

Midterm Project Proposal

Facial Keypoints Detection

Objective:

The objective is to train a model that can predict the keypoint positions on face images.

Precisely, 15 keypoints needs to be predicted:

left_eye_center, right_eye_center, 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

Motivation:

1. It is a big dataset with multiple features, so a deep CNN is needed.
2. It has practical application because the built model can be used for:
 - tracking faces in images and video
 - analyzing facial expressions
 - detecting dysmorphic facial signs for medical diagnosis
 - biometrics / face recognition

Dataset Description:

- Total Images: 8832 (7049 training, 1783 testing)
- Each predicted keypoint is specified by (x, y) pair.
- Images are 96x96 pixels.

Approach:

- Build a model using statistical machine learning methods like SVM to set a benchmark performance.
- Then, approach the problem by implementing a Deep Convolutional Neural Network.
- Implement ResNets[1] and DenseNet[2] neural networks and compare the performance.
- Build an end to end pipeline so that whenever a new image is inputted, the process resizes the image and predicts the keypoints on the face.
- Leverage the power of GPUs using AWS.

References:

- [1] Deep Residual Learning for Image Recognition. <https://arxiv.org/pdf/1512.03385.pdf>
- [2] Densely Connected Convolutional Networks <https://arxiv.org/pdf/1608.06993.pdf>