BAB 1 INTRODUCTION

Around 80% of the population in developing nations lives in rural areas, where agriculture serves as the primary source of livelihood [1]. As the demand for agricultural products continues to rise, ensuring crop health and productivity has become increasingly important. Apples are among the most widely consumed fruits globally, ranking as one of the top four produced fruits [2]. Not only are a major dietary staple consumed, but an economically valuable commodity is also recognized. However, apple plants are highly vulnerable to a variety of diseases that can significantly reduce yield and fruit quality. These infections can be classified into abiotic and biotic categories [3]. Abiotic infections are caused by non-living factors such as extreme weather and pollutants, while biotic infections are caused by living organisms such as fungi, bacteria, and viruses, all of which pose serious threats to the productivity and profitability of apple orchards. The traditional methods used to identify these diseases, such as Fluorescence In-Situ Hybridization (FISH), Polymer Chain Reaction (PCR), and Immunofluorescence (IF), require significant expertise and can only be performed by trained phytopathologists in laboratory environments [3]. This reliance on experts can delay disease identification and subsequent treatment, leading to severe economic losses for farmers.

Recent studies have demonstrated the advantages of deep learning techniques for plant disease classification, showcasing their superiority over traditional machine learning models. For instance, researchers showed that deep learning models like InceptionV3, VGG-16, and VGG-19 achieved the best accuracy of 89.5%, surpassing traditional approaches like Random Forest (RF), Stochastic Gradient Descent (SGD), and Support Vector Machine (SVM), which achieved an accuracy of only 87% [4]. Another study compared MobileNet, InceptionV3, and ResNet152, achieving accuracies of 73.50%, 75.59%, and 77.65% respectively [5]. Despite these promising results, a common limitation is shared by many studies: datasets with simplified environmental conditions, such as uniform or uncomplicated backgrounds, are used.

Various approaches have been undertaken to develop plant disease classification systems using deep learning techniques. In 2019, researchers in India employed a CNN-based model using the Plant Village dataset and attained an accuracy of 98.54% [6]. The Plant Village dataset is significant due to its large and diverse collection of labeled plant leaf images, but it lacks the environmental complexity of real-world agricultural settings. In 2021, another study used VGG16 to classify diseases affecting citrus plants, achieving an accuracy of 89.5%, but again the dataset lacked environmental complexity [7]. More recently, in 2022, researchers proposed a plant disease classification method based on an optimized lightweight YOLOv5 model, incorporating Ghostnet and WBF to reduce model weight and BiFPN for enhanced feature fusion, resulting in an accuracy of 92.57% [8]. In 2023, an updated version of MobileNet known as MobileNetV2 achieved 99.36% accuracy [9], but, similar to earlier studies, this high accuracy was obtained using a dataset with simple backgrounds.

Advances in image processing technology, artificial intelligence, and computational resources such as graphics processing units (GPUs) could revolutionize the process of plant disease detection [10]. This study aim to address this gap by using a dataset with complex backgrounds to better represent the practical challenges of apple disease classification [11]. By incorporating images with more natural, varied environments, the goal is to develop a model that is robust to different conditions. The model will be trained on a complex background dataset and subsequently tested on simpler datasets [12] to evaluate its generalizability and effectiveness across diverse environments, demonstrating its versatility and robustness.

This work utilize a modified version of EfficientNet as the backbone for apple disease classification. EfficientNet [13] has become well-known for its ability to scale effectively while achieving higher accuracy compared to basic CNN architectures [14]. Its scaling strategy uniformly adjusts depth, width, and resolution, providing an efficient balance between accuracy and computational requirements. This makes EfficientNet an excellent candidate for plant disease classification, particularly when resources are limited. In this research, EfficientNet is further modified by removing specific blocks to achieve faster computation while maintaining accuracy. This modification is intended to streamline the model, making it lighter and more suitable for practical deployment in settings where computational resources are constrained, such as rural farms or portable devices.