CS6501: Deep Learning for Visual Recognition
Recognizing People in Images
Today’s Class

- Face Detection
- Face Matching - and any type of matching
- Pose estimation
Face Detection
Face Detection: Viola-Jones Face Detector circa 2001

1. Compute these types of features across the image

2. Use a shallow classifier – e.g. ADA Boost

3. Non-Max Suppression
Face Detection: Any Object Detector

https://towardsdatascience.com/faced-cpu-real-time-face-detection-using-deep-learning-1488681c1602
Face Detection can be Hard

<table>
<thead>
<tr>
<th>Scale</th>
<th>Pose</th>
<th>Occlusion</th>
<th>Expression</th>
<th>Makeup</th>
<th>Illumination</th>
</tr>
</thead>
</table>

[WIDER FACE dataset](https://www.wIDERface.org).
Person Identification: Simplest Case

Classify Among k-people in your database
Face Matching and just Matching Things

Are these pairs of images, instances of the same?
Matching Things: Siamese Networks

Find a neural network such that if two instances of the same thing are fed into the network, the outputs are similar under some simple distance metric.

Also called the embedding problem

Learning a Similarity Metric Discriminatively, with Application to FaceVerification
Chopra, Hadsell, and LeCun.
Matching Things: Siamese Networks

$x_1$  $\rightarrow$  $f(x_1)$

$x_2$  $\rightarrow$  $f(x_2)$

FaceNet: A Unified Embedding for Face Recognition and Clustering

Matching Things: Siamese Networks

If $x_1$ and $x_2$ are the same person then minimize:

$$|f(x_1) - f(x_2)|$$

FaceNet: A Unified Embedding for Face Recognition and Clustering
Matching Things: Siamese Networks

if $x_1$ and $x_2$ are the same person then minimize:

$$|f(x_1) - f(x_2)|$$

Beware of Trivial Solutions!

FaceNet: A Unified Embedding for Face Recognition and Clustering
Matching Things: Siamese Networks

if $x_1$ and $x_3$ are not the same person then minimize:

$$-|f(x_1) - f(x_3)|$$

FaceNet: A Unified Embedding for Face Recognition and Clustering
Better Idea: Triplet Loss. e.g. FaceNet

Minimize the following loss for every possible triplets

$$\sum(|f(x_1) - f(x_2)| - |f(x_1) - f(x_3)| + \alpha)$$

FaceNet: A Unified Embedding for Face Recognition and Clustering
Better Idea: Select Triplets that are Hard

Minimize the following loss for every possible triplets

$$\sum \left( |f(x_1) - f(x_2)| - |f(x_1) - f(x_3)| + \alpha \right)$$

FaceNet: A Unified Embedding for Face Recognition and Clustering
Pose Estimation

Deep Pose

[Diagram of Deep Pose model with dimensions and layers]

Deep Pose

Results
Pose Model II: HourGlass Network

Hourglass Module
Pose Model II: HourGlass Network
Pose Model II: HourGlass Network

Hourglass Network
Pose Model II: HourGlass Network
Dense Pose

http://densepose.org/

DensePose-RCNN Results

DensePose COCO Dataset
Dense Pose

http://densepose.org/

DensePose COCO Dataset
Dense Pose

http://densepose.org/
Dense Pose  http://densepose.org/
Questions?