difference Between a Batch and an Epoch in a Neural Network. *Machine learning mastery* 20 (2018), 1–5.
- [5] Dharani Dhar Burra, Jack Hildebrand, James Giles, Thuy Nguyen, Eva Hasiner, Kateryna Schroeder, David Treguer, Armine Juergenliemk, Alexandra Horst, Andy Jarvis, et al. 2021. Digital Agriculture Profile: Viet Nam. (2021).
- [6] L. G. Divyanth, Peeyush Soni, and Rajendra Machavaram. 2021. Cassava Leaf Disease Dataset. https://doi.org/10.17632/3832tx2cb2.1
- [7] Shivani Gupta and Atul Gupta. 2019. Dealing with noise problem in machine learning data-sets: A systematic review. *Procedia Computer Science* 161 (2019), 466–474.
- [8] David Hughes, Marcel Salathé, et al. 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060* (2015).
- [9] Riaz Ullah Khan, Xiaosong Zhang, Rajesh Kumar, and Emelia Opoku Aboagye. 2018. Evaluating the performance of resnet model based on image recognition. In *Proceedings of the 2018 international conference on computing and artificial intelligence*. 86–90.
- [10] Sachin Kumar, Saurabh Pal, Vijendra Pratap Singh, and Priya Jaiswal. 2023. Performance evaluation of ResNet model for classification of tomato plant disease. *Epidemiologic Methods* 12, 1 (2023), 20210044. https://doi.org/doi:10.1515/em-2021-0044
- [11] Konstantinos G Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. 2018. Machine learning in agriculture: A review. *Sensors* 18, 8 (2018), 2674.
- [12] Samir Majumdar. 1997. Classification of cereal grains using machine vision. (1997).
- [13] Efthimia Mavridou, Eleni Vrochidou, George A Papakostas, Theodore Pachidis, and Vassilis G Kaburlasos. 2019. Machine vision systems in precision agriculture for crop farming. *Journal of Imaging* 5, 12 (2019), 89.
- [14] Shorif Uddin Mohammad and Jagdish Chand Bansal. 2021. Computer vision and machine learning in agriculture. (2021).
- [15] Sharada P Mohanty, David P Hughes, and Marcel Salathé. 2016. Using deep learning for image-based plant disease detection. *Frontiers in plant science* 7 (2016), 1419.
- [16] Charlotte Pelletier, Silvia Valero, Jordi Inglada, Nicolas Champion, Claire Marais Sicre, and Gérard Dedieu. 2017. Effect of training class label noise on classification performances for land cover mapping with satellite image time series. *Remote Sensing* 9, 2 (2017), 173.
- [17] Ramanjot, Usha Mittal, Ankita Wadhawan, Jimmy Singla, NZ Jhanjhi, Rania M Ghoniem, Sayan Kumar Ray, and Abdelzahir Abdelmaboud. 2023. Plant disease detection and classification: A systematic literature review. *Sensors* 23, 10 (2023), 4769.
- [18] Arun Thundyill Saseendran, Lovish Setia, Viren Chhabria, Debrup Chakraborty, and Aneek Barman Roy. 2019. Impact of noise in dataset on machine learning algorithms. *Mach. Learn. Res* 1 (2019), 1–8.
- [19] Davinder Singh, Naman Jain, Pranjali Jain, Pratik Kayal, Sudhakar Kumawat, and Nipun Batra. 2020. PlantDoc: A dataset for visual plant disease detection. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*. 249–253.
- [20] Yuan Yuan, Lei Chen, Yuchen Ren, Shimei Wang, and Yin Li. 2022. Impact of dataset on the study of crop disease image recognition. *International Journal of Agricultural and Biological Engineering* 15, 5 (2022), 181–186.
- [21] Xingquan Zhu and Xindong Wu. 2004. Class noise vs. attribute noise: A quantitative study. *Artificial intelligence review* 22 (2004), 177–210.