CHAPTER I
INTRODUCTION

1.1 Background

Leukemia is a malignant disease that affects the bone marrow and blood. Because it increases the rates of sickness (morbidity) and death (mortality), it continues to be a major global health burden. [1]. The two primary classifications of leukemia are the kind of blood cells involved (lymphoid or myeloid) and the pace of disease development (acute or chronic). Acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), acute myeloblastic leukemia (AML), and chronic myeloid leukemia (CML) are the four primary forms of leukemia that arise from the combination of these two categorization criteria [2]. In the bone marrow, the process of making blood cells starts with stem cells, which then develop into different kinds of blood cells. Depending on how quickly aberrant cells emerge and proliferate inside the body, leukemia may manifest as either acute or chronic [3].

Leukemia is classified into acute and chronic forms according to the rate of disease advancement and the level of cell maturation. The rapid advancement of acute leukemia, including acute myeloblastic leukemia (AML) and acute lymphoblastic leukemia (ALL), is marked by the proliferation of immature cells (blasts), leading to bone marrow failure. Conversely, chronic leukemias, including chronic myeloid leukemia (CML) and chronic lymphocytic leukemia (CLL), are distinguished by more mature or developing cells and a gradual disease development [4].

The neoplastic condition known as chronic lymphocytic leukemia (CLL) is characterized by the clonal proliferation of mature $CD5^+$ B cells in the peripheral circulation. A crucial immunophenotypic feature for diagnosing CLL is $CD5^+$, a cell surface marker usually present on T cells but abnormally expressed on neoplastic B cells. One of the most prevalent forms of leukemia in the older population, CLL is most often diagnosed in these individuals. Additionally, it is more common in males, and ethnic and genetic variants affect its frequency and clinical features [5].

A study entitled *"Diagnosis of Chronic Lymphocytic Leukemia Using iwCLL 2018 Compared with NCI-WG96 Criteria at Cipto Mangunkusumo Hospital: A Practical Consideration in a Resource-Limited Setting"* reveals that most Chronic

Lymphocytic Leukemia patients in Indonesia are diagnosed at advanced stages, with 63.4% classified as Binet C and approximately 70% as Rai III–IV at initial diagnosis. The limited availability of flow cytometry, the standard assay for confirming B-cell clonality, in just a few major Indonesian cities, clearly correlates with the deficiency of diagnostic facilities. Furthermore, many new patients pursue medical care only after their clinical symptoms have become severe, owing to financial and geographic limitations, which hinders early detection and diminishes the probability of obtaining optimal treatment [6].

Globally, there were 474,519 new cases of leukemia and 311,594 deaths in 2020, according to GLOBOCAN 2020 statistics utilized in a research titled "Disease Burden, Risk Factors, and Trends of Leukemia: A Global Analysis." These results show that leukemia continues to be a major public health burden, with national differences in incidence and fatality rates. Additionally, the research demonstrates that males often have greater rates of leukemia incidence and death [7].

Although CLL is much less frequent in Asia than in Western nations, leukemia still ranks ninth worldwide, according to a research titled "Chronic Lymphocytic Leukemia: A Literature Review." Leukemia is the most frequent kind of cancer in Indonesia, with 13,959 new cases in 2022. Despite the lack of specific statistics on CLL in Indonesia, the continually increasing overall burden of leukemia highlights the need of creating a more effective and fair diagnosis system. Given that the average age of diagnosis for CLL is between 65 and 72 years, these results suggest that the burden of hematologic illnesses in Indonesia is still substantial and will probably rise in tandem with the country's aging population. Additionally, this illness has a complicated clinical history. Patients may develop thrombocytopenia, anemia, recurrent infections, or even develop aggressive large B-cell lymphoma (Richter syndrome), which has a bad prognosis. This intricacy raises the demand for long-term healthcare services and complicates patient management. Additionally, a number of diagnostic procedures, including peripheral blood smears, FISH testing, bone marrow aspiration, and lymph node biopsy, are necessary for the diagnosis of CLL, although their accessibility varies across Indonesian healthcare institutions [8].

Acute myeloblastic leukemia (AML) accounted for 32.1% (42 out of 131 studies) of the evaluated literature, making it the most studied of the numerous forms of hematologic malignancies examined. Acute lymphoblastic leukemia (ALL) came in second. On the other hand, the incidence of using artificial intelligence to chronic leukemias is much lower, with just 1.5% (2/131) of all papers including chronic myeloid leukemia (CML) and 9.9% (13/131) involving chronic lymphocytic leukemia (CLL). This suggests that there is still a dearth of research on automating CLL identification and classification, which calls for greater study to create more thorough and efficient AI-based methods [9].

Leukemia is often diagnosed by a doctor or hematologist manually analyzing microscopic blood smear pictures to look at blood cell shape. This process is often carried out in combination with cytogenetic analysis, bone marrow aspiration, and a complete blood count (CBC). Inter-observer variability may occur from the manual method's high degree of accuracy, time-consuming nature, and reliance on the examiner's competence, despite its relative simplicity and cost-effectiveness. In order to increase the speed, consistency, and accuracy of leukemia detection, automated diagnostic techniques based on digital imaging and artificial intelligence have been developed [10].

Accurate classification of leukemia subtypes is essential for enhancing patient survival rates and determining the optimal therapeutic approach. Unlike conventional methods reliant on manual microscopy, a deep learning-based approach has demonstrated enhanced accuracy in diagnosing leukemia and classifying its subtypes, as indicated in the *paper "Deep Learning-Based Diagnosis and Subtype Classification of Acute Lymphoblastic Leukemia*". Artificial intelligence technology serves as a potential diagnostic instrument in hematology, enabling automated systems to accurately and efficiently discern the morphological characteristics of white blood cells.[11].

According to the journal paper *"Swin Transformer: Hierarchical Vision Transformer using Shifted Windows"* the Transformer is better at identifying intricate visual patterns than traditional techniques because it can learn long-term dependencies via the self-attention mechanism. The article then presents the Swin Transformer, a hierarchical architecture with a shifting window mechanism

intended to overcome CNNs' drawbacks, especially with respect to changes in object size and processing burden on high-resolution pictures. [12].

This work demonstrates how the architectural configuration of the Swin Transformer—which distinguishes the Swin-T, Swin-S, Swin-B, and Swin-L variants—such as the number of layers, patch size, embedding dimensions, and number of attention heads, significantly affects its performance. Different computer vision tasks perform differently as a consequence of these setting variations [12].

Applying the Swin Transformer to medical pictures may result in great classification performance, according to the paper "*Shifted Windows Transformers for Medical Image Quality Assessment.*" On cardiac MRI (LVOT) pictures, the Swin-Tiny model's accuracy was 95.59%, but Swin-Base's accuracy on chest X-ray images was 87.1%. These findings show that the Swin Transformer's shifting window attention mechanism effectively increases the precision and effectiveness of medical picture processing. [13].

Additionally, a study titled *"Classification of Lung Diseases in X-Ray Images Using Transformer-Based Deep Learning Models"* contrasted CNN models from earlier research with the Swin Transformer and Vision Transformer in the categorization of lung illnesses based on X-ray pictures. The results demonstrate the efficacy of the shifted window attention mechanism in enhancing the accuracy of medical image-based diagnosis, with the Swin Transformer achieving the highest accuracy of 96.10%, surpassing the Vision Transformer (95.10%) and several traditional CNN architectures like ResNet50 and MobileNetV2 [14].

Using the Swin Transformer architecture as the main paradigm, this work attempts to create a classification method for chronic lymphocytic leukemia. The shortcomings of manual diagnostic techniques, which are still often used in a variety of healthcare institutions and frequently encounter issues with subjective evaluation and time efficiency, are anticipated to be addressed by this strategy. This research aims to increase accuracy and consistency by using Transformer-based artificial intelligence technologies. Therefore, it is anticipated that the findings of this research will aid in the future development of improved digital hematological decision support systems as well as automated medical image-based diagnostic systems.

1.2 Problem Statement

Based on the background described above, the research questions are as follows:

1. How can the Swin Transformer algorithm be applied to classify chronic lymphocytic leukemia?
2. How does the Swin Transformer method perform based on the confusion matrix evaluation metrics of accuracy, precision, F1 score, and recall?
3. What is the effect of the Swin Transformer architectural parameters on model accuracy?

1.3 Research Objectives

This project seeks to diagnose Chronic Lymphocytic Leukemia (CLL) via the Swin Transformer approach to autonomously differentiate between pictures of normal white blood cells and those suggestive of leukemia by pattern recognition in microscopic images. Utilizing the shifting window attention method, the model is anticipated to proficiently extract both local and global information, facilitating precise classification.

1.4 Research Benefits

This study is expected to provide both practical and academic benefits. The benefits of this study are as follows:

1. Practical Benefits

   This study is expected to benefit practitioners in the medical field and hematology laboratories, particularly in improving the efficiency and accuracy of leukemia diagnosis through the application of an automatic classification system based on the Swin Transformer.

2. Academic Benefits

   This study is expected to make a scientific contribution to the development of artificial intelligence and medical image analysis, particularly regarding the application of the Swin Transformer for biomedical image classification.

1.5 Scope of the Problem

To ensure that this study remains focused and does not stray from the established topic, this study sets the following scope limitations:

1. The dataset used in this study consists of 10,216 single white blood cell microscopic images sourced from the official Mendeley Data website. https://data.mendeley.com/datasets/wydknkk5k2/2.
2. The classification process focuses solely on binary classification between normal cells and cells indicative of chronic lymphocytic leukemia, excluding other types of leukemia.
3. This study uses the Swin Transformer method for classifying chronic lymphocytic leukemia.
4. Model performance is evaluated using accuracy, precision, F1 score, and recall as the primary metrics.
5. The system implementation aspect is limited to model development and does not include system deployment in a clinical environment.
6. This study only covers the classification of chronic lymphocytic leukemia based on single white blood cell microscopic images, without considering other aspects beyond visual image analysis.