MIDS 207 Final Project

Anup Jha
Anish Phillip

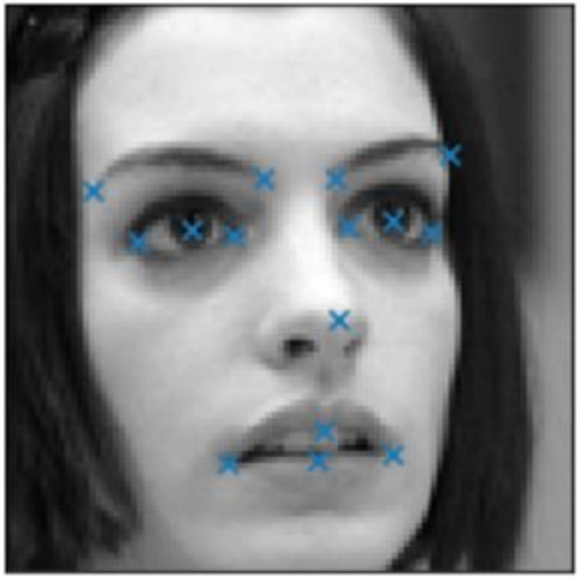
Description

The objective of this project is to identify the location of facial keypoints on the images of face in grayscale

- There are 15 key points on the face such as :
- Left_eye_center right_eye_cente
- Left_eye_inner_corner left_eye_outer_corner
- Right_eye_inner_corner right_eye_outer_corner
- Left_eyebrow_inner_end left_eyebrow_outer_end_
- Right_eyebrow_inner_end right_eyebrow_outer_end
- Nose_tip mouth_left_corner
- Mouth_right_corner mouth_center_top_lip
- Mouth_center_bottom_lip

- Each of the feature has X and Y co-ordinates
- The image itself is 96 X 96 grayscale image

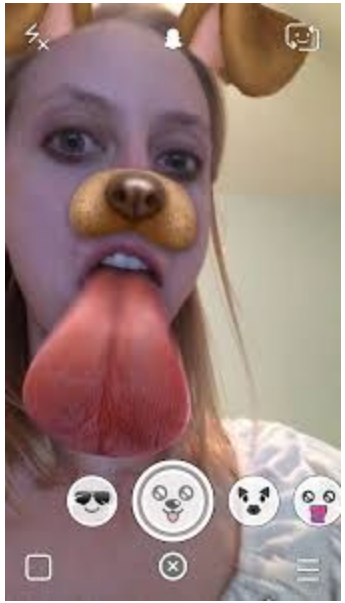
Description



Applications

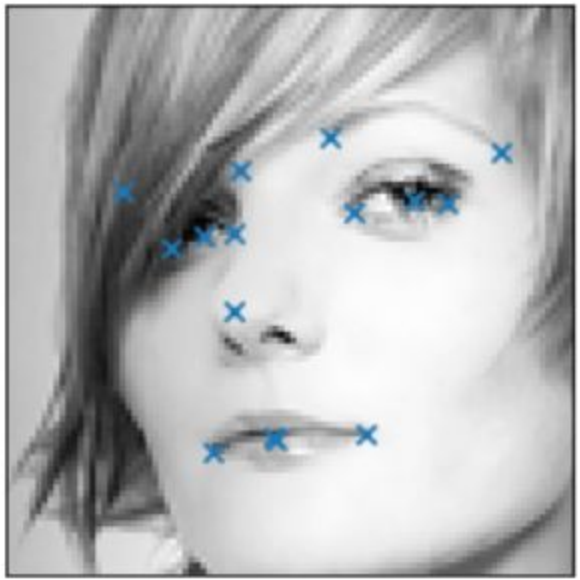
This can be used as a building block in several applications, such as:

- tracking faces in images and video
- analysing facial expressions
- detecting dysmorphic facial signs for medical diagnosis
- biometrics / face recognition



Features Available

- The value of 9216 (96 X 96) pixels of the image are available for each training example
- The values of pixels vary from 0 to 255



Feature Engineering

We intend to do the following feature engineering to make the predictions accurate

- For Random forest and SVM regression we will include gaussian blurring
- Normalization of the pixel values to be scaled between 0 and 1
- Normalization of the locations of the keypoints would be done between -1 and 1
- Since we plan to use the neural nets for the predictions we don't focus much on feature engineering for neural nets only normalization of features and predictions which will make the gradient decent faster .
- We will also have dropouts of the neurons

Expected Outcome

- We plan to create a model which can predict the location of the key features in terms of pixel position

- The model will be fed a gray scale image of face of size 96X96

Algorithms

We plan to use :

- Random Forest regression
- SVM regression
- Fully connected neural network
- Convolutional neural network

Baseline

- For Baseline performance we will take the mean value of location of each keypoint from the training set as prediction and check the RMSE of the Dev Set

We Got the following :

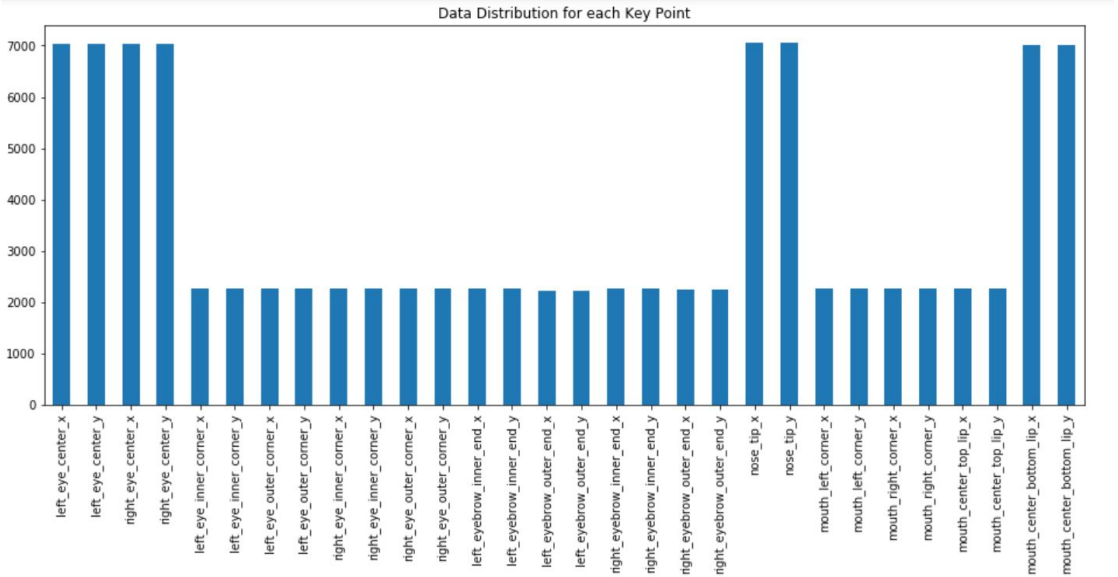
Baseline RMSE = 0.093240 in scaled prediction space

We have normalized the predicted value between -1,1 so original baseline RMSE is 48 X RMSE = 4.475508 in 96 X 96 space

Data Set

- For the training set we have 7049 example rows but only 2140 examples have all the facial key points
- We will split the training set into training set and validation set
- We have test set set with 1783 images of 96X96 pixel images for which we need to submit our predictions for the competition
- The training set has 7049 X 31 shape
- The last column has space

  separated values of pixels



Hyper Parameters

- For Random forest we will tune the number of estimators,max depth
- For SVM regressor we will tune the regularization penalty and the sigma and gamma of the rbf kernel
- For the neural nets we will tune the number of neurons in each layer and droputs and the learning rate

Dimension reduction

- For Random Forest and SVM regression we will consider PCA dimension reduction

Loss Function

- We would be using the RMSE error as output is a location of the keypoint and RMSE would be appropriate as it will give distance of the error

Data Augmentation

- Since the training set would be low in number we would also flip the images to double training data