# CHAPTER I
# INTRODUCTION

## 1.1 Background

Lumpy Skin Disease (LSD) has become a serious threat to the cattle farming industry in Indonesia. The disease was first reported in 2022 in Riau Province and quickly spread to various regions. According to data from the National Animal Health Information System (ISIKHNAS) as of November 2022, 11.474 cases of LSD have been recorded across six provinces in Indonesia. As of February 2023, 32 regencies/cities in Central Java, 7 regencies/cities in East Java, and 4 regencies/cities in the Special Region of Yogyakarta have been confirmed positive for LSD [1]. This rapid and widespread spread indicates that LSD has the potential to cause significant economic losses if it is not supported by an effective early-warning system.

LSD is caused by a virus of the genus Capripoxvirus. The onset of LSD is marked by the appearance of nodules or lumps on the cow's skin that can spread throughout the animal's body [2]. The impact of this disease is highly detrimental to livestock farmers, including reduced milk and meat production, decreased hide quality leading to lower market value, and the risk of livestock mortality. From July 2023 to early June 2024, a total of 6.803 cases of LSD were recorded in East Java, with seven cows reported dead as a result of the disease. Globally, LSD outbreaks have been reported to cause a drop in milk production of up to 85% [3], highlighting the significant potential for economic losses for farmers.

The diagnosis of LSD in the field to date is still dominated by clinical examinations and laboratory tests such as Polymerase Chain Reaction (PCR). Although these methods are highly accurate, the process is relatively time-consuming, costly, and requires specialized personnel, particularly in rural areas [4]. This situation leads to delays in case management and increases the risk of disease transmission between regions. Therefore, a complementary approach is needed that enables rapid and efficient early detection based on digital imagery.

The challenges associated with field diagnosis highlight the need for innovations based on artificial intelligence technology. Advances in artificial intelligence, particularly in the fields of computer vision and deep learning, open up significant opportunities for the development of automated diagnostic systems based on digital images. Various studies and research indicate that deep learning methods, particularly Convolutional Neural Networks (CNNs), are capable of effectively extracting visual features and achieving high accuracy in disease classification based on digital images, including skin diseases in humans and animals [5]. With sufficient image data, artificial intelligence-based systems can be developed to consistently recognize patterns of visual symptoms, thereby accelerating the disease identification process. Furthermore, the use of digital images allows the diagnostic process to be conducted in a more standardized and documented manner, thereby facilitating the monitoring and evaluation of conditions in the field.

Several previous studies have examined the use of deep learning models for classifying skin diseases in both animals and humans. Triwerdaya and Utami (2025) proposed a hybrid model that combines the Vision Transformer (ViT) with several CNN architectures, namely DenseNet121, ResNet50, and InceptionV3. On the PAD-UFES-20 and Monkeypox Skin Lesion Dataset (MSLD), the results showed that the combination of ViT and DenseNet121 yielded the best performance with an accuracy of 95.5%[6]. These findings indicate that combining CNNs and ViT can improve classification performance by simultaneously leveraging local features and global context.

Another study, conducted by Sohail (2024), developed an LSD detection system for cattle using ViT, which was designed to process low-resolution images from smartphone cameras. The results of the study showed that ViT was able to understand the global spatial relationships in images of cattle skin and achieved 99% accuracy in cross-validation testing [7]. This demonstrates that the transformer approach holds great potential for application as a practical diagnostic system in the field.

In addition, studies utilizing CNN architectures such as DenseNet121 have also been conducted for LSD classification. With its dense connectivity, DenseNet121 is capable of efficiently extracting nodule texture features. The results of the study show that this model achieved an accuracy of 85.23% without segmentation and 84.99% with image segmentation [8]. Although quite good, the single CNN approach still has limitations in comprehensively understanding the distribution patterns of lesions on the cow's body. CNNs are more effective at recognizing local details such as the texture and shape of nodules, but are less optimal at representing the global spatial relationships between lesions that appear in various locations on the body.

ViT offers a different approach by utilizing a self-attention mechanism to capture global relationships between parts of an image [9]. However, ViT tends to be less effective at extracting detailed local features when used on its own. On the other hand, DenseNet121 is a CNN architecture known for its dense connectivity mechanism, where each layer receives information from all preceding layers [10]. This mechanism enables more efficient local feature extraction. DenseNet121 has fewer parameters compared to DenseNet169 and DenseNet201, making it more suitable for datasets with limited data. Previous research has shown that on relatively small medical image datasets, using deeper networks does not always significantly improve model performance and may increase the risk of overfitting [11][12]. Therefore, combining ViT and DenseNet121 into an ensemble model can serve as one approach to improve LSD classification performance in cattle.

Based on the identified issues, this study proposes an ensemble model combining Vision Transformer (ViT) and DenseNet121 for the classification of LSD in cattle using digital images. In this model, images of cattle skin containing LSD nodules are divided into several patches of fixed size and processed using the self-attention mechanism in ViT to understand the global distribution pattern of lesions. Meanwhile, DenseNet121 is used to extract local features in the form of texture and shape of nodules on cattle skin. The combination of these two approaches is expected to produce more accurate and consistent classifications regarding variations in the size, number, and location of LSD nodules. The developed model has the potential to be used as a rapid and efficient early-diagnosis tool in the field without relying solely on laboratory examinations, thereby enabling earlier disease management and reducing the spread of LSD as well as economic losses in the livestock sector.

## 1.2 Problem Formulation

Based on the background explanation, the formulation of this research problem is as follows:

1. How does the ViT – DenseNet121 ensemble model perform in classifying LSD in cattle based on digital images of cattle skin?

2. How does the performance of the ViT – DenseNet121 ensemble model compare to the single architecture based model in LSD classification tasks in cattle?

### 1.3 Research Objectives

Based on the formulation of the above problem, the purpose of this study is to develop and evaluate a model of LSD classification in cattle based on digital imagery using the ensemble architecture of ViT and DenseNet121, in order to produce an effective classification model in recognizing LSD and providing potential support in the early identification process.

### 1.4 Research Benefits

The results of this study are expected to provide an initial overview of the potential for using image processing technology as a tool for the early diagnosis of LSD in cattle, thereby helping cattle farmers and animal health professionals (veterinarians) conduct initial examinations more quickly and efficiently.

### 1.5 Problem Limitations

1. The dataset used consists of secondary data from Kaggle with certain variations in conditions, so it does not fully represent all real-world conditions.
2. This research focuses on the recognition and classification of images into two main classes: “normal cows” and “cows infected with LSD”.
3. The evaluation is limited to measuring model performance based on test results without addressing computational aspects such as training time or model complexity.