

Tahmina Rahman Jidny

ID: 2021-2-60-167

Section: 2

LAB 7: Final Lab

Lab Report on: Image Classification using ResNet50 and Grad-CAM

1. Introduction

The purpose of this lab is to train and evaluate a deep learning model for image classification using the Caltech-101 dataset. A pre-trained ResNet50 model is fine-tuned to classify images into 101 categories. Additionally, Grad-CAM is used to provide visual explanations for the model's predictions.

2. Methods

2.1 Environment Setup

The required Python packages are installed, including PyTorch, Torchvision, and Grad-CAM. The code is executed in Google Colab with GPU acceleration enabled.

2.2 Dataset Preparation

The Caltech-101 dataset is loaded and preprocessed using the following transformations:

- **Resize:** Images are resized to 224×224 pixels.
- **Grayscale Conversion:** Converted to 3-channel RGB.
- **Normalization:** Pixel values are normalized to mean = [0.5, 0.5, 0.5] and std = [0.5, 0.5, 0.5].
- **Splitting:** The dataset is divided into training (80%), validation (10%), and test (10%) sets.

2.3 Model Selection and Modification

A **ResNet50** model, pre-trained on ImageNet, is used. The final fully connected layer is replaced to match the 101 output classes.

2.4 Training

The model is trained for **10 epochs** using:

- **Loss Function:** CrossEntropyLoss

- **Optimizer:** Adam with learning rate 0.001
- **Batch Size:** 32
- **Device:** GPU (if available)

During training, loss values are computed and printed for each epoch.

2.5 Validation and Testing

- The trained model is evaluated on the validation dataset, and accuracy is calculated.
- The model is then tested on unseen data, and predictions are compared against ground truth labels.
- **Evaluation Metrics:**
 - Confusion Matrix
 - Classification Report (Precision, Recall, F1-score)

2.6 Grad-CAM Visualization

Grad-CAM is used to generate heatmaps highlighting the most important regions in images for model predictions. The visualization is applied to **5 test images**.

2.7 Model Saving

The trained model is saved to Google Drive for future use.

3. Results

- **Training Loss:** The loss gradually decreases over 10 epochs.
- **Validation Accuracy:** The model achieves a specific percentage accuracy on the validation dataset.
- **Test Performance:**
 - Confusion matrix and classification report provide insights into the model's performance.
- **Grad-CAM Outputs:** Heatmaps visually highlight important features used by the model.

4. Conclusion

- A ResNet50 model was successfully fine-tuned for **Caltech-101 classification**.
- The trained model achieved **reasonable accuracy**.

- Grad-CAM provided **explainability** by highlighting image regions influencing predictions.
- Further improvements could be made by **hyperparameter tuning, data augmentation, or using a more complex model.**