Image Processing

Facial Attribute Recognition

Guided by: Mr. Ly Quoc Ngoc

Group: VPN

Luu Thi Hong Ngoc	**19127051**
Ho Thien Phuoc	19127517

Motivation

Challenges

Problem statement

Performance table

Principle – Method – Algorithm

Installation

Motivation

- **Scientific significance**
  - We propose to take advantage of attribute relationships in three ways: by using MCNN sharing the lowest layers amongst 40 attributes(groups) before splitting the network into 6 branches, each one of which focuses on a certain attribute group; sharing the higher layers for related attributes(different attributes belonging to the same group); and by building an auxiliary network(AUX) on top of the MCNN which utilizes the scores from all attributes to improve the final classification of each attribute.
- **Application**
  - HCI required information about gender require information about gender.
  - Expression in order to determine the mood of the user.
  - Identity verification in low quality imagery to automatically search for suspects in surveillance video.

Problem statement - Input, output

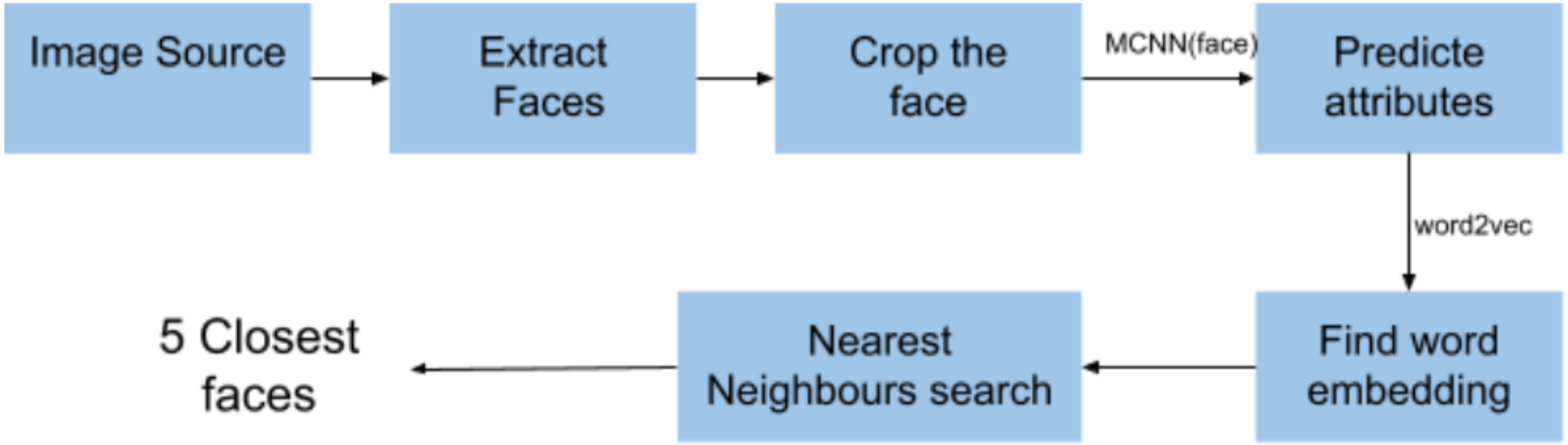
Input: a face image (an attribute vector)

Output: set of face images for each attribute vectors which represent each images, are similar to attribute vector of input.



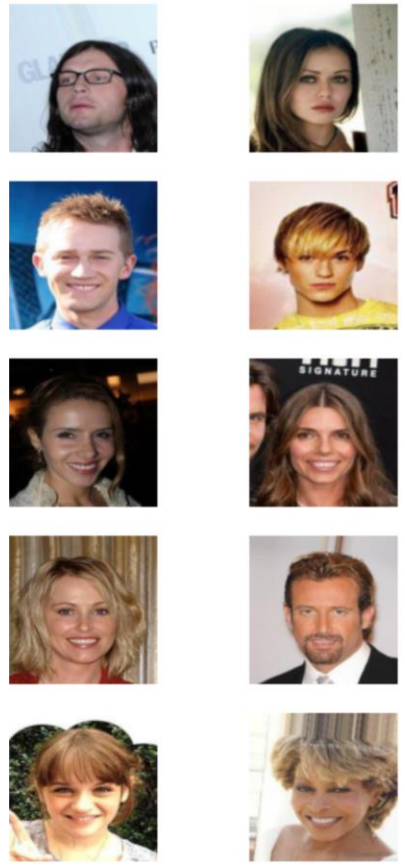



Problem statement - Framework

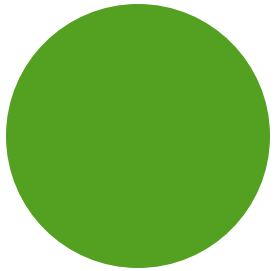


Problem statement - Standard Dataset

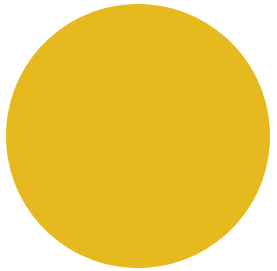
- CelebA
  - CelebA dataset consists of 200k images respectively with 160k, 20k, and 20k images for training, validation, and testing sets. CelebA dataset provide the same 40 binary attributes.



Problem statement - Challenges



Datasets are extremely challenging, with large variations in subject pose, illumination and image quality.



Improving the accuracy of attribute classifiers

Performance table

(Document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qlx6s3SElQan3SHJhnY/edit?usp=sharing)

MCNN – Principle

The input image is cropped to 227x227 and the training mean is subtracted. The image is then passed through the convolution layers and the fully connected layers to produce attribute scores. The attribute scores are then thresholded to give a yes or no answer. The red attributes indicate a lack of the attribute and the green attributes indicate a positive instance.

MCNN and AUX - Method

There are two parts to this solution

1. Face Attributes: Design a CNN that gives the facial attributes, given a face image
2. Face Similarity: Represent the face attributes in a vector space and use this to find similar faces

MCNN and AUX - Method

M (Multi-task learning)

- Used to facial landmark localization, pose estimation, action recognition, face detection, and many more
- 40 attributes share the lower layers in the CNN, so that information common to 40 the attributes can be learned. Applying M to attribute prediction is very natural given the strong relationships among the facial attributes.

(See more detail in document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qlx6s3SElQan3SHJhnY/edit?usp=sharing)

MCNN and AUX - Method

CNN

utilized a large dataset and applied both a siamese deep CNN and a classification CNN in order to maximize the distance between impostors and minimize the distance between true matches.

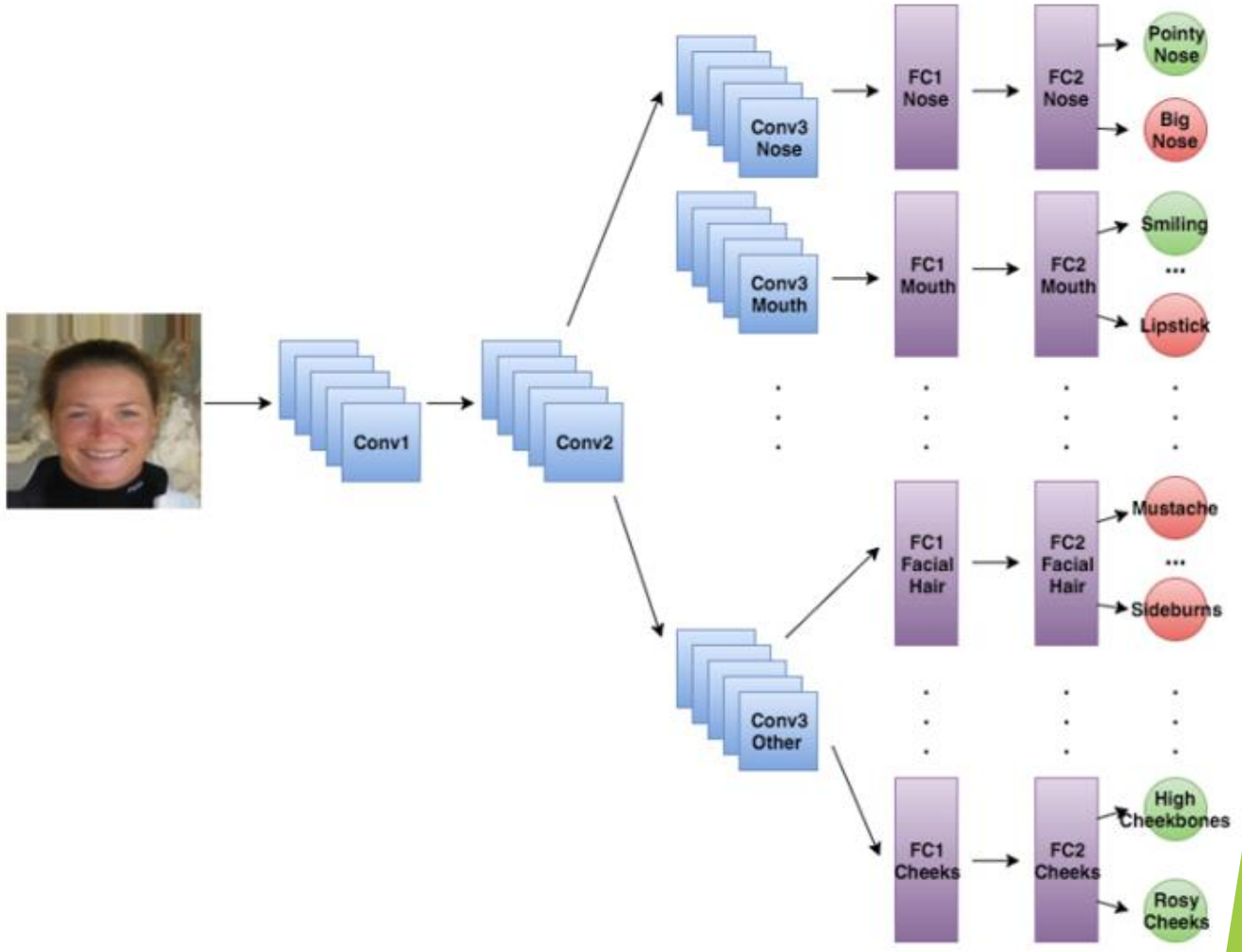
(See more detail in document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qlx6s3SElQan3SHJhnY/edit?usp=sha
g)

MCNN and AUX-Method

MCNN

A multi-task deep CNN for attribute classification.

Take an image as input and outputs 40 separate attribute scores, which are then thresholded to obtain binary outputs.



(See more detail in document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qIx6s3SElQan3SHJhnY/edit?usp=sharing)

MCNN and AUX- Method

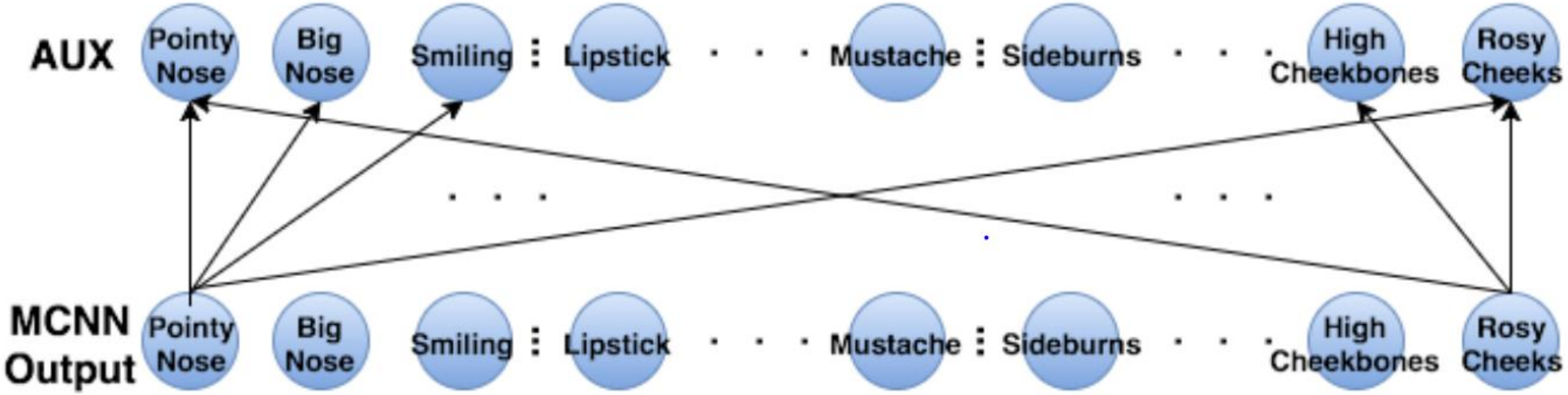
AUX → an auxiliary network for MCNN which allows for learning of attribute relationships at the score level

(See more detail in document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qlx6s3SElQan3SHJhnY/edit?usp=sharing)

MCNN and AUX-Method

Combine MCNN and AUX to create MCNN-AUX, a multi-task attribute network which utilizes implicit and explicit attribute relationships for improved classification.

(See more detail in document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qlx6s3SElQan3SHJhnY/edit?usp=sharing)



MCNN and AUX Algorithm

Our idea for looking for similar faces is to use the attributes of each face to get a vector representation for it. For this I have trained a word2vec using training set. Given the vector space for each face, it's easy to find nearby faces in the vector space.
(Source code attachment)

Installation - Environment

Programming language: Python

Tool:

- Anaconda Installer: Installing on Windows — Anaconda documentation
- Install Jupyter Notebook: Project Jupyter | Installing Jupyter Software
- Install torch, pytorch

Installation - Environment

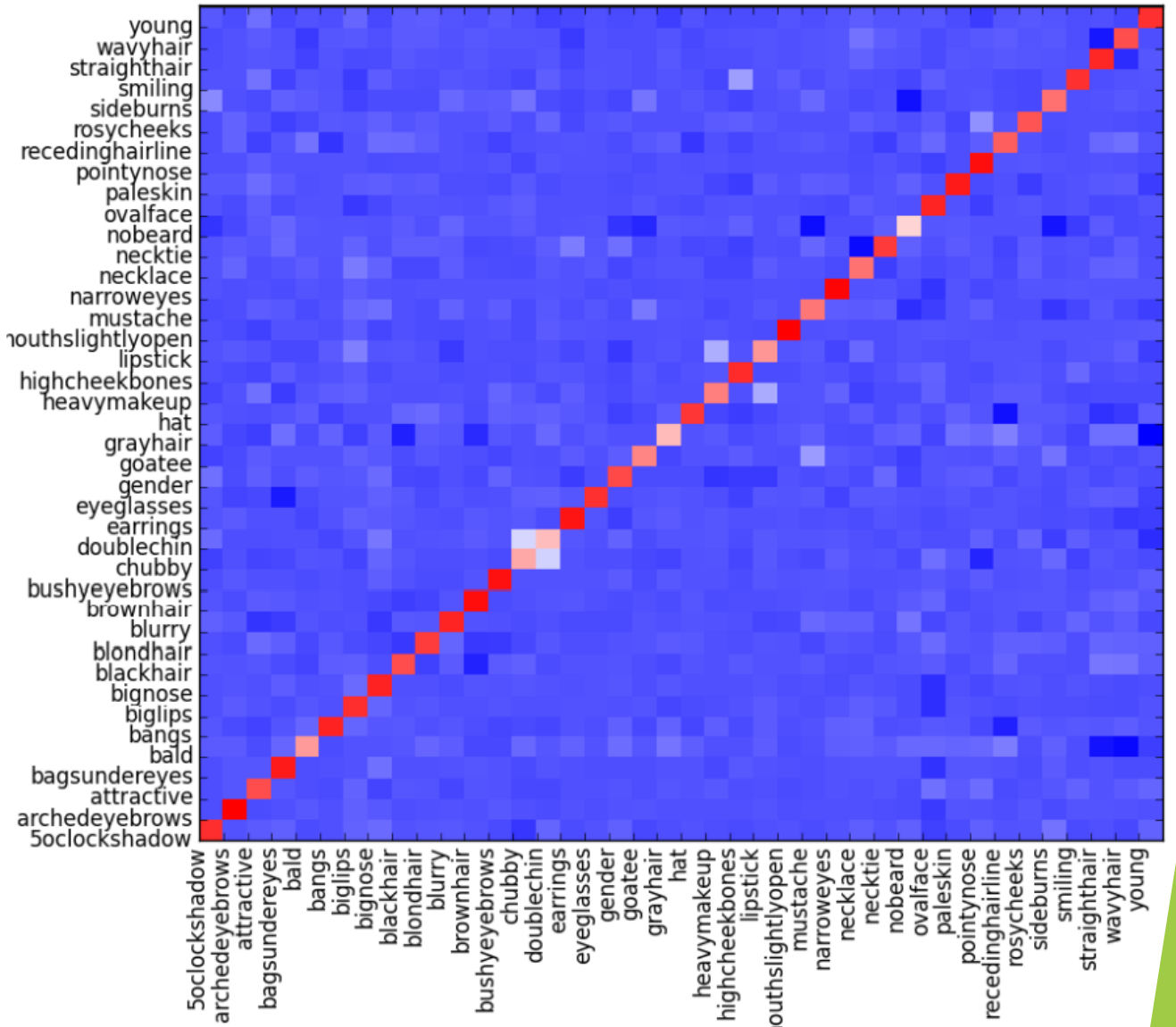
Open Anaconda Prompt, follows below command to open Jupyter Notebook:

```
jupyter notebook
```

Installation - Result

Attribute	Baseline	Liu et al.	Independent	MCNN	MCNN-AUX
5 o'clock Shadow	90.01	91	93.94	94.41	**94.51**
Arched Eyebrows	71.55	79	83.16	**83.55**	83.42
Attractive	50.41	81	82.22	82.94	**83.06**
Bags Under Eyes	79.73	79	84.83	84.89	**84.92**
Bald	97.88	98	98.85	98.87	**98.90**
Bangs	84.42	95	95.99	96.04	**96.05**
Big Lips	67.29	68	70.80	71.20	**71.47**
Big Nose	78.79	78	84.47	84.50	**84.53**
Black Hair	72.83	88	89.41	**89.87**	89.78
Blond Hair	86.67	95	95.88	95.97	**96.01**
Blurry	94.94	84	96.07	96.08	**96.17**
Brown Hair	82.03	80	88.75	88.99	**89.15**
Bushy Eyebrows	87.04	90	**92.87**	92.80	92.84
Chubby	94.69	91	95.55	95.66	**95.67**
Double Chin	95.42	92	**96.43**	96.41	96.32
Earrings	79.33	82	90.35	90.32	**90.43**
Eyeglasses	93.54	99	**99.67**	99.63	99.63
Goatee	95.41	95	97.13	**97.30**	97.24
Gray Hair	96.81	97	98.07	**98.20**	**98.20**
Hat	95.79	99	98.97	99.04	**99.05**
Heavy Makeup	59.50	90	90.95	91.37	**91.55**
High Cheekbones	51.81	**88**	87.34	87.55	87.58
Lipstick	52.18	93	93.80	93.95	**94.11**
Male	61.34	98	98.02	98.16	**98.17**
Mouth Slightly Open	50.49	92	**93.99**	93.74	93.74
Mustache	96.13	95	96.67	**96.93**	96.88
Narrow Eyes	85.13	81	87.22	87.16	**87.23**
Necklace	86.20	71	86.41	**86.82**	86.63
Necktie	92.99	93	**96.71**	96.53	96.51
No Beard	85.36	95	95.93	**96.11**	96.05
Oval Face	70.43	66	74.70	75.81	**75.84**
Pale Skin	95.79	91	**97.07**	97.01	97.05
Pointy Nose	71.42	72	**77.47**	**77.47**	**77.47**
Receding Hairline	91.51	89	93.41	**93.81**	**93.81**
Rosy Cheeks	92.82	90	95.02	95.13	**95.16**
Sideburns	95.36	96	97.77	97.82	**97.85**
Smiling	50.03	92	92.65	92.66	**92.73**
Straight Hair	79.01	73	82.62	83.39	**83.58**
Wavy Hair	63.59	80	83.24	**83.92**	83.91
Young	75.71	87	87.98	88.30	**88.48**

Installation - Result



Installation - Result

Attribute	Positive Influences	Negative Influences
Bald	Receding Hairline	Straight Hair, Wavy Hair
Bangs	N/A	Receding Hairline
Black Hair	Straight Hair, Wavy Hair	Blond Hair, Brown Hair
Blond Hair	Attractive	Black Hair, Brown Hair, Bushy Eyebrows
Chubby	Double Chin	Pointy Nose
Double Chin	Chubby, Big Nose	Young
Eyeglasses	N/A	Bags Under Eyes
Male	5 o'clock Shadow, Necktie	Earrings, Heavy Makeup, High Cheekbones, Lipstick
Goatee	Mustache	5 o'clock Shadow, No Beard
Gray Hair	Receding Hairline	Black Hair, Brown Hair, Young
Hat	Black Hair, Blond Hair	Bald, Receding Hairline
Heavy Makeup	Attractive, Lipstick	Bags Under Eyes
High Cheekbones	Smiling	N/A
Lipstick	Heavy Makeup	Male
Mustache	Goatee	No Beard
Necklace	N/A	Necktie
Necktie	Male	Necklace
No Beard	N/A	5 o'clock Shadow, Goatee, Male, Mustache, Sideburns
Receding Hairline	Bald	Bangs, Hat
Sideburns	5 o'clock Shadow, Goatee	No Beard
Smiling	High Cheekbones	Big Lips
Straight Hair	N/A	Wavy Hair
Wavy Hair	N/A	Straight Hair
Young	Attractive	Gray Hair

Installation - Result

Publication	Approach	Dataset	Accuracy
Hand and Chellappa [33]	Multi-task CNN features (3 Conv. layers and 2 FC layers); Joint regression of multiple binary attributes	CelebA (public) $(180K, 20K)$ LFWA (public) $(6,263; 6,970)$	**CelebA** 91% (Avg. of 40 attributes) **LFWA** 86% (Avg. of 40 attributes)

Installation – Result Comment

- achieve state-of-the-art performance for many attributes, some showing up to a 15% improvement over other methods
- skipping alignment or part extraction in the preprocessing stage which are expensive and error-prone processes.
- significantly decrease the number of parameters – over four times - and the amount of training time - over 16 times - required for the attribute classifier.

(See more result comments in Document attachment https://docs.google.com/document/d/1Xp2_NOgoSk9KfOiv79cZ6bH2qlx6s3SElQan3SHJhnY/edit?usp=sharing)

Responsibility

(Document attachment Personal information - Google Drive)