CPS803 - Final Project

Face detection and blurring through Machine Learning

Nikolas Trivanovic

Craig Pinto

Ahmet Cengiz

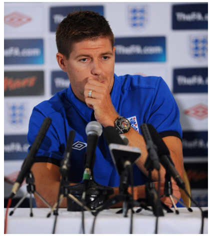
Mohammed Qasim Khan

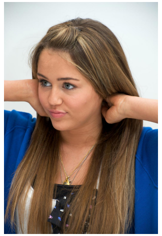
INTRODUCTION

The prevalence of public surveillance raises concerns for the privacy of citizens and instances of potential noncompliance with anonymity laws for corporations.

Our goal is to develop deep-learning algorithms that we learn through the CPS 803 course that can detect a face and subsequently apply a dynamic face-blurring filter. Our algorithms should be able to do this for both still and live images.

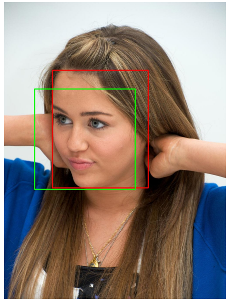
PROJECT OVERVIEW





1) **Dataset**: Wider Faces dataset
   a) From the University of Hong Kong

2) **Models Used**
   a) Haar Cascade
   b) HOG SVM
   c) CNN model (inc.)
   d) DNN w/res10
   e) DNN w/caffe



3) **Blurring Algorithm (Gaussian)**

DATA PREPROCESSING

- Normalizing the image to the CV_8U data type (unsigned 8 bit integer), ensuring colour values are between 0 to 255.
- Ingesting the data set’s ground truth bounding boxes from a text file
- Haar-Cascades: Required converting the images to grayscale.
- DNN: Required resizing images to 300x300 pixels.
- CNN: Required resizing images to a constant dimension of [218, 178, 3] and normalizing by dividing them by 255.

FEATURES SELECTION

- The **Haar Cascades model** achieves feature selection via a modified AdaBoost of weak learners on a single feature
- Our **HOG SVM model** made use of HOG gradients, where cells of pixels in images where used to compute gradients that were used as a precursor to facial landmark location.

- Our **DNN** and **CNN models** made use of facial landmark locations as input for its features

DATASET SPLIT

- The dataset used for the pre-trained models did not require splitting into training and validation sets as the models were ready to be tested.
- As a result the testing dataset was split into two categories:
  - Images containing a single visible face
  - Images containing no visible faces
- This allowed us to test for True Positives, True Negatives, False Positives and False Negatives.

PERFORMANCE METRIC

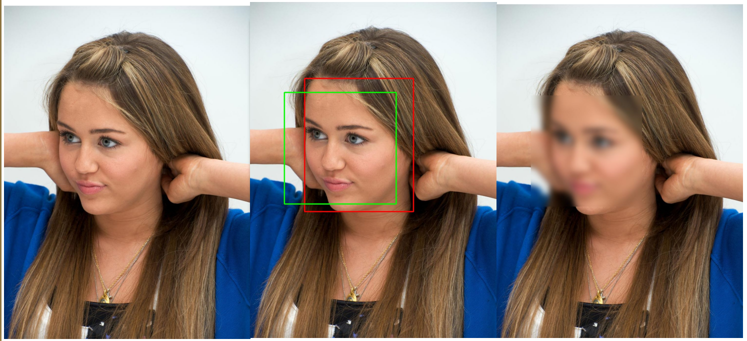
- Mean Average Precision: The Intersection over Union of the bounding box results. IoU threshold of 0.5 to determine whether result is a False Positive or True Positive.
- Ground Truth Accuracy: The area of coverage of the detected bounding box in comparison to the ground truth, given as a percentage.

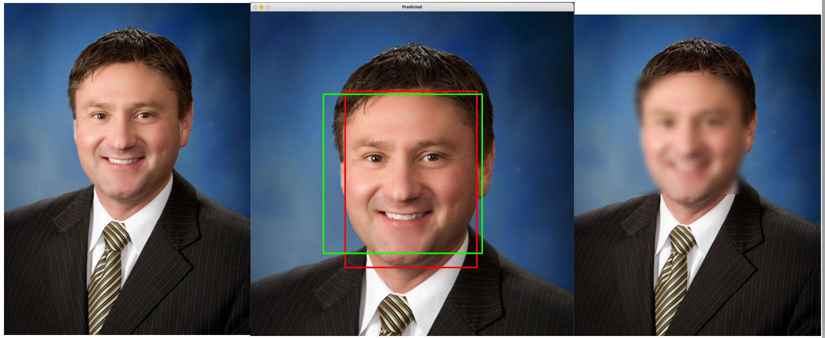
POST PROCESSING - GAUSSIAN BLUR

- Gaussian blur with a pixel mask of 75x75 was applied to the regions of the image within the detected bounding box.
- Works by sampling all pixels within a +-75/2 pixel range in the x and y direction of the center pixel and averaging the intensity.
- Works well for lower resolution images but does not blur adequately for higher resolution images as pixel mask is too small.
- Future work includes an adjustable pixel mask that is proportional to the image resolution.

Haar Cascades

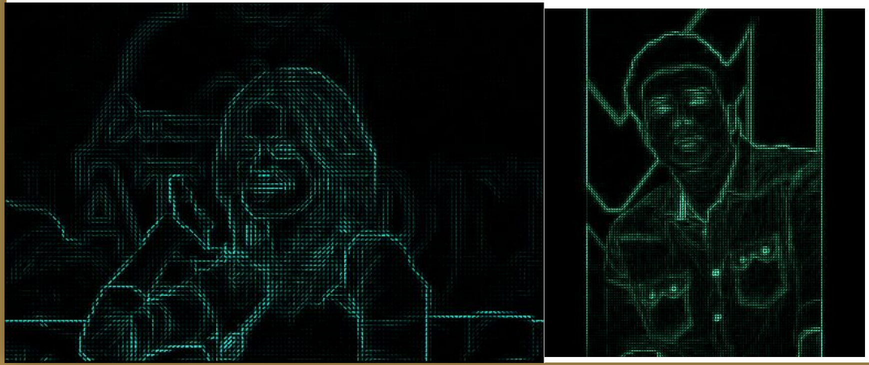
- Object detection through the use of Integral Image.
- Computationally cheap and fast in comparison with other object detection models.
- Pixel Values are determined via the sum of previous pixels to left and above.
- Prone to False Positives at a rate much higher than other models.

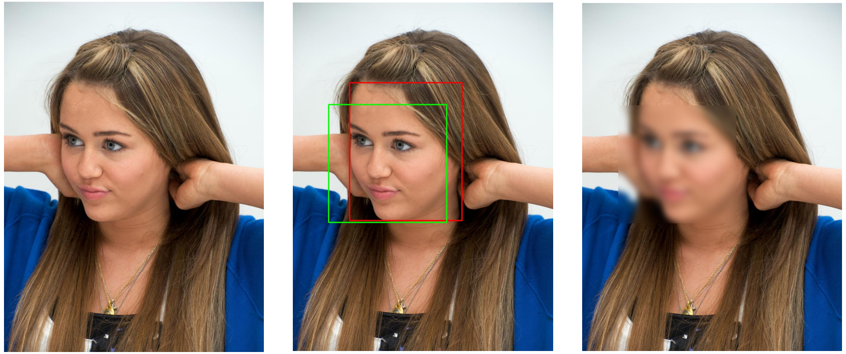




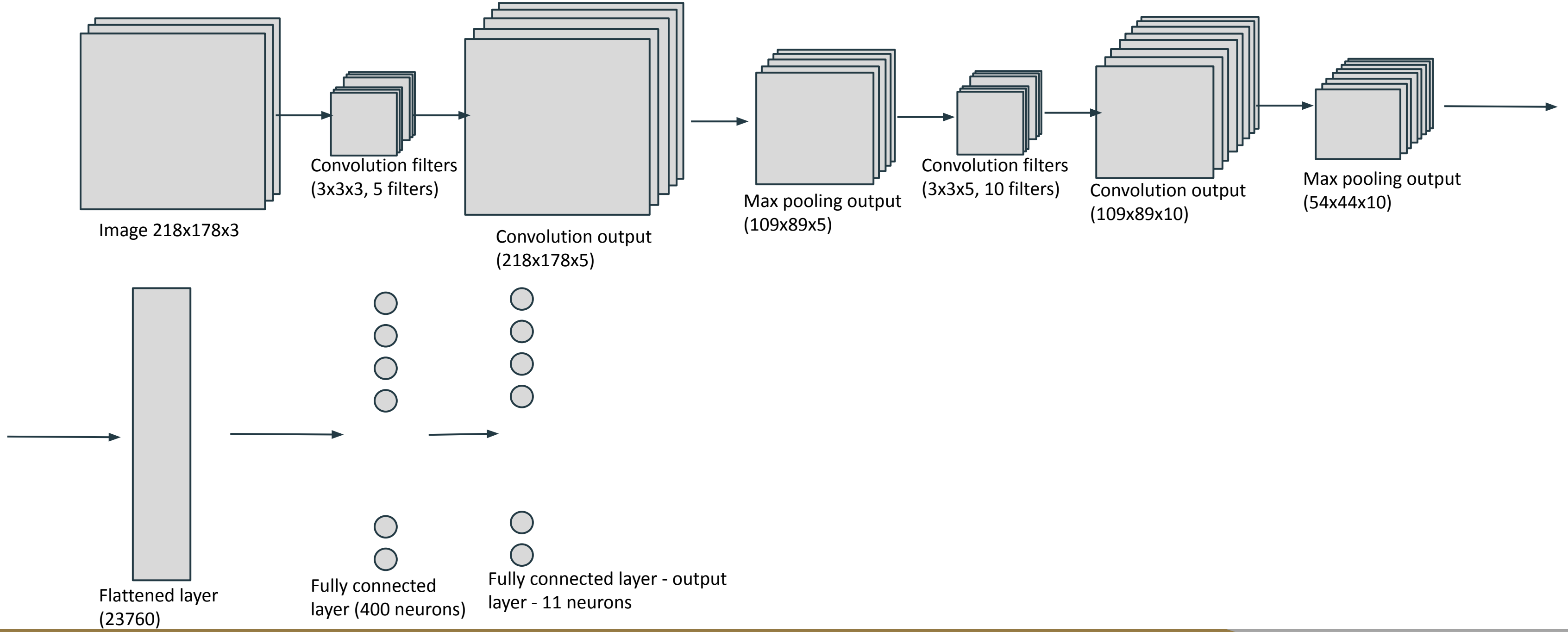
HOG-SVM

- HOG SVM: Uses Histogram oriented gradients divided into cells and bins to determine the shape as well as the edges (direction preserving) in the image.
- Gradients are grouped into bins, with larger gradients carrying more weight in their respective bin.





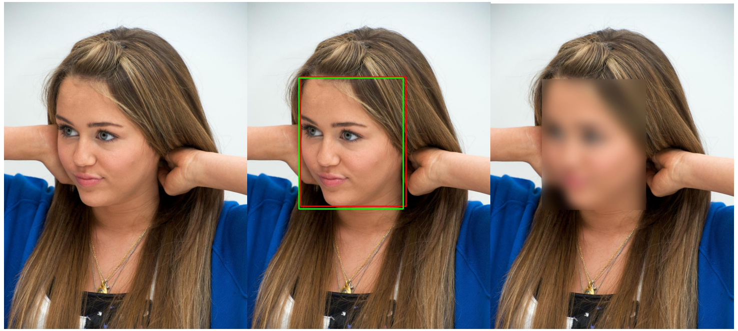
Convolutional Neural Network (CNN)



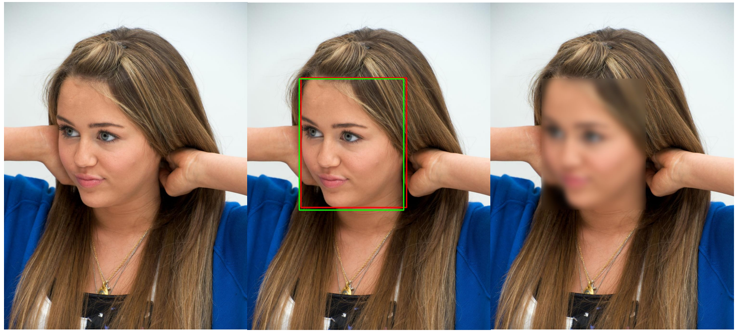
Deep Neural Network

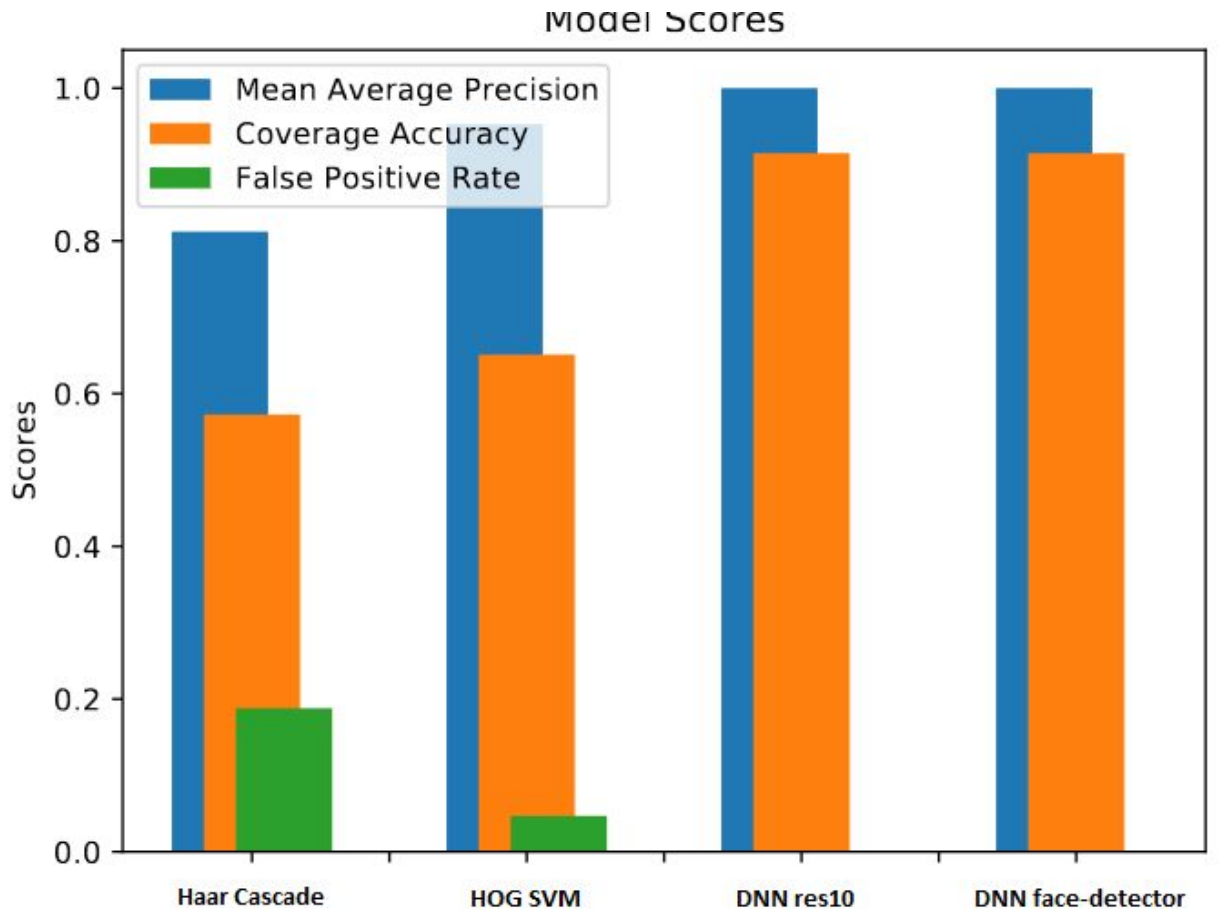
- OpenCV’s Res10 and Face-Detector Model, both models are based off of the DNN structure but are trained separately.
- The inclusion of two separately trained models of the same type was to test for consistency in results as both DNN models boasted very high precision.

res10



face-detector





Model Scores
Mean Average Precision
Coverage Accuracy
False Positive Rate
Scores
1.0
0.8
0.6
0.4
0.2
0.0
Haar Cascade
HOG SVM
DNN res10
DNN face-detector