Adaptive Learning Model for Detecting Wheat Diseases

Mohammed Abdalla, Osama Mohamed, Elshaimaa M. Azmi
Faculty of Computers and Artificial
Intelligence, Beni-Suef University, Egypt

***Abstract*—Nowadays, the wheat plant has been considered a crucial source of protein, energy, and micronutrients for people. The motivation behind this study comes from how to increase the wheat crop growth and prevent wheat diseases as this plant plays a significant impact on food security all over the world. Wheat plant diseases can be divided into fungal, bacterial, viral, nematode, insect pests, physiological and genetic anomalies, and mineral and environmental stress. Digital images containing the wheat plant disease are collected from different public sources like Kaggle and GitHub. In this study, an adaptive deep-learning model is developed to classify and detect various types of wheat diseases collected digitally in an efficient accurate manner. The dataset is split into two sets: approximately 80% of the data (8,946 images) for the training set and 20% (2,259 images) for the validation set. The training set is composed of 1445, 1478, 1557, 1510, 1424, and 1532 images of healthy, leaf rust, powdery mildew, septoria, stem rust, and stripe rust while the validation set contains 357, 360, 404, 402, 353 and 383 images respectively. The suggested method achieved 97.47% validation accuracy on the training set of images and a testing accuracy of 98.42% on the testing set. This study offers a method of training for the classification and detection of wheat diseases using a mix of recently established pre-trained convolutional neural networks (CNN), DenseNet, ResNet, and EfficientNet integrated with the one-fit cycle policy. In comparison to the current state of the art, the proposed model is accurate and efficient.**

Keywords*—Food security; image recognition; deep learning; conventional neural networks; digital agriculture; agriculture sustainability

I. Introduction

Recently, there have been many production and economic losses around the world due to several agricultural crop diseases. Indeed, the wheat plant is one of the primary crops grown worldwide and a major source of food for humans, considering that it is the second largest crop in the world, providing 19% of people's calorie intake [1], [2]. In study [3], the authors emphasized that between 26 to 30 percent of the world's yearly wheat crop is lost to wheat diseases. Additionally, they mentioned that wheat disease losses can account for up to 70% of the wheat output if plant protection technologies are not used to manage fields.

Indeed, this paper is motivated by the desire to handle and detect wheat disease which can lead to high crop growth increase by using deep learning techniques. This paper demonstrated how deep learning ideas in artificial intelligence and computer vision have become a potential remedy for a variety of issues in agriculture. Convolution neural networks (CNN) have recently studied the use of digital imagery for autonomous disease detection in crops. By using a convolutional neural network (CNN), the characteristics and features will be learned automatically rather than by human presence, and this will save time, and costs and help the farmer take quick action to treat the wheat disease in the early stage. CNNs apply several convolutions to extract important features from images [4], [5], [6], [7], [8], [9].

This paper addressed only the fungal wheat diseases which include the following: Powdery Mildew, Leaf Rust, Stem Rust, Stripe Rust, and Septoria. Table I is a summarized table of discussed wheat fungal diseases, their pathogens with scientific names, and the visual symptoms observed on infected plants. This table is designed to provide a quick reference for readers interested in wheat pathology.

This paper proposes a revolutionary model that utilizes the transfer learning concept rather than training CNN from scratch which requires a massive amount of data and robust computer hardware (GPUs) to be trained. A CNN model is trained on a sizable dataset to become a pre-trained model in the proposed method. Next, learned features by this pre-trained model are transferred to the new model. After that, the fit-one-cycle policy technique is used to adjust deep learning models' hyperparameters. Tuning CNN hyperparameters is a challenge because it requires more time and experience to tune them. A fit-one-cycle policy shortens the training period while enhancing performance[4].

1) Contributions: The following is a summary of this paper's significant contributions:

- We develop a deep learning model that identifies wheat plant fungi diseases with the best accuracy achievement.
- The proposed model utilizes a real dataset collected from various sources which contain five types of wheat fungi diseases and healthy ones.
- The proposed model handles the data imbalance common issue which is a known issue in several deep learning techniques by using a robust data augmentation technique.
- A detailed comparison between different CNN pre-trained models applied on the real dataset to demonstrate the performance differences, evaluating the generalization ability and training error of these models.
- Finally, the proposed model employs the fit-one-cycle policy method which automates hyperparameter learning to select the best value in the learning

TABLE I. OVERVIEW OF WHEAT FUNGAL DISEASES

Disease Name	Scientific Name	Visual Symptoms on Wheat Plants
Powdery Mildew	Blumeria graminis f.sp. tritici	White, powdery spots on leaves and stems; can lead to yellowing and drying of the tissue.
Leaf Rust	Puccinia triticina	Orange-red pustules on leaves and stems; leads to premature leaf senescence and dropping.
Stem Rust	Puccinia graminis f.sp. tritici	Large, brick-red pustules on stems and leaves; severely infected plants may produce less grain.
Stripe Rust	Puccinia striiformis f.sp. tritici	Yellow-orange stripes or streaks on leaves; can cause significant yield loss.
Septoria	Zymoseptoria tritici (formerly Septoria tritici)	Brown spots with yellow halos on leaves; spots often coalesce causing large areas of dead tissue.

process. This leads to high-performance achievement and optimal training time.

2) Roadmap: Following is the breakdown of the remaining sections: Section II discusses some related work and previous studies in this domain. Section III describes the proposed model and a detailed description of the work methodology. Presentation of the experimental findings and analyses in Section IV. Finally, the paper is concluded in Section V.

II. BACKGROUND

According to the study in [10], fungal, bacterial, viral, nematode, insect pests, physiologic and genetic abnormalities, and mineral and environmental stress are some of the several categories of crop pathogens. These pathogens can lead to damage to any part of the plant whether above or below the ground. Indeed, the major challenge is to identify symptoms and know when and how to effectively control diseases. For this reason, diagnosis of wheat diseases and managing the spread of disease in the early stage is essential for producing healthy wheat products and improvement of wheat yield and quality. Indeed, there is another notable challenge which is to diagnose wheat disease the expert or farmer depends on observing the symptoms of the disease manually, so it takes more time and cost to diagnose a large space and treat the disease. The accuracy of the manual prediction depends upon the experience and knowledge of the person so the unavailability of experts can obstruct the accurate diagnosis and treatment of the diseases in the early stages [11].

One of the risky diseases infecting the wheat yield is fungi diseases. The fungus diseases include powdery mildew, rust, and septoria of the leaves and ears [12]. Fungi indeed represent a separate kingdom of life, distinct for their unique biological and ecological characteristics beyond the absence of photosynthesis. This kingdom encompasses a diverse range of organisms, including molds, yeasts, and mushrooms, which play crucial roles in natural ecosystems and have significant implications for agriculture. The fungi can develop in a variety of ways, including from seeds or soil, or they can be spread by wind, water (either rain or irrigated), and other insects and animals. The overwatering of the host plant region, the host's weak density, and the ambient temperature all affect the fungal infection. Additionally, the fungi did not always destroy the entire crop but rather affected its growth, and the interplay between diseased and healthy plants determines how quickly the disease spreads [13].

Furthermore, there is another risky disease infecting the wheat yield which is Wheat rust disease. This disease can be divided into three rust categories: leaf, stem, and stripe. Indeed, rust diseases can be distinguished from each other based on some symptoms like the color, size, and arrangement of blisters on the plant surface and the plant part that is affected [14]. The rust diseases can be described as follows:

- Small, orangish-brown spots on leaves are symptoms of the disease leaf rust. The leaf sheath, which stretches from the base of the leaf blade to the stem node, can develop round or oval lesions, which are most frequently found on leaves.
- Stem rust disease is characterized by reddish-brown lesions that are oval and extended with tattered edges clearly, on leaves, leaf sheaths, and stems, Stem rust creates lesions that are more extensive than those caused by leaf rust.
- Stripe rust disease is most prevalent on leaves, and it produces yellow blister-like lesions that are grouped in stripes.

Additionally, one of the fungal diseases that infect wheat yield is Powdery mildew which is caused by the fungus Blumeria graminis, and it is most commonly overwinters. Powdery mildew is characterized by white to gray lesions on leaves, and leaf sheaths, It has several quick life cycles over a growing season and, once established, can be quite challenging to control. Septoria is considered a fungal disease that causes tan that is extended on wheat leaves. Although the degree of yellowing varies between kinds, lesions may have a yellow edge [15], [16].

Fig. 1 describes the five types of wheat fungi disease.

III. RELATED WORK

This section describes earlier research using deep learning and machine learning methods to evaluate, segment, and categorize illnesses of wheat crops using digital images.

The prior work in this direction may be divided into three categories: segmentation techniques for Wheat Crop disease, deep learning models to classify Wheat Crop diseases, and machine learning models to classify Wheat Crop diseases.

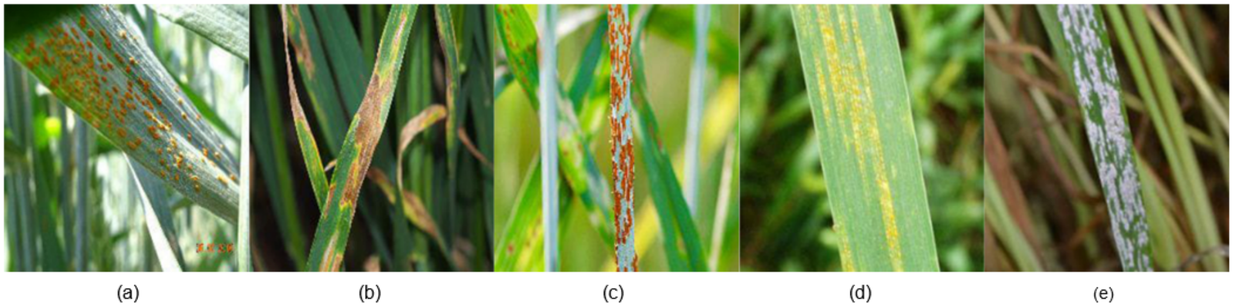


Fig. 1. Examples of wheat fungi diseases.

A. Deep Learning Models to Classify Wheat Crop Diseases from Digital Images

Overall, this section describes the deep learning methods, techniques, and approaches that are proposed to classify the different Wheat diseases from digital images.

In research [17], the classification of Powdery Mildew Wheat Disease was offered by the authors using 450 wheat photos that were gathered from primary (using a camera) and secondary (websites) sources. They used a normalization technique for data preprocessing, and The preprocessed normalized images were input to CNN achieving an accuracy of 89.9% for Powdery Mildew wheat disease. Next, the pre-trained model is applied to the CIAGR pictures dataset using the transfer learning technique, and it achieves 86.5 percent classification accuracy.

In research [18], the authors demonstrated a brand-new deep-learning model that has been trained to categorize 10 different wheat illnesses. The model outperforms two well-known pre-trained deep learning models, VGG16 and RESNET50, in terms of testing accuracy, with a score of 97.88%.

In [19], using 2000 photos of wheat plants for training and testing, the authors presented a Deep Convolutional Neural Network (DCNN) to classify Wheat Rust illnesses. This DCNN obtains an accuracy of 97.16% for wheat rust diseases.

In research [20], the authors presented multi-task and the pre-trained model VGG16 to distinguish between two types of wheat leaf diseases and three types of rice leaf diseases. The multi-task learning method is *alternate learning*. The idea of *alternate learning* makes use of various data sets, each of which has a distinct objective. Mutual training is used to train each job within an epoch, and each time, the parameters of the common layer are modified. Data augmentation is necessary to increase the variety because the data sets for wheat leaf disease and rice leaf disease are both modest and independently collected. For rice leaf diseases, the model's accuracy is 97.22 percent, and for wheat leaf diseases, it's 98.75 percent.

In study [21], the authors described different CNN Models such as ResNet50, DenseNet121, MobileNet, and MobileNetV2 to classify four classes of wheat images: (1) tan spot, (2)fusarium head blight, (3) stem rust, and healthy wheat. They applied Data augmentation to expand the dataset. The maximum accuracy of ResNet50 is 98%.

In study [22], based on the CGIAR dataset, the authors presented the VGG16 model to classify three types of wheat rust diseases: stem rust, leaf rust, and healthy wheat. With an initial learning rate ranging from 0.01 to 0.0001, the suggested model has a classification accuracy of 99.54 percent during training and 77.14 percent during validation on 80 epochs. The authors explain that even though the model had acceptable training accuracy, classifying stem and leaf rust was not appropriate since certain photos in this dataset contained several diseases, which meant that one image comprised the characteristics of both leaf and stem rust.

In study [23], the authors suggested a brand-new CNN model called *CerealConv* that is trained using a dataset of wheat photos captured in actual growth conditions and divided into five categories: "healthy," "yellow rust," "brown rust," "powdery mildew," and "Septoria leaf blotch." With batch normalization, maximum pooling, and dropout, the *CerealConv*'s 13 convolutional layers were able to achieve an accuracy of 97.05%. Four pre-trained networks were used in their experiments: MobileNet, InceptionV3, VGG16, and Xception. On a portion of the dataset's photos, they performed tests against experienced pathologists. In conclusion, the model produced an accuracy score that was 2% greater than that of the top pathologist. Finally, they employed image masks to demonstrate that the model was generating its classifications using the right data.

In study [24], to automatically identify Wheat rusts, the authors presented the EfficientNet model. They created a dataset known as WheatRust21 that included 6556 pictures of healthy and three different types of Wheat rust infections. The EfficientNet-B4 model has a testing accuracy of 99.35% even though they tried many CNN-based models.

In study [25], five fungal diseases of wheat crops, including (1) leaf rust, (2) stem rust, (3) yellow rust, (4) powdery mildew, and (5) septoria, were proposed by the authors. These diseases may be recognized both individually and in cases of multiple infections. In this work, duplicate photos were removed from the training data using the image hashing algorithm. The recognition process makes use of the EfficientNet pre-trained model. The accuracy of the model is (0.942). The recognition strategy was created as a bot for Telegram.

In study [26], for their UAV to be able to recognize three different forms of wheat leaf diseases, the authors developed a two-stage classifier. They first found individual plant leaves using an object detection model, such as the

YOLOV4 or EfficientDet models, and then cropped the image using bounding box coordinates. The cropped photos are then fed into the classification network in the second stage, which will identify the type of disease on the leaf. The EfficientNet-B0 model performs with an accuracy of 99.72 percent, outperforming the YOLOV4-tiny model in object detection.

B. Machine Learning Models to Classify Wheat Crop Diseases from Digital Images

Generally, this section presents the different machine learning techniques that are proposed to classify and analyze the different Wheat diseases from digital images.

In research [27], by combining spectral vegetation indices data from spectrum sensors in Random Forest models, the authors proposed a high-throughput plant phenotyping technique to automate disease scoring of yellow rust in a large plant breeding field trial.

In study [28], for diagnosing wheat leaf diseases and their severity, the authors suggested a method based on Elliptic Maxima Margins Criteria metric training learning. They used information on wheat leaf diseases such as powdery mildew and stripe rust. The gradient rise approach and the greatest margin criteria are used to alter the feature space and decrease feature correlation before creating the elliptical metric matrices. Additionally, using photographs of wheat leaves, the Otsu method is utilized to separate the disease spots according to the characteristics of disease distribution. Their technique outperforms other learning algorithms and conventional support vector machines. They were 94.16 percent accurate.

In research [29], the effectiveness of various Machine Learning and Deep Learning algorithms for identifying plant disease was compared by the authors. In terms of disease prediction accuracy, Deep Learning models surpass Machine Learning models as follows: The following models were successful: VGG-16, Inception-v3, VGG-19, SVM, SGD, and RF (89.5, 89, 87.4, 87.5, 86.5, and 76.8%, respectively). According to the findings, VGG-16 has the highest classification accuracy, while random forest has the lowest.

In research [30], the authors suggested using machine learning to identify the illnesses that cause brown-and-yellow streaks in wheat harvests. By shrinking and segmenting the data, this study pre-processed it. Additionally, to extract elements including shape, texture, and color, they used three feature descriptors: Histogram of Oriented Gradient (HOG), Local Binary Pattern (LBP), and Hue- Moment (HM). They used a number of different models, but the RFC performance delivered the best results when compared to the other models, which had an accuracy rate of 99.8%. A two-stage classifier was also suggested by the study to help the UAV detect plant diseases. After cropping the image with the bounding box coordinates and finding individual plant leaves using an object detection network, the model then utilizes a second classifier to identify the type of illness on the leaf.

In research [31], the authors used classification methods (Artificial Neural Network, Support Vector Machine (SVM), and k-Nearest Neighbor (k-NN) are trained based on morphological features like shape and size to identify wheat crop seed that was derived from the singleton wheat kernel images to identify wheat seeds from three different types: (1) Canadian, (2) Rosa, and (3) Kama. The k-NN classifier outperformed the other two classifiers, producing the greatest classification accuracy of 94.23 percent.

C. Segmentation Methods for Wheat Crop Diseases

This section describes several segmentation methods that are used to extract interest regions from digital images and generate segmented data.

In study [32], the authors demonstrated a deep learning-based semantic segmentation method for Wheat Stripe Rust pictures. They tested four different models: PSPNet, DeepLabv3, U-Net, and Octave-UNet. The Octave-UNet model produced the greatest results of all the models; its accuracy was 96.06 percent, its mean pixel accuracy was 94.58 percent, and its mean intersection over a union was 83.44 percent. The original images were roughly 1000 x 4000 pixels; to avoid significant information being lost due to direct resizing, each original image was divided into several 512 x 512 pixels local images to increase the amount of data, followed by filtering images. They divided the image into three categories: (1) background, (2) leaf, and (3) spore.

In study [33], by using a variety of segmentation techniques, such as Watershed, Grab Cut, and U2-Net, the authors explored the classification of wheat stripe rust into three infection kinds, including healthy, resistant, and susceptible. Multiple segmented datasets are created using these techniques, and the region of interest is then extracted by cropping the segmented images. Then segmented data is produced using the pre-trained ResNet-18 model. On the U2-Net-segmented dataset, the maximum classification accuracy (96.196%) is attained.

In study [34], the authors suggested a Res-capsule network, which was designed to be a segmentation model by replacing the AveragePooling layer of the upgraded ResNet34 with a Capsule network, which can preserve deeper semantic information. This network can segment wheat plantation rows that were photographed by a UAV. They create a threshold after the convolution operation, which they refer to as threshold convolution, in addition to decreasing redundant features and improving effective features. By extracting the textural features (TF), grayscale features (GF), and hue saturation value features (HSV), they increase the accuracy of segmentation. They then input the three extracted features into their enhanced ResNet34.

Table II shows the main characteristics of the reviewed work.

D. Summary

To sum up, this paper differentiates itself from the previous studies by the following: (1) we employed the one-fit cycle to adapt the hyperparameters in an efficient manner which improved the learning process, (2) we studied major types of wheat diseases, and investigated the collected images from different data sources which make our study comprehensive, (3) we prevented the imbalance data issue from occurring in the model developed by introducing a data augmentation approach to fix this, (4) we experimented a large number of different types of deep learning models, and compared them from different perspectives like; accuracy, precision, and recall.

IV. Materials and Methods

This section presents the methodology and phase for the proposed solution.

A. Datasets

This study categorized the images that were gathered into healthy and five different forms of fungal infections, including powdery mildew, septoria, leaf rust, stem rust, and yellow rust types of wheat rust disease.

The following are the many image sources:

- Five fungal diseases of wheat crops are included in the dataset (leaf rust, stem rust, yellow (stripe) rust, powdery mildew, and septoria), both individually and when several diseases are present [35].
- The dataset contains images of yellow-rust(stripe rust), brown-rust (leaf rust) wheat leaf diseases, and healthy wheat leaf[36].
- Images from the CGIAR (Computer Vision for Crop Disease) dataset are included. This collection includes pictures of wheat leaf diseases like stem rust and leaf rust as well as healthy ones.[37], [38].
- There is a wheat leaf dataset on Kaggle that includes pictures of both diseased and healthy wheat leaves, including those with stripe rust and septoria [39].

Table III lists the classes that are gleaned from various sources, along with the number of images gleaned from each source.

B. Data Preprocessing

This phase is a major step in building the proposed learning model. This step includes several tasks which are fundamentals to build a learning model with high accuracy and best performance. These tasks include: data ingest, data cleaning, and data standardization. The data ingest means collecting raw data from diverse sources for further processing. The data clean means removing inconsistencies from collected datasets, handling missing values, and addressing any quality issues. Finally, data standardization means transforming data into a consistent format for seamless processing and Organizing and structuring data for effective feature engineering and model development.

A crucial stage in a model pipeline to find diseases is image pre-processing. as images could contain noise or different sizes. Because the images in the collected dataset come from many sources and have varying sizes and formats, all of the images were initially reduced to 224×224 pixels (resolution) and saved as (.jpg).

There are some images have The height is greater than the width or vice versa, so when resizing this image, the image will expand, and maybe some important features will be lost like the leaf shape which may cause low accuracy, so adding the black border to the image may save the image presentation Fig. 2 describes these stages. Some images of the collected dataset have noises that cause the loss of some important features, contain human hands, or have multiple diseases in the same image which impacts the model accuracy, so the dataset was filtered from these images which give high accuracy.

1) Data augmentation: Because there are more images in certain classes than others (some classes have extremely few occurrences compared to other classes), the suggested dataset is unbalanced. The deep-learning models' performance would be impacted and overfitting would result from this mismatch in the amount of photos in the classes. When a model performs well on the training dataset but poorly on new data, it is said to be overfit. Consequently, a data augmentation strategy is applied to avoid this issue. Data augmentation is the process of creating additional samples from existing datasets by modifying the original images, which enlarges the dataset or increases its volume. To create new photos for the classes of fewer photographs, transformation techniques such as rotation (90 degrees), flipping, and zooming between [0.5,1.5] range were applied.

After pre-processing (filtering) the dataset and adding further data, the suggested dataset has a total of (11,205) images. The proposed dataset was divided into two sets: a training set with about 80% of the data (8,946 photos) and a validation set with around 20% (2,259 images). The validation set includes 357, 360, 404, 402, 353, and 383 images of healthy, leaf rust, powdery mildew, septoria, stem rust, and stripe rust while the training set consists of 1445, 1478, 1557, 1510, 1424, and 1532 images of these conditions.

The number of images in the suggested dataset after pre-processing and data augmentation is displayed in Table IV

C. Proposed Model

1) Convolutional Neural Network (CNN): CNN is frequently utilized in computer vision applications like segmentation, pattern identification, and classification issues. CNN reduces the number of neurons and achieves better learning. Indeed, CNN recognizes the content of the images in three-dimensional volume without converting it to a one-dimensional vector such as multi-layer perceptron(MLP) which becomes computationally expensive because of the huge number of neurons that are needed to recognize small images.

The convolutional layer, activation layer, and pooling layer are the three layers that make up a CNN in general. These layers primarily extract characteristics, which are later used for classification by fully connected layers.

- One of the components of a CNN, the *convolutional layer*, is used to extract significant information from an image using a convolution process. One value is produced by the convolution process, which is a dot product between two matrices. Every input image is represented by a matrix of pixel values and another matrix called the filter matrix. The filter matrix is also known as a kernel, or a filter made up of learnable weight values. The kernel is a small matrix, and it is sliding over the input matrix by one pixel which creates a new matrix called a feature map or activation map that represents the extracted features. Utilizing many

TABLE II. PRIMARY CHARACTERISTICS OF THE RELATED WORK THAT WAS REVIEWED (SUMMARY)

Paper	Method	Accuracy	Dataset	Volume
[18]	CNN architecture.	97.88%	LWDCD2020 (10 classes)	12000
[17]	Normalization technique for preprocessing and CNN	89.9% and 86.5% CIAGR images	CIAGR images	450 wheat images, 101 (CGIAR)
[19]	Deep convolutional Neural network (DCNN)	97.16%	CGIAR dataset & secondary resources	2000
[20]	Alternate learning and VGG16	98.75% for wheat leaf diseases	public data sets in the UCI machine learning database and public images found on the Internet	200
[21]	AResNet50, DenseNet121, MobileNet, and MobileNetV2	998%, 90%, 91% and 89%	Kaggle dataset	2015
[22]	VGG16	99.54% in training and 77.14% in validation	CGIAR dataset	863
[23]	CNN deep learning model	97.05%	from several different sites throughout the UK and Ireland	19160
[32]	Octave-UNet	96.06%	(CDTS) dataset and collected images using mobile devices	33238
[24]	EfficientNet-B4	99.35%	WheatRust21 dataset	6556
[29]	Various models of ML and DL	The high accuracy 89.5% of VGG16	wheat seed dataset	2700
[31]	(k-NN), (BPNN), and (SVM)	94.23% of K-NN	Citrus leaf disease dataset	609
[25]	EfficientNet	94.2%	WFD2020	2414
[26]	YOLOV4 EfficientNet	99.72%	Wheat Disease Detection	3672

TABLE III. OVERVIEW OF THE SOURCES OF IMAGES IN DATASET (SUMMARY)

Dataset	Classes	Images Count
Wheat Fungi Diseases (WFD2020)	Septoria, yellow rust, powdery mildew, leaf rust, stem rust, and healthy	1695
Wheat Disease Detection	healthy, brown rust (leaf rust), and yellow rust (stripe rust)	3672
CGIAR (Computer Vision for Crop Disease)	healthy, stem rust, and leaf rust	876
Wheat leaf	stripe rust, Septoria, and healthy	407

TABLE IV. OVERVIEW OF PROPOSED DATASET

Classes	Number of Images
Healthy	1807
Leaf Rust	1848
Stem rust	1796
Yellow rust	1915
powdery mildew	1947
Septoria	1912
Total	11205

filters in the convolutional layer, numerous feature maps are produced by extracting various features from the image. The starting values of the filter metrics are chosen at random, then backpropagation is used to learn the best values for the filter matrix, which may then be used to extract the most crucial characteristics from the photos. After the convolutional layers, an activation layer is added to use an activation function to introduce non-linearity to the output.

- The feature maps are created following the convolution layer. The feature maps have an excessively wide dimension. Because of this, the *pooling procedure* is used to minimize the size of feature maps, maintaining only the pertinent information and deepening feature maps to produce a highly compressed feature vector in the end. Different pooling operations exist, including average and maximum pooling. When using maximum pooling, the filter is slid over the matrix and the maximum value from the slid filter is used.

- After collecting the features from the convolution network, *fully connected* layers are built to classify the image and train the network. A 2D tensor

(a) (b) (c)

Fig. 2. Example of images with different height and width.

is created by flattening the final feature map. In multiclass situations, the softmax activation function transforms the neural network's outputs into a vector of probabilities, each probability belonging to a certain class, and deep feedforward networks for classification utilizing cross-entropy loss.

2) Transfer Learning: Transfer learning is a method where a model is created and trained on one task, then utilized as the basis for another task [40].

This method transfers the weights that a network has learned at one task to another rather than starting the learning process from scratch by first training a base model on a particular task or dataset, and then applying the knowledge of an already trained model to a different task or dataset. As a result, this technique shortens the time needed to train such models, which can be laborious.

In this study, pre-trained CNN models EfficientNetB0, ResNet152, and DenseNet161 were trained.

- *EfficientNet* is a CNN model that is an efficient computationally and achieved state of art results on the ImageNet dataset. the core idea of this model is Model scaling. Model scaling is about scaling the existing model depth-wise, width-wise, or scaling input image resolution to get better results. Model scaling is used to enhance the model's performance. The most common is depth-wise scaling. To effectively scale up the model, EfficientNet employs a method called compound coefficient. It equally scales each dimension using a set of predetermined scaling factors. The architecture makes use of a larger MobileNet-V2-like mobile inverted bottleneck convolution. EfficientNet's creators created seven models with varied dimensionalities. The baseline network of the EfficientNet family is EfficientNetB0 [41].
- Building CNN deeper by increasing the number of layers may cause a common problem called the Vanishing/Exploding gradient. The Vanishing/Exploding gradient causes the gradient to become 0 or too large a value, which causes an increase in training and test error rate. So, to solve the problem of the vanishing/exploding gradient, the *ResNet* [42] convolutional neural network model was introduced which is based on the concept called Residual Blocks that used a technique called skip connections. The skip connection takes the output from one layer and adds it to others by skipping some levels in between; regularization will skip any layers that have poor performance. The network can now reach considerably deeper thanks to this. For image recognition and classification, the model took home the ILSVRC ImagNet-2015 and MS COCO 2015 awards.
- *DenseNet* is used to solve the vanishing while going deeper but at the same time avoiding the vanishing problem by using shorter connections between the layers. ResNet and DenseNet vary in that ResNet uses summation to connect all previous feature maps, whereas DenseNet concatenates them all. Each layer in DenseNet obtains inputs from all the previous layers and passes on its output to all the layers that will come after it. DenseNet consists of Dense blocks that are composed of composition layers that contain batch normalization, RELU activation function, and 3*3 conv layer these Blocks are connected by 1×1 Conv followed by 2×2 average pooling layers that are used as the transition layers between blocks. DenseNet achieved the greatest classification performance in 2017 on ImageNet and CIFAR-10 datasets [43], [44].

3) One Fit policy cycle method: CNN hyperparameters are parameters that are used to regulate the model's behavior. It is crucial to improve the performance of the model. The *Learning Rate* is one of the hyperparameters, and it may be the most significant hyperparameter in deep learning.

How many gradients will be back-propagated will depend on the learning rate. The model slowly diverges when the learning rate is high, but it quickly converges when it is low. To find the right learning rate, the learning rate must be tweaked, which takes some time and effort.

The typical approach is to experiment with various learning rates and select the one that results in the minimum loss value, allowing the model to swiftly adapt to the situation. In study [45], this study established a new method called fit-one-cycle which is a way of tuning the learning rate.

After each mini-batch, the learning rate should be increased from a low starting point. The formulae below in Fig. 3 update it following each mini-batch:

$$Max\_lr = init\_lr * q^n$$

$$q = \left(\frac{Max\_lr}{init\_lr}\right)^{\frac{1}{n}}$$

$$lr_i = init\_lr * \left(\frac{Max\_lr}{init\_lr}\right)^{\frac{i}{n}}$$

Fig. 3. Maximum learning rate and lower learning rate.

Max_lr and Init_lr stand for the maximum learning rate and lower learning rate, respectively, where n is the number of iterations. The test range's initial value is the lower learning rate. Let q be the agent that raises the learning rate after each mini-batch. The research also suggests that for a full run, the learning rate should cycle between the lower bound and the higher bound. An iterative process where we move from a lower bound learning rate to a higher bound and then back to the lower bound is called a cycle.

To conclude, this method saves the time and effort of running multiple full cycles with different momentum values. Additionally, it yields more stable results and requires fewer epochs to train our model to completion. This study [46] confirms the improvement in validation accuracy when comparing the naive learning rate policy with the one-cycle policy. Besides that, using this strategy avoids having to conduct numerous full cycles with various momentum levels. Additionally, it produces more consistent results and takes fewer training epochs to fully train our model. The improvement in validation accuracy when contrasting the one-cycle policy with the naive learning rate policy is supported by the study [46].

V. RESULTS AND DISCUSSION

Five different models were used in the experiments presented in this section, two of which are extensions of EfficientNet (EfficientNetB0 and EfficientNetB1), two of which are extensions of DenseNet (DenseNet161 and DenseNet169), and one of which is ResNet152. These models were trained using data from five different types of fungal infections as well as healthy leaves from wheat. A training set of 8,946 photos and a validation set of 2,259 images make up this dataset, which is used to test and validate procedures.

A. Experimental Settings

This paper employed the *Fast ai framework*[47] in building learning models. It is a high-level framework over Pytorch for training machine learning models and achieving state-of-the-art performance. This framework is mostly employed for image classification, object recognition, and image segmentation. It offers faster computations than rivals and comes with data purification widgets, providing a very user-friendly workflow and making debugging easier. Additionally, Google Colab was used to conduct the trials.

B. Evaluation Criteria

Along with the receiver operating characteristic (ROC) curve and the area under the curve (AUC), the accuracy, precision, and recall/sensitivity are the performance measures chosen to assess and analyze the performance of the created model.

The performance of classification models is evaluated using a matrix called the confusion matrix. The True positive (TP), True negative (TN), False positive (FP), and False negative (FN) factors are computed for each class using the confusion matrix. The metrics for evaluation can be summed up as follows:

- The percentage of all samples that were properly identified by the classifier is used to determine the *accuracy* number.
- The true positives are divided by the total samples that were projected to be positive (TP + FP) to determine the *precision* value.
- The true positives are divided by samples that should be predicted as positive (TP, FN) to get the *recall* value.
- The *F1-score* is regarded as the harmonic average of recall and precision.
- By averaging the metrics that are obtained for each class, the *Macro-F1 (macro-averaged F1-score)* is calculated.
- A graphical depiction called the *ROC curve* (receiver operating characteristic curve) shows how well a classification model performs at every classification threshold. The True Positive Rate (TPR), which stands for the recall measure, and the False Positive Rate (FPR), are plotted on this curve at various categorization levels. AUC (Area Under the ROC Curve), a sorting-based algorithm, is used to calculate the points in a ROC curve. The probability that a model would rank a random positive instance higher than a random negative instance is shown by the AUC, which offers an overall measure of performance overall potential classification thresholds.

The following section describes the evaluation of these parameters against the learning model.

C. Accuracy and Loss Evaluation Results

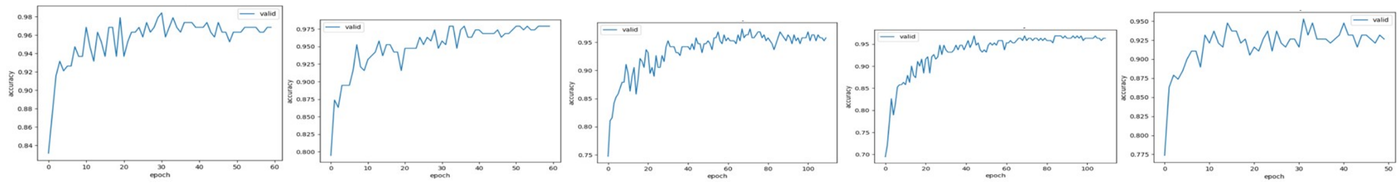
In this study, the following five models are evaluated EfficientNetB0, EfficientNetB1, DenseNet161, DenseNet169, and ResNet152 across several experiments. As a result of experiments, it is discovered that these models achieved 97.37%, 96.84%, 98.42%, 97.89%, and 95.2% classification accuracy in the validation stage and 94.15%, 93.86%, 93.76%, 93.33%, and 93.10% in the testing stage, respectively.

In conclusion, the DenseNet161 model had the best validation accuracy, at 98.42 percent, but EfficientNetB0 had the highest testing accuracy, at 94.15 percent, as opposed to DenseNet161, which had a testing accuracy of 93.76 percent. During the training and validation operations, the accuracy and loss plot curves are built as a function of epochs.

Fig. 4 shows the validation accuracy for each model.

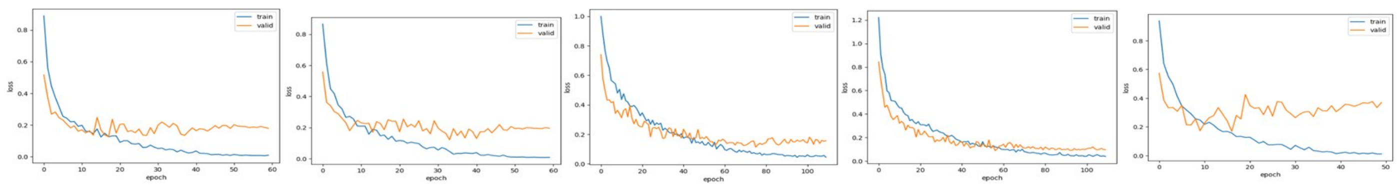
Fig. 5 shows the loss through the training and validation process.

Table V demonstrates the validation accuracy, precision, and Recall values measured during the validation process and testing. These values are calculated over all classes for each model and the number of epochs.



(a) (b) (c) (d) (e)

Fig. 4. Validation accuracy.



(a) (b) (c) (d) (e)

Fig. 5. Validation loss.

TABLE V. ACCURACY, PRECISION, AND RECALL MEASURES OVER ALL CLASSES FOR EACH MODEL (SUMMARY)

Classifier	**Accuarcy**	**Recall**	**Precision**
EfficientNetB0	**97.37%**	**97.37%**	**97.53%**
EfficientNetB1	**96.84%**	**96.84%**	**97.11%**
DenseNet161	**98.42%**	**98.42%**	**98.48%**
DenseNet169	**97.89%**	**97.89%**	**97.97%**
ResNet152	**95.26%**	**95.26%**	**95.22%**

D. Confusion Matrix Parmatters Evaluation Results

Indeed, the main role of the confusion matrix parameters is to show how the model detects instances correctly and the relatively incorrect classifications of the instances. This matrix identifies confusion between classes of datasets.

The relevant performance measures that have been calculated based on the confusion matrix are precision, Recall, F1-Score, and Macro average of each measure. These values are calculated for each class to determine how well the classifier can identify different classes.

True positive, True Negative, False positive, and False Negative are regarded as a one-vs-all problem in multi-class classification problems. As a result, the positive class is a certain class, while the negative class is every other class.

The evaluation of the confusion matrix parameters for each model is shown in Fig. 6. The graphic demonstrates that the EfficientNetB0 model correctly identified just two healthy samples as powdery mildew, but it incorrectly identified 12 samples as stripe rust, which is higher than other models did.

Fig. 7 presents the precision, recall, and f1-scores for models used in the experiment.

Fig. 8 describes the ROC curves for these models are shown. This demonstrates that both EfficientNetB0 and DenseNet161 had similar AUC for all the given classes except the healthy class EfficientNetB0 had the highest AUC of 100%. All models had the same AUC of 95% of the leaf rust class except EfficientNetB1 had the lowest AUC of 94% but EfficientNetB1 had the highest AUC of 97% of the stripe rust class. DensNet169 and DensNet161 had similar for all the given classes except powdery mildew and septoria, DensNet161 AUC is higher. Also, ResNet152 had a similar AUC to DensNet161 except for stripe rust and powdery mildew, DensNet161's AUC is higher.

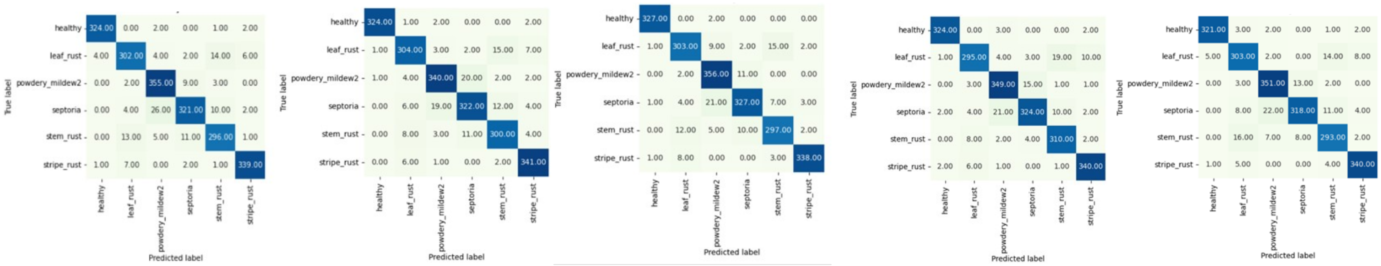
In this experiment, EfficientNetB0 is compared with the work of study [26], [33] by applying EfficientNetB0 on (leaf rust, stem rust, and healthy) classes from the proposed dataset in study [33] and on (leaf rust, stripe rust, and healthy) classes as [26]. The results concluded that EfficientNetB0 is a high-performer model with high accuracy.

E. Impact of applying fit-one-cycle policy Results

This experiment studied the impact of not applying the fit-one-cycle policy on accuracy scores. Table VI presents the accuracy results without using a fit-one-cycle policy. By comparing the results highlighted in Table V and the results presented in Table VI, the overall accuracy is decreased by 4% without applying a one-fit-cycle policy.

F. Summary

This study compared different CNN learning models to classify five wheat fungal diseases based on RGB images. This work uses three classes of diseases: stripe rust, leaf rust, and healthy, the dataset captured from [1]. Additionally, the CGIAR and wheat leaf datasets were captured from this study and used



(a) (b) (c) (d) (e)

Fig. 6. Confusion matrix of the models.

(a)

class	precision	Recall/Sensitivity	F1-Score
Healthy	0.98	0.98	0.98
Leaf rust	0.92	0.91	0.92
Powdery mildew	0.91	0.96	0.93
septoria	0.93	0.88	0.91
Stem rust	0.91	0.91	0.91
Stripe rust	0.97	0.97	0.97
average	0.94	0.94	0.94

(b)

class	precision	Recall/Sensitivity	F1-Score
Healthy	0.99	0.98	0.99
Leaf rust	0.92	0.92	0.92
Powdery mildew	0.92	0.92	0.92
septoria	0.91	0.89	0.90
Stem rust	0.91	0.92	0.91
Stripe rust	0.95	0.97	0.96
average	0.93	0.93	0.93

(c)

class	precision	Recall/Sensitivity	F1-Score
Healthy	0.99	0.99	0.99
Leaf rust	0.92	0.91	0.92
Powdery mildew	0.91	0.96	0.93
septoria	0.93	0.90	0.92
Stem rust	0.92	0.91	0.92
Stripe rust	0.98	0.97	0.97
average	0.94	0.94	0.94

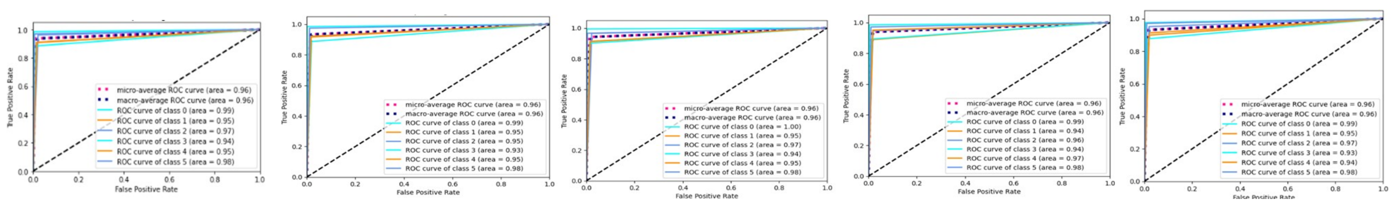
(d)

class	precision	Recall/Sensitivity	F1-Score
Healthy	0.98	0.98	0.98
Leaf rust	0.93	0.89	0.91
Powdery mildew	0.92	0.95	0.93
septoria	0.94	0.89	0.91
Stem rust	0.91	0.95	0.93
Stripe rust	0.95	0.97	0.96
average	0.94	0.94	0.94

(e)

class	precision	Recall/Sensitivity	F1-Score
Healthy	0.98	0.98	0.98
Leaf rust	0.90	0.91	0.90
Powdery mildew	0.91	0.95	0.93
septoria	0.94	0.88	0.91
Stem rust	0.90	0.90	0.90
Stripe rust	0.96	0.97	0.96
average	0.93	0.93	0.93

Fig. 7. Recall, precision, and F1-score measured.



(a) (b) (c) (d) (e)

Fig. 8. ROC curve of the models.

TABLE VI. ACCURACY MEASURES OVER ALL CLASSES FOR EACH MODEL WITHOUT FIT-ONE-CYCLE

Classifier	Accuarcy
EfficientNetB0	**93.68%**
EfficientNetB1	**91.58%**
DenseNet161	**89.47%**
DenseNet169	**92.63%**
ResNet152	**95.26%**

in this work[2]. The CGIAR contains three classes which are leaf rust, stem rust, and healthy, while the wheat leaf dataset contains (stripe rust, Septoria, and healthy). Moreover, Wheat Fungi Diseases (WFD2020) presented in [3] are used also in this work which contains classes of five types of wheat fungi diseases and healthy wheat leaf, but the number of images is small containing 2414 images through all classes of the dataset. Finally, in this work, the experimental dataset was collected from all the mentioned datasets and contains six classes of five types of wheat leaf fungal diseases and healthy ones.

In this work, there are five CNN pre-trained models constructed based on datasets collected from various sources. These models include: EfficientNetB0, EfficientNetB1, DenseNet161, DenseNet169, and ResNet152. This work employs the one-fit-cycle method to enhance the proposed models' accuracy and reduce the time needed to train the models. Overall, the EfficientNetB0 and EfficientNetB1 achieved a high testing accuracy, however, DenseNet161 achieved a high validation accuracy.

VI. CONCLUSION

Eventually, to sum up, this paper discusses the different wheat diseases that can impact wheat crop growth and therefore will harm food security all over the world. Mainly, the paper describes how wheat disease can be detected and recognized efficiently. For this purpose, this paper employs convolutional neural network models as deep learning models. Moreover, this work compares different learning models such as ResNet, DensNet, and EfficientNet to select the best model that achieves the highest accuracy. Additionally, this study used a one-fit-policy method that automates the selection of the best value of hyperparameters, this led to a great performance achievement. Experimental evaluation for all models performed on different real datasets collected from various. Experimental results proved that the EfficientNet learning model is an effective model more than ResNet and DensNet and the justification behind that is that EfficientNet needs fewer hyperparameters to train and learn than ResNet and DensNet.

REFERENCES

[1] A. Rady, "Inheritance as a base for future improvement of new egyptian bread wheat cultivars," *Journal of Plant Production*, vol. 13, no. 12, pp. Pages 847–850, 2022.

[2] J. A. L. Head and A. Gates, "Wheat as food, wheat as industrial substance; comparative geographies of transformation and mobility," *Geoforum*, vol. 41, no. 2, pp. Pages 236–246, 2010.

[3] Ákos Mesterházy, Judit Oláh and J. Popp, "Losses in the grain supply chain: Causes and solutions," *Sustainability*, pp. Pages 236–246, 2020.

[4] E. C. T. L. Y. S. Njuki and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, pp. Pages 272–279, 2019.

[5] M. V. A. Kumar and V. Sharma, "Deep learning solutions for pest identification in agriculture," *Object Detection with Deep Learning Models*, pp. Pages 199–214, 2022.

[6] P. MahiDar and D. Ghai, "Bio-inspired optimization algorithms for machine learning in agriculture applications," *Smart Agriculture*, pp. Pages 53–60, 2021.

[7] H. S. A. R. E.Sheriff and F. Mahieddine, "Convolution neural network in precision agriculture for plant image recognition and classification," *2017 Seventh International Conference on Innovative Computing Technology (INTECH)*, 2017.

[8] R. K. M. A. K. M. A. A. N. Dhingra and N. Bhati, "Machine learning-based agriculture," *Application of Machine Learning in Agriculture*, pp. Pages 3–27, 2022.

[9] M. S. N. I. A. C. N. D. K. M. Ramachandran and A. Kumar, "Plant disease detection using hybrid deep learning architecture in smart agriculture applications," *Artificial Intelligence and Smart Agriculture Applications*, pp. Pages 213–232, 2022.

[10] Ümit Atila, Murat Uçar, Kemal Akyol and E. Uçar, "Plant leaf disease classification using efficientnet deep learning model," *Ecological Informatics*, vol. 61, pp. Pages 236–246, 2021.

[11] M. D. R. M. S. B. L. W. J. Whish and S. Savary, "Modelling the impacts of pests and diseases on agricultural system," *Agricultural Systems*, vol. 61, pp. Pages 213–224, 2017.

[12] S. S. C. A. K. U. P. Singh and S. Jain, "Bacterial foraging optimization based radial basis function neural network (brbfnn) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology," *IEEE Access*, pp. Pages 8852–8863, 2018.

[13] M. F. K. E. Hammond-Kosack and P. S. Solomon, "A review of wheat diseases-a field perspective," *Molecular Plant Pathology*, vol. 19, no. 9, pp. Pages 1523–1536, 2017.

[14] H. W. Z. K. X. L. Y. L. Y. L. S. Wang and D. Liu, "Identification of wheat leaf rust resistance genes in chinese wheat cultivars and the improved germplasms," *Plant Disease*, vol. 104, no. 10, pp. Pages 2669–2680, 2020.

[15] M. Albahar, "A survey on deep learning and its impact on agriculture: Challenges and opportunities," *Agriculture*, vol. 13, no. 3, pp. Pages 2669–2680, 2023.

[16] Ümit Atila, Murat Uçar, Kemal Akyol and E. Uçar, "Plant leaf disease classification using efficientnet deep learning model," *Ecological Informatics*, vol. 13, no. 3, pp. Pages 2669–2680, 2021.

[17] D. Kumar and V. Kukreja, "N-cnn based transfer learning method for classification of powdery mildew wheat disease," *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 2021.

[18] L. G. C. M. S. A. Singh and P. K. Singh, "Leaf and spike wheat disease detection & classification using an improved deep convolutional architecture," *Informatics in Medicine Unlocked*, vol. 25, 2021.

[19] V. Kukreja and D. Kumar, "Automatic classification of wheat rust diseases using deep convolutional neural networks," *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2021.

[20] Z. J. Z. D. W. Jiang and Y. Yang, "Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning," *Computers and Electronics in Agriculture*, vol. 186, 2021.

[21] N. Kumari and B. Saini, "Fully automatic wheat disease detection system by using different cnn models," *Advances in Intelligent Systems and Computing*, pp. Pages 351–365, 2023.

[22] S. S. H. Singh and S. Jindal, "Rust disease classification using deep learning based algorithm: The case of wheat," *Sustainable Development*, 2022.

[23] M. L. M. H. R. J. Morris and J. K. Brown, "Classification of wheat diseases using deep learning networks with field and glasshouse images," *Plant Pathology*, vol. 72, no. 3, pp. Pages 536–547, 2023.

[24] S. N. R. J. S. M. A. A. M. A. H. A. Dheeraj and V. K. Singh, "Deep transfer learning model for disease identification in wheat crop," *Ecological Informatics*, vol. 75, 2023.

[25] M. G. E. S. E. G. E. O. N. Bechtold and D. Afonnikov, "Image-based wheat fungi diseases identification by deep learning," *Plants*, 2021.

[26] B. S. Y. Alborzi and E. Najafi, "Automated wheat disease detection using a ros-based autonomous guided uav," *ArXiv Preprint ArXiv:2206.15042*, 2022.

[27] A. K. F. O. M. A. T. Henriksson and A. Chawade, "Predicting yellow rust in wheat breeding trials by proximal phenotyping and machine learning," *Plant Methods*, vol. 18, no. 1, 2022.

[28] W. B. J. Z. G. H. D. Z. L. Huang and D. Liang, "Identification of wheat leaf diseases and their severity based on elliptical-maximum margin criterion metric learning," *Sustainable Computing: Informatics and Systems*, vol. 30, 2021.

[29] R. S. J. M. C. N. Jhanjhi and S. N. Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection," *Microprocessors and Microsystems*, vol. 80, 2021.

[30] H. K. I. U. H. M. M. S. U. Khan and M. Y. Lee, "Automated wheat diseases classification framework using advanced machine learning technique," *Agriculture*, vol. 12, no. 8, 2022.

[31] S. C. R. Y. N. N. Malvade and N. Gadagin, "Automated classification of wheat varieties using soft computing techniques," *Communications in Computer and Information Science*, vol. 12, no. 8, pp. Pages 89–98, 2022.

[32] Y. L. Q. W. L. W. J. J. L. Y. L. Y. T. X. M. H. L. Q. Hu and Q. Yao, "Semantic segmentation of wheat stripe rust images using deep learning," *Agronomy*, vol. 12, no. 12, pp. Pages 89–98, 2022.

[33] H. R. B. R. M. S. I. U. S. I. U. H. S. M. H. ZAIDI and M. HAFEEZ, "Assessing the impact of segmentation on wheat stripe rust disease classification using computer vision and deep learning," *IEEE Access*, vol. 9, pp. Pages 164 986–165 004, 2021.

[34] W. C. Z. W. Y. S. M. Li and X. Yang, "Residual-capsule networks with threshold convolution for segmentation of wheat plantation rows in uav images." *Multimedia Tools and Applications*, vol. 80, pp. Pages 32 131–32 147, 2021.

[35] M. A. G. S. S. E. I. G. O. E. A. O. N. P. Bechtold and andDmitry A. Afonnikov, "Image-based wheat fungi diseases identification by deep learning." *Plants*, vol. 10, no. 8, pp. Pages 32 131–32 147, 2021.

[36] B. S. Y. Alborzi and E. Najafi, "Automated wheat disease detection using a ros-based autonomous guided uav," *rXiv Preprint ArXiv:2206.15042.*, 2022.

[37] S. S. H. Singh and S. Jindal, "Rust disease classification using deep learning-based algorithm: The case of wheat," *Sustainable Development*, 2022.

[38] cgiar Dataset, "Kaggle," https://www.kaggle.com/shadabhussain/cgiar-computer-vision-for-crop-disease, 2023, [Online; accessed Augest 2023].

[39] wheat-leaf dataset, "Kaggle," https://www.kaggle.com/datasets/olyadgetch/, 2023, [Online; accessed Augest 2023].

[40] M. A. G. S. S. E. I. G. O. E. A. O. N. P. Bechtold and andDmitry A. Afonnikov, "Transfer learning," *Handbook of Research on Machine Learning Applications and Trends*, pp. Pages 242–264, 2010.

[41] Y. H. K. C. Y. Song and Z. Chen, "A multi-scaling reinforcement learning trading system based on multi-scaling convolutional neural networks," *Mathematics*, vol. 11, no. 11, 2023.

[42] S. T. D. Almeida and K. Lyman, "Resnet in resnet: Generalizing residual architectures," *Mathematics*, 2016.

[43] F. I. M. M. S. K. R. G. T. Darrell and K. Keutzer, "Densenet: Implementing efficient convnet descriptor pyramids," *ArXiv Preprint ArXiv:1404.1869.*, 2014.

[44] C. Z. P. B. . D. M. A. . S. L. J. K. . F. R. . J.-C. Bazin and I. S. Kweon, "Resnet or densenet? introducing dense shortcuts to resnet." *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. Pages 3550–3559, 2021.

[45] L. N. Smith, "Cyclical learning rates for training neural networks." *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. Pages 464–472, 2017.

[46] L. N. Smith and N. Topin, "Super-convergence: Very fast training of neural networks using large learning rates," *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, 2019.

[47] J. Howard and S. Gugger, "Fastai: A layered api for deep learning," *Information*, vol. 11, no. 2, 2020.