**CSE514 Data Mining: Review Essay of Human Emotion Classification**

November 14, 2018

**Team member**

Zhibo Liang: ID: 467574 Email: [zliangaf@gmail.com](mailto:zliangaf@gmail.com)

Xiangzhi Cao: ID: 465424 Email: [xcao22@wustl.edu](mailto:xcao22@wustl.edu)

Zubin Zhang: ID: 466774 Email: [zhangzubin123@outlook.com](mailto:zhangzubin123@outlook.com)

**Problem description**

We decided to build an Emotion Classifier to solve human emotion classification problem. Our classifier should be able to classify multiple human natural emotions using images as input.

**Data collection**

We will use an image database with total 13690 different images. The images are divided into three different sets: training set, validation set and test set. These percentages may change according to our actual training result. Besides, we have a csv file named label. It includes the user.id, the name of images and the kind of emotion in images. Here is the screenshot of our csv file:

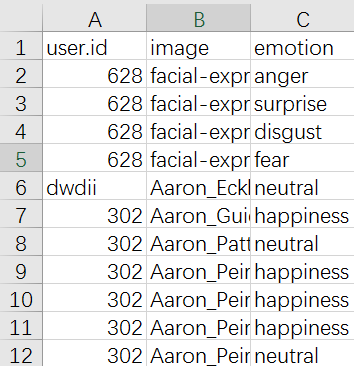


Figure 1: the screenshot of our csv file

We divide 7 kinds of emotion, which is surprise, happiness (positive emotion), neutral (neutral emotion), and anger, disgust, fear, sadness (negative emotion). After the training, we use validation set to optimize the parameter of CNN. After optimization, we use both training data and validation data to train the final model. Finally, we will use test data to test our final model. The evaluation criteria is accuracy. We will calculate the overall accuracy, the recall as well as the false positive rate of different labels.

**The program we have written**

1. **Data Processing**

First, we set the number of training, validation and test data. Then, in order to prepare the training, validation and test data, we construct the database for model training. We resize all the detected faces into 350 X 350, load images and convert it to RGB. After reading image name and label from the label.csv, we do the randomized perturbation in order to better overall accuracy.

The reason we need to perturbate is below:

As we checked the emotion label in label.csv. I have found the number of each emotion is that:

|  |  |
| --- | --- |
| emotion | number |
| surprise | 368 |
| happiness | 5696 |
| neutral | 6868 |
| anger | 257 |
| disgust | 208 |
| fear | 21 |
| sadness | 309 |

Form 1 The number of different emotions in label.csv

You can see that “happiness” and “neutral” emotion is significantly more than other emotions. Therefore, the database in unbalanced, which does bad to overall accuracy. Our object is to make the number of seven emotion images are all the same. So, we need perturbate to enlarge the number of other five emotions. The principle is that:

1. Get the number of images with each specific emotion.
2. Compare and get the emotion label which has the greatest number of images.
3. Enlarge the images of other six emotion labels. The method we use is rotation. We set a random number from -45 to +45, and let each image belonging to other six emotion labels to rotate a random degree until the number of seven emotion images are the same.

After we rotate, we also use Horizontal mirror flip and GaussianBlur in order to increase the overall accuracy. We set a random number from 0 to 1. If the number is larger than 0.5, the image will have its horizontal mirror flip partner or GaussianBlur partner.

1. **The module we use**

The classification module we use is the ensemble of multiple deep convolutional neural networks (CNN). Here is the screenshot of our CNN model.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d (Conv2D) (None, 348, 348, 32) 320

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d (MaxPooling2D) (None, 174, 174, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_1 (Conv2D) (None, 172, 172, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 86, 86, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 84, 84, 128) 73856

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 42, 42, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 40, 40, 128) 147584

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_3 (MaxPooling2 (None, 20, 20, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 18, 18, 256) 295168

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_4 (MaxPooling2 (None, 9, 9, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_5 (Conv2D) (None, 7, 7, 256) 590080

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_5 (MaxPooling2 (None, 3, 3, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten (Flatten) (None, 2304) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense (Dense) (None, 512) 1180160

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 1) 513

=================================================================

Form 2 The summary of CNN model

You can see we have five convolutional layers, five max pooling layers and two fully connected layers. Convolutional layers are used to feature extraction. The max pooling 2D layers is to apply a maximum pooling to the spatial domain signal and it can gradually reduce the pixel value of images. By the way, the fully connected layers contain dropout, which is another method of randomization. And it can make the classification come true.

**Initial test results**

Here is the screenshot of our test results. (training number=1000, validation number=100, test number=60)

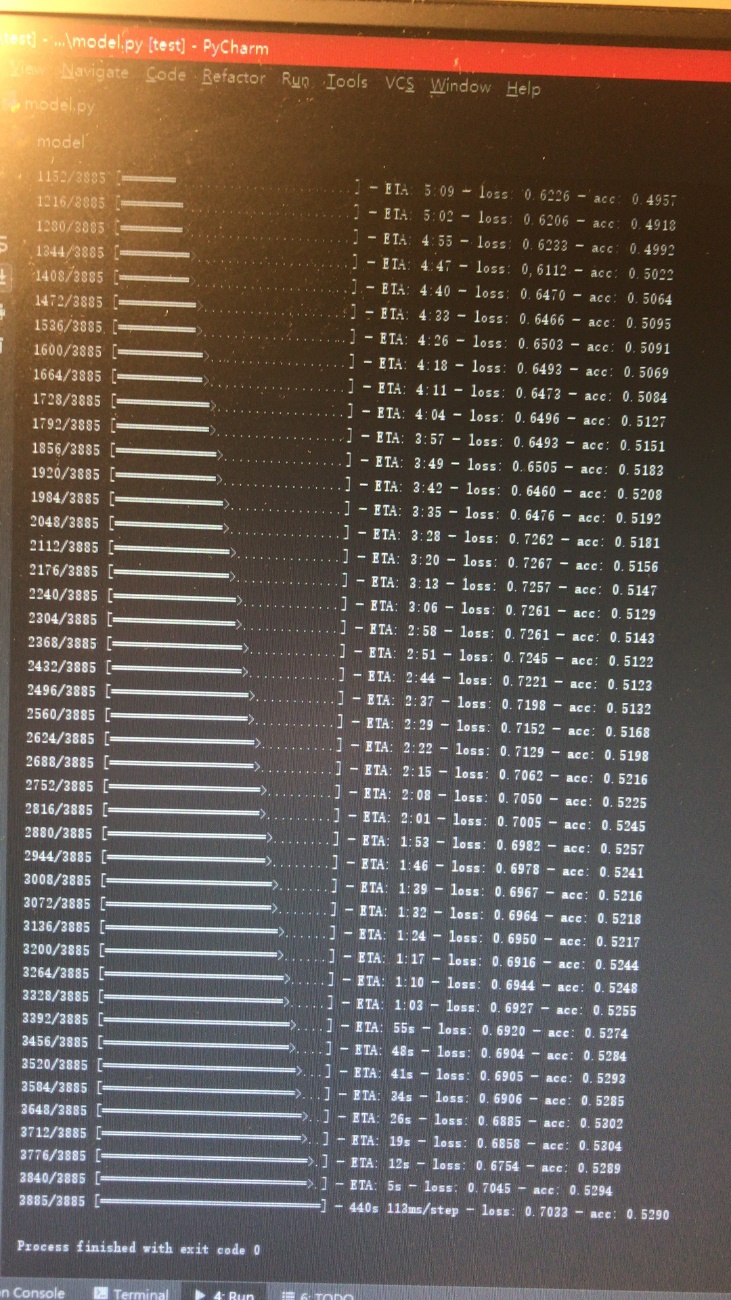
****

Figure 2 The Screenshot of initial test results

You can see that the total number of images in training set is 3885, which means that after perturbation the number of images in training set has increased from 1000 to 3885=555 X 7. Therefore, the number of emotion label which has the greatest number of images is 555. Before perturbation, the overall accuracy is nearly 0.4. And after perturbation, the overall accuracy is nearly 0.5. That means the method of perturbation is successfully increase the overall accuracy.

**The remaining tasks**

Until now, we have done the basic work for building the CNN. In the future, we will mainly focus on these four tasks:

(1) Optimize the parameter of the CNN and the pre-process of input image data

(2) Train multiple CNNs (about 5 - 10) and use the method similar to linear regression to combine these CNNs into a whole classifier

(3) Work on auto-encoder and use it on the building of CNN

(4) Try different kinds of combination (For example, using or not using auto-encoder) to evaluate their performance and find out the best structure for human emotion classification. We will also perform some data visualization.

Since there is a long holiday next week, we believe we can spend more time on the project. We believe we can finish the project on time.