

Prediction and classification of wheat plant disease in Pakistan using deep learning techniques

Muhammad Usman Malik ^{a, b*}, Dr Muhammad Faisal Hayat ^a

^a Department of Computer Engineering, University of Engineering and Technology, Lahore, Pakistan

^b Department of Computer Engineering, FAST University, Peshawar

* Corresponding author: Muhammad Usman Malik, Email: malikusman485@yahoo.com

Received: 08 June 2022, Revised: 24 June 2025, Accepted: 30 June 2025, Published: 01 July 2025

KEYWORDS	ABSTRACT
Wheat plant disease Deep Learning Convolution Neural Network Classical data augmentation Wheat Database	Agriculture is the backbone of the economy in many developing countries, including Pakistan. The most economically important crops in Pakistan are wheat and cotton. However, the production of these crops is severely affected by diseases caused by pests, pathogenic bacteria, and viruses, which significantly impact a country's economic growth. Over the last decade or so, numerous image-processing techniques and deep-learning models have been developed by researchers for the identification of plant diseases; however, wheat plant disease has been very rarely addressed. In this research, several deep learning models were trained on a dataset of wheat plant disease images, which were collected from research labs and by taking images through Android mobile cameras from different regions of Khyber Pakhtunkhwa and Punjab. The dataset contains 650 images with 5 classes of wheat diseases, which is increased to 3500 images by using the classical augmentation technique to resolve the problem of overfitting. Among all the deep learning models, ResNet50 and VGG16, with fine-tuning, achieve outstanding accuracies of 95% and 93% on test data, respectively, which outperform the results of previous research conducted by the researcher using WDD17 and other datasets.

1. Introduction

Agriculture is considered the backbone of the economy in the least developed countries like Pakistan, which is heavily dependent on crops like wheat, sugarcane, and cotton. The contribution of Pakistan's agriculture to GDP (Gross Domestic Product) is 21% with 2.7% of annual growth. [1]. This research primarily focused on wheat plants. In Pakistan, 80% of the farmers cultivate wheat crops on land of around 9 million hectares in the 'Rabi' season. The contribution of wheat to the GDP growth is

around 3.1% [2]. The production of wheat is expected to be 27 million tons in 2021-2022, an increase of 8% is noticed in the production as compared to the previous year. With the increase in the country's population, day by day, the demand for wheat also increases. But in Pakistan, this demand is not fulfilled. To fulfill this demand, the Government of Pakistan has imported up to 3 million duty-free in the year 2021-22 to fulfill the country's domestic requirements. [3]. The major cause of the lower production of wheat is the diseases that are caused by different pests and insects. In Pakistan, the most common diseases that affect

wheat production are yellow or brown rust, black bacteria, and powdery mildew, as shown below:

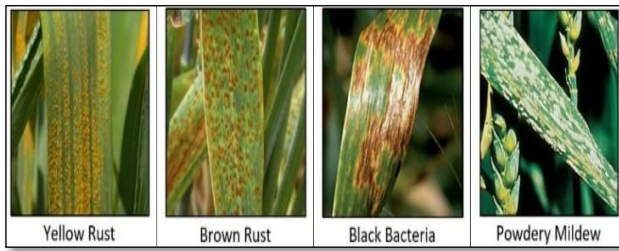


Fig. 1. Wheat Plant Disease images

Early and precise identification of these diseases is an important step in the crop protection system. [4]. In traditional practices, diagnosing plant disease is conducted by visual observation and through laboratory testing. Visual observation of disease requires expertise, and it may vary subject on the person, which may lead to inaccuracies. Conversely, diagnosing through the laboratory is a time-consuming process. [5]. For the better production of a crop, it is important to develop an algorithm that identifies the disease at early stages and then sprays the relevant pesticides correctly and appropriately. An Image-based machine learning approach is the best technique for the early and accurate detection of plant diseases. The detailed study of machine learning and deep learning techniques used by the researchers is discussed in the literature survey section. Limited research has been done on the multiclass identification of wheat disease; most of the research has been done on binary class classification. Furthermore, the datasets used in previous research are either PlantVillage or Google-generated datasets, which contain a lot of raw data and a limited number of wheat plant disease images.

The major contribution of this research is: The dataset is made by collecting affected wheat plant images from different research labs and by taking in-field images in the region of Punjab and KP using an Android mobile camera. The dataset contains 650 images with 5 different classes. The images in the dataset are real images with background information to make the model identify real-time images. In contrast, many known datasets contain either just stripes of wheat leaf or images with a plain background.

Training of several deep-learning models using a multiclass dataset developed for this research. The proposed model, ResNet50, achieves the highest accuracy of 95% on test data, which outclasses the results of previous research.

The remaining paper is organized as follows: a literature survey of previous research done on the prediction and classification of plant disease using machine learning and deep learning models is discussed in Section 2. The proposed methodology of our research for the identification of wheat plant disease is explained in Section 3. Results and discussion of our experimental work are presented in Section 4. Finally, Section 5 presents the conclusion (summarizing the research) and future work (offering suggestions for future improvements).

2. Literature Survey

Over the last decade or so, many researchers have applied deep learning and machine learning techniques for the detection of plant disease. Image-based deep-learning techniques are considered the best for the diagnosis of a disease. The most common machine learning algorithms used are K-nearest-neighbour (K-NN) classifier, K-means clustering, and Support vector machine (SVM), and the most popular for image-based recognition and classification are CNN models.[6].

Several deep-learning models have been developed for the diagnosis of plant disease. Johannes et al proposed a computer vision-based technique for the early and accurate detection of wheat disease in the early stage of growth. The dataset used in this research was made by the author by collecting 3500 affected images of wheat plants from two different sites: Spain and Germany, using different mobile cameras, and achieved an accuracy of 80% on test data. [7].

Picon et al predict wheat leaf disease using a residual convolutional neural network and get the optimal performance by training their model using a dataset of 8178 images taken from two different pilot sites in Germany and Spain. They focus on three diseases, i.e., tan spot, Septoria, and rust. Before the training of the proposed model training dataset was labeled by some experts, every infected spot was segmented, and the region containing the leaf image was also segmented to get the best training and testing accuracy. The author has also developed a mobile application for automatic disease detection. The proposed model gives an accuracy of 87% [8].

Ferentinos et al trained their proposed CNN model using a dataset of 87848 images containing 25 different classes of plants with healthy and affected images. Various deep-learning models were trained using a training dataset. By comparing the accuracy of different models, VGG16 achieves the best accuracy

of 99% on the training dataset, which is quite effective. [9].

Kumar et al. proposed a novel deep learning model, N-CNN, for the diagnosis of powdery mildew wheat disease, which is mainly caused by fungi, using a dataset of 450 images collected through different sources. The author has used the Normalization technique as pre-processing for the enhancement of images. The dataset was split into training and testing, having 360 images for training and 90 images for testing purposes. Through the proposed model, the author gets an accuracy of 89% [10].

Bao et al. work on two classes of wheat disease, powdery mildew and stripe rust, and develop an E-MMC (Elliptical-Maximum Margin / Criterion) algorithm for the identification of the disease and then classifying the severity level of that disease. As compared to other matrices, the elliptic matrix, along with MMC, describes the nonlinear transformation that shows the spatial structure or gives semantic detail of the affected wheat leaf image, which can increase the distance between different classes and hence complete the identification task. The experimental results show that the proposed algorithm performs better as compared to traditional ones. The algorithm achieves the highest accuracy of 94.16% for the identification of wheat leaf diseases. [11].

Lu et al. proposed two CNN models, VGG-CNN-VD16 and VGG-CNN-5, for wheat disease identification and achieving a remarkable accuracy of 97% and 95% on testing data. The WDD17 multiclass images dataset was used for the training of models. Images were collected through mobile and resized by pre-processing. Features were extracted through a convolutional neural network framework. Whereas the models were not fully reliable, as the quality of images was very bad, and the dataset contained a lot of raw images. However, the results obtained were better than those obtained by the old methods of classification and identification. The proposed work is also packed in a real-time mobile application supporting smart agriculture disease detection. [12].

3. Methodology

The proposed methodology is divided into two stages. The first stage is data engineering, in which images are collected for the creation of a dataset and then pre-processed, whereas in the second stage, I will train CNN models on a self-made dataset to get the highest

accuracy. Fig. 2 illustrates the flowchart of the proposed methodology.

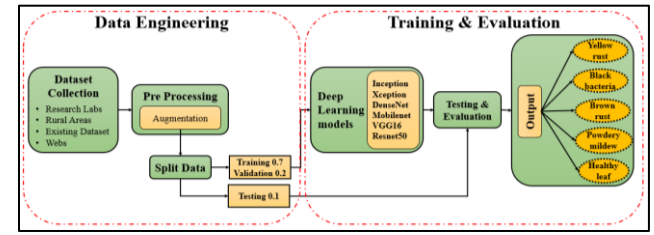


Fig. 2. Flow chart of the research work

3.1 Environment Setting

The experiment was conducted using TensorFlow and Keras in the Google Colab environment.

3.2 Dataset Acquisition

The Dataset is created by collecting wheat leaf disease and healthy images through primary and secondary sources. Primary Sources include in-field images taken through different mobile cameras in multiple cities of two different provinces, i.e., Punjab and Khyber-Pakhtunkhwa, including Peshawar, Dera Ismail Khan, Lahore, and Northern Areas of KP. Additionally, images were obtained from the agriculture research departments of KP, which include the Agriculture University Peshawar and NIFA (Nuclear Institute for Food and Agriculture). The dataset also includes high-quality images from secondary sources like Kaggle, UCI, Plant Village, and other authentic online sites. Sample training images are shown in the Fig. below:



Fig. 3. Wheat Diseases sample training images.

For the training of models, 5 classes of wheat disease are used, namely: Yellow Rust, Brown Rust, Black Bacterial, Powdery Mildew, and Healthy images. The images in the dataset are of different sizes and were later normalized in image augmentation.

Table 1 shows the number of images before and after augmentation. Initially, the dataset contained 650 images of five different classes. The Classical augmentation method was applied to overcome the problem of data scarcity and overfitting, resulting in a total of 3500 images, which will be later divided into Training, Validation, and Test sets with a ratio of 70%, 20%, and 10% respectively.

Table 1

Images Before and After Augmentation

Class Name	Image After Augmentation		
	Original Images	Augmented Images	Total Images
Yellow Rust	150	700	3500
Brown Rust	150	700	
Black Bacteria	100	700	
Powdery Mildew	100	700	
Healthy Leaf	150	700	

3.3 Data Pre-processing

The dataset used in this research contains a lot of noisy data because images were taken from different fields and research labs, and also, for the training of a deep learning model, a huge amount of data is required. So, from existing images, more images were generated by using the data augmentation technique.

3.3.1 Data Augmentation

Data augmentation is part of Pre-processing, which is applied to image datasets to overcome the problem of overfitting and helps in improving the performance of training different deep learning models. Augmentation is required in two cases: either the data is very low, or the data contains noisy images. So, to address the limited size of the dataset, classical augmentation techniques such as flipping, zooming, and rotations were applied using image processing standard functions.

3.3.2 Dataset Splitting

The dataset was further divided into three parts: Training, Testing, and Validation sets to evaluate the performance of models. This splitting is done after the augmentation of data for the better performance of models. The splitting of the dataset is shown in Table 2.

Table 2

Data Splitting After Augmentation

Class Name	Data Splitting After Augmentation		
	Training	Validation	Testing
Yellow Rust	490	140	70
Brown Rust	490	140	70
Black Bacteria	490	140	70
Powdery Mildew	490	140	70
Healthy Leaf	490	140	70

The dataset contained images of different sizes as they were collected through different sources, but were resized to the default size of 224 x 224 x 3 in the process of normalization to make the training and testing of models efficient.

3.4 Deep Learning-based Classification Approach.

In this research, seven CNN models were trained, such as ResNet50, VGG16, InceptionV3, DenseNet, Xception, and InceptionResnetV2, using a self-collected dataset which are based on Convolution, Max-pooling layers, and ReLu or SoftMax as activation function. These pre-trained models give different layers and weights related to them. These weights are used in the formation of deep learning models. The ReLU activation function is used to remove the non-linearity from the data. The combination of different convolution layers followed by the dense and fully connected layers provides different features of the input images. In the last layer, SoftMax is used as an activation function because this research focused on 5 classes, but models like ResNet50, VGG, and others classify 1000 classes of different images. The proposed methodology with the CNN models as classifiers is shown in Fig. 4:

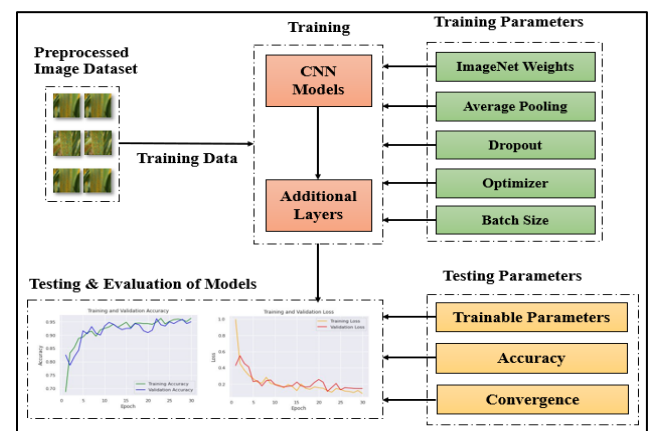


Fig. 4. Graphical description of the CNN model

3.5 Machine Learning Models

Several machine learning models were observed, which were previously used in research done in the

field of Agriculture by Researchers. The dataset was trained using selected models with fine-tuning to get efficient results. The detail of machine learning models is discussed below:

3.5.1 Xception

Xception is also known as “Extreme Inception”, having an architecture of 34 convolution layers, max pooling of 3x3, and a stride of 2x2 with 14 modules of residual blocks, ending with a fully connected layer of dimensions 2048. ReLU is used as an activation function in all 71 convolution layers.

3.5.2 DenseNet121

DenseNet is a newly proposed Convolution Neural Network model in the field of object recognition and is similar to ResNet in its architecture, with some differences. The previous layer of output is concatenated to the next layer of the model. [15].

3.5.3 InceptionResNetV2

InceptionResNet-V2 is a deep-learning CNN model trained on the ImageNet dataset, which contains millions of images. The network has deep layers of 164 and classifies over 1000 classes of different objects. The network learns rich features from a huge number of images of the ImageNet dataset. The input size for InceptionResNet-V2 is 299 x 299. It is the combination of two parts: the first part is known as Inception, and the second is the residual part. Residual networks are combined with convolution filters of multiple sizes in the InceptionResNet block. The main usage of the residual network is to avoid degradation issues occurring in deep architecture and minimize the training duration. The architecture of InceptionResNet-V2 is shown in the Fig. below:

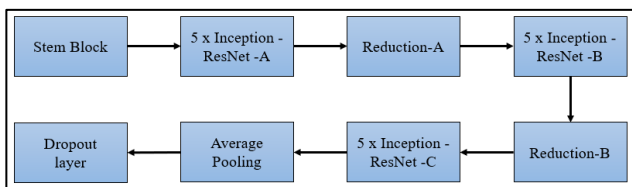


Fig. 5. The basic architecture of Inception-Resnet-v2.

3.5.4 VGG16

VGG16 is one of the more widely used deep learning models used by researchers in the field of agriculture for the detection and classification of different diseases. Its architecture is based on the fact of increasing the number of layers of the neural network. It is considered the best performance model. The advancement in this model is increasing the number of layers but decreasing the size of filters for building models. By decreasing the size of the filter, we can

detect a small portion of the required image, hence giving more detailed features. Regularization of the model is required after an increase in the depth of the model. VGG16 model has filter size and max-pooling dimensions of 3x3 and 2x2, respectively [16]. It works on the principle of stride and padding of one, which works on Windows. When the sliding window moves, a new feature chart is developed, which will act as input to the model through which the model will be trained.

3.5.5 ResNet50

It is one of the popular deep learning models, whose input layer dimensions are 224x224x3. It consists of 7x7 convolution and 3x3 max pooling kernel sizes. Weights are initialized for the training of new models for different datasets. Sometimes, models converge during training to a specific value, which means there is a huge gap between training and testing errors, which indicates that the model does not optimize to the given dataset. The errors in training a model are reduced by Backpropagation algorithms, which help in balancing weights for the model.

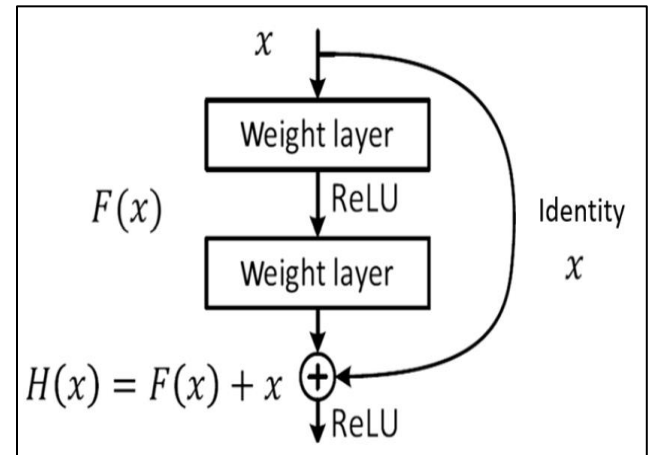


Fig. 6. Resnet50 [22]

3.6 Experimental Setup

The top layer from the proposed model was removed, whereas the other layers are already trained on ImageNet and are retrained to save time and effort. The fully connected layer of the model has 1000 classes, but in this research, only 5 classes were considered, so this final layer is replaced with a new layer having 5 output neurons. Adam optimizer with a learning rate of 0.001 was considered for all the models, which converges the model towards the minima with a fast speed. Categorical cross-entropy was used as a loss function for the multi-class wheat disease classification. The training of the model is

halted when the validation accuracy stops improving with upcoming epochs, which will help in the improvement of the training process and prevent the model from overfitting. Two models, ResNet50 and VGG16, achieve an accuracy of 98% and 96% on training data after 20 epochs, respectively.

For better convergence of the model, a batch size of 16 was set for all the models. All the images are normalized by dividing their pixel values by 255. For image processing, a built-in function called ImageDataGenerator() is used, which augments and resizes images in the dataset.

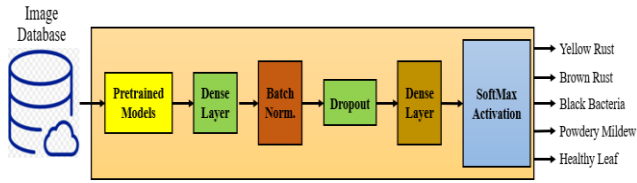


Fig. 7. Proposed Deep Learning Model Architecture

4. Simulation and Results

Many CNN models have been used by researchers for the detection and classification of wheat plant disease. This research proposed some deep learning models trained on self-collected datasets with fine-tuning to get the best-performing model. ResNet50 and VGG16 give an outstanding accuracy of 95% and 92 % on testing data, respectively. Table 3 shows the accuracy of different models. Fig. 8 shows the accuracy and loss graph of Resnet50, which shows that after 30 epochs accuracy and loss are stable and show a straight line.

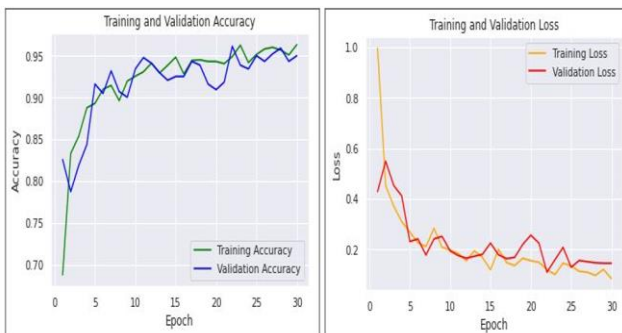


Fig. 8. The left shows the Accuracy graphs, and on the right side are the Loss graphs of ResNet50. [17]

The proposed model, ResNet50, was trained on a newly created dataset after augmentation to minimize the problem of overfitting and achieves the highest performance among all seven models in terms of accuracy, as shown in Table 3.

Table 3

Results of Deep Learning models on a self-collected dataset

Models	Trainin g Accurac y	Test Accurac y	Precisio n	Reca ll
Xception	0.77	0.65	0.63	0.68
MobileNet	0.90	0.71	0.731	0.80
InceptionResN et-V2	0.85	0.67	0.742	0.83
EfficientNetB0	0.93	0.86	0.85	0.88
DenseNet121	0.75	0.71	0.69	0.73
VGG16	0.93	0.89	0.90	0.91
ResNet50	0.96	0.95	0.95	0.94

The confusion matrix of ResNet50 is shown in Fig. 9. The confusion matrix is used for the evaluation of a model. Accuracy, Recall, and Precision of any model are calculated using the equations:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total number of Samples in Dataset}} \quad (3.1)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (3.2)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3.3)$$

In Fig. 9, the diagonal values are accurate, or in other words true positive values of all classes. For example, brown rust has 36 images that are predicted correctly by the model, whereas 6 of the brown rust images are predicted as yellow rust and are known as true negative values; also, 1 image of yellow rust is predicted as brown rust, which is called a False-positive value. Hence, the confusion matrix of any model gives complete details about the model and helps in the evaluation of a model.

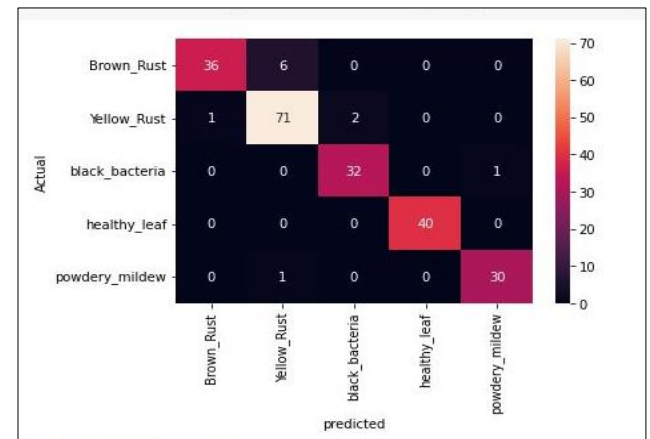


Fig. 9. Confusion matrix of ResNet50

The performance of the proposed models is shown in Fig. 10 by implementing the bar graph of the Test,

Recall, and Precision accuracies of the models. ResNet50 with the highest Recall and Precision accuracy shows that it is best fitted on the Self-collected dataset.

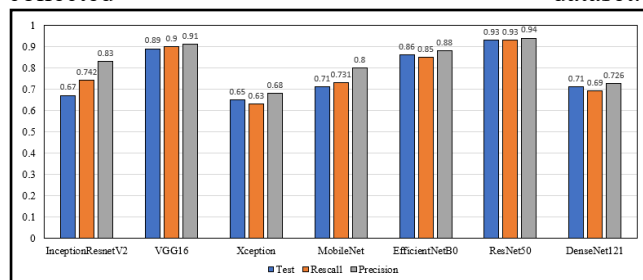


Fig. 10. Performance analysis of deep learning models

The comparison of test accuracies of all the models is shown in Fig. 11, and from the Fig., the CNN models were analysed, having ResNet50 giving the highest test accuracy among all the models. In previous research, most of the researchers have done binary classification using CNN models, and limited research has been done on multiclass classification using different datasets. Those datasets contain either raw images of wheat disease or laboratory images, whereas the dataset developed for this study contains original clear field images, and the results of the proposed model outclass the results of previous research.

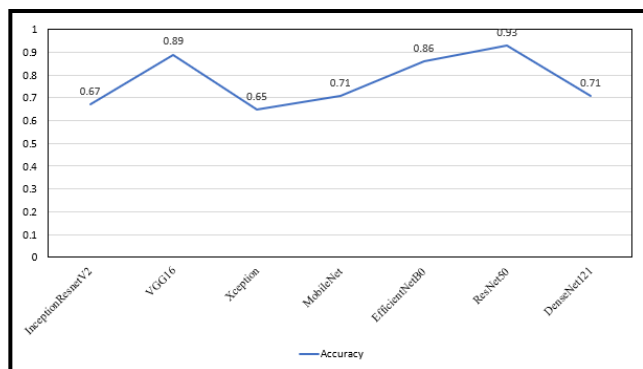


Fig. 11. Test Accuracies of deep learning models

5. Conclusion

This research presents an efficient method for the prediction and classification of Wheat plant disease using deep learning algorithms. The training of the deep-learning model is made by collecting affected images of wheat plants from different research labs of Khyber Pakhtunkhwa and by taking images of wheat disease using Android mobile cameras in the regions of Khyber Pakhtunkhwa and Punjab. 600 images were collected, which were later increased to 3500 using the classical augmentation technique to overcome the problem of overfitting. In this research, seven CNN models were trained on a dataset collected specifically for this research i.e. Xception, MobileNet, InceptionResnetV2, EfficientNetB0, DenseNet121, VGG16, and ResNet50 which gives an accuracy of 65%, 71%, 67%, 86%, 72%, 71%, 93% and 95% on

test data respectively as shown in table 5. From these results, it is concluded that among the seven models evaluated, the ResNet50 CNN model with fine-tuning achieves the highest accuracy on test data. In the future, the dataset can be expanded by taking in-field images from all over Pakistan. Additionally, the trained model can be deployed as an API, which can later be integrated into Android applications or drone cameras for remote prediction and classification of wheat disease in the field, and taking precautionary measures on time for better production of wheat crops.

6. References

- [1] Malik, S.J., H. Nazli, and E. Whitney, The official estimates of poverty in Pakistan—What is wrong and why?—Illustrations using the Government of Pakistan’s Household Integrated Economic Survey 2010-11. *Economic Survey*, 2010. 11.
- [2] Gardner, B.L., et al., *Handbook of Agricultural Economics: Agricultural Development: Farmers, Farm Production and Farm Markets*. Vol. 3. 2001: Elsevier.
- [3] Culas, R.J., and I.A. Baig, Impacts of irrigation water user allocations on water quality and crop productivity: The LCC irrigation system in Pakistan. *Irrigation and Drainage*, 2020. 69(1): p. 38-51.
- [4] Su, T., et al., A CNN-LSVM model for imbalanced images identification of wheat leaf. *Neural Network World*, 2019. 29(5): p. 345-361.
- [5] Bhardwaj, S.C., et al., Status of wheat rust research and progress in rust management-Indian context. *Agronomy*, 2019. 9(12): p. 892.
- [6] Lin, Z., et al., A unified matrix-based convolutional neural network for fine-grained image classification of wheat leaf diseases. *IEEE Access*, 2019. 7: p. 11570-11590.
- [7] Johannes, A., et al., Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Computers and electronics in agriculture*, 2017. 138: p. 200-209.
- [8] Picon, A., et al., Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, 2019. 161: p. 280-290.
- [9] Ferentinos, K.P., Deep learning models for plant disease detection and diagnosis. *Computers and*

Electronics in Agriculture, 2018. 145: p. 311-318.

- [10] Kumar, D. and V. Kukreja. N-CNN-Based Transfer Learning Method for Classification of Powdery Mildew Wheat Disease. In the 2021 International Conference on Emerging Smart Computing and Informatics (ESCI). 2021. IEEE.
- [11] Bao, W., et al., Identification of wheat leaf diseases and their severity based on elliptical-maximum margin criterion metric learning. Sustainable Computing: Informatics and Systems, 2021. 30: p. 100526.
- [12] Lu, J., et al., An in-field automatic wheat disease diagnosis system. Computers and electronics in agriculture, 2017. 142: p. 369-379.
- [13] Carneiro, T., et al., Performance analysis of Google Collaboration as a tool for accelerating deep learning applications. IEEE Access, 2018. 6: p. 61677-61685.
- [14] Mbugua, J.K., Deep Learning for Early Detection, Identification, and Spatiotemporal Monitoring of Plant Diseases Using Multispectral Aerial Imagery. 2021, The Claremont Graduate University.
- [15] Mi, Z., et al., Wheat stripe rust grading by deep learning with attention mechanism and images from mobile devices. Frontiers in Plant Science, 2020. 11.
- [16] Aboneh, T., et al., Computer Vision Framework for Wheat Disease Identification and Classification Using Jetson GPU Infrastructure. Technologies, 2021. 9(3): p. 47.
- [17] Guo, A., et al., Identification of wheat yellow rust using spectral and texture features of hyperspectral images. Remote Sensing, 2020. 12(9): p. 1419.