Handwritten Digit Recognition

Liyan Chen
Stevens Institute of Technology
lchen39@stevens.edu

Shangxing Jiang
Stevens Institute of Technology
shangxingjiang1@gmail.com

***Abstract*—This paper is a mid-term report for our handwritten digit recognition project. The project focus on building machine learning model to solve a special image classification task by implementing K-NN, SVM and CNN algorithm. The experiment will be conducted on MNIST digit database.**

***Index Terms*—handwritten digit recognition, k-Nearest Neighbor, Support Vector Machine, Convolution Neural Networks, TensorFlow**

I. Introduction

Recognition is a popular topic with commercial importance in machine learning area. It covers various fields such as face recognition, images recognition, character recognition, etc. Handwritten digit recognition today has been an indispensable component in many applications: check verification, postal address reading in envelope, pin code recognition and so on. Most of the digit recognition systems are based on printed front, yet the success has been extended to handwritten digits. Unlike printed digits, the recognition for handwritten digits is more difficult due to the unique writing styles of different writers which leads to various size and shape of the digits. Recognition is also a typical classification task, the machine learning model for recognition is to classify the data into a right group according to several features. In this case, the main mission of handwritten digits recognition system is to classify the input digit image into the most related digit group. There are many different machine learning methods for classification.

A. K-Nearest Neighbor

K-Nearest Neighbor (KNN) is instance-based method used for classification. This algorithm assumes all instances correspond to points in the n-dimensional space R^n. The nearest neighbors of an instance are defined in terms of the standard Euclidean distance. By calculating the distance, input instances are divided into different categories with their nearest neighbors. [1]

B. Support Vector Machine

Support Vector Machine (SVM) is a supervised learning model used for classification and regression analysis. [2] The fundamental idea behind SVM is to mark the training dataset into separate categories which divided by widest possible gap, the model then maps new examples into related categories based on which side of the gap they fall. [3]

C. Convolution Neural Networks

Over last few years, deep learning techniques are also widely applied on classification problems. Deep learning has many architectures such as Convolution Neural Networks (CNN). CNN is a multi-layer feed-forward neural network that extract properties from the input data. CNN trained with neural network back-propagation algorithm. CNN have the ability to learn complex, high-dimensional, non-linear mappings from very large number of data (images). Moreover, CNN shows an excellent recognition rates for characters and digits recognition. [4]

This project aims to build a handwritten digit recognition system which implemented in Python 3. The study will start from the pre-process based on KNN and SVM algorithm, then extend to the application of CNN. The project will also conclude the comparison of different learning models, not only limited to the accuracy but also their advantages and disadvantages.

II. Related work

In recent years various methods have been proposed to solve the handwritten digit recognition problem with high accuracy. In 2012, Niu and Suen applied Convolutional Neural Network (CNN) and Support Vector Machine (SVM) on a handwritten digit recognition experiment. There Experiments have been conducted on MNIST digit database. They achieve recognition rate of 94.40% and 5.60% with rejection. [5]

In 2014, Ravi and Venkateswarlu successfully used K-Nearest Neighbour Classifier to recognize the handwritten digit images from MNIST database. They presented a new approach based on four different types of structural features namely, number of holes, water reservoirs in four directions, maximum profile distances in four directions, and fill-hole density for the recognition of digits. 5000 images are used to test the proposed method a total 5000 numeral images are tested and got 96.94% recognition rate. [6]

In 2015, Eva and Nebojsa described an algorithm for handwritten digit recognition based on projections histograms. Classification is facilitated by carefully tuned 45 support vector machines (SVM) using One Against One strategy. Their proposed algorithm was tested on standard benchmark images from MNIST database and it achieved remarkable global accuracy of 99.05%. [7]

In 2016, Yan and JunMin introduced Ncfm (No combination of feature maps), a novel technique to improve the performance of CNNs and applied this technique to digit recognition problem. The Ncfm technique converges faster and performs better than Cfm (Combination of feature maps) with fewer filters. Through the type of feature map, experimental evaluation show that the performance was improved and they achieved the state-of-the-art performance with 99.81% accuracy rate on the MNIST datasets. [8]

In 2017, Tohidul and Ghulam implemented a multi-layer fully connected neural network with one hidden layer for handwritten digits recognition. The testing has been conducted from publicly available MNIST handwritten database. Their multi-layer artificial neural network has an accuracy of 99.60% with test performance. [9]

Summarizing from the related papers, we find out several useful learning models to build the recognition system. The project will also collect training and test set from MNIST digit database, which is the most common handwritten digits database in machine learning field. Detail description of this database will be presented in next section.

III. Dataset and visualization

The MNIST database (Modified National Institute of Standards and Technology database) is a subset of a larger set available from NIST. It has a training set of 60,000 examples, and a test set of 10,000 examples. The training dataset was taken from American Census Bureau employees, while the testing dataset was taken from American high school students. The black and white images from NIST were normalized to fit into a 28x28 pixel bounding box and anti-aliased, which introduced grayscale levels. [10]

By using Numpy and Matplotlib, we decompressed the data from MNIST database and read into Numpy array. We first finished the vector-to-image process, the digit picture we got as figure 1.

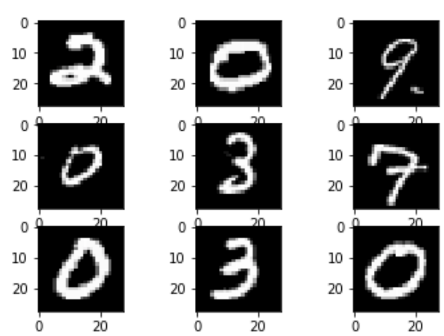

Fig. 1. Images of some training data.

Considering for reuse and later classification, we saved the data into 4 .csv files. They are:

train_set.csv (109.5 MB)
train_label.csv (120 KB)
test_set.csv (18.3 MB)
test_label.csv (20 KB)

By the end of the project, we wish that we can successfully recognize the handwritten digits of ourselves. Thus, we also created a image-to-vector function that can transmit the photo into the same format of our training dataset.

TABLE I
STATISTIC OF DATASET

Digit	***Training set***	***Test set***
0	5923	980
1	6742	1135
2	5958	1032
3	6131	1010
4	5842	982
5	5421	892
6	5918	958
7	6265	1028
8	5851	974
9	5949	1009
Total	60000	10000

IV. Implementation

A. K-Nearest Neighbor

We first chose to implemented the K-NN algorithm on the dataset. Loading in the dataset, each sample can be seen as a vector with size of (1,784). Supposed α and β are both N-dimensional vector, to calculate the Euclidean distance of this two vectors, a equation can be used:

$$d(\alpha, \beta) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \tag{1}$$

Here we set k=30, using majority vote from the 30 nearest neighbors to decide the catagory of input data. We run the model for three times and increased the size of test set in turns.

TABLE II
KNN TRAINING RESULT

Test size	**Error Count**	**Accuracy**
100	2	0.9800
1000	57	0.9430
10000	406	0.9594

B. Support Vector Machine

Then we turn to SVM and added polynomial features. For handwritten digit recognition, we have 10 catagories for the whole dataset. The problem now should be considered as a multi-catagory classification. Each digit can be assigned to a binary classification. Here we chose a classification strategy called one-vs-rest (OVR). Guiding by OVR, only one digit catagory is treat as positive example in each classify when the rest are all negative. In the testing stage, if just one classify comes out positive result, corresponding positive catagory is the final decision of the test data. For this testwe imported the SVM model from sklearn and parameters were set as following: C=100.0, kernel=poly, gamma=0.03.

TABLE III
STATISTIC OF DATASET

Digit	Test Size	Error Count	Accuracy
0	980	8	0.9918
1	1135	9	0.9921
2	1032	26	0.9748
3	1010	23	0.9772
4	982	17	0.9827
5	892	25	0.9720
6	958	21	0.9781
7	1028	28	0.9728
8	974	24	0.9754
9	1009	32	0.9683
Total	10000	213	0.9785

V. Feature work

Although KNN and SVM experiment run out pretty results of 95.94% and 97.85% recognition rate. Some weak points should not be ignored at this stage. No doubt that both two algorithms lead to tedious computation task. KNN is a lazy learning, when the training sample of different catagories is not equal, k-nearest neighbors of a new example may have greater probability to be the majority catagory. SVM shows lack of data sensitivity and really relies on the choice of Kernel function when solving nonlinear problem. Meanwhile, we can't extract features from the raw data when using these two methods, or saying that we can't well explain the output of KNN and SVM.

Later we will begin to set up multi-layer Convolution Neural Network using TensorFlow. Extracting feature maps from training dataset will be an important task for us to build an accurate recognition system.

References

[1] Mitchell T M. Machine learning. 1997[J]. Burr Ridge, IL: McGraw Hill, 1997, 45(37): 870-877.

[2] Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks". Machine Learning. 20 (3): 273297. doi:10.1007/BF00994018.

[3] Yang G, Yu S, Dong H, et al. Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction[J]. IEEE transactions on medical imaging, 2018, 37(6): 1310-1321.

[4] Maitra, D.S., Bhattacharya, U., Parui, S.K.: CNN based common approach to handwritten character recognition of multiple scripts. In: 2015 13th International Conference on Document Analysis and Recognition (ICDAR), pp. 10211025 (2015)

[5] Niu, X.-X., Suen, C.Y.: A novel hybrid CNNSVM classifier for recognizing handwritten digits. Pattern Recogn. 45(4), 13181325 (2012)

[6] Babu U R, Venkateswarlu Y, Chintha A K. Handwritten digit recognition using K-nearest neighbour classifier[C]//Computing and Communication Technologies (WCCCT), 2014 World Congress on. IEEE, 2014: 60-65.

[7] Tuba E, Bacanin N. An algorithm for handwritten digit recognition using projection histograms and SVM classifier[C]//Telecommunications Forum Telfor (TELFOR), 2015 23rd. IEEE, 2015: 464-467.

[8] Yin Y, Wu J M, Zheng H X. Ncfm: Accurate handwritten digits recognition using Convolutional Neural Networks[C]//Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE, 2016: 525-531.

[9] Islam K T, Mujtaba G, Raj R G, et al. Handwritten digits recognition with artificial neural network[C]//Engineering Technology and Technopreneurship (ICE2T), 2017 International Conference on. IEEE, 2017: 1-4.

[10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11):2278-2324, November 1998.