Detecting and Separating Overlapping Digits

1. Introduction

MNIST is a very popular dataset for the deep learning study. CNN, as one of the most successful neural network method of analyzing visual imagery, can accurately detect handwriting digits with mineral errors. This research is going to detect overlapping digits and separate digits beneath with CNN, and also give overall analysis for each section to evaluate results and present how the neural network developed the outcome step by step.

2. Analytics and Visualization

2.1 The initial neural network model

We used Keras generate the original network model. The heat map below showed the performance of those two digits detection.

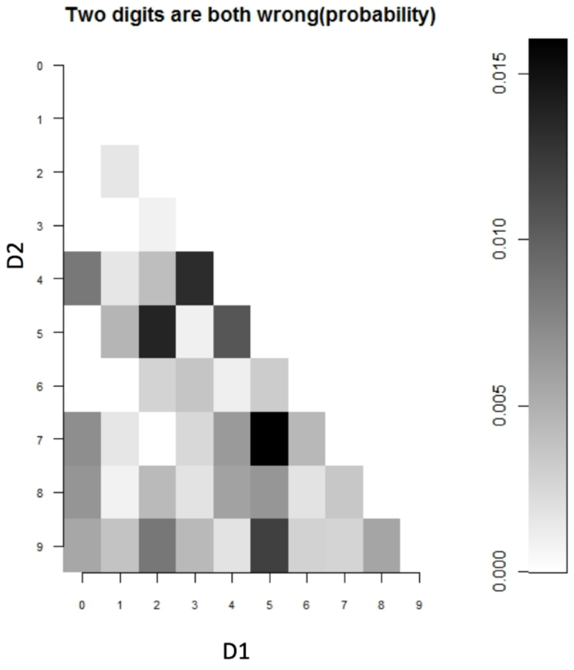

Figure 1.

Figure 1. shows both digits get incorrect detection. The darker the fill color is, the more error those pair digits have. (5,7), (2,5) and (3,4) have the significant results.

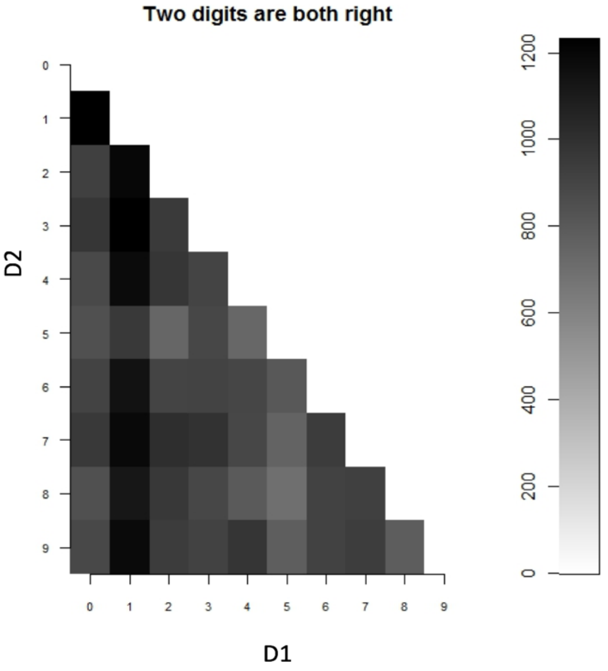

Figure 2.

Figure 2. shows both digits get accurate detection. (0,1) has the most significant result.

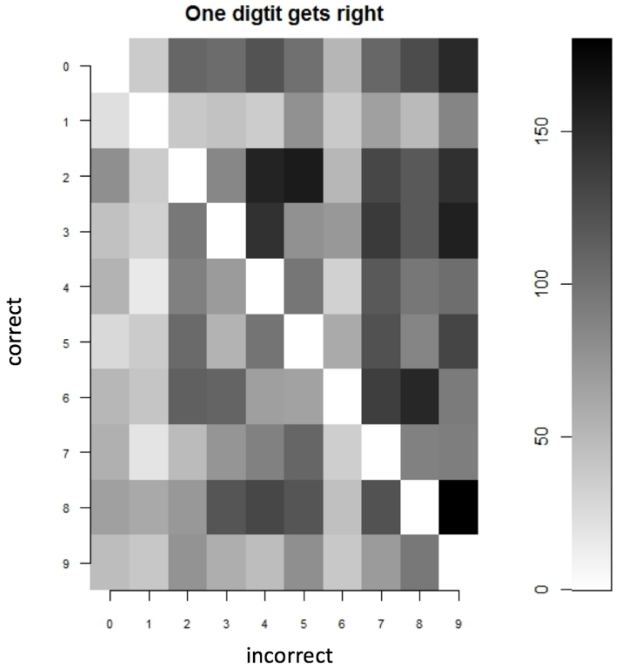

Figure 3.

Figure 3. shows the situation that one of the digits get correct. Two axes represent incorrect and correct results. Each row shows the correct number of single digit in the overlapping digits. Each column shows the incorrect number of those nine digits.

2.2 Generalization

After the rough model we build, we decided to enforce the model by CNN and also excluded (8,9) from the training model to improve the ability of generalization. The Figure 4. gave the scale of the result by sorting the value of ten digits. Since we used sigmoid function as the activation, values were all from 0 to 1. In general, the last two positions will have very high values, and the value of the 8th position will drop down enormously; however, it has too many outliers on the last three positions after excluding 8 & 9 from the training set.

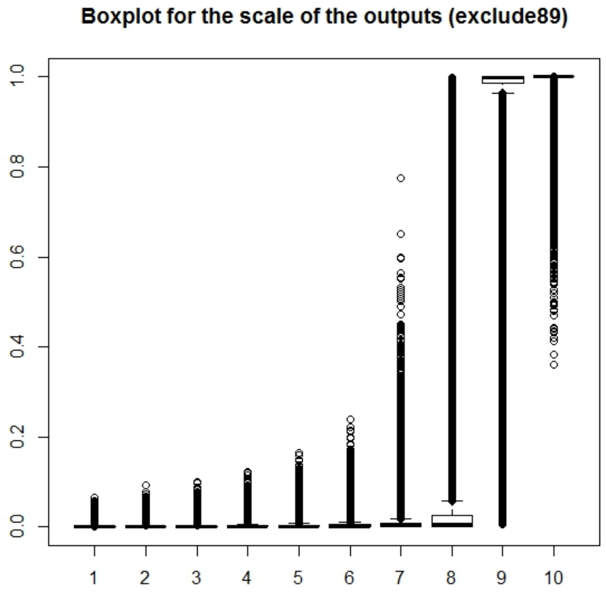

Figure 4.

2.3 Understanding Neural Network

For visually understand Convolutional Neural Network, we tried several methods like PCA and autoencoder to demonstrate hidden layers' results in sequence. The purpose of this visualization is to present how the network converges gradually. Finally, we decided to compute the average distance of same digits pairs, digits pairs that share one digit and entirely different digits across all hidden layers. We extracted input layer, fourth convolutional layer, and last dense layer initially, then used Euclidean distance to compute average distance. The image format will be flattened as a vector (convolutional layer will also be flattened). Each vector is one pair of overlapping digits. The calculation follows this formula below.

$$\mathrm{Avg(D)} = \frac{\sum \binom{n}{2} \mathrm{d(p,q)}}{n^2 - n/2}$$

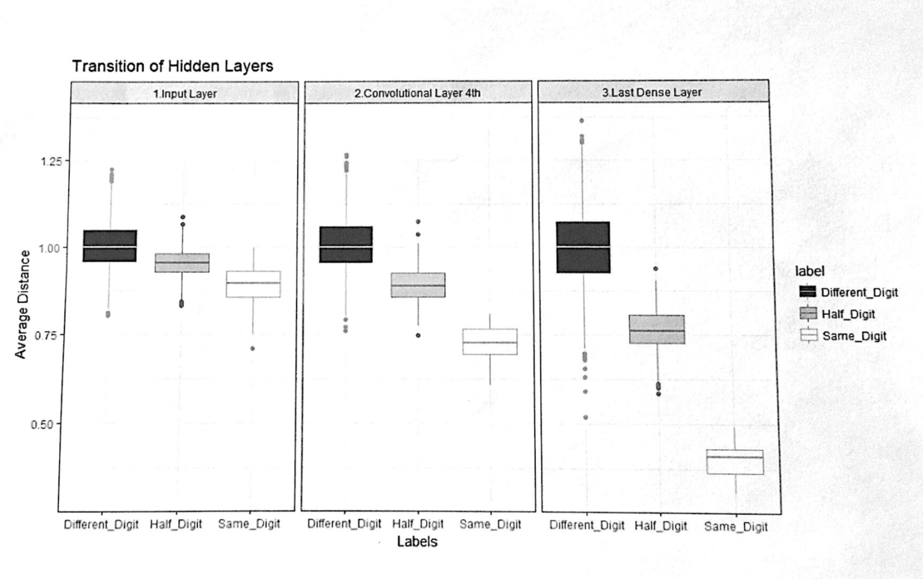

Figure 5.

Figure 5. is the boxplot of three layers, including the input layer, fourth convolutional layer, and last dense layer. All three layers were normalized by the median value of Different pair, which made three layers comparable. The plot clearly shows that the average distance between three conditions is turning to be larger from the input to the last dense layer.

3. Conclusion

All the analytics work derived from digit detection model. A more comprehensive the analysis could facilitate the improvement of the future model. For the digits separation model, though the architecture of hidden layers is more complicated than the detection's, it follows the same pattern as these analysis shows.