Fish Classification Using ResNet-50 Technical Report

Department of Mathematics, Universitas Indonesia
Komputasi Intelegensia
Fauzan Adzhima Alamsyah
1906376804

***Abstract*—This report presents the application of ResNet-50, a deep convolutional neural network, to classify images of fish species. The dataset consists of fish images from nine different classes. Techniques such as data augmentation and transfer learning were utilized to improve classification accuracy. The model achieved high accuracy during testing, validating the effectiveness of fine-tuning strategies.**

I. Introduction

Image classification is a critical task in computer vision with applications in numerous fields, including marine biology. This report focuses on classifying fish species using ResNet-50. The dataset consists of images from nine fish classes, and the problem involves accurately identifying the species from these images.

II. Problem Definition

The problem addressed in this report is the automatic classification of fish species from images. Manual identification is time-consuming and error-prone, and an automated solution is essential for scaling research efforts in marine biology.

III. Proposed Solution

The solution involves using ResNet-50, pretrained on ImageNet, as the base model. The fully connected layers were replaced with custom layers suitable for the nine-class classification task. The following steps were implemented:

- Data preprocessing and augmentation.
- Transfer learning by freezing initial layers of ResNet-50.
- Fine-tuning the entire model with a reduced learning rate.

IV. Research Methodology

A. Dataset Preparation

The dataset was preprocessed to remove extraneous entries and balance class distributions. It was split into training (80%), validation (10%), and test (10%) subsets. Data augmentation techniques included random rotations, horizontal flips, and resizing to 224x224 pixels.

B. Model Training

- Initial training was performed by freezing the pretrained layers and training only the classification head.
- Fine-tuning was conducted by unfreezing all layers and training with a reduced learning rate.
- Adam optimizer and cross-entropy loss were used for optimization.

C. Evaluation

The model was evaluated on the test set using metrics such as accuracy, precision, recall, and F1-score. A confusion matrix was generated to analyze classification errors.

V. Results and Analysis

A. Model Without Fine-Tuning Strategy

When the ResNet-50 model was trained without fine-tuning, only the fully connected layers were modified and trained. The pretrained layers remained frozen throughout the process. While the model achieved reasonable accuracy, its performance was slightly lower compared to the fine-tuned version due to the lack of adaptation to the specific fish dataset. Figures 1 shows the loss trends for this setup.

Fig. 1. Training and Validation Loss (No Fine-Tuning)

The training and validation loss for each epoch are summarized as follows:

- Epoch 1, Train Loss: 0.4568, Validation Loss: 1.1715
- Epoch 2, Train Loss: 0.1478, Validation Loss: 0.0915
- Epoch 3, Train Loss: 0.0816, Validation Loss: 0.4890
- Epoch 4, Train Loss: 0.1369, Validation Loss: 0.0849
- Epoch 5, Train Loss: 0.0620, Validation Loss: 0.2721
- Epoch 6, Train Loss: 0.0339, Validation Loss: 0.0014
- Epoch 7, Train Loss: 0.0513, Validation Loss: 0.2713
- Epoch 8, Train Loss: 0.0577, Validation Loss: 0.0356
- Epoch 9, Train Loss: 0.0311, Validation Loss: 0.5981
- Epoch 10, Train Loss: 0.0284, Validation Loss: 0.0349

The plot of training and validation loss for the model without fine-tuning illustrates several key points about the model's behavior and learning process:

- **Initial Drop in Loss:** Both the training and validation losses show a significant decrease during the first epoch. This indicates that the model is rapidly learning from the dataset, capturing some of the basic patterns present in the images.
- **Oscillation in Validation Loss:** The validation loss fluctuates significantly across epochs, showing spikes at certain intervals. This behavior suggests that the model is having difficulty generalizing to unseen data due to limited fine-tuning of the pretrained layers, which remain fixed and are not adapting to the specific fish dataset.
- **Smooth Decline in Training Loss:** The training loss decreases steadily over epochs, indicating that the model continues to fit the training data effectively. However, without fine-tuning, the lower layers of the network might not be fully optimized for the fish dataset's unique features.
- **Validation Loss Divergence:** The validation loss occasionally rises while the training loss continues to decrease. This divergence could point to overfitting, where the model becomes too tailored to the training data and struggles to perform as well on validation data.
- **Final Stability:** Toward the later epochs, the training loss stabilizes at a low value, and the validation loss also shows improvement in some epochs. However, the lack of fine-tuning limits the model's ability to fully minimize validation loss.

The fluctuations in validation loss and the steady decline in training loss indicate that while the model is learning patterns in the training data, it struggles to generalize well to validation data without fine-tuning. Fine-tuning the pretrained layers would likely improve the model's ability to adapt its features to the specific fish classification task, reducing validation loss and improving generalization.

The model achieved a nearly perfect accuracy of 99.56% on the test dataset, which showed the model performed well without having fine-tuned.

Evaluation Results:

- **Accuracy:** 1.00
- **Precision:** 1.00
- **Recall:** 1.00
- **F1-Score:** 1.00

Classification Report:

TABLE I
CLASSIFICATION REPORT (NO FINE-TUNING)

Class	Precision	Recall	F1-Score	Support
Sea Bass	0.97	1.00	0.98	59
Trout	1.00	1.00	1.00	54
Hourse Mackerel	1.00	1.00	1.00	47
Shrimp	1.00	1.00	1.00	40
Gilt-Head Bream	1.00	1.00	1.00	44
Red Sea Bream	1.00	1.00	1.00	52
Striped Red Mullet	1.00	1.00	1.00	43
Red Mullet	1.00	1.00	1.00	52
Black Sea Sprat	1.00	0.97	0.98	59
Accuracy	–	–	1.00	450
Macro Avg	1.00	1.00	1.00	450
Weighted Avg	1.00	1.00	1.00	450

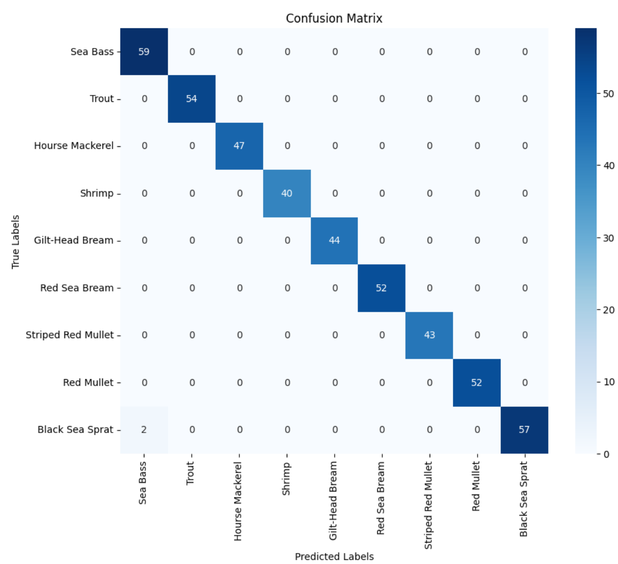

Fig. 2. Confusion Matrix Evaluation (No Fine-Tuning)

B. Model With Fine-Tuning Strategy

The fine-tuned ResNet-50 model achieved exceptional performance through a two-step process:

- **Initial Training:** The pretrained layers were frozen, and only the classification head was trained with a learning rate of 0.001. This step enabled the model to learn class-specific features effectively.
- **Fine-Tuning:** All layers were unfrozen, and the entire model was fine-tuned with a reduced learning rate of 0.00001 to adapt the pretrained weights to the fish dataset.

The results of the initial training phase are summarized as follows:

- Epoch 1, Train Loss: 0.8767, Train Acc: 0.8039, Val Loss: 0.3119, Val Acc: 0.9822
- Epoch 2, Train Loss: 0.2551, Train Acc: 0.9639, Val Loss: 0.1650, Val Acc: 0.9889
- Epoch 3, Train Loss: 0.1747, Train Acc: 0.9719, Val Loss: 0.1132, Val Acc: 0.9889
- Epoch 4, Train Loss: 0.1206, Train Acc: 0.9808, Val Loss: 0.0903, Val Acc: 0.9844

- Epoch 5, Train Loss: 0.1084, Train Acc: 0.9803, Val Loss: 0.0651, Val Acc: 0.9956

After fine-tuning all layers, the training results are as follows:

- Epoch 1, Train Loss: 0.0395, Train Acc: 0.9944, Val Loss: 0.0079, Val Acc: 1.0000
- Epoch 2, Train Loss: 0.0104, Train Acc: 0.9992, Val Loss: 0.0034, Val Acc: 1.0000
- Epoch 3, Train Loss: 0.0058, Train Acc: 0.9994, Val Loss: 0.0025, Val Acc: 1.0000
- Epoch 4, Train Loss: 0.0039, Train Acc: 1.0000, Val Loss: 0.0011, Val Acc: 1.0000
- Epoch 5, Train Loss: 0.0030, Train Acc: 0.9997, Val Loss: 0.0009, Val Acc: 1.0000
- Epoch 6, Train Loss: 0.0027, Train Acc: 0.9994, Val Loss: 0.0006, Val Acc: 1.0000
- Epoch 7, Train Loss: 0.0015, Train Acc: 0.9997, Val Loss: 0.0005, Val Acc: 1.0000
- Epoch 8, Train Loss: 0.0011, Train Acc: 0.9997, Val Loss: 0.0004, Val Acc: 1.0000
- Epoch 9, Train Loss: 0.0009, Train Acc: 1.0000, Val Loss: 0.0003, Val Acc: 1.0000
- Epoch 10, Train Loss: 0.0011, Train Acc: 1.0000, Val Loss: 0.0003, Val Acc: 1.0000

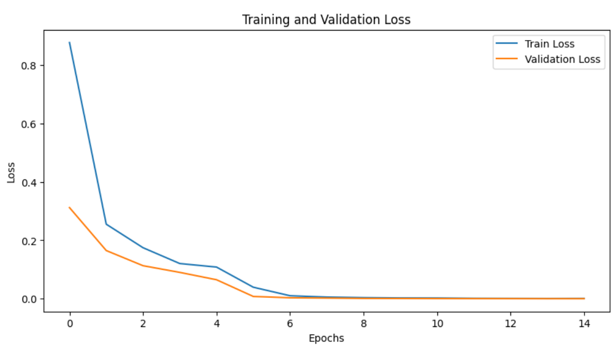

Fig. 3. Training and Validation Loss (Fine-Tuning)

Interpretation on Training and Validation Loss:

- **Rapid Initial Decline:** Both the training and validation losses show a sharp decrease in the first few epochs, indicating that the model quickly learns to represent the fish dataset.
- **Smooth Convergence:** The training and validation losses continue to decrease smoothly, and both converge near zero. This demonstrates that the model successfully minimizes both types of errors, suggesting excellent performance.
- **No Overfitting:** There is no significant divergence between the training and validation losses. This indicates that the fine-tuned model generalizes well to unseen data without overfitting.

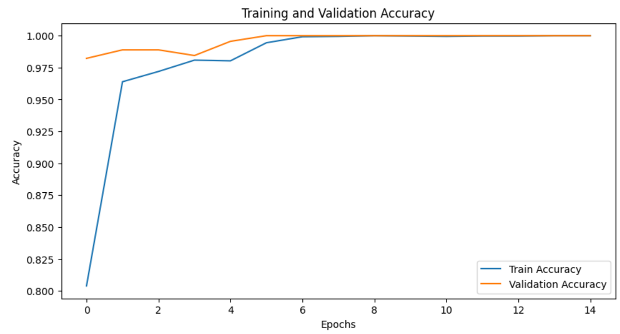

Fig. 4. Training and Validation Accuracy (Fine-Tuning)

Interpretation on Training and Validation Accuracy:

- **Rapid Accuracy Growth:** The training and validation accuracies increase significantly in the first few epochs, reflecting the model's ability to learn patterns from the fish dataset effectively.
- **Convergence to 100%:** Both the training and validation accuracies approach 100%, confirming that the model has achieved near-perfect performance.
- **Consistency:** The close alignment of training and validation accuracies further emphasizes the model's robust generalization capabilities.

The plots of training and validation loss and accuracy demonstrate the effectiveness of the fine-tuning strategy. The lack of overfitting, rapid convergence, and high performance on both training and validation sets validate the model's suitability for the fish classification task.

The model was further evaluated on the test dataset, achieving a perfect test accuracy of 100.00%. This result highlights the efficacy of fine-tuning in adapting the pretrained ResNet-50 to the fish classification task.

Evaluation Results:

- **Accuracy:** 1.00
- **Precision:** 1.00
- **Recall:** 1.00
- **F1-Score:** 1.00

Classification Report:

TABLE II
CLASSIFICATION REPORT (FINE-TUNING)

Class	Precision	Recall	F1-Score	Support
Sea Bass	1.00	1.00	1.00	59
Trout	1.00	1.00	1.00	54
Hourse Mackerel	1.00	1.00	1.00	47
Shrimp	1.00	1.00	1.00	40
Gilt-Head Bream	1.00	1.00	1.00	44
Red Sea Bream	1.00	1.00	1.00	52
Striped Red Mullet	1.00	1.00	1.00	43
Red Mullet	1.00	1.00	1.00	52
Black Sea Sprat	1.00	1.00	1.00	59
Accuracy	–	–	1.00	450
Macro Avg	1.00	1.00	1.00	450
Weighted Avg	1.00	1.00	1.00	450

Results showed consistent improvements in both training and validation metrics, as shown in Fig. 3 and Fig. 4, with minimal overfitting due to the use of regularization techniques like data augmentation.

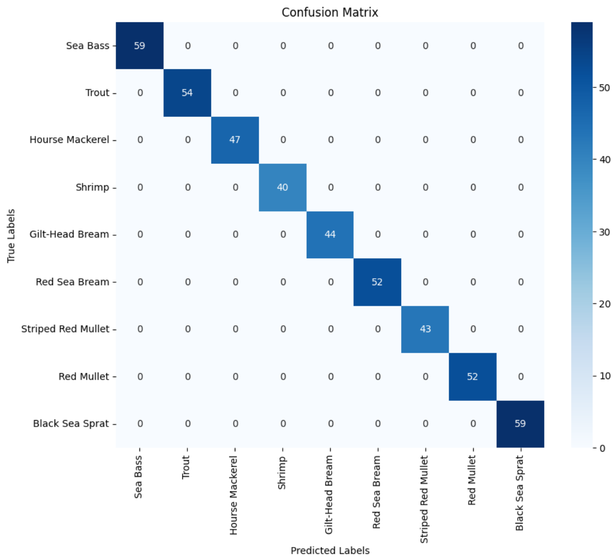

Fig. 5. Confusion Matrix Evaluation (Fine-Tuning)

VI. Potential Improvements to the Architecture

While the ResNet-50 architecture achieved excellent performance in the fish classification task, further improvements can be explored to enhance its robustness and adaptability:

- **Experimenting with Loss Functions:** Testing alternative loss functions like Focal Loss or Label Smoothing Loss could improve performance, especially for imbalanced datasets or when addressing harder-to-classify species.
- **Using Alternative Architectures:** Exploring newer architectures like EfficientNet, Vision Transformers (ViT), or Swin Transformers could yield better accuracy and computational efficiency.
- **Incorporating Attention Mechanisms:** Adding attention layers, such as the SE (Squeeze-and-Excitation) block or self-attention mechanisms, might help the model focus on the most critical regions of the images.
- **Hyperparameter Tuning:** Conducting extensive tuning of hyperparameters such as learning rate, batch size, and optimizer selection could optimize performance further.
- **Data Augmentation Strategies:** Leveraging advanced data augmentation techniques like CutMix, MixUp, or AutoAugment could increase model robustness to variations in the data.
- **Training on Larger Datasets:** Expanding the training dataset with additional fish species or using synthetic data generated through GANs (Generative Adversarial Networks) could improve generalization.
- **Ensemble Models:** Combining predictions from multiple models using ensemble techniques might enhance overall accuracy and reliability.

VII. Conclusions and Recommendations

The fine-tuned ResNet-50 model effectively classified fish images with high accuracy, demonstrating the potential of transfer learning for similar tasks. Future work could focus on:

- Expanding the dataset to include more fish species.
- Enhancing the model with attention mechanisms.
- Deploying the model in real-time applications.

References

[1] "DLVR Platform," Available: https://dlvr.rantai.dev/.

[2] "LMVR Documentation," Available: https://lmvr.rantai.dev/docs/lmvr/.

[3] A. Zhang, Z. Lipton, M. Li, and A. Smola, "Dive into Deep Learning," Available: https://d2l.ai/.

[4] "Source Code and Data Source," Available: https://github.com/NatureBase/EndTermTask.