Landscape Classification

DATS 6103 Data Mining

Jeongwon Yoo

1. Introduction

1.1. Overview

This project investigates automated classification of natural landscape images using deep learning. Landscape images often contain visually similar features, and scene categorization can be challenging due to overlapping patterns such as snow, water, vegetation, and mountains. The goal is to build a model that can classify five landscape categories: Coast, Desert, Forest, Glacier, and Mountain.

1.2. Objectives

The main objective is to develop a classification model that can accurately distinguish visually similar landscape types. Additional objectives include identifying common misclassification patterns and evaluating how transfer learning affects performance.

1.3. Research Questions

1. Can deep learning models accurately classify landscape images into five fine-grained categories?
2. Does transfer learning significantly improve performance compared to a traditional convolutional baseline?
3. What visual characteristics cause recurring misclassification errors?

The questions arose from the practical challenges encountered during the project. Initial experiments showed difficulty in separating visually similar landscapes, leading to questions about model accuracy and the benefits of transfer learning. Misclassification patterns also appeared consistently, motivating a question focused on understanding the visual causes behind these recurring errors.

1.4. Significance of the Study

Automated landscape classification supports large-scale environmental monitoring, tourism image tagging, and geospatial analysis. Accurate identification is useful for ecological studies and remote sensing tasks. Understanding misclassification patterns also provides insight into model limitations when differentiating visually similar natural scenes.

2. Dataset and Exploratory Analysis

2.1. Dataset Description

The dataset is the Kaggle Landscape Recognition dataset, consisting of approximately 12,000 images across five evenly balanced categories. The data is split into:

- 10,000 training images
- 1,500 validation images
- 500 test images

Images vary widely in size, ranging from roughly 150 px to 500 px on each dimension. Resizing to a uniform resolution was necessary for model training.

2.2. Exploratory Data Analysis

Several observations emerged during EDA:

- Image dimensions cluster around 250 to 300 px in width and height.
- The five classes exhibit meaningful visual diversity, but Glacier and Mountain show substantial overlap in color tone and structure.
- Many images contain mixed features, such as water within mountain regions or rocky beaches that resemble deserts.

- These overlaps suggest that fine-grained classification would be challenging for simpler models.
- Random samples from each class confirm substantial within-class variability and between-class similarity.

3. Methodology

3.1. Preprocessing Pipeline

All images were resized to 224 × 224, normalized, and augmented. The augmentation included random rotation, flip, zoom, and contrast adjustments to improve model robustness. EfficientNet built-in preprocessing was applied to ensure compatibility with the pre-trained backbone.

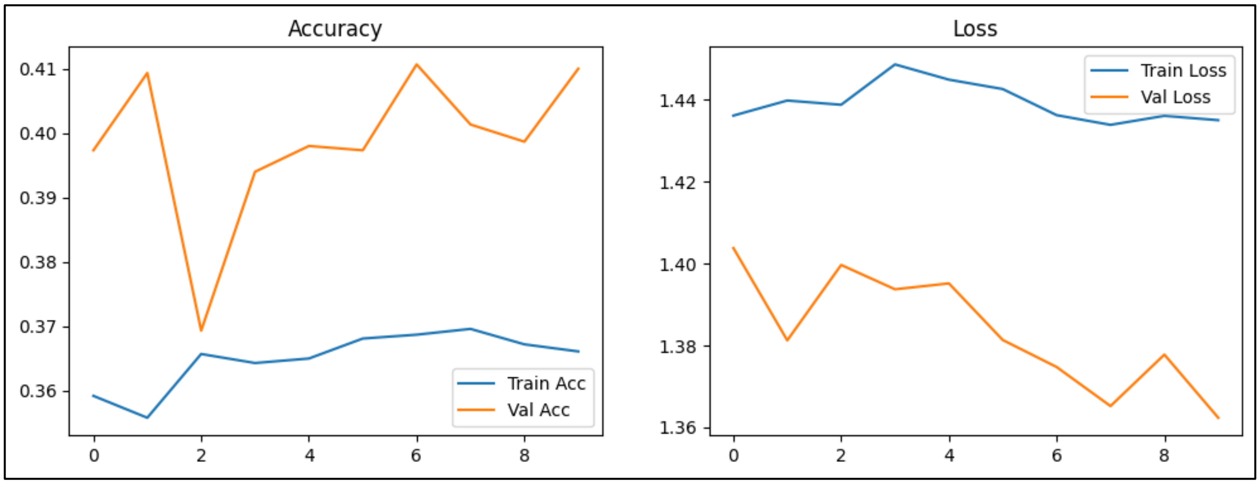
3.2. Data Pipeline

The TensorFlow pipeline used:

- cache() to keep batches in memory
- prefetch() for efficient GPU streaming
- Batch size set to 32

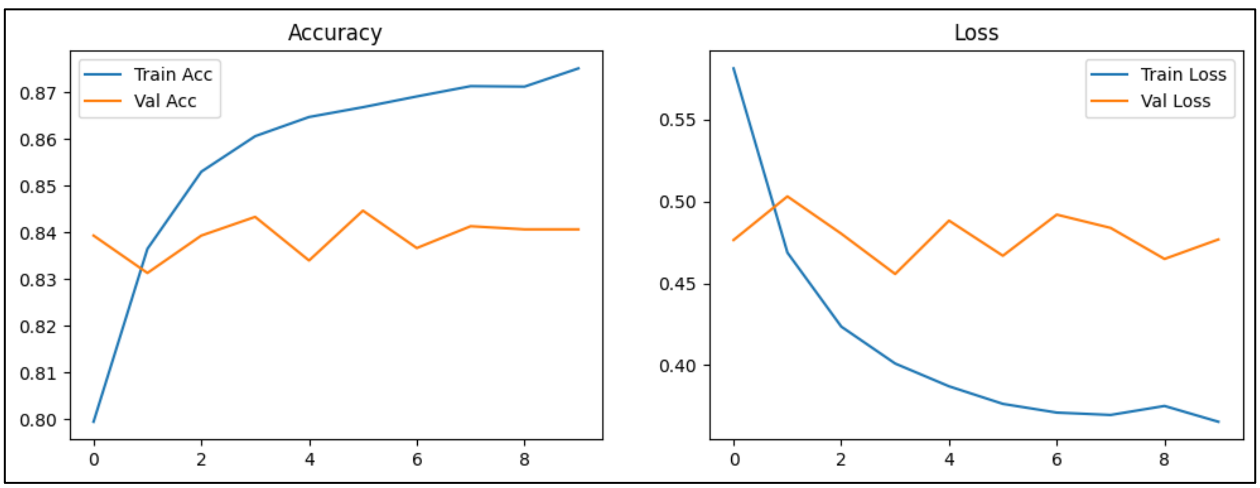
These steps reduced training bottlenecks and ensured smooth data flow.

3.3. Model Architecture and Training Set



Baseline Model: ResNet50

The project began with a pre-trained ResNet50 model. Despite its strong reputation, the model showed underfitting, achieving only about 0.40 validation accuracy. Loss curves indicated that the model struggled to capture fine-grained visual differences between classes.



Improved Model: EfficientNetB0

EfficientNetB0 was selected for its lighter architecture and efficient scaling properties. It uses compound scaling for width, depth, and resolution and performs well on natural image features.

The model demonstrated:

Validation accuracy = ≈ 0.83

Test accuracy = 0.884

Stable training and improvement across epochs

4. Models and Experiments

4.1. Experiment Setup

- **Epochs**: 10
- **Optimizer**: Adam (learning rate = 1e-3)
- Loss: Sparse categorical crossentropy
- **Batch size**: 32

Data augmentation applied to all models

4.2. Baseline Performance: ResNet50

Validation accuracy plateaued near 0.40. Training and validation curves diverged, indicating the model was not expressive enough for the dataset's complexity.

4.3. Improved Performance: EfficientNetB0

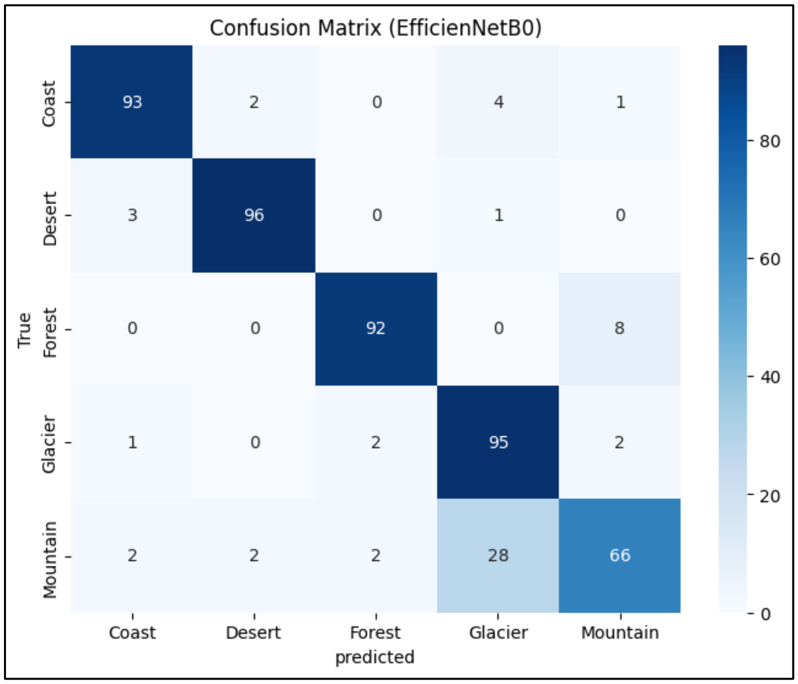
EfficientNetB0 improved performance across all metrics. The model learned stable representations, and both accuracy and loss curves improved consistently through training.

5. Results and Evaluation

5.1. Model Comparison

Model	Test Accuracy	Macro F1	Notes
ResNet50	~0.40	low	Underfitting
EfficientNetB0	0.884	0.88	Strong generalization

5.2. Confusion Matrix Analysis

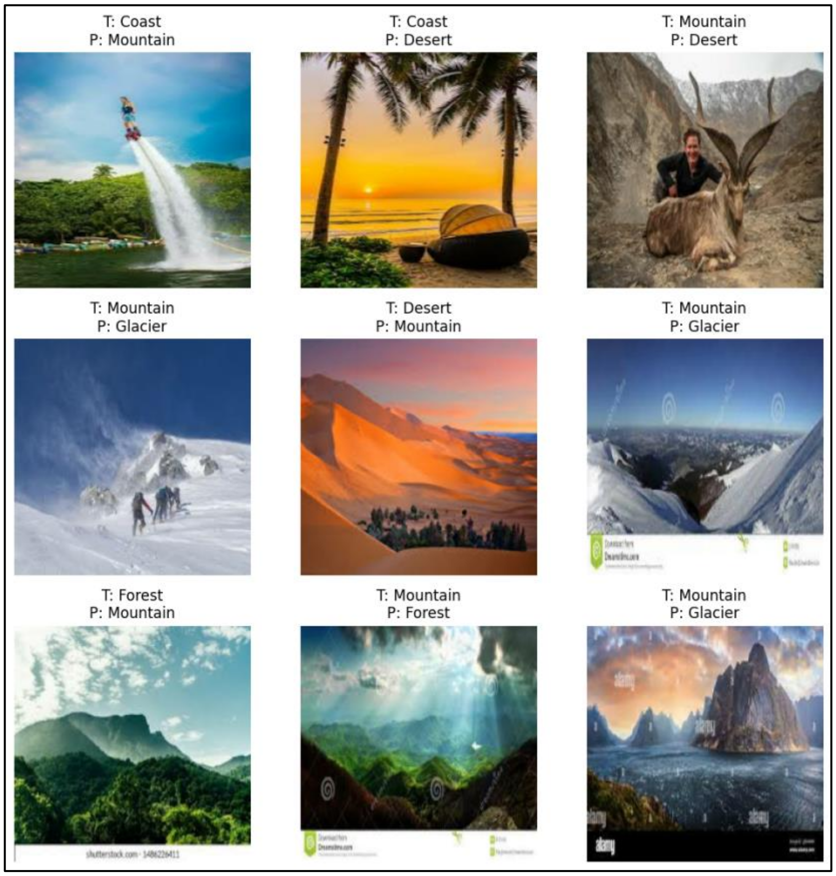


The confusion matrix reveals strong performance for Coast, Desert, Forest, and Glacier. The greatest difficulty appears between Mountain and Glacier, where structural and color similarities confuse the model.

5.3. Classification Report

Precision and recall are above 0.90 for most classes. The Mountain class has lower recall due to overlap with Glacier-like features such as snowy ridges, blue-toned peaks, and cloud patterns.

5.4. Key Error Patterns



Misclassifications typically occurred in images that contain:

- Snowy terrain that resembles both high mountains and glaciers
- Mixed water and mountain scenery

These errors are expected in fine-grained natural image classification, where class boundaries are not always visually distinct.

6. Conclusion and Future Work

6.1. Practical Use

The model can predict the landscape category of new, previously unseen images, which enables practical uses such as automatically tagging large environmental image collections and supporting ecological monitoring. For example, applying the model to periodic satellite images could help track changes in forests, glaciers, or coastlines over time. These predictions make it possible to sort, filter, and analyze natural scenes more efficiently without manual labeling.

6.2. Conclusion

EfficientNetB0 achieved a test accuracy of 0.884, clearly outperforming the ResNet50 baseline, which confirms that deep learning models can effectively classify the five landscape categories and that transfer learning provides a significant performance advantage. The model handled most categories reliably, although Glacier and Mountain scenes remained challenging due to their visual similarity, which shows that the third research question is only partially resolved. Overall, the results demonstrate solid progress on fine-grained landscape classification while also indicating where further improvement is needed.

6.2 Future Work

Several directions can further enhance performance:

1. Fine-tuning deeper EfficientNet variants such as B1 to B3.
2. Applying targeted augmentations focused on snowy or icy landscapes.
3. Incorporating metadata such as elevation or GPS information to improve class separation.
4. Experimenting with attention-based networks that learn localized features.

References

1. DEEPNETS. (n.d.). *Landscape recognition image dataset: 12k images*[Data set]. Kaggle. https://www.kaggle.com/datasets/utkarshsaxenadn/landscape-recognition-image-dataset-12k-images
2. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. arXiv : 1512.03385
3. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. arXiv:1905.11946