Rice Leaf Disease Classification Using Deep Transfer Learning Convolutional Neural Network: MobileNet + Bidirectional GRU

Sarah Yim Seoul International School, South Korea sarahyim2023@gmail.com

Abstract

Agriculture, the supporting backbone to the economic growth of many countries, has played a significant global role in the economy. With rapid technological advancements in various fields today, it is important to invest time and effort into developing advanced methods to preserve agricultural practices. However, the lack of technology and expertise in crop disease identification is a notorious problem troubling the agricultural industry, especially in developing countries. This often leads to severe crop loss and waste, affecting not only farmers' yield but also consumers' food intake. Rice leaf disease identification, in particular, rises as an important issue as rice is a staple food for a large proportion of the global population. Specifically, timely and accurate diagnosis of rice leaf diseases is crucial. To address this issue, this paper implements an image-based deep learning approach to identify and classify rice leaf diseases presented in a dataset derived from a rice field in Sherta located in Gujarat, India. This dataset contains 120 images belonging to three distinct classes: Brown Spot, Leaf Smut, and Bacterial. Evaluation performances followed by statistical analysis are conducted using eight different convolutional neural network (CNN) models: Inception V3, MobileNet, DenseNet121, Vgg16, Vgg19, ResNet101, NASNetMobile, and MobileNet+Bi-GRU. The best performing model was the MobileNet+Bi-GRU model with an accuracy score of 87.24%. The experimental results from the performance evaluations revealed great potential in incorporating deep learning techniques for rice leaf disease identification and classification.

Keywords: agriculture, rice leaf disease identification, deep learning

Introduction

Background:

Rice, a starchy cereal grain and grass species Oryza Sativa, is a staple food for roughly half of the world's population. With more than 90% of rice grown in Asia, practically all of East and Southeast Asia are dependent on rice as a staple food. Rice yields vary with different conditions, ranging from 700 to 4,000 kilograms per hectare (Encyclopedia, 2021). India, a global agricultural powerhouse with leading production in milk, wheat, and rice (India Brand Equity Foundation, 2021), ranks second highest in rice production after China, producing about 110 million metric tons of rice per year (Vukotić et al., 2016). Agriculture is a livelihood for approximately 58% of India's population.

With the growing world population, global rice demand continues to increase; with a global rice demand of 439 million tons in 2010, 496 million tons in 2020, to a predicted 555 million tons in 2035, projections speculate a demand increase of 26% in the next 25 years. Hence, it is highly likely the 150 million hectares of rice fields

available around the world will not suffice for the rice demands in the next few years. Because land is scarce and expansion is unlikely, global rice yields must increase at faster rates of at least 1.2-1.5% over the next decade, equivalent to 8-10 million tons more paddy rices each year (Ricepedia, n.d.). However, this is extremely difficult due to the frequent damages that are overlooked in paddy fields from various diseases, pesticides, and poor harvest management. Every year, farmers lose approximately 37% of their rice crops from diseases and pests (Rice Research, n.d.). In the world's most agricultural regions and developing nations, unsustainable and deficient rice management results in decreasing merchantability and rice yield, thereby increasing malnutrition and poverty. To make matters worse, many farmers lack expertise in agricultural management, hence cannot detect or identify the potentially harmful diseases in their rice crops.

In response to the exacerbating dilemma of rice diseases, farmers today utilize digital agriculture such as Artificial Intelligence (AI), Satellite Imagery, and Machine Learning, and other advanced analytics resulting in higher crop yield

and more efficient crop management. In particular, in collaboration with Microsoft, farmers in India have implemented artificial intelligence and machine learning based sowing advisories that send automated voice calls that inform farmers on the optimal date to sow and alerts prior to pest infestation. Another innovation is the agricultural multivariate commodity price forecasting model to predict crop yields, prices, and commodity arrivals at every farming stage. This model uses input data from remote sensing, geo-stationary satellite images, weather, and other datasets to predict measurements with accuracy (Microsoft India, 2017).

Artificial intelligence and machine learning have only recently marked their potential in agriculture. Such forms of digital agriculture have great prospects in providing stability for agricultural communities, especially in agriculture-dependent developing nations such as India. Additionally, with far-reaching impacts of climate change on a global scale today, digital agriculture is more imperative than ever before. The graph below shows the total rice consumption worldwide from 2008/2009 to 2020/2021 (Shahbandeh, 2021).



Objective

The objective of this research is to complement rice productivity through early prediction of rice leaf diseases such as bacterial blight, leaf smut, and brown spot diseases. With early alertness, farmers could take precautions to manage the spread of diseases to other crops, saving several tons of rice from going to waste and increasing total rice yield and merchantability. To do so, deep learning frameworks such as Inception V3, Vgg19, MobileNet, DenseNet121, Vgg16, ResNet101, NASNetMobile, and MobileNet+Bi-GRU were employed to detect potential rice diseases in rice crops. Next, a comparative study analysis of the performance of each algorithm was conducted. Lastly, utilizing Gradient-weighted Class Activation Mapping (Grad-CAM) and the heatmap feature, images could be processed and visualized, identifying notable patterns. This would overcome the limitations of deep learning's black box model, thereby increasing the study's applicability in agriculture.



FIGURE 2: Visualization of the given image

Related Works

In this section, previous work on incorporation of deep learning and neural networks on rice disease identification and classification, along with the accuracy scores for each algorithm, are discussed.

In literature, numerous research papers have been published for rice disease identification and classification such as using support vector machines, automated feature engineering, comparisons of various deep-learning models, and transfer learning of deep convolutional neural networks.

Firstly, Sethy et al. carried out performance evaluations of 13 number CNN models in transfer learning and deep feature plus support vector

machine (SVM) approach. Sethy et al. classified four types of rice-leaf diseases, namely bacterial blight, blast, brown spot and tungro. Because the feature of fc6 had the most significant correlation towards classification compared to fc7 and fc8 features of AlexNet, vgg16, and vgg19, only fc6 was considered when choosing the best classification model. Next, through statistical analysis of accuracy, sensitivity, specificity, FPR, F1 score, and training time, Sethy et al. concluded that the ResNet50 plus SVM model was the best classification model in the deep feature approach, while in the transfer learning approach, no statistical difference among the CNN models were exhibited. Among the small CNN models and comparable to the ResNet50 plus SVM model was the deep feature of mobilenetv2 plus SVM (Sethy et al., 2020).

Das et al. developed a deep-learning based automated feature engineering for early rice leaf disease prediction for diseases such as leaf blast, brown spot, and bacterial leaf blight. After images portraying specified portions of various rice leaf diseases were identified from a dataset of 10.500 infected leaves, they were fed into the convolution neural network (CNN) model consisting of four convolution layers, two fully connected layers, and one softmax output layer. Performance evaluations were conducted to assess the effectiveness of each classifier for rice leaf disease prediction. It was observed that the CNN, NB, and LR showed relatively better performance than other classifiers, having accuracy scores of 91.07%, 92,16%, and 90.35% respectively. Next, Das et al. compared the CNN classifier with other rice disease classifiers from previous works, depicting how the CNN method nonetheless had superior performance than all the other classifiers (Das et al., 2020).

Burhan et al. performed five different deep learning models (Vgg16, Vgg19, ResNet50, ResNet50V2, and ResNet101V2) using an artificial data set classified into four classes, namely Hispa, Healthy, Brown Spot, and LeafBlast, and a binary classification of Healthy Vs. Unhealthy using a dataset of images from rice fields in Pakistan. From the experiment using the artificial data set, the ResNet50, demonstrating an accuracy score of 75.0, showed the best performance. From the experiment using the real data set, ResNet101V2, demonstrating an accuracy score of 86.799, showed the best performance. However, limitations such as having a limited number of images in the datasets and visible shadows in many of the images hindered the efficiency and reliability of the results (Burhan et al., 2020).

Shrivastava et al. utilized a pre-trained deep convolutional neural network (CNN) and support vector machine (SVM) as their feature extractor and classifier, respectively. Experiments were conducted with varying ratios of training-testing sets, and the training-testing division of 80%-20% demonstrated the highest classification accuracy of 91.37% (Shrivastava et al., 2019).

Lu et al. proposed a novel rice disease identification disease method based on convolutional neural network (CNN) techniques. CNNs were trained to identify 10 rice diseases using 500 images portraying healthy and diseased rice leaves/stems. The proposed model had an accuracy score of 95.48%, much higher than that of conventional machine learning models such as the BP, SVM, and particle swarm optimization (PSO) methods (Lu et al., 2017).

Methods

Data Description

This rice leaf diseases dataset, obtained from the Shertha locality in Gujarat, India and uploaded on Kaggle, contains images of disease-infected rice leaves, each belonging to one of three classes i.e. Brown Spot, Leaf Smut, and Bacterial Leaf Blight. Bacterial blight, caused by Xanthomonas oryzae, is one of the most serious diseases in rice. The bacteria spreads through ooze droplets on lesions of infected plants and normally causes wilting and yellowing of leaves. The earlier the disease occurs, the higher the yield loss (Rice Knowledge Bank, n.d.). Leaf smut, a fungal disease caused by fungus *Entyloma oryzae*, is a widely distributed rice disease that produces angular, black spots on both sides of the leaves. The fungus is spread by airborne spores and over winters in soil, particularly in diseased leaf debris. Brown spot is a fungal disease caused by fungus Cochliobolus miyabeanus that infects coleoptiles, panicle branches, glumes and grains. It results in brown, circular spots on coleoptile leaves of seedlings, indicating plant inability to use nitrogen and weakened plants (Groth & Hollier, n.d.).

With a total of 120 rice leaf disease images in the dataset, 40 images account for each class. The leaves in these images are derived using a digital camera and empirical evaluation for background removal and segmentation (Prajapati et al., 2017). This was a very convenient and accessible dataset, for all images had a uniform, clean background with one leaf in each image.

GRU

GRU was proposed by Cho et al. in 2014 and its architecture is similar to long short term memory(LSTM). LSTM was developed to solve the existing problem of recurrent neural network(RNN). RNN suffers from a vanishing gradient problem; first received information has a strong influence on learning and then gradually diminishes, eventually failing to influence learning. Cell state from the LSTM prevents that drawback by storing previous steps of information in a memory cell and sending it out. The cell state utilizes three gates : input gate, forget gate and output gate to determine whether the information is reflected. The GRU consists of two gates : reset gate and update gate. The reset gate determines whether to combine the new input with previous memory and the update gate determines the amount of the memory to remember. The overall computational process of the GRU is carried out with the following formulas (Cho et al., 2014):

rt=σ(Wrht-1,xt+br)	(1)
ht=tanh(Whrtht-1 ,xt+bh	(2)
zt=o(Wzht-1,xt+bz)	(3)
ht-1=(1-zt)⊙ht-1+ztht	(4)



FIGURE 3: Overall architecture of GRU.

Convolutional Neural Network

Deep neural networks have a limitation that if it gets image data as input, it should flatten the input image. It means transferring the three dimensions into one dimension. The input layer of the convolutional neural network (CNN) gets the image in three dimensions, and the shape of the output is also three dimensional data. Then, the output is passed to the next layer. CNN consists of three particular layers, which are convolution layer, pooling layer, and fully connected layer. Convolution layer, applies a filter to the input image, and the filter, which is also known as kernel, passes through the input image and then makes the feature map as output. In technical terms, pooling reduces the amount of data processed in turn. This process can significantly reduce the total number of parameters in the model. The pooling includes Max-Pooling and Average-Pooling, which is the method of finding the maximum value in the area, and Average-Pooling is the method of calculating the average value of the area. Fully connected layer is also known as a deep neural network and its role is to classify the labels. For binary classification, sigmoid function is used for the activation function and for the multi label classification, soft max is utilized (Albawi et al., 2017).



FIGURE 4: Overall architecture of CNN

MobileNet

MobileNet is a pretrained network, which is trained by the large dataset named imagenet. MobileNet utilized the Depthwise separable convolution to make the model lighter. The reason why MobileNet focused on lightning is to apply the models low deep learning to memory environments such as mobile phones or embedded systems. MobileNet utilizes an outstanding convolution method, which is depthwise separable convolution. Depthwise convolution produces a feature map for each input channel by performing one 3x3 conv filter operation. Pointwise convolution adjusts the number of channels generated by Depthwise convolution to 1x1conv. Depthwise separable convolution is the application of pointwise convolution after depthwise convolution. Depthwise separable convolution considers both spatial feature and channel-wise feature to lighten the model (Howard et al., 2017).

Proposed Model

Our proposed model consists of MobileNet and bidirectional GRU(Bi-GRU). MobileNet draws out the features from the input image and passes the output to the Bi-GRU. We utilize the Bi-GRU instead of the fully connected network because prior research had shown that combining CNN and other deep learning networks such as long short term memory (LSTM), recurrent neural network (RNN), and GRU showed better classification performance (Gu et al., 2018, p. 774-784). Furthermore, we implemented the bidirectional GRU to train the whole parameters more efficiently.

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224\times224\times3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112\times112\times32$
Conv / s1	$1\times1\times32\times64$	$112\times112\times32$
Conv dw / s2	$3 \times 3 \times 64 \text{ dw}$	$112\times112\times64$
Conv / s1	$1\times1\times64\times128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128 \text{ dw}$	$56\times 56\times 128$
Conv / s1	$1\times1\times128\times128$	$56\times 56\times 128$
Conv dw / s2	$3 \times 3 \times 128 \text{ dw}$	$56\times 56\times 128$
Conv / s1	$1\times1\times128\times256$	$28\times 28\times 128$
Conv dw / s1	$3 \times 3 \times 256 \text{ dw}$	$28\times28\times256$
Conv / s1	$1\times1\times256\times256$	$28\times28\times256$
Conv dw / s2	$3 \times 3 \times 256 \text{ dw}$	$28 \times 28 \times 256$
Conv / s1	$1\times1\times256\times512$	$14\times14\times256$
5 Conv dw / s1	$3 \times 3 \times 512 \text{ dw}$	$14\times14\times512$
^o × Conv / s1	$1\times1\times512\times512$	$14\times14\times512$
Conv dw / s2	$3 \times 3 \times 512 \text{ dw}$	$14\times14\times512$
Conv / s1	$1\times1\times512\times1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024 \text{ dw}$	$7 \times 7 \times 1024$
Conv / s1	$1\times1\times1024\times1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1\times1\times1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

FIGURE 5: Summary of the MobileNet architecture

Results

The deep-learning software library Keras was used to calculate and evaluate the performances of the models. Using Google Colab along with GPU as the hardware accelerator, the rice leaf diseases dataset was imported onto Google Drive onto Google and mounted Colab. The performance of each algorithm for rice leaf disease classification was measured and evaluated in terms of accuracy, and a comparative study analysis of the performance was conducted. Figure 7 shows the average accuracy scores of different performing models. The best performing model was MobileNet+Bi-GRU with an accuracy score of 87.24%. The poorest performing model was ResNet101 with an accuracy score of 33.33%.

The Figure 8 graphs display that initially, although the accuracy was relatively low and loss was relatively high, upon increasing the number of epochs, the loss notably decreased while the accuracy increased. The greater the accuracy and smaller the loss, the better the classifier models are at modeling the correlations between the inputs and output targets with fewer errors.



FIGURE 7: Comparison of performance scores for each algorithm

RGB colorized heatmaps were created for the three rice leaf disease images, i.e. Brown Spot, Leaf Smut, and Bacterial, to colorize the intensity of the heatmap and identify the area in each leaf in which the disease was most dispersed. From a color scale of purple to red exhibiting lower to higher intensity of disease dispersal, performing heatmaps allowed for a more specific and accurate identification of the leaf's location in which each disease was found. The heatmap for the Brown Spot and Bacterial rice leaf disease image indicated that the inner center of the rice leaf was most infected with the disease, whereas the heatmap for the Leaf Smut rice leaf disease image indicated that the top center half was most infected with the disease. Areas depicted with lower color intensity level such as blue or purple indicated areas of less infection by each disease.



FIGURE 8: Accuracy and loss graph for training and validation set

Discussions

Principle Finding

After evaluating the performances (accuracy scores) of the eight different CNN models, i.e. MobileNet+Bi-GRU, Inception V3, Vgg16, Vgg19, MobileNet. DenseNet121, ResNet101, and NASNetMobile, to determine the best performing model for identification of rice leaf diseases, it was revealed that performances from best to worst MobileNet+Bi-GRU. Inception V3. were MobileNet and NasNetMobile, DenseNet, Vgg19, Vgg16, ResNet101, in that order. Overall, among all the CNN models. MobileNet+Bi-GRU demonstrated the best performance with the highest accuracy score of 87.24%.

Our accuracy score of the best performing model (MobileNet+Bi-GRU), 87.24%, is relatively high compared to other related works. For example, when Burhan et al. performed five different deep learning models for multiclass rice leaf disease classification, the accuracy score of his best performing model (ResNet50) was 75.0% (Burhan et al., 2020). This may be because the images in our dataset consisted of more clear, unshadowed images contributing to higher efficiency and reliability of the results, compared to many other related works utilizing indistinct, shadowed rice leaf images.



FIGURE 9: Heatmap of Brown Spot rice leaf disease image



FIGURE 10: Heatmap of Leaf Smut rice leaf disease image



FIGURE 11: Heatmap of Bacterial rice leaf disease image

Limitations

One limiting factor in this Indian dataset was its small size. This dataset consisted of 120 images of disease infected rice leaves, with only 40 images in each class. With more images to represent each class in the dataset, the model could extract each class' features more precisely, possibly increasing the accuracy scores for the performing models. Another limitation of the dataset was the fact that out of the many more rice leaf diseases found in rice crops, only three types of rice leaf diseases (brown spot, leaf smut, and bacterial) were identified. Although these three may have been the most prominent ones found in the area in which this dataset was collected, the Shertha rice field in Gujarat, India, many more types of rice leaf diseases such as eyespot and sheath rot could be included in the dataset for broader application.

In this study, the accuracy scores did not go over the score of 90(%), possibly because of the small size of the dataset or the limited number of rice leaf diseases identified in the dataset. This limited the ability of the models to classify the rice leaf disease images into their respective classes with accuracy and precision.

Finally, only image classification was performed with the images in this dataset, limiting us to classification of what was contained in each image. An improvement would be to perform object detection as the next step, helping specify the location and identify the object in each image more thoroughly and quickly. This would especially be convenient when using a mobile application, for images uploaded into the application could be processed and checked in real time for fast and efficient classification of rice leaf diseases for farmers.

Conclusions

In this paper, deep learning techniques were utilized to classify rice leaf diseases. Eight distinct CNN models were applied to the rice leaf diseases dataset from Shertha in Gujarat, India. The experiments were conducted by portioning the dataset into a training-testing ratio of 70%-30%. The best proposed model, MobileNet+Bi-GRU, was able to classify rice leaf diseases with a classification accuracy score of 87.24%, whereas the worst performing model, ResNet101, was able to classify rice leaf diseases with a classification accuracy score of 33.33%. With a larger dataset constituted of various types of rice leaf diseases globally found, MobileNet+Bi-GRU with a 70%-30% training-testing ratio could be a

remarkably resourceful model to deploy for rice leaf disease classification in rice fields.

With the statistical analysis of the performances conducted in this study, we can conclude that applying deep learning algorithms to rice leaf disease classification and digital agriculture as a whole is a potential solution to decrease rice loss and increase global rice productivity to meet the increasing global rice demand. This approach can also be further incorporated into a mobile application for farmers, especially in developing countries with poor crop management technology. Deep neural networks are often criticized for being a black box model which makes the process between the input and output untraceable, thereby limiting the practicality when applied to applications. However, this study overcame that limitation by utilizing Gradientweighted Class Activation Mapping (Grad-CAM) to better understand the model's predictions and facilitate the applicability to agriculture. The application could scan images of rice crops and not only detect diseases that may have infected the crops but also identify the specific area of infection, establishing advanced agricultural technology. With digital agriculture in its nascent stage today, it is crucial that more researchers implement artificial intelligence to more effectively address problems in the agricultural industry.

References

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. International Conference on Engineering and Technology (ICET). Published. https://doi.org/10.1109/ICEngTechnol.2017.8308186

Burhan, S. A., Minhas, S., Tariq, A., & Hassan, M. N. (2020). Comparative Study Of Deep Learning Algorithms For Disease And Pest Detection In Rice Crops. International Conference on Electronics, Computers and Artificial Intelligence (ECAI). Published. https://doi.org/10.1109/ECAI50035.2020.9223239

Chung, J., Gulcehre, C., Cho, K. H., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. Neural and Evolutionary Computing. Published. https://arxiv.org/abs/1412.3555

Das, A., Mallick, C., & Dutta, S. (2020). Deep Learning-Based Automated Feature Engineering for Rice Leaf Disease Prediction. Computational Intelligence in Pattern Recognition. Published. https://doi.org/10.1007/978-981-15-2449-3_11 Groth, Don, & Hollier, Clayton. (n.d.). Brown Spot of Rice. LSUAgCenter, https://www.lsuagcenter.com/~/media/system/4/6/2/f/462fd54617eaff51c 0fdeb50d5feccdd/pub3108brownspothighres.pdf

Gu, S., Zhang, L., Hou, Y., & Song, Y. (2018). A Position-aware Bidirectional Attention Network for Aspect-level Sentiment Analysis. Proceedings of the 27th International Conference on Computational Linguistics. Published. https://aclanthology.org/C18-1066.pdf

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. Computer Vision and Pattern Recognition. Published. https://arxiv.org/abs/1704.04861

India Brand Equity Foundation. (2017). Analysis About Indian Agriculture Industry, Market Size, Export & Investment Opportunity. Published. https://www.ibef.org/industry/agriculture-india.aspx

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

Microsoft News Center India. (2017, November 7). Digital Agriculture: Farmers in India are using AI to increase crop yields. Microsoft. https://news.microsoft.com/en-in/features/ai-agriculture-icrisat-upl-india/

Prajapati, H. B., Shah, J. P., & Dabhi, V. K. (2017). Detection and classification of rice plant diseases. Intelligent Decision Technologies. Published. https://doi.org/10.3233/IDT-170301

Rice. (n.d.). Britannica. https://www.britannica.com/plant/rice Rice Knowledge Bank. (n.d.). Bacterial blight. http://www.knowledgebank.irri.org/decision-tools/rice-doctor/rice-doctorfact-sheets/item/bacterial-blight

Rice Research: Open Access. (n.d.). Rice Research: Open Access. https://www.omicsonline.org/scholarly/rice-diseases-journals-articles-ppts-list.php

Ricepedia. (n.d.). Food security. https://ricepedia.org/challenges/food-security

Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Deep feature based rice leaf disease identification using support vector machine. Computers and Electronics in Agriculture, 175. https://doi.org/10.1016/j.compag.2020.105527

Shahbandeh, M. (2021, April 22). Total rice consumption worldwide from 2008/2009 to 2020/2021 (in 1,000 metric tons). Statista. https://www.statista.com/statistics/255977/total-global-rice-consumption/

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. Published. https://doi.org/10.5194/isprs-archives-XLII-3-W6-631-2019

Vukotić, V., & Gravier, G. (2016). A Step Beyond Local Observations with a Dialog Aware Bidirectional GRU Network for Spoken Language Understanding. Interspeech 2016. Published. https://doi.org/10.21437/Interspeech.2016-1301