Yuzuncu Yil University Journal of Agricultural Sciences, Volume: 32, Issue: 4, 30.12.2022

Yuzuncu Yil University
Journal of Agricultural Sciences
(Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi)
https://dergipark.org.tr/en/pub/yyutbd

ISSN: 1308-7576
e-ISSN: 1308-7584

Research Article

Identification of Some Paddy Rice Diseases Using Deep Convolutional Neural Networks

Hakkı Fırat ALTINBİLEK[1], Ünal KIZIL*[2]

[1]İpsala District Directorate of Agriculture and Forestry, İpsala, Edirne, Türkiye
[2]Çanakkale Onsekiz Mart University, Agricultural Structures and Irrigation Dept., Çanakkale, Türkiye

[1]https://orcid.org/0000-0001-6761-1445 , [2]https://orcid.org/0000-0002-8512-3899

*Corresponding author e-mail: unal@comu.edu.tr

Article Info

Received: 05.07.2022
Accepted: 30.09.2022
Online published: 15.12.2022
DOI: 10.29133/yyutbd.1140911

Keywords

Deep learning,
Image classification,
Paddy rice disease

Abstract: In modern digital agricultural applications, automatic identification and diagnosis of plant diseases using artificial intelligence is becoming popular and widespread. Deep learning is a promising tool in pattern recognition and machine learning and it can be used to identify and classify diseases in paddy rice. In this study, 2 different paddy rice diseases, including rice blast and brown spot, were investigated in the district of İpsala in the province of Edirne between the 2020 and 2021 production seasons by collecting 1569 images. These diseases are very common and important in Edirne province and surrounding rice production areas. Therefore, practical methods are needed to identify and classify these two diseases. A Convolutional Neural Network (CNN) model was created by applying pre-processing techniques such as rescaling, rotation, and data augmentation to the paddy rice disease images. The classification model was created in Google Colab, which is a web-based Python editor using Tensorflow and Keras libraries. The CNN model was able to classify rice blast and brown spot diseases with high accuracy of 91.70%.

To Cite: Altınbilek, H F, Kızıl, Ü, 2022. Identification of Some Paddy Rice Diseases Using Deep Convolutional Neural Networks. *Yuzuncu Yil University Journal of Agricultural Sciences,* 32(4): 705-713. DOI: https://doi.org/10.29133/yyutbd.1140911

1. Introduction

There are many fungal diseases that affect the yield and quality of paddy rice. Two major diseases that can cause economic losses in Türkiye are rice blasts (*Pyricularia oryzae*) and brown spot (Helmithosporium oryzae). If these diseases spread, the plant may dry out completely, and sometimes, no crop may be harvested. With the developments in computer science in parallel with technology, the detection of such diseases in a shorter time and with minimal experience has become possible (Arnal Barbedo, 2013).

Artificial intelligence (AI) or machine learning is one of the important tools that is used in image-based identification and classification. During the classification of images, computational complexity should be avoided, and the working time should be greatly reduced. For this purpose, Deep learning (DL) is used to learn characteristic hierarchies, increase representation capacity and provide better prediction performance (Affonso et al., 2017; Zhang et al., 2019; Wang et al., 2020). The coding of the model can be done through computer models that offer interactive as well as web-based platforms. Google Colaboratory is the web-based framework for writing and executing Deep Learning and Machine Learning codes. It can upload larger datasets directly from Google Drive servers at very high speed (Kanani and Padole, 2019).

Convolutional Neural Network (CNN) is generally used in image-based AI applications. It is a neural network approach that consists of image acquisition, pre-processing, fragmentation, extraction of features from images, detection, and classification (Rawat and Wang, 2017; Kadhim and Abed, 2018). By optimizing the model parameters, the difference between the actual and predicted output values can be minimized in the classification. Many studies have been carried out on plant diseases using the advantages of CNN (Kawasaki et al., 2015; Boulent et al., 2019; Ferentinos, 2019; Priyadharshini et al., 2019; Gokulnath and Usha Devi, 2021).

Due to the inadequacy of the appropriate methods to detect rice plant leaf diseases, paddy rice production has been decreasing gradually in recent years. Thus, a user-friendly and fast recognition system is needed. CNN is a useful tool that can be considered in this sense (Pinki et al., 2017; Mique and Palaoag, 2018; Vardhini et al., 2020). Lu et al., (2017) reported the performance of CNN in paddy rice disease detection by using 500 images. They showed that comparing the alternative approaches to predict diseases, and CNN is the best of the classification models considered. Asfarian et al. (2013) diagnosed four significant rice plant diseases such as brown spots, leaf blight, leaf blast, and tungro. According to the results of the study, CNN is the fastest and most accurate classifier. Xiao et al. (2018) showed that CNN is superior to the current machine learning technique with great accuracy in differentiating the impact of rice diseases. Vanitha (2019) used CNN to detect leaf blight, bacterial blight, and stem rot diseases in paddy rice. They conducted the study with 500 images, both diseased and healthy. They stated that the model could effectively detect three rice diseases with an accuracy of 99.53%. Tawde et al. (2021) used CNN as a feature extractor and support vector machine (SVM) as a classifier for the identification of rice diseases. They reported that rice diseases could be successfully identified early with this approach with an accuracy rate of 96%. In this study, a model based on a deep CNN was developed to identify two 2 major fungal rice diseases in the Edirne province of Turkey.

2. Material and Methods

The study was carried out on paddy rice fields in the İpsala district of Edirne province/Turkey in 2020 and 2021. İpsala district is located at latitude 40.8865 and longitude 26.3712 (Figure 1). Approximately 14% of the rice production in Turkey is made in İpsala. Images of healthy and infected rice samples from cultivars such as Casanova, Güneş-Cl, and Keşhan were acquired to identify blast (*Pyricularia oryzae*) and brown spot *(Helmithosporium oryzae*) diseases. The symptoms of the blast in the plant are seen in leaves, nodes, clusters, peduncles, and husks. These symptoms are generally lozenge-shaped with two pointed ends, gray in the middle, and brown-reddish spots around it. On the stem, it looks like an oil stain, and green mold develops. It causes no grains to form in the cluster and the formation of white husks (Devi and Neelamegam, 2018). Leaves infected with brown spots have many large spots on the leaves that may kill the entire leaf. The rapid spread and development of the disease are fostered by continuous rainfall, clouds, and higher daytime temperatures. The yield can be reduced by up to 45% in cases of severe illness and by 12% in cases of mild infection (IRRI, 2012).

Image acquisition started from the second week of July (the first appearance of the diseases) in 2020 and 2021 and continued until the end of August in both years. Images were obtained from 15 infected, and 5 healthy rice fields were used. A total of 1569 RGB images were acquired using Redmi Note 9 Pro.

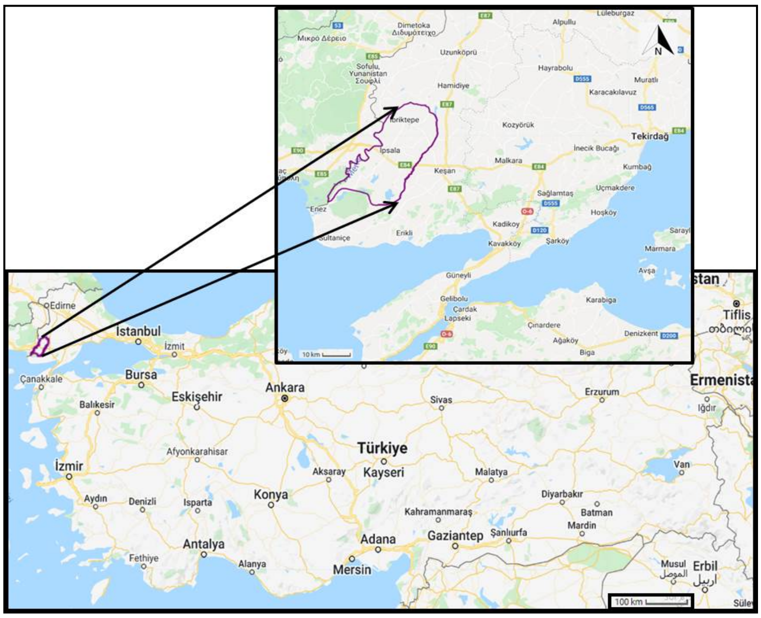

Figure 1. Study area.

2.1. Pre-processing of images

To reduce model operation time and noise disturbance, paddy rice images were passed through a number of pre-processing steps. Firstly, each image was labeled as healthy, rice blast, or brown spot (Figure 2). The image size was decreased from 4640 × 3472 to 256 × 256 in order to reduce the size of the total image data. Since the entire paddy pictures are colored, they are compressed to the 0-1 range, and their backgrounds are colored gray in order to gain stability and prevent noise. Later, in order for the CNN model to process the data set more comprehensively, the horizontal and vertical ratio of the images was chosen as 0.2, and the images were inverted.

Healthy Rice Blast Brown Spot

Figure 2. Sample paddy rice images.

2.2. Modelling

The rice diseases detection CNN model was developed in the Google Colaboratory (Colab). The TensorFlow, matplotlib.pyplot, IPython.display, GPU, Sequential, compile, model.fit, and sklearn.metrics libraries were used in classification. The dataset consisting of 1569 images for the proposed CNN model was uploaded to Google Drive and later transferred into Google Colab. The percentages of training, validation, and testing were 70, 20, and 10%, respectively. The reason for choosing these ratios is to increase the learning speed and accuracy score of the model. The epoch number was set to 100 in order to ensure that the model was not over-fitting and appropriately reflected the validation accuracy. The flowchart of the model is given in Figure 3.

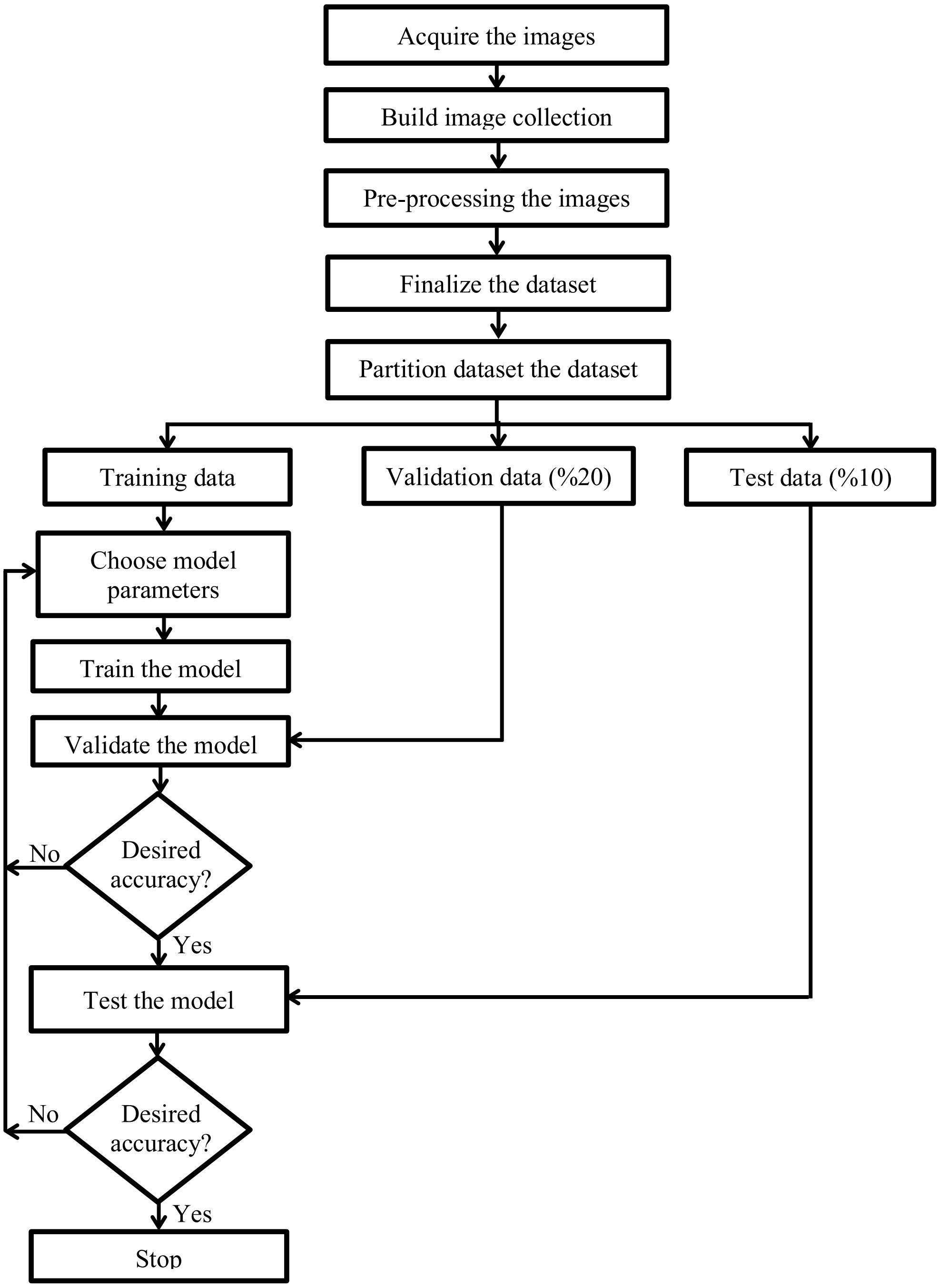

Figure 3. Flowchart of the model.

Values that yield the highest classification accuracy were applied in the selection of model parameters. In CNN, the simplest technique to build a model is sequential. In this technique, the model is built layer-by-layer (Russakovsky et al., 2015). Conv2D was used to add layers to the model. A total of 8 layers was used, including 6 convolutions + rectification layers (Figure 4). The input images were processed by Conv2D layers (Rajaraman et al., 2018). The number of nodes in each layer was determined as 32 and 64. The number can be adjusted to be larger or smaller in accordance with the model's accuracy score by trial and error method (Russakovsky et al., 2015). The ReLU activation function was used as a rectifier to increase the non-linearity of the images. Also, this function only activates a few neurons at a time. So it makes the model sparse and easy to perform (Siddiqi, 2019). In the modeling 2×2 maximum pooling method was applied. Pooling increases the NN's flexibility to recognize features that belong to the same class but are rotated, distorted, squeezed, etc. It also gets rid of the majority of the information that is not related to the feature that should be recognized. However, it still preserves the textural or locational information of the feature even if they are different in various locations. By applying pooling, the size of the image can be reduced about %75. In max-pooling, the maximum value is selected from all available values in the window, thereby incorporating only the highest feature of this pixel (Scherer el al., 2010). The softmax function compresses the output of each class between 0 and 1. Thus, it provides the probability that an input belongs to a particular class. After sequential modeling, two dense layers were added to the end of the CNN (Figure 4). Dense layers extract characteristics from the convolution and pooling layers (Siddiqi, 2019). The second dense layer generated the classification. Given that the model was designed to solve a classification problem, Sparse Categorical was selected entropy as the loss function, and adam was selected as the optimizer (Anonymous, 2022).

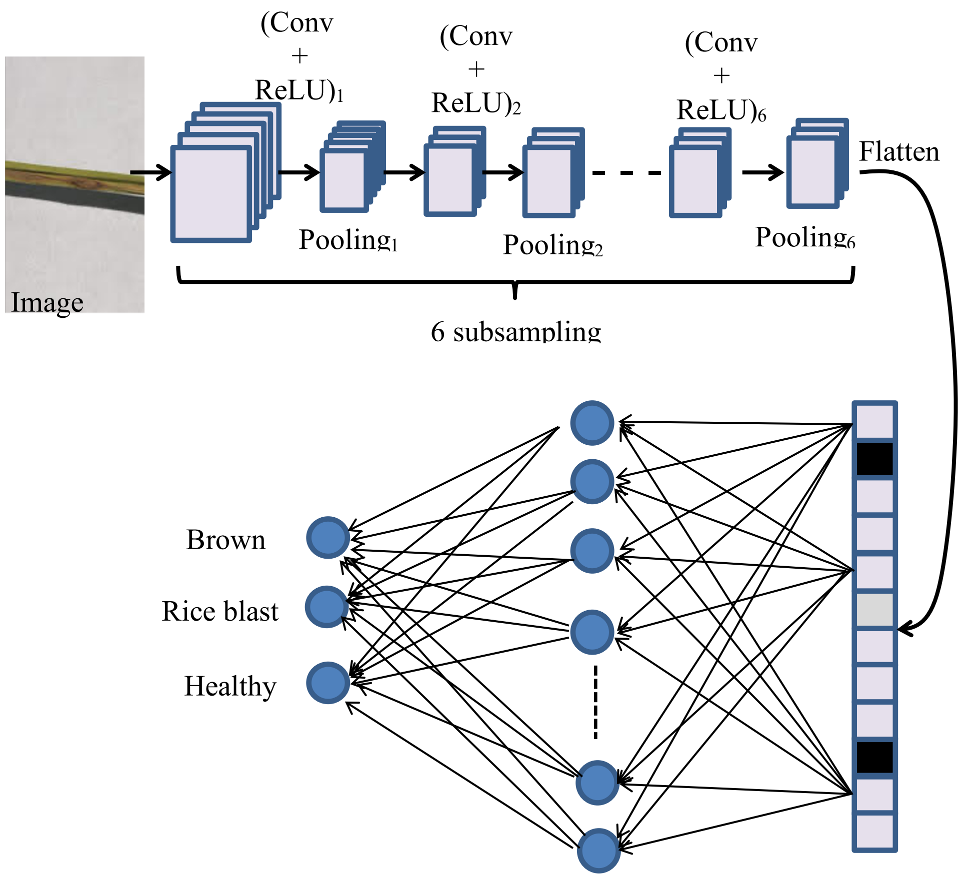

Figure 4. CNN model structure.

3. Results and Discussions

In this study, 70/20/10 percentages of training, testing, and validation were applied for the best results in generalizing the model. A total of 1098 images were used for the training process. The model was validated after the completion of the training phase. Twenty percent of the dataset was randomly selected for validation. The training step value (epoch) for this model 100, batch size 32 was chosen. These numbers can be changed until the highest accuracy is achieved. The change in model accuracy and loss factor as the epochs increase in training and validation steps are plotted in Figure 5. As expected, the accuracy increases and stabilizes after the 90th epoch while the loss factor gains the lowest values for each step. A higher accuracy of 92% was achieved in the training and validation steps.

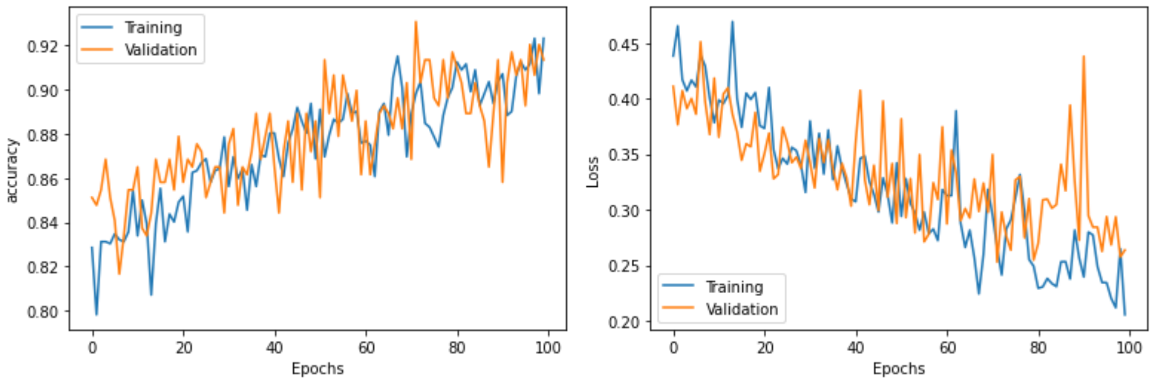

Figure 5. Relationship among epoch, accuracy, and loss factor.

One of the methods used to determine the classification performance of each algorithm is the creation of a confusion matrix. The matrix gives information about how often an observation belonging to a certain class is detected correctly and how often it is determined as another class (Ruuska et al., 2018). The confusion matrix of the model is given in Figure 6.

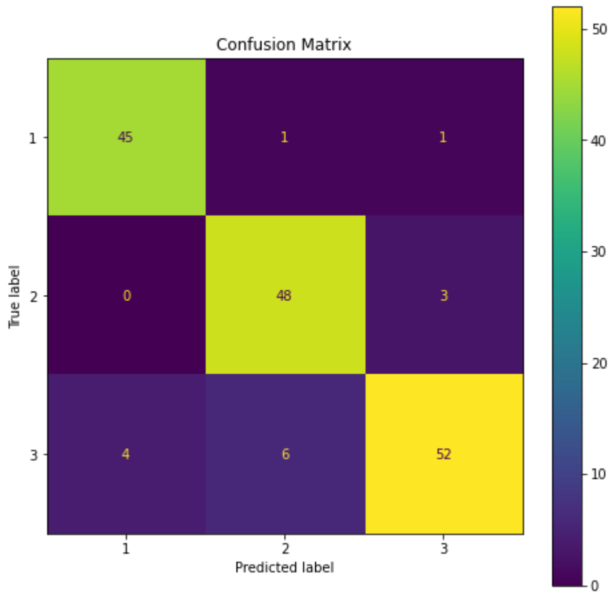

Figure 6. Confusion matrix of the model.

In the creation of the confusion matrix, the test data, which corresponds to 10% of the database, was used. A total of 160 images was evaluated, including 47, 51, and 62 rice blasts, brown spots and healthy, respectively. Based on the data obtained from the confusion matrix, 45 of 47 rice blasts, 48 of 51 brown spots and 52 of 62 healthy images were correctly classified. Classification accuracies of each class are also summarized in Table 1. Higher classification accuracies were obtained for two diseases (>0.94). The overall classification performance of the model on test data was also calculated to be 0.91.

Table 1. Summary of model accuracies

Class	Accuracy
Rice blast	0.96
Brown spot	0.94
Healthy	0.84

In Table 2, various studies with the CNN method on the detection of rice diseases are given. The overall accuracies of the studies given in Table 2 are compatible with this study. Vanitha 2019 found the highest model result and used 350 datasets. The reason why the accuracy rate of the proposed study is lower than other studies is that the number of an epoch is limited to 100, and the dataset consists of 1569 images. The accuracy rate of the model can be increased by using a higher epoch number and a less comprehensive dataset, but this may cause overfitting in the model. This problem lowers the predictive score of the model in the test dataset.

Table 2. Comparative analysis of CNN techniques for rice diseases

Author	Dataset	Diseases	Overall accuracy
Lu et al., 2017	500	Rice Blast and Brown Spot	0.95
Shrivastava et al., 2019	619	Rice Blast, Bacterial Leaf Blight, and Sheat Bligh	0.91
Vanitha, 2019	350	Brown Spot, Sheath Rot, and Bacterial Blight	0.99
Anadhan and Singh, 2021	Over 1500	Rice Blast, Bacterial Leaf Blight, Sheat Blight, Leaf Streak	0.87
Tawde et al., 2021	Not specified	Rice blast, Rice blight, Brown spots, leaf smut, tungro, and sheath blight	0.96

Conclusion

Deep CNN has been used in various applications in agriculture and other disciplines. In this study, it was aimed to re-evaluate the classification performance of a 8-layer CNN model in major paddy rice diseases in the İpsala district of Edirne province in Türkiye. The local cultivars were used as the study material. The rice blast and brown spot fungal diseases were identified using RGB images collected from paddy rice fields in the district. The results showed that the CNN algorithm could be successfully applied to local paddy rice cultivars. Developing of a database for training, validating, and testing the model is relatively easy with use of current technologies. Also, cloud-based modeling environments such as Google Colab provide useful, fast and effective tools in model development. Such work can be conducted without downloading larger sophisticated software on the computer free of charge. However, there are still challenges in the use of deep learning algorithms such as CNN. The main challenge is the decision and optimization of model parameters, structure (number of layers, etc.), epoch numbers, etc. The use of the trial error method is often time consuming and requires former experience in model building. Our future work includes expanding the number of diseases observed in the region affecting the local cultivars that can be identified by CNN. Even though classification using a CNN model stored in a could system performs well, it is not yet practical. A tool should be developed that doesn't require the use of the Google Colab environment. Therefore, in the next step, a user-friendly mobile smartphone application will be developed that can be used by the local farmers. It is also planned to use aerial imagery acquired by a drone to identify such diseases on field scale.

References

Affonso, C., Rossi, A. L. D., Vieira, F. H. A., de Carvalho, & de Leon Ferreira de Carvalho, A. C. P. (2017). Deep learning for biological image classification. *Expert Systems with Applications*, 85, 114–122. https://doi.org/10.1016/j.eswa.2017.05.039

Anadhan, K., & Singh, A. S. (2021). *Detection of paddy crops diseases and early diagnosis using faster regional convolutional neural networks*. Paper presented at the International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 898-902, March 4-5, India.

Anonymous, (2022). *Multi-hot sparse categorical cross-entropy*. https://cwiki.apache.org/confluence/display/MXNET/Multi-hot+Sparse+Categorical+Cross-entropy. Access date: 06:06:2022.

Arnal Barbedo, J. G. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. *Springer Plus*, 2, 660. https://doi.org/10.1186/2193-1801-2-660

Asfarian, A., Herdiyeni, Y., Rauf, A., & Mutaqin, K.H. (2013). *Paddy diseases identification with texture analysis using fractal descriptors based on fourier spectrum*. Paper presented at International Conference on Computer, Control, Informatics and Its Applications (IC3INA), 77-81, November 19-21, Indonesia.

Boulent, J., Foucher, S., Théau, J., & St-Charles, P. L. (2019). Convolutional neural networks for the automatic identification of plant diseases. *Frontiers in Plant Science*, 10, 941. https://doi.org/10.3389/fpls.2019.00941

Devi, T.G., & Neelamegam, P. (2018). Image processing based rice plant leaves diseases in Thanjavur Tamilnadu. *Cluster Computing*, 22, 13415- 13428. https://doi.org/10.1007/s10586-018-1949-x

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, 311–318.

Gokulnath, B. V., & Usha, D. G. (2021). Identifying and classifying plant disease using resilient LF-CNN. *Ecological Informatics*, 63,1, 101283. https://doi.org/10.1016/j.ecoinf.2021.101283

IRRI. (2012). Rice facts. International Rice Research Institute. Manila, Philippines.

Kadhim, M. A., & Abed, M. H. (2019). Convolutional neural network for satellite image classification. *Studies in Computational Intelligence*, 165–178. https://doi.org/10.1007/978-3-030-14132-5_13

Kanani, P., & Padole, M. (2019). Deep learning to detect skin cancer using google colab. *International Journal of Engineering and Advanced Technology*, 8,6, 2176-2183. doi:10.35940/ijeat.F8587.088619

Kawasaki, Y., Uga, H., Kagiwada, S., & Iyatomi, H. (2015). *Basic study of automated diagnosis of viral plant diseases using convolutional neural networks*. Lecture Notes in Computer Science, 638-645. https://doi.org/10.1007/978-3-319-27863-6_59

Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*, 267, 378-384. https://doi.org/10.1016/j.neucom.2017.06.023

Mique, E. L., & Palaoag, T. D. (2018). *Rice pest and disease detection using convolutional neural network.* Paper presented at Proceedings of the 2018 International Conference on Information Science and System - ICISS '18, 147–151, April 27-29, Republic of Korea. https://doi.org/10.1145/3209914.3209945

Pinki, F.T., Khatun, N., & Islam, S.M.M. (2017). *Content based paddy leaf disease recognition and remedy prediction using support vector machine*. Paper presented at 20th International Conference of Computer and Information Technology (ICCIT), 1-5, December 22-24, Bangladesh. doi: 10.1109/ICCITECHN.2017.8281764

Priyadharshini, R. A., Arivazhagan, S., Arun, M., & Mirnalini, A. (2019). Maize leaf disease classification using deep convolutional neural networks. *Neural Computing and Applications*, 31, 8887-8895. https://doi.org/10.1007/s00521-019-04228-3

Rajaraman, S., Candemir, S., Kim, I., Thoma, G., & Antani, S. (2018). Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Applied Sciences*, 8,10, 1715. https://doi.org/10.3390/app8101715

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: a comprehensive review. *Neural Computation*, 29,9, 2352–2449. https://doi.org/10.1162/neco_a_00990

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115,3, 211–252. https://doi.org/10.48550/arXiv.1409.0575

Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., & Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural Processes*, 148, 56-62. https://doi.org/10.1016/j.beproc.2018.01.004

Tawde, T., Verekar, L., Aswale, S., Deshmukh, K., Reddy, A., & Shetgaonkar, P. (2021). Rice plant disease detection and classification techniques: a survey. *International Journal of Engineering Research & Technology*, 10,7, 560-567.

Scherer, D., Muller, A., & Behnke, S. (2010). *Evaluation of pooling operations in convolutional architectures for object recognition*. Paper presented at In Proceedings of the 20th International Conference on Artificial Neural Networks, 92-101, September 15–18, Thessaloniki, Greece. https://doi.org/10.1007/978-3-642-15825-4_10

Shrivastava, V. K., Pradhan, M. K., Minz, S., & Thakur, M. P. (2019). *Rice plant disease classification using transfer learning of deep convolution neural network*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W6, 631-635. DOI:10.5194/isprs-archives-XLII-3-W6-631-2019

Siddiqi, R. (2019). *Automated pneumonia diagnosis using a customized sequential convolutional neural network*. Paper presented at Proceedings of the 3rd International Conference on Deep Learning Technologies, 64-70, July 5-7, Xiamen, China. https://doi.org/10.1145/3342999.3343001

Vanitha, V. (2019). Rice disease detection using deep learning. *International Journal of Recent Technology and Engineering*, 7, 534-542.

Vardhini, P. A. H., Asritha, S., & Devi, Y. S. (2020). *Efficient disease detection of paddy crop using CNN*. Paper presented at 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), 116-119, October 9-10, Bengaluru, India. doi: 10.1109/ICSTCEE49637.2020.9276775

Wang, P., Fan, E., & Wang, P. (2020). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*, 141(11): 61-67. doi:10.1016/j.patrec.2020.07.042

Xiao, M., Ma, Y., Feng, Z., Deng, Z., Hou, S., Shu, L., & Lu, Z. (2018). Rice blast recognition based on principal component analysis and neural network. *Computers and Electronics in Agriculture*, 154, 482–490. https://doi.org/10.1016/j.compag.2018.08.028

Zhang, J., Xie, Y., Wu, Q., & Xia, Y. (2019). Medical image classification using synergic deep learning. *Medical Image Analysis*, 54, 10-19. doi: 10.1016/j.media.2019.02.010