Recognition of Handwritten numerical digits Using Deep learning Network

Liuwei, Zheng (Joanna) 201918020402
Supervisor: Dr. Grace Ugochi Nneji

Abstract

Handwritten numerical digit recognition using deep learning is an effective way to recognize numbers written by hand. The challenge of the project is that it is a complex task due to the variability in handwriting styles and the similarity between certain numbers. This project used deep learning algorithms, based on the Tensorflow library to construct Visual Geometry Group (VGG) model and Mixed National Institute of Standards and Technology (MNIST) dataset to train the model, to recognize patterns in handwritten digits and to make predictions. The final recognition model performed effectively in measurements such as achieved a 99.4% accuracy rate. OpenCV to perform target detection and stroke processing on numbers on photographs

Introduction

In recent years, deep learning has become widely used due to its ability to accurately and efficiently solve complex problems. The classification and recognition of digital images have also become increasingly sophisticated. Handwritten digital image classification and recognition has made significant contributions to industries such as finance, healthcare, transportation, and retail, making them more profitable. Convolutional neural networks are important algorithms in image pattern recognition that can achieve deep learning. These networks mimic biological neural networks and consist of multiple layers of nodes, including input nodes, hidden layer nodes, and output nodes. They can be used for handwritten digital image classification and recognition.

Dataset

This project used the Mixed National Institute of Standards and Technology (MNIST) dataset which contains a training set of 60,000 training set and 10,000 test set with a 28×28 pixels.

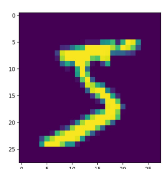

Figure 1: An example diagram from MNIST dataset

Image Pre-processing

The MNIST dataset uses a version of npz, the array storage method provided by Numpy. The MNIST dataset size was changed to 32×32 by padding and stacking because the resize image too large causes too much RAM exhaustion.

Futher Work

To address these limitations, the current parameters of VGG16 could be adjusted inthe future to obtain models with better experimental results, for example byspeaking of the images of the dataset as grey-scale images, changing the size ofthe resize, changing the optimisation algorithm, etc. Future work will focus ondebugging the order of the numbers displayed and exploring the use of datasetswith more complex backgrounds for training the model. Simultaneously, While the VGG16 model has shown promising results, there are many other architectures thatcould be explored for handwritten digit recognition. It is worth to test other models tosee if they can improve upon the current VGG model.

Methodology

This project used deep learning algorithms based on Tensorflow to construct the Visual Geometry Group (VGG) model. VGG is structured with a deeper network, making the feature map wider and more suitable for large datasets. It is smaller convolution kernel, which ensures a perceptual field of view, and its ability to reduce the parameters of the convolution layers.

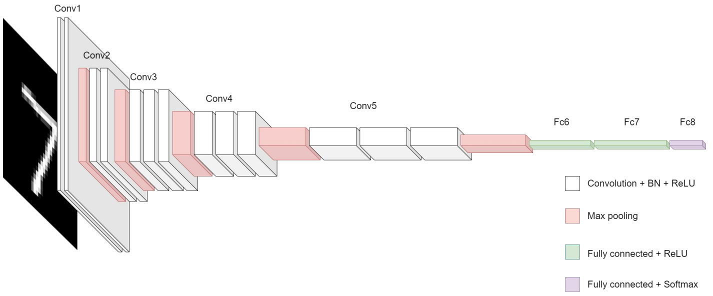

Figure 2: The VGG16 Architecture for the Project

Pre-processing of digital images, uploaded by user, includes greyscaling of digital images, handwritten digital smoothing, handwritten digital binarisation, digital image removal of distracting parts of the image, digital image normalisation and digital image refinement.

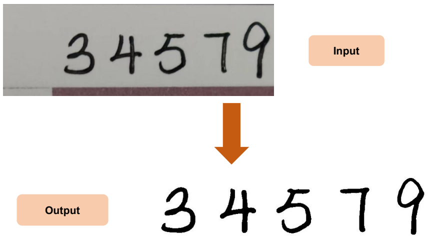

Figure 3: Input and output for target detection

Result & Discussion

The final recognition model was evaluated by the model.evaluate() function to obtain an accuracy of 99.2% and a loss of 5.4% when the epochs were 50.

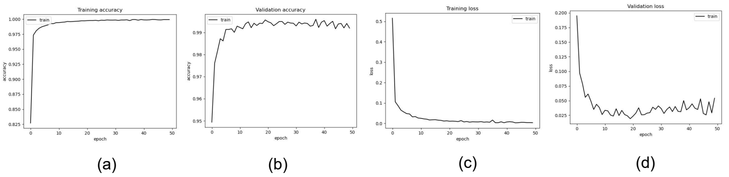

Figure 4: Training accuracy, validation accuracy and training loss and validation loss

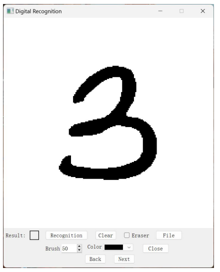

Figure 5: Digital recognition UI and one of the results of target detection

The results show that the proposed model achieved high accuracy on the MNISdataset. It is concluded that these technologies have demonstrated high accuracyand efficiency in recognizing and classifying handwritten digits. Additionally, theimplementation of target detection to process user uploaded images and GUI hasimproved the usability and accessibility of the system for end-users. However, thelimitations of the project were discovered, which is the difficulty in recognizingcolored numbers and some confusion in the order of the numbers output.

Reference

[1] Simonyan, K. and Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition.* 3rd International Conference on Learning Representations (ICLR 2015), 1–14. Available at: http://arxiv.org/abs/1409.1556.

成都理工大学中英合作办学
CDUT Sino-British Collaborative Education
OXFORD BROOKES UNIVERSITY