VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY

UNIVERSITY OF INFORMATION TECHNOLOGY

COMPUTER SCIENCE DEPARTMENT

PROJECT'S REPORT

Computational Thinking – CS117.N21

IMAGE-BASED CLASSIFICATION OF RICE PADDY DISEASES ON IMBALANCED DATASET

LECTURER:

PhD. Ngô Đức Thành

AUTHORS:

Nguyễn Huỳnh Minh Triết – 21520497

Nguyễn Minh Thư – 21520472

Lê Châu Giang – 21520213

Hoàng Minh Hiếu – 21520232

Ngô Phúc Danh – 21521924

Trương Văn Khải – 21520274

Ho Chi Minh City, 07/2023

TABLE OF CONTENTS

Contents

Chapter 1. PROBLEM IDENTIFICATION ... 3

1.1. Problem Introduction ... 3

1.2. Problem Description ... 4

Chapter 2. THE PROCESS OF APPLYING COMPUTATIONAL THINKING ... 6

Chapter 3. OUR SOLUTION ... 12

3.1. Method ... 12

3.1.1. Traditional Image Data Augmentation ... 12

3.1.2. Convolutional Neural Network ... 12

3.2. Evaluation ... 15

3.2.1. Accuracy ... 15

3.2.2. Precision and Recall ... 16

Chapter 4. CONCLUSION ... 17

REFERENCE ... 18

Chapter 1. PROBLEM IDENTIFICATION

1.1. Problem Introduction

Rice paddy, also known as rice (Oryza sativa), is a vital staple crop worldwide, providing sustenance for a significant portion of the global population. However, the production of rice faces numerous challenges, with the prevalence of diseases posing a significant threat to crop yield and quality. Early detection and accurate classification of these diseases are crucial for implementing effective control measures and ensuring food security.

One of the challenges in developing an image-based classification system for rice paddy diseases lies in dealing with imbalanced datasets. Imbalanced datasets occur when certain disease classes have significantly fewer samples compared to others, making it difficult for classification models to learn and accurately classify the minority classes. This issue can lead to biased and unreliable predictions, which hinders the effectiveness of disease management strategies.

Traditionally, disease diagnosis in paddy crops has relied on visual inspection by trained agronomists, which is a time-consuming and subjective process. However, with the advancements in computer vision and machine learning techniques, there has been a growing interest in developing automated systems for paddy disease classification based on image analysis. This approach offers the potential to improve the efficiency and accuracy of disease identification, enabling prompt intervention and targeted management strategies.

The objective of this study is to address the challenges posed by imbalanced datasets in the image-based classification of rice paddy diseases. By leveraging machine learning techniques and advancements in computer vision, we aim to develop a robust and reliable system capable of accurately identifying and classifying different types of diseases that affect paddy crops. Specifically, we will focus on developing strategies to overcome the limitations imposed by imbalanced datasets, ensuring that the classification models can effectively handle all disease classes, regardless of their sample sizes and detect and categorize various diseases, such as bacterial blight, blast, dead heart, and brown spot.

The availability of high-resolution images captured using digital cameras or drones, combined with recent advances in image processing techniques, provides an excellent opportunity to analyze plant health at a detailed level. Leveraging these advancements, our research will focus on extracting relevant features from paddy disease images and training machine learning models to classify them accurately. We will explore various techniques, such as data augmentation, resampling methods, and model adjustments, to address the challenges posed by imbalanced datasets.

The successful implementation of an automated paddy disease classification system on imbalanced datasets would have several benefits. It would enable early detection of diseases, allowing farmers to take immediate action to mitigate the spread of infections and minimize yield losses. Moreover, it would provide valuable insights into disease patterns and distribution, helping researchers and policymakers make informed decisions regarding disease management strategies and crop improvement programs.

In conclusion, this study aims to develop a reliable and efficient system for image-based classification of rice paddy diseases on imbalanced datasets. By leveraging machine learning techniques and advancements in computer vision, we strive to provide a valuable tool for farmers and agricultural stakeholders to accurately identify and classify diseases affecting paddy crops. Ultimately, this research has the potential to contribute to improved disease management practices and increased agricultural productivity in the paddy farming sector, specifically addressing the challenges associated with imbalanced datasets.

1.2. Problem Description

In this problem, we use a dataset called "Paddy Doctor". It provides a training dataset of 10,407 (75%) labeled images across ten classes (nine disease categories and normal leaf). Moreover, they also provide additional metadata for each image, such as the paddy variety and age.

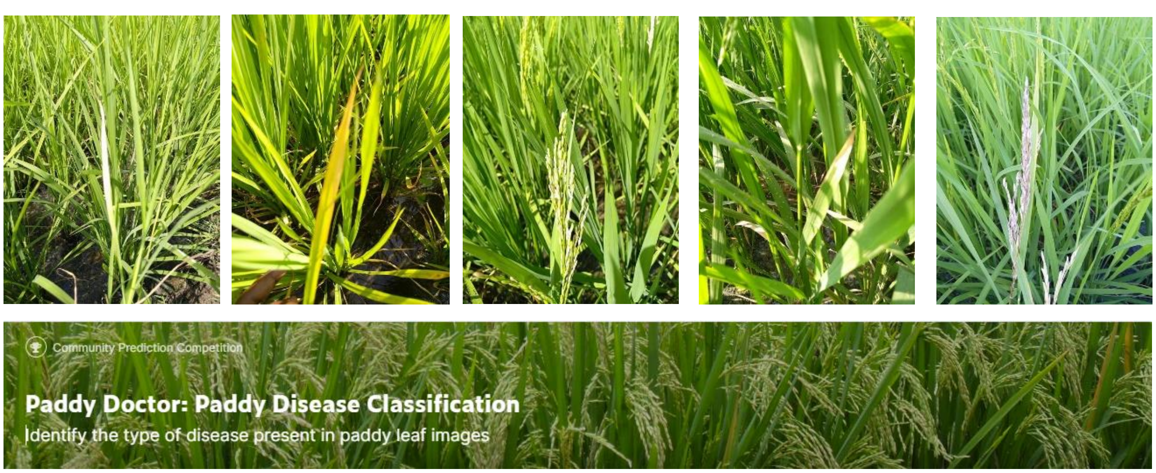


Figure 1. Images of paddy disease by Paddy Doctor from Community Prediction Competition

We will preprocess the input image of the paddy plant and then analyze its characteristics to determine the specific disease affecting it, thereby giving the best assessment. The results will show the type of paddy disease predicted by the model. As follow:

- **Input**: Color photo of a region on a rice plant captured under good lighting conditions, frontal view, and clear.
- **Output**: The input image is accompanied by a red text line in the bottom right corner indicating the type of disease or "**Normal**". For example: **Hispa**, **Blast**, **Tungro**,…

For Example:

Figure 2. Image of a region on a rice plant

Figure 3. The input image accompanied by a red text line

Chapter 2. THE PROCESS OF APPLYING COMPUTATIONAL THINKING

Step 1. We use abstraction skill

In the context of Image-Based Classification of Rice Paddy Diseases on an Imbalanced Dataset, abstraction techniques are utilized to gain a comprehensive understanding of the problem. Abstraction facilitates a clear delineation of the challenges involved in classifying rice paddy diseases, allowing for focused problem-solving approaches.

To address the issue of imbalanced data, data balancing techniques are applied. These techniques aim to mitigate the unequal distribution of disease classes within the dataset, ensuring that the classification model is not biased toward the majority class. By balancing the dataset, the model can learn from a representative sample of each disease class, leading to more accurate and unbiased classification results.

Furthermore, a robust paddy disease classification model is developed using deep learning algorithms, such as convolutional neural networks (CNN). The CNN model leverages its inherent architecture to automatically extract relevant features from the images, which significantly contribute to the accurate identification and classification of various rice paddy diseases.

By combining data balancing techniques with the power of CNN models, this project aims to improve classification accuracy and assist in the effective management of rice paddy diseases. The utilization of abstraction techniques allows for a clear understanding of the problem, while data balancing and CNN-based classification facilitate accurate disease diagnosis, enabling timely interventions and ultimately leading to improved crop yield and quality.

After applying the abstraction skills of computational thinking, we successfully abstracted the complex problem into a simplified representation by utilizing the image presented in Figure 6.

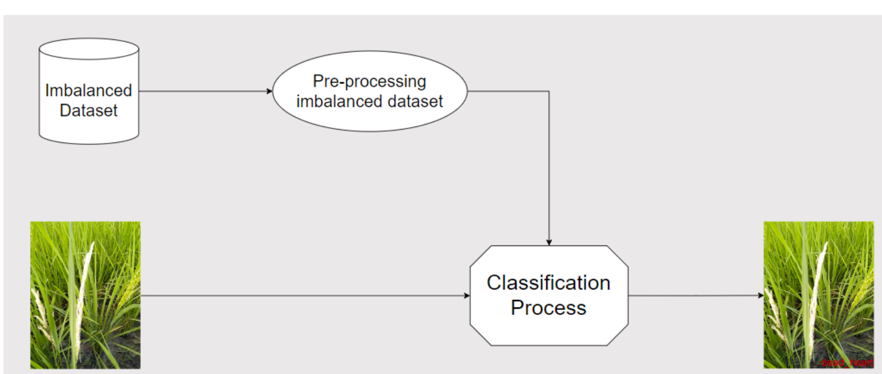


Figure 4. The process flowchart of Image-Based Classification of Rice Paddy Diseases on an Imbalanced Dataset

Step 2. We use Decomposition skill

After utilizing abstraction techniques to abstract and clarify the problem, we proceed with decomposition to break down the problem into smaller components. Specifically, to address the problem, we perform the following subproblems:

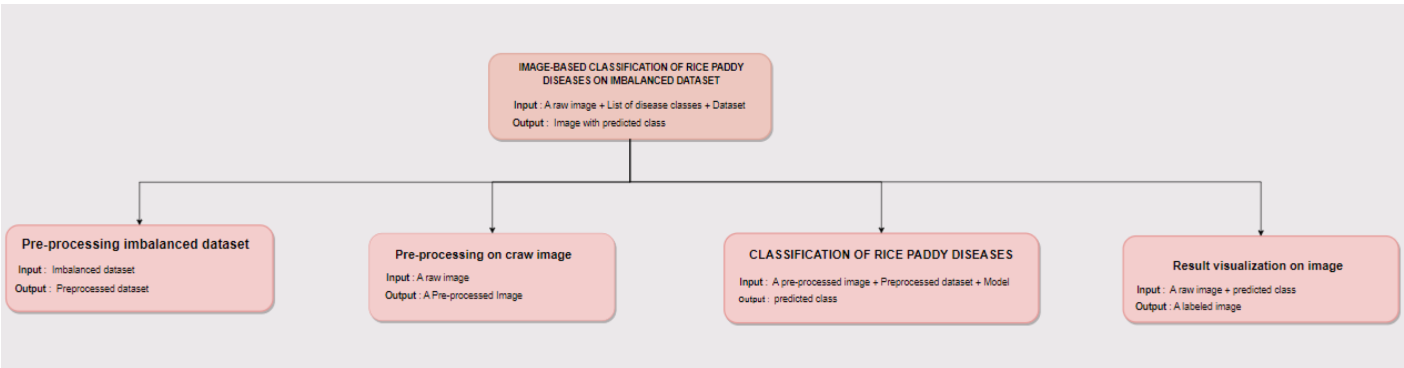


Figure 5. Components after breaking down the problem

Imbalanced dataset pre-processing: Before training the classification model, we need to handle the imbalanced dataset. This involves techniques such as oversampling, undersampling, or utilizing data augmentation algorithms to create a balanced dataset in terms of the number of rice paddy disease classes.

Classification of rice paddy diseases: After data pre-processing, our focus is on developing a disease classification model. This is a crucial step for accurately identifying and classifying different types of rice paddy diseases based on the provided images. We employ machine learning or deep learning algorithms such as convolutional neural networks (CNN) to build a powerful classification model.

Pre-processing on raw images: Enhance the clarity and visibility of plant images to accurately identify disease areas. This involves techniques such as denoising, contrast enhancement, and image segmentation.

Result visualization on images: Display the disease labels on the original input images for clear identification. This involves overlaying text or labels indicating the specific type of disease on the paddy images.

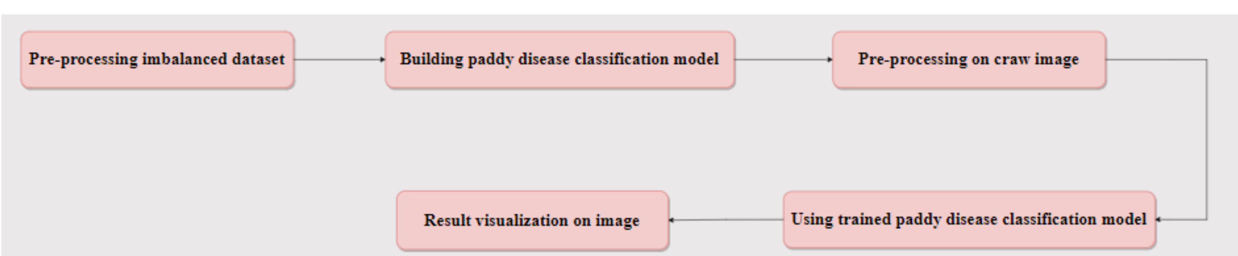


Figure 6. A full activity sequence of rice leaf disease classification

Step 3. We use Pattern Recognition skill (Mapping)

In the subproblem of "Imbalanced dataset pre-processing," we employ pattern recognition to address the issue using traditional image data augmentation techniques. Instead of techniques like Synthetic Minority Oversampling Technique (SMOTE) or ensemble methods, we focus on utilizing traditional image data augmentation to balance the dataset.

Traditional image data augmentation involves applying various transformations to the existing data, thereby generating additional samples. These transformations can include rotations, translations, flips, zooms, and changes in brightness or contrast. By augmenting the dataset with these variations, we increase the number of samples for minority classes, thereby reducing the class imbalance.

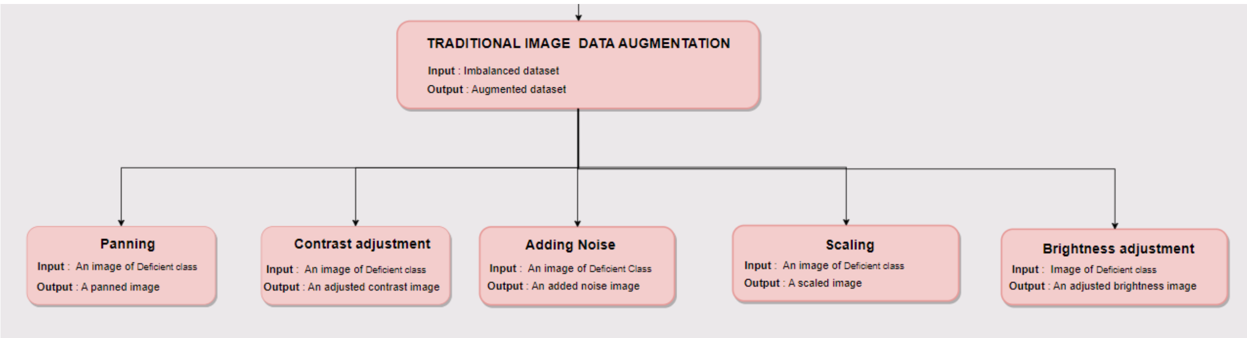


Figure 7. Some techniques used in traditional image data augmentation

Through traditional image data augmentation, we harness the power of pattern recognition to create a more balanced dataset, enabling the model to learn from a more representative sample distribution. This approach improves the performance of the classification model in accurately predicting rice paddy diseases on the imbalanced dataset.

Step 4. We use Pattern Recognition skill (Mapping)

In the sub-problem of "Paddy disease classification," we apply pattern recognition skills to map it to a similar problem in our knowledge domain, which is classification on image. By utilizing pattern recognition, we can identify patterns and structures within the image data, thereby transforming the original paddy disease classification problem into an image classification problem.

The task of classification on image focuses on categorizing images based on their visual information. In this case, we can use machine learning algorithms such as Convolutional Neural Networks (CNN), Support Vector Machines (SVM), or Decision Trees (DT) to build classification models. These algorithms have the capability to analyze specific features within the images and determine the corresponding disease types in paddy plants.

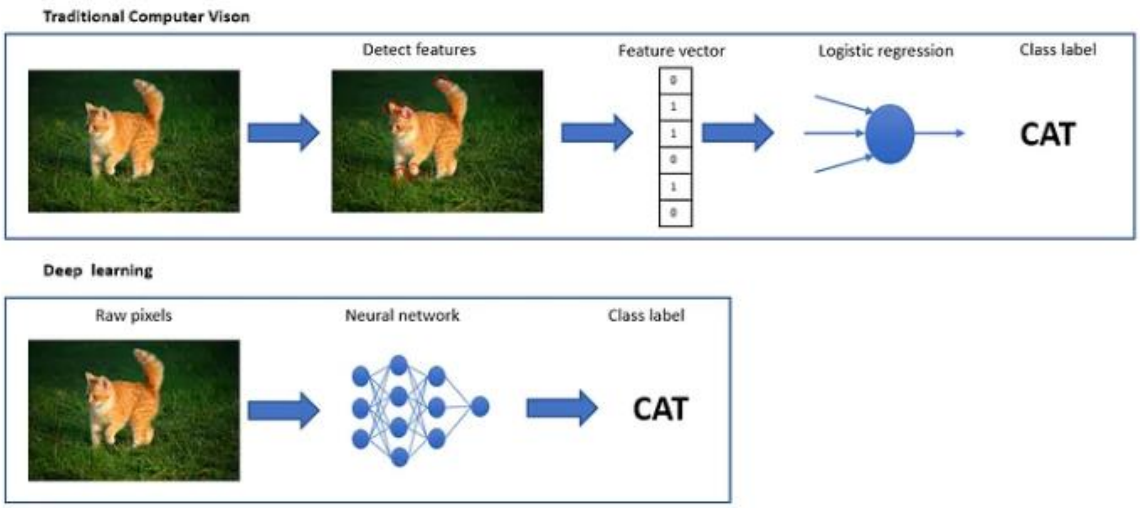


Figure 8. Examples of some classification algorithms

To summarize, in the sub-problem of "Paddy disease classification," we employ pattern recognition skills to map it to the classification on image problem. By recognizing patterns and structures within the image data, we construct classification models using machine learning algorithms such as CNN, SVM, or DT to address the paddy disease classification problem. The application of pattern recognition skills and classification algorithms on images empowers us to achieve precise and efficient classification outcomes for our problem.

Step 5. We use Pattern Recognition skill (Mapping, Grouping)

Pre-processing on images is a crucial step in various computer vision tasks, such as object detection, image recognition, and image segmentation. It involves applying a series of operations to raw images to enhance their quality, reduce noise, and extract meaningful features. This sub-problem focuses on discussing the key aspects and techniques involved in pre-processing images.

One essential step in pre-processing is resizing the image to a suitable resolution. Image resizing allows for uniformity in the dataset and helps reduce computational complexity. Techniques like nearest-neighbor interpolation, bilinear interpolation, or bicubic interpolation can be employed to resize the image while preserving its content.

Images captured from various sources often contain unwanted noise, which can degrade the accuracy of subsequent image processing algorithms. Noise reduction techniques like median filtering, Gaussian filtering, or bilateral filtering are employed to reduce noise while preserving important image details. These filters help to smoothen the image and improve its overall quality.

Pre-processing on images is an important sub-problem in computer vision that involves several techniques to enhance image quality, reduce noise, and extract meaningful features. The discussed techniques, including image resizing, enhancement, noise reduction, normalization, segmentation, and feature extraction, play a vital role in preparing images for various computer vision applications. By employing appropriate pre-processing techniques, the accuracy and performance of subsequent image-processing algorithms can be significantly improved.

Traditional image data augmentation is a popular technique in conventional image processing, where various transformations such as rotation, scaling, cropping, flipping, lighting adjustments, and many other transformations are applied to generate multiple different versions of the same original image. The techniques used in "pre-processing on image" are similar to traditional transformations in traditional image data augmentation. Therefore, using pattern recognition grouping can bring them together under a common form. This approach allows for the identification of similarities and helps save time and costs.

Step 6. Visualization with text on image

Visualization of images is a technique used to enhance the communication of visual information by incorporating various visual elements. This method aims to present information in a clear, engaging, and easily understandable manner through the addition of visual components, such as text, charts, data representations, graphical icons, or other graphic elements.

From the available visualization methods, we conducted an analysis and chose Visualization with text on an image. We believe that adding text to images creates an interactive combination of visual and textual content, thereby facilitating the easy and effective transmission of information to users

Visualization with text on the image was selected due to its flexibility and interactivity. We believe that incorporating text within images helps users comprehend context, explain critical elements, and establish a connection between the image and the conveyed message. Additionally, this method provides a visually appealing and impactful experience for viewers.

For example:

Figure 9. Examples of some Visualization with text on the image

Decomposition Tree

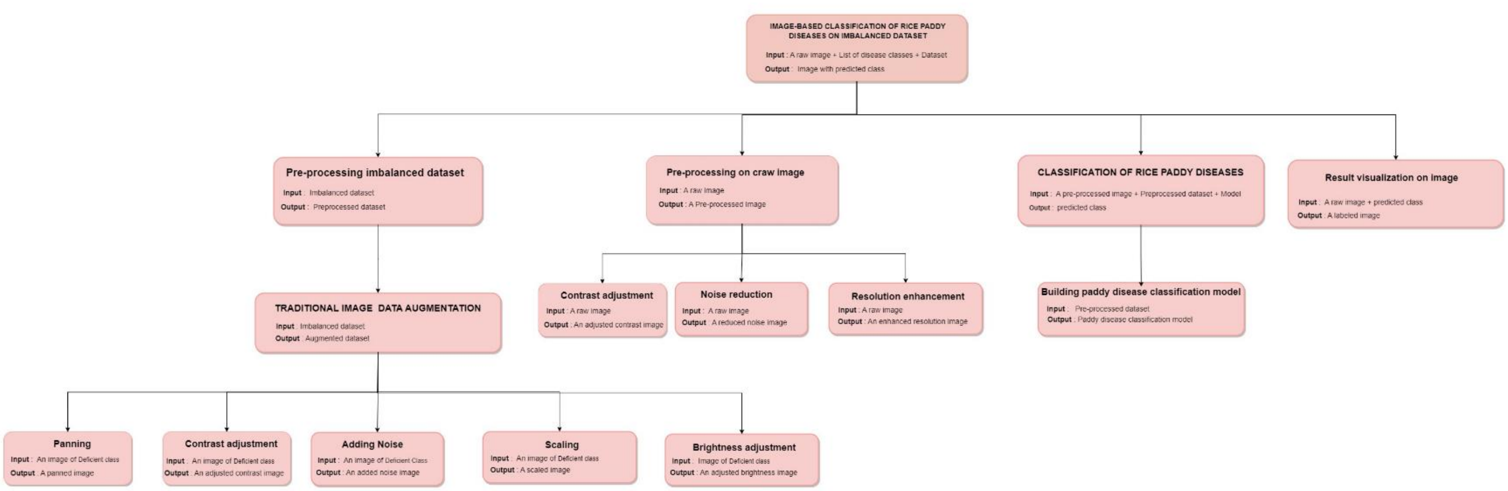


Figure 10. Decomposition Tree

Chapter 3. OUR SOLUTION

3.1. Method

3.1.1. Traditional Image Data Augmentation

CNN is an effective model for extracting characteristics from unstructured data. However, they lack image invariance due to the downsampling process that alters the image [1]. To enhance the performance of neural networks, certain transformations can be applied to the dataset in order to generate a diverse set of samples, thus increasing the network's robustness. This is achieved through data expansion and increasing the number of training sessions. To achieve invariance to affine transformations in the samples, the network is typically trained using the Traditional Image Data Augmentation (TIDA) method. This involves applying rotations, translations, scaling, brightness adjustments, contrast adjustments, and adding noise to the images. The transformed images are then used to augment the original dataset, as depicted in Figure 4.

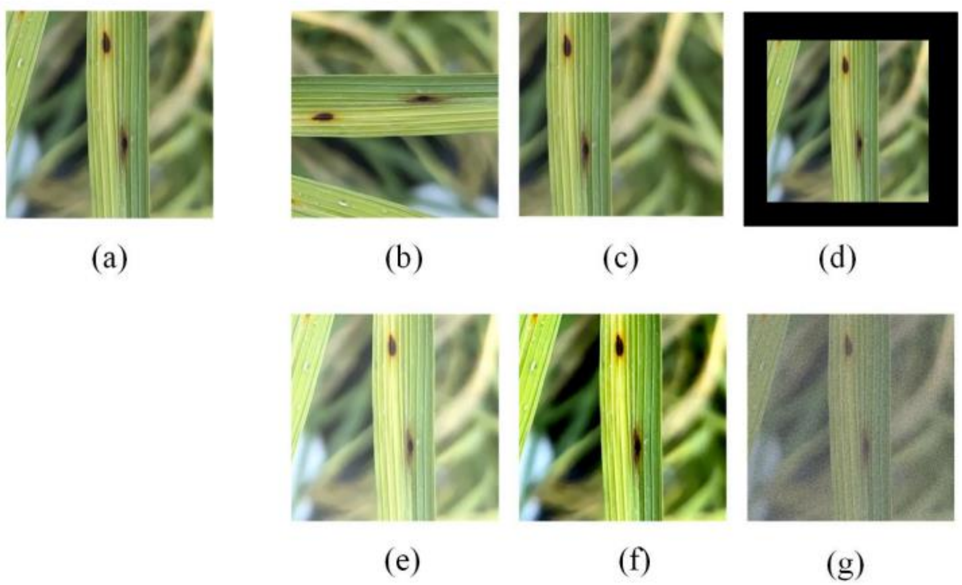


Figure 11. TIDA approach, where (a) is the original image, (b) rotation, (c) panning, (d) scaling, (e) brightness adjustment, (f) contrast adjustment, and (g) adding noise.

3.1.2. Convolutional Neural Network

3.1.2.1. Why we choose CNN to solve problems

ANNs have been used for image classification tasks, but they have some limitations. One of the main challenges of using ANNs for image classification is that they do not take into account the spatial structure of the image. ANNs treat each pixel as an independent feature, leading to poor performance on tasks that require spatial reasoning, such as object recognition. To handle the spatial structure of pictures, neural networks can be classified as convolutional neural networks (CNNs). CNNs employ convolutional layers, which subject the input picture to several teachable filters. Each filter provides a feature map that indicates the locations of the features it has detected in the picture, such as an edge or texture.

Criteria	Artificial Neural Networks (ANNs)	Convolutional Neural Networks (CNNs)
Architecture	Multilayer perceptron	Convolutional layers, pooling layers, and fully connected layers
Feature extraction	Hand-crafted or learned features	Learned features through convolutional layers
Spatial information	Not specifically designed to capture spatial structure	Specifically designed to capture the spatial structure of images
Parameter sharing	No parameter sharing	Parameter sharing through convolutional layers
Scalability	Less scalable due to the high number of parameters and overfitting	More scalable due to shared weights and hierarchical representations
Training data	Requires a large amount of training data to avoid overfitting	More efficient use of training data due to parameter sharing
Transfer learning	Less effective for transfer learning	Effective for transfer learning due to pre-trained models on large datasets
Computational efficiency	Less computationally efficient, especially for large images or deep architectures	More computationally efficient, especially for large images or deep architectures
Accuracy	Can achieve high accuracy on image classification tasks	Can achieve higher accuracy than ANNs on image classification tasks

***Table 1:** Comparision between ANN and CNN*

When it comes to image classification jobs, CNNs have several benefits over ANNs. They are more adapted to capture the spatial structure of pictures and automatically extract pertinent characteristics from pictures, which increases their efficiency for challenging image classification tasks like locating certain items in an image.

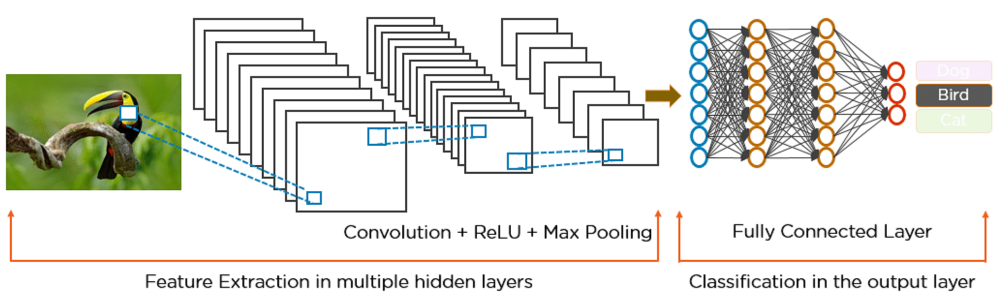


Figure 12. How Convolutional Neural Network work

To conclude, CNNs are typically preferred over ANNs for image classification tasks because they are designed to capture the spatial structure of photos and automatically extract relevant characteristics. Moreover, CNNs are better suited for large datasets and challenging image classification tasks than ANNs since they are more efficient and scalable.

3.1.2.2. Our CNN

Convolutional Neural Networks (CNNs) have emerged as a powerful tool in the field of computer vision and image classification. With their ability to automatically learn and extract meaningful features from raw image data, CNNs have revolutionized the way we approach complex image recognition problems.

The proposed CNN architecture is specifically designed to address the problem of classifying paddy disease. By leveraging the power of deep learning and convolutional neural networks, this architecture aims to accurately classify different types of paddy diseases based on input images.

The architecture consists of 6 convolutional layers, which are responsible for detecting local patterns related to paddy diseases in the input images. These convolutional layers utilize filters to learn and extract disease-specific features, such as discoloration, lesions, or spots, that are indicative of various types of paddy diseases. By using multiple convolutional layers, the model can capture both subtle and prominent disease-related patterns, allowing for more accurate classification, as shown in Figure 5.

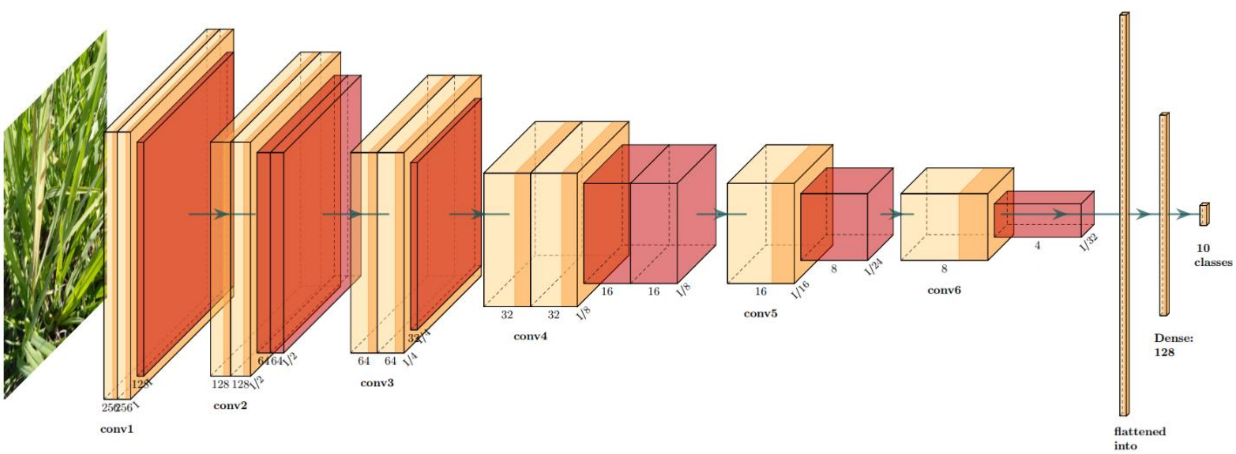


Figure 13. Convolutional Neural Network Architecture

To capture more abstract and higher-level features, the architecture includes pooling layers. These pooling layers reduce the spatial dimensions of the feature maps generated by the convolutional layers while preserving important information. By downsampling the feature maps, the pooling layers help the model focus on the most discriminative disease-related features, irrespective of their exact spatial locations. This enables the network to learn and generalize patterns associated with different paddy diseases, improving its classification performance.

Additionally, the architecture incorporates fully connected layers towards the end. These layers take the extracted features from the convolutional and pooling layers and process them to make the final classification decision. The fully connected layers further enhance the model's ability to capture complex relationships between the extracted features and different paddy disease classes. Dropout layers are also employed to mitigate overfitting by randomly deactivating a portion of the neurons during training.

Overall, the proposed CNN architecture for paddy disease classification combines the power of convolutional layers to detect disease-related patterns and pooling layers to capture abstract features. This combination allows the model to learn and recognize diverse paddy disease patterns, facilitating accurate and efficient disease diagnosis in the agricultural domain.

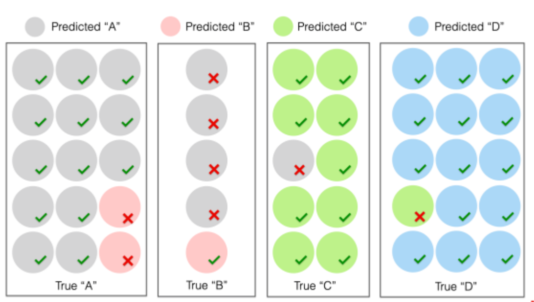
3.2. Evaluation

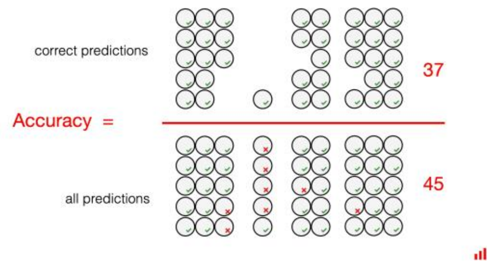
3.2.1. Accuracy

Accuracy measures the proportion of correctly classified cases from the total number of objects in the dataset. To compute the metric, divide the number of correct predictions by the total number of predictions made by the model.

For example:

- After we trained our model and generated the predictions for the validation dataset, we can evaluate the model quality. Here is the result we received:





- To calculate accuracy, divide all correct predictions by the total number of predictions. In our case, the accuracy is 37/45 = 82%.

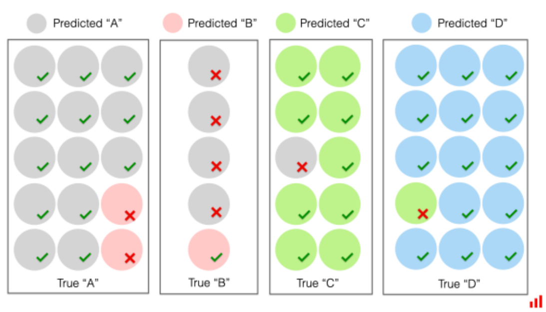
3.2.2. Precision and Recall

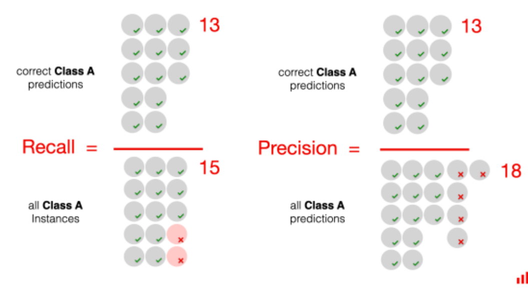
Precision for a given class in multi-class classification is the fraction of instances correctly classified as belonging to a specific class out of all instances the model predicted to belong to that class.

Recall in multi-class classification is the fraction of instances in a class that the model correctly classified out of all instances in that class.

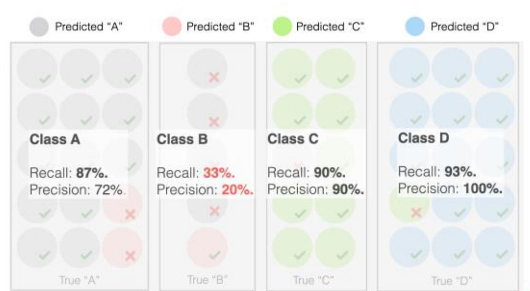
For example:

- After we trained our model and generated the predictions for the validation dataset, we can evaluate the model quality. Here is the result we received:





  - Recall is the ratio of correct predictions of Class "A" to the total number of Class "A" objects, while precision is the ratio of correct predictions of Class "A" to the total number of Class "A" predictions.



  - For other classes, we follow a similar approach.

Chapter 4. CONCLUSION

In conclusion, the image-based classification of rice paddy diseases on imbalanced datasets presents a significant challenge in modern agriculture. By addressing the issues posed by imbalanced datasets, researchers strive to develop effective methods for accurate disease classification. The utilization of machine learning techniques, such as data augmentation and resampling methods, plays a crucial role in improving the performance of classification models on imbalanced datasets.

The machine learning model trained on the imbalanced dataset has shown promising results in accurately identifying and classifying various types of diseases that affect paddy crops. This provides valuable insights for farmers and agricultural stakeholders, enabling early detection and timely intervention to mitigate the spread of infections and minimize yield losses.

However, challenges still exist, including color variations and overlapping symptoms that can affect classification accuracy. Ongoing research focuses on enhancing the robustness of the model by exploring advanced techniques, such as ensemble learning and fine-tuning pre-trained models. Additionally, the integration of emerging technologies, such as hyperspectral imaging and drone-based monitoring systems, holds great potential in improving disease detection accuracy.

Furthermore, efforts are being made to develop user-friendly interfaces that allow farmers and agricultural professionals to easily utilize machine learning models for disease diagnosis. These interfaces provide instant recognition and treatment recommendations, empowering farmers to make informed decisions and effectively manage diseases affecting paddy crops.

In conclusion, the application of machine learning techniques in the image-based classification of rice paddy diseases on imbalanced datasets offers significant advantages for modern agriculture. It contributes to proactive disease management, reduced crop losses, and increased agricultural productivity. With ongoing advancements in this field, the vision of an efficient and reliable system for paddy disease classification on imbalanced datasets is within reach, providing valuable tools for farmers and agricultural stakeholders in their quest for improved disease management and sustainable agricultural practices.

REFERENCE

1. Azulay, A., & Weiss, Y. (2018). Why do deep convolutional networks generalize so poorly to small image transformations?. *arXiv preprint arXiv:1805.12177*.
2. Yu, H. J., & Son, C. H. (2020). Leaf spot attention network for apple leaf disease identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 52-53).
3. Petchiammal, Kiruba, B., Murugan, & Arjunan, P. (2023, January). Paddy doctor: A visual image dataset for automated paddy disease classification and benchmarking. In *Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD)* (pp. 203-207).
4. O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.