Lab Week 14 – CNN Image Classification

Name: Pinisetti Sudhiksha

SRN: PES2UG23CS916

Course: Machine Learning

Submission Date: 19/11/2025

Section: F

Introduction

The objective of this lab was to design, implement, and train a Convolutional Neural Network (CNN) capable of classifying images of hand gestures into three categories: rock, paper, and scissors. Using PyTorch and a labeled dataset, the task involved preparing the data, defining an appropriate CNN architecture, training the model, and evaluating its performance on unseen images.

Model Architecture

The CNN architecture consisted of three convolutional blocks, each containing a 3×3 convolutional layer followed by a ReLU activation and a MaxPool2d layer with a pooling size of 2. The network increased its channel depth from 3 to 16, then 32, and finally 64 across the convolutional stages while progressively reducing the spatial dimensions from 128×128 down to 16×16. After the convolutional feature extractor, the model used a fully connected classifier that first flattened the 64×16×16 tensor into a vector of 16,384 elements. This was passed through a linear layer of size 256, followed by ReLU activation and a dropout layer with p=0.3, before reaching the final linear layer that outputs class logits for the three gesture categories.

Training and Performance

Key hyperparameters:

- Optimizer: Adam
- Learning rate: 0.001
- Loss function: CrossEntropyLoss
- Epochs: 10

Performance results:

- Training loss reduced steadily from 0.7475 → 0.0029 across 10 epochs
- Final test accuracy: 97.95%

Conclusion and Analysis

The model demonstrated strong performance, achieving nearly 98% accuracy on unseen data.

Loss decreased smoothly during training, indicating stable learning.

Individual image tests and example predictions confirmed that the model made correct classifications.

A few potential improvements include:

- Adding data augmentation to improve robustness
- Increasing model depth or experimenting with different architectures like BatchNorm to enhance generalization