DEPARTMENT OF COMPUTER SCIENCE

UNIVERSITY OF DELHI

DATA MINING PROJECT

DOODLE CLASSIFICATION

REPORT FILE

SUBMITTED TO :

DR. BHARTI RANA MA'AM.

(DEPARTMENT OF COMPUTER SCIENCE)

SUBMITTED BY :

ARYAN THAPLIYAL (13)

MANPREET SINGH (26)

MSC COMPUTER SCIENCE , FIRST YEAR.

DOODLE CLASSIFICATION

1. Introduction

- The goal of this project is to develop a Convolutional Neural Network (CNN) model to classify hand-drawn doodles into predefined categories. Using the QuickDraw dataset, which includes various doodle categories, we aim to create a model that can learn distinguishing features and accurately classify doodles based on learned patterns.

2. Dataset and Preprocessing

Dataset Description

The QuickDraw dataset contains doodles categorized into different classes. For this project, we selected six classes:

1. Airplane
2. Apple
3. Banana
4. Bird
5. Bicycle
6. Clock

Each class contains a set of `.npy` files that store image data in a NumPy format, where each image is 28x28 pixels in grayscale.

Data Downloading and Loading

Using a list of URLs, we downloaded the required `.npy` files into the working directory. A function `download_files` was implemented to handle downloading and storing each file. The `load_npy_files` function was used to load each class's `.npy` files into NumPy arrays, which were then processed and combined into a `CustomDataset` class, compatible with PyTorch's `DataLoader`.

Data Preprocessing

1. Each doodle image is reshaped to 28x28 pixels.

2. The data is organized into training (70%), validation (15%), and test (15%) sets.
3. The `DataLoader` utility is used to create data loaders with a batch size of 64 for efficient mini-batch processing.

3. Model Architecture

The classification model is a Convolutional Neural Network (CNN) defined in the `CNN` class. The architecture includes:

- **Convolutional Layers**:
    a. Three blocks of convolutional layers with ReLU activations, each followed by a max pooling layer for down-sampling.
    b. Final output shape from the convolutional blocks is 7x7x16.
- **Fully Connected Layers**:
    a. A linear layer that maps the flattened feature map (784 units) to 100 hidden units.
    b. An output layer maps 100 units to 6 units, corresponding to the 6 classes.

This architecture leverages convolutional layers to learn spatial hierarchies, enhancing the model's ability to recognize doodle features.

4. Training and Evaluation

- ### Hyperparameters

    1. **Learning Rate**: 1×10^-3
    2. **Batch Size**: 64
    3. **Epochs**: 5
    4. **Loss Function**: Cross-Entropy Loss
    5. **Optimizer**: Adam

Training Process

1. The model was trained over five epochs. For each batch:
2. The CNN model processed the images, computing predictions.
3. Cross-Entropy loss was calculated, followed by backpropagation and an optimization step.
4. The `get_accuracy` function was used to calculate accuracy for both the training and validation sets after each epoch.

RELATED WORK

PAPER 1 : FREE HAND SKETCH RECONGNITION CLASIFICATION.

- **Objective**: Improve free-hand sketch recognition accuracy using deep CNNs, particularly modified ResNet architectures.

- **Data Source**: 20,000 sketches across 250 classes from Eitz et al., collected through Amazon Mechanical Turk.

- **Data Preprocessing**:
  1. Resized images to 128x128 pixels.
  2. Converted images to grayscale.
  3. Augmented training samples by flipping images.

- **Data Mining Techniques**:
  1. Employed various CNN architectures with residual networks.
  2. Tested the effects of dropout, batch normalization, and different network widths and depths.
  3. Used the Adam optimization algorithm for training.

- **Result**: Best model achieved a test accuracy of 65.6%, which is competitive with traditional methods but lower than recent state-of-the-art models.

- **Advantages**:
  1. Demonstrated that residual networks improve training stability and classification accuracy.
  2. Established that depth is more impactful than width for sketch recognition tasks.

- **Disadvantages**:
  1. Accuracy still lower than state-of-the-art methods.
  2. Limited exploration of architectures deeper than 25 layers.

PAPER 2 : A Neural Representation of Sketch Drawings.

- **Objective**: Develop a model (*sketch-rnn*) for generating and completing vector-based sketches, both conditionally and unconditionally.

- **Data Source**: *Quick, Draw!* dataset, containing 70K training samples per class, representing sketches as sequences of pen strokes.

- **Data Preprocessing**: Stroke sequences were simplified using the Ramer–Douglas–Peucker algorithm, and offsets were normalized for consistency across the dataset.

- **Data Mining Techniques**:
  1. Sequence-to-Sequence Variational Autoencoder (VAE) with a bidirectional RNN encoder and autoregressive RNN decoder.
  2. Temperature scaling for output randomness and hybrid loss function with Reconstruction and KL Divergence terms.

- **Result**: *Sketch-rnn* generates coherent sketches and performs well on tasks like sketch completion, latent space interpolation, and analogy-based sketch modification.

- **Advantages**:
  1. High versatility in sketch generation and manipulation.
  2. Open-source model and dataset, promoting further research in sketch generation.
- **Disadvantages**:
  1. Difficulty in handling complex sketches with high detail.
  2. Limited capability to model large numbers of classes simultaneously without interclass feature blending.

EXPERIMENTAL RESULT

The accuracy obtained by CNN on this dataset is :

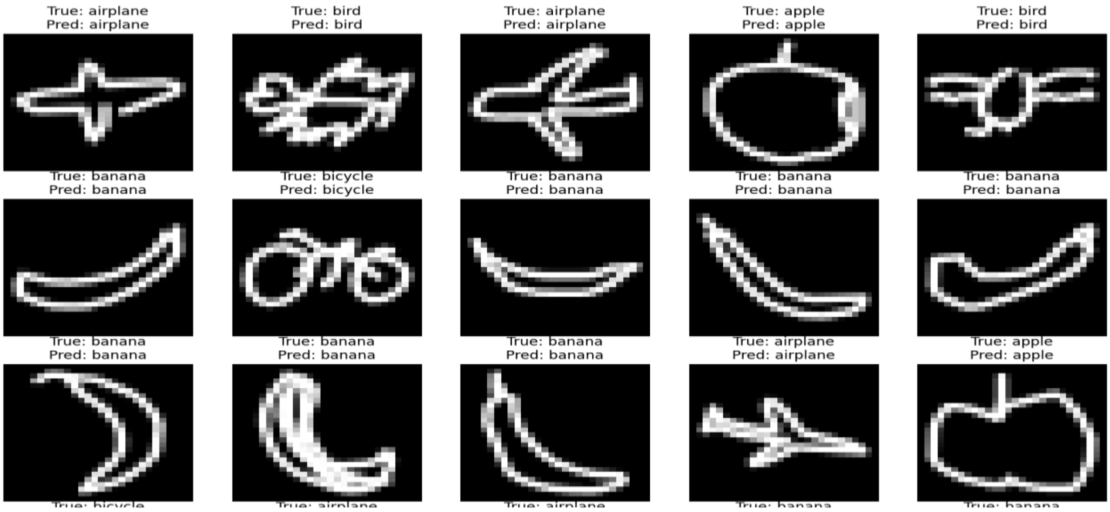
CNN :

Training Accuracy : 96.93%

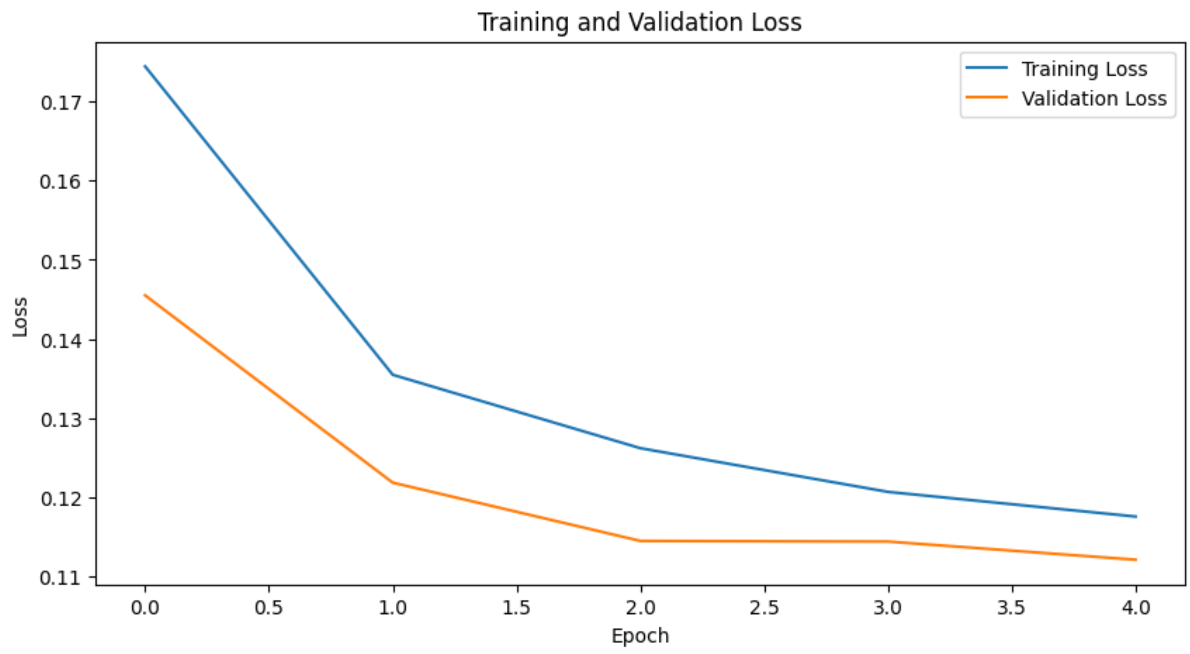
Validation Accuracy : 96.43%

Test Accuracy : 96.48%

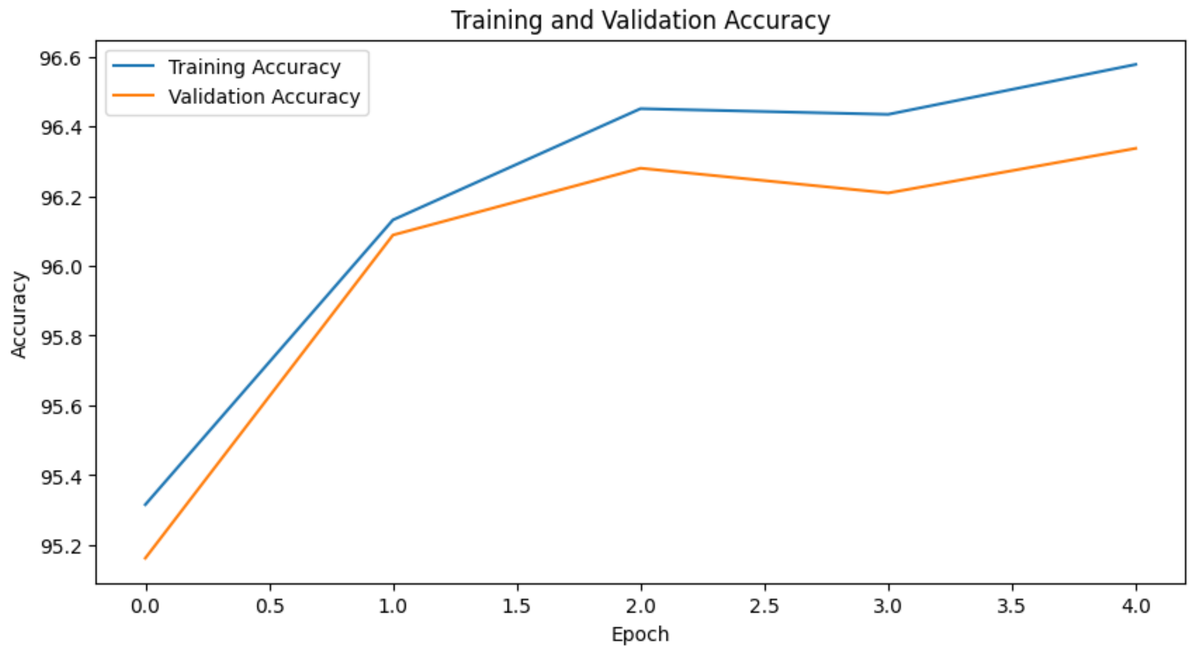
Predictions on Test Data :



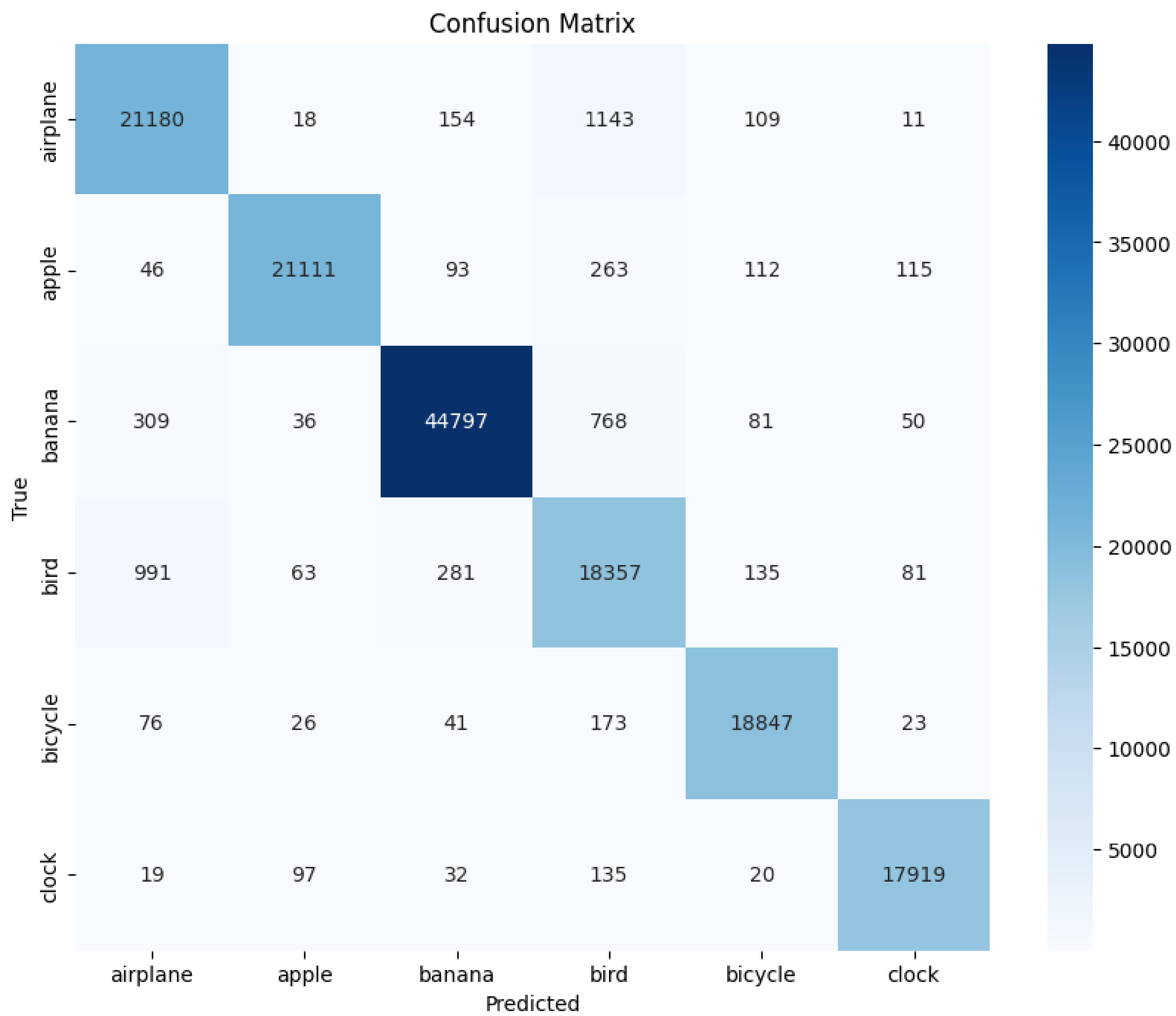
Training and Validation Loss :



Training and Validation Accuracy :



Confusion Matrix :



Confusion Matrix
airplane
apple
banana
bird
bicycle
clock
True
Predicted
21180
18
154
1143
109
11
46
21111
93
263
112
115
309
36
44797
768
81
50
991
63
281
18357
135
81
76
26
41
173
18847
23
19
97
32
135
20
17919
40000
35000
30000
25000
20000
15000
10000
5000

Conclusion

PAPER 1 : FREE HAND SKETCH RECONGNITION CLASIFICATION.

The study highlights that deeper CNN architectures improve sketch classification accuracy, though significant challenges remain due to intraclass variability and class overlap. Future work could explore deeper networks, additional regularization methods, and alternative datasets to enhance model performance.

PAPER 2 : A Neural Representation of Sketch Drawings.

The *sketch-rnn* model demonstrates effective sketch generation, completion, and manipulation capabilities using RNN-based architectures. By making available the model and dataset, the authors encourage further research in vector image generation and manipulation.

REFERENCES :

[1] Wayne Lu , Elizabeth Tran , Stanford University , **(2017)** Free-hand Sketch Recognition Classification.

[2] David Ha , Douglas Eck , Google Brain , (2017) A Neural Representation of Sketch Drawings.