Intel Image Classification using MobileNetV2

1. Introduction

This project focuses on implementing an image classification system using Transfer Learning with MobileNetV2, a pre-trained Convolutional Neural Network (CNN). The objective is to accurately classify landscape images from the Intel Image Classification dataset into one of six categories: buildings, forest, glacier, mountain, sea, and street. The project involves training, evaluating, and deploying the model using TensorFlow and Streamlit.

2. Dataset Description

The Intel Image Classification dataset is a widely used benchmark dataset for multi-class image classification tasks. It contains labeled RGB images of six natural scene classes:

- Buildings
- Forest
- Glacier
- Mountain
- Sea
- Street

Dataset statistics:

- Training Images: 11,230
- Validation Images: 2,804 (20% split from training set)
- Testing Images: 3,000

3. Data Preparation

To ensure the model receives consistent input, all images were resized to 224x224 pixels. Normalization was applied by scaling pixel values to the [0, 1] range. Data augmentation techniques such as rotation (20 degrees), zoom (20%), and horizontal flipping were used to improve generalization. The Keras ImageDataGenerator was used for loading and preparing the data with real-time augmentation.

4. Model Architecture

The model is based on the MobileNetV2 architecture, which is optimized for efficiency and performance. The base model was loaded without the top classification layers, and its weights were frozen during the initial training phase.

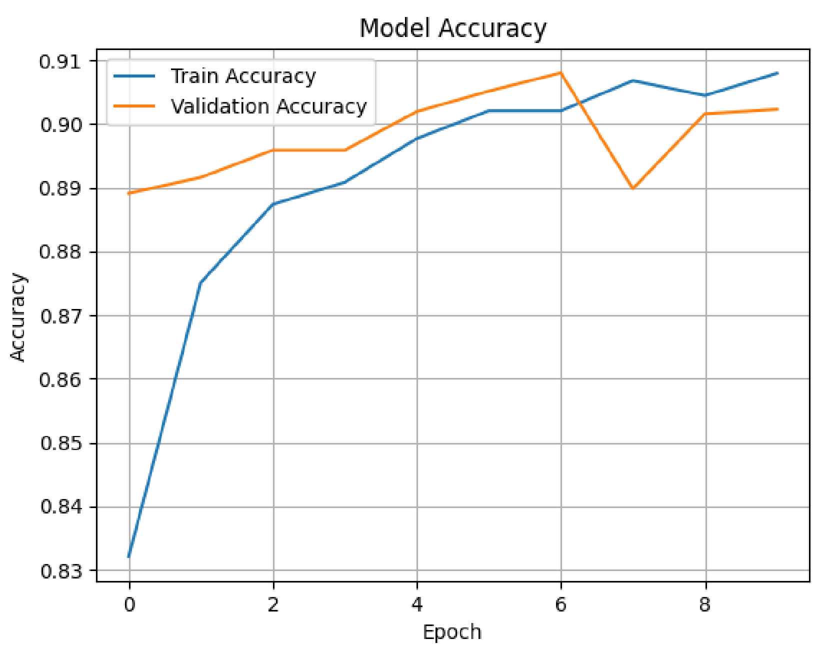
A custom classification head was added on top of the base model:
- GlobalAveragePooling2D layer
- Dense layer with 128 units and ReLU activation
- Dropout layer with 30% rate
- Final Dense layer with softmax activation (6 output units for 6 classes)

Model Summary:
- Total parameters: 2,422,726 (9.24 MB)
- Trainable parameters: 164,742 (643.52 KB)
- Non-trainable parameters: 2,257,984 (8.61 MB)

5. Training

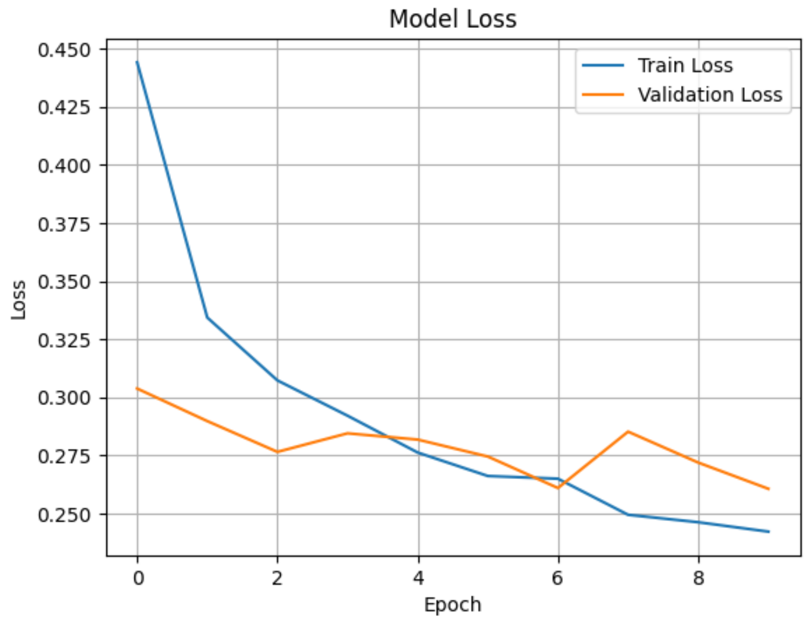
The model was compiled using the Adam optimizer and categorical crossentropy loss. Training was conducted over 10 epochs with a batch size of 32. Both training and validation accuracy steadily increased while loss decreased, indicating successful learning.

Final Results:
- Training Accuracy: 91.22%
- Validation Accuracy: ~90.8%
- Training Loss: 0.2349
- Validation Loss: ~0.27



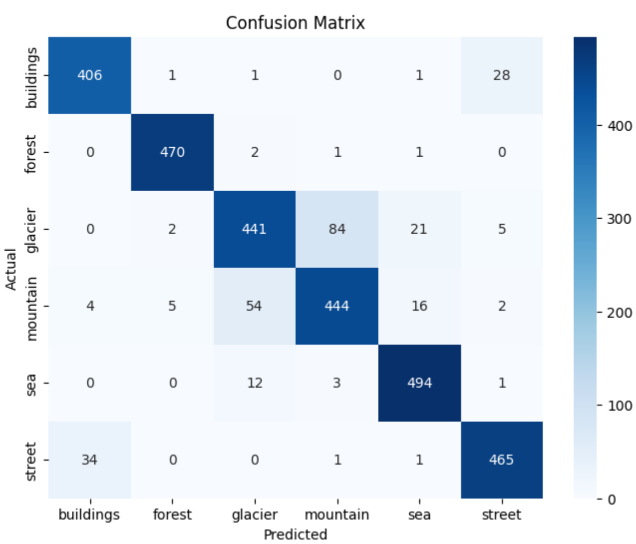
6. Evaluation

Evaluation was performed on 3,000 unseen test images. Metrics included confusion matrix and accuracy/loss visualizations. The model demonstrated strong performance across all six categories, maintaining consistent predictions with high confidence.



7. Deployment

The trained model was saved in the HDF5 format (`model.h5`). Deployment is facilitated using a Streamlit web app, where users can upload an image file and receive real-time predictions. The app displays the predicted class with confidence and optionally the top 3 predictions. It provides an intuitive interface for showcasing the model's capabilities.



8. Conclusion

This project demonstrates the effectiveness of transfer learning with MobileNetV2 for natural scene classification. Using a relatively small training set and efficient pre-trained model, we achieved high classification accuracy. The deployment step makes this solution accessible to non-technical users through an interactive web interface built with Streamlit.