Prasanna Kumar Srinivasachar

Sham Prasad PS

Vibhor Mishra

**Project Final reports**

**Introduction:**

We worked on two projects: Gender prediction based on handwriting and Emotion detection.

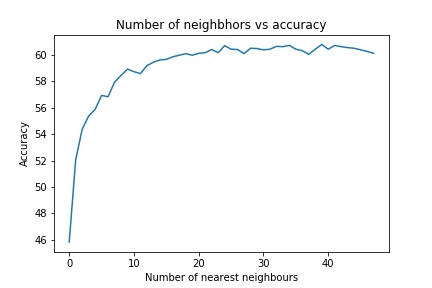
For Emotion Detection, we are trying to differentiate 6 distinct types of emotions: Anger, Disgust, Fear, Happy, Sad, Surprise and Neutral.

**Approach:**

* Gender prediction based on handwriting
  + Dataset : Set of images having handwritten texts and a csv file that has writer id and gender(labels) which need to be predicted. (<https://www.kaggle.com/c/icdar2013-gender-prediction-from-handwriting/data>)
  + **Techniques explored**: For the baseline, we have used KNN for the classification task. We also tried other algorithms like SVM with a linear kernel (LinearSVC)) and SVC with rbf kernel. We used 200 training samples and 50 test samples. The accuracy reported are as follows :

|  |  |  |  |
| --- | --- | --- | --- |
|  | KNN (neighbors = 10) | LinearSVC | SVC(kernel=rbf,gamma=1.5,C=2) |
| Training data | 0.74 | 0.73125 | 0.7600 |
| Test data | 0.615 | 0.6600 | 0.6700 |

* Emotion detection: We try to predict the emotion expressed by a person in the given image.
  + Dataset: csv files containing emotions column along with pixel values of facial images (labels) . (<https://inclass.kaggle.com/c/facial-keypoints-detector/data>)
  + We segmented facial parts using facial landmark detector, an implementation of the “[One Millisecond Face Alignment with an Ensemble of Regression Trees](https://pdfs.semanticscholar.org/d78b/6a5b0dcaa81b1faea5fb0000045a62513567.pdf)”, included in the dlib library. As features, we have used width and length of these facial parts, i.e. both eyes, both eyebrows and lips, as of now which is doing pretty well and giving us a decent accuracy.
  + After performing gridSearch with a cross validation of 5. We got 41 neighbours as the best param and attained an accuracy of 60.531. The following graph shows mean of the accuracies for this 5-fold cross validation versus the number of nearest neighbours.



Initially for KNN and SVM, we had to extract features from our training images for which we used the following approach:

1. detecting and segmenting facial parts using facial landmark detector, an implementation of the “One Millisecond Face Alignment with an Ensemble of Regression Trees”, included in the dlib library.
2. using width and length of each of the facial parts, i.e. both eyes, both eyebrows and mouth, as features. This is based on the intuition that for every facial expression there is change in dimensions of each facial parts, such as in laugh, the lips extend on the both sides increasing the width and height of the mouth overall and eyes usually shrink thus decreasing their dimensions.

We fed these features to SVM and KNN and finely tuned the parameters for best accuracy.

Later to explore further techniques, we trained Convolutional Neural Network on our dataset of Emotion detection. The facial images in our dataset consists of cropped face images of 48x48 dimensions. So, we don’t require to pre-process our images before we feed them to our CNN.

**Analysis:**

Our testing approach for this project is inspired by “Facial Emotion Recognition in Real Time” paper.