# Convolution Neural Network Lab - 6 Submission

**NAME: MITHA M K**

**SRN:PES2UG23CS339**

**COURSE: BTECH(CSE)**

# INTRODUCTION

The objective of this lab was to design, implement, and evaluate a Convolutional Neural Network (CNN) capable of classifying hand gesture images into three categories: rock, paper, and scissors. Using the Rock–Paper–Scissors dataset, the task involved performing data preprocessing, constructing a CNN model from scratch using PyTorch, training the model, evaluating its performance on unseen test images, and finally analyzing the effectiveness of the model based on the achieved accuracy.

# MODEL ARCHITECTURE

The CNN model used in this lab consists of three convolutional blocks, each composed of a convolution layer followed by a ReLU activation and MaxPooling:

- Conv Block 1:
  - Conv2d(3 → 16) with kernel size 3×3 and padding 1
  - ReLU activation
  - MaxPool2d(2) halves spatial resolution (128 → 64)
- Conv Block 2:
  - Conv2d(16 → 32) with kernel size 3×3, padding 1
  - ReLU activation
  - MaxPool2d(2) (64 → 32)
- Conv Block 3:
  - Conv2d(32 → 64) with kernel size 3×3, padding 1
  - ReLU activation
  - MaxPool2d(2) (32 → 16)

After these three blocks, the input image of size 128×128 becomes a feature map of size 64×16×16.

Fully Connected Classifier

The flattened feature vector (size 64 × 16 × 16 = 16384) is passed through:

- Linear (16384 → 256)
- ReLU activation
- Dropout (p = 0.3) to reduce overfitting

- Linear (256 → 3) output layer for the three gesture classes

This architecture extracts progressively richer spatial features in the convolutional layers and then uses a simple feed-forward classifier to map them to class labels.

# TRAINING AND PERFORMANCE

Hyperparameters Used

- Optimizer: Adam
- Loss Function: CrossEntropyLoss
- Learning Rate: 0.001
- Epochs: 10
- Batch Size: 32
- Train/Test Split: 80% / 20%

Model Performance

After training for 10 epochs, the model achieved a final test accuracy of:

Test Accuracy: 97.95%

# CONCLUSION AND ANALYSIS

Overall, the CNN performed well on the Rock–Paper–Scissors dataset and was able to learn meaningful features from the images. The model showed consistent improvement across epochs and achieved a strong test accuracy by the end of training.

Challenges faced included handling dataset paths (especially when transitioning from Colab-style paths to local machine paths), ensuring all preprocessing transforms were applied correctly, and selecting the appropriate image size for training. Training on CPU also increases training time compared to GPU.

Potential improvements for future work include:

- Adding data augmentation (random flips, rotations, brightness changes) to reduce overfitting.
- Increasing the number of epochs or adding a learning rate scheduler for better convergence.
- Trying a deeper CNN or using a pretrained model (e.g., ResNet18) to boost accuracy further.