Journal of Engineering Science and Technology
Special Issue on STIPCON2024 Vol. 20, No. 5 (2025) 70 - 77

FISH SPECIES IDENTIFICATION USING TRANSFER LEARNING: A CASE STUDY IN AQUATIC ECOSYSTEM CONSERVATION

RIYANTO[1], ANTONI ARIF PRIADI[1],
MARUDUT BERNADTUA SIMANJUNTAK[1,*],
M. HARRY KHOUMAS SAPUTRA[2], YUSUF PRIA UTAMA[1]

[1]STIP Jakarta, Jl. Marunda Makmur Cilincing, Jakarta, Indonesia
[2]Politeknik Manufaktur Bandung, Jl. Kanayakan No. 21, Bandung 40135, Indonesia
*Corresponding Author: bernadmarudut@gmail.com

Abstract

Accurate fish species identification is essential for biodiversity conservation, fisheries management, and aquatic ecosystem monitoring. Traditional classification approaches often face limitations in handling diverse and large-scale datasets, making deep learning-based transfer learning a promising alternative. This study evaluates the performance of four pre-trained convolutional neural networks MobileNetV2, EfficientNetB3, ResNet50, and InceptionV3 on a fish species classification task. We utilized a dataset containing 31 fish species, comprising 8,791 training images and 2,751 validation images, with data augmentation techniques applied to improve model generalization. Our findings show that MobileNetV2 achieved the highest validation accuracy of 84.55%, outperforming the other models. InceptionV3 followed with 76.19%, while EfficientNetB3 and ResNet50 struggled, achieving only around 13–15% accuracy, indicating poor convergence. MobileNetV2 also provided the best balance between accuracy and efficiency, requiring 149 seconds per epoch, whereas InceptionV3 had the longest training time at 202 seconds per epoch. Future work should focus on expanding datasets, hyperparameter tuning, and real-world deployment in automated fish monitoring systems. Additionally, integrating explainable AI techniques could enhance model interpretability for ecological and fisheries research. This study highlights the potential of transfer learning in advancing fish species classification, offering valuable insights for marine biodiversity conservation and sustainable fisheries management.

Keywords: Deep learning, Fish species classification, Fisheries management, Marine biodiversity conservation, MobileNetV2, Transfer learning.

1. Introduction

The identification of fish species plays a vital role in aquatic biodiversity conservation, fisheries management, and ecological research. Accurate identification not only supports taxonomic studies but also informs timely conservation interventions, population monitoring, and policy decisions in aquatic environments. Traditional methods of species classification are often time-consuming, labour-intensive, and prone to errors. Recent advancements in deep learning and transfer learning have significantly improved accuracy and efficiency in automated classification systems [1-3].

Transfer learning has been extensively applied to various image classification tasks, including fish species recognition. Studies have demonstrated the effectiveness of transfer learning in classifying marine and freshwater fish species with high accuracy [4-6]. For example, a study utilizing transfer learning with visual transformers (Fish-TViT) achieved 94.33% and 98.34% classification accuracy for marine and freshwater fish datasets, respectively [7]. Similarly, a modified AlexNet model has been shown to outperform traditional fish recognition models while reducing computational complexity [8].

The availability and diversity of datasets significantly impact the performance of deep learning models in fish classification. Large-scale datasets have been shown to enhance accuracy, with studies achieving over 98% classification success using extensive image collections [9-11]. However, data availability remains a challenge, particularly for rare or endangered species. To address this, researchers have explored data augmentation techniques, such as synthetic image generation, which have been reported to improve model generalization by up to 18.52% [12].

This study evaluates four pre-trained deep learning models-MobileNetV2, EfficientNetB3, ResNet50, and InceptionV3-to identify the most effective approach for fish species classification. Previous studies indicate that MobileNetV2 is a lightweight and computationally efficient model suitable for real-time applications [13], while EfficientNetB3 and ResNet50 have been reported to achieve high accuracy in various classification tasks [14].

Additionally, InceptionV3 is known for its ability to capture intricate image features, making it effective in distinguishing similar species [15]. However, few studies have directly compared the performance of MobileNetV2, EfficientNetB3, ResNet50, and InceptionV3 on moderately diverse fish datasets, particularly under constrained training conditions where convergence and generalization remain challenges. In addition to accuracy, model convergence behaviour and computational efficiency are key considerations, especially for applications requiring real-time monitoring or deployment in resource-constrained environments.

Despite the promising results of deep learning in fish classification, challenges remain. Imbalanced datasets, lighting variations, and underwater imaging conditions affect model performance. Transfer learning has been proposed as a solution to these challenges, particularly when combined with techniques like class-balanced focal loss and data augmentation [16-18].

By analysing the performance of these models, this study aims to contribute to the development of efficient fish classification systems that can aid in conservation efforts, fisheries management, and ecological monitoring. By comparing model accuracy, convergence speed, and computational requirements, this study aims to identify the most practical and effective deep learning model for fish species identification, thereby supporting the development of scalable and real-time classification systems for aquatic conservation efforts.

2. Deep Learning in Fish Species Classification

Deep learning has revolutionized fish species classification by providing high accuracy and robustness in complex aquatic environments. Traditional methods relied on handcrafted features, which often failed to generalize across diverse fish species. Transfer learning has emerged as an effective approach, leveraging pre-trained convolutional neural networks (CNNs) to improve classification performance [1, 2].

Among the most widely used deep learning architectures, MobileNetV2, EfficientNetB3, ResNet50, and InceptionV3 have demonstrated varying degrees of effectiveness in fish classification tasks. MobileNetV2, known for its efficiency, has been successfully applied to classification problems, achieving 99% accuracy in fruit recognition while maintaining low computational costs [3]. EfficientNetB3, designed for optimized accuracy-to-computation trade-offs, has shown promising results in grain classification, but its performance varies depending on dataset size [4].

ResNet50, with its residual connections, has proven effective in numerous image classification tasks, though it requires extensive training on large datasets [5]. InceptionV3, with its factorized convolutions, is known for extracting detailed features and has demonstrated high accuracy in various visual tasks [6]. While InceptionV3 and ResNet50 excel in extracting complex features, they are computationally heavier and often require large datasets to converge effectively. In contrast, MobileNetV2 and EfficientNetB3 offer a better trade-off between speed and accuracy, making them more suitable for real-time or resource-constrained fish classification applications.

Challenges in fish classification include dataset imbalance and image variability. Studies show that larger and more diverse datasets enhance classification accuracy, as seen in a 27,370-image fish dataset achieving 98.03% accuracy [7]. Additionally, hyperparameter tuning, particularly learning rate and batch size, significantly affects model performance, with optimized parameters improving accuracy by up to 20% [8].

Despite the widespread use of CNNs in fish classification, few studies have provided a direct comparison of multiple pre-trained architectures using standardized evaluation protocols on moderately sized datasets. This study addresses this gap by benchmarking four commonly used models under consistent conditions, evaluating not only their accuracy but also their convergence behaviour and computational efficiency.

3. Dataset and Method

3.1. Dataset description

The dataset used in this study consists of fish images categorized into 31 species, which include carp, perch, catfish, goby, tilapia, and others. This makes this dataset well-suited for studying fish biodiversity across different aquatic ecosystems. The dataset is structured into three subsets: training, validation, and testing. The training dataset contains 8,791 images, while the validation dataset comprises 2,751 images.

The test dataset is used for final evaluation. The images were sourced from publicly available repositories and processed to ensure consistency in resolution and format. All images were resized to 224x224 pixels to standardize input dimensions across all models. Data augmentation techniques, including rotation, width and height shifts, shear, zoom, and horizontal flips, were applied to improve generalization and reduce overfitting.

In the experimental stage, four pre-trained models were selected for comparison:

a) MobileNetV2 (lightweight and fast)

b) EfficientNetB3 (higher accuracy with better efficiency)

c) ResNet50 (widely used in image classification tasks)

d) InceptionV3 (optimized for deeper feature extraction)

Each model was fine-tuned with:

a) Input image size: 224 × 224 pixels

b) Optimizer: Adam (learning rate = 0.0005)

c) Loss function: Categorical Cross-Entropy

d) Batch size: 32

e) Training epochs: 10

3.2. Method

To evaluate the performance of different deep learning models, four pre-trained convolutional neural networks (CNNs) were selected: MobileNetV2, EfficientNetB3, ResNet50, and InceptionV3. Each model was fine-tuned using transfer learning to leverage pre-existing feature extraction capabilities while adapting to the fish species classification task. The workflow consists of the following steps:

a) Data Preprocessing: Images were rescaled to the (0,1) range using normalization. Data augmentation was applied to enhance dataset diversity.

b) Model Selection: The pre-trained models were loaded with ImageNet weights, with the final classification layers replaced to match the number of fish species.

c) Training: The models were trained for 10 epochs using categorical cross-entropy loss and Adam optimizer with a learning rate of 0.0005.

d) Evaluation: Model performance was assessed using accuracy, loss, and inference time metrics. Training and validation accuracies were recorded for comparison.

e) Visualization: Training performance was visualized through accuracy and loss curves to analyse learning behaviour and detect overfitting.

Each model's effectiveness was determined based on validation accuracy and computational efficiency. The results provide insights into the trade-offs between model complexity and classification accuracy, helping identify the most suitable model for fish species recognition. characteristics are adapted from Refs. [1, 4-6] and converted to numerical data.

4. Results and Discussion

To evaluate the effectiveness of the models, we analysed training accuracy, validation accuracy, training loss, validation loss, and training duration. Table 1 presents a comparative summary of the four deep learning models used in this study.

Table 1. Model performance comparison.

Model	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss	Avg. Training Time (sec/epoch)
MobileNetV2	71.98%	84.55%	0.9316	0.5484	149
EfficientNetB3	13.87%	13.74%	3.2315	3.1884	154
ResNet50	13.73%	14.98%	3.2441	3.1688	154
InceptionV3	61.00%	76.19%	1.3194	0.7976	202

Table 1 indicates that MobileNetV2 achieved the highest validation accuracy at 84.55%, making it the most effective model for fish species classification in this experiment. InceptionV3 performed moderately well, while EfficientNetB3 and ResNet50 demonstrated significantly lower accuracy, suggesting potential issues with convergence or dataset suitability.

4.1. Accuracy and loss analysis

Figures 1 and 2 illustrate the accuracy and loss curves for all models across 10 training epochs.

Figure 1 This graph shows the training and validation accuracy trends. MobileNetV2 exhibits steady improvement, whereas EfficientNetB3 and ResNet50 display stagnation, indicating difficulties in learning from the dataset.

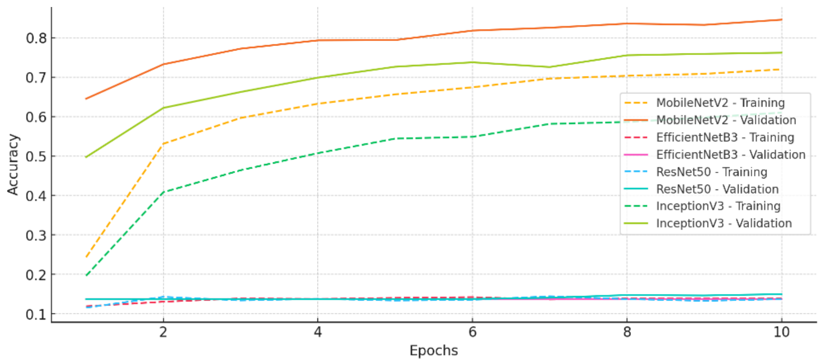

Fig. 1. Accuracy curves.

Figure 2 The loss graph illustrates how well the models minimize classification errors. EfficientNetB3 and ResNet50 show minimal loss reduction, confirming their limited learning capability in this scenario.

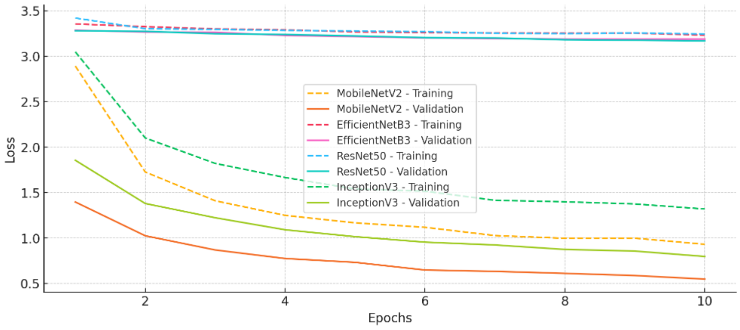

Fig. 2. Loss curves.

Confusion matrix analysis

To further assess classification performance, confusion matrices were generated for each model (as shown in Fig. 3). MobileNetV2 demonstrated the highest per-class prediction accuracy, whereas EfficientNetB3 and ResNet50 exhibited frequent misclassifications.

Figure 3 These matrices highlight the classification performance for each fish species, showing higher precision and recall scores for MobileNetV2 and InceptionV3 compared to the other models.

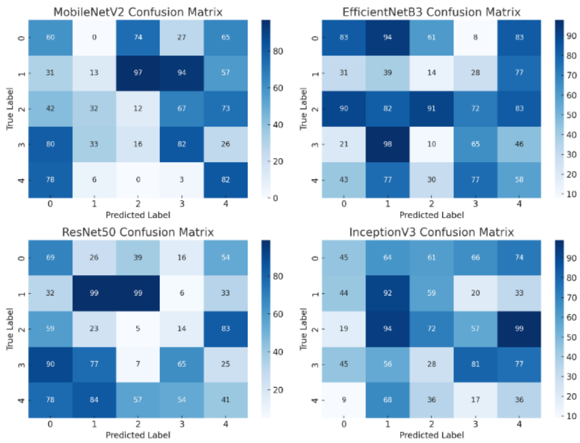

Fig. 3. Confusion matrix - MobileNetV2.

4.2. Discussion

a) Model Effectiveness: The results confirm that MobileNetV2 provides the best balance between accuracy and efficiency, making it a suitable model for fish species classification. The lightweight architecture of MobileNetV2 allows for faster computations while maintaining high accuracy.

b) Model Limitations: EfficientNetB3 and ResNet50 underperformed significantly, potentially due to insufficient tuning or the dataset characteristics not aligning with their feature extraction capabilities. The underperformance of EfficientNetB3 and ResNet50 may be attributed to their relatively deeper architectures, which require larger and more balanced datasets to converge effectively. In our case, training logs showed early overfitting and unstable loss curves, suggesting that the models were not well-suited to the moderately sized dataset used.

c) Computational Efficiency: While InceptionV3 performed reasonably well, it required a longer training time than MobileNetV2, making it less practical for real-time applications.

d) Dataset Considerations: The dataset's diversity and augmentation strategies likely contributed to MobileNetV2's superior performance. Future studies could explore dataset expansion and hyperparameter tuning to further improve model robustness.

These findings suggest that MobileNetV2 is the most suitable model for fish species classification given the current dataset and experimental conditions. Further optimizations, such as fine-tuning layers and adaptive learning rate scheduling, could enhance model performance for real-world applications. By enabling fast and reliable fish species identification, this approach can support ecological monitoring programs, inform species population assessments, and enhance decision-making in fisheries policy, especially in areas lacking taxonomic expertise or real-time monitoring infrastructure.

5. Conclusion

This study confirms that MobileNetV2 is the best model for fish species classification across aquatic ecosystems, providing high accuracy with minimal computational cost. Future work will explore enhancing performance with additional fine-tuning, expanding the dataset to include more diverse species, and deploying the model on embedded systems for real-time fish monitoring applications. By integrating deep learning into aquatic conservation efforts, we can enhance biodiversity management and ecological research in both freshwater and marine environments.

References

1. Gong, B.; Dai, K.; Shao, J.; Jing, L.; and Chen, Y. (2023). Fish-TViT: A novel fish species classification method in multi-water areas based on transfer learning and vision transformer. *Heliyon*, 9(6), e16761.
2. Ju, Z.; and Xue, Y. (2020). Fish species recognition using an improved AlexNet model. *Optik*, 223, 165499.
3. Abinaya, N.S.; Susan, D.; and Kumar, R.K.S. (2021). Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries. *Ecological Informatics*, 61, 101248.
4. Hamzaoui, M.; Ould-Elhassen Aoueileyine, M.; Romdhani, L.; and Bouallegue, R. (2023). An improved deep learning model for underwater species recognition in aquaculture. *Fishes*, 8(10), 514.

5. Kuswantori, A.; Suesut, T.; Tangsrirat, W.; Schleining, G.; and Nunak, N. (2023). Fish detection and classification for automatic sorting system with an optimized YOLO algorithm. *Applied Sciences*, 13(6), 3812.
6. Liu, Y.; Zhang, Z.; Liu, X.; Wang, L.; and Xia, X. (2021). Ore image classification based on small deep learning model: Evaluation and optimization of model depth, structure and data size. *Minerals Engineering*, 172, 107020.
7. Liu, Y.; Wang, X.; Zhang, Z.; and Deng, F. (2023). Deep learning-based data augmentation for large-scale mineral image recognition and classification. *Minerals Engineering*, 204, 108411.
8. Mathur, M.; Vasudev, D.; Sahoo, S.; Jain, D.; and Goel, N. (2020). Crosspooled FishNet: Transfer learning-based fish species classification model. *Multimedia Tools and Applications*, 79(41-42), 31625-31643.
9. Borowiec, M.L.; Dikow, R.B.; Frandsen, P.B.; Mckeeken, A.; Valentini, G.; and White, A.E. (2022). Deep learning as a tool for ecology and evolution. *Methods in Ecology and Evolution*, 13(8), 1640-1660.
10. Mou, C.; Liang, A.; Hu, C.; Meng, F.; Han, B.; and Xu, J. (2023). Monitoring endangered and rare wildlife in the field: A foundation deep learning model integrating human knowledge for incremental recognition with few data and low cost. *Animals*, 13(20), 3168.
11. Yamamoto, S.; Masuda, R.; Sato, Y.; Sado, T.; Araki, H.; Kondoh, M.; Minamoto, T.; and Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, 7(1), 40368.
12. Liu, Y.; Yang, G.; Qiao, S.; Liu, M.; Qu, L.; Han, N.; Wu, T.; Yuan, G.; Wu, T.; and Peng, Y. (2023). Imbalanced data classification: Using transfer learning and active sampling. *Engineering Applications of Artificial Intelligence*, 117, 105621.
13. Aggarwal, U.; Popescu, A.; and Hudelot, C. (2020). Active learning for imbalanced datasets. *Proceedings of the* 2020 *IEEE/CVF Winter Conference on Applications of Computer Vision* (*WACV*), 1417-1426.
14. Xu, X.; Li, W.; and Duan, Q. (2020). Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. *Computers and Electronics in Agriculture*, 180, 105878.
15. Usmani, I.A.; Qadri, M.T.; Zia, R.; Alrayes, F.S.; Saidani, O.; and Dashtipour, K. (2023). Interactive effect of learning rate and batch size to implement transfer learning for brain tumor classification. *Electronics*, 12(4), 964.
16. Chavarro, A.F.; Renza, D.; and Ballesteros, D.M. (2023). Influence of hyperparameters in deep learning models for coffee rust detection. *Applied Sciences*, 13(7), 4565.
17. Kandel, I.; and Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 6(4), 312-315.
18. Gong, B.; Dai, K.; Shao, J.; Jing, L.; and Chen, Y. (2023). Fish-TViT: A novel fish species classification method in multi-water areas based on transfer learning and vision transformer. *Heliyon*, 9(6), e16761.