CHAPTER 1
INTRODUCTION

This chapter discusses the introduction of the research which includes the urgency of the research, the direction of the research, and the scope of the research. The discussion in this chapter includes the background, problem formulation, objectives, research benefits, and problem limitations.

1.1 Background

Coffee is the most widely consumed beverage in the world, coffee is the main export commodity produced predominantly in the tropics and is an economic pillar for millions of farmers [1]. However, coffee productivity is consistently threatened by various biotic barriers, especially pest and plant disease attacks that are able to reduce the quality and quantity of crops [2] [3]. Coffee Leaf Rust remains the most detrimental disease in the world [4]. Estimated production losses due to leaf rust (Rust) can reportedly reach a range of 20% to 70% at severe infection rates, depending on environmental factors and the resistance of the variety [5]. In addition to leaf rust, other diseases such as Cercospora leaf spot also worsen plant health and the competitiveness of coffee commodities in the international market [6].

Disease detection and pest symptoms have an important role in control strategies that allow for more precise disease control, reducing production costs through reducing excessive use of pesticides [6]. Unfortunately, manual visual inspection methods by farmers are limited by physiological factors (eye fatigue), psychological (subjectivity), and limited time availability [7]. In addition, the initial symptoms of the disease are often vague and difficult to distinguish from the condition of healthy leaves by the human eye, so there is often a delay in diagnosis resulting in widespread outbreaks [8].

In line with technological developments in the agricultural sector, artificial intelligence technology and computer vision offer automation solutions for the diagnosis of plant diseases. The object detection technique has become particularly relevant because of its ability not only to classify the type of disease, but also to localize the position of lesions on the leaf surface [9]. Although conventional methods such as SVM, KNN, and standard CNN have been widely used, they often face constraints in real-time applications [10]. The Deep Learning model of the YOLO (You Only Look Once) family overcomes these limitations and has become the industry standard for real-time applications [11]. Comparative studies of YOLOv1-YOLOv8 have been conducted, which show that YOLOv8 has the most consistent and superior performance compared to previous generations in various scenarios[12].

Specifically in the domain of coffee commodities, the latest study [6] conducted a systematic evaluation of various YOLO-based detection architectures ranging from YOLOv8 to YOLOv11, to detect Arabica coffee leaf pests and diseases on the improved BRACOL dataset. The results of the study confirm that YOLOv8s provides the best balance between accuracy and computational efficiency. However, the study also reported that YOLO performance decreased significantly in cases of small lesions, which was reflected in the low mAP value despite the relatively high recall. Symptoms of diseases such as rust often appear in the form of very small objects and have a high visual texture resemblance to other necrotic patches [6]. This phenomenon causes a high level of ambiguity between disease classes [13].

YOLO's limitations in detecting small objects are closely related to the downsampling mechanism in the backbone and neck [14]. In standard YOLOv8 architectures, the feature resolution is reduced to stride 32 (P5), so the spatial details of small objects can potentially be lost before reaching the detection head [15]. To overcome this decline in information, several studies have proposed the addition of a high-resolution detection head (P2) [16]. In the PyramidYOLOv8 study for rice disease detection, this study added a detection head at the stride level 4 (P2) so that the model was able to take advantage of high-resolution features that are more sensitive to small objects. This approach has been shown to improve the detection performance of small lesions compared to standard YOLOv8 structures that use only P3–P5 [17].

The use of high-resolution features also increases sensitivity to background noise [18] and irrelevant textures such as leaf fibers or leaf surface color variations [19]. Therefore, an attention mechanism is needed to help the model focus the feature extraction process on informative areas [19]. Previous research has shown that Convolutional Block Attention Module (CBAM) is able to improve the detection performance of small objects by filtering features through two stages, namely Channel Attention to determine "what" features are important and Spatial Attention to determine "where" the feature is located [19]. The effectiveness of CBAM in the context of plant diseases has also been proven to report an increase in mAP and a decrease in false positives in the detection of apple leaf disease [20].

The problem of class imbalance is also an obstacle in the BRACOL dataset. The Leaf Rust class has a much larger number of bounding boxes than other classes, which has the potential to cause model bias against the majority class [6]. Previous research confirms that the dominance of the majority class can cause the model to ignore the minority class, hence the need for adaptive augmentation strategies to enrich the visual variation of the minority class [21]. In addition, the quality of annotations also plays an important role in model learning. In object detection tasks, bounding boxes are used as localization targets during the training process so that bounding box inconsistencies can result in less than optimal supervision [23]. Previous research has shown that annotation localization errors can degrade prediction quality especially on small objects because relatively small changes in the *position of the bounding box* can adversely affect *the Intersection over Union* (IoU) value.

This study proposes the development of the YOLOv8s model with two main modifications, namely the addition of *a high-resolution detection head* (P2) and the integration of CBAM. The addition of P2 aims to preserve the spatial details of small lesions, and CBAM serves to improve the discrimination of disease features against the background of the leaves. In addition, *a balancing* strategy was applied to reduce bias due to class imbalances in the BRACOL dataset. With this approach, this study is expected to be able to fill the gaps in previous research, especially the weaknesses identified related to the low detection performance in small lesions and diseases with high visual similarity [6].

1.2 Problem Formulation

Based on the description of the background, several main problems can be identified that are the basis of this research, namely:

1. How to design and implement the YOLOv8s architecture with the addition of a high-resolution detection head (P2) and CBAM integration for coffee leaf disease detection?
2. How does the addition of a high-resolution detection head (P2) and CBAM integration affect the detection performance of coffee leaf disease on small objects?
3. How does the performance of the proposed YOLOv8s-P2-CBAM model compare to the standard YOLOv8s model based on evaluation metrics?

1.3 Research Objectives

To answer the formulation of the problem described, the objectives of the research can be arranged as follows:

1. Designed and implemented a modified YOLOv8s architecture with the addition *of* a high-resolution detection head (P2) and CBAM integration for coffee leaf disease detection.
2. Analyze the effect of the addition of *P2 detection heads* and CBAM integration on improving detection performance, especially in detecting small objects.
3. Compare the performance of the YOLOv8s-P2-CBAM model with the standard YOLOv8s using object detection evaluation metrics, to determine the effectiveness of the proposed model.

1.4 Research Benefits

This research is expected to provide benefits both theoretically and practically. Methodologically and theoretically, this study provides a scientific contribution in the form of an approach to modify the YOLOv8s architecture through a combination of adding a high-resolution detection head (P2) and CBAM integration to overcome the constraint of loss of spatial features in small objects. This approach enriches the study of the application of attention mechanisms to object detection models for the precision agricultural domain, especially in plant disease problems with complex visual backgrounds. Practically, the results of this research are expected to be the foundation of precision technology, thereby helping coffee farmers in detecting diseases as early as possible to reduce the risk of crop losses and improve the efficiency of pest control.

1.5 Problem Limitations

In order for this research to be directed and focus on the goals to be achieved, several research limitations are set as follows:

1. This study uses the YOLOv8s algorithm as the basic model. The architectural modification is limited to only two areas, namely the addition of *a high-resolution detection head (P2) in the* Head*, and the integration of an attention module in the* Neck.
2. The attention mechanism applied is limited to the *Convolutional Block Attention Module* (CBAM). The module is *plug-and-play integrated* on the feature fusion path in the *Neck section*.
3. The object of the study was focused on the detection of Arabica coffee leaf disease lesions using the revised BRACOL dataset. The detection class is limited to four main categories, namely *Rust*, *Cercospora*, *Miner*, and *Phoma*.