Project Progress Report

Austen Ang Ming Xuan, Cassandra Chew Ying Ying, Diren D. Bharwani, Leonard Lee Hon Han, Lee Ming Yang

Abstract

The goal of this project is to create an image recognition algorithm that can detect fixed length handwritten binary numbers and convert them to decimals.

Project Overview

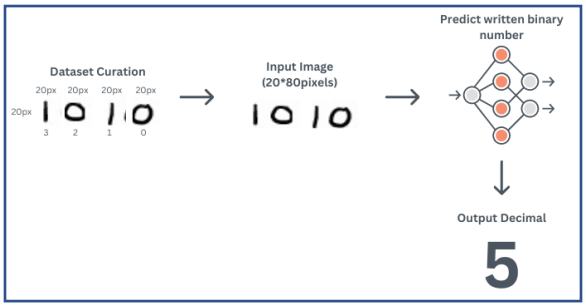

Figure 1 - Project Overview

As seen in *Figure 1*, the project can be segmented into the following phases:

1. Dataset Curation
2. Binary Classification

Machine Learning Lifecycle

1. Data Collection

To meet our algorithm's requirements, we'll create a dataset using the MNIST database, adapting the images to be 20x40 pixels. This simplifies the project, as the MNIST database examples are already normalized to 20x20 pixels, and we only need to adjust the size.

source: http://yann.lecun.com/exdb/mnist/

2. Data Preparation

The MNIST database of handwritten digits, has a training set of 60,000 examples and a testing set of 10,000 examples, where each digit has been normalized and centered in a fixed-sized image of 20x20 pixels.

The general idea of curating our own dataset is as follows:

1. *Drop all examples which are not labeled as 0 or 1.*
2. *For each of the position [0, 3],*
    a. *Randomly select an example from the dataset.*
    b. *Append the data set a M x 2 matrix.*
        i. *M represents the size of the image data(20x20px).*
        ii. *First column contains the image data.*
        iii. *Second column contains the encoded position [0,3].*
3. *Repeat N times. Where N represents the desired number of examples.*

Note: We are considering this as the naïve method of generating our dataset as there are better ways to go about it.

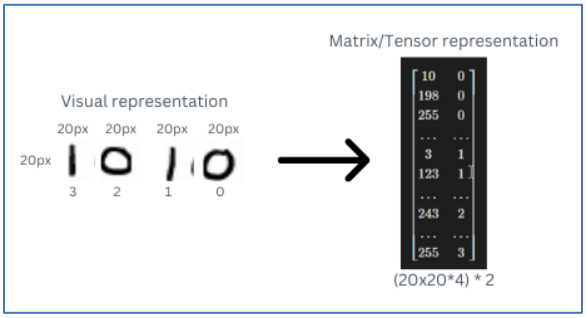

Figure 2 - Representation of one train/test example

3. Model Training

The plan is to use a Convoluted Neural Network to predict the handwritten binary number and classify the image into one of the 16 classes. Details of the CNN, including its layers, activations, and purposes, will be discussed later.

4. Model Evaluation

The following are some evaluation metrics would be employed:

- Confusion Matrix
- Accuracy

Proposed Frameworks

- NumPy
- Matplotlib
- Seaborn
- TensorFlow

Risks

If the model cannot accurately predict the position of each digit, we will mitigate this by predicting the digit of each 20*20px segmented image and manually reconstruct the predicted binary number.

Conclusion

The objective of this project is to provide us with hands-on experience with industry frameworks and showcase the knowledge we have gained throughout the course.