CHAPTER 1
PRELIMINARY

1.1 Background

Humans are one of the living beings in the world, just like other living things, over time, the physical condition of humans will decrease. This decline in physical condition makes the human body more susceptible to disease, but the disease does not only infect the old, even the young. One of the health complications suffered by humans that do not only attack the elderly and even toddlers is lung disease. According to data from the World Health Organization [WHO 2019], lung disease is one of the highest causes of death in the world.

Lungs are important organs in the human respiratory system, respiratory organs (breathing) associated with the respiratory system, and circulation (blood circulation) in the body of air-breathing vertebrates. Its function is to exchange oxygen from the air with carbon dioxide from the blood. If the function of the lungs is disturbed, then the health of the human body can be affected as a whole.

Common lung problems suffered by humans are problems due to viral, and bacterial infections. Pulmonary Infectious Diseases include the Corona Virus (SARS-CoV-2), Tuberculosis (TBC), and Pneumonia. This disease is generally characterized by one symptom, namely a continuous cough. In Humans from Toddlers to the Elderly cases of Coronavirus (SARS-COV-2) suffered by 250 million people in the world in 2021, then Pneumonia suffered by more than 150 million people in the world Per year in 2021, and Tuberculosis (TBC) with the disease with the 13th death in the world.

In diagnosing this disease, generally, the diagnostic tool or technique used is to perform an X-ray with Radiology rays or X-rays on the patient's chest. The image data generated from this technique is called a chest x-ray (CXR) image, this technique was chosen because this technique is economical and so easy to use (Qin et al., 2018). On CXR, an image of the inner surface of the patient's chest can be seen which is used as consideration by an expert to determine whether a patient has pneumonia or tuberculosis.

Along with the rapid advancement of technology, especially in the field of machine learning, problems in the image, such as classification, continue to be developed to become more dynamic for various problems, because it helps in distinguishing an object in the image can be done with the help of machine computing. So, the predictions generated by machine learning models in terms of detecting a disease can help a pulmonologist in diagnosing possible diseases suffered by patients who suffer from it.

Several studies have been carried out related to identifying lung disease using machine learning algorithms on CXR data. (Hwang et al., 2016) were the first to suggest the use of Convolutional Neural Networks in an automatic tuberculosis disease detection system using transfer learning to improve system performance, the accuracy of the system that was successfully created by them was 90.3% using a data set KIT, MC, and Shenzhen.

While research on a combination of several lung diseases was carried out by (Bar et al., 2015) by using a combination of features extracted by the Convolutional Neural Network (CNN) method with ImageNet architecture they succeeded in detecting cardiac enlargement, right pleural effusion, healthy vs. another abnormal disease with AUC values 93%, 89%, and 79%, respectively. However, (Kermany et al., 2018) with their own CNN architecture managed to beat the CNN model with the

GIST, VGG16, and VGG19 architectures in classifying 14 types of lung disease with a total accuracy of 92.4%.

In general, every single x-ray image on the chest using CXR will produce a scan and pulmonology needs to analyze each scan to know the results of the image manually to provide a diagnosis to the patient. This process analyzes CXR results take a long time and has a high tendency toward subjective results (Deng & Yu, 2013). Through this problem, a more efficient method of analyzing CXR image results is needed to be able to detect lung disease, one of these methods is the use of deep learning (Choi et al., 2005).

Currently, the use of deep learning as part of machine learning is increasing. This is fueled by the usefulness of deep learning that can help identify features from data, the increasing number of libraries of learning algorithms that can be used, and the ease of use of deep learning (Gliner et al., 2021).

Convolutional Neural Network (CNN) is a method of deep learning to quickly learn and identify features of images. This method makes the use of deep learning in the world of medical imaging increasingly widespread (Maeda-Gutiérrez et al., 2020). Often with deep learning research and development, CNN methods that use simple architectures are developed into several more complex architectures to improve performance using deep learning. Some of the architectures developed are Inception which is claimed to improve the performance of deep learning models and can reduce the number of metrics (Lee et al., 2017).

Several previous studies were carried out in clarifying lung disease using deep learning, in this study (Asif et al., 2020) was carried out with a CXR scan using Architecture v3 and using accuracy. However, the classification only uses 3 classes, namely Normal, Covid and Pneumonia, so it does not classify TB. Then the model does not measure the performance of the deep learning model with metric accuracy, sensitivity, precision, specificity, and F1-Scores.

Therefore, a deep learning model was designed for the detection of lung disease for 4 classes, namely normal people, Coronavirus (SARS-VOV-2), Tuberculosis (TBC), and Pneumonia. The model uses a CNN architecture, namely Inception V3 because it has layers with high complexity but produces a minimal number of parameters due to the use of the Inception Module which can make the feature extraction process in several layers parallel (Bankar & Gavai, 2018). The metric used to measure the performance of the deep learning model is accuracy, sensitivity, precision, specificity, and F1-Scores through the confusion matrix of the model results.

Based on the description of the problem above, the researcher is interested in conducting a study under the title **"Classification of Lung Diseases Using the V3 Inception Model"**.

1.2 Problem Identification

Based on the background of the problems that have been stated above, the identification of problems in the preparation of this Final Thesis can be identified as follows:

1. Health is very important. Especially in the lung section, identification of diseases in the lungs is one way to overcome problems in the lungs or diseases of the lungs where technology still needs to be optimized for accurate detection.
2. In the world of health, doctors have an important role in identifying diseases, but doctors are humans where humans must make mistakes, with technology aiming to reduce errors while evaluating the performance of technology in detecting diseases in the lungs.

1.3 Formulation of The Problem

Based on the identification of the problems stated above, the formulation of the problem solved in this study is:

1. How to detect lung disease using a deep learning model?
2. How is the performance of detecting 4 classes of lung disease with the Inception architecture and the Correlation between Evaluation Metrics with the number of parameters for each architecture?
3. What is the effect of using additional neural network layers after using a deep learning architecture?

1.4 Purpose Of Thesis

Based on the background and problem formulation stated above, the objectives of this study are as follows:

1. Design and test a deep learning model to detect 4 lung diseases using the Inception V3 architecture.
2. Analyzing the classification results from the deep learning model through evaluation metrics, namely accuracy, sensitivity, precision, specificity, and F1-Score using a confusion matrix.
3. Knowing the Effect of added layers on deep learning.

1.5 Thesis Benefit

Based on the background stated above, the benefits of this research are as follows:

1. the use of a convolutional neural network on 4 types of lung data.
2. Be a source of information about the performance of the Inception-v3 architecture
3. Become a source of information about the use of convolutional neural networks with the Inception-v3 architecture.

1.6 Scope of Thesis

Based on the background stated above, the scope can be identified as follows:

1. This study aims to analyze the performance of deep learning using a dataset of lung disease.
2. This study uses 4 variables from the convolutional neural network including evaluation metrics, namely accuracy, sensitivity, precision, specificity, and F1-Score using a confusion matrix.
3. The final result of the Deep learning Model Research using CNN with V3

Inception testing as a learner in the world of health.

1.7 Research Methods

This study uses the following methods:

1. Literature Study

   Literature study is the process of collecting information related to the discussion material from books and journals related to dry Coronavirus (Sars-Cov2), Tuberculosis (TBC), Pneumonia, deep learning, convolutional neural network (CNN), and Inception V3.

2. Deep Learning Model Design

   The design stage is carried out by creating a deep learning model schema, a trial plan schema, and a deep learning model analysis schema.

3. Deep Learning Model Testing

   The testing stage is carried out by running a deep learning model program on the provided Notebook.

4. Deep Learning Model Testing and Analysis

   The testing stage is carried out by running an analysis scheme on the deep learning model to assess the model's performance.

1.8 Thesis Writing Systematics

Systematics of writing can be divided into three parts, namely the beginning, content and end. The following is the systematic writing:

CHAPTER I PRELIMINARY

This chapter, it is explained the background of the problem, the identification and formulation of the problem, the objectives and benefits of the research, the scope of the research, research methods, and the systematics of writing the final Thesis.

CHAPTER II THEORETICAL BASIS

This chapter contains sources of knowledge that are the basis for supporting the thesis argument "*Deep Learning Model for Classification of Diseases in the Lungs with V3 Inception Testing* and *Convolutional Neural Network Algorithm Approach*".

CHAPTER III RESEARCH METHODS

This chapter contains an overview of the research object to be studied and the approach that will be used to complete the thesis according to the context

CHAPTER IV RESULTS AND DISCUSSION

This chapter contains a plan of research results and describes what techniques are used in data collection and the methods used in making the Final Thesis.

CHAPTER V CONCLUSIONS AND RECOMMENDATIONS

	This chapter states that the thesis has been completed and alternative views if there are opportunities for suggestions for the development of further research to the next level.
REFERENCE LIST	This chapter contains the sources cited in the thesis