Vinay Verma

CS 4391.002

Ranran Feng

11 May 2025

**Term Project Report: Scene Recognition**

**1. Project Overview**

The goal of this project was to implement basic scene understanding using image processing, feature extraction, clustering, and classification methods. I was provided two sets of images (Train and Test) across four categories: bedroom, desert, landscape, and rainforest.

I experimented with multiple classification methods, including:

* Raw pixel representation
* SIFT feature extraction
* Histogram of grayscale intensities
* Convolutional Neural Networks (CNN), including a Transfer Learning approach using MobileNetV2.

**2. Implementation Details**

**2.1 Preprocessing**

* All images were converted to grayscale.
* Brightness adjustment was performed:
  + If average brightness < 0.4, brightness was increased.
  + If average brightness > 0.6, brightness was reduced.
* Each image was resized to:
  + 50×50 for Pixel, SIFT, and Histogram methods
  + 200×200 for CNN models

**2.2 Feature Extraction**

* SIFT features were extracted using OpenCV, and feature vectors were averaged for each image.
* Histograms (32-bin grayscale) were computed and normalized.

**2.3 Classification Methods**

* Nearest Neighbor Classifier (k=1) was used for:
  + Pixel values (50×50)
  + SIFT features
  + Histogram features
* CNN (Transfer Learning) was implemented:
  + Base model: MobileNetV2 pretrained on ImageNet.
  + Custom classification head added.
  + Grayscale images were adapted by replicating the single channel to 3 channels.

**3. Results**

| **Classifier** | **Accuracy** | **False Positive Rate** | **False Negative Rate** |
| --- | --- | --- | --- |
| Pixel (50x50) + NN | 38.00% | 62.00% | 62.00% |
| SIFT Features + NN | 48.00% | 52.00% | 52.00% |
| Histogram Features + NN | 51.50% | 48.50% | 48.50% |
| Transfer Learning CNN (MobileNetV2) | 77.00% | 23.00% | 23.00% |

**4. Discussion and Analysis**

**4.1 Performance Comparison**

* Histogram + Nearest Neighbor performed best among traditional methods, achieving 51.5% accuracy.
* SIFT features improved upon direct pixel comparison but did not outperform histograms.
* Transfer Learning with MobileNetV2 achieved 77.0% accuracy, a substantial improvement over all previous methods.

**4.2 Why Transfer Learning Worked Better**

* MobileNetV2 was pretrained on millions of images and learned general-purpose features.
* Even though our dataset was small, MobileNetV2 was able to transfer its feature extraction ability effectively.
* This allowed the CNN model to avoid overfitting and generalize well, unlike our earlier simple CNN which overfit rapidly.

**4.3 Factors Affecting Accuracy**

* Small dataset size made it difficult for CNNs trained from scratch to generalize.
* Grayscale images limited the amount of information compared to full RGB images.
* Scene similarities between categories (e.g., landscape and rainforest) increased confusion rates in simpler classifiers.

**5. Conclusion**

In this project, we successfully implemented multiple scene classification approaches.  
All models achieved more than the required 25% classification accuracy, with the best performance achieved using Transfer Learning via MobileNetV2, reaching 77% test accuracy.

The project highlights the importance of feature selection and the power of transfer learning when working with small datasets.