Article

Detection of a Potato Disease (Early Blight) Using Artificial Intelligence

Hassan Afzaal [1], Aitazaz A. Farooque [1,*], Arnold W. Schumann [2], Nazar Hussain [1], Andrew McKenzie-Gopsill [3], Travis Esau [4], Farhat Abbas [1] and Bishnu Acharya [1,†]

[1] Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE C1A4P3, Canada; hafzaal2@upei.ca (H.A.); nhussain@upei.ca (N.H.); fabbas@upei.ca (F.A.); bsp874@mail.usask.ca (B.A.)
[2] Citrus Research and Education Center, University of Florida, Lake Alfred, FL 33850, USA; schumaw@ufl.edu
[3] Agriculture and Agri-Food Canada, Research and Development Centre, Charlottetown, PE C1A4N6, Canada; andrew.mckenzie-gopsill@canada.ca
[4] Engineering Department, Faculty of Agriculture, Dalhousie University, Truro, NS B2N5E3, Canada; tesau@dal.ca
* Correspondence: afarooque@upei.ca; Tel.: +1-(902)-566-6084-550
† Current address: Department of Chemical and Biological Engineering, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A2, Canada.

Abstract: This study evaluated the potential of using machine vision in combination with deep learning (DL) to identify the early blight disease in real-time for potato production systems. Four fields were selected to collect images (n = 5199) of healthy and diseased potato plants under variable lights and shadow effects. A database was constructed using DL to identify the disease infestation at different stages throughout the growing season. Three convolutional neural networks (CNNs), namely GoogleNet, VGGNet, and EfficientNet, were trained using the PyTorch framework. The disease images were classified into three classes (2-class, 4-class, and 6-class) for accurate disease identification at different growth stages. Results of 2-class CNNs for disease identification revealed the significantly better performance of EfficientNet and VGGNet when compared with the GoogleNet (*FScore* range: 0.84–0.98). Results of 4-Class CNNs indicated better performance of EfficientNet when compared with other CNNs (*FScore* range: 0.79–0.94). Results of 6-class CNNs showed similar results as 4-class, with EfficientNet performing the best. GoogleNet, VGGNet, and EfficientNet inference time values ranged from 6.8–8.3, 2.1–2.5, 5.95–6.53 frames per second, respectively, on a Dell Latitude 5580 using graphical processing unit (GPU) mode. Overall, the CNNs and DL frameworks used in this study accurately classified the early blight disease at different stages. Site-specific application of fungicides by accurately identifying the early blight infected plants has a strong potential to reduce agrochemicals use, improve the profitability of potato growers, and lower environmental risks (runoff of fungicides to water bodies).

Keywords: deep learning; disease classification; PyTorch; EfficientNet; image processing; machine vision; smart sprayer

1. Introduction

Canada stands fifth in the list of potato (*Solanum tuberosum* L.) exporters with its exports amounting to $1.60 billion per year [1]. Atlantic Canada contributes 37% of Canadian potato production [2]. Typically, 20 applications of agrochemicals are uniformly applied during a growing season, without considering the spatial variability of disease infestations within a potato field. The desire of growers to increase potato production has resulted in a 150% increase in the uniform usage of agrochemicals during the past 20 years leading to an increased cost of production and environmental degradation [3]. Efficient crop management through the precise and accurate application of agrochemicals

can enhance potato quality and tuber yield while providing economic and environmental benefits [4]. The development of smart sprayers, capable of identifying spatially variable weed and disease distributions within potato fields, can help apply agrochemicals in the infected localities only. Such *precision* agriculture technologies, in their infancy stage in Canada, continue to gain popularity worldwide.

Early blight—a common potato disease, caused by *Alternaria solani Sorauer*, is found throughout North America [5]. Like other plant leaf diseases, it typically targets less productive and older foliage first, followed by steady movement upwards in the plant canopy causing leave senescence [5]. Visible symptoms of this disease include small 1–2 mm black or brown lesions, at early stages, turning into dark pigmented concentric rings under conducive environmental conditions, at mature stages of the disease [6]. Currently, management of early blight involves the uniform application of fungicides, ignoring its spatial distributions, which not only increases the production cost but also adversely impacts the environment. Therefore, an intelligent classification system to differentiate diseased plants from healthy plants and achieve the targeted application of fungicides can improve economic and environmental sustainability.

The identification of plant diseases is a long-researched topic [7]. An accurate and real-time disease detection system may help in developing mitigation strategies to ensure economical crop protection on a small scale and food security on a large scale. Furthermore, an accurate disease classification through machine vision and deep learning (DL) could provide the basis to accomplish the site-specific application of agrochemicals. Abdullah et al. [8] proposed a method of disease detection by using a neural network. Their algorithm applied principal components analysis to a set of low-resolution images followed by feeding the processed images to a neural network for plant disease detection. Sena et al. [9] suggested a thresholding and image segmentation model for disease detection. They divided greyscale images of diseased leaves into different blocks and applied the programmed threshold levels to detect disease in each image block. Camargo and Smith [10] presented an alternative image segmentation method for disease quantification. Their developed algorithm transformed color images to HSV (hue, saturation, value) and then in I1I2I3 color space. With this method, they were able to quantify diseases of various plant species. Boese et al. [11] used an unsupervised color-based disease quantification technique, which divided the image into several classes followed by training of each class by a probabilistic supervised learning technique. According to Boese et al. [11], this method had higher accuracy in comparison with previous methods. Although the above-discussed algorithms provide the basis of disease detection, most of these methods are not feasible for real-time application of agrochemicals as they require a very long inference time for image processing.

As per recent developments in the graphical processing unit (GPU)-embedded processors, the applications of artificial intelligence have gained exponential growth leading to the development of novel methodologies and models called DL [12]. GPU-based parallel computing has supported the development and use of CNNs for various applications. Krizhevsky et al. [13] tested large datasets of multiple image classes with high accuracy using various CNNs. More recently, CNNs have become one of the most successful DL techniques in image processing. Typically, a CNN consists of several layers such as convolutional, pooling, and fully connected layers [12]. A convolutional layer is composed of several mathematical operations, which play key roles in CNN functions, such as convolution—a special linear operation [14]. The convolutional layer lifts the heavy load of the computation and is a critical building block of a CNN. Pooling layers usually come in between consecutive convolutional layers. The main function performed by pooling layers is to reduce the spatial computational parameter load. The fully connected layers consist of the activation functions by which the final output, such as object classification, is extracted [14]. The most common type of activation function is called a rectified linear unit, which replaces the negative numbers in pixels by zero values to minimize the computation time and exposes the targets for identification [15]. A digital image stores the data in

the form of two-dimensional grids; i.e., an array, and a small unit of an array called a kernel [14]. In CNNs, training is the process to reduce the error between predicted output and ground truths by the process of convolution.

The CNNs have been successfully implemented in disease detection problems. For example, Oppenheim and Shani [16] classified four different potato tuber diseases by using a visual geometry group. Ferentinos [7] was able to attain a 99.5% success rate for detecting 58 diseases in 25 types of plants by training the plant disease images with different CNNs. Carranza-Rojas et al. [17] proposed a species identification system for botanists using CNN for herbarium datasets that are comprised of multimillion images of various species. With the high accuracy of CNN, the DL technique has been suggested for semi or fully automated species identification systems. Fuentes et al. [18] recognized nine types of disease and pests in tomato plants by comparing several CNNs. The CNN, based on faster regional CNN with Visual Geometry Group (VGG) as a feature extractor, attained the highest accuracy for disease detection. A smartphone-assisted disease detection system supported by a trained CNN was developed by Mohanty et al. [19] to identify 14 crop species and 26 diseases with 99.4% accuracy. Recently, a new family of CNNs was developed based on a systematic design called EfficientNet, which is designed based on optimized network width, depth, and resolution. The optimized design of EfficientNet has resulted in a higher accuracy as well as a smaller number of learning parameters.

The objective of this study was to evaluate the potential of using artificial intelligence-based convolutional neural networks to identify the early blight disease in real-time for potato production systems. This research compares three ages of CNNs (GoogleNet—old, VGGNet—mid-age, and EfficientNET—new) for identification of early blight disease in its different growth stages (initial, mid, last) and their possible combinations in potato production. Furthermore, inference times for the selected CNNs were computed to assess their adequacy in developing a smart sprayer for the targeted application of fungicides on an as-needed basis. The disease identification using the integration of CNNs, machine vision, and assessing potential and feasibility to have integrated them into a smart sprayer adds novelty and innovation to this research study being reported here.

2. Materials and Methods

2.1. Study Site and Data Collection

For the present study, two fields in New Brunswick, Canada (Florenceville 46.318161 N, 67.617139 W and McCain's Farms 47.103106 N, 67.855958 W) and two fields in Prince Edward Island, Canada (Winsloe North 46.359412 N, 63.201919 W and 225 Highway 46.255502 N, 63.201203 W) were selected. Atlantic Canadian provinces climate is considered humid, strongly influenced by the surrounding seas and oceans throughout the year. The summer season is relatively shorter with the long winter season. In Atlantic Canadian provinces, the majority of soils are sandy loam. Images of healthy and diseased potato plants (*Russet Burbank* cultivar) were captured using digital cameras namely Canon Power-Shot SX540 and Logitech C270 HD Webcam. Visual symptomology of early blight was discussed and identified with a local plant pathologist. These observations coincided with Kemmit [6] for the identification of early blight disease stages throughout the growing season. Images were captured based on visual symptoms to categorize for the initial, mid, and last stages of the disease in potato fields. Literature suggested that the visual symptoms of early blight (pale, yellow-colored leaves) start appearing in early July, i.e., 35 days after planting (DAP)—a period counted for observing the initial stage of the disease. Mid-stage of the disease occurs at 65 DAP (brown color concentric circles on potato plant leaves). The late stage of the disease evolves at 80 DAP (Figure 1). As no curative measures exist, accurate identification of disease at an early stage using machine vision is a key to mitigate the damage. The data points in the captured frame may vary real-time as it is the function of disease intensity. The more intense disease may have more disease data points than the less intense disease plants. Furthermore, the disease appeared on potato plants in the form of patches because of disease spread.
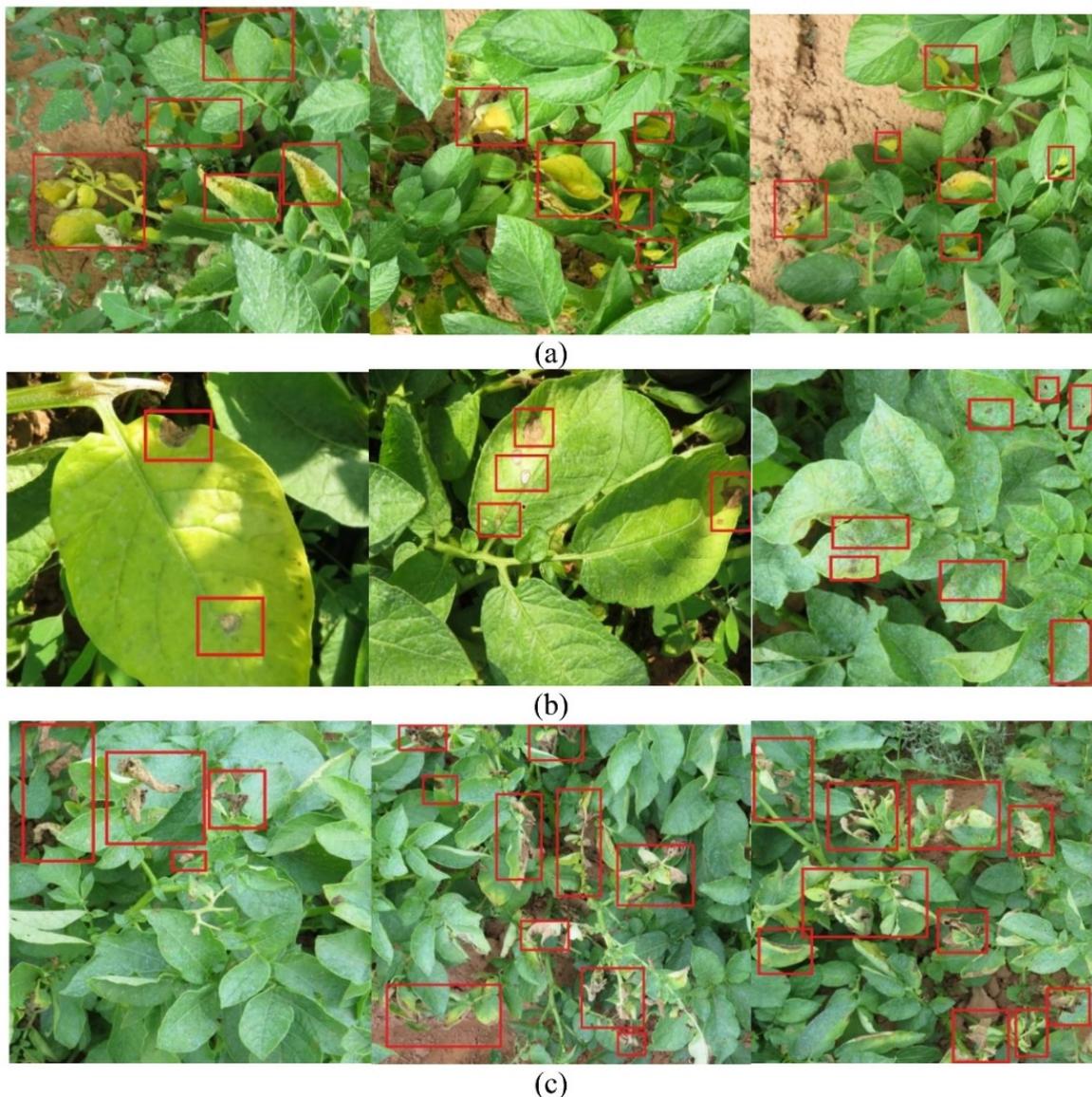
Figure 1. Sample images for identification of early blight disease stages on potato leaves; (**a**) pale yellow color leaves at the initial stage of the disease, (**b**) concentric brown circles in the leaves at mid-stage, and (**c**) senescence of leaves at the last stage of the disease.

The images were taken 140 cm above the soil surface, under natural sunlight between 9:00 a.m. and 4:00 p.m. in 2018. The collected images were resized (1280 × 720 pixels) using a custom-developed program written in the Python programming language. The pixel resolution of 1280 × 720 was selected to extend the application of this model for real-time applications. Moreover, the selected resolution was chosen to assess the feasibility of the developed CNN model to be integrated with the hardware, to develop a smart variable rate sprayer using a real-time video feeding from a webcam.

The disease stages were identified by collecting the images throughout the potato-growing season from June to October 2018. Images of each disease stage, with its symptoms, were trained with different CNNs. A dataset of 532 images was reserved for testing and for evaluation of statistical measures after the CNNs are trained (Table 1). The remaining images were used for training (70%) and validation (30%) of the developed CNNs. The accuracy of CNNs in identifying the early blight disease stages was assessed using statistical measures including *precision, recall*, and *FScore*.

Table 1. Number of images used for training and validation to identify the disease stages for selected fields.

Disease Stage	Days after Planting	Disease Symptoms	Label	Training/Validation/ Testing	Images Used	Total Images
Early	30	Yellow color leaves	Diseased	Training	622	842
				Validation	127	
				Testing	93	
			Healthy	Training	614	854
				Validation	151	
				Testing	89	
Mid	55	Brown concentric leaves	Diseased	Training	524	728
				Validation	132	
				Testing	72	
			Healthy	Training	658	913
				Validation	164	
				Testing	91	
Last	80	Senescence of leaves	Diseased	Training	702	976
				Validation	176	
				Testing	98	
			Healthy	Training	638	886
				Validation	159	
				Testing	89	

2.2. Multistage Disease Comparison

Plant disease progression is a function of environmental variables (temperature, humidity, wind, etc.). If the CNNs are trained for only three stages (mid, early, and late), the overlap between disease stages will compromise the accuracy of the CNNs. This will reduce the CNNs' applicability for real-time smart sprayer applications due to low disease identification accuracy. To compensate for variating disease stages, the overlap between stages, and to clearly distinguish the disease from potato plants, CNN were divided into three separate experimental groups for accurate identification. These steps have simplified the tasks of disease detection at various stages of the disease and potato growth. The 2-Class CNN compared one disease and one healthy stage. The 4-Class CNN was used to compare two disease and two healthy plant stages. The 6-Class CNN compared three disease and three healthy plant stages simultaneously to differentiate healthy potato plants accurately and reliably from diseased plants. Healthy plant stages were defined based on leaf size on the respective phenological stages of a potato plant. The analogy behind these simulations was to expose the CNNs to all possible scenarios to avoid fungicide spraying on healthy plants. Furthermore, the agrochemical application can be optimized based on the severity of disease infestation, picked by variable-rate sprayer integrated with DL CNNs and machine vision approach solicited by this research.

The 2-, 4-, and 6-Class CNNs

The architecture of the 2-Class CNN was trained on one disease and one healthy potato plant stage, i.e., initial, mid, and last early blight (Figure 2). All the validation images were passed through the trained model to evaluate its ability to distinguish various diseases and healthy potato plant leaves at their various growth stages. The 4-Class CNN was trained using two stages of disease symptoms and two healthy potato plants (initial + mid, mid + last, and initial + last). The performance of the 4-class CNN was evaluated by combining the two groups of images and running these images with the trained model to

classify the healthy and diseased plants. The 6-Class CNN was trained on three groups of diseased and healthy plant stages (initial + mid + last). The 6-Class CNN's performance was evaluated in terms of differentiation between healthy and diseased plants using all possible combinations. These interactions using various combinations of diseased and healthy plants in conjunction with CNNs enabled robust and comprehensive evaluation of the CNNs. These architectures, extensive training, testing, and validations evaluated the efficacy of the CNNs and machine vision in disease classification for a potato cropping system.
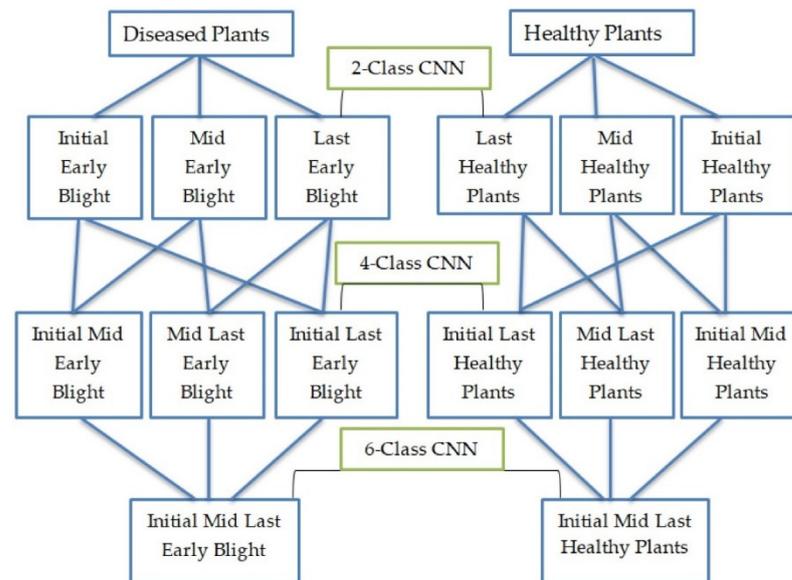
Figure 2. Concept flowchart of combinations of disease and healthy plant stages to assess and compare the classification accuracy of 2-Class, 4-Class, and 6-Class convolutional neural networks (CNNs).

2.3. Deep Learning Frameworks and Convolutional Neural Networks

The rapid surge in DL and machine vision has provided easier frameworks to run CNNs. Several frameworks were developed such as Caffe, Torch, Theano, and others, which facilitate the computational process and efficiency for various applications. These frameworks have gone through many development phases to attain the desired accuracy for a particular application. The DL frameworks can build large and complex computational analyses efficiently. The DL frameworks also help in calculating the gradient losses to reduce the error between predicted and ground truth data. Due to heavy computational loads associated with CNNs, a GPU-enabled workstation saves time in training and validation of an image using DL frameworks. Furthermore, the DL frameworks enable networking in central processing unit (CPU), GPU, and memory by using DL libraries. The PyTorch framework was used for all CNNs (training, validation, and testing) to differentiate diseased and healthy potato plants because of its readily available resources and ease of use.

PyTorch, developed by the Facebook AI research lab, is a fast-maturing framework with good GPU support [20]. PyTorch has gained popularity in the DL community because of its in-depth features such as a dynamic graph, easily available resources, and research-oriented focus. In this study, three CNNs, namely GoogleNet, VGGNet, and EfficentNet, were selected to identify early blight disease at different stages. These models were selected based on literature-reviewed accuracies, real-time performance, and applicability in production for commercial applications. Szegedy et al. [21] suggested the use of GoogleNet for accurate predictions. GoogleNet has been successfully used in plant disease identification and plant parts identification problems [22,23].

The VGGNet architecture consists of 3×3 convolutional filters to improve its accuracy and reduce overfitting problems [24]. The VGGNet is available in different depths of layers

such as Vgg-11, Vgg-13, and Vgg-16. In this study, the most recent one, VGGNet-16, was tested for disease classification stages. EfficientNet is a recently developed model by Tan and Le [25] which has been designed on a systematic approach of network depth, width, and resolution. EfficientNet is claimed to be a light as well as an accurate model based on the compound scaling approach. In this study, we used the lightest EfficientNet model B0, which is considered to be an accurate model considering image results. Overall, this study used the old (GoogleNet), mid-age (VGGNet-16), and the recent (EfficientNet) CNNs using the PyTorch framework to evaluate their potential and accuracy in the classification of early blight diseased and healthy potato leaves.

2.4. Deep Learning Workstation and Hyperparameter Tuning of CNNs

A custom-designed DL workstation was used to train the selected CNNs. The workstation consisted of Intel core CPU i9-7900X (3.3Ghz, SkylakeX 10 core), two GPUs (Nvidia GTX 1080 GeForce, 11GB), two memory sticks of Ram (Vengeance 32 GB), and 1 TB solid state drive. Ubuntu 16.04 operating system was installed with PyTorch (Version 1.12), NumPy (Version 1.18.1), Matplotlib (version 3.1.3), Pandas (1.0.1), and associated libraries. PyTorch provided several code snippets to get the reproducible results for similar configurations of CNNs and datasets. In this study, we selected the random seeds of 20 for deep learning libraries including NumPy and cuDNN. Furthermore, all data loader workers were set to zero to achieve similar results.

As DL is computationally expensive, it is important to set the optimal hyperparameters for CNN models to improve accuracy. The hyperparameters were determined with several trials of different optimizers, learning rates, and weights initialization. The GoogleNet performed better if trained from scratch in comparison with pre-trained weights and fine-tuning. The learning rate was set to 0.01 for 2-Class CNN and 0.001 for 4- and 6-Class CNNs to reduce over-fitting. Adam optimizer with step down learning rate and 0.95 momenta performed well on our image datasets. VGGNet-16 is computationally a more expensive CNN than GoogleNet because of a higher number of learning parameters, e.g., 138 million. For VGGNet-16, the batch size was set to 64 using both GPUs. Because of the large number of parameters, these CNNs were trained using ImageNet pre-trained weights to optimize the training time and accuracy. Adam optimizer performed well with weight decay, the momentum of 0.95, and a learning rate of 0.001 for 2-Class CNN and 0.0001 for 4- and 6-CNNs. The CNN's convergence was faster with lower learning rates because of pre-trained weight initialization in these models. EfficientNet was trained using pre-trained weights of lightest model B0. The lightest model of EfficientNet was selected to assess its application for real-time image classification. This CNN worked well with a batch size of 256. Learning rates with Adam optimizers were selected from 10^{-4} to 10^{-6} for different classes of disease. Furthermore, a dropout effect of 20% was introduced in earlier layers of CNN to overcome the over-fitting issues for EfficientNet.

2.5. Statistical Analysis and Inference Time

The performance of all CNNs using the PyTorch framework was evaluated using the well-established statistical measures of *precision*, *recall*, and *FScore* [22]. Results were also assessed based on loss and accuracy during the training and validation phases of the CNNs. A confusion matrix provided the parameters to calculate these statistical measures. Inference time is a key to assess the applicability of the CNNs for real-time application (e.g., real-time detection of disease by smart sprayers before application of agrochemicals). The neural network must be efficient in terms of both computation time and memory during the inference [26] to qualify for real-time applications. Adequate accuracy in conjunction with inference time is considered one of the key elements in practical applications of the recommended CNN for efficient resource allocation and power consumption. Inference time was recorded for each CNN using a Dell Latitude 5580 (8 GB ram, Core I7 7 generation processor) on GPU mode using DL framework libraries. One image was used three times to calculate the average inference time for all CNNs. A batch size of 1 was selected to

calculate the inference time for real-time application with less powerful devices with no GPU for applicability and lower power consumption.

3. Results

3.1. Training Results

The 2-, 4- and 6-Class CNNs

In the training of initial early blight detection, validation accuracy values of all the CNNs were recorded in the range of 0.95 to 0.97. GoogleNet achieved slightly lower validation accuracy in comparison with VGGNet and EfficientNet. Results reported that the validation accuracy for all CNNs ranged from 0.99 to 1.00 for mid-stage early blight detection (Table 2). Similarly, validation accuracy ranged between 0.99 and 1.00 for the last stage early blight. The performance accuracy of CNNs was non-significant for the detection of mid and late stages of early blight. Results of the 4-Class CNNs revealed validation accuracy of 0.92. The validation accuracy of joint stages (initial + mid) was observed to be lower when compared with the individual stages, i.e., initial, mid, and late separately (Table 2). For the mid + last early blight disease stage, the validation accuracy ranged from 0.84 to 0.86. The narrow range of accuracy revealed that all the CNNs performed similarly. The classification accuracy ranged from 0.92 to 0.94 for the initial + last stage of early blight disease. Results depicted slightly higher validation accuracy for the initial + last stage of the disease in comparison with initial + mid and mid + last stages for 4-Class CNNs. The validation accuracy for 6-Class CNN ranged from 0.81 to 0.85 for the initial + mid + last disease stage. Because of the high number of similar classes, low validation accuracy values were observed for 6-Class CNN with relatively high training and validation losses (Table 2).

Table 2. Training and statistical evaluation of the 2-, 4- and 6-Class convolutional neural networks in classifying the early blight disease stages for the potato production system.

Class	Model	Growth Stage	[1] Val Acc	[2] Val Loss	[3] Tr Loss	Precision	Recall
2 Class CNNs	GoogleNet	Initial	0.95	0.33	0.01	0.8	0.91
		Mid	1.00	0.03	0.01	0.92	0.94
		Last	1.00	0.00	0.00	0.83	0.87
	VGGNet	Initial	0.96	0.61	0.00	0.88	0.87
		Mid	0.99	0.04	0.01	0.97	0.95
		Last	0.99	0.06	0.00	0.88	0.9
	EfficientNet	Initial	0.97	0.07	0.01	0.88	0.88
		Mid	0.99	0.02	0.00	0.99	0.97
		Last	1.00	0.01	0.00	0.87	0.87
4 Class CNNs	GoogleNet	Initial + Mid	0.92	0.28	0.03	0.87	0.87
		Mid + Last	0.86	0.62	0.20	0.83	0.83
		Initial + Last	0.92	0.01	0.00	0.93	0.93
	VGGNet	Initial + Mid	0.92	0.27	0.01	0.84	0.85
		Mid + Last	0.86	0.13	0.07	0.79	0.78
		Initial + Last	0.93	0.25	0.12	0.88	0.88
	EfficientNet	Initial + Mid	0.92	0.01	0.01	0.88	0.88
		Mid + Last	0.84	0.79	0.30	0.85	0.85
		Initial + Last	0.94	0.03	0.00	0.95	0.95
6 Class CNNs	GoogleNet	Initial + Mid + Last	0.81	0.21	0.08	0.74	0.74
	VGGNet	Initial + Mid + Last	0.82	0.32	0.12	0.73	0.72
	EfficientNet	Initial + Mid + Last	0.85	0.12	0.02	0.74	0.75

[1] Val acc = validation accuracy; [2] Val Loss = validation loss; [3] Tr loss = training loss.

3.2. Statistical Evaluation of the 2-, 4- and 6-Class CNNs, and Inference Time

In the detection of the initial stage of the disease, the 2-Class CNNs (VGGNet and EfficientNet) performed well by generating and computing the highest number of true positives (82) (Figure 3). The VGGNet *precision* and *recall* scores were 0.88 and 0.87, respectively. EfficientNet achieved the highest *precision* score of 0.99 with 71 true positive and 1 false-positive image for the -mid-stage of the disease. The VGGNet performed well by achieving *precision* and *recall* scores of 0.88 and 0.90 for the last stage of early blight.
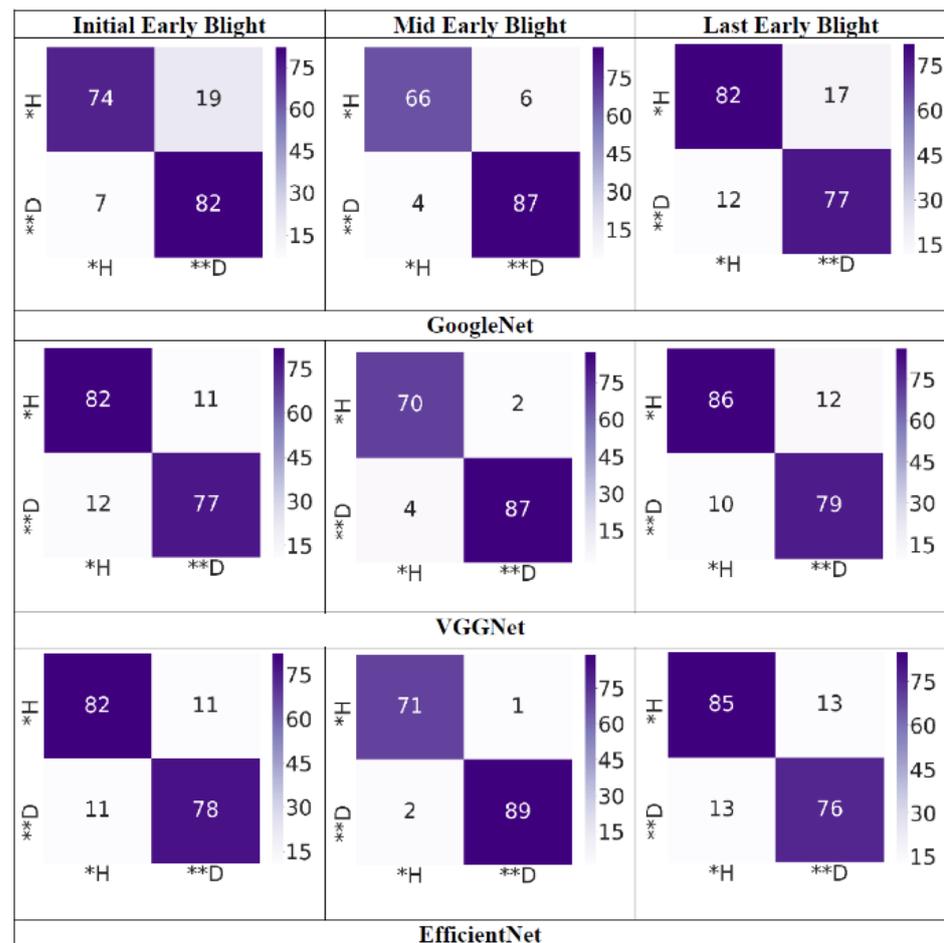
Figure 3. Confusion matrices of 2-Class convolutional neural networks, where * H = the number of actual and predicted healthy plants; ** D = the number of actual and predicted diseased plants.

The 4-Class CNN (EfficientNet) performed comparatively better with *precision* and *recall* scores of 0.88 and 0.88, respectively for the initial + mid stage of disease detection (Table 2) with an average of 77.8 true positives and 10.5 false negatives per class (Figure 4). At the mid + last disease stage, EfficientNet performed comparatively better with *precision* and *recall* scores of 0.85 and 0.85, respectively (Table 2) with an average of 74.5 true positives and 13 false negatives per class (Figure 4). During the detection of the initial + last stage of the disease symptoms, EfficientNet performed comparatively better concerning *precision* and *recall* scores of 0.94 and 0.94 (Table 2) with an average of 87 true positives and 5.75 false negatives per class (Figure 4). Overall, the results suggested that the 4-Class EfficientNet performed better in early blight disease detection using multiple combinations when compared with the GoogleNet and VGGNet (Table 2; Figure 3).
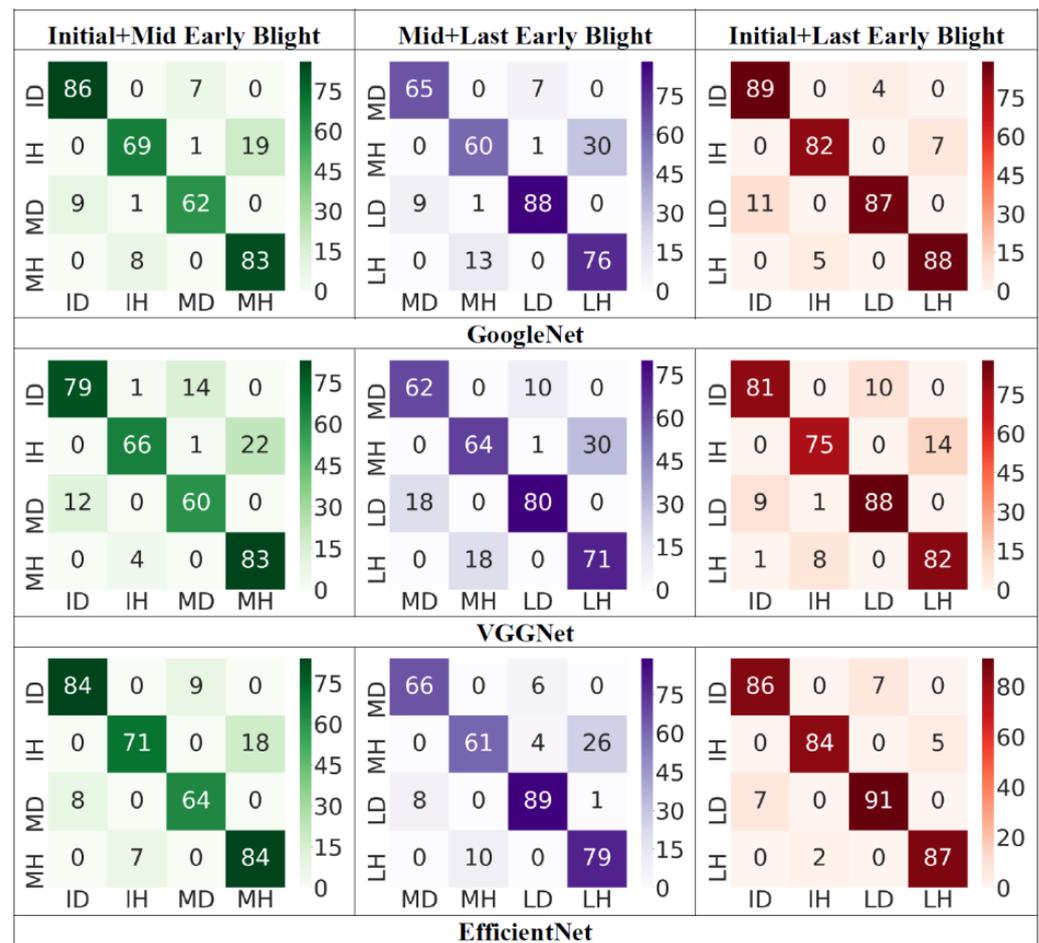
Figure 4. Confusion matrices of four class convolutional neural networks, where IH = number of actual and predicted initial healthy plants; ID = number of actual and predicted initial diseased plants; MH = number of actual predicted mid healthy plants; MD = number of actual and predicted mid diseased plants; LH = number of actual and predicted healthy plants; LD = number of actual and diseased plants.

The 6-Class CNN (EfficientNet) performed better when compared with CNNs at the initial + mid + last stage of early blight by achieving *precision* and *recall* scores of 0.74 and 0.75, respectively (Table 2). For the combined combination of all disease stages (initial + mid + last stage), EfficientNet achieved an average of 66 true positives and 22.67 false-positive per class (Figure 5). In summary, results of the confusion matrix in conjunction with *precision* and *recall* revealed that the 2-, 4- and 6-Class EfficientNet performed better in detecting early blight disease at individual and combined stages when compared with GoogleNet and VGGNet.

FScore is another criterion to evaluate the performance accuracy of the CNNs. Results of 2-Class CNNs revealed the *FScore* ranged from 0.853 to 0.905; 0.935 to 0.975; and 0.844 to 0.950 for initial, mid, and last stages of early blight disease, respectively (Figure 6). Results suggested that the CNNs performed comparatively better in detecting early blight disease with higher *FScore* during mid and late stages when compared with initial stages. During all individual stages, i.e., early, mid, and late, VGGNet and EfficientNet performed better than the GoogleNet for disease detection (Figure 6).
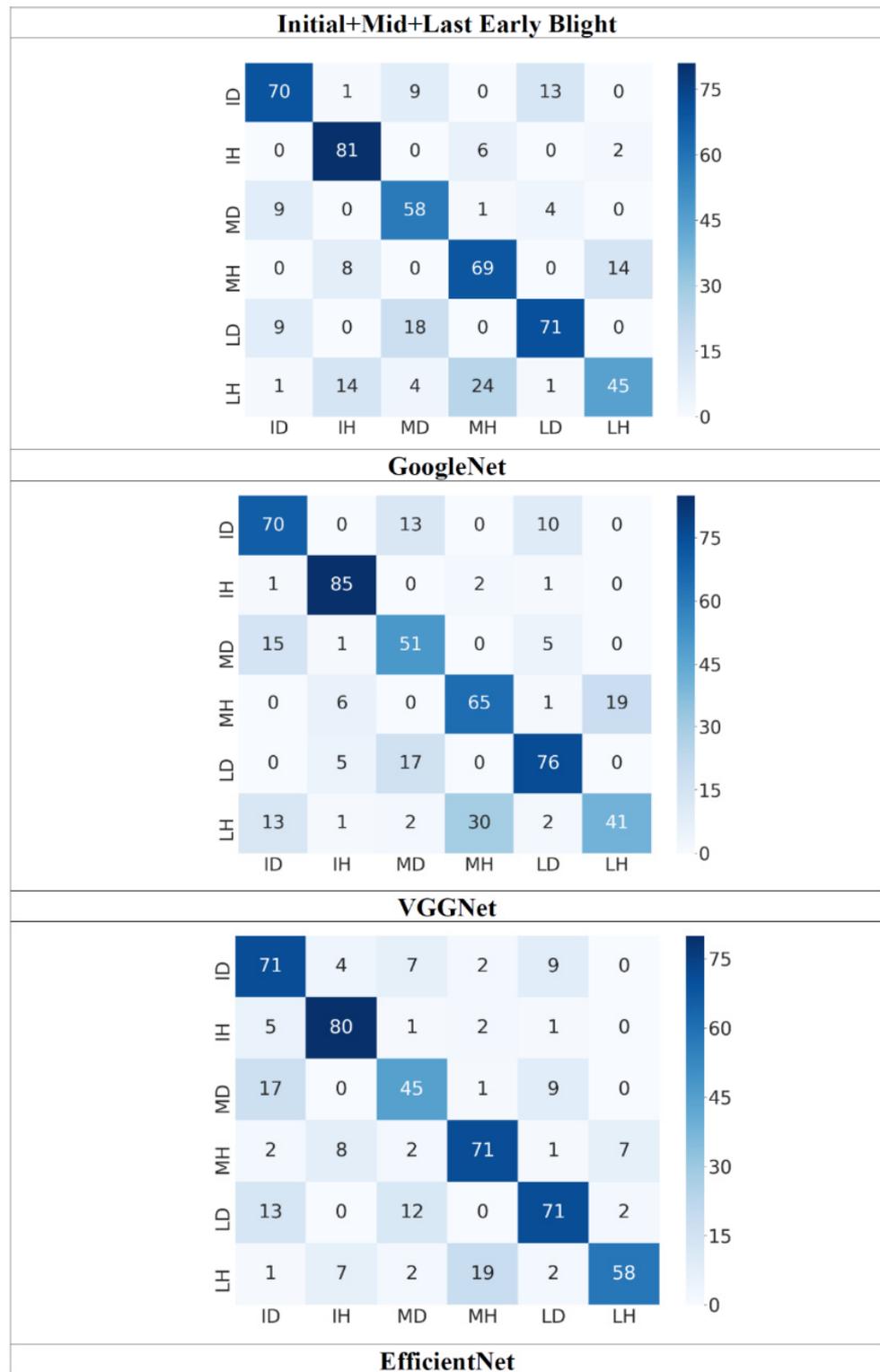
Figure 5. Confusion matrices of 6-Class convolutional neural networks, where IH = number of actual and predicted initial healthy plants; ID = number of actual and predicted initial diseased plants; MH = number of actual predicted mid healthy plants; MD = number of actual and predicted mid diseased plants; LH = number of actual and predicted healthy plants; LD = number of actual and diseased plants.
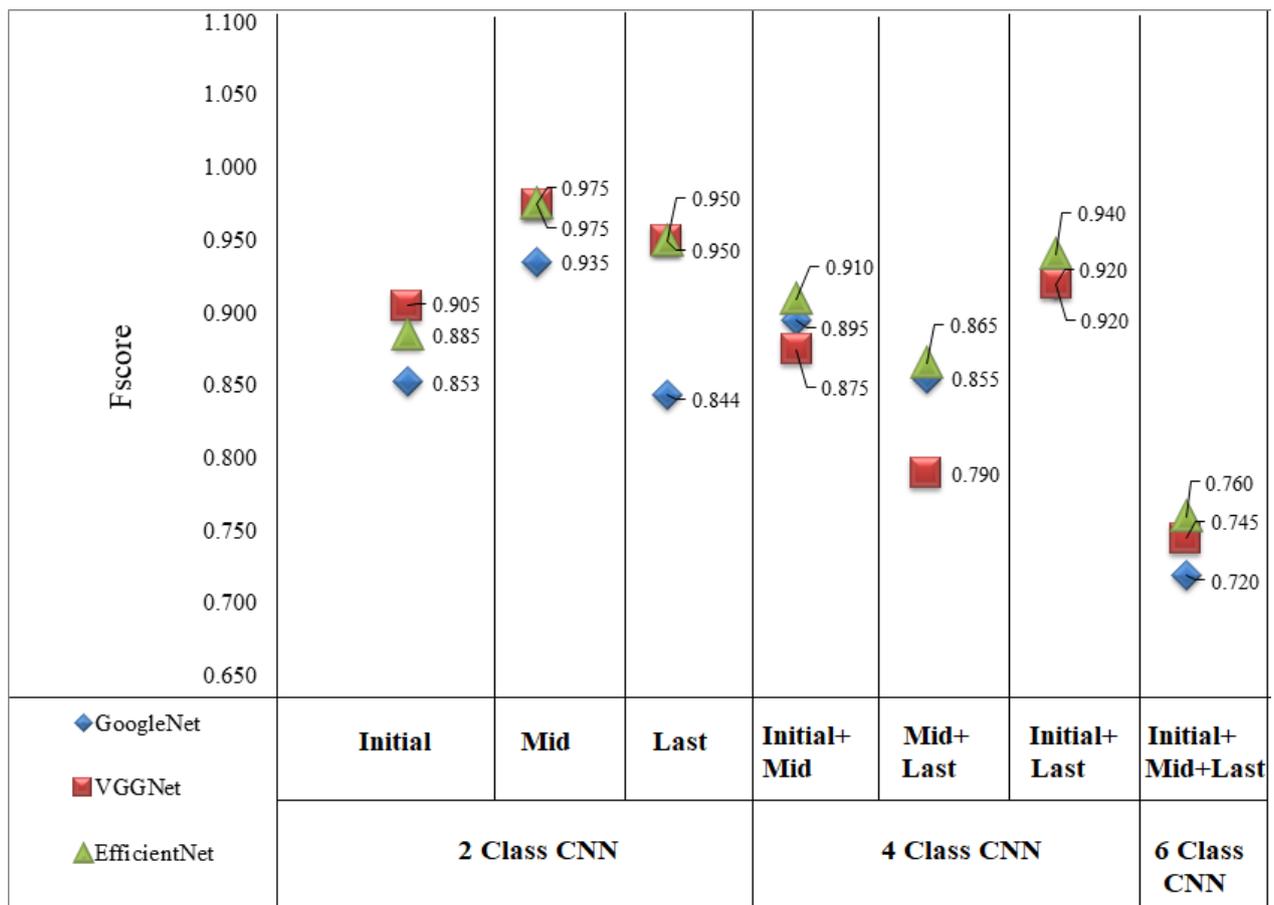
Figure 6. *FScore* comparison of various convolutional neural networks to evaluate their performance accuracy in early blight disease detection in the potato cropping system.

Results of 4-Class CNNs showed the *FScore* ranged from 0.875 to 0.910; 0.790 to 0.865; and 0.920 to 0.940 for initial + mid, mid + last, and initial + last stages of early blight disease, respectively (Figure 6). The recorded *FScores* ranged between 0.720 and 0.760 for initial + mid + last stage early blight disease detection for 6-Class CNNs (Figure 6). During individual and combination stages of the disease progression, EfficientNet was found to be the best performing CNN when compared with the GoogleNet and VGGNet. The results of VGGNet, however, were non-significantly different when compared with EfficientNet. Results of *FScore* support the trends of other statistical measures including the confusion matrices, *precision*, and *recall*.

The inference time is one of the crucial factors to induce the implementation of the developed CNNs for real-time applications (e.g., smart sprayer development to achieve targeted application if fungicides). The computed inference time ranged from 155–168, 121–145, and 398–458 milliseconds for EfficientNet, GoogleNet, and VGGNet, respectively (Figure 7). The lowest inference time was recorded for GoogleNet when compared with other CNNs. There were non-significant differences in inference time when computed for various disease stages, both individually and in combination. The inference time was observed to be highest for VGGNet when compared with other CNNs (Figure 7). The highest disease detection accuracy for EfficientNet (as described by the *precision*, *recall*, and *FScore*) in combination with lower inference time make it an excellent candidate for further processing and real-time application.
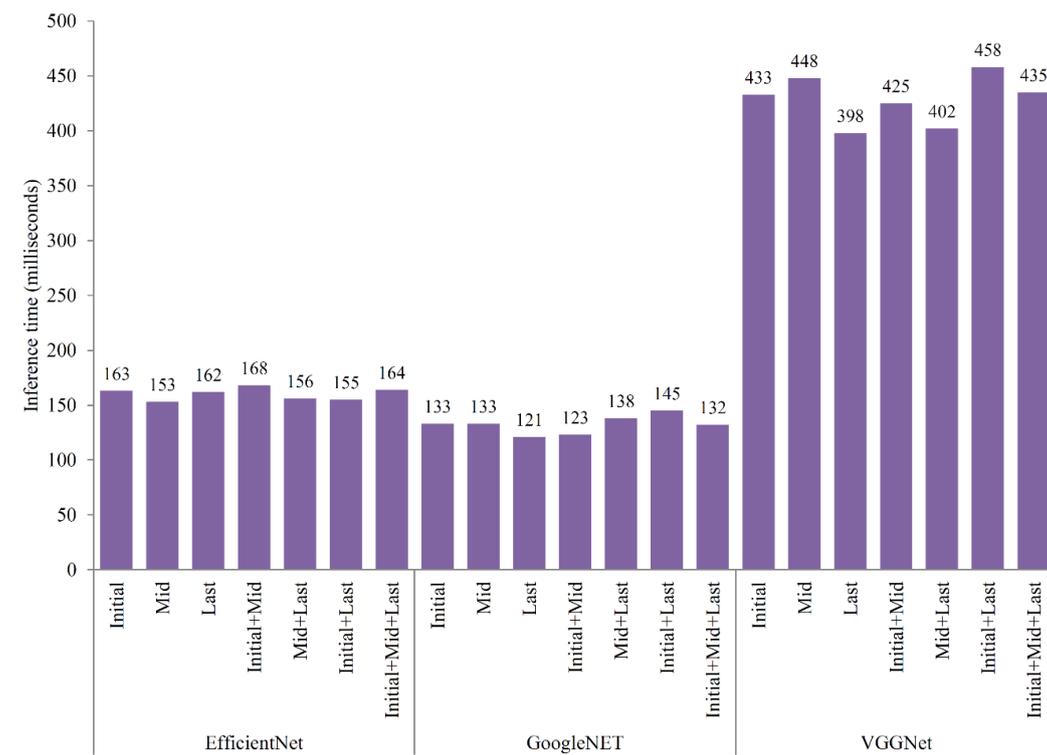
Figure 7. Inference time comparison for combinations of various convolutional neural networks.

4. Discussion

Symptoms of initial stages of early blight disease in potato plants include pale-yellow chlorotic leaves of the infected plant (Figure 1). The values of validation accuracy of CNNs in detecting initial early blight disease ranged from 0.95 to 0.97. These validation accuracy values were lower as compared to the mid and last stages of early blight disease. This may be due to the relative difficulty in detecting disease symptoms at this stage. Slightly lower *precision* scores were recorded in comparison with *recall* scores due to higher false positives during the initial stage (Figure 3). The lower *precision* score depicts the CNNs' inability to correctly classify the disease images and confusion with healthy plant images. GoogleNet was found to be less accurate in detecting disease infestation as indicated by the low *precision* and *recall* scores when compared with other CNNs. The architecture of GoogleNet that starts with a big receptive field to reduce computation requirements [21] might have resulted in the poor detection of small-sized disease symptoms for the initial stage of early blight disease. These results are in concurrence with the findings of Russakovsky et al. [27] who also reported the inability of GoogleNet in detecting small features. VGGNet achieved the highest *FScore* in detecting symptoms of initial early blight most likely because the architecture of VGGNet uses a small filter size of 3×3, which helps the CNN to identify small and complex objects [24]. These results suggested non-significant differences between VGGNet and EfficientNet in terms of disease detection at the early stages.

The validation accuracy values for CNNs during the mid-stage of early blight disease were observed to be slightly higher (e.g., 0.99 to 1.00) when compared with the early stage where the symptoms were confused with leaves. The *precision* and *recall* scores for the chosen CNNs were non-significantly different. The *FScore* displayed a similar statistical pattern as *precision* and *recall*. Similar values of validation accuracy were observed for the last stage of early blight disease; however, a slightly lower *FScore* was achieved for the last stage of the disease. In the last stage early blight detection, a similar performance of GoogleNet was observed in detecting the initial stage of early blight due to the reasons quoted above for GoogleNet's architecture. In the 4-Class CNN at the mid + late stage, EfficientNet optimized architecture did not perform well in comparison with VGGNet.

This might be due to overlapping symptoms at both stages due to environmental factors. The convolution and max-pooling filters throughout the VGGNet architecture might have helped VGGNet to perform better in the detection of all disease stages of 2- and 4-Class CNNs.

During the detection of initial + mid stage, early blight disease symptoms, the pale-yellow color, brown concentric circles, and plant development stages were the main targets for the CNNs to pick up during DL processing. The validation accuracy of 0.92 was recorded for the three CNNs. There were non-significant differences between *precision* and *recall* scores for initial + mid early blight disease CNNs. The initial healthy class was problematic for initial + mid early blight CNNs as it was highly confused with the mid healthy class (Figure 4). In the higher class CNNs, the performance of GoogleNet was lower in comparison with VGGNet and EfficientNet. EfficientNet's architecture was able to classify the complex and difficult targets with higher accuracy as suggested by *precision*, *recall*, *FScore*, and confusion matrices (Table 2; Figures 4–6). In general, EfficientNet CNN was found to be the best performer in terms of early blight disease detection.

In the mid + last stage of early blight disease detection, the brown concentric leaves, leave senescence and plant development stages were the main targets for 4-Class CNNs using DL. The validation accuracy values for the mid + last stage CNNs ranged from 0.84 and 0.86, which were slightly lower than the values of initial + mid and initial + last stages. Mid healthy plant stages and last disease stages were problematic for these CNNs as the models were confused by the last healthy and initial disease stages (Figure 4). A similar trend of *FScore* was observed for various CNNS as shown in the initial + mid stage (Figure 6). Non-significant differences were observed in terms of *precision* and *recall* scores (Table 2).

The brown pale-yellow leaves senescence, and plant development stages were the main detection targets for 4-Class CNNS at initial + last stage early blight disease detection. In this stage, the values of validation accuracy were in the range of 0.92 and 0.94, which were the highest values of accuracy ranges among 4-Class CNNs. The detection accuracy of the 4-Class CNN at the initial + last stage of the disease evolution was supported by the highest statistical measures of *precision*, *recall*, and *FScore*. This may be due to the clear distinction between healthy plant development stages and disease symptoms of the initial and last stage of early blight. The 6-Class CNNs performed significantly lower in terms of detection accuracy when compared with the 2- and 4-Class CNNs. The possible reason for this finding could be a higher number of similar healthy and disease classes. Moreover, overlapping of symptoms during different disease stages might have contributed to the lower accuracy. Initial + mid + last early blight validation accuracy values were in the range of 0.81 to 0.85. Mid-diseased, last-diseased, and last healthy classes were problematic in 6-Class CNN (Figure 5). EfficientNet was the highest performing CNN for 6-Class CNN. However, there were non-significant differences about *FScore* for 6-Class CNNs. The idea of training the CNNs with various combinations of healthy and diseased plants at different stages was to identify and develop robust CNN for practical applications and assess its potential to be integrated into a smart sprayer to achieve the site-specific application of fungicides within potato fields.

GoogleNet inference time was recorded to be 121–145 milliseconds on a Dell Latitude 5580 with all three CNNs. However, the values of inference time for all classes' CNNs were in a similar range with no major effect of inference time on CNN classes. The highest inference time was observed for VGGNet 398–458 milliseconds. This was mainly due to its higher number of learning parameters in one forward pass in comparison with GoogleNet and EfficientNet. EfficientNet inference ranged from 5.95–6.53 FPS. The lowest inference time was recorded for GoogleNet when compared with other CNNs. However, the detection accuracy of the GoogleNet was lower when compared with other CNNs. VGGNet performed well in terms of disease detection accuracy as suggested by *FScore*, *recall*, and *precision*; however, higher inference time makes it lower priority candidate for real-time applications. GoogleNet and EfficientNet can be used with a tradeoff between model

accuracy and inference speed for real-time smart sprayer development. The inference times for all the CNNs were significantly different; however, the lower differences were observed between GoogleNet and EfficientNet. Bianco et al. [28] reported similar ranges of the inference time for GoogleNet and VGGNet. Overall, the results of accuracy in combination with the statistical measure for various CNN architectures and inference time suggested that the EfficientNet can be deployed and integrated to develop a smart sprayer for targeted application of fungicides in potato fields.

5. Conclusions

Several combinations of datasets with different CNNs were tested for the identification and classification of early blight disease in the potato production system. All CNNs were able to accurately identify diseased and healthy plants for collected data with variable accuracy, *precision*, *recall*, and *FScore*. Results suggested that the performance accuracy of EfficientNet and VGGNet was higher in terms of disease detection at different stages when compared with GoogleNet. GoogleNet was observed to have the lowest inference time compared with EfficientNet and VGGNet. Considering the disease detection accuracy of the EfficientNet during different stages of disease evolution, as illustrated by statistical measure (*recall*, *precision*, and *FScore*), in conjunction with reasonably lower inference time supported the use of EfficientNet for further processing integration into a variable rate smart sprayer for targeted applications of fungicides within potato fields. Site-specific applications of fungicides using a machine vision and DL-based sprayer, only in diseased areas of the potato fields, have a significant potential to lower the agrochemicals in agricultural fields. Targeted application of fungicides can directly be translated into farm profitability and reduced unintended risk to the environment such as runoff of fungicides into water bodies and negative impacts on aquatic habitats. In the case of more than one disease, the potato plant species, and ground surface cover, the sprayer nozzle may get mixed signals of detection (of other weeds and/or diseased plants). However, in the potato field of North America, early blight is the most common potato disease and has been addressed in this study. Future research will cover more diseases and varieties of potato and other cropping systems.

Future Work

The trained EfficientNET CNN will be deployed and integrated into hardware (nozzles, solenoid valves, relays, flow meters, wiring, toggle switches, controller, etc.) using custom-developed software to develop a machine vision and DL-based smart sprayer for targeted application of fungicides in potato fields. The components of the developed sprayer will be calibrated before installation into the smart sprayer. Real-time image acquisition, processing, and delay time will be calculated to accurately synchronize the solenoid valves in the areas where early blight disease has been detected. Extensive lab and field evaluations will be performed to testify the performance accuracy of the developed smart sprayer. The development of this smart sprayer using machine vision, CNNs, and DL will be a breakthrough for the potato industry. Other herbicides and fungicides applications will be explored during the later stages of this development process.

Author Contributions: Conceptualization: A.A.F.; methodology: A.A.F., H.A., and A.W.S.; software: A.A.F., H.A., A.W.S., and F.A.; validation: A.W.S., A.M.-G., and B.A.; formal analysis: H.A., T.E., and F.A.; investigation: A.A.F., H.A., and F.A.; resources: A.A.F.; data curation: H.A.; writing—original draft preparation: A.A.F., H.A., T.E., F.A., B.A., and N.H.; writing—review and editing: A.A.F., H.A., A.W.S., A.M.-G., T.E., F.A., B.A., and N.H.; supervision: A.A.F., A.W.S., A.M.-G., T.E., and B.A.; project administration: A.A.F.; funding acquisition: A.A.F. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Agriculture and Agri-Food Canada (AAFC). Potato Market Information Review 2016–2017. 2018. Available online: https://www5.agr.gc.ca/eng/industry-markets-and-trade/canadian-agri-food-sector-intelligence/horticulture/horticulture-sector-reports/potato-market-information-review-2016-2017/?id=1536104016530#a1.2.3 (accessed on 15 January 2020).
2. Statistics Canada. Area, Production and Farm Value of Potatoes. 2018. Available online: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210035801 (accessed on 15 January 2020).
3. Fernandez-Cornejo, J.; Nehring, R.; Osteen, C.; Wechsler, S.; Martin, A.; Vialou, A. Pesticide use in U.S. agriculture: 21 selected crops, 1960–2008. *USDA-ERS Econ. Inf. Bull.* **2014**. [CrossRef]
4. McKenzie, C.; Woods, S. Potential of precision farming with potatoes. In *Handbook of PA: Principles and Applications*; Food Products Press: New York, NY, USA, 2006; pp. 379–391.
5. Franc, G.D.; Christ, B.J. *Compendium of Potato Diseases*, 2nd ed.; American Pythopathological Society: St. Paul, MN, USA, 2001; pp. 22–23.
6. Kemmitt, G. Early blight of potato and tomato. *Plant Health Instr.* **2002**. [CrossRef]
7. Ferentinos, K.P. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **2018**, *145*, 311–318. [CrossRef]
8. Abdullah, N.E.; Rahim, A.A.; Hashim, H.; Kamal, M.M. Classification of rubber tree leaf diseases using multilayer perceptron neural network. In Proceedings of the 5th Student Conference on Research and Development, SCORED, Shah Alam, Selangor, Malaysia, 11–12 December 2007. [CrossRef]
9. Sena, D.G.; Pinto, F.A.C.; Queiroz, D.M.; Viana, P.A. Fall Armyworm Damaged Maize Plant Identification using Digital Images. *Biosyst. Eng.* **2003**, *85*, 449–454. [CrossRef]
10. Camargo, A.; Smith, J.S. An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosyst. Eng.* **2008**, *102*, 9–21. [CrossRef]
11. Boese, B.L.; Clinton, P.J.; Dennis, D.; Golden, R.C.; Kim, B. Digital image analysis of Zostera marina leaf injury. *Aquat. Bot.* **2008**, *88*, 87–90. [CrossRef]
12. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
13. Krizhevsky, A.; Sutskever, I.; Hinton, E.G. ImageNet Classification with Deep Convolutional Neural Networks. *J. Geotech. Geoenviron. Eng.* **2012**, *12*, 04015009. [CrossRef]
14. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [CrossRef] [PubMed]
15. Gu, J.; Bradbury, J.; Xiong, C.; Li, V.O.; Socher, R. Non-Autogressive Neural Machine Translation. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018. Available online: https://arxiv.org/pdf/1711.02281.pdf (accessed on 21 February 2019).
16. Oppenheim, D.; Shani, G. Potato Disease Classification Using Convolution Neural Networks. *Adv. Anim. Biosci.* **2017**, *8*, 244–249. [CrossRef]
17. Carranza-Rojas, J.; Goeau, H.; Bonnet, P.; Mata-Montero, E.; Joly, A. Going deeper in the automated identification of Herbarium specimens. *BMC Evol. Biol.* **2017**, *17*, 181. [CrossRef] [PubMed]
18. Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* **2017**, *17*, 2022. [CrossRef]
19. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. [CrossRef]
20. Ketkar, N. Introduction to PyTorch. In *Deep Learning with Python*; Apress: New York, NY, USA, 2017; pp. 195–208. [CrossRef]
21. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. 2015. IEEE. Available online: https://storage.googleapis.com/pub-tools-public-publication-data/pdf/43022.pdf (accessed on 14 February 2019).
22. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]
23. Ghazi, M.; Yanikoglu, B.; Aptoula, E. Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing* **2017**, *235*, 228–235. [CrossRef]
24. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
25. Tan, M.; Le, Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 2019. Available online: http://arxiv.org/abs/1905.11946 (accessed on 25 December 2019).
26. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]

27. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]
28. Bianco, S.; Cadene, R.; Celona, L.; Napoletano, P. Benchmark Analysis of Representative Deep Neural Network Architectures. *IEEE Access* **2017**, *6*, 64270–64277. [CrossRef]