Viola and Jones

Annotate bounding boxes myself for a represented dataset

Use the algorithm to generate the bounding boxes

Annotate bounding boxes myself for an underrepresented dataset

Use the algorithm to generate the bounding boxes

Calculate the average difference between my annotations and the algorithms on both datasets, and use this data to compare the accuracy..

Could extend this approach for the eye or face classifier

MTCNN and Retina Net

Annotate bounding boxes myself for a represented dataset

Use the algorithm to generate the bounding boxes and landmark locations

Annotate bounding boxes myself for an underrepresented dataset

Use the algorithm to generate the bounding boxes and landmark locations

Calculate the difference between the landmarks and bounding boxes the algorithm made and my own

Test these on photos that were ideally taken from a controlled dataset. Maybe from the same dataset

68 Points

Annotate the bounding boxes myself for specific points

Annotate the two outer eye points

Annotate the two mouth points

Annotate the nose point

Run the algorithm and select the 7 points from the 68. Compare against my calculated ground truth and see the difference between the two groups.

Run the algorithm for both datasets of represented and underrepresented people.

Calculate the average distance between the points for each dataset. Look for this difference to see if there are any points that cause a skew and which dataset this appears in

Datasets

Viola and Jones, differences in brightness values can cause the detector not to classify a region as a face, eye or mouth as on darker toned individuals.

DLIB, differences in gradient can effect what is classified as a face but in general it is only a good frontal face detector and will have real trouble trying to detect faces at an angle. The CNN detector offered by DLIB can detect faces at angles but requires much more computational power than the HOG and SVM method.

MTCNN uses a shallow network to detect faces instead of relying on other packages or methods to do so. Its meant to learn unbiased feature representations. It is stated in the paper that the performance might be limited because the filters in the convolution layers lack diversity and that may limit their discriminative ability. Also, the false classification of the ground truth which will affect the results of the bounding box and landmark locations.

Face net

Look into the datasets themselves, can’t really experiment but comb through the photos and get research in what the datasets are formed of.

Benchmark the different algorithms against a biased dataset like the 300W dataset.

100 photos – Comprised of widely available celebrity photos

50 Male

0-24 – Minorities

Darker skinned tone individuals

Celebrities

Asian individuals

Celebrities

25-49-Represented

Caucasian tone individuals

Celebrities

50 Female

0-24 – Minorities

Darker skinned tone individuals

Celebrities

Asian individuals

Celebrities

25-49-Represented

Caucasian tone individuals

Celebrities

Photos vary in illumination

Don’t vary in pose

Vary roughly in facial expression

Vary in ageing

Don’t vary in resolution

Filename: id\_gender\_minorites.jpeg