IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 February 2023, accepted 28 February 2023, date of publication 8 March 2023, date of current version 14 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3254430

RESEARCH ARTICLE

Embedded AI for Wheat Yellow Rust Infection Type Classification

UFERAH SHAFI[1], RAFIA MUMTAZ[1], (Senior Member, IEEE), MUHAMMAD DEEDAHWAR MAZHAR QURESHI[1,3], ZAHID MAHMOOD[2], SIKANDER KHAN TANVEER[2], IHSAN UL HAQ[1], AND SYED MOHAMMAD HASSAN ZAIDI[1]

[1]School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan
[2]Crop Sciences Institute, National Agricultural Research Centre, Islamabad 44000, Pakistan
[3]Graduate Business School, Technological University Dublin, Dublin, D02 HW71 Ireland

Corresponding author: Rafia Mumtaz (rafia.mumtaz@seecs.edu.pk)

This research is funded by National Center for Artificial Intelligence (NCAI), National University of Sciences and Technology (NUST), Islamabad, Pakistan.

ABSTRACT Wheat is the most important and dominating crop in Pakistan in terms of production and acreage, which is grown on 37% of the cultivated area, accounting for 70% of the total production. However, wheat yield is highly affected by stripe rust, which is considered the most devastating fungal disease, causing 5.5 million tonnes of loss per year globally. In order to minimize this loss, the accurate and timely detection of rust disease is crucial instead of manual inspection. Towards this end, we propose a system to detect wheat rust disease and classify its infection types into four classes, including healthy, resistant, moderate (moderately resistant to moderately susceptible), and susceptible. The wheat rust dataset is collected indigenously from the National Agricultural Research Centre, Islamabad. A pre-trained U^2 Net model is used to remove the background and extract the leaf containing the rust disease. Subsequently, two deep learning classifiers, including the Xception model and ResNet-50 are applied to classify the stripe rust severity levels, where the ResNet-50 model outperformed with the highest accuracy of 96%. This research presents a comparison between two state-of-the-art deep learning classifiers in terms of accuracy, memory utilization, and prediction time, which will assist the research community in selecting the most appropriate model for plant disease detection. Moreover, to assess the external validity, the performance of these classifiers is compared with the existing technique using a publicly available dataset, which confirms the validity of the results. Additionally, an intelligent edge computing rust detection device has been developed, where the trained ResNet-50 model is deployed, which facilitates the farmers to monitor the rust attack. The proposed research is aimed to assist the agricultural community to employ preventive measures in a site-specific manner based on the accurate diagnosis of rust disease & its severity, which is intended to improve the quality of the wheat as well as production.

INDEX TERMS Deep learning, classification, wheat stripe rust disease, segmentation, rust infection types, edge device.

I. INTRODUCTION

Wheat is the prime cereal crop that holds paramount importance worldwide due to its immense contribution to the human diet. It is the third most harvested crop in the world and contains carbohydrates, dietary fiber, protein, and vitamins. Its average annual production is around 680 million tonnes per year globally [1]. In Pakistan, wheat is considered the most significant staple crop, accounting for 37.1% of the cultivated area and 65% of the food grain acreage. It is cultivated by 80% farmers and covers almost 9 million hectares of agricultural land. Its favorable temperature is 21° C to 24° C, where it requires a moist and cool temperature in the early stages and a warm temperature in the mature stages with substantial sunshine.

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-Keung Fung.

Several diseases can affect the wheat crop, which results in a serious loss of wheat production and damage to the quality of the yield. Among these diseases, rust is the most damaging disease that can lead to 20% to 50% yield loss if it is not diagnosed and controlled at an early stage. It has three main types, including leaf, stripe, and stem rust. The rust severity is assessed as the percentage infected leaf region on visual basis where leaf samples are collected separately according to their infection types and % severity for further image analysis as discussed in [2]. However, the infection type is computed from its severity, which is normally mapped into ten levels (0–9) as discussed in [3]. Infection type 0 indicates that there are no rust symptoms on the leaf, infection type 1 shows that there are necrotic flecks with no sporulation, infection type 2 indicates that there are stripes or chlorotic blotches with no sporulation, and infection type 3 contains stripes with trace sporulation. However, there are chlorotic blotches with light sporulation in infection type 4, intermediate sporulation in infection type 5, and high sporulation in infection type 6. In infection type 7, there are stripes with abundant sporulation, and infection type 8 holds chlorotic blotches with sufficient sporulation. In infection type 9, there is large sporulation without chlorosis or necrosis [4].

The wheat yellow rust spreads rapidly, which makes it quite challenging to contain the disease. Therefore, early and precise detection of this disease is crucial to take remedial actions to minimize economic loss. For this purpose, advanced techniques and tools are used to automatically diagnose the rust attack. Towards this end, artificial intelligence, the Internet of Things(IoT), edge computing, machine, and deep learning are playing an integral role in crop disease detection [5]. For this purpose, heterogeneous data is captured from various sources such as soil parameters, plant canopy, climate data, IoT data, and vegetation indices extracted from the remote sensing images [6]. The famous machine learning techniques used to process agricultural data are Support Vector Machine (SVM), Decision Tree, Random Forest, k-Nearest Neighbour (kNN), Artificial Neural Network (ANN), Principle Component Analysis (PCA), Naive Bayes, Logistic Regression, and k-mean Clustering [7], [8], [9].

Wheat rust disease surveillance and diagnosis typically involve imagery data that provides deep insights into the particular disease. These images are captured using digital cameras or some imaging systems, which contain noise, impurities, and unwanted distortions. Thus, the preliminary step after data acquisition is to preprocess the data to make it suitable for any machine or deep learning technique. Data augmentation, contrast enhancement, noise removal, brightness adjustment, and image resizing are some common preprocessing steps [10]. After image preprocessing, the region of interest is extracted using several segmentation techniques such as GrabCut, WaterShet, Fuzzy based methods, Thresholding, Region growing, etc. as discussed in [11] and [12]. The next step is to use an appropriate machine/deep learning approach to identify crop disease and its severity levels. In this context, deep learning techniques, along with transfer learning, are gaining more success in imagery data for crop disease diagnosis. The most famous deep learning architectures are AlexNet, GoogleNet, ResNet18, ResNet-50, Deep CNN, and VGG16 as discussed in [7].

Most of the existing research in crop disease detection focuses on the detection and classification of different types of diseases, where publicly available datasets are utilized. The literature reveals that there is very little research to deeply analyze a single disease and classify it into its severity levels, which is difficult due to the similarity in single disease patterns. Moreover, small studies have been found in which a real-time practical solution is presented by developing hardware for plant disease detection. Toward such end, we propose a system to diagnose the wheat yellow rust and its infection types where different deep learning techniques are used (Figure 1). The dataset is collected using mobile cameras, which is further preprocessed and segmented using the U^2 net model. Subsequently, two deep learning models are applied including ResNet-50, and the Xception models which are used to classify wheat yellow rust severity levels. These models are selected due to their high capability to perform real-life complex classification tasks. Moreover, these models are among the top five models that achieved the highest accuracy on the benchmark dataset (ImageNet). In order to validate the performance of these models, a publically available wheat yellow rust dataset is utilized where models have been tested out on a standard, and unseen dataset. Further, the trained model with the best performance is deployed on the edge device for the purpose of on-device inferences that help to detect rust disease and its infection types in real-time by tanking images in the wheat fields. The main contributions of the current research are given below:

- Development of a complete solution for wheat stripe rust detection by addressing the research gap where a single disease is considered and thoroughly investigated. For this purpose, the collected dataset of wheat stripe rust and the publicly available dataset are used. Two deep learning models are evaluated, including ResNet-50 and the Xception model, which will assist agricultural scientists in selecting the most suitable model for rust severity level classification.
- Data collection: The wheat stripe rust dataset has been collected locally from the NARC fields throughout the disease life cycle. The collected dataset will help researchers to continue their investigation in wheat rust detection and mature their research.
- Development of an edge device and on-device inference by deploying the most appropriate deep learning model to identify the wheat rust infection type. This edge device is portable and quite handy to perform ground surveys which will assist farmers, agronomists, and researchers to detect wheat rust and its severity levels.

The rest of the paper is organized as follows: Section II presents the related work; Section III discusses the materials and methods; Section IV presents the results and

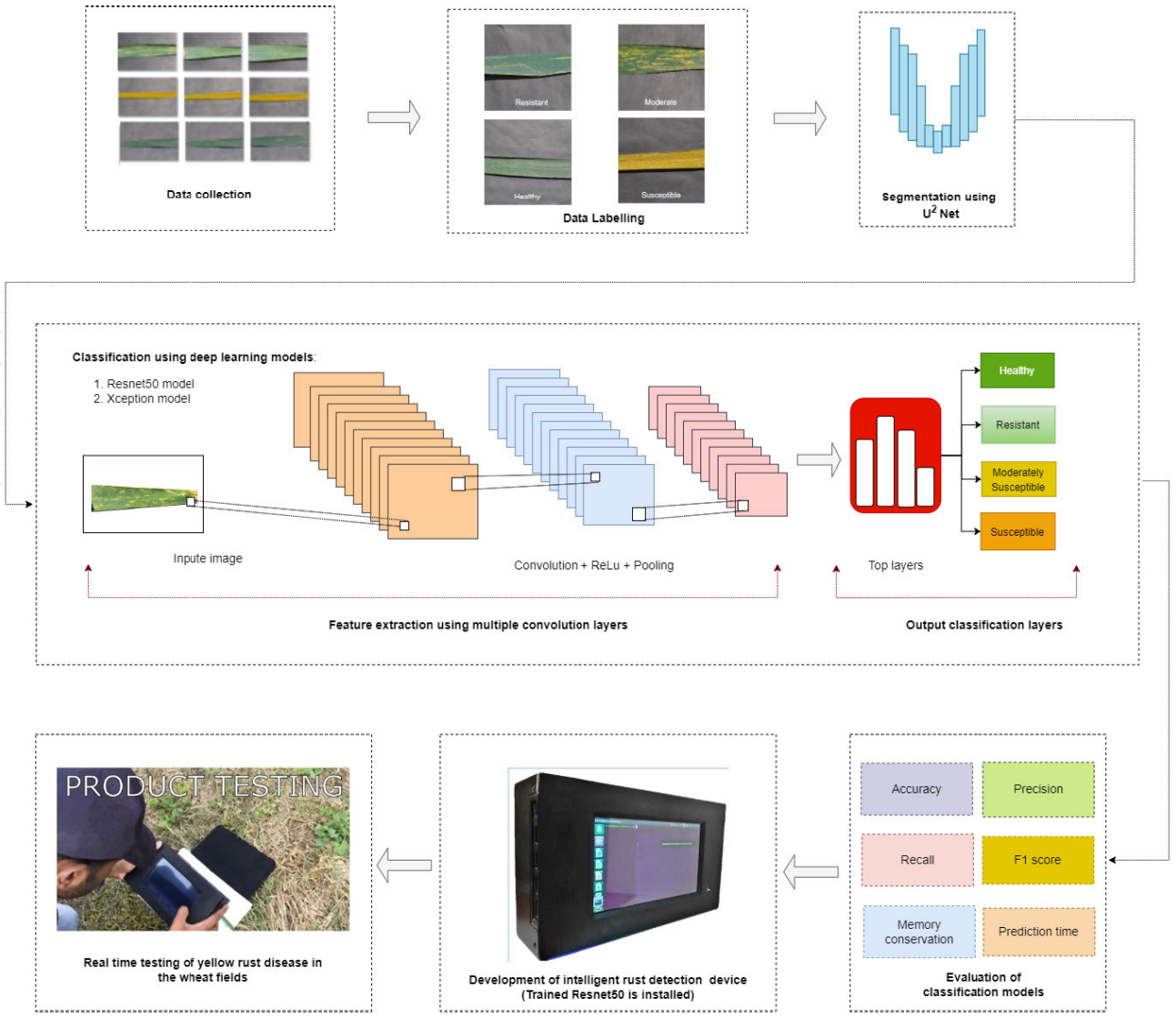

FIGURE 1. Proposed framework for wheat stripe rust infection type classification.

discussion, and Section V discusses the conclusion and future work.

II. RELATED WORK

There have been several lately studies in the plant disease detection domain in which machine and deep learning techniques are used. The performance of any machine or deep learning model mainly depends on the dataset. In [13], wheat disease classification is performed using texture and color features, where images are collected by RGB camera along with the public dataset. The normal wheat leaves are differentiated from the rust leaves and nitrogen-deficient leaves using an ensemble learning-based technique. The results indicate that an accuracy of 96% is observed for nitrogen deficient vs normal, and 98.25% is observed for leaf rust vs normal. In [14], the impact of the variety and size of the dataset on the effectiveness of the deep learning model is discussed. The dataset containing different plant diseases is used, where each plant disease dataset varies in size and number of classes. GoogleNet is applied to these datasets to classify plant diseases, where the highest accuracy of 97% is recorded in the case of sugarcane disease detection. In [15], wheat yellow rust is detected using a CNN, where, a dataset comprising 2000 images is collected using a digital camera. The proposed deep learning model obtained the highest testing accuracy of 97.3% on classifying the images into two classes i.e. healthy leaf and yellow rust. Similarly, CNN is used to diagnose crop disease which achieved an accuracy of more than 98% as discussed in [16].

In [17], wheat spike and leaf disease detection is performed in which the publicly available dataset (LWDCD2020) is used along with some manually collected images from the wheat fields. The dataset contains 10 classes representing different types of wheat diseases. A deep learning CNN model is proposed which has 21 convolution layers followed by 7 pooling layers with 3 fully connected dense layers. The performance

of the proposed deep learning model is compared with the famous deep learning architectures such as VGG16 and ResNet-50. The proposed model outperformed with the highest testing accuracy of 97.88%, where VGG16 and ResNet-50 obtained a testing accuracy of 90.87, and 81.96% respectively. Similarly, in [18], ResNet-50 and VGG16 are used to detect wheat leaf and stem rust. These deep learning models are applied to different sizes of datasets to assess the effect of dataset size on model accuracy. The results indicate that VGG16 obtained the highest validation accuracy of 98.33% on the large dataset containing 6000 images.

In [19], wheat rust disease classification is performed using a deep CNN model in which 1486 wheat disease images are obtained from the publicly available dataset (CGIAR), and 514 images of wheat stripe rust are obtained from the secondary source. The dataset contains 2000 wheat rust images which are labeled into four classes such as healthy plant, leaf rust, stripe rust, and stem rust. The proposed deep CNN model is trained with the batch size = 64, dropout rate = 50%, learning momentum = 0.9, and optimizer = Adaptive Moment Estimation (adam); which achieved the highest accuracy of 97.16%. In [20], AlexNet is used to classify wheat diseases into stem rust, yellow rust, powdery mildew, and normal (healthy). In order to train the model, 7062 wheat images are used for training, and 1766 images are used to test the trained model. The highest accuracy of 84.54% is observed, where the batch size is set to 16, and the learning rate is set to 0.0001 with the Rectified Linear Unit (ReLU) activation function.

In [21], wheat stripe rust is detected using Resnet18, where the dataset is collected using a DSLR camera. In order to collect the images, the camera is fixed on a boom which is mounted on a carrier so that images are acquired with the specific setting. The distance of the camera from the ground is set to 2 meters. Additionally, drone imagery is acquired with the same imaging sensor mounted on the drone, which is further used to compute the triangular greenness index. The computed index has a good correlation with the photosynthetic activity which helps to assess the rust attack. The Resnet18 model obtained the highest accuracy of 97% at the patch level, where a classification accuracy of 77% is observed at the image level.

In [22], a deep CNN model is proposed to classify plant diseases in which a publicly available dataset (PlanVillage) is used. The dataset contains 70295 images of different plant diseases which have 38 different classes. The proposed model obtained a testing accuracy of 98.3%, where the batch size is set to 32 with 25 epochs. Similarly, another deep CNN model is presented in [23] to classify 5 different apple diseases. The proposed model is trained on the dataset of 26,377 images which obtained a mean average accuracy of 78.80%.

In [24], high-resolution hyperspectral imagery is used to automatically detect the yellow rust disease. A deep CNN model is proposed in which several Inception Resnet layers are introduced. The performance of the proposed model is compared with the Random Forest classifier. The results show that the proposed model outperformed with an accuracy of 85%, whereas the Random Forest obtained an accuracy of 77%.

It is observed from the literature that most of the existing research focuses on the detection of different types of plant diseases, where each disease has some distinct patterns which make a classification task much easier. Consequently, high performance is achieved using machine and deep learning techniques. Additionally, these techniques are mostly applied to publicly available datasets, where no practical hardware solution or disease detection device is proposed. However, it is quite challenging to consider a single disease and diagnose its severity levels due to the similarity in the disease patterns. Moreover, indigenous data collection and the development of a real-time solution for plant disease detection require substantial effort. Toward such end, we developed a system to detect a single wheat rust disease and map it into four severity levels including healthy, resistant, moderate (moderately resistant to moderately susceptible), and susceptible. For this purpose, multiple ground surveys are performed to collect the dataset locally through the entire life cycle of the rust disease. Subsequently, two deep learning classification models are used including ResNet-50 and Xception models, where these models are tested on a publically available dataset to assess the external validity. Additionally, an embedded edge device is developed, where the most suitable model is deployed which will help the farmers to diagnose the wheat rust attack and its severity levels.

III. MATERIALS AND METHODS

A. DATA ACQUISITION

1) STUDY AREA

The study area for the proposed research is the National Agriculture Research Center (NARC), Islamabad, which is situated at 33.67°N latitude and 73.13°E longitude. Wheat rust data is collected from several wheat experimental fields listed in Table 1 and shown in Figure 2. There are several small plots in each experimental field with different wheat varieties, germplasm from Genebank, and landraces (Figure 3).

The wheat disease images are collected using mobile phones, where the wheat leaves are placed on a uniform background. Initially, the collected dataset is comprised of 2000 images, which are further reduced due to cluttered backgrounds, illumination variation, and compromised camera quality. The final dataset contains 1640 wheat images, which are labeled with the help of experts from NARC. In this research, the dataset is labeled into four infection types including (i) healthy, (ii) resistant, (iii) moderate, and (iv) susceptible. There are 645 images of healthy, 348 of resistant, 238 of moderate, and 409 of susceptible class (Table 2).

The collected dataset used for this research work is available on Github[1]

[1]Github Repository: https://github.com/Uferah/Wheat-Stripe-Rust-Dataset

TABLE 1. Experimental fields of wheat yellow rust disease.

Sr no	Experimental Field	Location	Detail
1	Wheat Programme, NARC experimental Fields	33°40'28.5"N 73°07'44.2"E	Containing different wheat trials and experiments
2	Plant Genetic Resources Institute, NARC Experimental fields	33°40'26.7"N 73°07'36.0"E	Containing different wheat germplasm from genebank
3	Wheat Experimental Fields of NIGAB, NARC	33°40'25.3"N 73°07'28.0"E	Containing different wheat varieties and landraces

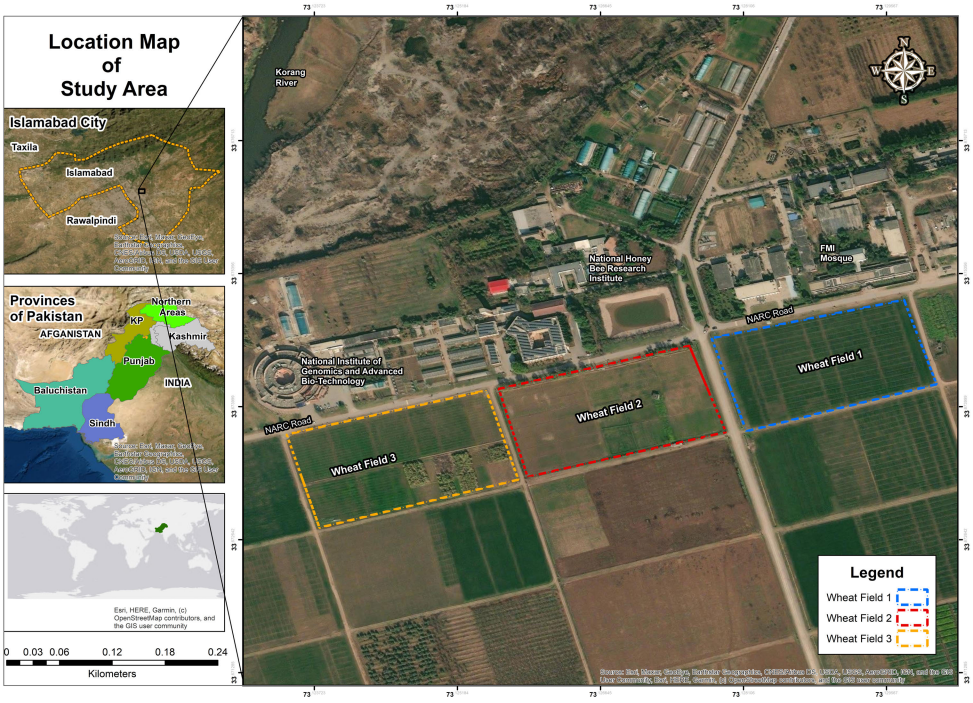

FIGURE 2. Study area location map.

FIGURE 3. Aerial view of wheat experimental fields containing small plots with different wheat varieties, germplasm from genebank and landraces.

The collected data is insufficient to train a deep learning model and achieve a good performance. To boost the model performance, the dataset is enhanced by utilizing a publically

TABLE 2. Datasets used for wheat yellow rust detection.

Dataset	Healthy	Resistant	Moderate	Susceptible
Collected dataset	645	348	238	409
Dataset [26])	2500	2500	2500	2500

available wheat yellow rust dataset [25] which contains six infection types including healthy, resistant, moderately resistant, moderately resistant moderately susceptible, moderately susceptible, and susceptible. There are 2500 images of each infection type and a total of 15,000 images. However, images of four infection types are merged with our collected dataset, including healthy, resistant, moderate, and susceptible. Table 2 shows the class-wise distribution of images in both datasets used for wheat rust infection type classification.

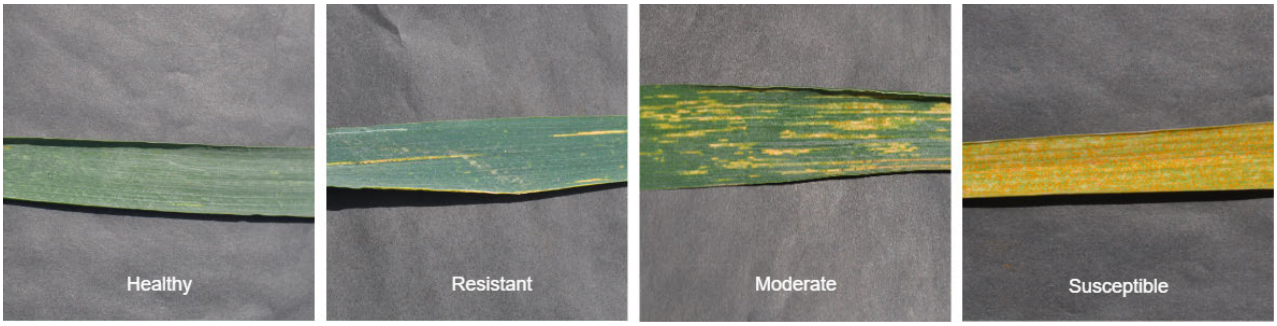

FIGURE 4. Wheat yellow rust disease infection types.

B. SEGMENTATION USING U^2- NET

Segmentation refers to partitioning an image into multiple groups or segments by assigning each pixel to a specific category. Generally, it is used to localize different objects in the image, which further helps to perform advanced image analysis. Several image segmentation techniques have been proposed in the literature, including region growing, thresholding, K-Means clustering, watershed method, graph cuts, and model-based techniques as discussed in [27]. However, deep learning models developed in recent years have shown remarkable results in image segmentation tasks. Among these models, U-Net is a widely used segmentation technique to extract salient objects from the image. Its architecture consists of two sections (i) the contraction path, which captures the contextual information; and (ii) the expansion path, which learns precise localization [28].

Several variants of the U-Net model have been introduced, where the U^2-Net model is used in this research work. It is a nested U-Net architecture [29] rather than a cascaded one, which is shown by the exponential use of '2' in the title. U^2-Net is primarily used for Salient Object Detection (SOD) which focuses on separating salient object(s) in the image from the background by increasing the foreground intensity and decreasing the background intensity. The architecture of the network emphasizes going deeper without adding a massive overhead on memory and computation costs. This is accomplished by nesting the normal U-Net structure on the bottom level with a ReSidual U-block (RSU) which allows it to extract intra-stage multi-scale features from the image without necessarily degrading the feature map resolution. At the top level, there is a U-Net architecture, where each stage is filled by an RSU block.

The presence of different backgrounds across the images can be a source of noisiness in the dataset. This makes classification inferences difficult. To deal with this challenge, this paper uses the pre-trained U^2-Net architecture from the Ailia SDK (Ailia 1.2.3) provided by the original paper [29]. This version is optimized in terms of its weights for segmenting a foreground object from the background as discussed in [12]. The results of segmentation are shown in Figure 5.

C. CLASSIFICATION

Wheat yellow rust disease detection and its infection-type classification is performed on an Intel(R) Xeon(R) CPU with a clock speed of 2.20GHz. The GPU used is an Nvidia Tesla K80 GPU with 4GB memory. Two pre-trained deep learning models, including ResNet-50, and Xception are applied to the dataset which is explained below.

1) ResNet-50 MODEL

In order to tackle more complicated tasks, deep learning architectures are becoming deeper and more complex, which has also boosted classification and recognition task performance and strengthened their robustness. However, if we keep on adding extra layers to the network, it becomes considerably more difficult to train the model because it starts to saturate due to vanishing gradients. This problem is solved by the ResNet model in which residual blocks are introduced, as discussed in [30]. The core concept of this model is to skip connections by adding original data to the output of convolution blocks.

To perform wheat yellow rust infection type classification, the top of the model is replaced with two dense layers, each with a depth of 1024. The ReLU is used as an activation function, which is computed by Eq 1.

$$y = max(0, x) \tag{1}$$

The dense layers are followed by a dropout layer where the dropout ratio is set to 0.5, which reduces the number of connections going into the classification layer. This increases computational efficiency by only using a subset of the total connections to the classification layer. The dropout layer is then followed by the classification layer containing four nodes that carry the softmax function.

The model has been compiled with the adam optimizer which aims to combine the properties of AdaGrad and RMSProp, where Categorical Cross Entropy is used as the loss function which can be computed using Eq 2.

$$CE = -\sum_{i}^{C} t_i log(f(s)_i) \tag{2}$$

2) XCEPTION MODEL

The Xception model is another deep convolutional neural network comprising 71 layers, which makes it more powerful to perform complex image classification and recognition tasks [31]. However, to reduce the computational cost,

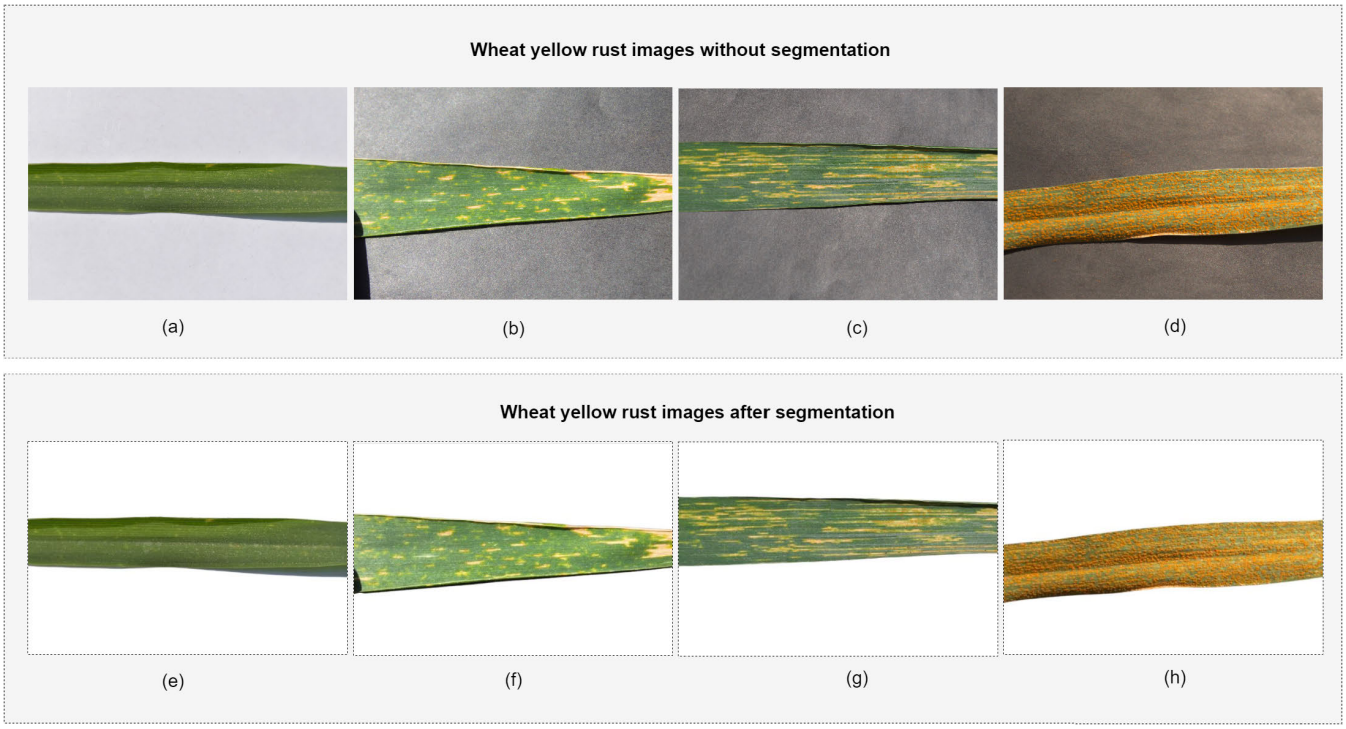

FIGURE 5. U^2 Net segmentation results (a) Healthy image and (e) Segmented image of (a); (b) Resistant image and (f) Segmented image of (b); (c) moderate image and (g) Segmented image of (c); (d) Susceptible image and (e) Segmented image of (d).

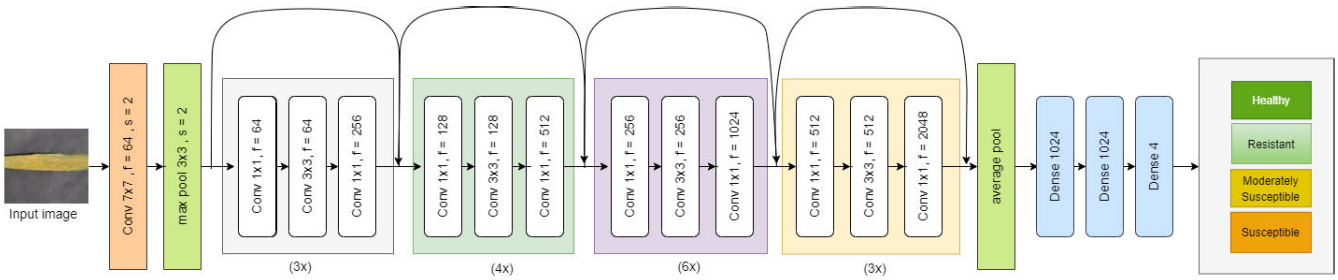

FIGURE 6. Redesigned ResNet-50 model for wheat rust infection type classification.

depth-wise separable convolutions have been introduced in this model. Additionally, there are skip connections between convolutional blocks as shown in Resnet. It initially applies the filters to each depth map before utilizing 1×1 convolution to eventually compress the input space across the depth. This leads to the production of similar results with less computation, making this model extremely efficient in terms of its prediction time and memory. Its architecture has three flows including (i) entry flow, (ii) middle flow, and (iii) exit flow, as discussed in [31].

For wheat rust infection type classification, a pre-trained Xception model is used where model weights are optimized on the ImageNet dataset. The input layer of the model is set as a three-channel image of dimensions 250×250. Meanwhile, the top of the model is where the resultant image from the convolutional layers are flattened and then two dense layers of 1024 nodes activated by the ReLU function are added at the top of the model, as shown in Figure 7. These layers are followed by a dropout layer with a dropout ratio of 0.5, which makes the model more efficient. The final classification layer has 4 nodes and is activated by the softmax function, which is computed by the Eq 3.

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3}$$

Here, the numerator computes the score of a particular node, which represents one of the four classes and is divided by the sum of the squares of all the nodes in the classification layer. This gives a class-based probability of a particular input image belonging to that class. In the softmax function, the sum of all the probabilities in the layer is always equal to 1.

3) PERFORMANCE EVALUATION METRICS

There are several evaluation metrics that are being used in different applications to assess the performance of classifiers including specificity, sensitivity, F1-score, precision, recall,

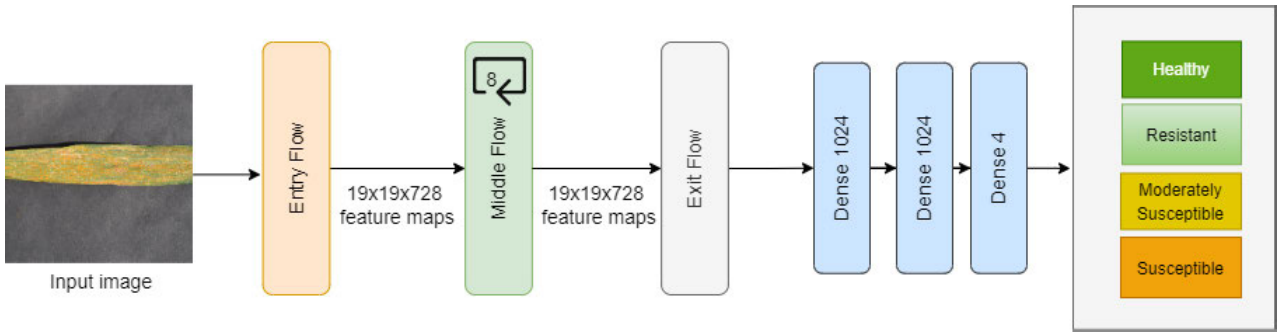

FIGURE 7. Schematic architecture of Xception model for wheat rust infection type classification.

FIGURE 8. Components used in an intelligent wheat rust detection device.

accuracy, and others as discussed in [32] and [33]. In this research, the performance of each classification model is evaluated using accuracy, precision, F1 score, recall, and efficiency as discussed in [34] and [35].

D. ON-DEVICE INFERENCE USING EDGE COMPUTING/EMBEDDED AI

An intelligent edge computing/embedded AI system for wheat stripe rust detection has been developed with the goal of quickly and accurately identifying the disease, and having real-time access to its severity levels. It consists of a camera connected to a Single Board Computer (SBC) and a screen that will capture high-resolution images of wheat, detects the wheat rust disease severity level, and display the results on the screen. The device is powered up with the help of a rechargeable power bank so that it can be used in remote areas and agricultural fields. Figure 8 shows all components used in the wheat rust detection device, where the details of the components are given below:

- Nvidia Jetson Nano: It is a small computer that enables users to run several neural networks simultaneously for image classification, object recognition, segmentation, and audio processing applications (shown in Figure 8). It contains a 128-core NVIDIA Maxwell architecture-based GPU, Quad-core ARM® A57 CPU, 4 GB 64-bitDR4, and 25.6 gigabytes/second memory with Gigabit Ethernet connectivity and operating support of Linux for Tegra®.
- Touch Screen (7 inches): This type of display supports various systems such as Raspberry Pi/ Banana Pi/ NVIDIA Jetson. The built-in HDMI interface enables the display to work as a computer monitor just like any other HDMI screen. The dimension of this screen is 194mm x 110mm x 20mm with a screen resolution of 800 x 480 pixels (shown in Figure 8). It is connected to the Jetson nano developer kit using an HDMI port.
- Logitech Webcam C930e: It is a high-resolution HD 1080p webcam that can work in low-light and severely illuminated environments. Its pixel resolution is 1920 x 1080 pixels with a frame rate of 30 frames per second.
- Power Bank: The developed wheat rust detection device is charged with the help of a power band that has a 20000 mAh capacity. It also has an upgraded digital display which consists of battery percentage, charging voltage, and charging current.

Two state-of-the-art deep learning models are trained on the collected wheat yellow rust dataset, where the most suitable model is deployed on the developed edge device. The ResNet-50 model contains 158,860,164 parameters with a size of 100MB, which is pretty heavy for a small edge device. In order to deploy it on the edge device, the Tensorflow Lite library is used, which enables the machine learning models to run on the embedded systems, mobile or end devices [36]. Subsequently, the wheat rust detection device installed with the ResNet-50 model is tested in a real-time environment where the wheat rust severity level is measured by capturing the wheat rust images as shown in Figure 9.

IV. RESULTS AND DISCUSSION

In order to classify wheat rust into four infection types, two pre-trained deep-learning classifiers are applied. These classifiers are trained on the combined dataset, including the

FIGURE 9. Wheat rust disease detection device testing (a) Testing in the Lab (b) Image capturing in wheat fields, (c) Testing device with constant white background (d) Result of (c).

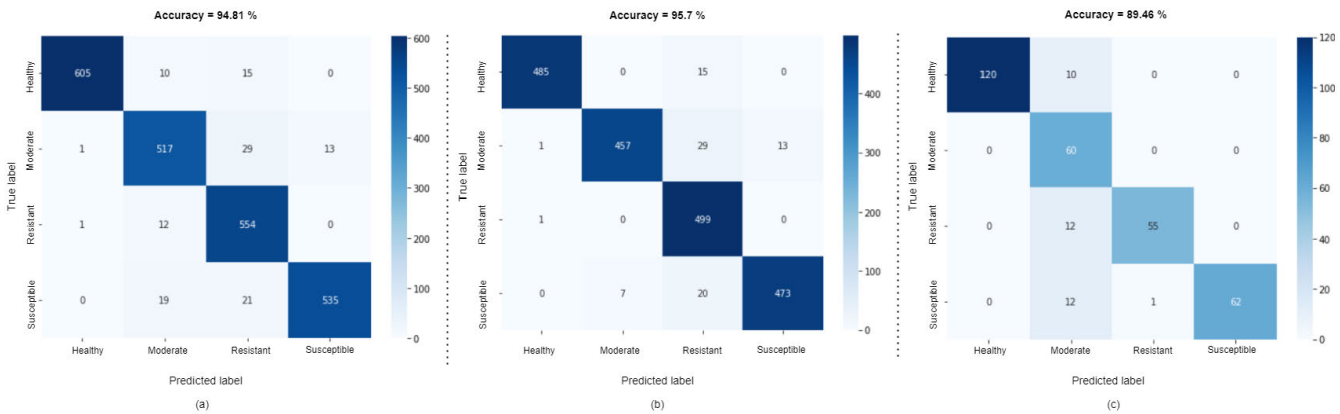

FIGURE 10. Confusion matrix of Xception model (a) Confusion matrix on combined data, (b) Confusion matrix on publically available data, (c) Confusion matrix on collected data.

collected dataset and the publicly available dataset. However, each classifier is evaluated on three different testing datasets, i.e., (i) testing split of our collected data, (ii) testing split of publicly available data, and (iii) combined testing split (containing collected and publicly available datasets). The distribution of each class in training, testing, and validation splits is given in Table 3.

Figure 10 and Figure 11 show the confusion matrices of the Xception model and ResNet-50 model on three testing splits. The diagonal of the matrix shows all correctly classified predictions, whereas all other cells represent miss-classified predictions. This elaborates an in-depth comparison of a model's performance in terms of each class. It is evident that the ResNet-50 model achieved the highest accuracy of 96% on the testing split containing the publicly available dataset and 95.4 % on the combined testing dataset. However, when applied to the collected dataset, the model performance is slightly reduced to 91.87%. Similarly, the Xception model achieved comparable performance on the combined dataset and publically available dataset. However, its accuracy is

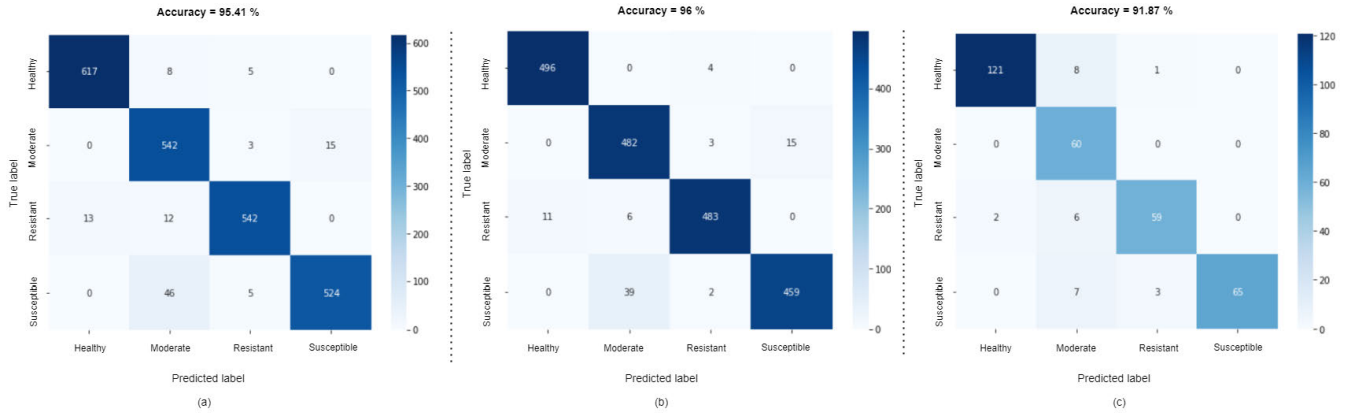

FIGURE 11. Confusion matrix of ResNet-50 model (a) Confusion matrix on combined data, (b) Confusion matrix on publically available data, (c) Confusion matrix on collected data.

TABLE 3. Dataset division into train, test, and validation splits for classification.

Dataset	Healthy	Resistant	Moderate	Susceptible
Training (dataset [26])	1843	1858	1822	1845
Training (collected dataset)	515	281	178	334
Validation split	157	142	178	155
Testing split (dataset [26])	500	500	500	500
Testing split (collected dataset)	130	67	60	75
Testing split (combined dataset)	630	567	560	575

compromised a little when applied to the collected dataset due to the quality of the dataset. The resolution of the collected dataset is not high as compared to the publicly available dataset, which has been captured using high-resolution cameras in a controlled environment. Additionally, the size of our collected dataset is substantially smaller than the publicly available dataset. Consequently, the models have been trained on the majority of images in the publicly available dataset, which makes these models more robust on this dataset, whereas the model's performance is reduced a bit on the collected dataset.

It is evident from the confusion matrices (Figure 10 and Figure 11) that most of the healthy, and susceptible images are correctly classified as compared to the resistant and moderate images. The resistant images are mostly misclassified as healthy whereas healthy images are misclassified in resistant class due to the similarity in leaf color because resistant images contain very small rust disease patterns. Similarly, there is a small difference between moderate and susceptible images, which is the main reason for the misclassification into these classes. Moreover, the collected data contains a smaller number of images in the resistant and moderate classes, which is another reason for the misclassification of these images.

In order to evaluate the model performance and assess the external validity, the performance of each model is tested on a publically available dataset and results are compared with the existing approach for wheat yellow rust detection in terms of accuracy, precision, F1 score, and recall, which is given in Table 4.

It is noticeable from Table 4 that our redesigned Xception and ResNet-50 models achieved the highest performance as compared to the existing approach which obtain the highest accuracy of 91%. The results show that the ResNet-50 model slightly outperforms the Xception model in terms of accuracy (96%) due to the skip connections, which allow contextual information from the previous layers to be transported into the next layer. These connections in the ResNet model enable faster learning because the gradients at different points are added rather than sole multiplication which results in a higher gradient value for the optimizer in each timestamp. Hence, we see that the ResNet-50 model manages to achieve convergence more quickly as compared to the Xception model. The comparison between the accuracy of each model at different epochs is shown in Figure 12. The figure shows how the training of each model progresses with respect to accuracy. Each model has been stopped at different epochs to ensure that the models do not end up as an overfit.

Table 5 shows the number of parameters, size, and prediction time of each model. The memory conservation of the model is based on several factors, including the architecture of the network and the inputs to it, which are reflected in the size of the model. Larger models require more space on

TABLE 4. Performance comparison of models.

Model / Class	Precision	Recall	F1-score	Accuracy
ResNet-50 model				**96%**
Healthy	0.99	0.98	0.99	
moderate	0.96	0.91	0.94	
Resistant	0.97	0.98	0.97	
Susceptible	0.92	0.97	0.94	
Xception model				**95.7%**
Healthy	0.97	1.0	0.98	
moderate	0.91	0.98	0.95	
Resistant	1.0	0.89	0.94	
Susceptible	0.95	0.97	0.96	
Existing work [26]				**91%**
Healthy	0.97	0.98	0.97	
Moderate	0.90	0.82	0.86	
Resistant	0.95	0.97	0.96	
Susceptible	0.93	0.87	0.90	

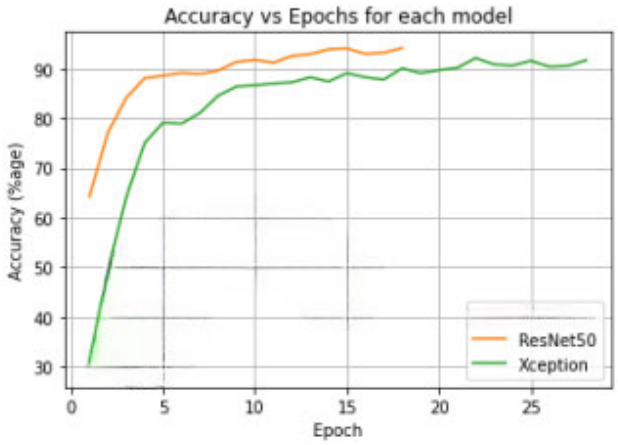

FIGURE 12. Comparison of Accuracy at each epoch.

any deployment device and also take more time to mount. Particularly, models with smaller sizes are more suitable for dealing with IoT microprocessors. Similarly, the number of parameters of the model is another strong indicator of memory optimization since it reduces the workload of the model. The number of parameters is shown in Table 5 which includes both trainable and non-trainable parameters. The total number of parameters represents the approximate number of calculations that each model has to make at each timestamp. Thus, there is an ultimate tradeoff between model size and the number of parameters that must be decided.

TABLE 5. Efficiency comparison of models.

Model	Size (MB)	No of parameters	Prediction time
ResNet-50	100	158,860,164	6 ms
Xception	95	156,133,932	5 ms

Table 5 demonstrates that the Xception model has fewer parameters and size as compared to the ResNet-50 model which results in a better prediction time but ResNet-50 provides a better classification score with minimal difference in prediction time.

V. CONCLUSION AND FUTURE WORK

Wheat rust disease is considered the most damaging disease that can lead to an acute loss in terms of wheat quality and yield. The precise and timely detection of rust disease and its severity is crucial to reduce this loss, which requires a technology-based solution instead of the traditional manual inspection. Towards this end, we proposed a system to detect wheat dust disease in four severity levels, including healthy, resistant, moderate, and susceptible. In order to remove the background, the collected images are segmented using a pre-trained U^2 net. Subsequently, two cutting-edge deep learning models are applied to the dataset including ResNet-50, and the Xception model, where ResNet-50 outperformed with the highest accuracy of 96% with a prediction time of 6 ms.

Additionally, the performance of these models is compared with the existing approach on a publically available dataset where our redesigned models (Xception and ResNet-50) outperformed. Subsequently, an embedded edge device is developed using Nvidia Jetson Nano, a 7-inch Touch Screen, and Logitech Webcam. The trained ResNet-50 model is installed on the wheat rust detection device, which is further tested in the wheat experimental fields. The system will help agricultural users to monitor the wheat crop's health and detect the rust attack and its severity with an accuracy of 93% so that remedial action can be taken on time.

Currently, we have collected the wheat yellow rust disease dataset containing four infection types, including healthy, resistant, moderate, and susceptible. In the future, the frequency of data collection will be increased to cover more classes, such as moderately resistant, and moderately susceptible. Moreover, the data will not only be captured through mobile cameras but drone imagery will also be used to address this problem. For this purpose, drone flights will be carried out at a low altitude to capture high-resolution images. Currently, the wheat rust images are captured in a controlled environment by placing the leaves on a uniform background to deal with the cluttered background and illumination issues. In the future, we will make the data collection phase more flexible so that leaves under rust attack can be segmented automatically. Moreover, we will make the model more robust so that blurred images with variations in illumination can be treated by the model efficiently. Subsequently, other deep learning models will be investigated to accurately segment out the affected wheat leaf and detect the rust attack & its severity levels.

ACKNOWLEDGMENT

This research study was conducted in the collaboration with the NARC.

REFERENCES

[1] P. R. Shewry and S. J. Hey, "The contribution of wheat to human diet and health," *Food Energy Secur.*, vol. 4, no. 3, pp. 178–202, Oct. 2015.

[2] Z. Mahmood, M. Ali, J. I. Mirza, M. Fayyaz, K. Majeed, M. K. Naeem, A. Aziz, R. Trethowan, F. C. Ogbonnaya, J. Poland, U. M. Quraishi, L. T. Hickey, A. Rasheed, and Z. He, "Genome-wide association and genomic prediction for stripe rust resistance in synthetic-derived wheats," *Frontiers Plant Sci.*, vol. 13, p. 66, Feb. 2022.

[3] U. Shafi, R. Mumtaz, I. U. Haq, M. Hafeez, N. Iqbal, A. Shaukat, S. M. H. Zaidi, and Z. Mahmood, "Wheat yellow rust disease infection type classification using texture features," *Sensors*, vol. 22, no. 1, p. 146, Dec. 2021.

[4] X. Chen and Z. Kang, *Stripe Rust*. New York, NY, USA: Springer, 2017.

[5] M. Ouhami, A. Hafiane, Y. Es-Saady, M. El Hajji, and R. Canals, "Computer vision, IoT and data fusion for crop disease detection using machine learning: A survey and ongoing research," *Remote Sens.*, vol. 13, no. 13, p. 2486, Jun. 2021.

[6] U. Shafi, R. Mumtaz, J. García-Nieto, S. A. Hassan, S. A. R. Zaidi, and N. Iqbal, "Precision agriculture techniques and practices: From considerations to applications," *Sensors*, vol. 19, no. 17, p. 3796, Sep. 2019.

[7] H. Orchi, M. Sadik, and M. Khaldoun, "On using artificial intelligence and the Internet of Things for crop disease detection: A contemporary survey," *Agriculture*, vol. 12, no. 1, p. 9, Dec. 2021.

[8] S. T. Jagtap, K. Phasinam, T. Kassanuk, S. S. Jha, T. Ghosh, and C. M. Thakar, "Towards application of various machine learning techniques in agriculture," *Mater. Today: Proc.*, vol. 51, pp. 793–797, Jan. 2022.

[9] Z. Chen, H. Seng Goh, K. Ling Sin, K. Lim, N. Ka Hei Chung, and X. Yu Liew, "Automated agriculture commodity price prediction system with machine learning techniques," 2021, *arXiv:2106.12747*.

[10] J. Chaki and N. Dey, *A Beginner's Guide to Image Preprocessing Techniques*. Boca Raton, FL, USA: CRC Press, 2018.

[11] W. Khan, "Image segmentation techniques: A survey," *J. Image Graph.*, vol. 1, no. 4, pp. 166–170, 2014.

[12] H. R. Bukhari, R. Mumtaz, S. Inayat, U. Shafi, I. U. Haq, 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. 164986–165004, 2021.

[13] A. K. Dewangan, S. Kumar, and T. B. Chandra, "Leaf-rust and nitrogen deficient wheat plant disease classification using combined features and optimized ensemble learning," *Res. J. Pharmacy Technol.*, vol. 15, no. 6, pp. 2531–2538, Jun. 2022.

[14] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Comput. Electron. Agricult.*, vol. 153, pp. 46–53, Oct. 2017.

[15] S. Nigam, R. Jain, S. Marwaha, A. Arora, V. K. Singh, A. K. Singh, R. K. Paul, and K. Immanuelraj, "Automating yellow rust disease identification in wheat using artificial intelligence," *Indian J. Agricult. Sci.*, vol. 91, no. 9, pp. 1–5, Sep. 2021.

[16] W. Haider, A. Ur Rehman, A. Maqsood, and S. Z. Javed, "Crop disease diagnosis using deep learning models," in *Proc. Global Conf. Wireless Opt. Technol. (GCWOT)*, Oct. 2020, pp. 1–6.

[17] L. Goyal, C. M. Sharma, A. Singh, and P. K. Singh, "Leaf and spike wheat disease detection & classification using an improved deep convolutional architecture," *Inform. Med. Unlocked*, vol. 25, Jan. 2021, Art. no. 100642.

[18] S. Sood and H. Singh, "An implementation and analysis of deep learning models for the detection of wheat rust disease," in *Proc. 3rd Int. Conf. Intell. Sustain. Syst. (ICISS)*, Dec. 2020, pp. 341–347.

[19] V. Kukreja and D. Kumar, "Automatic classification of wheat rust diseases using deep convolutional neural networks," in *Proc. 9th Int. Conf. Rel., Infocom Technol. Optim. (Trends Future Directions) (ICRITO)*, Sep. 2021, pp. 1–6.

[20] A. Hussain, M. Ahmad, and I. A. Mughal, "Automatic disease detection in wheat crop using convolution neural network," in *Proc. 4th Int. Conf. Next Gener. Comput.*, 2018, pp. 1–4.

[21] M. Schirrmann, N. Landwehr, A. Giebel, A. Garz, and K.-H. Dammer, "Early detection of stripe rust in winter wheat using deep residual neural networks," *Frontiers Plant Sci.*, vol. 12, p. 475, Mar. 2021.

[22] M. Chohan, A. Khan, R. Chohan, S. Hassan, and M. Mahar, "Plant disease detection using deep learning," *Int. J. Recent Technol. Eng.*, vol. 9, no. 1, pp. 909–914, 2020.

[23] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks," *IEEE Access*, vol. 7, pp. 59069–59080, 2019.

[24] X. Zhang, L. Han, Y. Dong, Y. Shi, W. Huang, and L. Han, "A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images," *Remote Sens.*, vol. 11, no. 13, p. 1554, 2019.

[25] *YELLOW-RUST-19*. Accessed: Jan. 14, 202. [Online]. Available: https://www.kaggle.com/datasets/tolgahayit/yellowrust19-yellow-rust-disease-in-wheat

[26] T. Hayit, H. Erbay, F. Varçın, F. Hayit, and N. Akci, "Determination of the severity level of yellow rust disease in wheat by using convolutional neural networks," *J. Plant Pathol.*, vol. 103, no. 3, pp. 923–934, 2021.

[27] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022.

[28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Munich, Germany. New York, NY, USA: Springer, 2015, pp. 234–241.

[29] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U2-Net: Going deeper with nested U-structure for salient object detection," *Pattern Recognit.*, vol. 106, Oct. 2020, Art. no. 107404.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[31] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258.
[32] T. B. Chandra, B. K. Singh, and D. Jain, "Disease localization and severity assessment in chest X-ray images using multi-stage superpixels classification," *Comput. Methods Programs Biomed.*, vol. 222, Jul. 2022, Art. no. 106947.
[33] T. B. Chandra, K. Verma, B. K. Singh, D. Jain, and S. S. Netam, "Coronavirus disease (COVID-19) detection in chest X-ray images using majority voting based classifier ensemble," *Expert Syst. Appl.*, vol. 165, Mar. 2021, Art. no. 113909.
[34] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques [Internet]*, vol. 745. San Francisco, CA, USA: Morgan & Kaufmann, 2012.
[35] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*. San Mateo, CA, USA: Morgan & Kaufmann, 2022.
[36] *Deploy Machine Learning Models on Mobile and Edge Devices*. Accessed: Sep. 12, 2022. [Online]. Available: http://https://www.tensorflow.org/lite

UFERAH SHAFI received the B.Sc. degree in mathematics from Bahauddin Zakaria University, Multan, Pakistan, the M.Sc. degree in information technology from Quaid-i-Azam University, Islamabad, Pakistan, and the M.S. degree in computer science from COMSATS Institute of Information Technology, Islamabad. She is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Computer Science, NUST, Pakistan. Her research interests include deep learning, remote sensing, image processing, and the Internet of Things (IoT).

RAFIA MUMTAZ (Senior Member, IEEE) received the Ph.D. degree in remote sensing and satellite image processing from the University of Surrey, U.K., in 2010. She is currently a Professor and the Director of the Internet of Things (IoT) Laboratory, NUST-SEECS. She was a recipient of several national and international research grants worth PKR 134 million. She was awarded the NUST-SEECS Best Researcher Award in 2019, the Women of Wonder Award in 2021, and the University Best Teacher Award in 2022.

MUHAMMAD DEEDAHWAR MAZHAR QURESHI received the bachelor's degree in software engineering from NUST-SEECS, in 2022. He is currently a Doctoral Researcher with Technological University Dublin, Dublin, Ireland, under the Science Foundation Ireland, Machine-Learning Laboratories Initiative, where his research domain focuses on the application of XAI (explainable AI) for social media moderation. His research interests include explainability and sustainability in machine learning and the application of machine learning for environmental and medical applications.

ZAHID MAHMOOD received the Ph.D. degree from Quaid-i-Azam University, Islamabad, Pakistan. He completed the Thesis Research in plant sciences from The University of Queensland, Australia. He is currently a Senior Scientific Officer with the Crop Sciences Institute, National Agricultural Research Centre, Islamabad, Pakistan. He has more than 19 years of professional experience in crop improvement, breeding, and plant genetics. He has published more than 40 research articles in international journals and conferences. He received several international fellowships/training in leading institutes and universities in USA, Australia, China, Mexico, Kenya, and India. His current research interests include the areas of high-throughput genotyping and phenotyping using UAV-based sensors for crop evaluation and development.

SIKANDER KHAN TANVEER received the B.Sc. and M.Sc. degrees (Hons.) in agronomy from the University of Agriculture, Faisalabad, Pakistan, and the Ph.D. degree in crop cultivation and farming system from the Northwest A&F University, Yangling, China. He is currently a Principal Scientific Officer and a Program Leader of the Wheat Program PARC, NARC, Islamabad. He has published a number of papers and book chapters.

IHSAN UL HAQ received the B.Sc. degree in electrical engineering from the University of Engineering and Technology (UET), Taxila, Pakistan, and the M.Sc. degree in electrical engineering from SEECS, NUST, Islamabad, Pakistan. He is currently a Research Assistant with the Internet of Things (IoT) Laboratory, SEECS-NUST, Islamabad, where he is working on multiple projects related to embedded systems, machine learning, and the IoT. His research interests include deep learning, computer vision, embedded systems, and the IoT.

SYED MOHAMMAD HASSAN ZAIDI received the Ph.D. degree from the University of South Florida, USA, in 1992. He is currently a former Principal and the Dean of the School of Electrical Engineering and Computer Science (SEECS), NUST. He is also a Senior Research Faculty with NUST-SEECS. He has won international research grants and has more than 100 research publications in reputed journals and conferences. His research interests include wireless and photonics networks, ultra high-speed communication systems, network security, and the Internet of Things (IoT). He was a recipient of several coveted awards, including the Best University Researcher and Teacher and the Presidents' Pride of Performance for his National Level Meritorious Service in IT education and research. He has been the pioneer steering committee International Co-Chair of IEEE sponsored international conference HONET-ICT, for the past 16 years.