Leveraging VGG-16 and Optuna for Carbon Stock Estimation with Drone Imagery in Indonesia

1st Caecarryo Bagus Dewanata
School of Computing
Telkom University
Bandung, Indonesia
caecarryo@student.telkomuniversity.ac.id

2nd Erwin Budi Setiawan
School of Computing
Telkom University
Bandung, Indonesia
erwinbudisetiawan@telkomuniversity.ac.id

3rd Gamma Kosala
School of Computing
Telkom University
Bandung, Indonesia
gammakosala@telkomuniversity.ac.id

Abstract—Accurate estimation of carbon stocks is vital for management. While field-based methods often face limitations in coverage, remote sensing approaches present a more effective alternative. Recent advancements in deep learning have enabled the application of Convolutional Neural Networks (CNNs) to analyze high-resolution drone imagery in carbon stock estimation. This study assesses the performance of VGG-16 and ResNet-20 for regression tasks, employing Optuna for hyperparameter optimization to enhance prediction accuracy. The experimental findings reveal that VGG-16 achieved an R^2 score of 0.645, with lower RMSE and MAE values than ResNet-20. Furthermore, the study highlights significant challenges, such as dataset imbalance and feature extraction in regions with high carbon stocks. Future research may investigate hybrid learning techniques, ensemble models, and multispectral data fusion to improve model estimation accuracy and generalization.

Keywords—Carbon stock estimation, drone imagery, VGG-16, ResNet-20, deep learning.

I. Introduction

Climate change is one of humanity's most critical challenges, driven primarily by rising greenhouse gas emissions, particularly carbon dioxide (CO_2). As one of the most significant contributors to global CO_2 emissions, Indonesia recorded 1.24 gigatons of CO_2 equivalent in 2022, a 10% increase from the previous year [1]. The country's extensive deforestation, land-use changes, and fossil fuel consumption are key contributors to this alarming trend [2], [3]. As natural carbon sinks, forests play a vital role in mitigating climate change by absorbing atmospheric CO_2, yet deforestation disrupts this cycle, intensifying global warming.

Accurate carbon stock estimation is essential to addressing climate change, providing actionable insights for policymakers to implement mitigation strategies. Traditional field-based measurements, though correct, are labor-intensive, costly, and unsuitable for covering large or remote areas. Remote sensing technologies, particularly drone imagery, offer a scalable alternative with higher spatial resolution than satellite imagery, enabling the capture of fine-grained details critical for forest ecosystem assessments. Drone imagery is particularly beneficial in tropical forests like Indonesia, where dense canopies and varied terrain hinder traditional methods [4], [5].

Recent studies have explored using Convolutional Neural Networks (CNNs) for carbon stock estimation, leveraging high-resolution aerial and satellite imagery[6]. VGG-16 and ResNet-20 have gained attention among various architectures due to their strong feature extraction capabilities. However, CNN performance highly depends on hyperparameter optimization, which can significantly impact prediction accuracy. This study employs Optuna and GridSearchCV to identify optimal configurations for carbon stock regression to enhance estimation precision [7].

Despite advancements in CNN-based carbon stock estimation, several challenges persist. One major limitation is dataset imbalance, where low-carbon stock regions are overrepresented, leading to underestimating high-carbon stock values. Additionally, dense vegetation [8], overlapping canopy structures, and lighting variations pose challenges in feature extraction, reducing model reliability. Furthermore, while deep learning models provide robust feature extraction, integrating additional data sources such as multispectral or LiDAR data remains underexplored [9].

Hyperparameter tuning plays a critical role in optimizing CNN performance. Traditional methods like manual tuning or GridSearchCV are computationally expensive and prone to suboptimal results. Optuna, a modern hyperparameter optimization framework, offers an efficient alternative by dynamically exploring and pruning the hyperparameter space, thereby improving model performance with reduced computational overhead [10]. This study integrates Optuna with VGG-16 and ResNet-20 to evaluate their effectiveness in estimating carbon stock [11].

This study aims to investigate the potential of VGG-16 and ResNet-20 architectures for regression-based carbon stock estimation using high-resolution drone imagery. VGG-16 serves as the primary model, with ResNet-20 providing a comparative baseline to assess the impact of residual learning on model performance. The objectives of this study include evaluating the performance of VGG-16 and ResNet-20 in predicting carbon stock values, leveraging Optuna to optimize hyperparameters efficiently, and highlighting the application of drone imagery in advancing carbon stock estimation methodologies [12]. This research contributes to the field of carbon stock estimation by demonstrating the advantages of CNN-based feature extraction combined with advanced hyperparameter tuning. The findings underscore the effectiveness of Optuna in optimizing CNN models for regression tasks while identifying key datasets and model limitations.

Because automated approaches can achieve better performance with less training time and resource usage, these inefficiencies highlight the benefits of adopting the automatic hyperparameter tuning approach offered by Optuna's

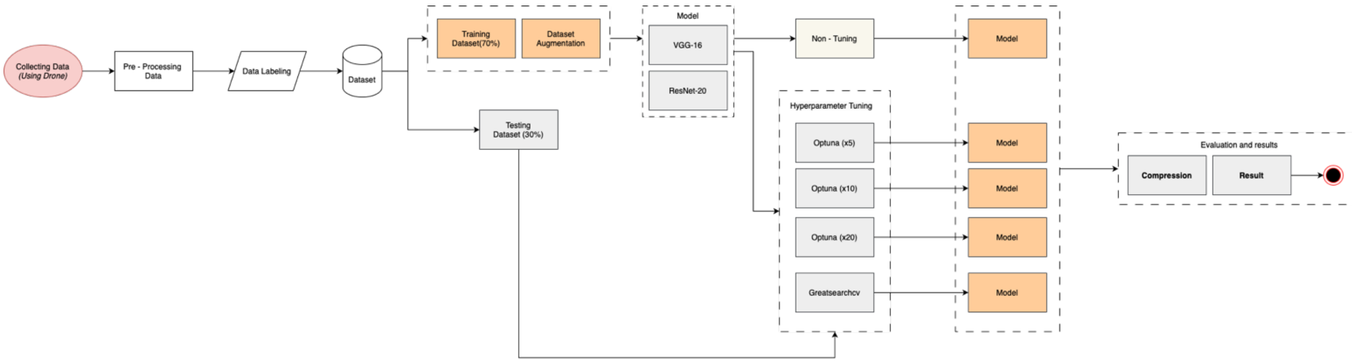

Figure 1 Architecture System Flowchart

Bayesian optimization framework, which helps to solve these difficulties. Effective search space exploration and pruning techniques to lower computational costs dramatically increase model accuracy. Optuna, for instance, has significantly raised its R2 scores, going from 0.8857 to 0.9524 [13]. Optuna's capacity to improve model performance has been demonstrated by its wide range of applications in various fields, significantly improving model efficiency [14]. For instance, fresh malware samples were created using the Generative Adversarial Network (GAN) sampling technique in malware detection.

The various variations of GANs necessitate hyperparameter optimization to determine the best settings for every dataset. By optimizing the GAN model's parameters using the autonomous hyperparameter tuning algorithm Optuna, this study achieved 98.06% accuracy, 99.00% precision, 97.23% recall, and 98.04% F1 score. After fine-tuning the hyperparameters of five supervised boosting algorithms—XGBoost, LightGBM, CatBoost, Extra Trees Classifier, and Gradient Boosting Classifier—the study applied a weighted ensemble technique to improve model performance further. Using Optuna to optimize the model's hyperparameter efficiency has significantly enhanced GAN performance and efficiency [15], [16]. Optuna has demonstrated its effectiveness in practical applications, achieving 85% and 80% accuracy rates in different real-world scenarios [17].

By combining drone images with a pre-trained VGG-16 model optimized with Optuna, this paper's primary contribution is the integration of sophisticated approaches for carbon stock estimation. Although deep learning models such as CNN have been investigated in the past for environmental applications [4], more is needed to integrate these components in this particular setting. To properly utilize drone imagery's potential for carbon estimation, most research either applied optimization to a small range of models or relied on manual hyperparameter tuning. This study fills this knowledge gap by proposing a new approach to improving the precision and efficiency of carbon stock forecasts. Furthermore, in contrast to previous research that primarily used Optuna for classification problems, this work focuses on regression, offering a novel perspective on using Optuna in environmental modeling.

II. METHODOLOGY

This section outlines the methodology employed in this study, which includes data collection, preprocessing, feature extraction, model training, and hyperparameter optimization. Data collection primarily consisted of high-resolution drone imagery captured from tropical forest regions in Indonesia. These images offered detailed visual insights, showcasing variations in vegetation density, canopy structure, and environmental conditions. The choice of drone imagery over satellite imagery was made due to its superior spatial resolution, enabling more accurate mapping of small-scale vegetation features crucial for estimating carbon stock.

A. Data Collecting & Preprocessing

High-resolution drone imagery was utilized as the primary dataset to capture vegetation characteristics, including biomass, density, and canopy structure [18]. The dataset was collected across various vegetation types and environmental conditions. Each image was labeled with carbon stock values (kg/ha) extracted using regular expressions from structured file naming conventions. Table 1 shows an example of the labeled dataset [19].

Table 1 Labelled Dataset

No	Name File	Plot Id	Carbon Amount (Kg/Ha)
1	Telkom_Z4H1P1-8181,394.PNG	10/05/2024	8181,394
2	Telkom_Z1H1P1-6146,061.PNG	11/05/2024	6146,061
3	Telkom_Z3H1P2-4330,987.PNG	10/05/2024	4330,987
4	Telkom_Z1HIP3-7072,629.PNG	02/05/2024	7072,629
5	Telkom_Z2H1P1_11920,73.PNG	15/05/2024	11920,73
	Total Carbon Summary		37,651.801

$$y = y_1 + \frac{(x - x_1)(y_2 + y_1)}{x_2 + x_1} \quad (1)$$

Normalization was performed to standardize pixel values across the dataset:

$$x = \frac{x - \mu}{\sigma} \quad (2)$$