CHAPTER I
INTRODUCTION

1.1 Background

Rice is a staple food that is widely consumed by about 95% of the population in Indonesia. West Java is responsible for a significant proportion of rice production, accounting for about 17 percent. It holds the third position in terms of rice output, following East Java and Central Java[1]. Seeds are essential living commodities that must be cultivated, processed, and harvested appropriately to produce the best yield and productivity in the agricultural market. The paddy rice seed Germination process will be studied at an observation interval of three to fourteen days during the certification procedure to determine the proportion of seed samples that usually sprout[2]. The seed certification is a systematic procedure that seeks to assess the quality of seed samples before they are supplied to farmers and accelerate seed production. The certification process relies on the compatibility of the seeds with their surrounding environmental factors. Additionally, certain plant growth processes are influenced by the initiation of dormancy breaking. The assessment of seed germination typically falls under the responsibility of seed producers. Seed germination involves reviving the embryonic axis development within a seed that has ceased to produce seeds (seedling). This transformation changes dormant seeds into seedlings that subsequently continue to grow. To achieve this, specific conditions must be met both within the seeds themselves and in the surrounding environment for successful germination.

Seed certification serves as a structured and systematic procedure aiming to evaluate the quality of seed samples before their distribution to farmers, thus facilitating seed production processes. The initial data collection for the sources was carried out by local farmers, who provided the relevant information to BPSBTPH for certification[3]. Currently rice seed germination evaluation is carried out manually by experienced persons in Balai Pengawasan dan Sertifikasi Benih Tanaman Pangan dan Hortikultura (BPSBTPH). This research will focus on classification the growth and quality rice seeds to speed up seed certification processes at BPSBTPH. Using a Computer Vision with models Deep learning an automated system to classification and detect the quality of rice seed germination. Deep Learning is a rapidly spreading approach that aids every industry in making informed decisions to maximize its applicability[4].

The rice seed germination process will be studied at an observation interval of three to fourteen days during the certification procedure to determine the proportion of seed samples [5] that usually sprout. Previous studies have used machine learning [6] to investigate a detection and germination classification approach using recent convolutional neural network architectures. The experiment involves analyzing 2400 seeds from three different crops. The objective is to accurately detect individual grains and distinguish between germinated and non-germinated seeds. The model yielded a success rate over 97%, indicating its efficacy in expediting seed germination across various seed cultivars through the utilization of machine learning techniques. CNN's models automatically extract and learn relevant features from raw images and have been applied to various image classification problems. Classifying seeds into two classes (germinated and non-germinated) uses the CNNs algorithm. The reason for their success in work can be attributed to their reduced reliance on various lighting conditions and obstacles, resulting in enhanced precision in the field of computer vision [7]-[8] The assignment involved the utilization of black paper sheets on a flatbed scanner to create a dark background for capturing the dataset. This approach was chosen due to its cost-effectiveness and easy availability in most stationery stores [9]. This makes the authors interested in conducting research by taking their datasets for classifying and predicting rice seed germination. The method has been developed for the categorization of rice seed varieties and the evaluation of rice seed germination.

This system utilizes a Digital Image processing system to identify the specific variety of rice seeds and assess their germination [10]. In the study referenced as [11], a data set was constructed and assessed using 24 extracted features. All available classification methods were employed for this purpose. Among these algorithms, the Ensemble classification algorithm demonstrated a notable accuracy of 91.6%. Germination is a set of phases that transform dormant seeds into seedlings that will remain to grow. Sources should fulfill conditions within the authorities and the environment to germinate. The CNN architecture demonstrates proficient learning capabilities when applied to the dataset, resulting in enhanced accuracy levels in the context of rice germination [12]-[13]. The paper presents a DNN using CNN architecture to classify Thai Hom Mali rice germination. The proposed CNN is an excellent technique for organizing the typical germination dataset, and for the testing data the total 35 images in the test set. This study uses very low data for testing data. This will affect the results of the performance of rice seed classification. The achieved accuracy rate stands at 0.89, with a corresponding loss of 0.11. However, this result falls short of the desired threshold of 90% for a satisfactory classification accuracy. Subsequent sections of this paper outline the upcoming research endeavors, which encompass the exploration of additional methodologies. Notably, methods like the Siamese Neural Network architecture, Inception-V3, and Xception will be implemented to address the multi-label rice germination classification challenge [14]-[15].

Utilizing computer vision assistance [16]-[17] system enables anyone to improve agricultural efficiency with germination prediction by applying deep learning algorithms to classify rice seeds. In recent times, the application of Deep Learning using convolutional neural networks in the field of computer vision has gained prominence for diverse purposes [7]. This study uses rice seed datasets from BPSBTPH to make it easier to classify and predict the germination. The initial data collection of the sources came from local farmers handed over to BPSBTPH for certification. If the rice seeds fulfill the criteria from BPSBTPH, they can be continued to be processed by farmers. Evaluate the prediction results using Deep Neural Network with architecture Convolutional Neural Network by evaluating the final target of the model that can deliver high accuracy. The approach should result in a system that can distinguish between normal germination, abnormal germination, Benih Segar Hidup (fresh germination), and dead, and also detect any factors that may slow the germination of rice seeds[18].

Inspired by prior research [14] that involved layer modifications for classifying multi-label rice seeds based on their growth and quality, this study proposes an architecture namely as STR-Net, with the term "STR-Net" derived from "SINTANUR Neural Network," where "SINTANUR" refers to the specific rice variety and "Neural Network" signifies the architecture and functionality of the model. The validation of this research entails the assessment of key parameters such as Accuracy, Precision, Recall, F1-score, and training time.

1.2 Problem Identification and Objective

This research was conducted to develop a classification system for Rice Seeds. Currently, rice seed germination evaluations are conducted manually by experienced individuals[7]. Based on the background, the following research problems can be formulated:

1. In the previous research [14], which involved modifications to the convolutional layers of the CNN, the accuracy achieved was below 90%. This lower accuracy could possibly be attributed to the presence of non-linearities within the dataset. Addressing this issue might involve increasing the diversity of image samples used for rice seed germination analysis.
2. The traditional method of assessing the quality of rice seeds relies on the expertise of BPSBTPH staff. However, this approach faces challenges due to human limitations in perception and a lack of technical support.
3. The system model’s efficiency in managing multi-label image data is evaluated by categorizing the germination of rice seeds into four classes as normal, abnormal, fresh (Benih Segar Hidup), and dead, based on a 14-day growth observation.

1.3 Research Objective

This research aims to create a system model capable of categorizing rice seeds into four different label classes, corresponding to their growth stages and quality over a 14-day observation period. This system is expected to provide significant support to BPSBTPH, a related organization, in the process of certifying the quality of rice seeds. The specific objectives of this research are as follows:

1. Enhance the accuracy of a previously developed model [14], by collecting more data on seed growth and the quality of rice seeds input for training the system. This approach aims to address the issue of non-linearity in data points by expanding the diversity of seed image samples.
2. Collecting seed data and developing a CNN-based classification model for rice seeds based on growth and quality, helping BPSBTPH staff streamline seed certification. The model categorizes seeds over a 14-day observation, automating classification by characteristics over time.
3. Evaluate the model's performance by adjusting hyperparameters in the CNN model. This adjustment aims to enhance the model's ability to automatically classify the germination of rice seeds.

1.4 Scope of Work

To ensure that this problem's scope does not extend to unrelated aspects, the problem's boundaries must be determined. The scope of this study's problem is as follows:

1. The study utilizes rice seed data obtained from BPSBTPH West Java. The specific variety used in the study is the SINTANUR variety, and the dataset consists of up to 4000 individual seeds.
2. The rice seed data is observed over a 14-day period, with four classes based on the observation day: "Day 3," "Day 5," "Day 7," and "Day 14." Additionally, each seed is labeled for quality as either "Normal," "Abnormal," "Fresh," or "Dead."
3. The Classification process is based on Deep Learning Method with CNN Architecture.
4. The system's input format for image data is specified as Photographic Group JPG File (.jpg). This format is used to represent the images to be processed by the system.
5. The study employs version 3.10.1 of the Python programming language, and Jupyter Notebook is utilized as the environment for implementing the system.

1.5 Research Methodology

1. Study Literature

    This process is learning stage about the theories of the Deep learning method, CNN with the architecture, and Rice Seed Growth from newest source such as paper, journal, and book.

2. Problem Identification

    In the problem identification stage, a literature study was conducted from previous research related to the classification of the seed germination to find out the problems that might be used as the material in this study. This stage also explored and identified the dataset problems of seed length of rice and found solutions that could be given.

3. Requirement Identification

    The process of requirement identification was subsequently undertaken, encompassing two main facets: research requisites and system specifications. This involved the delineation of necessary materials and methodologies essential for the accurate classification of rice seed germination. Within the research scope, it was imperative to amass pertinent data on rice seed germination. Concurrently, the identification of system prerequisites revolved around delineating the software and hardware specifications crucial to underpin the research endeavors.

4. System Design

    The system design was done to define the stages in creating a system classifying rice seed germination. The system built was a system that was able to classify seeds into four classes (day 3, day 5, day 7, and day 14).

5. Data Collection and Labelling

    Datasets were collected from Balai Pengawasan dan Sertifikasi Benih Tanaman Pangan dan Hortikultura (BPSBTPH). Furthermore, the data collected was carried out labeling process and validated by program.

6. System Implementation

    The implementation of a system designed previously was carried out in several stages: preprocessing data, dataset splitting, feature extraction using convolutional neural network.

7. System Testing & System Evaluation

    At this stage, the system built was tested using data testing to determine system performance.

8. Results Analysis

   The last stage, the results of previous tests were analyzed and conclusions from this study were made.

1.6 Research Timeline

The research timeline helps organize, and progress monitoring. The research timeline can be seen on **Table 1.1** and the output will be published in September.

Table 1. 1 Research Timeline

Timeline Activity	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Create system model in Jupyter Notebook									
Build dataset pre-processing									
Start writing the Thesis Section II & III									
Train the model of CNN Algorithm									
Screening Presentation									
Continue writing the draft Thesis Section IV									
Writing Paper									
Publish Paper, Review, and Acceptance									
Finalization of Thesis									

1.7 Structure of the Thesis

The subsequent sections of this thesis are structured as follows:

CHAPTER 1: INTRODUCTION

This chapter discusses the context of this research, including the problem in the field, related research, the scope of the work, and the research methodology employed.

CHAPTER 2: BASIC CONCEPTS

This chapter provides background information for this thesis, including a description of Digital Image, Deep Learning with CNN Architecture, the rice dataset, and the CNN method that will be utilized in this research.

CHAPTER 3: SYSTEM DESIGN

This chapter describes the system model, including the parameters and variables used in the thesis, the research flow, and how the algorithm's simulation functions.

CHAPTER 4: SIMULATION RESULT AND ANALYSIS

This chapter discusses the results of this thesis, beginning with the pre-processed dataset and moving on to the simulation's classification report, confusion matrix, and learning process time output.

CHAPTER 5: CONCLUSIONS AND FUTURE WORKS

This chapter concludes the thesis and makes recommendations for future research.