

CHAPTER I

INTRODUCTION

1.1 Background

Automated digital image feature extraction has undergone a fundamental transformation along with the adoption of Deep Learning architectures, particularly Convolutional Neural Networks (CNNs), which effectively replace the limitations of conventional manual image processing methods [1]. Convolutional neural networks are capable of extracting visual representations autonomously directly from raw pixel data without requiring extensive feature extraction pre-processing, establishing them as the primary standard for solving pattern recognition tasks in high-dimensional images [2]. In the development of image classification techniques, various modern CNN architectures have been extensively explored to achieve performance optimization. Comparative studies of network architectures demonstrate that DenseNet121 offers highly significant computational advantages, achieving accuracy levels that surpass other high-depth architectures due to computational time efficiency and a reduction in the number of parameters, which facilitates feature reuse across all network abstraction layers [1].

Although CNN architectures have proven superior in extracting visual representations, the performance optimization of these networks relies heavily on the class distribution within the training data space, where the phenomenon of imbalanced datasets often presents a persistent structural constraint [3]. This imbalance phenomenon is explicitly evident in the representation of the public PlantVillage dataset, commonly utilized as a computational object, where there is an extreme disparity in sample space quantity among classification classes [1]. In the specific details of this real data, the majority class, identified as Common Rust, dominates the distribution with 1,306 images, whereas the minority class, represented by Gray Leaf Spot, experiences a severe sample deficit with only 574 images [1]. This disproportionate data distribution fundamentally affects the model's learning process, creating an inductive bias that algorithmically strongly favors the majority class and reduces the network's capacity to generalize patterns for the minority class [3].

The computational impact of data imbalance manifests destructively when CNN architectures are trained using a standard evaluation function, namely Cross-Entropy Loss, which calculates error penalties equally across all samples without considering the ratio between classes [4]. The accumulation of gradient values from thousands of majority class samples dominates the direction of network parameter updates during the backpropagation process, thereby systematically steering the model to continuously prioritize the most abundant data category in order to rapidly minimize the global loss rate [4]. This computational condition triggers an anomaly known as the Accuracy Paradox, where the global classification accuracy value appears exceptionally high due to the dominance of correct guesses in the majority class, yet is inversely proportional to the Recall metric value, which drops drastically in the recognition of the minority class [3]. This decrease in network sensitivity directly triggers a surge in the False Negative rate in the final testing results, representing the failure of the computational algorithm to detect and classify visual features of minority class objects during the inference stage [2].

Resolving the challenges of class bias resulting from data imbalance has been attempted in several previous studies; however, conventional approaches have been evaluated as insufficient to robustly handle the functional weaknesses of the algorithm [3]. The application of pre-processing techniques through standard augmentation, such as random rotation modifications, flipping, and scaling (zoom) adjustments, is frequently implemented to engineer artificial spatial variations to increase training data samples [1]. Nevertheless, manipulation at the data level through such standard augmentation approaches is fundamentally incapable of mitigating the degradation of gradient propagation that is internally induced by weaknesses in the formulation of the standard loss function when confronting an imbalanced distribution [5]. Given these limitations, this research explicitly restricts the pre-processing stage solely to the use of standard conventional augmentation to shorten computational time, thereby shifting the focus of the problem-solving intervention entirely to modifications at the algorithm level without involving complex pre-processing techniques.

As a computational solution to these weaknesses, an intervention purely at the algorithmic level is proposed through the implementation of an adaptive loss

function in the form of Focal Loss to restructure the gradient evaluation mechanism within the neural network [5]. Mathematically, Focal Loss modifies the standard Cross-Entropy equation by injecting a focusing parameter that operates dynamically to induce the down-weighting of gradient contributions for samples from the majority class that are already easily predicted (easy examples) [4]. This adaptive penalty mechanism automatically forces the updating of the model's parameter weights to concentrate attention on samples from the minority class that are difficult to classify (hard examples), ensuring that the computational bias toward the majority space can be precisely reduced [2]. To validate this theoretical foundation, the computational experiments in this research are strictly designed to solely compare the classification performance between Cross-Entropy Loss and Focal Loss directly (head-to-head), utilizing the DenseNet121 architecture as the feature extraction backbone.

1.2 Problem Formulation

Based on the background of the problem previously outlined, the problem formulations in this study are as follows:

1. What is the impact of utilizing the standard evaluation function, Cross-Entropy Loss, on the model's performance metrics (particularly Recall and the False Negative rate) within the DenseNet121 architecture when confronted with extreme class data imbalance?
2. How does the effectiveness of implementing Focal Loss compare against the standard Cross-Entropy Loss in reducing algorithmic bias and enhancing the detection sensitivity for the minority class (Gray Leaf Spot) within the DenseNet121 architecture?

1.3 Scope of the Problem

To ensure that this research remains focused and comprehensive, the following scope and limitations of the study have been established:

1. **Dataset:** This study utilizes corn leaf disease image data sourced from the public PlantVillage dataset, focusing the testing on the extreme imbalance between the majority class (Common Rust, comprising 1,306 images) and the minority class (Gray Leaf Spot, comprising 574 images).

2. Network Architecture: The extraction of visual features is strictly limited to the utilization of the DenseNet121 Convolutional Neural Network (CNN) architecture.
3. Imbalance Handling Intervention: The intervention is purely focused on algorithmic-level modifications of the loss function, specifically a direct (head-to-head) comparison between Cross-Entropy Loss and Focal Loss.
4. Data Pre-processing: The pre-processing stage is explicitly restricted to standard conventional augmentation techniques (such as random rotation, flipping, and scaling/zooming) to ensure computational time efficiency, without involving complex synthetic data manipulations.
5. Evaluation Focus: The parameters for the model's success are not emphasized on global accuracy (to avoid the Accuracy Paradox), but rather on the improvement of the sensitivity metric (Recall), the precision of minority class detection, and the reduction of False Negatives.

1.4 Research Objectives

The objectives to be achieved through this research are:

1. To analyze and empirically demonstrate the computational impact of imbalanced datasets on the emergence of the Accuracy Paradox and the degradation of Recall values in the DenseNet121 model utilizing Cross-Entropy Loss.
2. To contrastively evaluate and compare the effectiveness of the Focal Loss function in forcing the model to concentrate gradient evaluation on the minority class (hard examples), thereby enhancing the robustness of detection performance on the Gray Leaf Spot class compared to using Cross-Entropy Loss.

1.5 Research Benefits

The results of this research are expected to provide beneficial contributions, both theoretically and practically:

1. Theoretical Benefits:
 - Providing literature contributions to the fields of Computer Vision and Deep Learning regarding the empirical validation of the

effectiveness of Focal Loss as an algorithmic-level solution for handling extreme dataset imbalances within efficient architectures such as DenseNet121.

- Providing a more profound computational understanding concerning the Accuracy Paradox phenomenon and how the down-weighting mechanism in Focal Loss is capable of improving the distribution of gradient parameter updates.

2. Practical Benefits:

- Producing a corn leaf disease classification model that is significantly more sensitive and precise in detecting rare (minority) disease classes, thereby mitigating the risk of detection failures (False Negatives) that could detrimentally affect early diagnosis.
- Offering a computationally efficient classification framework, where the handling of data imbalance can be resolved without the need to add processing time burdens during the synthetic augmentation data pre-processing stage.