Fine-Grained Bird Classification

1. Data Processing

The **CUB-200-2011** dataset was prepared by splitting it into **training, validation, and test sets**. The dataset was preprocessed to enhance generalization. The training set was augmented with transformations such as **resizing, normalization, horizontal flipping, rotation, and zooming** to improve robustness.

- **Total Training Images (Before Validation Split):** 5994
- **Total Test Images:** 5794
- **Training Samples After Split:** 5400
- **Validation Samples:** 594
- **Image Size:** 224×224 pixels

Data augmentation ensured diversity in the training set and helped the model learn better discriminative features.

2. Training Models and Methodology

The **EfficientNet-B3** model was chosen for transfer learning. Several models were considered, including **ResNet50, VGG16, and Xception**, but EfficientNet-B3 provided the best balance of performance and efficiency. The model was fine-tuned by modifying the final classification layer to match the **200 bird species**.

Approaches Considered

- **ResNet50** – Deep residual network, widely used for feature extraction.
- **VGG16** – A deep convolutional network with simple architecture.
- **EfficientNet-B3** *(Selected Model)* – Highly efficient and accurate for fine-grained classification.
- **Xception** – A deep learning model using depth-wise separable convolutions.

Hyperparameters Used

- **Optimizer:** Adam (Adaptive Moment Estimation)
- **Learning Rate:** 0.0005
- **Batch Size:** 32
- **Loss Function:** Cross-Entropy Loss
- **Epochs:** 50 *(early stopping applied at epoch 33)*

3. Training Performance

The model was trained for **50 epochs**, with the best model performance recorded at **epoch 33**.

- **Best Model Performance (Epoch 33):**
  - **Training Accuracy: 96.65%**
  - **Training Loss: 1.3205**
  - **Validation Accuracy: 77.78%**
  - **Validation Loss: 1.8289**

Fine-tuning was applied to deeper layers while freezing the initial layers to retain previously learned features from ImageNet.

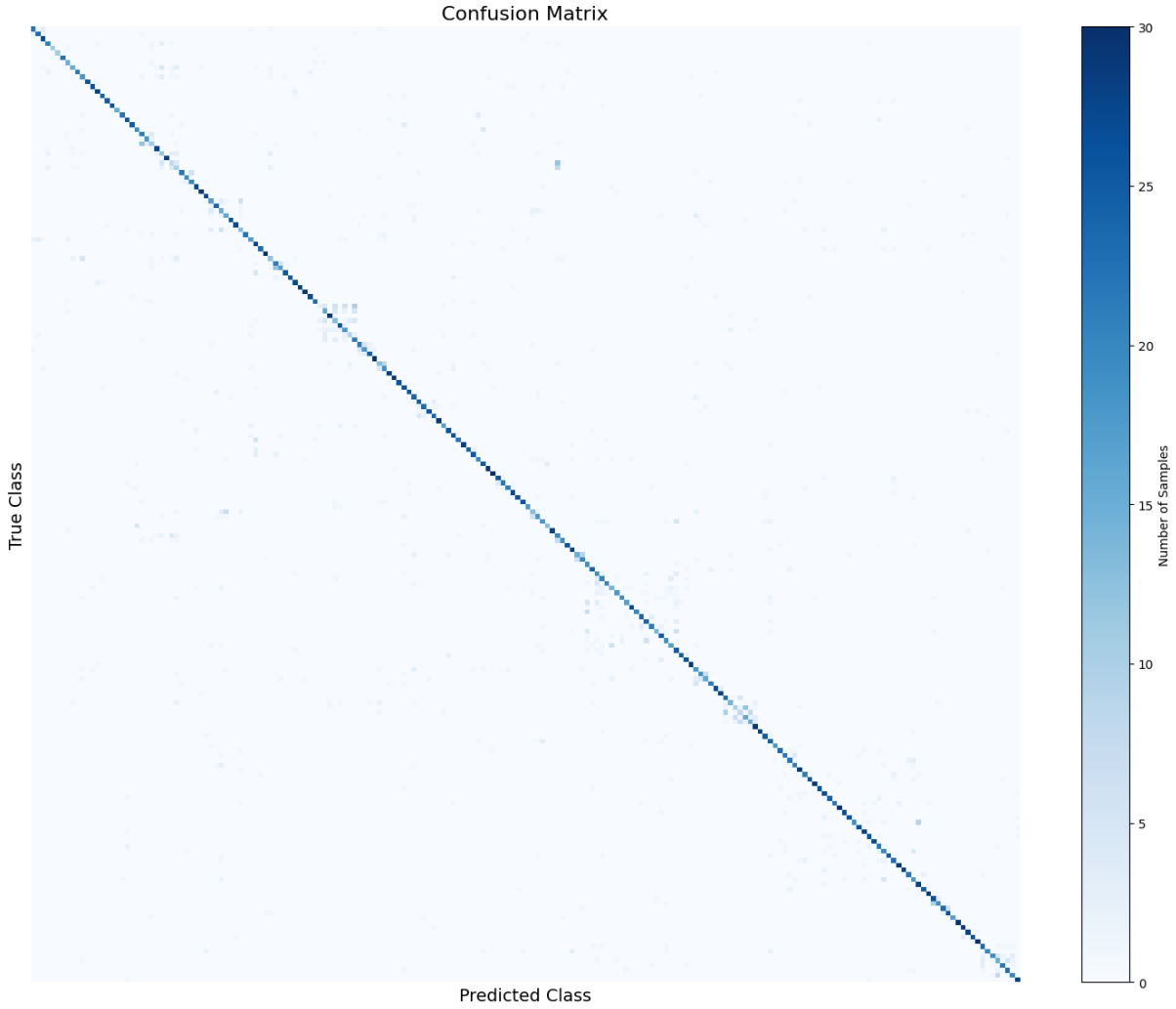
4. Test Performance & Metrics

The trained model was evaluated on the unseen test set, yielding the following results:

- **Test Accuracy: 76.63%**
- **Test Precision (Macro): 78.02%**
- **Test Recall (Macro): 76.88%**
- **Test F1-score (Macro): 76.85%**

Confusion Matrix

The confusion matrix below represents classification results across **200 bird species**:



This confusion matrix highlights areas of strong classification and misclassification patterns across different bird species.

5. Methods for Improving Performance

Several improvements can be applied to further optimize the model:

1. **Data Augmentation Enhancements:** Additional augmentations such as **brightness variation, contrast adjustment, and cutout augmentation** could improve generalization.
2. **Ensemble Learning:** Combining multiple models like **EfficientNet + ResNet50** could enhance classification robustness.
3. **Fine-Tuning More Layers:** Unfreezing additional layers for deeper feature extraction may boost accuracy.
4. **Higher Resolution Images:** Increasing resolution to **256×256** or **512×512** pixels could enhance feature detection.
5. **Hyperparameter Optimization:** Using **Bayesian Optimization or Grid Search** to fine-tune the learning rate, dropout, and weight decay.

6. Presentation of Results

The model's performance was assessed using key evaluation metrics:

- **Accuracy, Precision, Recall, and F1-score**
- **Confusion Matrix for error analysis**
- **Comparison of Training vs. Validation Performance**

The final trained model achieved **76.63% accuracy** on the test set, demonstrating strong performance in fine-grained bird classification. However, further optimization is required to exceed **80% accuracy**.

Conclusion

The **EfficientNet-B3** model successfully classified **200 bird species** from the **CUB-200-2011 dataset**. The final **test accuracy of 76.63%** and **macro F1-score of 76.85%** indicate strong classification performance. Future work will focus on **improving data augmentation, fine-tuning more layers, and optimizing hyperparameters** to further enhance the model's accuracy.