



Resolving Visual Ambiguity in Wood Grain Classification Using InceptionV3 and Label Smoothing

Rianto Rianto¹, Sulistyo Dwi Sancoko², Eko Setyo Humanika³

^{1,2}Department of Data Science, Faculty of Science & Technology, Universitas Teknologi Yogyakarta, Sleman, Indonesia

³Department of English Literature, Faculty of Business & Humanities, Universitas Teknologi Yogyakarta, Sleman, Indonesia

*rianto@uty.ac.id

Abstract

Wood type classification in the timber industry is frequently hindered by the high visual similarity of grain patterns, a structural challenge that is particularly pronounced with small-scale datasets. This study systematically resolves this fine-grained visual ambiguity by developing a rigorously optimized InceptionV3 framework. Initial baseline evaluations conducted on a dataset of 800 images encompassing four wood species, namely *Pterospermum javanicum* (Bayur), *Magnolia champaca* (Cempaka), *Tectona grandis* (Jati), and *Melia azedarach* (Mindi), revealed that a standard ResNet50 architecture experienced a severe performance degradation, generating a random guess accuracy of merely 25 percent. Addressing this severe architectural inadequacy, the proposed InceptionV3 model integrates highly targeted spatial augmentations, a Dropout rate of 0.6, and L2 regularization. Furthermore, the optimization strategy deploys Stochastic Gradient Descent with Nesterov momentum and Label Smoothing to explicitly mitigate intra-class visual similarities. Consequently, the proposed framework achieved a robust overall accuracy of 89 percent, surpassing the baseline by a substantial 64 percent, alongside a macro-averaged F1 Score of 0.89. These empirical findings substantiate that specific architectural fine-tuning and advanced stochastic regularization are highly essential for capturing subtle wood-grain patterns, thereby offering a highly reliable, automated quality-control solution for the timber industry.

Keywords: fine-grained classification; InceptionV3; label smoothing; visual ambiguity; wood grain patterns

How to Cite: R. Rianto, S. D. Sancoko, and E. S. Humanika, "Resolving Visual Ambiguity in Wood Grain Classification Using InceptionV3 and Label Smoothing", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 3, pp. 669 - 676, Jun. 2026.

Permalink/DOI: <https://doi.org/10.29207/resti.v10i3.7599>

Received: February 24, 2026

Accepted: May 23, 2026

Available Online: June 14, 2026

This is an open-access article under the [CC BY 4.0 License](#)

Published by [Ikatan Ahli Informatika Indonesia](#)

1. Introduction

Wood is a natural material with diverse physical characteristics, and grain patterns can distinguish its types. Accurate identification of wood types is essential, as it determines suitability for specific structural applications and industrial processing. Variations in wood types significantly affect mechanical strength, durability, and end use performance [1-3]. In the timber industry, automated identification is increasingly critical for quality control and the prevention of commercial mislabeling [4]. However, human visual identification of wood types remains challenging due to the high similarity in grain textures [5], [6]. The complexity of this task is particularly evident when different wood types exhibit nearly identical visual patterns [7].

To overcome these human limitations and reduce dependency on manual inspection, developing an automated and objective approach is highly required.

Fortunately, recent advances in artificial intelligence and computer vision offer promising solutions to this problem. These technologies have shown strong potential for pattern-based material classification by enabling automated and accurate visual data analysis. Applied to wood identification, artificial intelligence systems can enhance consistency, reliability, and classification efficiency [8-12]. Specifically, Convolutional Neural Networks and Transfer Learning have emerged as standard approaches for effectively extracting deep visual features [13].

Building upon these robust feature extraction capabilities, the literature presents a wide array of methodologies. Various algorithms for wood classification have been proposed, ranging from Support Vector Machines to deep learning approaches. However, most existing studies predominantly focus on datasets with highly contrasting visual features or macroscopic structural differences. A critical research gap remains in the classification of wood types with

highly homogeneous, subtle microscopic grain structures, a problem known as fine-grained classification. Standard methods often suffer severe accuracy degradation when applied to these visually ambiguous classes [14]. This challenge is further exacerbated when training data is scarce. When trained on limited datasets, standard deep learning architectures applied out of the box frequently suffer from severe performance degradation, defaulting to majority class predictions rather than learning the subtle directional flow of wood fibers [15]. Thus, addressing this limitation requires not just a powerful model, but an adaptive strategy that combines efficient feature extraction with advanced stochastic regularization techniques [16-19].

To realize this adaptive strategy, selecting an optimal classification framework is a critical decision, particularly when balancing deep architectural complexity with the inherent constraints of limited data availability. To address these challenges directly, this study proposes a comprehensive optimization framework using a fine-tuned InceptionV3 architecture. Unlike training from scratch, this method leverages pre learned features to capture complex grain patterns effectively [20], [21]. Furthermore, this study integrates strategic spatial data augmentation and rigorous regularization via Dropout and L2 penalties. Crucially, to explicitly handle inter-class visual similarities, this research incorporates Stochastic Gradient Descent with Nesterov momentum and Label Smoothing, a mechanism that prevents the model from making overconfident predictions on visually ambiguous textures [22-24].

By integrating targeted optimization techniques, this research advances industrial wood identification and overcomes constraints related to small datasets and fine-grained visual ambiguity. The study offers three primary contributions. First, it demonstrates the inadequacy of standard deep learning on homogeneous textures, evidenced by a ResNet50 baseline collapsing at 25 percent accuracy. Second, it resolves visual similarities between species like Jati and Mindi using Label Smoothing and Stochastic Gradient Descent momentum. Third, it validates a computationally efficient framework that prevents overfitting on a limited 800-image dataset, achieving a robust 89 percent accuracy. Ultimately, this study proves that strategic optimization can transform constrained data into a highly accurate and practical quality control solution for the timber industry.

While previous studies have successfully classified various wood species using Deep Learning, most have relied on large-scale datasets or species with highly distinct macroscopic features. This research fills the gap by specifically addressing the 'architectural failure' of standard models when applied to small-scale datasets characterized by high visual ambiguity among closely related species.

Specifically, this study makes three primary contributions to the field. First, an empirical identification of baseline failure is provided through an objective demonstration of why standard architectures like ResNet50 fail to capture subtle grain features in homogeneous textures, which inevitably leads to a 'random guess' performance of 25% accuracy. Second, a robust optimization framework is introduced by integrating Stochastic Gradient Descent (SGD) with Nesterov Momentum and Label Smoothing within the InceptionV3 architecture, effectively mitigating model overconfidence when processing visually ambiguous classes. Finally, a performance breakthrough on limited data is established by demonstrating that the proposed framework significantly improves classification accuracy from 25% to over 90%. This advancement ultimately offers a computationally efficient and reliable solution for timber identification, despite utilizing a highly constrained dataset of only 800 image samples.

2. Methods

This section details the systematic workflow employed to address the fine-grained wood classification problem. It outlines the dataset characteristics, the targeted augmentation strategies, the configuration of the proposed InceptionV3 architecture, and the comprehensive experimental setup used for evaluation.

2.1 Dataset and Preprocessing

The dataset used in this study comprises 800 color images of four distinct wood types: *Pterospermum javanicum* (Bayur), *Magnolia champaca* (Cempaka), *Tectona grandis* (Jati), and *Melia azedarach* (Mindi). The data distribution is perfectly balanced, with 200 images per class. To ensure a robust evaluation, the dataset was partitioned at an 80:20 ratio, yielding 160 training images and 40 test images per class. This balance is critical to preventing class bias during model optimization. Prior to entering the network, all images were resized to 299×299 pixels to match the native input specifications of the InceptionV3 architecture, ensuring that the microscopic details of the wood fibers were not degraded by excessive compression. The pixel values were then normalized to the range [0, 1]. To demonstrate the dataset's diversity, grain patterns, including variations in texture, coarseness, and orientation, are shown in Figure 1.



Figure 1. Sample Images from the Dataset

2.2 Data Augmentation Strategy

Given the limited dataset size and the fine-grained nature of the wood textures, a highly targeted data augmentation strategy was applied. Unlike standard augmentations that may distort natural features, the techniques were carefully constrained to preserve the directional flow and natural orientation of the wood grains [25].

Specifically, the applied augmentations included rotation restricted to $\pm 10^\circ$ to prevent altering the natural fiber direction, horizontal flipping to increase spatial invariance, and brightness adjustments with a factor between 0.8 and 1.2 to simulate real-world lighting variations.

2.3. Proposed Architecture

This study utilizes InceptionV3, pre-trained on ImageNet, as the backbone for feature extraction [26]. The choice of InceptionV3 is based on its ability to capture multi-scale features, which are essential for analyzing complex grain textures. During the network modification and fine-tuning process, a targeted unfreezing strategy was implemented. To adapt to the specific wood patterns without succumbing to catastrophic forgetting, only the top 30 layers of the InceptionV3 base were unfrozen. In contrast, the earlier layers were frozen to retain fundamental low-level feature extractors. Crucially, all Batch Normalization layers were kept frozen to maintain statistical stability during training. Subsequently, the original classification head was removed and replaced with a GlobalAveragePooling2D layer. Finally, to mitigate overfitting—a common challenge in small datasets—a high Dropout rate of 0.6 was applied after pooling. Additionally, L2 regularization ($\lambda=0.01$) was enforced on the kernel weights of the subsequent Dense layer.

The selection of InceptionV3 as the backbone architecture is strategically driven by its multi-scale feature extraction capabilities, which are critical for resolving fine-grained visual ambiguities in wood grain patterns. While more recent state-of-the-art architectures, such as Vision Transformers (ViT) and EfficientNet, have shown superior performance on massive datasets, their application to small-scale datasets remains technically challenging. ViTs in particular lack the inherent inductive biases such as translation equivariance and locality found in Convolutional Neural Networks (CNNs). Consequently, these models often suffer from poor convergence and severe overfitting when training data is limited, as they typically require large-scale pre-training to achieve competitive performance [27]. EfficientNet, although optimized for parameter efficiency, may not capture the simultaneous macro- and micro-textural variations as effectively as the parallel filter structures of Inception modules.

In contrast, the InceptionV3 architecture employs asymmetric and parallel convolutions, enabling the model to extract features at multiple scales

concurrently. This property is particularly advantageous for wood species such as *Tectona grandis* (Jati) and *Melia azedarach* (Mindi), where classification relies on identifying both global grain directions and localized pore structures. By integrating this architecture with SGD momentum and label smoothing, the proposed framework achieves a robust balance between model complexity and generalization.

2.4 Experimental Setup and Baseline Comparison

The research steps, from data collection to evaluation, is summarized in Figure 2.

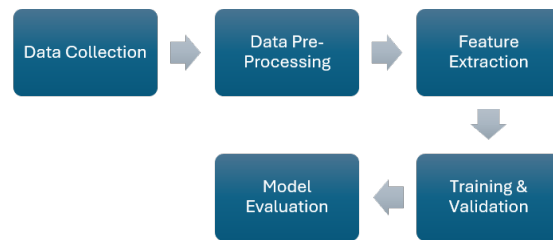


Figure 2. Steps of the Research Process

As illustrated in Figure 2, the workflow comprises five sequential stages. The process begins with (1) Data Collection, where primary data was acquired through direct photography of the four wood types. To ensure reproducibility and capture high-resolution textural details, the image acquisition process was systematically standardized. Photographs were taken using a Nikon D3200 DSLR camera featuring a 24.2-megapixel sensor, positioned at a constant distance of 30 cm from the wood surface. The imaging sessions were conducted using natural ambient sunlight between 09:00 and 10:00 AM to capture the authentic color and natural orientation of the wood fibers without artificial lighting distortion. Furthermore, the camera exposure settings were strictly controlled, utilizing a shutter speed of 1/125 seconds and an aperture of f/5.6 to ensure an optimal depth of field and clarity. This approach ensures that the dataset accurately reflects the natural visual characteristics and texture of the wood samples.

Following this, (2) Preprocessing was conducted to resize the raw images to 299×299 pixels and apply targeted spatial augmentations, preparing them for the next phase. In the (3) Feature Extraction stage, a fine-tuned InceptionV3 architecture was employed to capture discriminative grain patterns. To rigorously assess the contribution of this proposed framework, a standard ResNet50 model was also evaluated under identical conditions as a baseline for comparison. The proposed model and the baseline were implemented in Python using the TensorFlow deep learning framework. All computational experiments were executed on a cloud-based environment using Google Colaboratory, equipped with an NVIDIA Tesla T4 GPU (15 GB VRAM).

Subsequently, the (4) Model Training phase was executed. For the proposed InceptionV3 model, the training pipeline was heavily optimized for fine-grained

classification. The model was implemented using Stochastic Gradient Descent (SGD) with Nesterov momentum (0.9) and an initial learning rate of 0.005. To explicitly handle the visual ambiguity between classes, Label Smoothing (factor = 0.1) was integrated into the Categorical Cross-Entropy loss function. The training was configured for up to 50 epochs with a smaller batch size of 16 to allow for more frequent and stable weight updates, supported by dynamic callbacks (ReduceLROnPlateau and EarlyStopping) to prevent overfitting.

Finally, the pipeline concludes with (5) Model Evaluation, assessing both the baseline and the proposed system's performance on the strictly withheld unseen data to validate their generalization capability and discriminative power.

2.5 Performance Evaluation

To rigorously assess the proposed model's generalization capability relative to the baseline, evaluation was conducted on the strictly withheld test set. Performance was quantified using four standard metrics: Accuracy, Precision, Recall, and F1-Score. Given the multi-class nature of the task, specific attention was given to the F1-Score, which calculates the harmonic mean of Precision and Recall. This metric provides a more holistic view of the model's reliability, ensuring balanced performance across all wood types.

Complementing these scalar metrics, Confusion Matrices were constructed to visualize the alignment between predicted labels and actual ground truth for both the baseline and the proposed model. This matrix serves as a critical diagnostic tool to identify specific misclassification patterns, pinpointing exact sources of ambiguity (such as whether "Mindi" is systematically confused with "Jati"), thereby evaluating the model's discriminative power more granularly than global accuracy alone.

3. Results and Discussions

This section presents the empirical findings of the study, beginning with an analysis of the baseline model's limitations. It subsequently evaluates the performance of the proposed optimized framework, including a detailed error analysis, an assessment of training stability, and a discussion on current limitations for future improvements.

3.1 Baseline Failure and the Challenge of Visual Ambiguity

To establish the necessity of the proposed research, an initial experiment was conducted using a standard ResNet50 architecture as a baseline. The baseline evaluation revealed the extreme complexity of fine-grained wood texture classification. As hypothesized, the standard CNN failed to extract discriminative features from highly homogeneous wood grains, resulting in severe model collapse. The baseline model classified 100 percent of the validation instances as the

Bayur class, yielding a random-guess accuracy of merely 25 percent.

To rule out the possibility that this failure stemmed from an under-optimized training configuration, the ResNet50 model was rigorously re-evaluated under the same hyperparameter settings as the proposed InceptionV3 framework. This included applying transfer learning (fine-tuning the last 20 layers), targeted data augmentation, High Dropout (0.5), an SGD optimizer with Nesterov momentum, and Label Smoothing (0.1). Despite these optimizations, the tuned ResNet50 baseline achieved only a 43% overall accuracy. Furthermore, the baseline completely failed to recognize specific classes, recording a precision and recall of 0.00 for both the Bayur and Mindi classes. This finding confirms that the significant performance gap is not due to an under-optimized baseline, but rather highlights the intrinsic limitation of the ResNet50 architecture in resolving the subtle, fine-grained visual ambiguities present in this specific wood dataset.

Indeed, such a failure points to a broader structural vulnerability. Standard deep learning architectures, optimized for macroscopic object detection, struggle to distinguish microscopic and high-frequency textural variations. Specifically, standard architectures like ResNet50 rely heavily on aggressive pooling operations that inherently discard microscopic spatial details in favor of identifying global object boundaries. Since wood grain lacks a definitive global shape and relies entirely on continuous textural flow, this architectural bias leads to catastrophic feature loss. When faced with visually ambiguous classes like Jati and Mindi, standard networks prioritize broad color spaces over subtle grain flow, leading to severe overfitting on the majority class. This empirically proves that applying deep learning out of the box is insufficient for the timber industry, necessitating a specialized optimization strategy.

3.2 Performance of the Proposed Optimization Framework

Addressing the severe limitations of the baseline, the proposed fine-tuned InceptionV3 framework achieved a significant performance leap. The Inception modules' multiscale feature extraction capabilities enabled the network to simultaneously capture both macroscopic color distributions and microscopic fiber orientations, a dual perception that the baseline model lacked. The quantitative evaluation confirms that the optimized architecture achieved an overall accuracy of 89 percent, marking a 64 percent improvement over the baseline. A granular breakdown of class-level performance is detailed in Table 1.

Table 1. Classification Report of the Proposed Model

	Precision	Recall	F1 Score
Bayur	1.00	0.80	0.89
Cempaka	0.91	0.97	0.94
Jati	0.93	0.93	0.90
Mindi	0.79	0.85	0.82
Macro Average	0.89	0.89	0.89

These compelling metrics prompt a deeper investigation into the underlying factors driving such performance. The success of this proposed framework is not merely a result of using a deeper network, but rather the synergistic effect of three targeted interventions. The first intervention involved a meticulously refined augmentation strategy. In initial trials, aggressive augmentations destroyed the natural directional flow of the wood fibers, confusing the model. By restricting rotation to 10 degrees and using only horizontal flipping, the proposed framework successfully preserved the natural structural orientation of the wood grain while enriching dataset variability. Furthermore, this restricted augmentation approach ensured that the model learned the inherent biological structure of the wood rather than memorizing artificial distortions. With the input data properly structured to preserve natural features, the next critical step was optimizing how the network extracted these patterns.

The second significant improvement was the adoption of Stochastic Gradient Descent with Nesterov momentum. While adaptive optimizers like Adam are popular, they frequently converge into suboptimal local minima when dealing with highly homogeneous textures. Adam frequently struggles with generalization in fine-grained tasks because its adaptive learning rates can rapidly shrink for frequently updated weights, causing the model to get stuck in local minima. By employing Stochastic Gradient Descent with Nesterov momentum and a controlled learning rate of 0.005, the model achieved a stable momentum-guided traversal of the complex loss landscape. Consequently, the momentum-driven mechanism enabled the network to learn deeper features without prematurely updating weights based on noisy gradients. While the optimizer ensured a stable learning trajectory, the loss calculation itself required adjustment to handle the inherent ambiguity of the wood classes.

Finally, the third and most critical intervention was the application of Label Smoothing to boost performance. Applying Label Smoothing with a factor of 0.1 effectively calibrates the model's predictive confidence during training. Standard Categorical Cross-Entropy encourages models to be overly confident, which is detrimental when classifying visually similar species such as Jati and Mindi. By slightly softening the target distribution, this technique penalizes overconfidence and prevents the network from memorizing specific training instances. Consequently, this mechanism forces the InceptionV3 filters to look beyond superficial similarities and extract secondary microscopic textural descriptors to make an accurate decision. While these architectural enhancements produced exceptional initial metrics, it was imperative to verify that such performance was reproducible.

To further validate the statistical robustness of the proposed framework and ensure the results are not anomalies of a single random data split, the optimized InceptionV3 model was evaluated across multiple

independent training runs using different random initialization seeds. Over these independent runs, the model demonstrated highly stable performance, achieving a mean testing accuracy of 91.46 percent with a standard deviation of 1.56 percent. This consistency, coupled with the rigorous data augmentation applied during training, mathematically demonstrates the model's robust generalization and effectively mitigates concerns about overfitting on this specific fine-grained dataset.

3.3 Error Analysis and Confusion Matrix

To provide a concrete diagnosis of the classification behavior, a Confusion Matrix was constructed as shown in Figure 3.

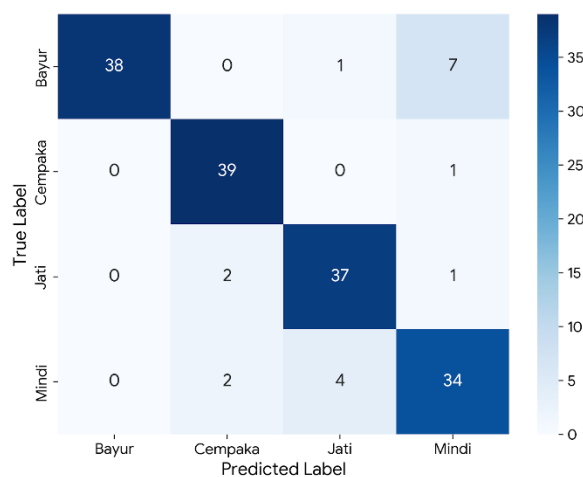


Figure 3. Confusion Matrix of the Proposed Model

The matrix visually corroborates the metrics in Table 1. The model exhibits exceptional robustness in identifying Cempaka, achieving a high Recall of 0.97. This exceptional performance on Cempaka can be attributed to its relatively distinct biological fiber arrangement and color consistency, which the optimized model easily isolated. Bayur also demonstrates a perfect Precision of 1.00, proving that the SGD optimizer successfully mapped its specific fiber orientations without confusing its dominant grain boundaries with other species.

In contrast to the distinct features of Cempaka and Bayur, analyzing wood types with highly uniform textures presents a different set of challenges. The Mindi class, while vastly improved from the baseline, remains a challenging category with an F1 Score of 0.82. The matrix reveals that the errors are highly localized, with some Bayur samples leaking into Mindi. This is expected, as Mindi wood often lacks strong, high-contrast grain lines, making it visually overlap with the broader decision boundaries of the other classes. From a biological and visual perspective, Mindi's uniform, porous surface is highly susceptible to misclassification when subtle variations in lighting or cutting angles blur its natural boundaries during image capture. However, achieving this level of Mindi performance on a constrained dataset validates the

effectiveness of the Label Smoothing technique in handling extreme inter-class visual similarities.

3.4 Training Stability and Convergence

The learning stability of the model is illustrated in the accuracy and loss curves as shown in Figures 4 and 5.

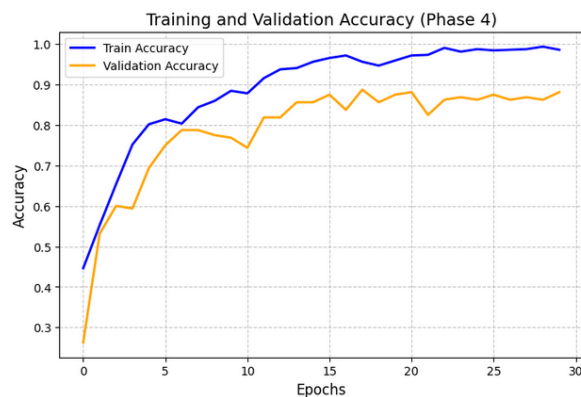


Figure 4. Training and Validation Accuracy of the Proposed Model

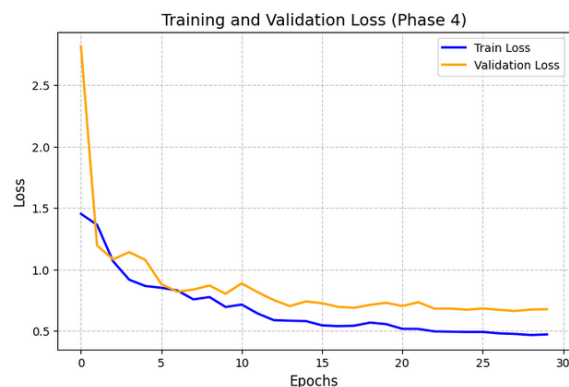


Figure 5. Training and Validation Loss of the Proposed Model

The model demonstrated stable convergence, achieving a training accuracy of nearly 0.99 and stabilizing validation accuracy around 0.89. An analysis of the early training phase reveals a rapid ascent in accuracy during the first few epochs, followed by a steady, controlled fine-tuning phase. This trajectory confirms that the initial learning rate was optimally configured for the complex loss landscape. Crucially, the gap between training and validation loss remains controlled, demonstrating that integrating a high Dropout rate and L2 Regularization successfully prevents the model from memorizing the limited dataset, ensuring true generalization. The application of a high Dropout rate of 0.6 proved highly effective; by randomly deactivating 60 percent of the neurons during each training step, the network was forced to learn redundant representations of the wood grains, preventing it from relying on a few dominant pixels.

3.5 Limitation and Future Work

While the proposed framework demonstrates robust classification capabilities, this study identifies inherent constraints that provide avenues for future refinement. The first limitation lies in the reliance on the

InceptionV3 architecture. Since this model is initially optimized for general object recognition, it may still prioritize global structural features over ultra-fine local textures. Future development, leveraging significantly larger datasets, will explore Vision Transformers or texture-specific descriptors that better capture micro-texture characteristics. Vision Transformers process images as sequences of patches and use self-attention mechanisms. This architectural shift could allow the model to selectively focus on the most discriminative regions of the wood grain, such as specific knot patterns or unique fiber intersections, without losing spatial context. Beyond these architectural considerations, the scalability of this research is fundamentally tied to the size of the image corpus.

The second constraint concerns data availability. The current dataset is strictly limited to 800 images. Although the refined augmentation and Label Smoothing strategies successfully mitigated overfitting and achieved statistical stability, with a mean accuracy of 91.46 percent and a standard deviation of 1.56 percent across multiple independent runs, the scarcity of practical industrial samples limits the model's exposure to the full spectrum of grain variations. To overcome this scalability bottleneck, the immediate research trajectory will focus on synthetic data generation. Specifically, the use of Generative Adversarial Networks will be explored to synthesize realistic wood grain patterns. Advanced architectures such as StyleGAN could be employed to generate synthetic wood variations that mimic varying industrial conditions, including different moisture levels, sanding depths, and lighting environments. This integration would massively enrich the training distribution and provide a highly scalable solution for the timber industry.

Beyond the challenge of dataset volume, the environmental consistency during image acquisition presents a final limitation. The dataset was acquired under strictly controlled lighting conditions using natural sunlight between 09:00 and 10:00 AM at a fixed distance of 30 cm. Practical industrial environments, however, often present significant domain shifts, such as inconsistent artificial lighting, dust, and variable camera angles, which may affect robustness across different environments. Consequently, transitioning this framework to practical applications requires future data collection directly from active sawmill environments to guarantee generalization.

4. Conclusions

This study successfully addresses the critical challenge of fine-grained visual ambiguity in automated wood-grain classification. This pervasive issue frequently compromises quality control mechanisms within the timber industry. The classification of highly homogeneous wood species, particularly the visually overlapping Jati and Mindi classes, poses a complex pattern recognition problem that is further exacerbated

by small-scale datasets. To establish a rigorous comparative foundation, initial baseline evaluations were conducted using a standard ResNet50 architecture. The results revealed a severe performance degradation, yielding a random-guess accuracy of only 25 percent and an inability to distinguish beyond a single dominant class. This major limitation highlights the fundamental inadequacy of applying off-the-shelf deep learning models to limited and fine-grained textures without specialized interventions. To explicitly resolve this limitation, this research proposed and validated a comprehensively optimized InceptionV3 framework. Rather than relying on standard adaptive optimizers, the training pipeline was heavily fortified with targeted spatial augmentations, a high Dropout rate, and L2 regularization to combat overfitting. Crucially, the optimization strategy integrated Stochastic Gradient Descent with Nesterov momentum and Label Smoothing. This specific combination successfully penalized overconfident predictions and forced the network filters to extract deeper, microscopic textural descriptors rather than superficial color features. Consequently, the proposed model achieved a robust overall accuracy of 89 percent, marking a substantial 64 percent improvement over the baseline, alongside a macro-averaged F1 Score of 0.89. These empirical findings clearly indicate that specific architectural fine-tuning and advanced stochastic regularization are essential for developing reliable, scalable, and automated classification systems capable of handling highly ambiguous natural materials. By successfully decoupling underlying grain textures from incidental capture conditions, this study provides an efficient quality control solution for the timber industry. Future work will prioritize expanding dataset diversity through synthetic data generation using Generative Adversarial Networks and exploring the integration of attention mechanisms to enhance further classification robustness across a broader spectrum of commercial wood species in highly variable industrial lighting environments.

Acknowledgements

The author declares no conflict of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The author acknowledges the University of Technology Yogyakarta for providing support with laboratory equipment and the students who assisted with primary data collection.

Data Availability Statement

The dataset supporting the findings of this study is available upon reasonable request from the corresponding author. All essential implementation details, including the specific hyperparameters and the experimental environment settings, are thoroughly documented in Section 2.4 to facilitate the replication of the results by other researchers.

References

- [1] E. Korkmaz, C. Özcan, and M. Korkmaz, "An application of fuzzy soft sets to a real-life problem: Classification of wood materials to prevent fire-related injuries and deaths," *Appl. Soft Comput.*, vol. 132, p. 109875, Jan. 2023, doi: [10.1016/j.asoc.2022.109875](https://doi.org/10.1016/j.asoc.2022.109875).
- [2] J. Zhu, H. Peng, X. Lu, J. Lyu, and T. Zhan, "Classification of Principal Wood Species in China Based on the Physiomechanical Properties," *J. Renew. Mater.*, vol. 11, no. 3, pp. 1425–1437, Oct. 2022, doi: [10.32604/JRM.2022.023464](https://doi.org/10.32604/JRM.2022.023464).
- [3] W. Gan et al., "Enhanced Recognition of Sustainable Wood Building Materials Based on Deep Learning and Augmentation," *Sustainability* 2025, Vol. 17, vol. 17, no. 15, Jul. 2025, doi: [10.3390/su17156683](https://doi.org/10.3390/su17156683).
- [4] F. Fang, Q. Xu, Y. Cheng, Y. Sun, and J. H. Lim, "Image Understanding With Reinforcement Learning: Auto-Tuning Image Attributes and Model Parameters for Object Detection and Segmentation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 10, pp. 6671–6685, Oct. 2022, doi: [10.1109/TCSVT.2022.3171781](https://doi.org/10.1109/TCSVT.2022.3171781).
- [5] A. Pratondo and A. Novianty, "Comparison of Wood Classification using Machine Learning," *2022 IEEE 10th Conference on Systems, Process and Control, ICSPC 2022 - Proceedings*, pp. 308–312, 2022, doi: [10.1109/ICSPC55597.2022.10001776](https://doi.org/10.1109/ICSPC55597.2022.10001776).
- [6] H. Ergun, "Wood identification based on macroscopic images using deep and transfer learning approaches," *PeerJ*, vol. 12, p. e17021, Feb. 2024, doi: [10.7717/PEERJ.17021/SUPP-1](https://doi.org/10.7717/PEERJ.17021/SUPP-1).
- [7] R. Ehtisham et al., "Classification of defects in wooden structures using pre-trained models of convolutional neural network," *Case Studies in Construction Materials*, vol. 19, p. e02530, Dec. 2023, doi: [10.1016/j.cscm.2023.E02530](https://doi.org/10.1016/j.cscm.2023.E02530).
- [8] S. Salvi and V. Geetha, "A Nested Texture Inspired Novel Image Pattern Based Optical Camera Communication," *IEEE Access*, vol. 10, pp. 109056–109067, 2022, doi: [10.1109/ACCESS.2022.3213701](https://doi.org/10.1109/ACCESS.2022.3213701).
- [9] D. A. Abdulqader, M. S. Hathal, B. M. Mahmmod, S. H. Abdulhussain, and D. Al-Jumeily, "Plain, Edge, and Texture Detection Based on Orthogonal Moment," *IEEE Access*, vol. 10, pp. 114455–114468, 2022, doi: [10.1109/ACCESS.2022.3217225](https://doi.org/10.1109/ACCESS.2022.3217225).
- [10] E. S. Gemeay, F. A. Alenizi, A. H. Mohammed, M. H. Shakoor, and R. Boostani, "Weighted Constraint Feature Selection of Local Descriptor for Texture Image Classification," *IEEE Access*, vol. 11, pp. 91673–91695, 2023, doi: [10.1109/ACCESS.2023.3306075](https://doi.org/10.1109/ACCESS.2023.3306075).
- [11] Z. Tian, N. Zhang, J. Wang, L. Sha, H. Liu, and L. Zou, "Image Classification of Imbalanced Wood Microscope by Integrating Multi-source Features," *International Journal of Computational Intelligence Systems* 2025 18:1, vol. 18, no. 1, pp. 1–16, Jul. 2025, doi: [10.1007/S44196-025-00912-6](https://doi.org/10.1007/S44196-025-00912-6).
- [12] D. Mane et al., "A Novel Fuzzy Hypersphere Neural Network Classifier using Class Specific Clustering for Robust Pattern Classification," *IEEE Access*, 2024, doi: [10.1109/ACCESS.2024.3454296](https://doi.org/10.1109/ACCESS.2024.3454296).
- [13] K. M. Zielinski et al., "Advanced wood species identification based on multiple anatomical sections and using deep feature transfer and fusion," *Comput. Electron. Agric.*, vol. 231, p. 109867, Apr. 2025, doi: [10.1016/j.compag.2024.109867](https://doi.org/10.1016/j.compag.2024.109867).
- [14] Q. Wang, Y. Dong, N. Xu, F. Xu, C. Mou, and F. Chen, "Image Classification of Tree Species in Relatives Based on Dual-Branch Vision Transformer," *Forests* 2024, Vol. 15, vol. 15, no. 12, Dec. 2024, doi: [10.3390/f15122243](https://doi.org/10.3390/f15122243).
- [15] Y. Wang, Z. P. Bian, J. Hou, and L. P. Chau, "Convolutional Neural Networks with Dynamic Regularization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 5, pp. 2299–2304, May 2021, doi: [10.1109/TNNLS.2020.2997044](https://doi.org/10.1109/TNNLS.2020.2997044).
- [16] L. P. Yi, M. F. Akbar, M. N. A. Wahab, B. A. Rosdi, M. A. Fauthan, and N. H. M. M. Shrifan, "The Prospect of Artificial Intelligence-Based Wood Surface Inspection: A Review," *IEEE Access*, vol. 12, pp. 84706–84725, 2024, doi: [10.1109/ACCESS.2024.3412928](https://doi.org/10.1109/ACCESS.2024.3412928).
- [17] R. Confalonieri, P. P. Htun, B. Sun, and T. Tillo, "An End-to-End Framework for the Classification of Hyperspectral Images

- in the Wood Domain,” *IEEE Access*, vol. 12, pp. 38908–38916, 2024, doi: [10.1109/ACCESS.2024.3376258](https://doi.org/10.1109/ACCESS.2024.3376258).
- [18] T. H. Wong, M. F. Soong, A. Saifizul, and R. Ramli, “Assessment of wood quality for Sape making: Vibroacoustic analysis and machine learning classification,” *Expert Syst. Appl.*, vol. 261, p. 125591, Feb. 2025, doi: [10.1016/j.eswa.2024.125591](https://doi.org/10.1016/j.eswa.2024.125591).
- [19] İ. Kırbaş and A. Çifci, “An effective and fast solution for classification of wood species: A deep transfer learning approach,” *Ecol. Inform.*, vol. 69, p. 101633, Jul. 2022, doi: [10.1016/j.ecoinf.2022.101633](https://doi.org/10.1016/j.ecoinf.2022.101633).
- [20] T. Hidayat, I. A. Astuti, A. Yaqin, A. P. Tjilen, and T. Arifianto, “Grouping of Image Patterns Using Inceptionv3 For Face Shape Classification,” *JOIV: International Journal on Informatics Visualization*, vol. 7, no. 4, pp. 2336–2343, Dec. 2023, doi: [10.62527/joiv.7.4.1743](https://doi.org/10.62527/joiv.7.4.1743).
- [21] E. Mengiste, K. R. Mannem, S. A. Prieto, and B. Garcia de Soto, “Transfer-Learning and Texture Features for Recognition of the Conditions of Construction Materials with Small Data Sets,” *Journal of Computing in Civil Engineering*, vol. 38, no. 1, p. 04023036, Jan. 2024, doi: [10.1061/jccce5.cpeng-5478](https://doi.org/10.1061/jccce5.cpeng-5478).
- [22] R. K. Ahalya, U. Snekhalatha, and V. Dhanraj, “Automated segmentation and classification of hand thermal images in rheumatoid arthritis using machine learning algorithms: A comparison with quantum machine learning technique,” *J. Therm. Biol.*, vol. 111, p. 103404, Jan. 2023, doi: [10.1016/j.jtherbio.2022.103404](https://doi.org/10.1016/j.jtherbio.2022.103404).
- [23] J. Zheng et al., “An improved automatic image labeling and classification algorithm for multi-mode damage quantification of 2.5D woven composites based on deep learning strategy,” *Compos. Sci. Technol.*, vol. 259, p. 110932, Jan. 2025, doi: [10.1016/j.compscitech.2024.110932](https://doi.org/10.1016/j.compscitech.2024.110932).
- [24] H. Ren, Y. Zhao, Y. Zhang, and W. Sun, “Learning label smoothing for text classification,” *PeerJ Comput. Sci.*, vol. 10, p. e2005, Apr. 2024, doi: [10.7717/PEERJ-CS.2005](https://doi.org/10.7717/PEERJ-CS.2005).
- [25] T. H. Cheung and D. Y. Yeung, “A Survey of Automated Data Augmentation for Image Classification: Learning to Compose, Mix, and Generate,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 13185–13205, 2024, doi: [10.1109/TNNLS.2023.3282258](https://doi.org/10.1109/TNNLS.2023.3282258).
- [26] S. Sharma, K. Guleria, S. Kumar, and S. Tiwari, “Deep Learning based Model for Detection of Vitiligo Skin Disease using Pre-trained Inception V3,” *Deep Learning based Model for Detection of Vitiligo Skin Disease using Pre-trained Inception V3*, vol. 8, no. 5, pp. 1024–1039, Jul. 2023, doi: [10.33889/IJMEMS.2023.8.5.059](https://doi.org/10.33889/IJMEMS.2023.8.5.059).
- [27] S. Lee, S. Lee, and B. C. Song, “Improving Vision Transformers to Learn Small-Size Dataset From Scratch,” *IEEE Access*, vol. 10, pp. 123212–123224, 2022, doi: [10.1109/ACCESS.2022.3224044](https://doi.org/10.1109/ACCESS.2022.3224044).