COGS 400 – Group 16

Hanzi Recognition with Neural Networks

Group members

Yifei	Zhaoyu	Ian	Pengyu
20054101	20010277	10190262	20024936
yifei.yin@queensu.ca	15zy19@queensu.ca	14yc65@queensu.ca	16pz1@queensu.ca

1 Problem and Motivation

The problem is identifying handwriting Chinese characters using several different artificial neural network techniques.

By solving this problem, system like writing to speech can be developed. The problem already has numerous solutions with room of improvements available. We chose this as the topic of our research because we could build upon current research using what we have learnt in this course. By experimenting different neural network models, we hope to see improvements in recognition accuracy as well as recognition time.

2.1 Dataset Usage

The full name of the dataset that will be used in this project is Harbin Institute of Technology Opening Recognition Corpus for Chinese Characters (HIT-OR3C). it was first created in April, 2010 and updated once only in January 2011. it can be download from **the website showed in the appendix** to do the offline training and testing. The dataset contains 5 subclasses containing digits, letter, GB1, GB2 and a documents for test use. The first four subsets of characters contain 6825 classes produced by 832,650 samples in total, and the last document data have been post-processed and split into individual characters stored sequentially in a single image and a single vector file.

Appendix(REF)
http://www.iapr-tc11.org/dataset/OR3C_DAS2010/v1.1/OR3C/offline/character.rar

2.2 ANN Model for Each Member

Yifei: Adaptive Multilayer Network. Use pruning to decrease the amount of insignificant connections in order to improve network efficiency while retaining the precision.

Zhaoyu: Convolutional neural networks, which is really practical in image recognition and object classification. For instance, regardless of image size, tiling regions of size 5 x 5, each with the same shared weights, requires only 25 learnable parameters. In this way, it resolves the vanishing or exploding gradients problem in training traditional multi-layer neural networks with many layers by using backpropagation.

Ian: Learning Vector Quantization (LVQ) a form of supervised learning neural network, that uses winner-takes-all hebbian learning. Similar to a k-nearest neighbour algorithm, LVQ is a classification algorithm, that classifies data points based off k of its most similar prototypes. Using euclidean distances, the prototype that is the closest to the data point is selected. Then, the prototype's output is compared with the data point's output, and moved closer if matches, or moved away if it doesn't match. The distance that the prototype is moved is based off the learning rate.

Eric Pengyu: A deep convolutional neural network can be implemented in this project. Since the Chinese handwritten character recognition is more complex than that of either the Alphabets or digits because of having more categories, sophisticated structures and different handwriting skills depend on individuals, a relatively deep neural network will be built to make sure the model is expressive enough to achieve a high accuracy. Within the network, the first few layers can be convolutional layers specifically called "Max-polling layer" that extracts the features from the Chinese handwritten characters. After doing the max-polling, FC layers will be attached to the bottom to get an array of scores that stand for every possible class. Finally, a softmax function will be implemented to let the model make the final decision.

2.3 Validation Plan

Validations for the project can be done using n-fold cross validation technique. The n-fold cross validation takes the training data set that was given and then divides the set in to n folds. Looping through n times, the neural network will train on n-1 folds of the data and select a fold to be the test data set. The results of the model is saved and compared with performance from other iterations. The iteration with the best performance will be selected as the final model. The n fold cross validation would serve as addition validation on top of the standard train and test procedure of the data. Another variation of the n-fold cross validation that could be considered would be using n-folds on the entire data set instead of just the training data set. Using the entire data set would show even more variations in neural network models, providing higher chances of finding the model with optimal performance.