

Week 14 Lab Report- CNN Image Classification

Name: Megha Dhanya

SRN: PES2UG23CS336

Section: F

1. Introduction

The objective of this lab was to design, build, and train a Convolutional Neural Network (CNN) using PyTorch to classify hand gesture images into three categories: rock, paper, and scissors. The dataset consisted of 2188 labeled images organized into respective folders. The task required completing a boilerplate notebook, training the model, evaluating accuracy, and preparing a final report.

2. Model Architecture

The constructed CNN includes three convolutional blocks. Each block contains a Conv2D layer with kernel size 3×3 and padding 1, followed by ReLU activation and MaxPooling with a stride of 2. The channel progression is as follows: 3 to 16, 16 to 32, and 32 to 64. After three pooling layers, the spatial dimension reduces from 128×128 to 16×16 . The fully connected classifier includes a Flatten operation, a Linear layer projecting from $64 \times 16 \times 16$ to 256 units, followed by ReLU activation, Dropout ($p=0.3$), and a final Linear layer producing three output class logits.

3. Training and Performance

The model was trained for 10 epochs using the Adam optimizer with a learning rate of 0.001 and CrossEntropyLoss as the loss function. The dataset was split into 80% training data and 20% testing data. After training, the model achieved a final test accuracy of 99.32%, demonstrating excellent classification performance.

4. Conclusion and Analysis

The CNN performed extremely well, achieving over 99% accuracy. The model successfully learned distinguishing features of hand gestures and generalized effectively to unseen data. Challenges were minimal; however, small batch size adjustments or additional data augmentation could further improve robustness. Future improvements could include experimenting with deeper architectures or applying regularization strategies to prevent overfitting.