**Abstract**

Rice is commonly found in every Indian plate either as a dish or as an essential ingredient of the dish, but its agriculture consists of a lot of challenges due to the diseases which can occur in rice leaves. A few of these diseases which were focused on with this research includes Brown spot, Leaf smut, and Bacterial leaf Blight. Few neural network models were given a chance for this cause from which Resnet18 proved to be the best with a high level of accuracy (95.47%). For, converting this model into a usable form a GUI was also made through the use of the Tkinter library.

1. **Introduction**

Agriculture is a debatable sector in India with rice as one of the common crops, especially in Odisha, and several diseases can affect these rice crops. Therefore, detecting them earlier would play an important role, it will improve agriculture, food security, etc. This research was conducted to solve these issues and was highly impressed with the features of deep learning models due to their high effectiveness for crop disease detection. A lot of papers were given a thorough check to gather some information on this topic and figure out the useful methods.

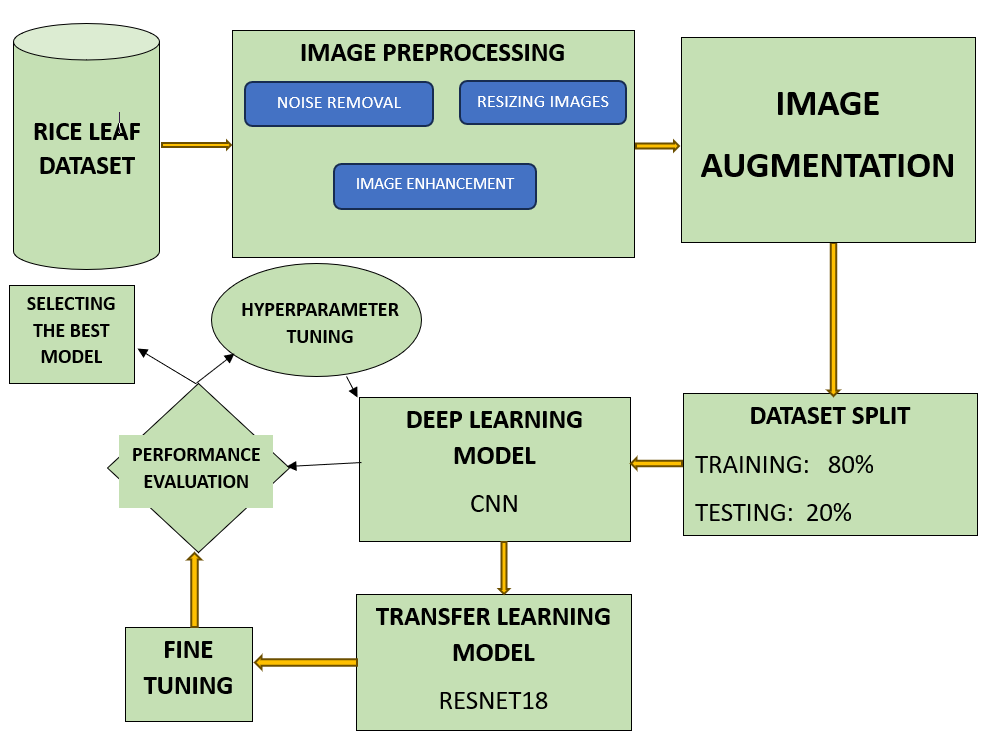
1. **Aim and Objectives**

This research grabbed the opportunity to solve the agricultural problems regarding rice leaf diseases and aimed to gather a dataset of rice leaf images, use image processing to improve the dataset, and accurately detect the diseases through the means of neural networks. It also aimed to make the model reliable enough that it aligns well with the GUI for telling the disease the rice leaf is facing through the images selected.

1. **Related Work**

By [1], the focus was on blast and tango; it also covered a few of the diseases that this research covered. It used various Convolutional Neural Network (CNN) models, with the integration of transfer learning. Almost 6000 images of disease-affected rice leaves were obtained from the agriculture sites of Sambalpur and Bargarh district, Odisha. Further, the performance was analyzed of small models like MobileNetv2 and Shufflenet (in both transfer learning). The deep feature of ResNet50 plus Support Vector Machine (SVM) proved to be better in competition with the transfer learning method with an F1 Score of 0.9838. Combining the effectiveness of SVM along with deep learning as in [2], rice blight diseases (sheath and stripe), as well as blast, were classified using datasets from rural farmland and rice leaf disease atlas. Some CNN-based programs were used for extracting the features, followed by SVM to further find out the disease. The achieved accuracy was almost 97% when the penalty parameter c was 1 and kernel parameter g was 50. Three types of rice leaf diseases and two types of wheat leaf diseases were studied in [3]. The Visual Geometry Group Network – 16 (VGG16) model was used and improved with the use of multi-task learning, transfer learning, and alternate learning techniques. Transfer learning was used from a pre-trained model on ImageNet to increase classification performance. Separate classification layers for each dataset were made to address multi-tasking. The models were then compared with the other methods like ResNet50 and DenseNet121. The accuracy was 97%+ in the case of rice disease classification and 98%+ in the case of wheat leaf disease classification, and it showed that the models consisting of multi-learning do better than single-task models, reuse-model methods in transfer learning, and various other models like ResNet50 and DenseNet121. Exploring the methodology of [4], the researchers used the CNN model of InceptionResNetV2 along with a transfer learning approach. The dataset in this study included a few rice leaf images of 3 categories (leaf blast, brown blast, and bacterial blight) which were increased to 5000+ through augmentation. The data needs in this study were fulfilled through the data available on Kaggle and Google Images. The attention-based mechanism helped to produce a mode that had high accuracy (95%+). In [5], Federated Learning played its role in preparing the model. This model was trained on remote datasets without the transfer of data from client nodes. This study also collected its data from fields, but it also used the agricultural pest and insect pests picture dataset and then this data was pre-processed including resizing images to 75x75 and applying grayscale. The model contained several fully connected layers, and 13 pre-trained models like VGG, ResNet, and Inception were used as inputs for training the classifier. The EfficientNetB3 was highly accurate (99%) during the classification, which outperformed other models like VGG16, VGG19, Resnet152, DenseNet, InceptionResnetV2, Xception, etc. The diseases that were aimed to be detected in [6] were the same as this study aimed and were detected using the AlexNet model. The dataset considered here was picked from Kaggle which included 120 images initially but after augmentation, it increased to 900 images. Further, it was divided so that 70% was for model training and the rest for testing. Alexnet was considered here for the prediction of classifying rice leaf diseases. A fully connected, Softmax, classification output layer was used instead of the last 3 layers. A high level of accuracy (almost 99.4%) was achieved while classifying the diseases using AlexNet neural networks. [7] decided to create a system to automate the process of finding the disease of rice images with CNNs, thereby ensuring better yield and rice agriculture. For the same, 100s of rice images including stems and leaves were captured from fields, these were pre-processed, followed by the training process, resulting in a highly accurate model (almost 95.5% accuracy) in disease identification which was higher than a lot of the other models. The goal of [8] included using the BLSNet method for automated recognition and segmentation of the rice disease called bacterial leaf streak lesions. Less than 200 images of rice leaves were collected from the fields around Shanji, Xuzhou, and JiangSu Province, China under different sunlight intensities and capture angles, and then increased to 1000+ images using image augmentation. The disease severity level was determined, after which training was done followed by evaluation. This resulted in high levels of accuracy with BLSNet taking a longer time to predict than UNet but shorter than DeepLabV3+, etc. A system was expected to be made to identify 2 types of rice leaf diseases: leaf blast disease and brown spot disease [9]. The dataset here included about 200 rice leaf images, and the YOLO (You Only Look Once) Algorithm was used. Leaf Blast disease was identified with an accuracy of 90%, while it was 70% when identifying brown spot diseases and 100% for the unknown disease. Hence, overall accuracy came to be almost 75% with almost 25% as the error of commission. Talking about [10], 1000+ images were gathered including the ones available online as well as from real-life agricultural fields with heterogeneous backdrops and variable lighting intensities were collected. Out of various CNNs, the confidence was the highest over MobileNetV2, which was pre-trained on ImageNet and was considered for identifying diseases. Few techniques were considered in this model for analyzing the inter-channel relationships and spatial points for input features. A bunch of methods were considered like transfer learning and augmentation, and CNN training and testing were conducted using Anaconda 3 (Python 3.6), Keras-GPU library, and OpenCV-python3 library. This showed 99%+ efficiency in prediction, which overall was almost 98.5%, this outperformed the other models. Researchers wanted to create a system as in [11], namely: “an attention-based depth-wise separable neural network with Bayesian optimization” (ADSNN-BO), which was able to detect rice diseases by analyzing images of its leaves with 94%+ accuracy. However, it might make mistakes in recognizing diseases and could be hard to use widely because of technical and resource limits. In [12] the importance of the proper production of rice was highlighted. CNNs as well as support vector machines (SVM) were considered for the purpose. The aim was that there should be no yield loss due to any kind of disease and they should be detected timely with some robust techniques. Even though it did well, sometimes it might misidentify diseases. A deep learning approach, DenseNet169-MLP, is highly useful as discussed in [13] to diagnose rice plant diseases, important for Asian agriculture. Preprocessing stages include channel separation, grayscale conversion, and noise removal. Fuzzy c-means (FCM) segmentation separated diseased areas, and DenseNet169 showed high importance during the extraction of necessary features. The last layer, which was changed to a Multilayer Perceptron (MLP), helped classify diseases and was very accurate (97%+). Exploring [14], the paper focused on rice plant disease identification, which had high needs in certain areas, especially in continents of Africa as well as Asia where rice is a big part of people's diet. Various techniques (image processing and deep learning techniques) were explored to identify rice plant diseases. Pre-existing techniques were also explored that dealt with rice disease detection like decision trees and neural networks. Variations in the environment and how diseases appear can create problems. Exploring the importance of precision in agriculture and the need to utilize technology as discussed in [15] to predict plant sickness based on leaf images, the study proposed certain methods, that can be effective at increasing crop yield due to their time efficiency. A lot of methods were given a chance, and it was also made sure that the dataset was kept properly like the colors in the images of the dataset were balanced and important features were pulled out. The methods that featured machine learning were the best. The study [16] aimed to find early rice disease detection methods which use rice images. A special setup was used consisting of sensors and a camera to take pictures, and a CNN model was selected to tell the disease that the rice leaf had. The models that were featured here showed great performance, especially after improving the dataset. Further, the author aimed to make a comparison of CNN with other methods and improve the setup, although it may not be that viable in different places and with different crops. Rice farming is important in India as discussed by [17], but certain diseases are important to tackle, lowering how much rice is grown and its quality. Diseases that were aimed target in this study were Leaf Blast, Hispa, and some which this paper is covering, which are common, and also affect the economy and food supply. The study included some techniques for the identification of these diseases and evaluated them. It might be hard to use this model in certain scenarios because resources can get over it and it's too complicated. As discussed by [18], In places like Nepal, where rice is a common ingredient in people’s diets, diseases in rice plants are a big problem for farmers. To help tackle this issue, some experts decided to provide a system using computers to inform users about such issues. Certain methods were used for improving the images and for training the model Twin Support Vector Machine (TSVM) was hired. Paddy sample images undergo RGB calculation and binary conversion, with normal samples classified automatically based on RGB range. The algorithm the researchers came up with in [19] aims to find diseases in rice plants early, which are big problems for rice crops. The experts in this study considered using the method of image detection which would spot the diseased segments in rice leaves, no matter what the conditions were. This task was a bit complex, mainly due to the challenge of separating disease spots from busy backgrounds. Discussing the techniques proposed by [20], it mainly covered hyperspectral imaging (HSI) which was quick and precise enough in detecting diseases for a variety of rice leaves. Some of the methods discussed in this study include CNN, transfer learning, and fine-tuning. Fine-tuning worked best, as it reached over 88% accuracy in most tasks. Deep CORAL achieved over 80% accuracy in four tasks. Also, multi-tasking and pairwise tasking were both useful. It used saliency maps to show the important wavelength ranges detected by the CNN, both with and without transfer learning, and it had similar results.

1. **Research Methodology**



***Figure 1****. Methodology*

Dataset

It started with a search for the appropriate dataset that ended on the Rice leaf diseases image dataset on Kaggle. This dataset contained 120 jpg images of disease-infected rice leaves as in *Figure 2*. There were 40 images in each class including Leaf smut, Brown spot, and Bacterial leaf blight that were clicked from a village. The dataset was first downloaded, after which the necessary actions were performed to convert it into usable form.

A close-up of a plant

Description automatically generated

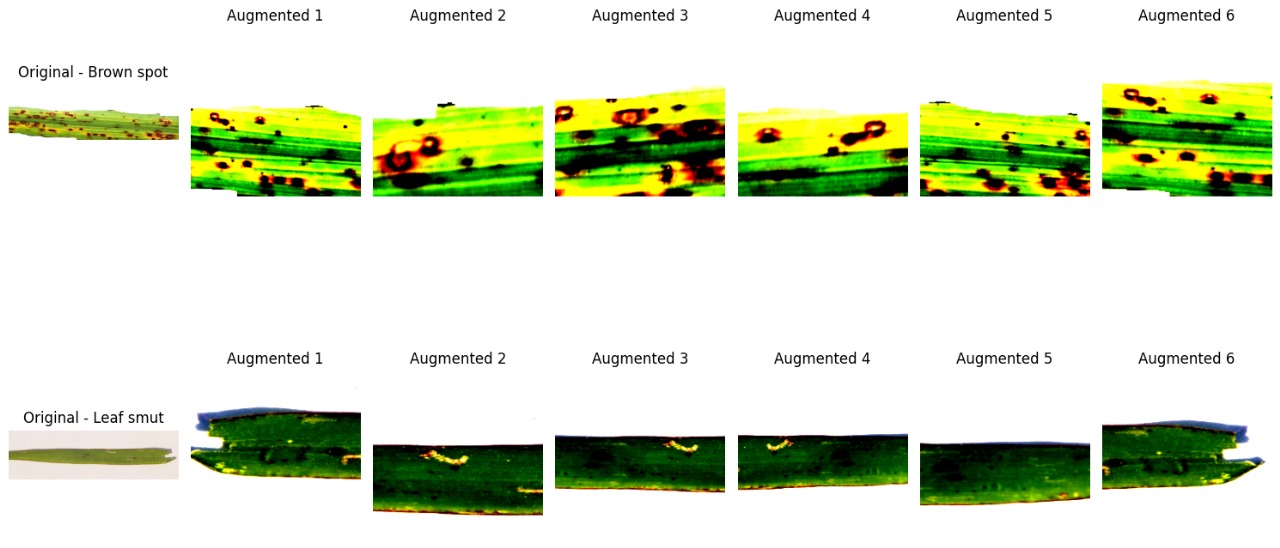
***Figure 2****. Affected Rice leaves*

Data Preprocessing and Visualization

A lot of measures were taken to convert the data into a format that would be efficient for training the model, the steps included were:

1. Dataset Inspection: The visual aspects of each rice leaf were analyzed to gain some information about any important feature or pattern.

2. Image Augmentation: Due to the shortness of images, image augmentation was considered with the hope that a large dataset would be generated which would be sufficient to train the dataset. Augmentation was possible by considering the ImageDataGenerator() function, as in *Figure 3*. It generated 1000+ images as discussed in *Table 1*. For storing these images, a separate folder was created.



***Figure 3****. Images of Brown spot and Leaf smut after augmentation*

|  |  |  |
| --- | --- | --- |
| Name of the  Disease | No. of Images before Augmentation | Images after Augmentation |
| Leaf Smut | 40 | 600 |
| Brown Spot | 40 | 585 |
| Bacterial Leaf Blight | 40 | 690 |

***Table 1****. Rice Leaf Images that were affected by diseases*

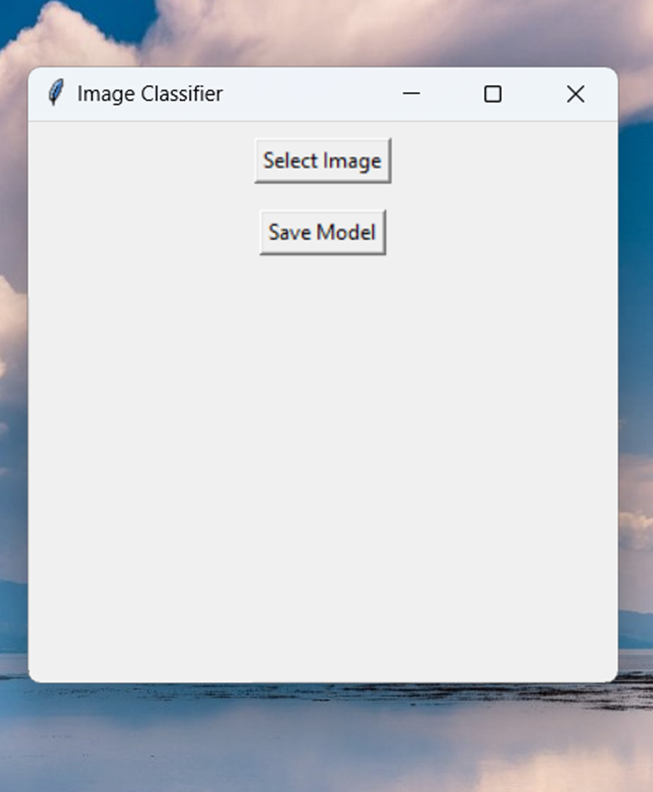
3. Data Cleaning: The dataset obtained after the augmentation consisted of a lot of inefficient images which were potential outliers to the model, for this the data was manually explored and such images were deleted from the dataset. After this, these images were split and separately stored to divide them into training and validation.

Model Selection and Evaluation

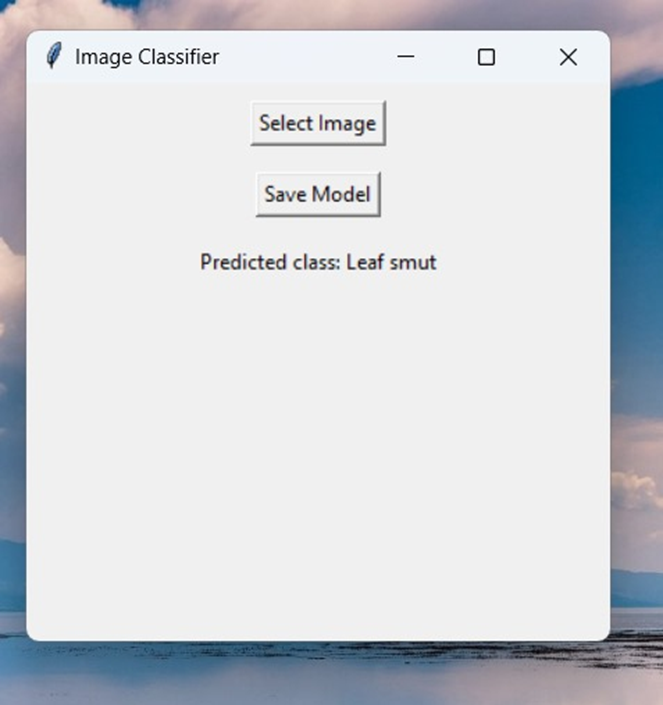
For predicting the diseases accurately, the approach of transfer learning was adopted in which several pre-trained models were imported using the Pytorch library and further trained over the dataset. For selecting the model which would be best fit here, all the models were evaluated based on their accuracies, losses, etc., along with some graphs.

Graphical User Interface (GUI) Development

After successfully training and evaluating the best model for the purpose it was saved as a pickle file using the pickle library so that it could be further used. The Tkinter library was highly useful for the GUI part as in *Figure 3. And Figure 4.*, and the Pillow library for handling the image part.



***Figure 4****. GUI of the rice disease classification app*



***Figure 5****. App telling the disease rice leaf has*

1. **Results and Discussions**

The models were compared to decide the one which would be best fit for the app, for which two tables were plotted, one for comparing the metrics of the models at their best validation accuracies as in *Table 2*., and the other for comparing the metrics of these models at equal intervals as in *Table 3*., *Table, 4*., *Table 5*. While training each of these models, a learning rate of 0.001 and momentum of 0.9 were specified to the optimizer, and at least 50 epochs were run.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model  Name | Number of Epochs | Training Loss | Training Accuracy | Validation  Loss | Validation Accuracy |
| AlexNet | 20 | 0.2960​ | 0.8598​ | 0.2388​ | 0.8966​ |
| ResNet50 | 11 | 0.3083 | 0.8552​ | 0.2331 | 0.9310 |
| ResNet18 | 30 | 0.2678 | 0.8853​ | 0.1376​ | 0.9547​ |

***Table 2****. Metrics of models at their highest validation accuracy*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number of Epochs | Training Loss | Training Accuracy | Validation  Loss | Validation Accuracy |
| 10h | 0.4069 | 0.8098 | 0.3477 | 0.7977 |
| 20th | 0.3376 | 0.8443 | 0.2666 | 0.8870 |
| 30th | 0.3300 | 0.8460 | 0.2707 | 0.8897 |
| 40th | 0.2963 | 0.8603 | 0.2518 | 0.8920 |
| 50th | 0.2906 | 0.8672 | 0.2372 | 0.8897 |

***Table 3****. Metrics of AlexNet*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number of Epochs | Training Loss | Training Accuracy | Validation  Loss | Validation Accuracy |
| 10th | 0.2965 | 0.8626 | 0.2231 | 0.8966 |
| 20th | 0.3032 | 0.8690 | 0.2173 | 0.8943 |
| 30th | 0.2899 | 0.8678 | 0.2069 | 0.8897 |
| 40th | 0.2949 | 0.8695 | 0.2104 | 0.8874 |
| 50th | 0.2736 | 0.8810 | 0.2096 | 0.8920 |

***Table 4****. Metrics of ResNet50*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number of Epochs | Training Loss | Training Accuracy | Validation  Loss | Validation Accuracy |
| 10th | 0.3341 | 0.8480 | 0.1669 | 0.936 |
| 20th | 0.2655 | 0.8793 | 0.1518 | 0.9387 |
| 30th | 0.2320 | 0.8913 | 0.1475 | 0.9467 |
| 40th | 0.2239 | 0.9047 | 0.1754 | 0.9227 |
| 50th | 0.1842 | 0.9280 | 0.1191 | 0.9387 |

***Table 5****. Metrics of ResNet18*

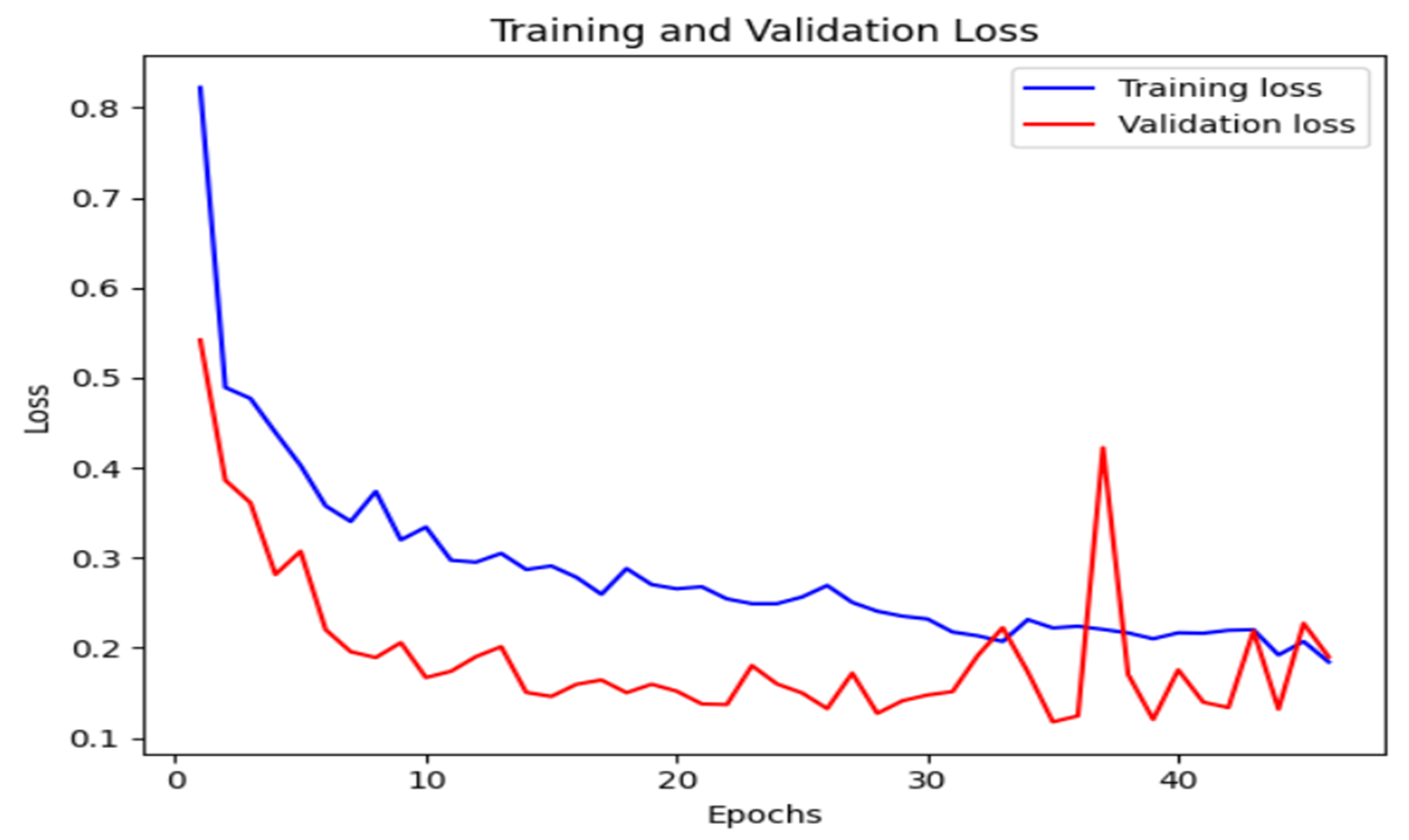
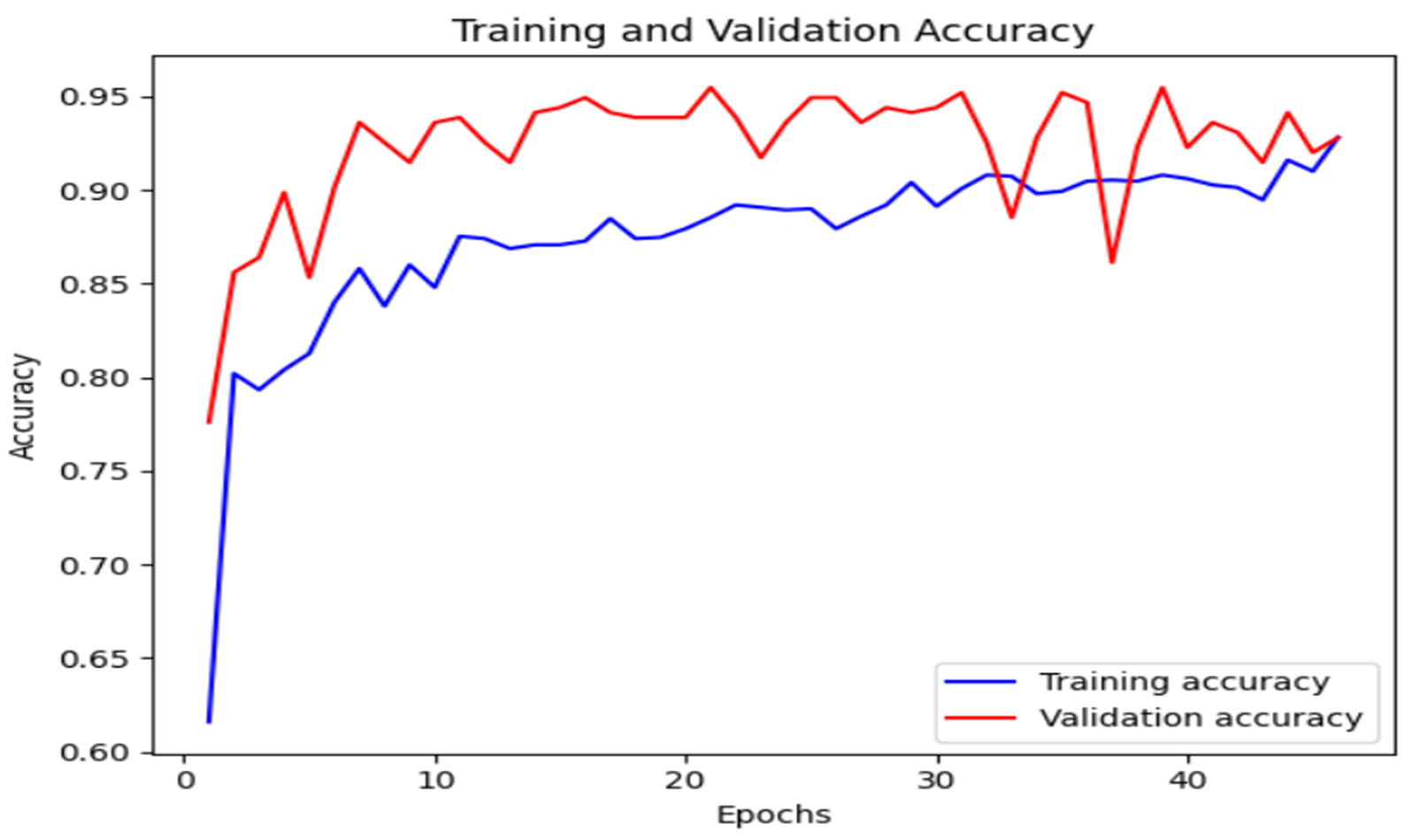
After looking at the accuracies of these models, it was clear that the model to consider for this app would be ResNet18. The first time this model showed a validation accuracy of 90% or more was at just the 7th epoch (93.6%), and the highest it reached was at the 30th epoch which was 95.47%, and a training accuracy of 88.53% with the lowest training loss of 0.2678 and validation loss of 0.1376. Other models including mainly ResNet50 (93.1% accuracy) were also relevant as a second option.

**

***Figure 6****. Graphs to visualize accuracy and loss with each epoch of AlexNet*

**

***Figure 7****. Graphs to visualize accuracy and loss with each epoch of ResNet50*

***Figure 8****. Graphs to visualize accuracy and loss with each epoch of ResNet18*

1. **Conclusion and Future Work**

The goal associated with this project can be considered successful since an interactive GUI was successfully created, where users can choose the images of their rice leaves and the model will tell the disease it is facing. The model associated with the GUI was accurate enough (95.47%) in identifying the disease rice leaf has. In the Future, more online and real-life datasets would be a good idea to train the model better, making the model more accurate and reliable.

**References**

[1] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, “Deep feature-based rice leaf disease identification using support vector machine,” *Computers and Electronics in Agriculture*, vol. 175, p. 105527, 2020, Doi: https://doi.org/10.1016/j.compag.2020.105527.

[2] F. Jiang, Y. Lu, Y. Chen, D. Cai, and G. Li, “Image recognition of four rice leaf diseases based on deep learning and support vector machine,” *Computers and Electronics in Agriculture*, vol. 179, p. 105824, 2020, doi: https://doi.org/10.1016/j.compag.2020.105824.

[3] Z. Jiang, Z. Dong, 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, p. 106184, 2021, doi: https://doi.org/10.1016/j.compag.2021.106184.

[4] K. N, L. V Narasimha Prasad, C. S. Pavan Kumar, B. Subedi, H. B. Abraha, and S. V E, “Rice leaf diseases prediction using deep neural networks with transfer learning,” *Environ Res*, vol. 198, p. 111275, 2021, doi: https://doi.org/10.1016/j.envres.2021.111275.

[5] M. Aggarwal, V. Khullar, N. Goyal, A. Alammari, M. A. Al Bahar, and A. Singh, “Lightweight Federated Learning for Rice Leaf Disease Classification Using Non-Independent and Identically Distributed Images,” *Sustainability (Switzerland)*, vol. 15, no. 16, Aug. 2023, doi: 10.3390/su151612149.

[6] Md. M. H. Matin *et al.*, “An Efficient Disease Detection Technique of Rice Leaf Using AlexNet,” *Journal of Computer and Communications*, vol. 08, no. 12, pp. 49–57, Nov. 2020, doi: 10.4236/jcc.2020.812005.

[7] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, “Identification of rice diseases using deep convolutional neural networks,” *Neurocomputing*, vol. 267, pp. 378–384, 2017, doi: https://doi.org/10.1016/j.neucom.2017.06.023.

[8] S. Chen *et al.*, “An approach for rice bacterial leaf streak disease segmentation and disease severity estimation,” *Agriculture (Switzerland)*, vol. 11, no. 5, 2021, doi: 10.3390/agriculture11050420.

[9] Ma. K. Agbulos, Y. Sarmiento, and J. Villaverde, “Identification of Leaf Blast and Brown Spot Diseases on Rice Leaf with YOLO Algorithm,” in *2021 IEEE 7th International Conference on Control Science and Systems Engineering (ICCSSE)*, 2021, pp. 307–312. doi: 10.1109/ICCSSE52761.2021.9545153.

[10] J. Chen, D. Zhang, A. Zeb, and Y. A. Nanehkaran, “Identification of rice plant diseases using lightweight attention networks,” *Expert Syst Appl*, vol. 169, p. 114514, 2021, doi: https://doi.org/10.1016/j.eswa.2020.114514.

[11] Y. Wang, H. Wang, and Z. Peng, “Rice diseases detection and classification using attention-based neural network and Bayesian optimization,” *Expert Syst Appl*, vol. 178, p. 114770, 2021, doi: https://doi.org/10.1016/j.eswa.2021.114770.

[12] T. S. Poornapriya and R. Gopinath, “Article ID: IJEET\_11\_10\_050 Artificial Intelligence Approaches,” *International Journal of Electrical Engineering and Technology (IJEET)*, vol. 11, no. 10, pp. 392–402, 2020, doi: 10.34218/IJEET.11.10.2020.050.

[13] R. P. Narmadha, N. Sengottaiyan, and R. J. Kavitha, “Deep transfer learning-based rice plant disease detection model,” *Intelligent Automation and Soft Computing*, vol. 31, no. 2, pp. 1257–1271, 2022, doi: 10.32604/iasc.2022.020679.

[14] V. Sandhya Venu, B. Kiranmai, S. Venu Vasantha, and S. Rama Krishna, “Techniques for Rice Leaf Disease Detection using Machine Learning Algorithms.” [Online]. Available: www.ijert.org

[15] H. Pallathadka *et al.*, “Application of machine learning techniques in rice leaf disease detection,” *Mater Today Proc*, vol. 51, pp. 2277–2280, 2022, doi: https://doi.org/10.1016/j.matpr.2021.11.398.

[16] T. Tawde, L. Verekar, S. Aswale, K. Deshmukh, A. Reddy, and P. Shetgaonkar, “Rice Plant Disease Detection and Classification Techniques: A Survey.” [Online]. Available: www.ijert.org

[17] P. Tejaswini, P. Singh, M. Ramchandani, Y. K. Rathore, and R. R. Janghel, “Rice Leaf Disease Classification Using CNN,” in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics, 2022. doi: 10.1088/1755-1315/1032/1/012017.

[18] B. Chawal and S. P. Panday, “RICE PLANT DISEASE DETECTION USING TWIN SUPPORT VECTOR MACHINE (TSVM),” 2019.

[19] K. S. Archana and A. Sahayadhas, “Automatic Rice Leaf Disease Segmentation Using Image Processing Techniques,” 2018. [Online]. Available: www.sciencepubco.com/index.php/IJET

[20] L. Feng, B. Wu, Y. He, and C. Zhang, “Hyperspectral Imaging Combined with Deep Transfer Learning for Rice Disease Detection,” *Front Plant Sci*, vol. 12, Sep. 2021, doi: 10.3389/fpls.2021.693521.