Department of Computer Science and Engineering
School of Electrical and Computer Sciences
IIT Bhubaneswar

Deep Learning Project
on

Street View House Number Recognition

•••

By
Sourav Makal-24AI06005
Dibyajyoti Nayak-24AI06012
Puvvula Nikhileswari-24AI06017
Polasani Likhitha Reddy-S25EC08002

Contents

- Problem Statement

- Dataset Description

- Pre-processing

- CNN Models

  - Custom Architecture

  - ResNet18

- Results

Problem Statement

- Recognition of street-level house numbers is a critical task for urban mapping, navigation systems, and infrastructure assessment.
- In this project, we aim to develop a deep learning model that can accurately recognize multi-digit house numbers from real-world images using the Street View House Numbers (SVHN) dataset.

Goal

- The goal is to automate visual number recognition in complex environments.

Dataset Description

- The **SVHN dataset** is a real-world image dataset for developing models that recognize digits from street view imagery.

- Images are in `.png` format with variable resolution and real-world distortions.

- Each digit in the image is labeled and has a bounding box stored in a `.mat` file.

File	Description	No of Digits
train.zip	Labeled training Digits	73,257
test.zip	Labeled test Digits	26,032

Dataset Description(Cont.)

Fig 1 : Sample Images

Labels

- **10 classes** for digits: '1' → 1, '9' → 9, '0' → 10 (not 0).
- Each image may contain **multiple digits**.

Annotation Format

- **digitStruct.mat:** Contains bounding box info per digit:
  - **name:** Image filename
 - **bbox:** Array of structs → height, width, label, top, left

Preprocessing

1. Extracted Data from .tar.gz archives (train/test).
2. Parsed digitStruct.mat for filenames, bounding boxes, and digit labels.
3. Converted labels (10 → 0) and created structured CSV files.
4. Filled missing digits with -1 for consistency.

Visualizations

- Digit length distribution.
- Digit frequency.

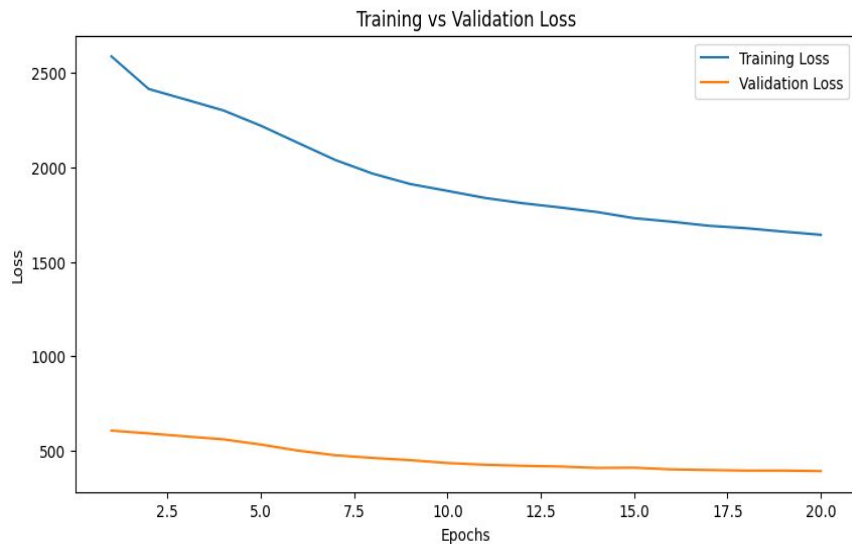**Output:** Clean CSVs ready for model training.

CNN Architectures
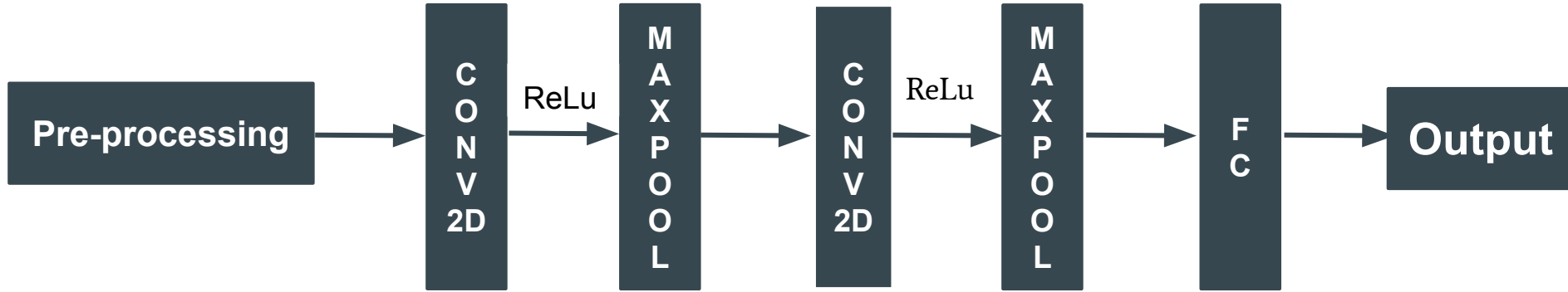
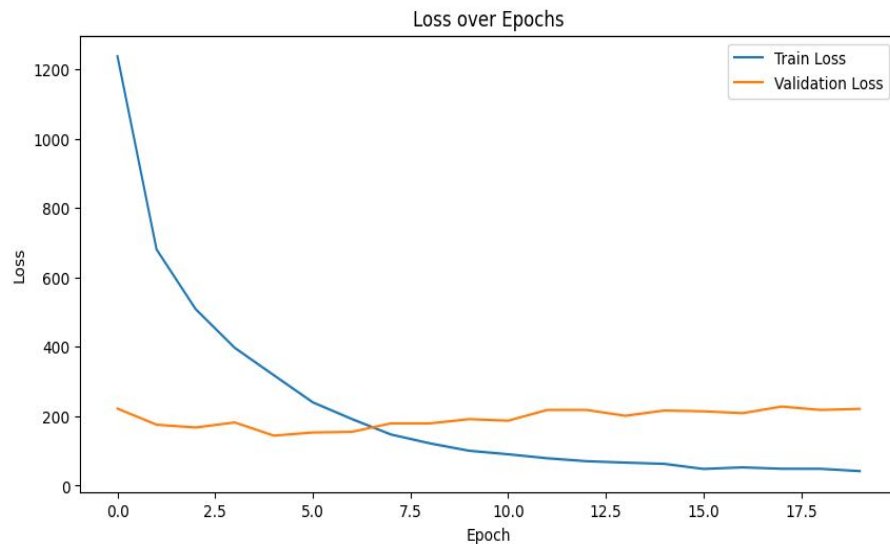Architecture 1 - Custom Architecture

Architecture 2 - Modified Pretrained Architecture

Results

Custom Architecture's
Epoch vs Loss Curve

Test Accuracy : 72.36%

ResNet Architecture's
Epoch vs Loss Curve

Test Accuracy : 86.33%