

UE23CS352A: MACHINE LEARNING

Week 14: CNN Image Classification

Name : CHARAN M REDDY

SRN : PES2UG23CS146

Section : C

Date : 20/11/2025

1. INTRODUCTION

In this lab we build and train a Convolutional Neural Network using PyTorch , to classify hand gestures in images into three categories: 'rock', 'paper', and 'scissors'. We used the Rock-Paper-Scissors dataset, which contains over 2,000 images, and tried to achieve high accuracy in classifying these gestures.

2. MODEL ARCHITECTURE

The CNN architecture has three convolutional blocks. Each block has a Conv2d layer, a ReLU activation and a MaxPool2d layer. The first block use 16 channels, the second use 32, and the third use 64. The convolutional layers use a 3x3 kernel with padding of 1, and MaxPooling layer uses 2x2 kernel.

After the convolutional layers, the features are flattened and passed through two fully connected layers. The first layer reduces the features to 256 dimensions, and the second layer will output 3 values . The 3 outputs will be one each for three classes : rock, paper, and scissors. A Dropout layer with a 0.3 probability is also added .

3. TRAINING AND PERFORMANCE

Optimizer: Adam optimizer , learning rate = 0.001.

Loss Function: Cross-Entropy Loss function.

Epochs: 10 epochs .

Test Accuracy: 98.63%

4. CONCLUSION AND ANALYSIS

The model performed nicely with a high accuracy of 98.63%. It was able to classify the gestures with good precision. No challenges were encountered during training, as the dataset was small and the architecture was simple.

Improvements could have been made . To improve generalization, we could have applied data augmentation techniques like random rotations or flipping to artificially increase the dataset size. The accuracy could have further been improved with more diverse training data or by using a larger dataset for more detailed training.