MIDS 207 Final Project - Team 4

Anup Jha
Anish Philip

Problem statement and scope

Baseline case

Next Steps

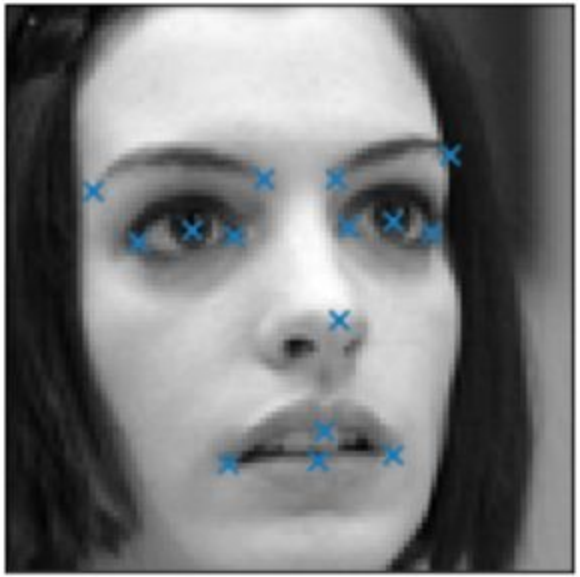
Description

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

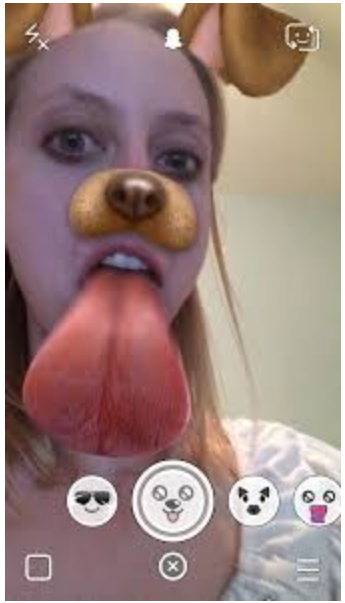
Description



Applications

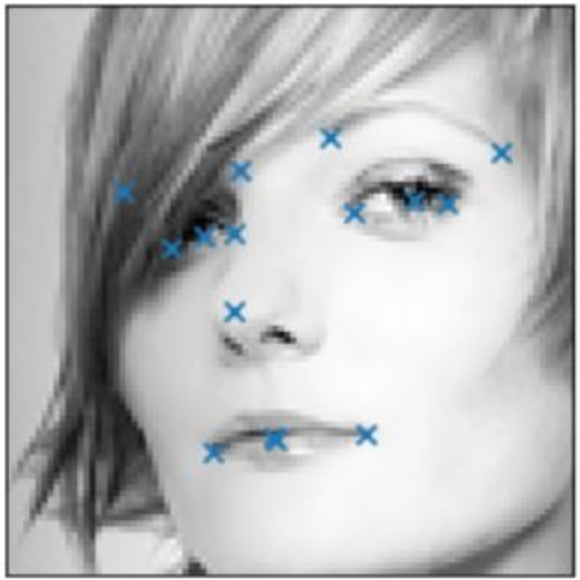
This can be used as a building block in several applications, such as:

- 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

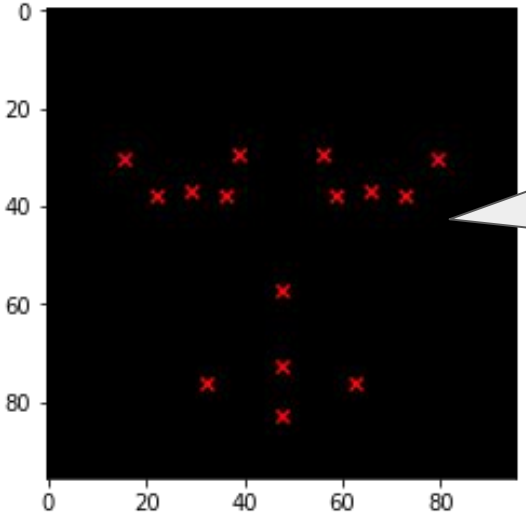
Project statement and scope

Baseline case

Next Steps

Baseline will provide meaningful reference point to which we can then compare our model(s)

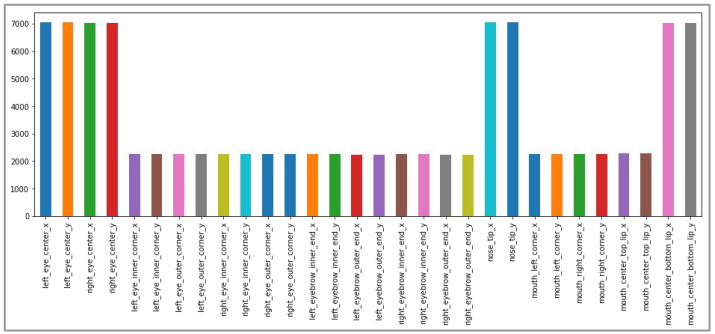
- For Baseline performance we took the mean value of location of each keypoint from the training set as prediction and compared it with the dev data set to arrive at baseline RMSE for our project



Average of all keypoint coordinates is overlayed over test data to get Baseline RMSE

There are 7,049 images with keypoint coordinates in the training data and 1,783 images in test data

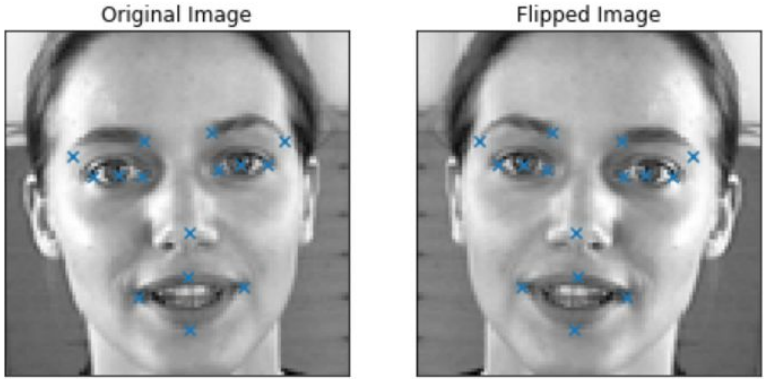
- There are 15 facial keypoint features with (x,y) coordinates



- Only **2,140 images** in training data have all 15 features
  a. Eyes (Center, Inner and Outer Corners) - 12 features
  b. Eyebrows (Corners) - 8 features
  c. Nose (Nose tip) - 2 features
  d. Lips (Corners, top, bottom) - 8 features

We decided to use the 2,140 images from training data which had all 15 keypoint features available to build our baseline case

- We flipped these 2,140 images to double the size of data we would use to build our baseline assumption and created a shuffled set of:
  - 3424 Training data with all 30 features
  - 856 Test data with all 30 features



Baseline RMSE for individual facial keypoint features indicate that we have the work cut out for us with Nose, Mouth and Eyebrow detection

Features	RMSE
Left eye Center	2.52
Left eye inner Corner	2.30
Left eye outer Corner	3.19
Left eyebrow inner end	3.28
Left eyebrow outer end	**4.06**

Features	RMSE
Right eye Center	2.58
Right eye inner Corner	2.39
Right eye outer Corner	3.19
Right eyebrow inner corner	3.39
Right eyebrow outer corner	**4.20**

Features	RMSE
Nose tip	**4.50**
Mouth left corner	**4.73**
Mouth right corner	**4.78**
Mouth center top lip	**4.70**
Mouth center bottom lip	**4.74**

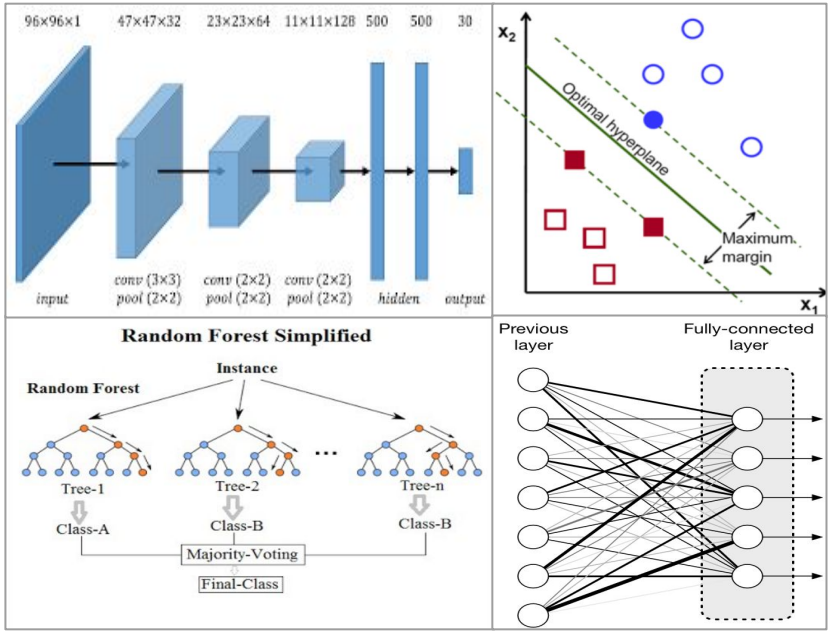
Project statement and scope

Baseline case

Next Steps

We plan to build multiple prediction models based on the following algorithms to achieve an RMSE lower than **4.47**

- Neural network
  a. Convolutional
  b. Fully Connected (If feasible)
- SVM regression
- Random Forest regression



Additional Slides

Hyper Parameters

- For the **neural nets** we will tune the number of layers, neurons in each layer, activation function, optimizer, Filter size and the learning rate

- 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

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