ECS 189G-001
Deep Learning
Spring 2025
Course Project: Stage 3

Team Information

Bayesian Bandits		
Student 1: Shevangae Singh	Student 1: 918163547	Student 1: svisingh@ucdavis.edu
Student 2: Vikram Penumarti	Student 2: 920928592	Student 2: vpenumarti@ucdavis.edu
Student 3: Rajat Gupta	Student 3: 922603941	Student 3: rmgupta@ucdavis.edu

Section 1: Task Description

The task that we studied in this project is object recognition using Convolution Neural Network, the aim of the study is image classification to train three types of datasets to develop CNN models, which are MNIST for handwritten digit recognition, ORL for face recognition, and CIFAR-10 for object classification. The task aim is to develop models based on the predefined code template and is trained on given training set and testing is evaluated on a pre-defined testing set .

Section 2: Model Description

Model Architecture Overview

This CNN model implemented in the stage 3 of the project is used for image classification tasks on MNIST, CIFAR-10, and ORL datasets.

Input Layer: Receives image inputs with shapes:

- MNIST: 1x28x28 (grayscale)
- CIFAR-10: 3x32x32 (RGB)
- ORL: 1x112x92 (grayscale)

- **First Convolutional Block**:
  - Conv2D: 64 filters, 3x3 kernel, stride=1, padding=1
  - BatchNorm2D(64)
  - ReLU activation
  - MaxPooling2D(2x2)
  - Dropout(0.25)

- **Fully Connected Layers**:
  - Fully connected Layer C1: Linear(Flattened_Size, 256) → BatchNorm1D → ReLU → Dropout(0.25)
  - Fully Connected Layer C2: Linear(256, 128) → BatchNorm1D → ReLU → Dropout(0.25)
- **Output Layer**:
  - Linear(128, num_classes)
  - num_classes = 10 for MNIST & CIFAR-10, 40 for ORL.
- **Training Settings**:
  - Max Epochs: 100
  - Learning Rate: 0.001
  - Batch Size: 128
  - Weight Decay: 5e-4 for CIFAR-10

Section 3: Experiment Settings

3.1 Dataset Description

Dataset	Train Instances	Test Instances	Image Size	Channels	Labels
ORL	360	40	112x92x3	Grayscale (R channel used)	Labels from 1 to 40 (person ID)
MNIST	60,000	10,000	28x28	Grayscale	Labels from 0 to 9 (digit)
CIFAR-10	50,000	10,000	32x32x3	RGB	Labels from 0 to 9 (object)

3.2 Detailed Experimental Setups

Parameter	Value
Convolutional Layers (conv_layers)	2
Kernel Size (kernel_size)	3x3
Stride	1
Padding	1
Pooling (pool_size)	MaxPooling 2x2 after each conv block
Fully Connected Layers (fc_layers)	2
FC Units	[256, 128]
Dropout Rate (dropout_rate)	0.25
In-Channels (in_channels)	1 (for ORL & MNIST), 3 (for CIFAR-10)
Number of Output Classes (num_classes)	10 (MNIST, CIFAR-10), 40 (ORL)

Detailed CNN Architecture

- **Input Layer:**
    - **ORL:** (1, 112, 92)
    - **MNIST:** (1, 28, 28)

    - **CIFAR-10:** (3, 32, 32)

- **Convolutional Layer 1 and 2:**
    - **Conv2D in layer 1:** in_channels (1 or 3) → out_channels 64 **Layer 2:** in_channels 64 → out_channels 128
    - **Kernel Size:** 3x3, Stride: 1, Padding: 1
    - **BatchNorm2d(64 layer 1 and 128 layer 2)**
    - **ReLU Activation**
    - **MaxPool2d(2x2)**
    - **Dropout(0.25)**
- **Flatten Layer:**
    - **Flattens feature maps into 1D vector based on dataset-specific dimensions.**
- **Fully Connected Layer 1& 2:**
    - **Linear:** flattened size → 256 units→ 128 units(layer 2)
    - **BatchNorm1d(256)→BatchNorm1d(128)**
    - **ReLU Activation**
    - **Dropout(0.25)**
- **Output Layer:**
    - **Linear:** 128 → num_classes (10 or 40)

Hyperparameters

Parameter	Value
Max Epochs	100
Batch Size	128
Learning Rate	0.001
Weight Decay	5e-4 (for CIFAR-10)

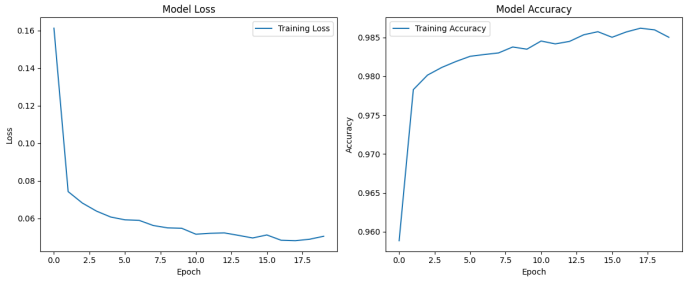
3.3 Evaluation Metrics

In this stage of the project, we evaluated the classification performance of the Convolutional Neural Network (CNN) models using standard performance metrics. These metrics are computed and visualized through the `Evaluate_CNN` class, which serves as a performance metric for result evaluation.

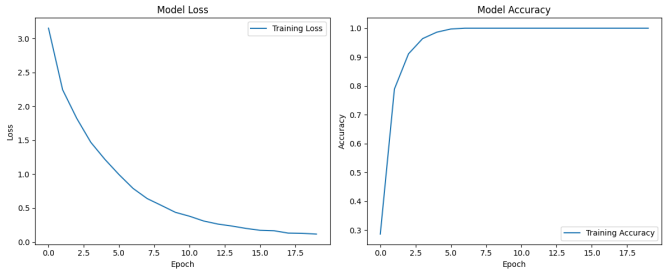
Metrics Used

1. **Accuracy**

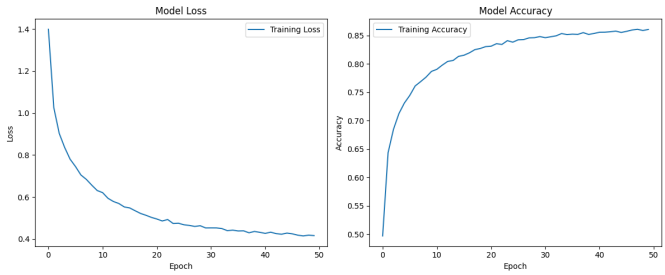
The proportion of correctly classified samples over the total number of samples. It provides an overall measure of how well the model performs .



MNIST Dataset



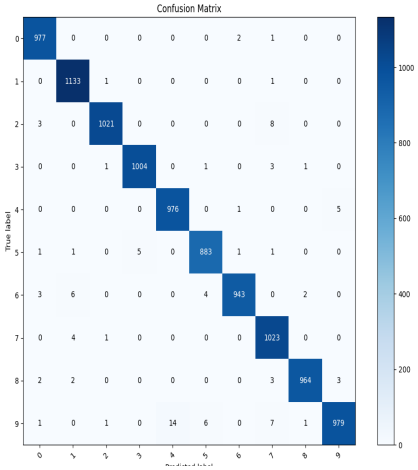
ORL Dataset

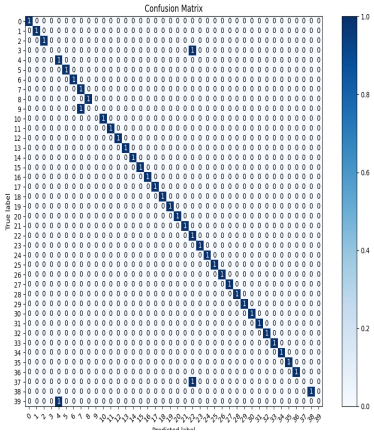


CIFAR Dataset

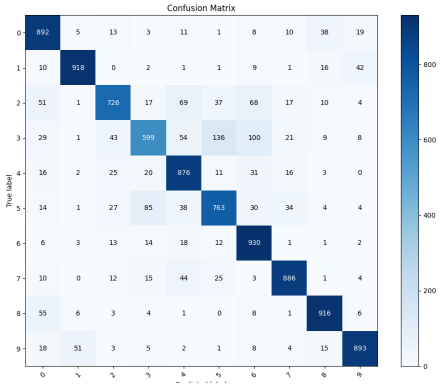
2. **Precision (Macro-Averaged)**
   Precision measures the ratio of true positive predictions to all positive predictions made by the model.
3. **Recall (Macro-Averaged)**
   Recall calculates the ratio of true positive predictions to all actual positive instances
4. **F1 Score (Macro-Averaged)**
   The harmonic mean of macro-averaged precision and recall.
5. **F1 Score (Weighted-Averaged)**
   This metric is also computed with weighted averaging, where each class's contribution is weighted by its actual instance count (support).
6. **Confusion Matrix Visualization**

In addition to numerical metrics, a confusion matrix is generated to visualize the model's prediction distribution across classes.





MNIST Dataset **ORL Dataset**

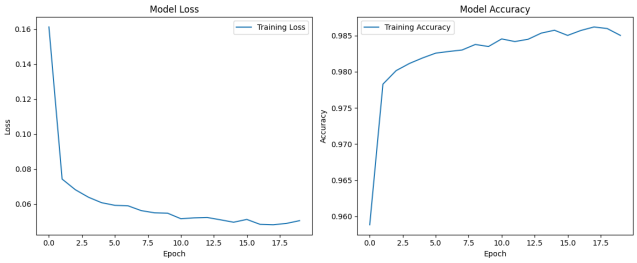


CIFAR Dataset

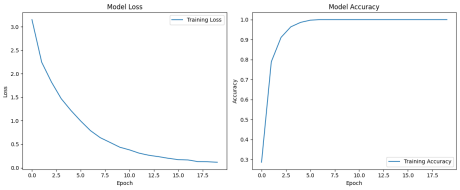
3.4 Source Code

https://gitfront.io/r/raj978/qw2MfpMwrHpw/ECS-189G-DeepLearning/

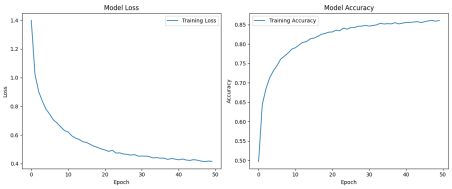
3.5 Training Convergence Plot



MNIST Dataset



ORL Dataset



CIFAR Dataset

3.6 Model Performance

3.6 Model Performance

The model performance was evaluated on three datasets: **MNIST**, **ORL**, and **CIFAR-10**, using standard metrics — **Accuracy, Precision, Recall, and F1-score** on the test set.
The results are summarized below:

Dataset	Accuracy	F1-Score
MNIST	99.03%	99.03
ORL	90.00%	87.08
CIFAR-10	83.99%	83.76