Antika Dhar; ID: 2104010202256
CSE 452: Neural Network & Fuzzy Logic Laboratory; Section: C

Self-Contribution Report: My Role in the Visual Question Answering Project

Our goal was to build a Visual Question Answering (VQA) system that could look at an image, understand a question about it, and provide an appropriate answer. We used deep learning models—ResNet-50, EfficientNet-B3, ViT-B/16 for images, and BERT for text. In this report, I'll explain my specific contributions.

My work focused on six key areas: model evaluation**,** model comparison**,** model selection and saving**,** inference on new data**,** error analysis**, and** submission preparation.

Model Evaluation and Comparison

I evaluated ResNet-50, EfficientNet-B3, and ViT-B/16 models using classification reports and confusion matrices, analyzing precision, recall, and f1-scores. I compared models based on accuracy and loss curves, highlighting the ViT-B/16 model's consistent performance across training and validation.

Model Selection and Saving

Based on performance, I selected the ViT-B/16 model as the final model. I saved the trained model using PyTorch's `torch.save()` function for future use.

Inference on New Data

I developed scripts for loading saved models and making predictions on unseen image-question pairs. Sample outputs were prepared to demonstrate the system's inference capabilities.

Error Analysis and Visualization

I performed error analysis by reviewing misclassified examples and created visualizations to better understand model limitations. Insights from this helped explain prediction inconsistencies.

Submission Preparation

I organized project files and conducted flexible sample tests to ensure a smooth, complete, and functional final submission.

My contributions helped ensure a thorough evaluation, reliable deployment, and a strong final submission for our VQA project.

IEE Format Report

Also, in the IEEE format report I authored the Hyper Parameter Tuning, Error Analysis, Discussion, Conclusion and References.

My contribution ensured that the project was thoroughly evaluated, visually explained, and well-prepared for demonstration and future extensions.