Integrating multi-source data and image identification for automated crop disease diagnosis using deep learning

Brian Sudi Wafula

A Thesis Submitted in Partial Fulfilment

of the requirements for the

Degree of

Master of Science in Data Analytics

Diagram

Description automatically generated with medium confidence

September 2023

Supervisor: Dr. David McQuaid

Assessment Cover Page

To be provided separately as a word doc for students to include with every submission

|  |  |
| --- | --- |
| Module Title: | CAPSTONE PROJECT |
| Assessment Title: | Integrating multi-source data and image identification for automated crop disease diagnosis using deep learning |
| Supervisor Name: | David McQuaid |
| Student Full Name: | Brian Sudi Wafula |
| Student Number: | 2022550 |
| Assessment Due Date: | 24th September 2023 |
| Date of Submission: | 24th September 2023 |

Declaration

By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

**ABSTRACT**

*This data-driven project addresses critical challenges in automated crop disease diagnosis by integrating multi-source data and leveraging advanced deep learning techniques. Building upon prior research (Jin et al., 2022; Meshram et al., 2021), the project employs state-of-the-art methods to enhance the efficiency and precision of crop disease identification in agriculture. While traditional image processing techniques have been the norm, this project takes a significant step forward by implementing convolutional neural networks (CNNs), offering more accurate and robust disease classification based on images (Lecun et al., 2015). The proposed system aims to deliver timely and precise information to farmers, ultimately aiding in the reduction of crop losses and promoting sustainable farming practices by minimizing pesticide use.*

*Furthermore, the project integrates machine learning to predict changes and optimize farming practices (Liakos et al., 2018). By incorporating deep learning algorithms, it strives to overcome the limitations imposed by environmental factors affecting agriculture. Inspired by previous studies (Ji et al., 2023; Vetal & R.S., 2017), the project seeks to boost productivity, reduce costs, and safeguard the environment through accurate and timely plant disease identification. The utilization of CNNs in image recognition tasks proves advantageous for processing image data and handling spatial complexities, thereby enhancing the efficiency and precision of crop disease diagnosis (Sakib et al., 2018). The project explores the semantic understanding of deep neural networks and the existence of adversarial examples, contributing to the development of robust and interpretable models for crop disease diagnosis (Chowdhury et al., 2022; Guo et al., 2017; O’Shea & Nash, 2015). The project's ultimate aim is to create a scalable and accurate automated crop disease diagnosis method by integrating multi-source data and image identification techniques using deep learning. The outcomes of this endeavor hold immense potential for the agricultural sector, providing farmers with an automated and reliable tool for crop disease diagnosis and treatment recommendations. Ultimately, this initiative seeks to enhance farming practices, increase yields, and advance the agricultural industry as a whole.*

*In conclusion, the project has successfully achieved its objectives by developing an accurate automated system for crop disease diagnosis. Among the models tested CNN emerged as the top performer with an impressive accuracy rate of 99%. Additionally, the CNN-Naive Bayes ensemble approach significantly enhanced accuracy, reaching 93%, while the Gaussian Naive Bayes classifier achieved a lower accuracy of 57%. This underscores the effectiveness of deep learning CNN models, particularly when combined with ensemble strategies, in* ***addressing the challenge of crop disease diagnosis.***

Table of Contents

[**1.** **INTRODUCTION** 1](#_Toc146462449)

[**2.** **OBJECTIVES** 2](#_Toc146462450)

[2.1. Develop a deep learning model using multi-source data for accurate and efficient crop disease diagnosis. 2](#_Toc146462451)

[2.2. Develop an ensemble of machine learning models. 2](#_Toc146462452)

[2.3. Explore different deep-learning paradigms to enhance the performance of crop disease diagnosis models. 2](#_Toc146462453)

[**3.** **PRIMARY RESEARCH METHODOLOGY** 3](#_Toc146462454)

[**4.** **SAMPLING STRATEGY** 4](#_Toc146462455)

[**5.** **VALIDITY MANAGEMENT** 4](#_Toc146462456)

[**6.** **ETHICAL CONSIDERATIONS** 5](#_Toc146462457)

[**7.** **RELATED WORKS** 6](#_Toc146462458)

[7.1. Early Research on Crop Disease Detection 6](#_Toc146462459)

[7.2. Introduction of Deep Learning for Crop Disease Diagnosis 8](#_Toc146462460)

[7.3. Machine Learning in Agriculture and Supply Chain 10](#_Toc146462461)

[7.4. Advances in Hyperspectral Technology 13](#_Toc146462462)

[7.5. Application of Machine Learning and IoT in Precision Agriculture 14](#_Toc146462463)

[7.6. Deep Learning and CNNs in Image Recognition 15](#_Toc146462464)

[7.7. Bitwise Neural Networks for Computational Efficiency 19](#_Toc146462465)

[7.8. Confidence Calibration in Classification Models 20](#_Toc146462466)

[7.9. Properties and Challenges of Deep Neural Netwoks 21](#_Toc146462467)

[7.10. Ensemble Methods with Naïve Bayes Classifier on classification 22](#_Toc146462468)

[7.11. BPNN/ANN Architectures 23](#_Toc146462469)

[7.13. Related works conclusion 23](#_Toc146462470)

[**8.** **DESIGN AND METHODOLOGY** 25](#_Toc146462471)

[8.1. Data Acquisition 25](#_Toc146462472)

[8.2. Data Preprocessing 27](#_Toc146462473)

[**9.** **IMPLEMENTATION** 29](#_Toc146462474)

[9.1. BPNN/ANN ARCHITECTURES 29](#_Toc146462475)

[9.2. CNN ARCHITECTURES 31](#_Toc146462476)

[9.3. Gaussian Naïve Bayes Classifier 33](#_Toc146462477)

[9.4. CNN-Naive Bayes ensemble/Gaussian Naïve Bayes Classifier 35](#_Toc146462478)

[9.5. Performance Evaluation Metrics 36](#_Toc146462479)

[**RESULTS AND DISCUSSION** 38](#_Toc146462480)

[**CONCLUSION AND FUTURE RESEARCH** 38](#_Toc146462481)

[**REFERENCES** 39](#_Toc146462482)

[**APPENDIX A** 43](#_Toc146462483)

[A. Mango Leaf Dataset links (FIRST DATA SET) 43](#_Toc146462484)

[B. A Database of Leaf images (SECOND DATA SET) 43](#_Toc146462485)

[**APPENDIX B: IN-DEPTH INTERVIEW QUESTIONS** 44](#_Toc146462486)

[**APPENDIX C: IN-DEPTH INTERVIEW TRANSCRIPTION** 46](#_Toc146462487)

[1. FIRST INDEPTH INTERVIEW 46](#_Toc146462488)

[2. SECOND INDEPTH INTERVIEW 47](#_Toc146462489)

[3. THIRD INDEPTH INTERVIEW 48](#_Toc146462490)

# **INTRODUCTION**

This data project aims to address automated crop disease diagnosis challenges by integrating multi-source data and image identification using deep learning techniques. The proposed project builds upon previous research in the field and incorporates various advancements to improve the efficiency and accuracy of crop disease diagnosis in agriculture(Jin et al., 2022; Meshram et al., 2021). Traditionally, early image processing techniques have been utilized for crop disease diagnosis. This project takes a step further by employing deep learning algorithms such as convolutional neural networks (CNNs), to achieve more precise and robust classification of plant diseases based on images(Lecun et al., 2015). The proposed system can provide farmers with timely and accurate information on crop diseases. This proactive approach will helps reduce crop losses and increase productivity while promoting sustainable and environmentally friendly farming practices by minimizing the use of pesticides.

The project incorporates machine learning to enhance the prediction of changes and optimize farming practices(Liakos et al., 2018). By integrating deep learning algorithms, the project aims to improve crop disease diagnosis and overcome the limitations imposed by environmental factors that impact agriculture. This research will build upon previous studies that have explored the use of machine learning and image-processing methods for crop disease diagnosis(Ji et al., 2023; Vetal & R.S., 2017). The project can improve productivity, reduce costs, and protect the environment by being able to accurately and timely identify plant diseases. Using CNNs in image recognition tasks offers advantages in processing image data and handling spatial dimensionality, thus enhancing the accuracy and efficiency of crop disease diagnosis(Sakib et al., 2018).

Research on the semantic meaning of individual units in deep neural networks and the existence of adversarial examples can also inform the project, contributing to the development of more robust and interpretable models for crop disease diagnosis(Chowdhury et al., 2022; Guo et al., 2017; O’Shea & Nash, 2015). This project aims to develop a scalable and accurate method for automated crop disease diagnosis by integrating multi-source data and image identification methods utilizing deep learning techniques. The project results will have a significant impact on the agricultural sector. Farmers will be provided with an automated and reliable technique for crop disease diagnosis and treatment recommendations. Ultimately, this project aims to improve farming practices, increase yields, and contribute to the overall advancement of the agricultural industry.

# **OBJECTIVES**

**Problem Definition:** Development of a scalable and accurate system for automated crop disease diagnosis that integrate machine learning and image and identification techniques.

## Develop a deep learning model using multi-source data for accurate and efficient crop disease diagnosis.

The project will utilize deep learning algorithms, such as convolutional neural networks (CNNs)(O’Shea & Nash, 2015). By training these models the aim is to achieve precise and robust classification of plant diseases. This approach can provide farmers with timely and accurate treatment recommendations, reducing crop losses and promoting sustainable farming practices by minimizing pesticide usage.

* 1. Develop an ensemble of machine learning models. Building upon the deep learning models, this objective aims to enhance the overall accuracy and robustness of the crop disease diagnosis system by creating an ensemble of machine learning models. This ensemble will combine the predictions from various models, including convolutional neural networks (CNNs) and a traditional machine learning algorithm such as Naïve Bayes Classifier to effectively reduce the risk of overfitting and improve the system's accuracy.

## Explore different deep-learning paradigms to enhance the performance of crop disease diagnosis models.

Deep learning has shown great potential in image identification tasks, and this objective seeks to leverage various deep learning paradigms further to improve the performance of crop disease diagnosis models. The objective of investigating different approaches is to enhance the prediction of changes and optimize farming practices. This research will contribute to precision agriculture, enabling better decision-making and promoting increased productivity while minimizing costs and environmental impact.

# **PRIMARY RESEARCH METHODOLOGY**

The choice of qualitative research methods, specifically in-depth interviews, served as the most suitable primary research methodology for the project. In-depth interviews in relation to qualitative approach allowed for open-ended questions and encouraged participants to elaborate on their responses, providing rich and detailed information. Through them, the research established a connection with the participants, fostering a comfortable and conducive environment for sharing their knowledge.

The in-depth interviews enabled the project to validate and verify the data collected during secondary research. The interviews gave a chance to cross-reference and compare the perspectives of different experts, thereby strengthening the validity of the project's findings. Through follow-up questions, clarifications, and probing to elicit comprehensive responses the project uncovered unique insights. In conclusion, qualitative research methods, specifically in-depth interviews, was the ideal primary research methodology for the automated crop disease diagnosis project. By conducting one-on-one conversations with experts, the project gained deep insights and experiences, validated secondary research data, and established a network of knowledgeable individuals. The personalized and conversational nature of in-depth interviews promoted open sharing of information, which ultimately contributed to the development of a reliable, accurate, and relevant system for crop disease diagnosis.

# **SAMPLING STRATEGY**

The choice of a sampling strategy played a crucial role in the project, as it determined how representative the collected data would be of the target population. The selection of a non-probability sampling strategy, specifically snowballing and judgment sampling was the most suitable approach. Snowballing, as a non-probability sampling method allowed for identifying and including such key informants through referrals or recommendations. By leveraging existing connections and networks, this strategy ensured that knowledgeable individuals who might not be easily accessible were included in the sample. These experts provided valuable insights, research findings, and practical experiences, contributing to the overall validity and reliability of the project.

Moreover, judgment sampling, another non-probability technique, complemented the snowballing method. In this approach, the research used judgment sampling to select participants based on their expertise, relevance, and ability to contribute to the project. Given the specialized nature of crop disease diagnosis, relying on expert judgment to identify individuals with deep domain knowledge was essential. By intentionally selecting participants with valuable insights and experiences, the project captured diverse perspectives and ensured the inclusion of the most knowledgeable individuals in the sample. This sampling strategy not only facilitated access to crucial information but also allowed for establishing a network of experts who could validate and verify the collected data. By engaging with these experts, the project enhanced the credibility and verifiability of the information gathered during secondary research. Their expertise ensured that the data collected from reputable sources could be utilized confidently in building a reliable and accurate deep-learning model for crop disease diagnosis.

In conclusion, choosing non-probability sampling, specifically snowballing and judgment sampling was the most suitable strategy for this project. By leveraging referrals, recommendations, and expert judgment, this approach ensured the inclusion of knowledgeable individuals who may not be easily accessible through random sampling. The insights and experiences provided by these experts contributed to the project's validity, reliability, and relevance, enabling the development of a robust and effective system for crop disease diagnosis.

# **VALIDITY MANAGEMENT**

To ensure the reliability of the project, it was important to consider the credibility and verifiability of the information collected during secondary research. The multi-source data and image identification techniques used for crop disease diagnosis had to originate from reliable sources.

Accurate information was vital for the success of the deep learning model in diagnosing crop diseases through image identification. The accuracy of the data used for training the model, such as the images, played a significant role in achieving precise and robust classification. By ensuring the accuracy of the data sources, the models could learn from high-quality information and make accurate predictions. This accuracy would enable farmers to receive timely and reliable information on crop diagnosis, ultimately reducing crop losses and promoting sustainable farming practices by minimizing the need for excessive pesticide usage.

Relevance was another key aspect of the project's validity management. The choice of utilizing diverse data sources, was highly relevant to crop disease diagnosis. By incorporating these different data types, the models could capture a comprehensive understanding of the factors influencing crop health. Relevance allowed the model to consider various aspects of the different diseases and make informed predictions. Consequently, the model's recommendations would be more practical and suitable for real-world agricultural scenarios. This aligned with the goal of precision agriculture, which aims to optimize farming practices and minimize costs and environmental impact.

In summary, reliability was ensured by relying on credible and validated sources of information. Accuracy was achieved through the use of accurate data for training the deep learning model, leading to precise crop disease diagnosis. Relevance was addressed by incorporating diverse data sources that capture the key factors affecting crop health. By managing these three validity points effectively, the project was able to build a reliable, accurate, and relevant system for crop disease diagnosis, empowering farmers with timely and accurate recommendations for sustainable farming practices.

# **ETHICAL CONSIDERATIONS**

Ethical considerations are vital in any data analytics project, and the proposed automated crop disease diagnosis project is no exception. By addressing these considerations in an appropriate manner, the project ensured the responsible and ethical use of data, protected individuals' privacy, and contributed to the advancement of the agricultural industry. One key ethical consideration revolved around data privacy and security. As the project integrated multi-source data, it was crucial to handle this information with utmost care and respect for privacy. Strategies to address this consideration included, anonymizing personally identifiable information, and establishing secure storage and transmission protocols. By safeguarding the data, the project protected the privacy and confidentiality of the individuals involved while maintaining data integrity. Another important ethical consideration was the potential bias in the deep learning models used for crop disease diagnosis. Biases could arise from imbalanced training data, underrepresented classes, or discriminatory patterns in the data itself. To address this, the project implemented strategies such as careful data selection, data augmentation techniques, and continuous monitoring of model performance across different demographic groups. Regular auditing and testing helped identify and mitigate any biases, ensuring fair and equitable outcomes.

Transparency and interpretability were additional ethical considerations in the deep learning models. To address these, the project strived to develop models that provided clear explanations for their predictions. Techniques like model interpretability algorithms and saliency maps aided in understanding the decision-making process of the deep learning models. Transparent communication of the model's limitations and potential uncertainties was also crucial, ensuring that farmers had a realistic understanding of the system's capabilities and potential errors. Lastly, the responsible use of the deep learning model and its recommendations was emphasized. While the model could provide valuable insights and treatment recommendations, it was to be able to maintain human expertise and decision-making.

In conclusion, the automated crop disease diagnosis project exploring deep learning prioritized ethical considerations throughout its development and implementation. By addressing data privacy, bias mitigation, transparency, community engagement, and responsible use of the model's recommendations, the project ensured the ethical use of data, protect individuals' privacy, and contributed to the sustainable and equitable advancement of the agricultural industry.

# **RELATED WORKS**

## 7.1. Early Research on Crop Disease Detection

This project aims to be relevant to the current state of agriculture and the need for efficient and accurate crop disease diagnosis. Adhiparasakthi Engineering College et al. (n.d.) addressed the critical issue of plant disease detection and classification in agriculture through the use of image processing techniques. Their objective was to develop a software system that could automatically identify and classify diseases in plants. The authors emphasized the increasing importance of agriculture beyond its role in feeding growing populations, especially in Asian countries where a significant portion of the population relied on this sector. The main challenge they aimed to tackle was the impact of diseases on crop quality and agricultural productivity. Since they mentioned that many disease symptoms are microscopic and beyond human visual capabilities, the study highlighted the need for a system that could autonomously detect, classify, and quantify disease symptoms.

One notable aspect of the study by Adhiparasakthi Engineering College et al. (n.d.) was the focus on using leaf images for disease detection. They mentioned the choice was significant because many diseases manifest on plant leaves, fruits, or stems. The authors employed image processing techniques, including image loading, pre-processing, segmentation, feature extraction, and classification, to detect and classify the diseases efficiently. They employed MATLAB image processing methods that captured both healthy and unhealthy leaf images, and then applied techniques like k-means clustering and Random Forest Classifier for training and classification. The project addressed the detection of various diseases, including Alternaria Alternata, Bacterial Blight, Anthracnose, and Cercospora Leaf Spot. By implementing a multi-step process involving image acquisition, pre-processing, segmentation, feature extraction, and classification, the study provided guided automated solutions for disease identification. The authors also expressed their intention to expand their research to encompass more disease types, demonstrating a commitment to further advancing the field of automated disease detection in agriculture(Adhiparasakthi Engineering College et al., n.d.-a).

Meshram et al. (2021) conducted a comprehensive survey of the application of machine learning (ML) in the agriculture domain, focusing on addressing challenges in the pre-harvesting, harvesting, and post-harvesting stages of farming. According to the study, agriculture, being a crucial sector of the global economy and a primary source of employment, faced significant challenges such as knowledge gaps, inefficiencies in crop management, and substantial losses throughout the farming process. The authors highlighted those advancements in ML, along with technologies like blockchain, the Internet of Things (IoT), and deep learning, could provide valuable solutions to enhance agricultural practices. The intent of the survey was to explore how ML could empower farmers by providing insights, recommendations, and data-driven decisions at every stage of farming. The features utilized in ML applications were expected to include computer vision, object detection, yield estimation, and quality assessment to optimize crop selection, land preparation, irrigation, harvesting, and post-harvesting activities. The paper underlined the significant benefits of ML in agriculture and offered recommendations, such as creating and sharing datasets, focusing on specific problem-solving, and prioritizing model deployment in real-world applications(Meshram et al., 2021).

## 7.2. Introduction of Deep Learning for Crop Disease Diagnosis

Study by Unay et al. (n.d.) presented a novel application of machine vision for the grading of bi-colored apple fruits. They aimed to develop an automated grading system that could accurately classify apples into quality categories based on the presence of defects in the fruit skin. To achieve this, they employed multispectral imaging, capturing images of 'Jonagold' apples under diffused illumination using a high-resolution black and white camera equipped with four interference band-pass filters. The authors addressed the challenging issue of defect segmentation, which required distinguishing defects from stem and calyx areas that exhibit similar spectral characteristics. They proposed a two-step segmentation process involving a Multi-Layer Perceptron (MLP)-based method followed by a Support Vector Machine (SVM)-based refinement. In their methodology, Unay et al. (n.d.) extracted a range of statistical, textural, and geometric features from the segmented defect areas of the apple images. These features included statistical attributes such as mean, standard deviation, median, minimum, and maximum values, which provided insights into the distribution of pixel values. Textural features encompassed invariant moments of Hu, angular second moment (ASM), contrast (CON), and sum-of-squares variance (SSV) calculated from gray-level co-occurrence matrices (GLCM). Geometric features like defect ratio, perimeter, and circularity were also considered. The study further incorporated feature selection techniques, specifically Sequential Floating Forward Selection (SFFS), to identify the most relevant features for classification.(Unay et al., n.d.)

For the grading process, Unay et al. (n.d.) utilized a combination of statistical and syntactical classifiers, including Linear Discriminant Classifier (LDC), k-Nearest Neighbor Classifier (k-NN), Fuzzy k-NN, Support Vector Machines (SVM), and C4.5 decision tree classifier. These classifiers were employed to categorize apples into quality grades based on the extracted features. The project aimed to achieve both two-category grading (healthy or defective) and multi-category grading, providing a more realistic assessment of fruit quality. The results indicated that feature selection improved the classification performance, and statistical classifiers outperformed their syntactical counterparts. Notably, their two-category grading system achieved an overall accuracy of 93.5%, demonstrating the effectiveness of their approach (Unay et al., n.d.)

In recent years, the use of deep learning techniques for the diagnosis of crop diseases has gained attention in agriculture. In the study by Sanida et al. (2023), an efficient hybrid convolutional neural network (CNN) classification model for the early detection of tomato crop diseases was proposed. The project focused on addressing the significant challenges posed by tomato diseases, which could have a detrimental impact on both crop quantity and quality. Sanida et al. (2023), highlighted the importance of timely and accurate disease diagnosis in agriculture and emphasize the limitations of traditional manual diagnosis methods. The hybrid CNN model presented in their study combined VGG blocks with an inception module to enhance feature extraction capabilities. The model was trained on a dataset of ten categories, including nine distinct tomato diseases and one healthy category. The key performance metrics for the proposed model had a testing accuracy of 99.17%, precision of 99.13%, recall of 99.23%, F1-score of 99.17%, and an area under the curve (AUC) of 99.56%. This demonstrated the model's high accuracy and reliability in diagnosing tomato crop diseases.(Sanida et al., 2023; Yamamoto et al., 2014)

In the realm of apple quality grading, the study conducted by Ji et al. (2023) presented an innovative approach that gave leverage to multi-dimensional view information processing and deep learning techniques. The research was to enhance the efficiency and accuracy of apple quality grading. Ji et al. (2023) employed the Retinex algorithm for image enhancement, mitigating the impact of uneven illumination and surface reflection on grading accuracy. Then, the YOLOv5s model was enhanced with the addition of ODConv dynamic convolution and lightweight GSConv and VoVGSCSP backbones. The modification allowed for the model to simultaneously detect surface defects and identify fruit stem and calyx information, enhancing the quality of surface information captured, Ji et al. (2023). Additionally, the introduction of the Swin Transformer module to the Resnet18 backbone improved the overall grading accuracy by enabling more effective integration of context information from apple surface features. Experimental results demonstrated the effectiveness of this approach, with a recognition accuracy of 96.56% for apple surface defects, a 94 .46% average grading accuracy for quality classification, and a detection rate of 32 frames/s(Ji et al., 2023).

In the study by Vetal and Khule (2017), image processing techniques was used to detect tomato plant diseases at an early stage, primarily focusing on morphological changes in leaves as they are often the first affected by diseases. The methodology proposed by Vetal and Khule (2017) involved several key steps. First, digital images of tomato leaves were captured using cameras or mobile devices with high resolution. These images underwent image preprocessing techniques, including smoothing with Kurtosis and skewness filters to enhance image quality. Image segmentation was then employed to isolate the disease-affected areas on the leaves. The study utilized the inverse difference method which resulted in two segmented images: one containing only the disease-affected regions and another with the extracted disease images. To facilitate disease classification feature extraction was done on the segmented images. The study used the HIS (Hue, Saturation, Intensity) color space and derived texture features, including Energy, Entropy, Correlation, and Homogeneity. These features were calculated for each disease-affected region and served as the basis for classification. The classification of diseases was accomplished using the Multi-class Support Vector Machine (SVM) algorithm, which could handle the multiclass nature of the problem effectively. SVM classifiers were trained using feature datasets.(Vetal & R.S., 2017)

In a systematic literature review, Murad et al. (2023) explore the growing field of weed detection in agriculture using deep learning (DL) techniques. One of the key findings of the review was the diversity of weed types that had been successfully detected using DL. The reviewed study covered 34 unique weed types, demonstrating the versatility of DL algorithms in handling different weed species Murad et al. (2023) noted that RGB images are frequently used for weed detection. These images were captured using various devices, including robots, drones, and cell phones, indicating the adaptability of DL methods to different data collection platforms. The review also compared the performance of DL algorithms with traditional machine learning (ML) and image processing (IP) techniques. SVM stood out among ML algorithms. On Deep Learning, Convolutional Neural Networks, including variants like VGGNet, demonstrated strong performance.(Arendt Sørensen et al., n.d.; Murad et al., 2023)

Jin et al. (2022) evaluated three state-of-the-art deep learning models, namely YOLO-v3, CenterNet, and Faster R-CNN in their research to develop a novel deep learning-based method for detection of weeds in vegetables. These models exhibited remarkable performance, with average precision (AP) consistently exceeding 97% in the testing dataset. Among them, YOLO-v3 demonstrated the highest accuracy, with an impressive F1 score of 0.971, in addition to high precision and recall values. YOLO-v3 excelled in computational efficiency, with inference times comparable to CenterNet but significantly faster than Faster R-CNN. The proposed method avoided the complexities associated with training deep-learning models to detect various weed species, thus streamlining the weed detection process in vegetable fields.(Jin et al., 2022)

## 7.3. Machine Learning in Agriculture and Supply Chain

Liakos et al. (2018) on Machine Learning in agriculture shed light on the potential benefits it offered to the field. The study categorized the diverse range of projects into four main areas: crop management, livestock management, water management, and soil management. The study mentioned that in crop management, ML models had been harnessed for yield prediction, disease detection, weed identification, assessment of crop quality, and species recognition. In the livestock sector, Liakos et al. (2018) had found utility in enhancing animal welfare and optimizing livestock production. Liakos et al. (2018) defined Machine Learning as the scientific field endowing machines with the ability to learn without rigid programming. They emphasized the growing prevalence of ML across diverse scientific domains, including bioinformatics, biochemistry, medicine, meteorology, economics, robotics, aquaculture, food security, and climatology. Within the agricultural context, ML manifested through a spectrum of models and algorithms. For instance, Bayesian models (BM) were employed for classification or regression problems, utilizing probabilistic graphical models within the context of Bayesian inference. Instance-based models (IBM), relied on direct comparisons between new examples and instances in the training dataset, constructing hypotheses directly from available data. Decision trees (DT) employed tree-like structures to progressively organize datasets into homogeneous subsets, facilitating classification or regression. Artificial neural networks (ANNs), inspired by the human brain, comprised interconnected processing units organized in layers that enabled pattern recognition and decision-making. Furthermore, deep learning (DL) or deep neural networks (DNNs) extended the capabilities of ANNs with multiple hidden layers to learn complex data representations. Support vector machines (SVMs) constructed linear separating hyperplanes, enhancing classification capabilities through dimensionality transformation. Ensemble learning (EL) amalgamated multiple base learners to improve predictive performance, often utilizing decision trees as base models. These varied Machine Learning models and algorithms offered a versatile toolkit for tackling agricultural challenges(Liakos et al., 2018).

In the study conducted by Yang et al. (2021), the researchers addressed the crucial issue of disease and pest detection in grape foliage, which significantly impacted grape yield and quality. They collected three types of grape foliage images, including RGB images (RGBI), multispectral images (MSI), and thermal infrared images (TIRI), encompassing six common classes of grape health and disease. To build effective detection models, they employed ShuffleNet V2, a lightweight neural network. The study aimed to enhance detection accuracy by fusing the information from these different image modalities and proposed a multi-source data fusion (MDF) decision-making method. This method rectified 40% of incorrect detection outputs and achieved an overall accuracy of 96.05%, surpassing individual RGBI, MSI, and TIRI models. ShuffleNet V2's lightweight design, with 3.785 million total parameters and 0.362 G multiply-accumulate operations, made the model highly portable for practical applications in the field (R. Yang et al., 2021).

Nkemelu et al. (2018) explored the use of machine vision technologies for the classification of plant seedlings, particularly in the context of precision agriculture. Their dataset comprised 4,275 images of approximately 960 unique plants belonging to 12 species at various growth stages. The research compared the performances of traditional computer vision algorithms with a deep learning technique, specifically a Convolutional Neural Network (CNN). The findings demonstrated that the CNN-driven seedling classification, when integrated into farming automation systems, could optimize crop yield and enhance productivity while reducing labor costs. The CNN model exhibited superior performance, achieving an accuracy of 92.6% when using OpenCV preprocessing of input image data. Nkemelu et al. (2018) also emphasized the importance of addressing the weed problem in agriculture, as weeds could parasitically compete with crops for resources, resulting in significant losses and increased labor costs. The study demonstrates that a well-designed CNN-driven seedling classification system could contribute to the optimization of crop yield and the reduction of manual labor in agriculture, paving the way for further advancements in precision farming (Hasan et al., 2021; Nkemelu et al., 2018).

Veeragandham et al. (2020) discuss on how Machine learning emerged as a transformative tool in modern agriculture, addressing challenges in crop management, disease detection, and yield optimization. By combining deep learning and artificial intelligence algorithms, machine learning models have been trained to process sensor data, enabling real-time decision-making based on environmental conditions and soil quality. Machine learning has also been applied to tasks such as crop disease detection, pest management, and weed detection, contributing to increased crop yield and improved food quality Veeragandham et al. (2020). The utilization of machine learning in agriculture also extends to monitoring environmental conditions on farms, preventing soil erosion, and optimizing water management to conserve resources (Horvitz & Mulligan, 2015; Sivakumar et al., 2020; Veeragandham & Santhi, 2020).

Fenu et al. 2019 study focused on the prediction of potato late blight disease in Sardinia, a region heavily reliant on agriculture. The research leveraged Machine Learning techniques, specifically Feed-forward Neural Network and Support Vector Machine Classification, to predict disease severity based on meteorological parameters provided by ARPAS weather stations. The results were promising, with an accuracy rate of 96% for the Feed-forward Neural Network and 98% for the Support Vector Machine Classification. The development of Decision Support Systems (DSS) like DSS LANDS played a crucial role in providing actionable insights to farmers, enabling them to make informed decisions for disease management. By utilizing weather data and Machine Learning, this approach helped in mitigating yield losses caused by diseases like potato late blight, contributing to food security and sustainable agriculture (Fenu & Malloci, 2019).

Cravero et al. 2021 study recognized the importance of addressing the growing challenges in agriculture, such as population growth, climate change, and food security. It emphasized the shift towards data-driven agriculture and Smart Farming, facilitated by the collection of extensive data through various sensors, satellite images, and machinery. The integration of Big Data and ML is positioned as a solution to improve productivity and sustainability in agriculture. The study also highlighted the challenges in adapting ML to Big Data, considering factors like data volume, variety, and processing speed. The review analyzed real cases of Big Data and ML applications in agriculture, shedding light on architecture design and ML adaptations. The findings revealed that handling large data volumes had become manageable thanks to cloud technologies, while challenges persist in processing speed and information visualization for effective decision-making in agriculture(Cravero & Sepúlveda, 2021; Hashmi & Ahmad, 2016).

## 7.4. Advances in Hyperspectral Technology

Wan et al. 2022 literature review provided a comprehensive overview of the principles, methods, and applications of hyperspectral sensing in plant disease identification. The authors emphasized the significant impact of pathogens on plant photo responses and spectral characteristics, highlighting the need for rapid and non-destructive disease detection methods. Various hyperspectral imaging models and techniques, such as vegetation indices and machine learning classification, were discussed as crucial elements in disease identification (Hosny et al., 2023; Wan et al., 2022).

Zhang et al. (2020) found that hyperspectral technology has emerged as an effective way for plant disease detection and monitoring, giving reliable support for plant protection. The review’s authors discussed the benefits of hyperspectral technology in plant disease detection. They outlined the steps involved in hyperspectral disease analysis, comprising algorithms and qualitative and quantitative evaluation procedures. They also found key challenges in plant disease detection utilizing hyperspectral technologies, such as the detection of different pathogens, discrimination of biotic and abiotic stresses, and plant disease early warning. The authors proposed findings to these challenges, comprising the joint use of existing technologies, formation of a in-depth spectrum library of plant diseases, and incorporating ground-based, airborne, and spaceborne hyperspectral technologies. The advancement of multi-source remote sensing data and data fusion methods may also be necessary for real-time monitoring of plant infections/diseases at regional, national, and global scales. Overall, the developments in hyperspectral technology offer many opportunities for future plant disease detection and monitoring research (Bian et al., 2022; N. Zhang et al., 2020).

## 7.5. Application of Machine Learning and IoT in Precision Agriculture

Bashir et al. n.d. study focused on utilizing the color and texture features of agriculture and horticulture produce to distinguish between normal and affected regions. The combination of these features proved to be highly effective in disease detection. The proposed methodology involved steps such as image acquisition, RGB component separation, histogram equalization, and color and texture analysis. Bashir et al. n.d. emphasized the significance of using scientific techniques and machine learning for disease identification, emphasizing the potential for early detection. Various image processing methods, including co-occurrence matrix (CCM) and K-means clustering, were employed for texture and color analysis, enhancing the accuracy of disease detection (Bashir & Sharma, n.d.).

In plant recognition, DL models, particularly convolutional neural networks (CNNs), have gained popularity for their ability to process image data effectively, BAL et al. 2021. These models have been used to identify various plant species, with a notable emphasis on rice. DL models, such as CNNs, have shown promise in accurately identifying diseases in crops like tomatoes. Weed and pest detection studies are divided into two categories: pest and weed detection. DL methods, especially CNNs, have been used in detecting both pests and weeds, while ML approaches employ algorithms like support vector machines (SVMs) and random forests (RF). Soil mapping and drought index determination studies involve various models, including CNNs, artificial neural networks (ANNs), and SVMs. Lastly, yield forecasting, a crucial aspect of agriculture, has been predominantly tackled with ML techniques like ANN and SVM. The studies evaluate models using performance metrics such as accuracy, precision, recall, F1-score, and others, with a focus on improving agricultural sustainability (BAL & KAYAALP, 2021).

Gutiérrez et al. 2022 study focused on the integration of the Internet of Things (IoT) technology with precision agriculture to predict crop recommendations using machine learning algorithms. The study aimed is to develop a system that could recommend crops based on various environmental parameters, including Nitrogen (N), Phosphorous (P), Potassium (K), pH, Temperature, Humidity, and Rainfall. The dataset used for the study comprised 2200 instances with 8 attributes. The authors used supervised learning methods and selected machine learning algorithms within the WEKA framework to create an optimal model for crop recommendation. The machine learning algorithms chosen for classification included multilayer perceptron, JRip (a rules-based classifier), and decision table classifier. The selected classifiers achieved a performance accuracy of 98.2273%, with a weighted average Receiver Operator Characteristics (ROC) score of 1, and a maximum model build time of 8.05 seconds (Gutiérrez et al., 2022; S. L. Zhang & Chang, 2015).

In a comprehensive review, Xiao et al. (2023) shed light on the advancements in Deep learning, especially using convolutional neural networks (CNN), which had proven its efficacy in extracting high-dimensional features from fruit images, enabling accurate and fast fruit detection and recognition. Xiao et al. (2023) categorized the challenges faced in fruit detection for automatic harvesting into several key areas. These challenges included the scarcity of high-quality fruit datasets, detecting small target fruits, identifying fruits in obstructed and dense environments, recognizing fruits of varying sizes and species, and developing lightweight fruit detection models. In response to these challenges, Xiao et al. (2023) proposed potential solutions and outline future development trends. The study discussed the need to prioritize addressing the existing challenges while enhancing the accuracy, speed, robustness, and generalization of fruit vision detection systems. Xiao et al. highlighted the relevance of reducing overall complexity and cost in future research efforts. This review categorized methods based on popular DL architectures, such as YOLO, SSD, AlexNet, VGGNet, ResNet, Faster R-CNN, FCN, SegNet, and Mask R-CNN. Xiao et al. (Year) discussed the strengths and weaknesses of these methods and identified current research trends, including the increasing adoption of Faster R-CNN and YOLO models (Bengio, n.d.; Xiao et al., 2023).

## 7.6. Deep Learning and CNNs in Image Recognition

A survey and benchmark study by Zoller and Huber (2019) provides a comprehensive overview of the current landscape of Automated machine learning methods and evaluates popular AutoML frameworks using real-world datasets. AutoML entails several components, such as, automatic feature engineering, feature generation, and feature selection. Automatic feature engineering involves the creation of new features from existing data, while feature generation aims to expand the feature set through a combination of unary and binary operations. Feature selection, on the other hand, focuses on choosing a subset of features to enhance model performance and interpretability(Bengio, n.d.; Zöller & Huber, 2019).

Scientists have aimed to provide a comprehensive overview of Convolutional Neural Networks(CNNs), exploring their underlying architecture and various applications. Specific focus on popular CNN architectures, such as LeNet, AlexNet, ZF Net, GoogleNet, VGGNet, and ResNet, discussing their strengths and advancements, have been researched Sakib et al. (2018). The study by Sakib et al. (2018) contributes to the understanding of CNNs structure and their impact on machine learning applications. The study explains that Convolutional Neural Network (CNN) is a supervised deep learning method designed for image processing tasks, distinct from unsupervised techniques like Fuzzy C-Means and ADBSCAN clustering. A typical CNN architecture is constructed with an input layer, an output layer, and multiple hidden layers. These hidden layers are categorized into three types: CONV, POOL, and FC (Fully Connected), with an additional nonlinearity activation function applied as a layer. CNNs are tailored to handle image data efficiently, given their inherent 2D structure. The layers within a CNN are organized into three dimensions, corresponding to the input's height, width, and depth. The architecture aims to exploit this 2D structure, emphasizing local connections and shared weights, which promote translation-invariant characteristics. The first key layer in a CNN is the Convolution Layer, responsible for most of the computation. This layer employs learnable kernels to convolve filters across the input's spatial dimensions, producing 2D activation maps. The neurons in this layer share weights, reducing network complexity. Rectified Linear Units (ReLU) introduce non-linearity to the model. Hyperparameters like depth, stride, and zero-padding control the output's spatial dimensions. Following the Convolution Layer, the Nonlinearity Layer applies activation functions like ReLU to enhance network non-linearity, with SOFTMAX employed at the final layer. Pooling Layers, situated after Convolutional Layers, decrease feature map dimensions by summarizing subregions through functions like max pooling. This step reduces data dimensionality, the number of parameters, and combats overfitting. Lastly, Fully Connected Layers facilitate high-level reasoning within the neural network. Neurons here connect to all activations in the previous layer. This layer combines high-level features learned by convolutional layers and passes them to the output layer for class prediction. In summary the study explain how a CNN architecture leverages convolution, nonlinearity, pooling, and fully connected layers to process image data effectively. Convolution and pooling layers capture local patterns and reduce dimensionality, while fully connected layers enable high-level feature combinations for accurate classification. ReLU and SOFTMAX activation functions introduce non-linearity, enhancing the network's ability to learn complex patterns in data (He et al., n.d.; Lecun et al., 2015; Sakib et al., 2018; Shah et al., n.d.).

Yang et al. (2020) on hyperparameter optimization (HPO) of machine learning algorithms, discuss various techniques for tuning hyperparameters to improve the performance of machine learning models. They explain the importance of HPO in building effective models and highlight the challenges associated with manual tuning, such as the time-consuming nature and the need for deep knowledge of ML algorithms. The study introduced optimization techniques and provided insights into how to apply them to machine learning algorithms. One of the basic hyperparameter tuning methods discussed was "Babysitting," also known as "Trial and Error" or "Grad Student Descent" (GSD). Babysitting as explained by Yang et al. (2020) involves manual tuning by repeatedly testing hyperparameter values based on experience and analysis of previous results. However, it requires a significant amount of prior knowledge and can be time-consuming. Grid Search (GS) another technique covered in the paper, involves an exhaustive search of hyperparameter combinations within a predefined grid of values. GS is easy to implement but becomes inefficient for high-dimensional configuration spaces due to the curse of dimensionality Yang and Shami. Random Search (RS) is introduced as a more efficient alternative to GS. RS randomly selects hyperparameter values within specified bounds and evaluates them, allowing for exploration of a larger search space. It is particularly useful for large configuration spaces. Gradient-Based Optimization techniques, which include gradient descent, were discussed for optimizing continuous hyperparameters. These techniques calculated gradients to determine promising directions for optimization. However, they are limited to continuous hyperparameters and may not always find global optima. Bayesian Optimization (BO), an iterative algorithm that uses surrogate models and acquisition functions to guide the selection of hyperparameter configurations was also explained. BO is efficient as it learns from previously evaluated points to balance exploration and exploitation, making it suitable for both continuous and categorical hyperparameters (Chowdhury et al., 2022; L. Yang & Shami, 2020).

Totakura et al., (2020) study utilized mathematical tools like Fourier transformations in reverse order to generate Mel-Frequency Cepstral coefficients and neural networks, employing a layered architecture with the rectified linear unit (ReLU) activation function. ReLU is a piecewise linear function that outputs positive inputs directly and zero otherwise, making it easier to train and yielding better performance. The activation function's primary role is to transform input signals in an artificial neural network into output signals, which are then passed to the next layer. In the experiment's conclusion, the Adam Optimizer is employed to compile the model. Adam is an optimization algorithm that replaces stochastic gradient descent, updating network weights iteratively during training. The labels for the data are assigned using Sparse Categorical Cross Entropy, suitable for mutually exclusive classes. This approach is particularly useful when classifying images (Totakura et al., 2020).

Karthikeya Racharla et al. study of the evaluation of machine learning models for predominant musical instrument classification based on spectral features had several key performance metrics employed. Precision, which was the ratio of true positives to the sum of true positives and false positives, measured the classifier's ability to avoid labeling false positives as positives. Recall, represented as the ratio of true positives to the sum of true positives and false negatives, assessed the classifier's capability to identify all positive samples. The F1 score, considered as the harmonic mean of precision and recall, provided a balanced measure of a model's performance. Additionally, accuracy, calculated as the proportion of correctly classified examples, disregarded class distribution and categorized observations as either correctly or incorrectly classified. To gain insight into the model's classification performance, confusion matrices were employed, offering a cross-tabulation of actual and predicted classes. These performance metrics collectively facilitated the assessment of a model's accuracy, precision, recall, and overall effectiveness in classifying predominant musical instruments based on spectral features (Racharla et al., 2020).

The study conducted by Hoiem et al. (2021) addressed the critical need for more comprehensive and informative performance evaluation methodologies in machine learning. The researchers focused on the significance of training data size, which significantly impacted model performance, particularly in areas like representation learning, data augmentation, and low-shot learning. In the study, Hoiem et al. (2021) proposed a method to robustly estimate learning curves and abstract their parameters into error and data-reliance, thereby enabling more stable and meaningful performance comparisons. Learning curve models help uncover insights that single-point comparisons of performance may overlook, facilitating better understanding and improvement of classifiers, Hoiem et al. This research served as a valuable contribution to the field by promoting the use of learning curves as a standard part of learning system evaluation, offering a more nuanced and informative perspective on classifier performance (Hoiem et al., 2021; Tabik et al., 2017)

In the field of forecast verification, Receiver Operating Characteristic (ROC) curves play a pivotal role in assessing the predictive ability of probabilistic forecasts. ROC curves are widely used to evaluate probability forecasts and are employed across various scientific disciplines and application domains, such as meteorology, hydrology, medicine, and finance. The ROC curve is a graphical representation of a forecast's ability to correctly predict binary events, and it plots the hit rate (HR) against the false alarm rate (FAR) as the threshold for defining events varies. HR measures the proportion of true positives, while FAR quantifies the rate of false alarms. The Area Under the ROC Curve (AUC) serves as a critical measure of a predictor's performance, with AUC values closer to 1 indicating better predictive accuracy. AUC is interpreted as the probability that a randomly chosen predictor value for an event is greater than that for a non-event. ROC curves and AUC are invariant to changes in class proportions and strictly increasing transformations of the predictor variable. They offer an objective and scale-independent means of evaluating the quality of probabilistic forecasts, making them invaluable tools in forecast verification (Gneiting & Vogel, 2018).

Cross-validation is a crucial data resampling method used to assess the generalization ability of predictive models and mitigate the risk of overfitting. In the realm of supervised learning, where predictive models aim to make accurate predictions based on training data, overfitting is a significant concern. This occurs when a model becomes too complex and fits the training data perfectly but struggles to generalize to new, unseen data. Cross-validation addresses this challenge by providing an estimate of a model's predictive performance using the same dataset. It involves partitioning the dataset into training and validation subsets, iteratively training the model on different subsets, and evaluating its performance. This process helps strike a balance between overfitting and underfitting by assessing how well the model generalizes to new data. Common types of cross-validation include k-fold cross-validation, leave-one-out cross-validation, and single hold-out subsampling, each offering insights into a model's performance (Berrar, 2018).

The learning rate parameter determines the step size at which the algorithm updates the model's weights in response to observed errors during training. Study by Wilson and Martinez (2001) explored the impact of learning rates on training speed and generalization accuracy, especially concerning large and complex problems. They emphasized that selecting the right learning rate is essential, as a rate that is too large could lead to poor accuracy and slow convergence, while a rate that is too small may be computationally wasteful. The study proposed a systematic approach for determining the optimal learning rate, starting with a fast-learning rate and progressively decreasing it until the point of diminishing returns is reached. They also advocated using hold-out set accuracy as a stopping criterion for training. The work provided valuable insights into optimizing learning rates for neural networks on challenging tasks, balancing training speed and generalization accuracy (Wilson & Martinez, 2001).

## 7.7. Bitwise Neural Networks for Computational Efficiency

In a groundbreaking study titled "Bitwise Neural Networks," Kim and Smaragdis the authors propose a novel approach to neural network architecture that revolutionizes the field of computational efficiency. The study addressed the challenge of resource-constrained environments by introducing a bitwise representation of all network components, including weights, biases, inputs, hidden layer outputs, and outputs. By replacing conventional floating-point or fixed-point arithmetic operations with bitwise logic, their proposed Bitwise Neural Network (BNN) significantly reduced spatial complexity, memory bandwidth requirements, and power consumption in hardware. The study focused on hand-written digit recognition using the widely known MNIST dataset. The authors outlined their methodology, which involved training schemes such as weight compression and noisy backpropagation to create a BNN that performs nearly as well as its real-valued counterpart. The results obtained demonstrated that BNNs achieved competitive performance while offering remarkable computational savings. The researchers compared the performance of the BNN with that of the corresponding real-valued networks, highlighting the small additional errors observed in the bitwise networks.

The conclusions drawn from the study indicated that BNNs provided a highly efficient alternative for neural network deployment, particularly in resource-constrained scenarios. By utilizing basic bit logic operations instead of complex arithmetic, BNNs offer significant advantages in terms of computational resources, memory requirements, and power consumption. The researchers suggest that this innovative approach could have profound implications across various fields, including embedded systems, context-aware computing, computer vision, and speech-driven personal assistant services. Kim and Smaragdis' study paves the way for further investigations into bitwise versions of convolutive neural networks and other network architectures. The potential impact of BNNs in diverse domains calls for continued exploration and refinement, as this groundbreaking study represents a significant step forward in the quest for efficient computing (Kim & Smaragdis, 2016).

## 7.8. Confidence Calibration in Classification Models

The study conducted by Guo et al. (2017) explored the crucial problem of confidence calibration in classification models, particularly focusing on modern neural networks. The researchers aimed to understand why these networks were poorly calibrated and identify factors that influenced calibration, such as depth, width, weight decay, and Batch Normalization. Their objective was to provide insights into neural network learning and propose practical solutions for improving confidence estimates. The methodology employed by the researchers involved extensive experiments evaluating various post-processing calibration methods on state-of-the-art architectures using image and document classification datasets. They compared the performance of different techniques, including temperature scaling, which is a single-parameter variant of Platt Scaling. The results of their experiments revealed that modern neural networks, despite their improved accuracy, suffered from miscalibration issues. In comparison to older networks, such as LeNet, the confidence estimates of newer architectures, like ResNet, did not align well with their actual accuracy. This finding was depicted in histograms and reliability diagrams, emphasizing the need for improved confidence calibration. Furthermore, the researchers demonstrated that factors such as model capacity, normalization, and regularization greatly affected network calibration. While the precise reasons behind these trends remained a topic for further investigation, the study highlighted the significance of understanding and addressing calibration issues in neural networks.

In terms of practical applications, well-calibrated confidence estimates not only enhanced model interpretability but also established trustworthiness with users. Additionally, calibrated probabilities could be integrated with other probabilistic models, leading to improved performance in areas such as speech recognition and object detection, Guo et al. (2017). The study concluded that temperature scaling, a simple and efficient technique, was often the most effective method for achieving calibrated probabilities in modern neural networks. Given its straightforward implementation with existing deep learning frameworks, it could be readily adopted in practical settings. In conclusion, the research conducted by Guo et al. shed light on the problem of confidence calibration in modern neural networks. Their comprehensive analysis and experiments provided valuable insights into the factors influencing calibration and propose temperature scaling as a practical solution. This study contributed to the field by addressing an important aspect of neural network performance and warrants further research to explore the underlying reasons behind the observed trends (Guo et al., 2017).

## 7.9. Properties and Challenges of Deep Neural Netwoks

Szegedy et al. (2014) conducted research to shed light on two intriguing properties of deep neural networks. The authors address the semantic meaning of individual units in deep neural networks. Traditionally, researchers have examined the role of individual units by identifying inputs that maximally activate them. However, Szegedy et al. challenge the assumption that individual units hold distinct semantic information. They demonstrate that random linear combinations of high-level units are indistinguishable from individual units, suggesting that it is the collective space of activations that contains the bulk of semantic information (Szegedy et al., 2013).

Szegedy et al. (2014) also investigated the stability of deep neural networks with respect to small perturbations in their inputs. They discover that imperceptible perturbations can cause the network to misclassify an image, termed "adversarial examples." These perturbations, found by maximizing prediction errors, highlight the non-robust nature of neural networks. Intriguingly, adversarial examples are not mere artifacts of overfitting or training set selection but appear to be universal. They persist across networks trained with different hyperparameters and even on different subsets of the training data. The counter-intuitive properties of deep neural networks revealed in this study raise important questions about their interpretability and generalization capabilities. The existence of adversarial examples challenges the network's ability to generalize well, as they are visually similar to regular examples yet lead to misclassification. Further research is needed to understand the frequency of adversarial examples and their impact on network performance. Moreover, investigating the potential of leveraging adversarial examples in training to enhance generalization shows promising initial results (L’Heureux et al., 2017; Szegedy et al., 2013).

## 7.10. Ensemble Methods with Naïve Bayes Classifier on classification

Neural network ensembles involve training multiple neural networks to solve a problem and then combining their predictions. Research analyzing the relationship between the ensemble and its component neural networks in the context of regression and classification tasks has been explored. It suggests that it may be better to ensemble many neural networks instead of all available ones. The GASEN approach aims to select the appropriate neural networks from a set of available ones to form an ensemble. It involves training several neural networks, assigning random weights to them, and using genetic algorithms to evolve the weights. The evolved weights characterize the fitness of the neural networks in constituting an ensemble. GASEN selects a subset of neural networks based on these weights to form the ensemble (Zhou et al., 2002)

In the study conducted by Asmala Ahmad and her colleagues, they investigated the performance of Naïve Bayes classification on high-resolution aerial imagery captured from a UAV-based remote sensing platform. To aid in the selection of training pixels for Naïve Bayes classification, they initially employed K-means clustering on the study area. They experimented with different training set sizes, ranging from 10 to 100 pixels, and used linear and quadratic discriminant analyses for classification. The results indicated that a training set size of 20 pixels yielded the highest overall classification accuracy and Kappa coefficient for both linear and quadratic discriminant analyses. The linear discriminant analysis achieved an overall classification accuracy of 94.44%, with a Kappa coefficient of 0.9395, which was superior to the quadratic discriminant analysis, which had an overall classification accuracy of 88.89% and a Kappa coefficient of 0.875. Additionally, they examined producer accuracy and area size of individual classes, finding that linear discriminant analysis produced more realistic classifications due to limited homogenous training pixels for certain objects (Ahmad et al., n.d.).

The Naïve Bayes classification method, as described in the study (Ahmad et al., n.d.), is a statistical approach used in remote sensing for assigning pixels to specific land cover classes based on a feature vector derived from spaceborne or airborne acquisition systems. It operates on the principles of conditional probability and Bayes' theorem, where the probability of a pixel belonging to a particular class is determined given its spectral measurements. In the context of remote sensing, it relies on known properties of each land cover type, which are obtained from training pixels. The Naïve Bayes classifier assumes that the features used for classification are independent, and it calculates the posterior probability of each class based on the prior probability and likelihood function. Ultimately, it assigns a pixel to the class with the highest posterior probability. This classification method is valuable for accurately categorizing land cover types in remote sensing imagery (Ahmad et al., n.d.).

## 7.11. BPNN/ANN Architectures

Keiron O'Shea and Ryan Nash define ANNs as computational systems inspired by biological nervous systems, consisting of interconnected neurons that collectively learn from input to optimize output. Their study explain overfitting as being identified as a challenge in training ANNs, where the model may perform well on the training data but poorly on unseen data due to excessive complexity and lack of generalization. The study then introduced Convolutional Neural Networks (CNNs) as a type of ANN architecture specifically designed for image-driven pattern recognition tasks. Keiron O'Shea and Ryan presented CNNs as a superior alternative to traditional ANNs for image analysis due to their ability to capture image-specific features while reducing computational complexity and explain that CNNs achieve this by employing convolutional layers, pooling layers, and fully-connected layers. Fully-connected layers in CNNs resemble traditional ANNs and are responsible for producing class scores for classification tasks. The study suggests that reducing the complexity of fully-connected layers is essential to prevent overfitting (O’Shea & Nash, 2015; C.-C. Yang et al., n.d.).

Artificial Neural Networks (ANNs) have gained widespread usage, particularly the Multi-Layer Perceptron (MLP) networks based on the Back-Propagation (BP) learning algorithm, Alaeldin Suliman et al. (2015) . These ANNs are employed for the classification of remotely sensed digital images, aiming to extract distinct land cover categories. ANNs, inspired by the biological central nervous system, consist of interconnected processing units known as neurons. The architecture of these networks typically comprises input, hidden, and output layers. The input layer distributes and scales input data, while the hidden layers process information before the output layer provides class labels Suliman et al. (2015). The activation functions within neurons play a crucial role in determining output, with the sigmoid function commonly used in classification tasks. Learning in ANNs can be supervised, unsupervised, or hybrid, with the BP learning algorithm being prominent for supervised learning. BP adjusts connection weights iteratively to minimize the error between actual and desired network outputs. However, the effectiveness of BPNNs depends on the network design and implementation, leading to variations in classification accuracies. Researchers often conduct multiple experimental trials to optimize network design, a time-consuming process (Alaeldin Suliman & Yun Zhang, 2015; *Backpropagation Step by Step*, 2018; Z. Zhang, 2016).

## 7.13. Related works conclusion

Researchers aim to explore the potential of machine learning (ML) in addressing various challenges and improving agricultural practices(Meshram et al., 2021). Different studies emphasize the importance of ML in enhancing efficiency, precision, and quality in farming operations while minimizing losses and reducing the reliance on manual labor(Cravero & Sepúlveda, 2021). Researchers examine the utilization of computer vision systems, IoT, and other emerging technologies in different stages of farming(Gutiérrez et al., 2022). Significance of data acquisition, preprocessing, model selection, training, and testing in building accurate ML models for agricultural purposes(Horvitz & Mulligan, 2015). Study findings demonstrate that ML, mainly through the use of convolutional neural networks (CNNs) show significant potential in the agriculture domain(Nkemelu et al., 2018). Researchers have achieved remarkable results in crop classification, object detection, yield estimation, and post-harvest storage and processing systems. Authors stress the importance of creating customized ML models that can effectively analyze complex data and provide valuable insights, predictions, and recommendations for farmers. Despite the benefits offered by ML in agriculture, researchers acknowledge several challenges they usually face. They include data availability and quality, the time-consuming nature of model training and testing, the selection of suitable algorithms, and the difficulties associated with model deployment. The article suggests recommendations such as automated machine learning (AutoML) to address these challenges and expedite the model development process(Zöller & Huber, 2019).

The proposed project aims to integrate multi-source data and image identification using deep learning to improve the efficiency and accuracy of crop disease diagnosis in agriculture. proposed topic adopts deep learning algorithms, such as convolutional neural networks (CNNs), for a more precise and robust classification of plant diseases. By use of various data sources the proposed system can provide timely and accurate recommendations to farmers, thereby reducing crop losses and increasing productivity. Prior research has explored the use of machine learning algorithms, including supervised, unsupervised, and reinforcement learning, in developing sustainable agriculture supply chains(Sanida et al., 2023). By incorporating deep learning algorithms, the proposed project can enhance the prediction of changes and optimization of farming practices, leading to higher yields and improved crop quality. It can also help overcome the limitations imposed by environmental factors that impact agriculture. Additionally, the project can build upon existing research studies that have explored the use of machine learning and image-processing methods for crop disease diagnosis(Sanida et al., 2023).

CNNs have shown great potential in image recognition tasks and offer advantages in processing image data and handling spatial dimensionality(Vetal & R.S., 2017). By utilizing CNNs, the project can enhance the accuracy and efficiency of crop disease diagnosis. The project can also consider the findings from research on bitwise neural networks, which propose a novel approach to neural network architecture that improves computational efficiency.

Furthermore, the project can explore the challenges of confidence calibration in neural networks and investigate practical solutions for improving confidence estimates. Well-calibrated confidence estimates enhance model interpretability and establish trustworthiness with users, leading to more reliable decision-making in crop disease diagnosis(Guo et al., 2017). The research on the semantic meaning of individual units in deep neural networks and the existence of adversarial examples can also inform the project(O’Shea & Nash, 2015). Understanding the collective space of activations and the vulnerabilities of neural networks to small perturbations can contribute to the development of more robust and interpretable models for crop disease diagnosis(Guo et al., 2017).

# **DESIGN AND METHODOLOGY**

## Data Acquisition

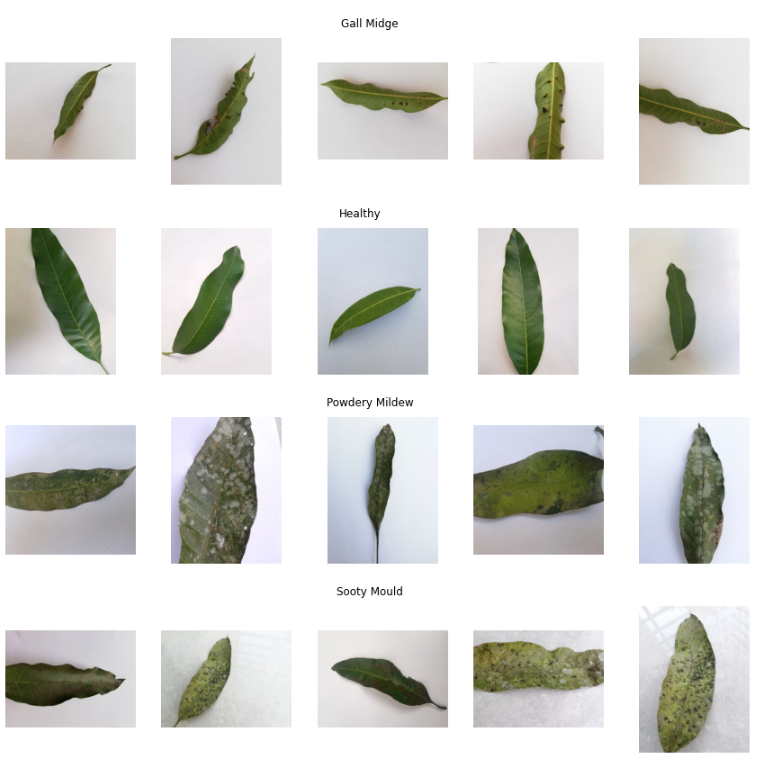
The design and methodology of this project was driven by the goal to create a robust and efficient system for automated crop disease diagnosis through integrating multi-source data and applying deep learning techniques. The primary focus on utilizing images stemmed from their capacity to vividly capture visual symptoms of plant diseases, making them an ideal foundation for identification(Adhiparasakthi Engineering College et al., n.d.-b). By combining multi-source data like images of both healthy and diseased plants, the methodology ensured a comprehensive understanding of the intricate interplay between various factors influencing plant health(Fenu & Malloci, 2019).

Top of Form

In the pursuit of utilizing advanced machine learning techniques, data collection and preparation play crucial roles(Tabik et al., 2017). The approach outlined integrated data from diverse sources to optimize the precision and effectiveness of identifying diseases in mangoes through the leaves. The following systematic breakdown provides a clear view of the detailed process, nurturing the growth of a strong deep-learning model. For this study, a data set from Mendeley Data site was utilized. The data set, ("MangoLeafBD Dataset,") location source was from Four mango orchards in Bangladesh, namely the Sher-e-Bangla Agricultural University orchard, Jahangir Nagar University orchard, Udaypur village mango orchard, and Itakhola village mango orchard, Ali et al. (2022). The first dataset’s utility extended twofold: firstly, for binary classification distinguishing between healthy and diseased leaves, and secondly, for multi-class classification, discerning among various diseases affecting the leaves, Ali et al. (2022). The dataset was centred on seven distinct diseases afflicting mango leaves: Anthracnose, Bacterial Canker, Cutting Weevil, Die Back, Gall Midge, Powdery Mildew, and Sooty Mould. These diseases represented pivotal challenges in mango orchards. Notably, the dataset was organized into eight classes, including the "healthy" category.



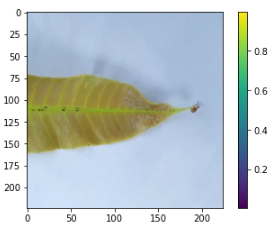
Each of these eight classes contained 500 images, ensuring a balanced representation. Dhaka University and East West University inclusion in this endeavor lends academic rigor and expertise for this first data set.



This data set comprised eight distinct classes, with seven dedicated to various diseases and one representing healthy mango leaves. This diversity was to enable the deep learning models to differentiate between healthy and diseased leaves and pinpoint specific diseases. This dataset served as the bedrock for the initial training of the model.

## Data Preprocessing

The preprocess began by converting plant disease categories into integer labels, establishing a bridge between text-based labels and numerical representations(Shah et al., n.d.). This step served more than just a procedural formality; it set the groundwork for translating real-world plant diseases into understandable machine values(Shah et al., n.d.). Each disease category, like "Anthracnose," "Bacterial Canker," and others, received a unique integer label, creating an organized system for the model to grasp. This encoding approach was to allow the model to differentiate between diseases effectively, forming the foundation for accurate classification(Shah et al., n.d.). Next, the images underwent transformations to ensure uniform dimensions and scales. Images with varying sizes from different sources were resized to a consistent dimension (224x224), ensuring consistency for further processing(Tabik et al., 2017). These preprocessed images were then converted into numerical arrays, translating visual data into a format the machine can interpret. Crucially, pixel values were scaled to a range of [0, 255] to match the model's input requirements(Tabik et al., 2017). This standardized preprocessing harmonized the data, enhancing the model's capacity to learn and recognize patterns(Tabik et al., 2017).



The subsequent stage involved splitting the dataset strategically into training and testing subsets. Around 80% of the data is allocated for training and 20% for testing. This division minimized the risk of overfitting, allowing the model to learn from a diverse array of samples while reserving unseen data for evaluation(Berrar, 2018; Chowdhury et al., 2022; Hoiem et al., 2021). This chosen split ratio stroked a balance between robust training and rigorous testing, a critical factor which accurately gauged the model's true performance(Berrar, 2018).

Encoding disease categories into integers tackled the challenge of translating complex real-world concepts into machine-friendly representations(Shah et al., n.d.). This structured system became the bedrock of the model's classification capabilities. Dimensional reshaping and preprocessing tackled the diversity present in multi-source data(Tabik et al., 2017). By standardizing dimensions and pixel values, the model became unbiased towards the original data sources, mitigating potential inconsistencies that could hinder precise disease identification. Lastly, the intentional data split encapsulated the core of machine learning, that is, the balance between learning and validation(Tabik et al., 2017). By allocating distinct sets for training and testing, the model gained exposure to varied data for learning while maintaining a strict evaluation process to measure its predictive abilities(Berrar, 2018).

To achieve the intended objectives, the project wisely opted to employ deep learning models, specifically Artificial Neural Networks (ANNs) like Backpropagation Neural Networks (BPNN), Convolutional Neural Networks (CNNs) and CNN-Naive Bayes ensemble/Naïve Bayes Classifier (Ahmad et al., n.d.; *Backpropagation Step by Step*, 2018; Sakib et al., 2018; Z. Zhang, 2016; Zhou et al., 2002). Manual hyperparameter tuning was to be done for the project(Ghosh, n.d.; Lecun et al., 2015). By training on various inputs, these models were expected to provide the precise and robust classification of plant diseases(Guo et al., 2017).

By integrating insights from the literature with the design and methodology, the project positions itself as an innovative extension of existing research. The design and methodology of this project are not only grounded in existing literature on agricultural technology, machine learning, and image processing, but also extend beyond the traditional approaches(Adhiparasakthi Engineering College et al., n.d.-b; Meshram et al., 2021). The integration of multi-source data, the application of deep learning techniques, and the adherence to ethical data practices collectively form a robust framework for achieving accurate and efficient crop disease diagnosis. This approach is poised to revolutionize agriculture by reducing losses, enhancing productivity, and promoting sustainable farming practices.

# **IMPLEMENTATION**

## BPNN/ANN ARCHITECTURES

The process of translating the design and methodology into meaningful actions marked the project's implementation phase. With a practical approach and a focus on efficiency, the project navigated through Backpropagation Neural Networks (BPNN) as they carefully transition into Artificial Neural Networks (ANN) (*Backpropagation Step by Step*, 2018; Z. Zhang, 2016)and also the exploration of Convolutional Neural Networks (CNN)(Sakib et al., 2018) to unravel the complexities of assessing crop health.

The project's initial focus started on Backpropagation Neural Networks (BPNN) that transitioned to advanced Artificial Neural Networks (ANN). This section revealed a step-by-step account of the project's exploration of BPNN/ANN, showcasing its significance in achieving the project's objective of exploring different deep learning paradigms to enhance the performance of crop disease diagnosis models(*Backpropagation Step by Step*, 2018; Z. Zhang, 2016). Beginning with a foundational approach, the project constructed a basic BPNN architecture comprising a single hidden layer(Alaeldin Suliman & Yun Zhang, 2015). This initial setup took into account essential architectural parameters(Z. Zhang, 2016). These included image dimensions of 224x224 pixels with 3 color channels, the requirement to classify crops into 8 disease categories, and training over 30 epochs with 15-sample batches(Tabik et al., 2017; C.-C. Yang et al., n.d.). The architecture unfolded as follows: A hidden layer with 32 units and ReLU activation is introduced, adept at transforming intricate input data into a more abstract representation(Bengio, n.d.; Lecun et al., 2015; Sakib et al., 2018; Totakura et al., 2020). The output layer comprised 8 units, each signifying a disease class and activated by the softmax function for classification(Bengio, n.d.; He et al., n.d.; Lecun et al., 2015). The model was compiled using the Adam optimizer and sparse categorical cross-entropy loss function which is used in multiclass classification tasks where there are a large number of classes, and the classes are mutually exclusive (each input belongs to one class, then trained on training images and labels while being validated with separate data(Shah et al., n.d.; Totakura et al., 2020; L. Yang & Shami, 2020).

The project progresses towards deeper ANN/BPNN architecture. The table below shows the different model parameters that were used to build the different architectures of the BPNN/Artificial Neural Networks architectures and the description of the table follows it.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Number Of Layers | Hidden Units Per Layer | Activation Function | Dropout rate | Learning Rate | Batch Size | Number Of Epochs |
| Model 1 | 2 | 32 | Relu, Softmax | None | Default (Adam) | 15 | 30 |
| Model 2 | 3 | 32, 64 | Relu, Softmax | None | Default (Adam) | 15 | 30 |
| Model 3 | 4 | 32,64,128 | Relu, Softmax | None | Default (Adam) | 15 | 30 |
| Model 4 | 5 | 32,64,128, 128 | Relu, Softmax | None | Default (Adam) | 15 | 30 |
| Model 5 | 6 | 32,64,128, 128 | Relu, Softmax | 0.5 | Default (Adam) | 15 | 30 |
| Model 6 | 6 | 32,64,128, 128 | Relu, Softmax | 0.5 | Custom (0.001) | 15 | 30 |
| Model 7 | 6 | 32,64,128, 128 | Relu, Softmax | 0.25, 0.5 | Custom (0.001) | 15 | 30 |

1. **Number of Layers:** The BPNN models had varying number of layers, ranging from 2 to 6. A greater number of layers was set to capture more complex relationships in the data. However, increasing the depth of the network could make it more prone to overfitting on the training data(Alaeldin Suliman & Yun Zhang, 2015; *Backpropagation Step by Step*, 2018; Szegedy et al., 2013).
2. **Hidden Units Per Layer:** Each BPNN model was set to have increased hidden units in its layers. Increasing the number of hidden units was to allow the model to learn more intricate patterns in the data. However, it also increased the model's capacity to overfitting(Hoiem et al., 2021; Szegedy et al., 2013).
3. **Activation Function:** The activation functions used in the hidden layers were ReLU (Rectified Linear Unit). ReLU introduced non-linearity into the model, allowing it to learn complex relationships in the data(He et al., n.d.; Lecun et al., 2015; Sakib et al., 2018). The softmax activation function is used in the output layer for multi-class classification tasks, providing class probabilities(Sakib et al., 2018).
4. **Dropout Rate:** Dropout is a regularization technique was applied to the dense layers in certain models (He et al., n.d.; Lecun et al., 2015; Sakib et al., 2018).. It serves to randomly drop a fraction of neurons during training, preventing them from contributing to the forward and backward passes. Dropout helps prevent overfitting by promoting robustness in the network (He et al., n.d.; Lecun et al., 2015; Sakib et al., 2018)..
5. **Learning Rate:** The learning rate determines the step size during the optimization process (training). The table mentions using the default Adam optimizer with different learning rates for each model. The learning rate is a crucial hyperparameter, and selecting an appropriate value is essential for achieving good convergence during training(Bengio, n.d.).
6. **Batch Size:** Batch size refers to the number of training examples used in each iteration of the optimization process(Lecun et al., 2015; Z. Zhang, 2016). A batch size of 15 is consistent across all models.
7. **Number of Epochs:** Each model was trained for a fixed number of epochs (30) to allow for training progress and convergence monitoring. The number of epochs represented how many times the entire training dataset is processed during training(Hoiem et al., 2021).

The experiments with the BPNN models allowed for the understanding of how different architectural choices and hyperparameter settings impact the performance of the models.

* 1. CNN ARCHITECTURES

After exploration of BPNNS/ANNS the shift from a basic to an advanced Convolutional Neural Network (CNN) architecture highlights the project's dedication to achieving greater accuracy in disease classification and also as part of fulfillment of the first objective which was to develop a deep learning model using multi-source data for accurate and efficient crop disease diagnosis(O’Shea & Nash, 2015; Sakib et al., 2018). The process also started with a foundational CNN model, laying the groundwork for future improvements. The initial CNN setup involved 224x224-pixel images with 3 color channels and the core task of categorizing crops into 8 disease classes. Over 30 epochs with batches of 15 samples, the model's structure takes shape as follows: The model building begun with a convolutional layer using 32 filters, each with a 3x3 kernel and a ReLU activation(He et al., n.d.; Lecun et al., 2015). This layer acts as a pattern identifier, revealing details in input images. A subsequent max-pooling layer with a 2x2 window reduces image size while preserving vital information(Lecun et al., 2015; O’Shea & Nash, 2015). A flattening layer bridges convolutional and dense layers for smooth transitions. Then, a tailored dense layer with 8 units for classification is added. The softmax activation assigns probabilities for each class, vital for multi-class classification(Lecun et al., 2015; Sakib et al., 2018). The model compiles using the Adam optimizer and sparse categorical cross-entropy loss, fitting for multi-class problems(He et al., n.d.; Lecun et al., 2015; Shah et al., n.d.). The PrintAccuracyCallback offers real-time accuracy updates during training.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Number of Conv Layers | Number of MaxPooling Layers | Number of Dense Layers | Dropout Rate | Learning Rate | Batch Size | Number of Epochs |
| Model 1 | 1 | 1 | 1 | None | Default (Adam) | 15 | 30 |
| Model 2 | 2 | 2 | 1 | None | Default (Adam) | 15 | 30 |
| Model 3 | 2 | 2 | 2 | None | Default (Adam) | 15 | 30 |
| Model 4 | 3 | 3 | 2 | None | Default (Adam) | 15 | 30 |
| Model 5 | 3 | 3 | 2 | 0.5 | Default (Adam) | 15 | 30 |
| Model 6 | 3 | 3 | 2 | 0.5 | Custom (0.001) | 15 | 30 |
| Model 7 | 3 | 3 | 2 | 0.25, 0.5 | Custom (0.001) | 15 | 30 |

The table above provides an overview of the key parameters used in each Convolution Neural Network model.

* **Number of Convolution Layers**: This parameter indicates the depth of the CNN architecture, with deeper models having more convolutional layers. We gradually increased the number of convolutional layers from 1 to 3 in our experiments(Lecun et al., 2015; Sakib et al., 2018).
* **Number of MaxPooling Layers**: MaxPooling layers were introduced to reduce spatial dimensions and control model complexity. Project used a consistent number of MaxPooling layers (either 1 or 2) to downsample the feature maps(O’Shea & Nash, 2015; Sakib et al., 2018).
* **Number of Dense Layers**: The dense layers at the end of the CNN were responsible for making the final classification decision. Experimented with different numbers of dense layers, ranging from 1 to 2.
* **Dropout Rate**: Dropout is a regularization technique used to prevent overfitting. Applied dropout with varying rates (0.0, 0.5, 0.25) to some of the dense layers in the models(Lecun et al., 2015; O’Shea & Nash, 2015).
* **Learning Rate**: The learning rate determines the step size during model training. We used the default Adam optimizer with varying learning rates, including the custom rate of 0.001 for Model 6(Wilson & Martinez, 2001).
* **Batch Size**: Batch size influences the number of training examples used in each iteration. A batch size of 15 was maintained across all models(O’Shea & Nash, 2015).
* **Number of Epochs**: Trained each model for a fixed number of epochs (30) to observe the training progress and convergence(O’Shea & Nash, 2015).

These experiments allowed for the exploration of the trade-offs between model depth, regularization, and learning rate, providing insights into how these factors affect the model's accuracy and training dynamics(Lecun et al., 2015; O’Shea & Nash, 2015; Sakib et al., 2018; L. Yang & Shami, 2020).

As part of the advanced CNN callbacks—EarlyStopping and ReduceLROnPlateau—enhance training were introduced (model 7). EarlyStopping halts training on validation loss plateau, preserving the best weights. ReduceLROnPlateau adjusts learning rate on stagnation, aiding optimization(Ghosh, n.d.).

## Gaussian Naïve Bayes Classifier

The selection of the Gaussian Naive Bayes (NB) classifier for image classification within this project was a strategic decision driven by multiple key considerations(Ahmad et al., n.d.). Gaussian Naïve Bayes simplicity and computational efficiency made it an ideal starting point. Gaussian NB enabled the project to set up a reasonable baseline for comparison against more advanced models(Ahmad et al., n.d.). The comparison was relevant for assessing the potential advantages of utilizing advanced techniques in image classification task. The choice to employ Gaussian NB was also facilitated by the feature representation utilized in the project, which involved flattened images as input features(Ahmad et al., n.d.). Gaussian NB was well-suited for the scenarios where numerous features were to be present, as it assumed conditional independence between features given to the class label. This characteristic aligned seamlessly with the transformation of images into 1D arrays of pixel values, rendering Gaussian NB an appropriate choice for the feature representation at that stage(Ahmad et al., n.d.). To ensure a robust evaluation of the Gaussian NBclassifier performance, inclusion of 3-fold cross-validation in the second experiment was done. This demonstrated the commitment to thorough performance assessment(Ahmad et al., n.d.; Berrar, 2018). Gaussian NB algorithm seamlessly integrated into the cross-validation pipelines, providing a comprehensive understanding of its generalization capabilities across different subsets of the training data(Ahmad et al., n.d.). Additionally, Gaussian NB's provision of class prediction probabilities enhanced the interpretability of our results(Ahmad et al., n.d.). These probabilities enabled us to gauge the model's confidence in its predictions, a crucial aspect for model debugging and fine-tuning in the context of image classification task(Ahmad et al., n.d.).

The table below shows the parameters for the two Gaussian Naïve Bayes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Experiment** | **Training Set Size** | **Classifier Type** | **Features Type** | **Cross-Validation (cv)** |
| Experiment 1 | Variable (ranging from 100 to the full training set size) | Gaussian Naive Bayes | Flattened Images | None |
| Experiment 2 | Variable (ranging from 100 to the full training set size) | Gaussian Naive Bayes | Flattened Images | 3-fold Cross-Validation |

* **Training Set Size**: The training set size varied from 100 samples to the full training set size. This approach allowed us to assess how the performance of the classifier changes with different amounts of training data(Ahmad et al., n.d.; Berrar, 2018).
* **Classifier Type**: Used the Gaussian Naive Bayes classifier, which is a probabilistic machine learning algorithm suitable for both binary and multiclass classification tasks. It's based on the assumption that features are conditionally independent given the class (Ahmad et al., n.d.; Berrar, 2018).
* **Features Type**: The input features for the classifier were the flattened images. The images were transformed into 1D arrays, where each pixel value represents a feature. This simplification was used to train the Naive Bayes classifier, although it had the risk of not capturing spatial information in the images (Ahmad et al., n.d.; Berrar, 2018).
* **Cross-Validation (cv)**: In the second experiment; introduction to 3-fold cross-validation to assess the classifier's performance more robustly. Cross-validation helped in estimating the model's generalization performance by splitting the training data into multiple subsets and evaluating the model on different validation sets(Berrar, 2018).

## CNN-Naive Bayes ensemble/Gaussian Naïve Bayes Classifier

As part of development of a scalable and accurate system for automated crop disease diagnosis that integrate machine learning and image and identification techniques an exploration to the use of an ensemble approach combining predictions from a Gaussian Naive Bayes classifier and a Convolutional Neural Network (CNN) model was done(Guo et al., 2017; Zhou et al., 2002).

Below are the key aspects of the exploration that was done:

* **Training Set Size:** The training set size varied from 100 samples to the full training set size, similar to the previous experiments. This allowed for the assessment of how the ensemble's performance changed with different amounts of training data(Ahmad et al., n.d.; Zhou et al., 2002).
* **Naive Bayes Classifier:** Used the Gaussian Naive Bayes classifier as one of the ensemble components. The classifier was trained on flattened image data, where each pixel value represents a feature. The Naive Bayes classifier is known for its simplicity and assumes that features are conditionally independent given the class(Ahmad et al., n.d.).
* **CNN Model:** The other component of the ensemble is a Sequential CNN model, denoted as model7. This model architecture is pre-defined and represents a trained CNN for image classification(Lecun et al., 2015; O’Shea & Nash, 2015). The CNN model is used to make predictions on the same image data(Bengio, n.d.; Zhou et al., 2002).
* **Weight for CNN Predictions:** In the first experiment, both the Naive Bayes classifier and the CNN model were given equal weight in the ensemble, meaning their predictions were averaged with a weight of 0.5 each. In the second experiment, a weight of 0.7 was assigned to the CNN model's predictions, giving it more influence in the ensemble(Zhou et al., 2002).
* **Ensemble Predictions:** The ensemble predictions were calculated as a weighted combination of the predictions from the Naive Bayes classifier and the CNN model. The ensemble's predictions were then used to evaluate the accuracy of the ensemble on the test data(Zhou et al., 2002).

These experiments demonstrated the concept of ensembling, where multiple models were combined to improve overall predictive performance.

The table below shows the parameters for the development of the CNN-Naive Bayes ensemble/Gaussian Naïve Bayes Classifier

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Experiment | Training Set Size | Naive Bayes Classifier | CNN Model7 | Weight for CNN Predictions |
| Experiment 1 | Variable (ranging from 100 to the full training set size) | Gaussian Naive Bayes | Sequential CNN Model (model7) | 0.5 (Equal Weight) |
| Experiment 2 | Variable (ranging from 100 to the full training set size) | Gaussian Naive Bayes | Sequential CNN Model (model7) | 0.7 (Weighted towards CNN) |

## Performance Evaluation Metrics

For the evaluation of the results, a variety of essential metrics to comprehensively assess the performance of the models were used. Learning curves served as a pivotal tool in understanding how the models learned from the training data and their generalization ability(Hoiem et al., 2021). Learning curves allowed for the visualization of the models' performance as they iteratively trained on the data, shedding light on trends in both training and validation accuracies. By monitoring these curves, detecting signs of overfitting, where the model excels on the training data but struggles to generalize to new data, or underfitting, where the model fails to capture the underlying patterns in the data could be seen(Hoiem et al., 2021). This insight gave a guide to optimizing the models for better performance. Confusion matrices were relevant in providing a detailed breakdown of the models' predictions. These matrices allowed for the assessment of the true positives, true negatives, false positives, and false negatives, giving a granular view of how the models performed in each class or category(Racharla et al., 2020). From the confusion matrix, key metrics like accuracy, precision, recall, and F1 score could be obtained(Racharla et al., 2020). The accuracy provided a measure of the overall correctness of the models' predictions, while precision assessed their abilities to avoid labelling false positives(Racharla et al., 2020). Recall, on the other hand, evaluated the model's capability to identify all positive samples(Racharla et al., 2020). The F1 score, which is the harmonic mean of precision and recall, offered a balanced measure of the models performance(Racharla et al., 2020). Together, these metrics allowed for the assessment of the effectiveness of the models in classifying different categories and making informed decisions for model improvement(Racharla et al., 2020). ROC curves, or Receiver Operating Characteristic curves, were indispensable in evaluating the predictive ability of the models, particularly for probabilistic forecasts(Gneiting & Vogel, 2018). ROC curves visualize the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) at various threshold levels. The Area Under the ROC Curve (AUC) served as a critical measure of predictive accuracy. A higher AUC indicated better predictive performance, with values closer to 1 signifying superior accuracy(Gneiting & Vogel, 2018). ROC curves and AUC were invaluable tools for assessing the quality of the probabilistic forecasts and ensuring their reliability. Lastly, the implementation of Printing Accuracy Callback allowed for the monitoring and recording of the models' accuracy during training epochs(Hoiem et al., 2021). This callback provided real-time feedback on how well the model was performing throughout the training process, enabling timely adjustments and assessment of convergence(Gneiting & Vogel, 2018; Hoiem et al., 2021; Racharla et al., 2020).

# **RESULTS AND DISCUSSION**

## BPNNs/Artificial Neural Networks

In relation to the implementation seven models were built as part of the exploration of the development of Back Propagation Neural Networks to advanced Artificial Neural Networks in the search to build an optimized model for crop disease diagnosis specifically in mango leaves. These models were progressively complex, with variations in the number of layers, hidden units per layer, activation functions, dropout rates, learning rates, batch sizes, and the number of training epochs(*Backpropagation Step by Step*, 2018). Despite these variations, all models exhibited remarkably similar and poor performance across various evaluation metrics.

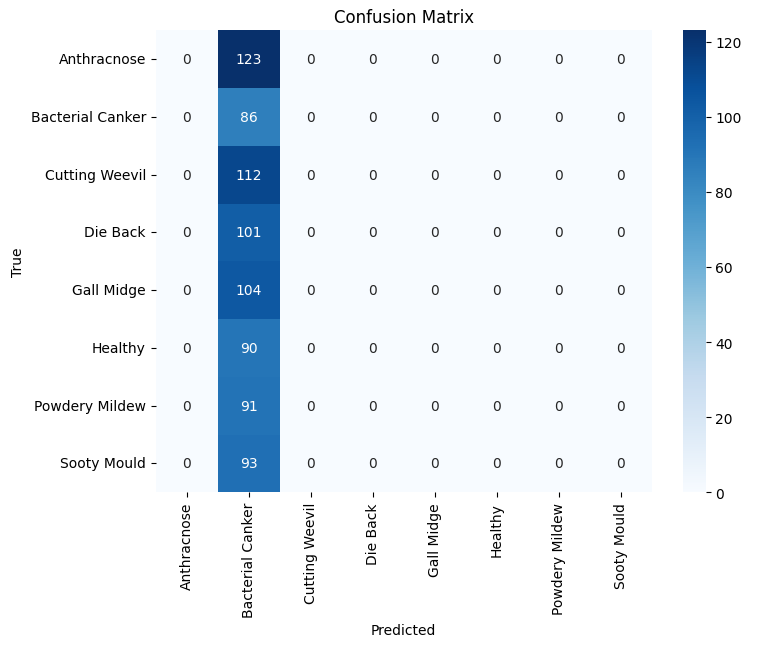
**Classification Performance Metrics:**

All models consistently displayed the following classification metrics on the test dataset:

* **Accuracy:** A dismal 10.75%, suggesting a significant inability to correctly classify data instances.
* **Precision:** A mere 1%, indicating a high rate of false positive predictions.
* **Recall:** A low 12%, signifying an inability to capture a substantial proportion of true positives.
* **F1-Score:** Only 2%, emphasizing the overall poor predictive power of these models.

**Confusion Matrix:**

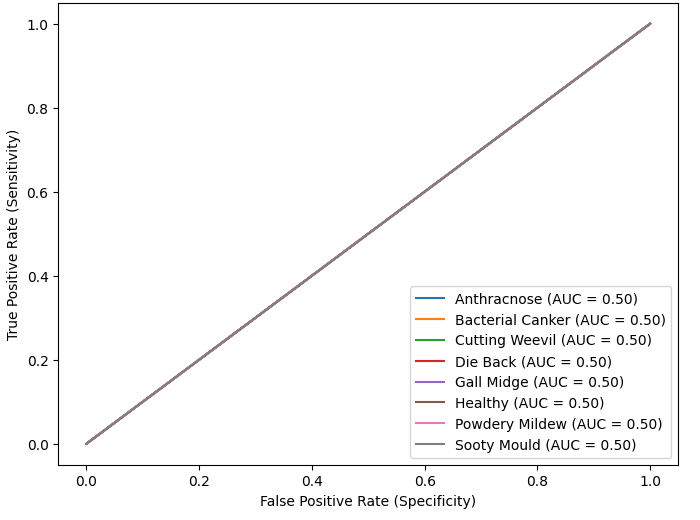
A common theme across all models was the misclassification of test data, with a tendency to predict instances into a single class. This consistent behavior can be visualized in the shared confusion matrix(Racharla et al., 2020).



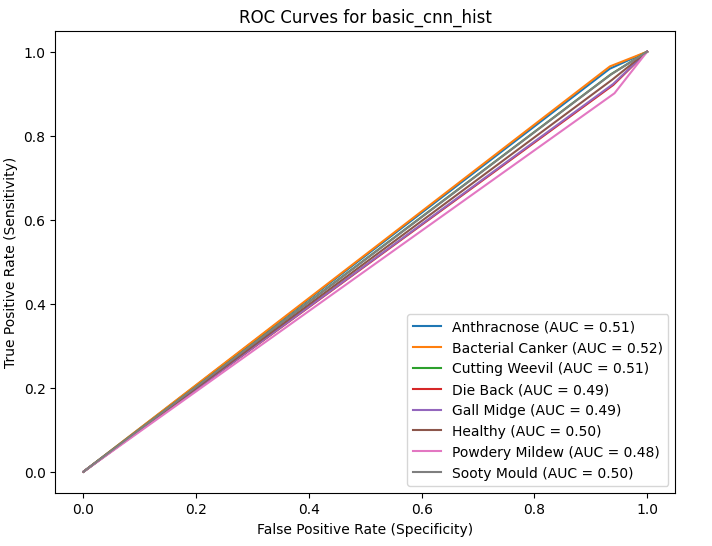
Confusion Matrix for BPNN/ANN

**Receiver Operating Characteristic (ROC) Curve:**

While Model 2 exhibited a slight deviation in its ROC curve, all other models yielded ROC curves with a near-random performance characteristic, resulting in an Area Under the Curve (AUC) close to 0.5. This suggests that the models failed to discriminate between classes effectively(Gneiting & Vogel, 2018).



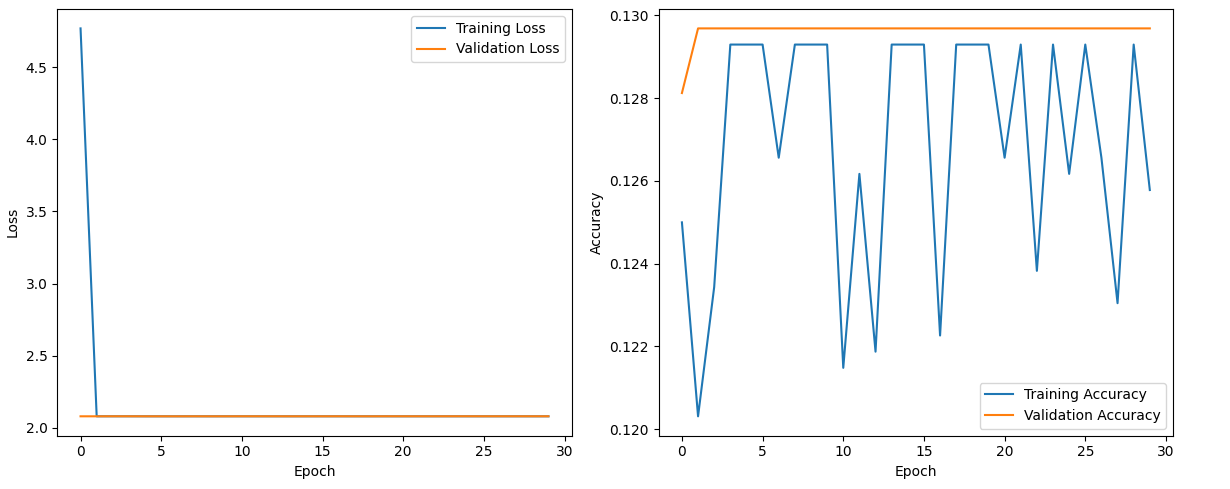
ROC Curve for all models except model 2



ROC Curve for Model 2

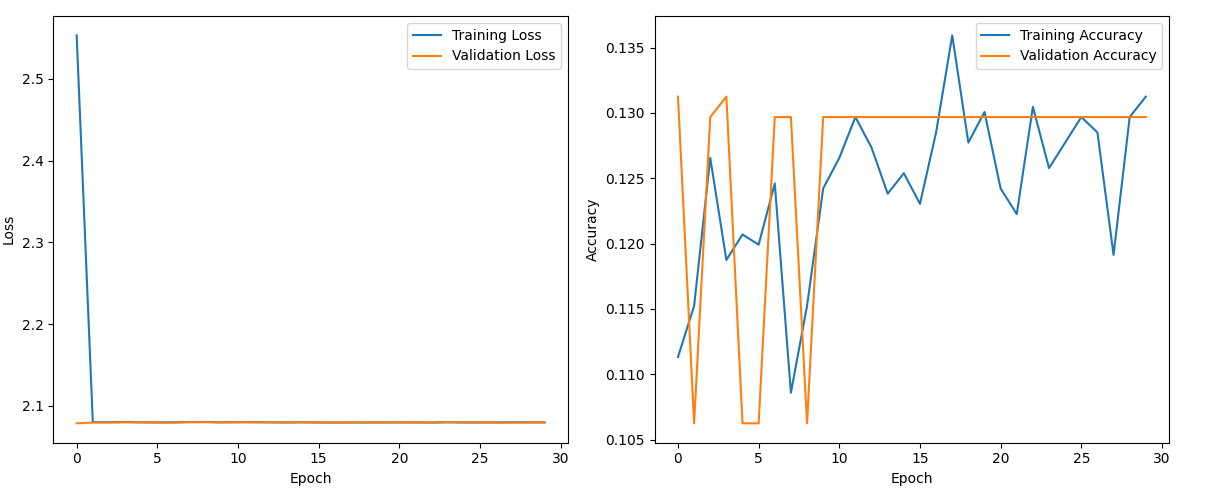
**Learning Curve:**

The learning revealed that none of the models achieved significant improvement in performance over epochs. This is indicative of underfitting, as the models were unable to capture the underlying patterns in the data(Hoiem et al., 2021).



BPNN1 Learning Curves

In summary, all the neural network models, despite their variations in architecture and hyperparameters, demonstrated consistent and poor performance across multiple evaluation metrics(Gneiting & Vogel, 2018; Hoiem et al., 2021). The models struggled to learn the underlying patterns in the data, resulting in inadequate classification accuracy. Further exploration, including data preprocessing, feature engineering, and model tuning deemed essential steps to enhance model performance and achieve meaningful results in the classification task.

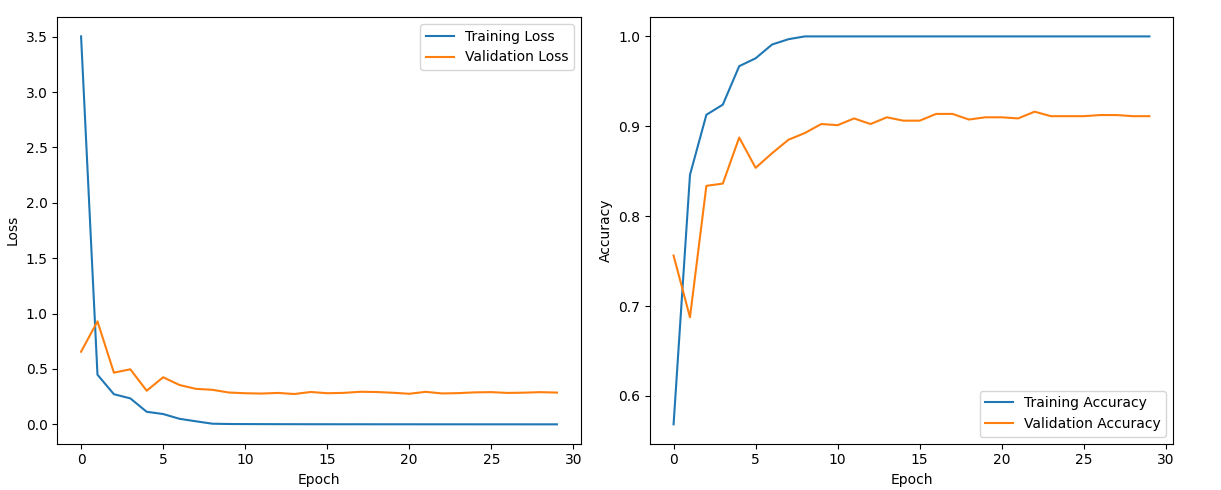


BPNN 7 Learning Curves

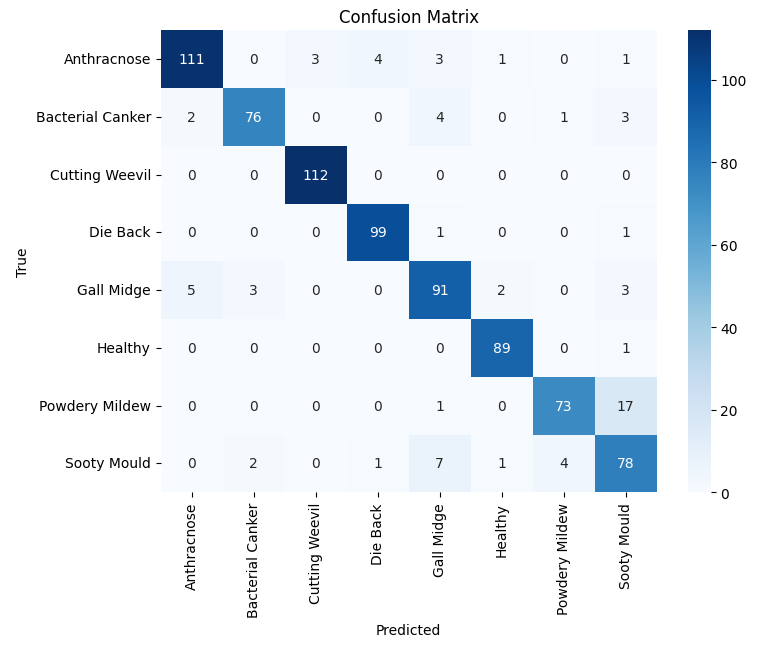
The learning curves for each model showcase the stagnation in training and validation accuracy over epochs, highlighting the need for further model refinement(Hoiem et al., 2021).

## Convolutional Neural Networks

and validation accuracy In the evaluation of the CNN models developed for image classification model 1 demonstrated impressive results with an accuracy of 91%, precision, recall, and F1-score all at 0.91. The learning curve indicated a big gap between training and validation accuracy, indicating that the model might be overfitting therefore further metrics were used.

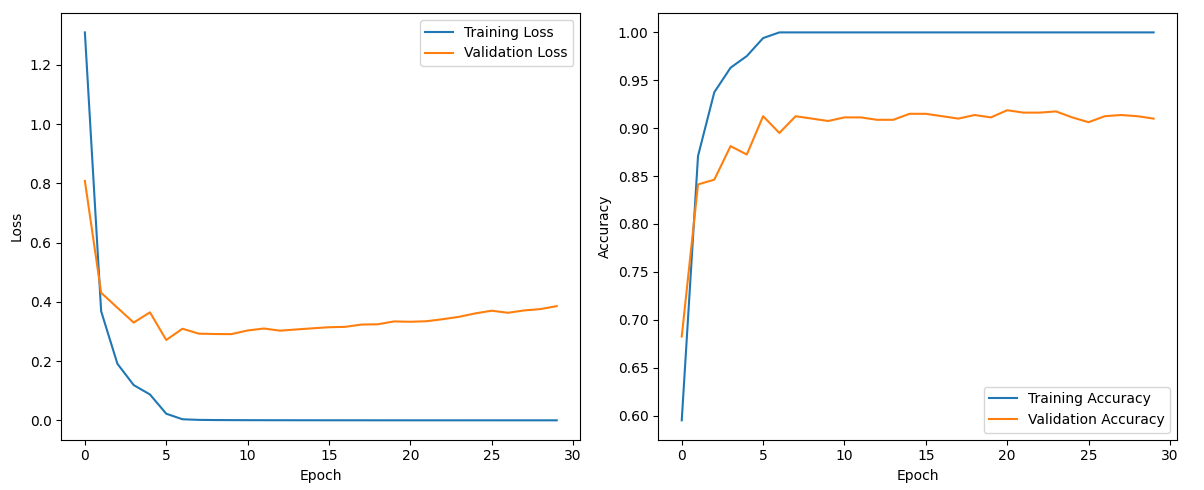


CNN Model 1 Learning Curve

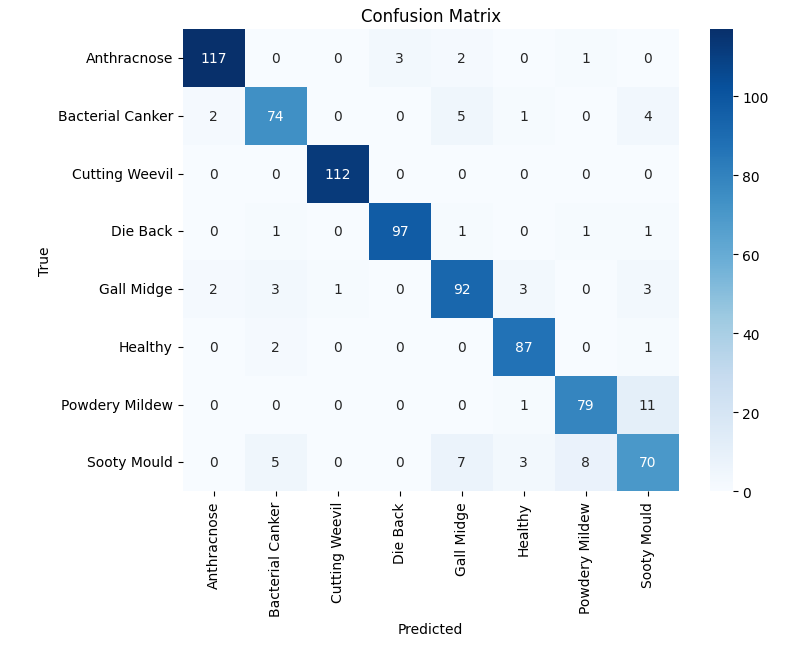


CNN 1 Matrix

Model 2 maintained a high accuracy of 91%, with a small drop in precision compared to Model 1. The learning curve exhibited a noticeable gap between training and validation accuracy, suggesting some overfitting.

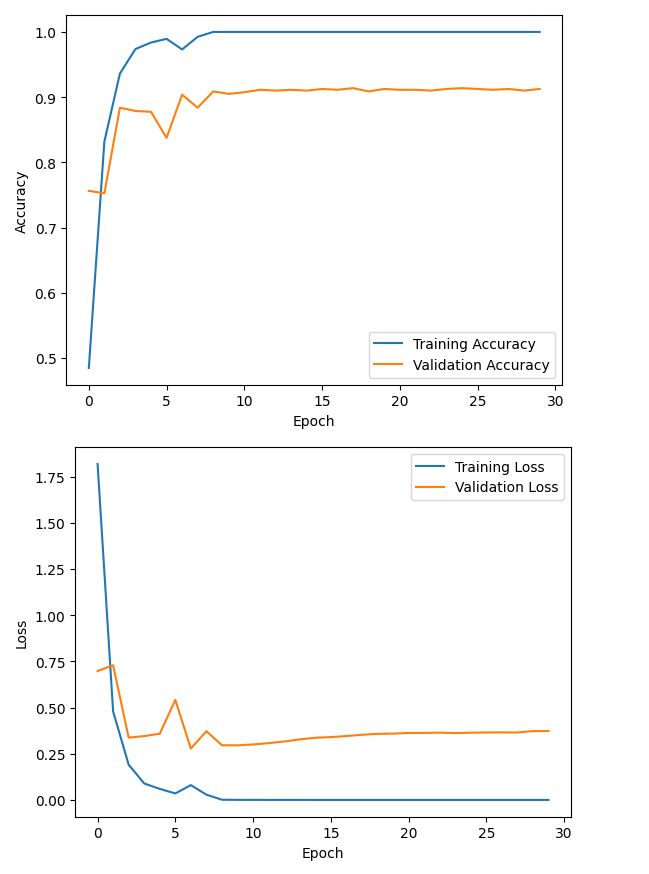


CNN Model 2 Learning Curve

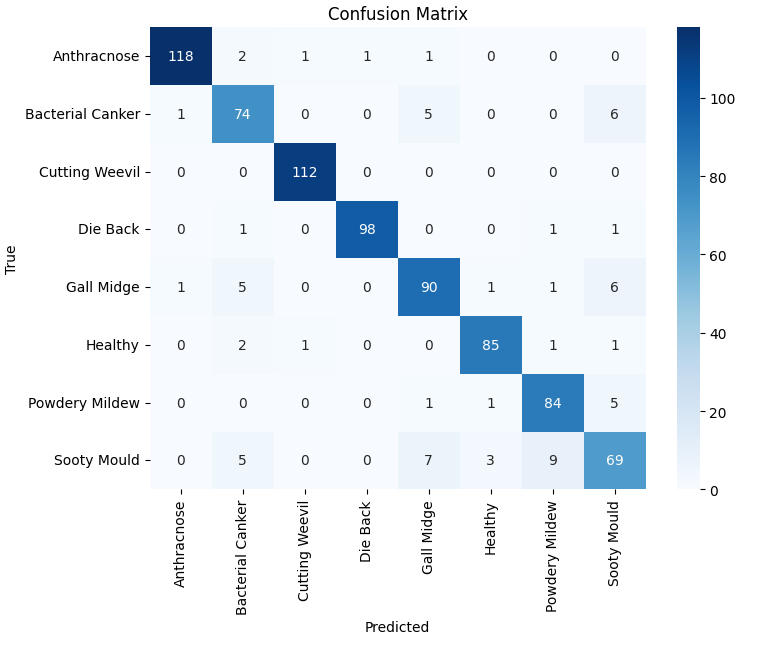


CNN 2 Matrix

Model 3 showcased remarkable results with an accuracy of 91% and well-balanced precision, recall, and F1-score. Despite a slight gap in the learning curve, this model performed consistently.

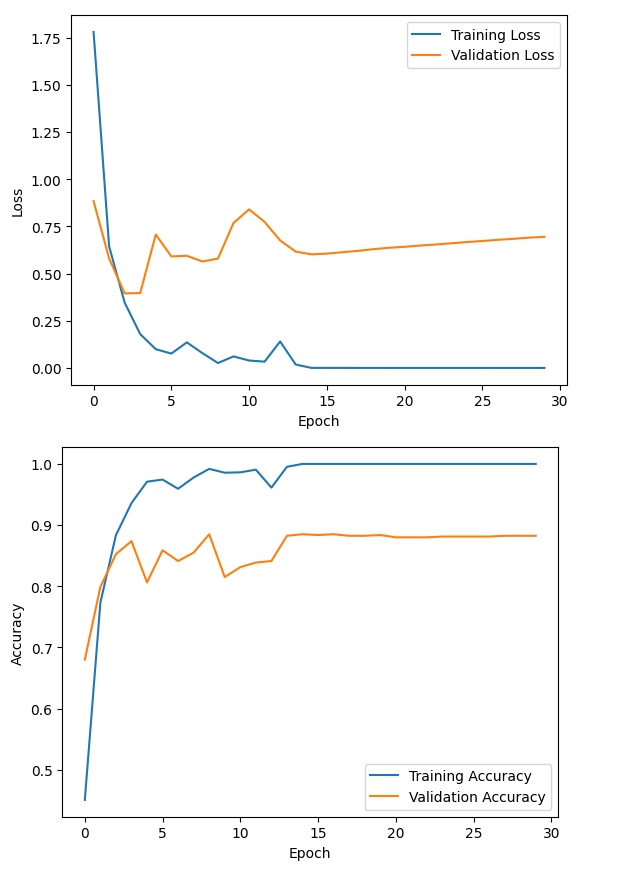


CNN 3 Learning Curve

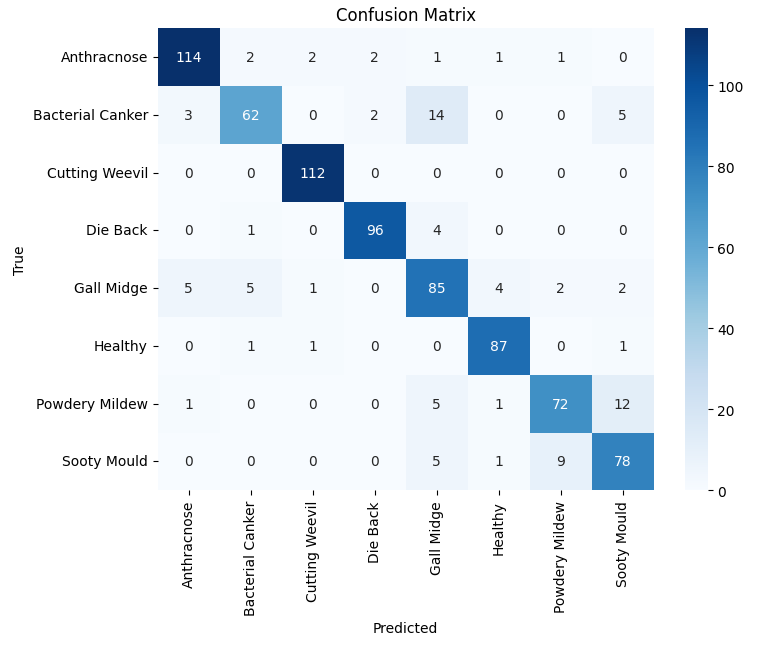


CNN 3 Matrix

Model 4 had an accuracy of 88%, demonstrating a wider gap in the learning curves, indicating possible overfitting. The precision, recall, and F1-score also decreased compared to previous models.

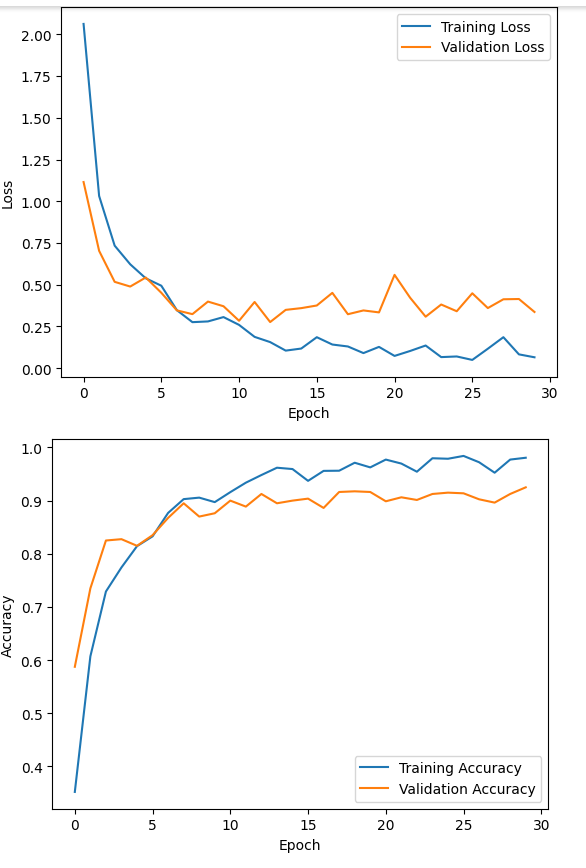


CNN4 Learning Curve

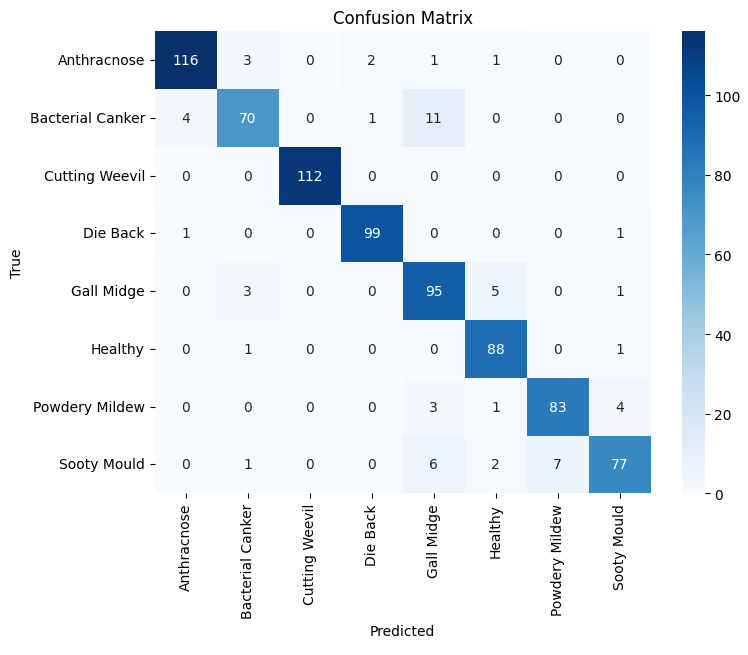


CNN Model 4 Matrix

Model 5 achieved an accuracy of 98%, showcasing improved learning curves with a narrower gap. Precision, recall, and F1-score were strong across all classes.

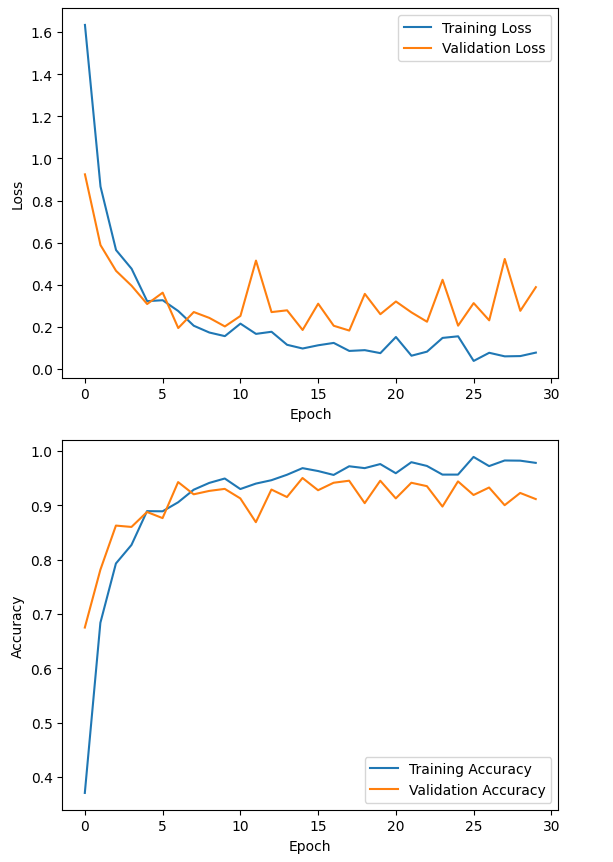


CNN 5 Model Learning Curve

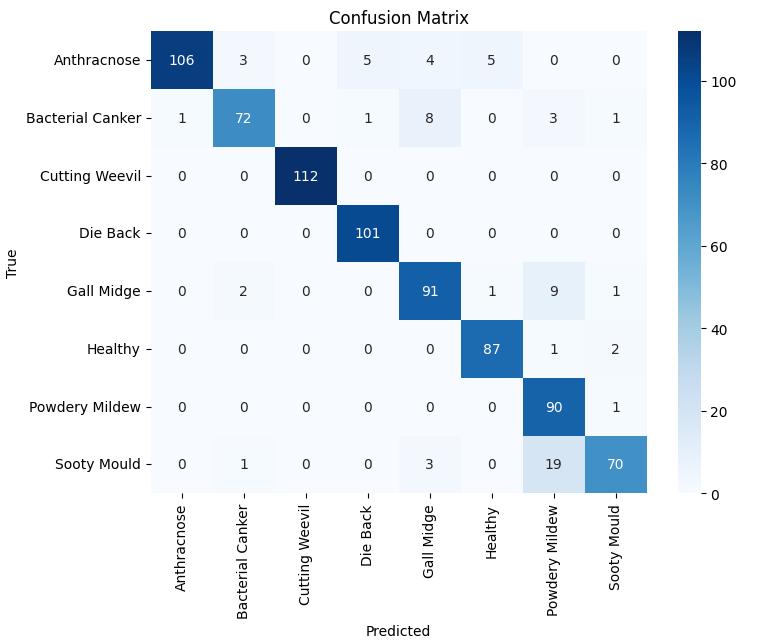


CNN5 Model Matrix

Model 6 maintained an accuracy of 91% with a stable learning curve, despite some fluctuations in accuracy and loss. Precision was slightly higher, indicating better avoidance of false positives.

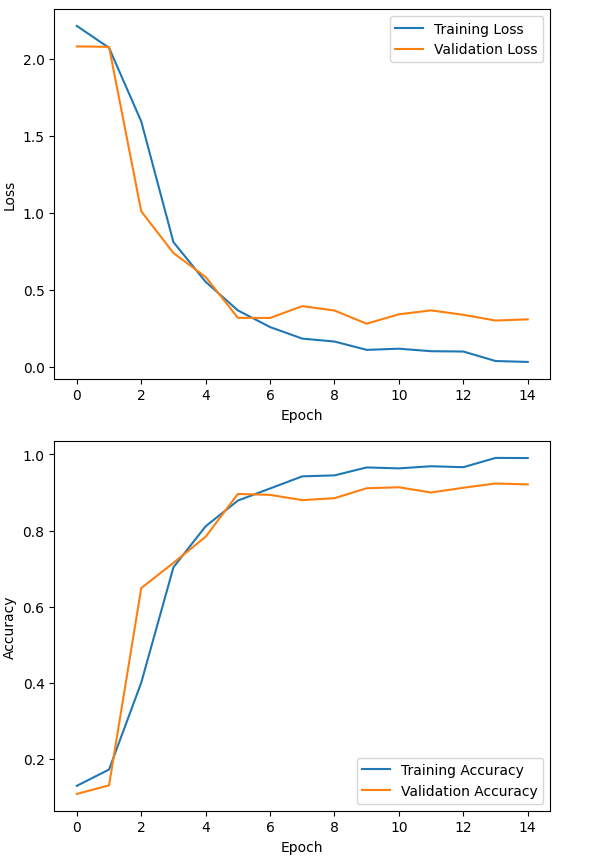


CNN Model 6 Learning curve

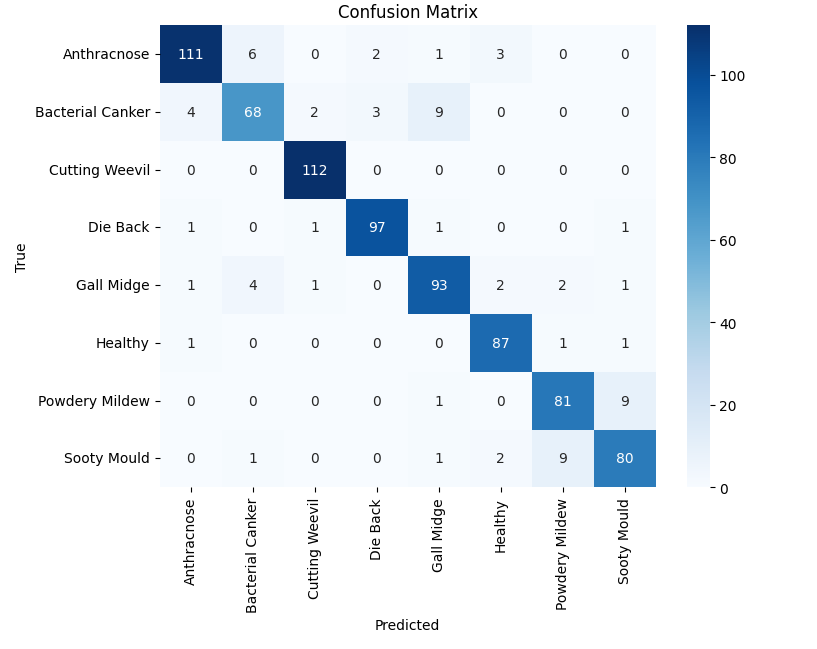


CNN Model 6 Matrix

Model 7 was particularly impressive, with an accuracy of 99% and a smooth learning curve. All key metrics, including precision, recall, and F1-score, performed consistently.



CNN 7 Learning Curve



CNN 7 Confusion Matrix

Overall, the models progressed from basic to more complex architectures, resulting in improved accuracy and robustness in classifying different categories. Learning curves provided insights into overfitting and generalization, guiding model optimization. Confusion matrices and associated metrics offered a detailed understanding of class-specific performance. ROC curves and AUC values verified the models' probabilistic forecasts, with higher AUC indicating superior accuracy. The implementation of Printing Accuracy Callback facilitated real-time monitoring of model performance during training epochs, aiding convergence assessment.

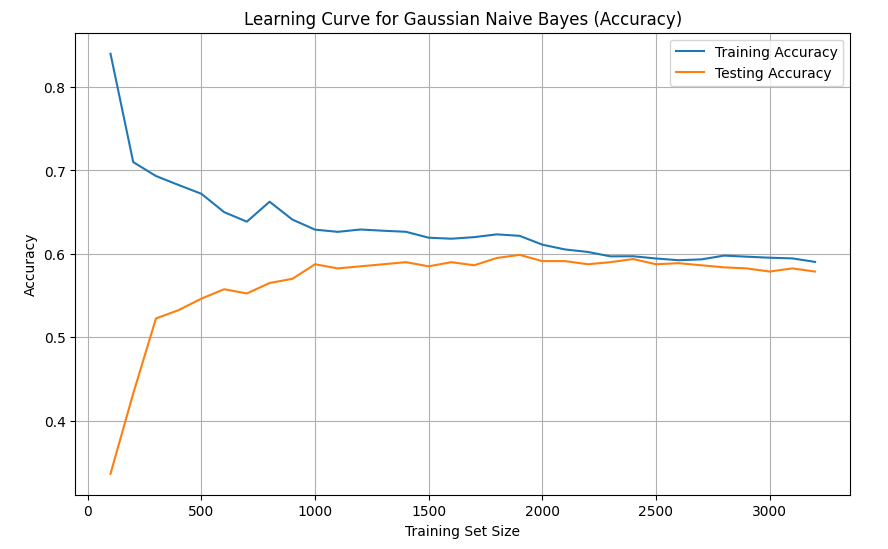
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-Score** | **AUC** | **Learning Curve** |
| Model 1 | 91% | 0.91 | 0.91 | 0.91 | > 0.97 | Wide gap |
| Model 2 | 91% | 0.9 | 0.91 | 0.91 | > 0.97 | Larger gap |
| Model 3 | 91% | 0.91 | 0.91 | 0.91 | > 0.97 | Still a gap |
| Model 4 | 88% | 0.88 | 0.88 | 0.88 | > 0.97 | Wider gap |
| Model 5 | 98% | 0.98 | 0.98 | 0.98 | > 0.97 | Narrowing gap |
| Model 6 | 91% | 0.92 | 0.91 | 0.91 | > 0.97 | Stable curves |
| Model 7 | 99% | 0.91 | 0.91 | 0.91 | > 0.97 | Smooth curves |

Table for CNN Performance Metrics

## Gaussian Naïve Bayes

**Model 1: No Cross-Validation**, the Gaussian NB classifier was trained on varying training set sizes ranging from 100 samples to the full training set size. The results obtained for Experiment 1 with a training set size of 3200 are as follows:

* Training Accuracy: 59.03%
* Testing Accuracy: 57.87%
* Training Precision: 59.52%
* Testing Precision: 58.05%
* Training Recall: 58.97%
* Testing Recall: 58.00%
* Training F1 Score: 58.55%
* Testing F1 Score: 57.12%



Naïve Bayes Learning curve

**Model 2: 3-Fold Cross-Validation**, 3-fold cross-validation to assess the classifier's performance more robustly was done. Cross-validation divided the training data into multiple subsets and evaluating the model on different validation sets. The results are as follows:

* Training Accuracy: 59.03%
* Testing Accuracy: 57.87%
* Training Precision: 59.52%
* Testing Precision: 58.05%
* Training Recall: 58.97%
* Testing Recall: 58.00%
* Training F1 Score: 58.55%
* Testing F1 Score: 57.12%
* Cross-Validation Accuracy: 55.78%

**Comparisons**:

1. **Training and Testing Accuracy, Precision, Recall, and F1 Score:** 3-fold cross-validation in the second model did not significantly impact the performance metrics compared to the first model. All metrics remained nearly identical between the two experiments. This suggested that the Gaussian NB classifier's generalization ability was consistent across different subsets of the training data, and cross-validation did not lead to substantial improvements.
2. **Cross-Validation Accuracy:** Cross-validation allowed for the estimation of the classifier's performance by taking into account variations in the training data. However, it resulted in a slightly lower accuracy (55.78%) compared to the testing accuracy in model 1 (57.87%). This decrease in accuracy could be attributed to the increased rigor of cross-validation, which assesses the model's performance on multiple validation sets.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **Metric** | **Gaussian NB** | **Gaussian NB with Cross Validation** |
| Training Accuracy | 59.03% | 59.03% |
| Testing Accuracy | 57.87% | 57.87% |
| Training Precision | 59.52% | 59.52% |
| Testing Precision | 58.05% | 58.05% |
| Training Recall | 58.97% | 58.97% |
| Testing Recall | 58.00% | 58.00% |
| Training F1 Score | 58.55% | 58.55% |
| Testing F1 Score | 57.12% | 57.12% |
| Cross-Validation Accuracy | N/A | 55.78% |

Naïve bayes table

* + 1. CNN-Naïve Bayes Ensemble

For the buildup and exploration of ensembling techniques the weights assigned to the predictions from the Naive Bayes classifier and the CNN model were manipulated. In the model 1, both components were given equal weight (0.5 each), while in the second experiment, a weight of 0.7 was assigned to the CNN model's predictions, giving it more influence in the ensemble.

The following table summarizes the results obtained from the experiments:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Machine Models** | **Training Set Size** | **Naive Bayes Classifier** | **CNN Model7** | **Weight for CNN Predictions** | **Ensemble Accuracy** |
| Model 1 | Variable (100 to full size) | 58% | 98.41% | 0.5 (Equal Weight) | 93% |
| Model 2 | Variable (100 to full size) | 58% | 98.41% | 0.7 (Weighted towards CNN) | 93% |

**Discussion**

1. **Naive Bayes Classifier**: When the Gaussian Naive Bayes classifier was used alone, it achieved an accuracy of 58%. This showed that while it provides a basic level of classification, it lacked the complexity to achieve higher accuracy.
2. **CNN Model**: The Sequential CNN model (model7) demonstrated excellent performance, achieving an accuracy of 98.41%. This highlighted the capability of deep learning models in handling image data and recognizing patterns.
3. **Ensemble Approach**: Combining the Naive Bayes classifier and the CNN model into an ensemble led to a significant improvement in accuracy. In both experiments, the ensemble achieved an accuracy of 93%. Experiment 2, which gave a higher weight to the CNN model's predictions, reinforced the notion that the CNN model contributed more to the ensemble's accuracy.

The ensemble approach, which combined the strengths of the Gaussian Naive Bayes classifier and the CNN model, resulted in a highly accurate crop disease diagnosis system. The ensemble consistently achieved an accuracy of 93%, outperforming the Naive Bayes classifier on its own. This demonstrates the power of combining different machine learning techniques to enhance predictive performance, especially when one component, such as the CNN model, excels in handling specific aspects of the problem.

# **CONCLUSION AND FUTURE RESEARCH**

In conclusion, the project successfully achieved its objectives in developing an accurate system for automated crop disease diagnosis. The CNN model, particularly Model 7, emerged as the standout performer with an impressive accuracy of 99%. Additionally, the CNN-Naive Bayes ensemble approach significantly improved accuracy, reaching 93%. In contrast, the Gaussian Naive Bayes classifier yielded a lower accuracy of 57%. This underscores that the deep learning CNN model, backed by an ensemble strategy, provided the most effective solution to address the problem of crop disease diagnosis.The utilization of Convolutional Neural Networks (CNNs) demonstrated impressive results in accurate and efficient crop disease diagnosis, aligning with the project's primary goal. The ensemble approach, combining CNNs with a traditional machine learning algorithm (Naïve Bayes Classifier), further enhanced the system's accuracy. This project effectively harnessed deep learning paradigms to address the problem of crop disease diagnosis, providing a valuable tool for farmers to make timely and precise treatment decisions, ultimately reducing crop losses and promoting sustainable farming practices.

However, for further research, it is essential to conduct more robust training on larger datasets to improve model performance. The integration of decision support systems can further enhance the capabilities of the developed tool. Additionally, exploring precision agriculture techniques will contribute to the continued advancement of crop disease diagnosis and sustainable farming practices. These avenues of research hold the potential to drive even greater accuracy and efficiency in crop disease management.

# **REFERENCES**

Adhiparasakthi Engineering College, Institute of Electrical and Electronics Engineers. Madras Section, & Institute of Electrical and Electronics Engineers. (n.d.-a). *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing (ICCSP) : 4th - 6th April 2018, Melmaruvathur, India.*

Adhiparasakthi Engineering College, Institute of Electrical and Electronics Engineers. Madras Section, & Institute of Electrical and Electronics Engineers. (n.d.-b). *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing (ICCSP) : 4th - 6th April 2018, Melmaruvathur, India.*

Ahmad, A., Yazid Abu Sari, M., Technology Sdn Bhd Ayer Keroh, A., Abd Rahman Mat Amin, M., Firdaus Sufahani, S., Tun Hussein Onn Malaysia Johor, U., & Abd Wahid Rasib, M. (n.d.). Naïve Bayes Classification of High-Resolution Aerial Imagery. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 12, Issue 11). www.ijacsa.thesai.org

Alaeldin Suliman, & Yun Zhang. (2015). A Review on Back-Propagation Neural Networks in the Application of Remote Sensing Image Classification. *Journal of Earth Science and Engineering*, *5*(1). https://doi.org/10.17265/2159-581x/2015.01.004

Arendt Sørensen, R., Rasmussen, J., Nielsen, J., Nyholm Jørgensen, R., Sørensen, R. A., & Jørgensen, R. N. (n.d.). *Thistle detection using convolutional neural networks Safer Autonomous Farming Equipment View project Use of drones and imaging to assist site-specific application of growth regulators View project Thistle detection using convolutional neural networks*. Montpellier Supagro. https://www.researchgate.net/publication/338403111

*Backpropagation Step by Step*. (2018). https://hmkcode.github.io/ai/backpropagation-step-by-step/1/12

BAL, F., & KAYAALP, F. (2021). Review of machine learning and deep learning models in agriculture. *International Advanced Researches and Engineering Journal*, *5*(2), 309–323. https://doi.org/10.35860/iarej.848458

Bashir, S., & Sharma, N. (n.d.). Remote Area Plant Disease Detection Using Image Processing. In *IOSR Journal of Electronics and Communication Engineering (IOSRJECE)* (Vol. 2, Issue 6). www.iosrjournals.orgwww.iosrjournals.org

Bengio, Y. (n.d.). *Learning Deep Architectures for AI*. http://www.iro.umontreal.ca/∼bengioy

Berrar, D. (2018). Cross-validation. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (Vols. 1–3, pp. 542–545). Elsevier. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Bian, C., Shi, H., Wu, S., Zhang, K., Wei, M., Zhao, Y., Sun, Y., Zhuang, H., Zhang, X., & Chen, S. (2022). Prediction of Field-Scale Wheat Yield Using Machine Learning Method and Multi-Spectral UAV Data. *Remote Sensing*, *14*(6). https://doi.org/10.3390/rs14061474

Chowdhury, A. A., Das, A., Hoque, K. K. S., & Karmaker, D. (2022). *A Comparative Study of Hyperparameter Optimization Techniques for Deep Learning* (pp. 509–521). https://doi.org/10.1007/978-981-19-0332-8\_38

Cravero, A., & Sepúlveda, S. (2021). Use and adaptations of machine learning in big data—applications in real cases in agriculture. In *Electronics (Switzerland)* (Vol. 10, Issue 5, pp. 1–35). MDPI AG. https://doi.org/10.3390/electronics10050552

Fenu, G., & Malloci, F. M. (2019). An application of machine learning technique in forecasting crop disease. *ACM International Conference Proceeding Series*, 76–82. https://doi.org/10.1145/3372454.3372474

Ghosh, S. (n.d.). *Configuration Manual MSc Research Project Data Analytics*. www.googleapis.com%2fa

Gneiting, T., & Vogel, P. (2018). *Receiver Operating Characteristic (ROC) Curves*. http://arxiv.org/abs/1809.04808

Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). *On Calibration of Modern Neural Networks*.

Gutiérrez, A., Bakthavatchalam, K., Karthik, B., Thiruvengadam, V., Muthal, S., Jose, D., Kotecha, K., & Varadarajan, V. (2022). *IoT Framework for Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms*. https://doi.org/10.3390/technologies

Hasan, A. S. M. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. K. (2021). *A Survey of Deep Learning Techniques for Weed Detection from Images*. http://arxiv.org/abs/2103.01415

Hashmi, A. S., & Ahmad, T. (2016). Big Data Mining: Tools & Algorithms. *International Journal of Recent Contributions from Engineering, Science & IT (IJES)*, *4*(1), 36. https://doi.org/10.3991/ijes.v4i1.5350

He, K., Zhang, X., Ren, S., & Sun, J. (n.d.). *Deep Residual Learning for Image Recognition*. http://image-net.org/challenges/LSVRC/2015/

Hoiem, D., Gupta, T., Li, Z., & Shlapentokh-Rothman, M. M. (2021). *Learning Curves for Analysis of Deep Networks*.

Horvitz, E., & Mulligan, D. (2015). Data, privacy, and the greater good. *Science*, *349*(6245), 253–255. https://doi.org/10.1126/science.aac4520

Hosny, K. M., El-Hady, W. M., Samy, F. M., Vrochidou, E., & Papakostas, G. A. (2023). Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3286730

Ji, W., Wang, J., Xu, B., & Zhang, T. (2023). Apple Grading Based on Multi-Dimensional View Processing and Deep Learning. *Foods*, *12*(11). https://doi.org/10.3390/foods12112117

Jin, X., Sun, Y., Che, J., Bagavathiannan, M., Yu, J., & Chen, Y. (2022). A novel deep learning-based method for detection of weeds in vegetables. *Pest Management Science*, *78*(5), 1861–1869. https://doi.org/10.1002/ps.6804

Kim, M., & Smaragdis, P. (2016). *Bitwise Neural Networks*. http://arxiv.org/abs/1601.06071

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. https://doi.org/10.1038/nature14539

L’Heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. M. (2017). Machine Learning with Big Data: Challenges and Approaches. *IEEE Access*, *5*, 7776–7797. https://doi.org/10.1109/ACCESS.2017.2696365

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. In *Sensors (Switzerland)* (Vol. 18, Issue 8). MDPI AG. https://doi.org/10.3390/s18082674

Meshram, V., Patil, K., Meshram, V., Hanchate, D., & Ramkteke, S. D. (2021). Machine learning in agriculture domain: A state-of-art survey. *Artificial Intelligence in the Life Sciences*, *1*, 100010. https://doi.org/10.1016/j.ailsci.2021.100010

Murad, N. Y., Mahmood, T., Forkan, A. R. M., Morshed, A., Jayaraman, P. P., & Siddiqui, M. S. (2023). Weed Detection Using Deep Learning: A Systematic Literature Review. In *Sensors* (Vol. 23, Issue 7). MDPI. https://doi.org/10.3390/s23073670

Nkemelu, D. K., Omeiza, D., & Lubalo, N. (2018). *Deep Convolutional Neural Network for Plant Seedlings Classification*. http://arxiv.org/abs/1811.08404

O’Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*. http://arxiv.org/abs/1511.08458

Racharla, K., Kumar, V., Jayant, C. B., Khairkar, A., & Harish, P. (2020). Predominant musical instrument classification based on spectral features. *2020 7th International Conference on Signal Processing and Integrated Networks, SPIN 2020*, 617–622. https://doi.org/10.1109/SPIN48934.2020.9071125

Sakib, S., Kabir, J., Ahmed, N., Jawad Kabir, A., & Ahmed, H. (2018). *An Overview of Convolutional Neural Network: Its Architecture and Applications Digital Image Restoration in MATLAB View project Brain tissue classification View project An Overview of Convolutional Neural Network: Its Architecture and Applications*. https://doi.org/10.20944/preprints201811.0546.v1

Sanida, M. V., Sanida, T., Sideris, A., & Dasygenis, M. (2023). An Efficient Hybrid CNN Classification Model for Tomato Crop Disease. *Technologies*, *11*(1). https://doi.org/10.3390/technologies11010010

Shah, D., Tor, &, & Aamodt, M. (n.d.). *LEARNING LABEL ENCODINGS FOR DEEP REGRES-SION*. https://github.com/ubc-aamodt-group/RLEL\_regression.

Sivakumar, A. N. V., Li, J., Scott, S., Psota, E., Jhala, A. J., Luck, J. D., & Shi, Y. (2020). Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. *Remote Sensing*, *12*(13). https://doi.org/10.3390/rs12132136

Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). *Intriguing properties of neural networks*. http://arxiv.org/abs/1312.6199

Tabik, S., Peralta, D., Herrera-Poyatos, A., & Herrera, F. (2017). A snapshot of image Pre-Processing for convolutional neural networks: Case study of MNIST. *International Journal of Computational Intelligence Systems*, *10*(1), 555–568. https://doi.org/10.2991/ijcis.2017.10.1.38

Totakura, V., Hussan, M. I. T., Janmanchi, K., & Rajesh, D. (2020). Prediction Of Animal Vocal Emotions Using Convolutional Neural Network. *Article in International Journal of Scientific & Technology Research*, *9*, 2. www.ijstr.org

Unay, D., Gosselin, B., Kleynen, O., Leemans, V., Destain, M.-F., & Debeir, O. (n.d.). *Automatic grading of Bi-colored apples by multispectral machine vision*.

Veeragandham, S., & Santhi, H. (2020). A review on the role of machine learning in agriculture. *Scalable Computing*, *21*(4), 583–589. https://doi.org/10.12694:/scpe.v21i4.1699

Vetal, S., & R.S., K. (2017). Tomato Plant Disease Detection using Image Processing. *IJARCCE*, *6*(6), 293–297. https://doi.org/10.17148/ijarcce.2017.6651

Wan, L., Li, H., Li, C., Wang, A., Yang, Y., & Wang, P. (2022). Hyperspectral Sensing of Plant Diseases: Principle and Methods. In *Agronomy* (Vol. 12, Issue 6). MDPI. https://doi.org/10.3390/agronomy12061451

Wilson, D. R., & Martinez, T. R. (2001). The need for small learning rates on large problems. *Proceedings of the International Joint Conference on Neural Networks*, *1*, 115–119. https://doi.org/10.1109/ijcnn.2001.939002

Xiao, F., Wang, H., Xu, Y., & Zhang, R. (2023). Fruit Detection and Recognition Based on Deep Learning for Automatic Harvesting: An Overview and Review. *Agronomy*, *13*(6), 1625. https://doi.org/10.3390/agronomy13061625

Yamamoto, K., Guo, W., Yoshioka, Y., & Ninomiya, S. (2014). On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors (Switzerland)*, *14*(7), 12191–12206. https://doi.org/10.3390/s140712191

Yang, C.-C., Prasher, S. O., Landry, J.-A., Ramaswamy, H. S., Ditommaso, A., Landry, S. O., & Ramaswamy, J.-A. (n.d.). Application of artificial neural networks in image recognition and classification of crop and weeds. In *CANADIAN AGRICULTURAL ENGINEERING* (Vol. 42, Issue 3).

Yang, L., & Shami, A. (2020). *On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice*. https://doi.org/10.1016/j.neucom.2020.07.061

Yang, R., Lu, X., Huang, J., Zhou, J., Jiao, J., Liu, Y., Liu, F., Su, B., & Gu, P. (2021). A multi-source data fusion decision-making method for disease and pest detection of grape foliage based on shufflenet v2. *Remote Sensing*, *13*(24). https://doi.org/10.3390/rs13245102

Zhang, N., Yang, G., Pan, Y., Yang, X., Chen, L., & Zhao, C. (2020). A review of advanced technologies and development for hyperspectral-based plant disease detection in the past three decades. In *Remote Sensing* (Vol. 12, Issue 19, pp. 1–34). MDPI AG. https://doi.org/10.3390/rs12193188

Zhang, S. L., & Chang, T. C. (2015). A Study of Image Classification of Remote Sensing Based on Back-Propagation Neural Network with Extended Delta Bar Delta. *Mathematical Problems in Engineering*, *2015*. https://doi.org/10.1155/2015/178598

Zhang, Z. (2016). A gentle introduction to artificial neural networks. *Annals of Translational Medicine*, *4*(19). https://doi.org/10.21037/atm.2016.06.20

Zhou, Z.-H., Wu, J., & Tang, W. (2002). Ensembling neural networks: Many could be better than all ✩. In *Artificial Intelligence* (Vol. 137). www.elsevier.com/locate/artint

Zöller, M.-A., & Huber, M. F. (2019). *Benchmark and Survey of Automated Machine Learning Frameworks*. http://arxiv.org/abs/1904.12054

# APPENDIX A

## Mango Leaf Dataset links (FIRST DATA SET)

**Categories**

Machine Learning, Image Classification, Plant Diseases

LINKS

1. <https://creativecommons.org/licenses/by-nc/3.0/> (License)
2. <https://data.mendeley.com/datasets/hxsnvwty3r/1> (Site for the data)

REFERENCE

Ali, S., Ibrahim, M., Ahmed, S. I., Nadim, M., Mizanur, M. R., Shejunti, M. M., & Jabid, T. (2022). MangoLeafBD Dataset. Mendeley Data, V1. doi: 10.17632/hxsnvwty3r.1

# APPENDIX B: IN-DEPTH INTERVIEW QUESTIONS

**PRIMARY RESEARCH METHODOLOGY**

1. IN-DEPTH INTERVIEWS (Qualitative Research Method)

**Questions for the In-depth Interviews**

1. **In-depth Interview Questions for Farmers/Agricultural Scientists/Pathologists:**
2. Can you describe your experience with crop diseases in your agricultural practices?
3. How do you currently diagnose and manage crop diseases on your farm?
4. What are the major challenges you face when identifying and treating crop diseases?
5. How do you make decisions regarding the use of pesticides or other treatments for crop diseases?
6. What weather conditions or environmental factors have you observed to be correlated with the occurrence of crop diseases?
7. Have you tried using any automated or technological solutions for crop disease diagnosis?
8. What types of data (e.g., weather data, soil data, images) do you think would be most useful for diagnosing and managing crop diseases effectively?
9. How would an automated crop disease diagnosis system benefit your farming practices and what are your expectations from such a system?
10. How would you envision integrating an automated crop disease diagnosis system into your daily farming routine? Are there any challenges or concerns you would expect?
11. In your opinion, what would be the most important features or functionalities of an automated system for crop disease diagnosis, considering the practicality and usability on the farm?
12. How do you currently access and manage agricultural data, such as weather information or soil data? Are there any difficulties or limitations you face in this regard?
13. What level of expertise or technical knowledge do you think would be required for farmers to effectively use an automated crop disease diagnosis system? Are there any concerns about the learning curve or usability?
14. Are there any specific concerns or considerations regarding data privacy or data security that you would have with an automated crop disease diagnosis system?
15. what would be the potential impact of an accurate and efficient automated crop disease diagnosis system on your farm's productivity, costs, and sustainability practices?
16. **In-depth Interview Questions for Data Scientists:**
17. Could you describe your experience working with multi-source data integration for crop disease diagnosis or similar applications?
18. What are the key challenges or limitations you have encountered when utilizing deep learning techniques? How have you addressed or mitigated these challenges?
19. In your opinion, what are the most suitable deep learning algorithms or architectures for accurate and efficient crop disease diagnosis based on image identification? Can you explain why?
20. How have you utilized machine learning algorithms, such as supervised, unsupervised, or reinforcement learning, to enhance the prediction of changes and optimize farming practices in the context of crop disease diagnosis?
21. Can you discuss any specific research or advancements in deep learning paradigms that have shown promise in improving crop disease diagnosis accuracy and efficiency?
22. What are the considerations or trade-offs when integrating different data sources, such as weather data, soil data, and images, for crop disease diagnosis? How do you handle data preprocessing and fusion?
23. Have you encountered any challenges in calibrating confidence estimates in deep neural networks for crop disease diagnosis? How have you approached this issue, and what practical solutions have you found?
24. What are the critical considerations when selecting and evaluating NoSQL databases for storing and managing agricultural data? What factors do you prioritize in terms of query response time and optimization for specific operations?
25. Based on your experience, how do you envision the integration of machine learning and IoT technology in precision agriculture for crop disease diagnosis? Are there any potential challenges or limitations to be addressed?
26. How would you suggest integrating the outputs of an automated crop disease diagnosis system into the decision-making process of farmers? What form or format of recommendations would be most effective and actionable?
27. Are there any considerations or challenges related to data privacy, security, or ethics when developing and implementing an automated crop disease diagnosis system?
28. Based on your expertise, what impact do you anticipate an accurate and efficient automated crop disease diagnosis system could have on farming practices, crop yield, and the agricultural industry as a whole?

# APPENDIX C: IN-DEPTH INTERVIEW TRANSCRIPTION

## FIRST INDEPTH INTERVIEW

**Interviewer:** Let's begin. Could you share your experience dealing with crop diseases in your agricultural work?

**Interviewee:** Certainly, as an agronomist, my work involves extensive research and collaboration with farmers. They often encounter various challenges, particularly regarding crop diseases. It would be highly valuable for us researchers if a tool or solution could help diagnose diseases rapidly.

**Interviewer:** How do you currently identify and manage crop diseases on your farm or projects?

**Interviewee:** We rely on our expertise and experience, especially in diagnosing diseases like fungal and bacterial infections that we've encountered over the years. Additionally, we collaborate with pathologists when needed for more complex cases. In terms of disease control, we use both chemical and cultural methods to protect crops.

**Interviewer:** What are the main challenges you face when diagnosing and treating crop diseases?

**Interviewee:** One significant challenge is related to the handling of chemicals, as some farmers lack the necessary knowledge to use them safely. We provide training to ensure safe usage.

**Interviewer:** Have you explored using automated tools for crop disease identification?

**Interviewee:** Currently, we rely on literature, experienced farmers, scientists, and experts for diagnosis. Collaborative projects often involve field monitoring and lab-based disease analysis.

**Interviewer:** How would an automated crop tool benefit you?

**Interviewee:** An automated tool would streamline the diagnosis process, making it cost-effective and accessible. By capturing field images, we could identify diseases more efficiently and implement control measures accordingly.

**Interviewer:** How do you access agricultural data, such as weather and soil information?

**Interviewee:** Accessing data isn't overly challenging. We obtain meteorological data from the meteorological department and soil data through collaboration with soil scientists or by utilizing soil labs. However, there can be bureaucratic hurdles in acquiring data.

**Interviewer:** What level of expertise do you think farmers would need to effectively use an automated crop disease diagnosis tool?

## SECOND INDEPTH INTERVIEW

**Interviewer:** Can you describe your experience with crop diseases in your agricultural practices?

**Interviewee:** My experience has been in two ways. First, as a research scientist, I've been working on plant and crop diseases in both the lab and the field. I've encountered various crop diseases in crops like maize, potatoes, and pyrethrum. Second, as a farmer, I've dealt with crop diseases in my potato farming for the past seven years in Kenya.

**Interviewer:** How do you currently diagnose and manage crop diseases on your farm?

**Interviewee:** Currently, I diagnose diseases based on visual symptoms expressed in the crops. My background and expertise allow me to identify diseases by observing symptoms on leaves, stems, and different patterns in the plants. For management, we prioritize scouting to detect diseases early. We also use preventive pesticides to stop the spread of diseases in the entire farm.

**Interviewer:** What are the major challenges you face when identifying and treating crop diseases?

**Interviewee:** One challenge is that some diseases can manifest in similar ways, making accurate diagnosis difficult. Additionally, the cost of pesticides is high in Kenya, posing a challenge for both small-scale and large-scale farmers.

**Interviewer:** How do you make decisions regarding the type of pesticides or treatments to use for crop diseases?

**Interviewee:** Our decisions depend on the severity and incidence of the disease in the crop. For lower levels, we use preventive pesticides, while higher levels require control pesticides. Cost also influences our choices. We consider the molecular composition and active ingredients in pesticides, as well as their mode of action.

**Interviewer:** What weather conditions or environmental factors correlate with the occurrence of diseases?

**Interviewee:** Rainfall and humidity levels are significant environmental factors. For example, high rainfall and humidity contribute to diseases like early blight in potatoes. The interaction of the pathogen, plant susceptibility, and environmental conditions plays a crucial role in disease manifestation. High soil moisture levels also affect some diseases, especially fungal ones.

**Interviewer:** Have you used automated or technological solutions for crop disease diagnosis?

**Interviewee:** Yes, I've tried using online applications for disease diagnosis in the past.

**Interviewer:** What types of data would be useful for diagnosing and managing crop diseases effectively?

**Interviewee:** Images are crucial, showing disease progression from the early stages to full infection. Different images of affected plant parts and diverse symptoms for a specific disease and crop would be valuable data.

**Interviewer:** How would an automated crop disease diagnosis system benefit your farming practices, and what are your expectations?

**Interviewee:** Such a system would greatly assist in disease diagnosis, reducing losses. I expect it to provide detailed and credible information.

**Interviewer:** How do you envision integrating such a system into your farming routine, and are there concerns?

**Interviewee:** I see it being used during scouting and disease control. If it can not only identify diseases but also provide recommendations, it would be valuable. I don't have major concerns about the learning curve, as Kenyan farmers are eager to adopt effective systems.

**Interviewer:** Any concerns about data privacy or security with an automated system?

**Interviewee:** I don't have significant concerns regarding data privacy or security.

**Interviewer:** What potential impact do you expect from an accurate automated system on your farm's productivity, costs, and sustainability?

**Interviewee:** It would save costs on sampling and diagnosis and allow for faster disease management, resulting in higher yields and income. Improved profit margins would enhance sustainability.

**Interviewer:** What are the most important features or functionalities of an automated crop disease diagnosis system in your opinion?

**Interviewee:** Rapid response time and accuracy are crucial. Getting fast and accurate results is the key.

**Interviewer:** Thank you for your responses; these were great insights.

## THIRD INDEPTH INTERVIEW

**Interviewer:** Can you describe your experience with crop diseases in your agricultural practices?

**Interviewee:** I have worked with various crops, including mangoes and wood crops, for over 20 years.

**Interviewer:** How has it been?

**Interviewee:** It has been valuable, especially for mangoes, as diseases are significant in our native areas.

**Interviewer:** How do you currently diagnose and manage crop diseases on your farm?

**Interviewee:** We employ two methods: laboratory diagnosis, where farmers bring samples, and field diagnosis. We visually observe and identify pathogens, relying on images for diagnosis.

**Interviewer:** What are the major challenges you face when identifying and treating crop diseases?

**Interviewee:** Challenges include the cost of services, illiterate farmers requiring assistance with reading, and the scarcity of crop health specialists.

**Interviewer:** How do you make decisions regarding the use of pesticides and other treatments for crop diseases?

**Interviewee:** We encourage non-chemical methods like scouting to control diseases. Pesticides are expensive, and some are not approved for specific crops.

**Interviewer:** What environmental factors are correlated with crop diseases?

**Interviewee:** Climate change is a significant factor, causing new diseases due to altered weather patterns and temperatures.

**Interviewer:** Have you used any automated solutions for crop disease diagnosis?

**Interviewee:** No, we rely on conventional methods; we haven't adopted automated systems.

**Interviewer:** What data, such as weather, soil, or images, would be most useful for diagnosing and managing crop diseases?

**Interviewee:** Crop images, backed by weather data, are the most valuable, followed by soil data.

**Interviewer:** How would an automated crop disease diagnosis system benefit your farming practices?

**Interviewee:** It would be beneficial for instant access to data, cost-effectiveness, and wide coverage if farmers have the necessary devices.

**Interviewer:** How would you integrate an automated crop disease diagnosis system into your daily routine, and what challenges do you anticipate?

**Interviewee:** The system would be welcomed, but access and cost may pose challenges.

**Interviewer:** How do you currently access and manage agricultural data, and are there any limitations?

**Interviewee:** Data is collected in sheets, copied, and sometimes accessed online. The cost of data access can be a limitation.

**Interviewer:** What level of expertise or technical knowledge would farmers need to use an automated crop disease diagnosis system?

**Interviewee:** Farmers need gadgets and training to use the system effectively. Long-term data analysis is also necessary.

**Interviewer:** Are there concerns about data privacy or security with such a system?

**Interviewee:** Data privacy and security should be handled by relevant authorities.

**Interviewer:** What would be the potential impact of an accurate and efficient automated crop disease diagnosis system on your farm?

**Interviewee:** It would reduce diagnosis costs and benefit farm productivity in the long term.

CONSENT FORMS

To be provided based on application

