Technical report

Title:

Deep Learning Approach for detecting nature images

Name: **Ashraf Alham**

Date: **15/5/2025**

1. Introduction:

1.1 Problem Statement:

my problem is to classify the inlet image to 6 categories:

(Sea – Mountain – Glacier – Forest – Street – Building)

1.2 Objective:

the goal of my project is to classify any image to these categories this will help us to good recognition to the nature and use these models for any future tasks and you can use it on mutable classes like detection of place for someone and you can detect how he can live from the nature of his life, also I have another objective is to compare the results of multiple models like (VGG16 – VGG19 – My customized model – Inception – ResNet) and get the best accuracy for each one and also use ensemble technique to get the best model prediction for each image to reduce the loss and error can happens.

1.3 Motivation:

this task is very important as I can use it on application like:
Urban Planning –Tourism and Travel Recommendations-Educational Tools-Social Media Filtering and Tagging-Climate Change Research-Environment Monitoring …. etc.

1.4 Dataset:

I get this dataset from Kaggle which has about 14,000 training images and about 3000 test image and about 1000 for prediction, this dataset almost enough for models.

2. Methodology

2.1 Dataset

The dataset consists of [18000] images divided into six classes: mountain, sea, forest, glacier, building, and street. Each class contains [Y] images, and the dataset was split into training (80%), validation (10%), and test (10%) sets. Data augmentation techniques such as rotation, scaling, shear, brightness, shifting, flipping, and zooming were applied to increase dataset diversity.

2.2 Model Architectures

- **VGG16**: A 20-layer CNN with 13 convolutional layers and 7 fully connected layers.
- **Customized**: A 11-layer CNN with 9 convolutional layers and 2 fully connected layers.
- **VGG19**: A 27-layer CNN with 16 convolutional layers and 11 fully connected layers.
- **InceptionV3**: A 98-layer CNN with 94 convolutional layers and 4 fully connected layers.
- **ResNet**: A 156-layer CNN with 152 convolutional layers and 4 fully connected layers.

2.3 Training Process

- All models were fine-tuned using transfer learning.

- The base layers were frozen, and only the fully connected layers were trained.

- Training parameters:

  - Optimizer: Adam
  - Learning Rate: 0.001 and reduces if needed
  - Batch Size: 64
  - Epochs: 30

2.4 **Evaluation Metrics**

Accuracy: Percentage of correctly classified images.

Precision: Proportion of true positives among predicted positives.

Recall: Proportion of true positives among actual positives.

F1-Score: Harmonic mean of precision and recall.

3. Results

3.1 Performance Comparison

Network	Accuracy	Recall	Precision	F1-Score
Your CNN model	0.82	0.8153	0.8548	0.8164
VGG16	0.90	0.9020	0.9034	0.9010
VGG19	0.90	0.90	0.90	0.90
Inception_V3	0.90	0.90	0.90	0.90
RasNet152V2	0.89	0.8943	0.8944	0.8938

3.2 **Confusion Matrix**

- **Inception**:
  (array ([[409, 1, 0, 0, 3, 24],
  [0, 461, 1, 6, 3, 3],
  [2, 2, 448, 70, 29, 2],
  [2, 0, 66, 446, 10, 1],
  [2, 1, 4, 13, 487, 3],
  [44, 0, 0, 3, 2, 452]])

- **VGG16**:
  (array ([[377, 2, 0, 1, 12, 45],
  [0, 470, 0, 0, 1, 3],
  [0, 2, 486, 40, 20, 5],
  [0, 8, 86, 406, 24, 1],
  [2, 1, 14, 4, 489, 0],
  [17, 2, 0, 0, 4, 478]])

- **VGG19**:
  (array ([[403, 0, 1, 0, 0, 33],
  [1, 472, 0, 1, 0, 0],
  [4, 5, 441, 89, 12, 2],
  [6, 8, 32, 467, 10, 2],
  [12, 8, 15, 17, 453, 5],
  [32, 2, 1, 1, 0, 465]])

- **ResNet**:
  (array ([[370, 4, 0, 4, 5, 54],
  [1, 469, 0, 2, 1, 1],
  [0, 5, 465, 59, 22, 2],
  [0, 2, 50, 443, 30, 0],
  [4, 6, 7, 10, 482, 1],
  [27, 4, 1, 10, 5, 454]])

- **Customized**:
  (array ([[433, 0, 1, 0, 0, 3],
  [17, 446, 2, 3, 0, 6],
  [18, 1, 447, 78, 7, 2],
  [15, 3, 46, 444, 16, 1],
  [44, 3, 27, 27, 406, 3],
  [224, 1, 0, 4, 2, 270]])

Conclusion

This study compared the performance of customized model, VGG16, VGG19, ResNet152, and InceptionV3 for nature image classification. Results show that **customized** shows lower than other models in terms of accuracy, precision, recall, and F1-score. However, **InceptionV3** offers a good balance between performance and computational efficiency. ensemble methods used to further improve performance, it really approved best accuracy.
Thank you.