ML Lab Week 14: CNN Image Classification

Name	Diya D Bhat
SRN	PES2UG23CS183
Section	C
Date	20/11/25
Course name	Machine Learning

1. Introduction

In this lab, we built and trained a Convolutional Neural Network (CNN) to classify images of rock, paper, and scissors. The goal was to understand how image preprocessing, convolutional layers, pooling, and fully connected layers work together to perform image recognition. We also learned how to train the model, evaluate it, and test it on new images.

2. Model Architecture

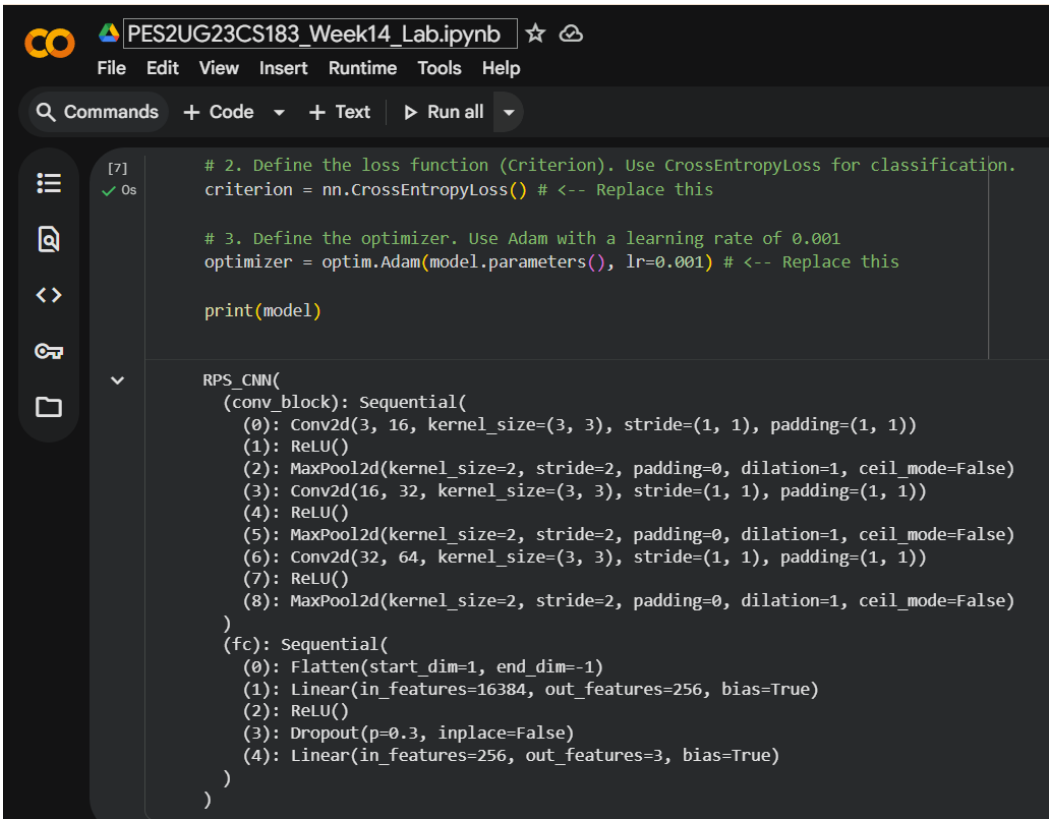
The CNN consists of three convolutional blocks, each using a 3x3 kernel, ReLU activation, and MaxPool2d(2) to reduce spatial size.

- Block 1: 3 → 16 channels
- Block 2: 16 → 32 channels
- Block 3: 32 → 64 channels

After pooling three times, the image reduces from 128×128 → 16×16, which is flattened and passed into the fully connected classifier which has the following features:

- Linear layer from 64x16x16 → 256
- Activation = ReLU activation
- Dropout = 0.3
- Final linear layer 256 → 3 for class prediction.

Screenshot of model structure:



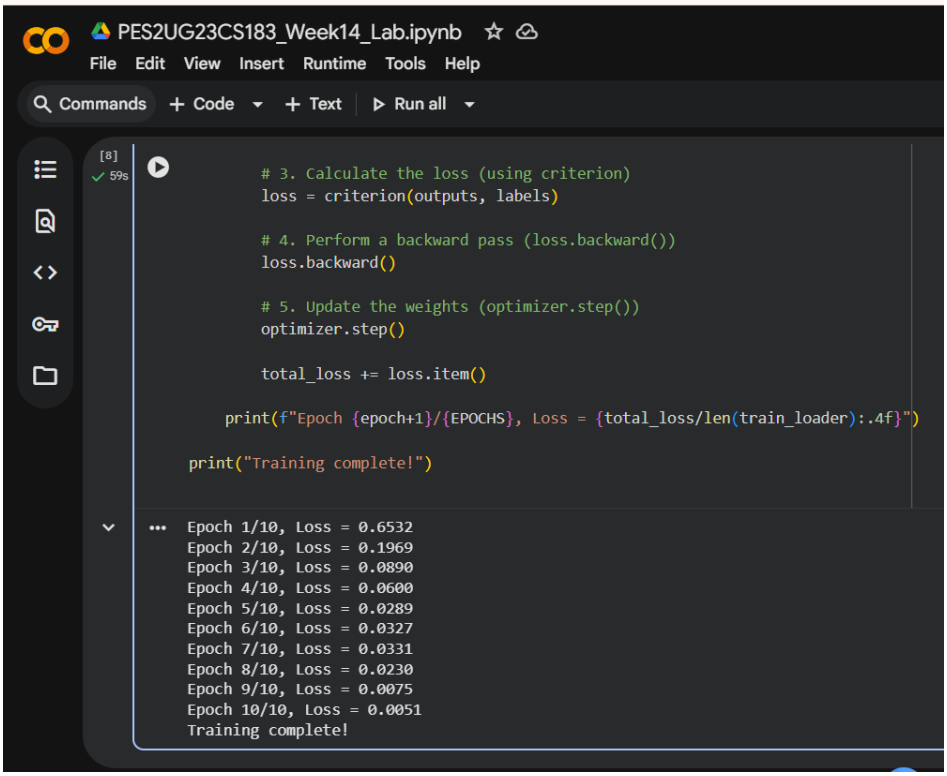
3. Training and Performance

The model was trained using:

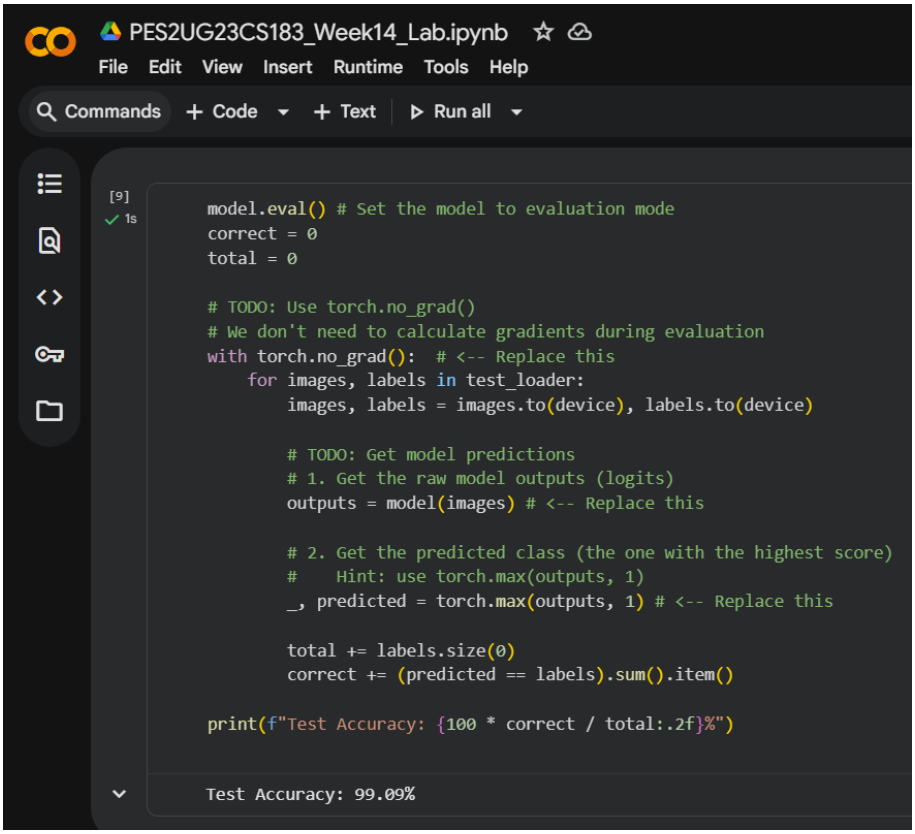
- Optimizer: Adam
- Learning rate: 0.001
- Loss function: CrossEntropyLoss
- Epochs: 10

After training, the model achieved a final test accuracy of 99.09% (replace with your actual value).

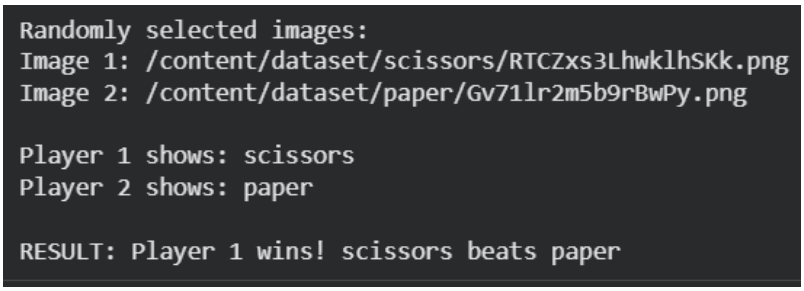
Screenshot of model training:



Screenshot of model evaluation and test accuracy:



Screenshot of the final output:



4. Conclusion and Analysis

The model performed well for a simple CNN, showing that the architecture can learn meaningful features from the RPS images. A challenge was ensuring the dataset was correctly loaded and transformed. To further improve accuracy, we could add data augmentation (rotation, flips) or use a deeper network such as a pre-trained model. These changes would help the model generalize better.