Attribute Classification in Low Resolution Images

Akarsha Sehwag[1], Dr. Richa Singh[1,2], Dr. Mayank Vatsa[1,2]
[1]IIIT-Delhi, India, [2]West Virginia University, USA

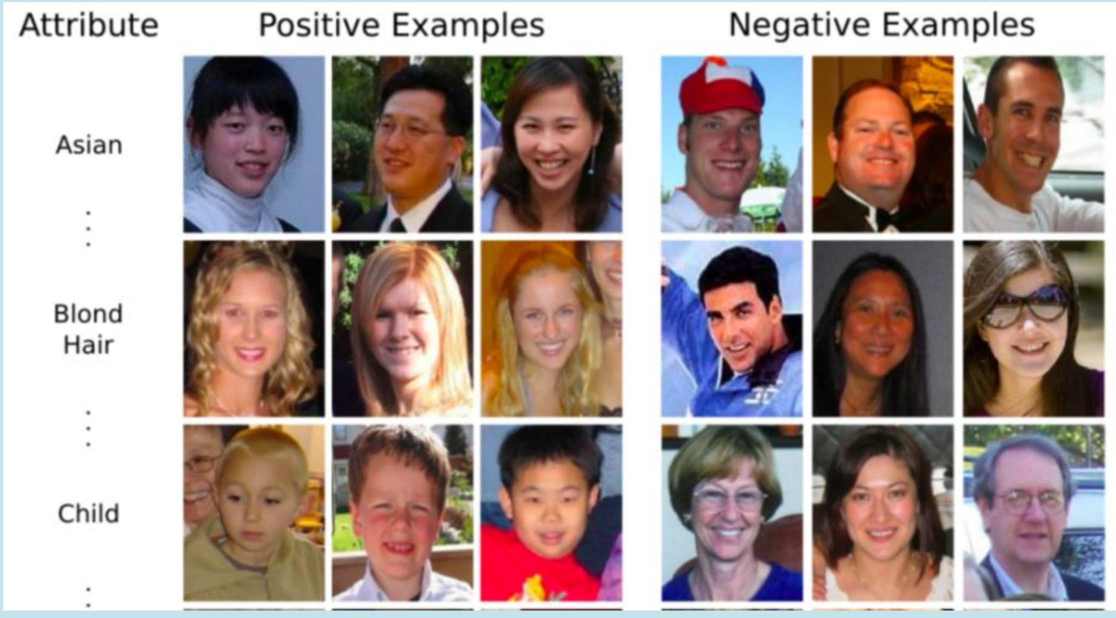
Introduction

Semantic Features commonly known as Attributes have been actively used in Biometrics industry for face identification and verification for more than ten years. Metrics show that the current state of the art is more than 90% and have been achieved for most datasets. All the current attribute classification algorithms have been trained on high resolution datasets, and thus perform poorly on datasets like missing people, where the quality of the image is extremely low. Low resolution images have only been considered for attribute classification on complete human body, not just face. We, through this project aim to classify the major attributes, so as to be able to eliminate the manual process of attribute classification for low-res data.
Currently the researchers have identified 73 facial attributes for classification out of which 40 have been considered in most cases.

Datasets

We use the two most challenging datasets available: Labeled Faces in the Wild and CelebA

SLabeled Faces in the Wild dataset : 13233 images of 5749 people and 76 attribute scores.



Large-scale CelebFaces Attributes (CelebA) Dataset: 10,177 identities, 202,599 face images, 5 landmark locations and 40 binary attributes annotations per image protocol

Implemented Algorithms

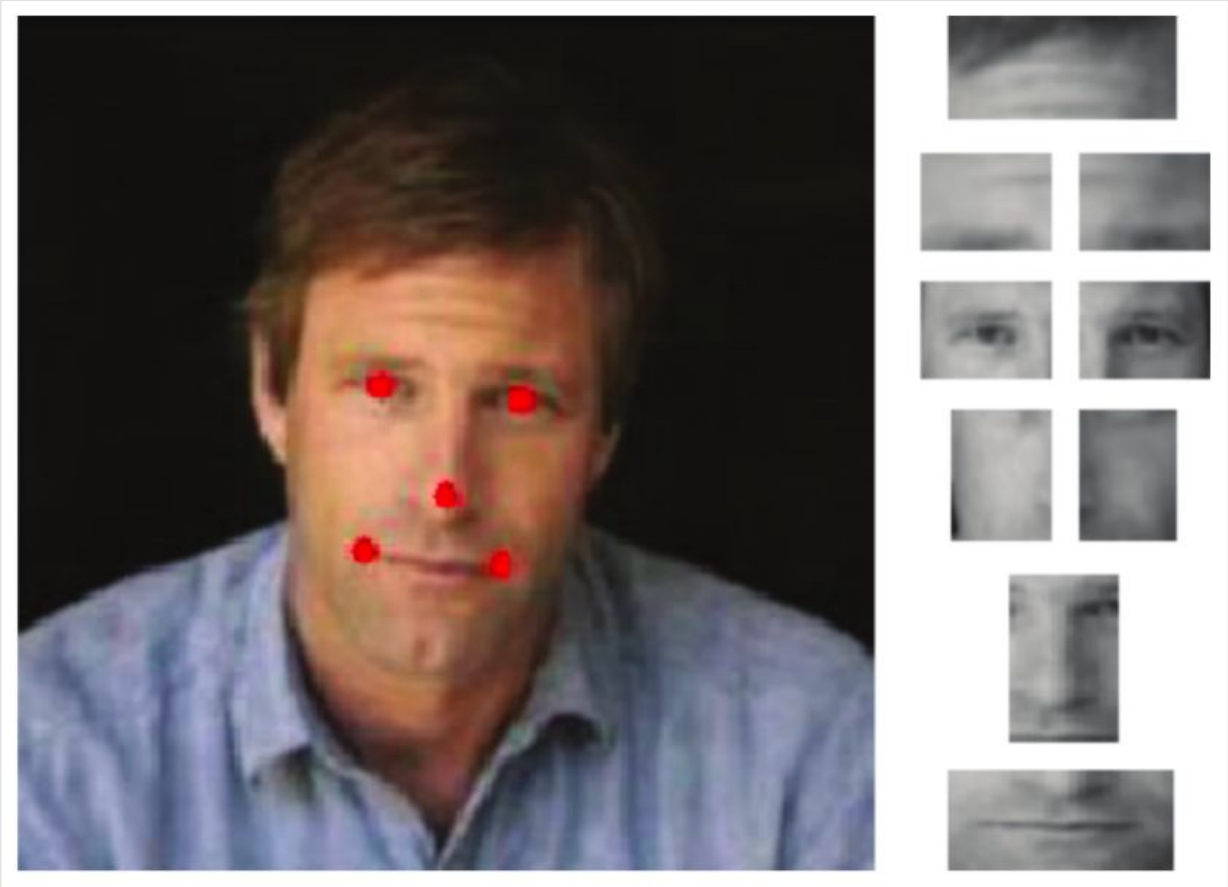


Fig. 3. Fiducial points and facial region extraction for training an SVM

SVM Classifier

A binary classifier has been trained to recognize the presence and absence of a particular attribute eg. gender, hair color, age, race etc. which is insensitive to environment changes. This has been compared with human performance. We start with extracting Low level features from the image. First, we detect fiducial point locations using OpenCV. Then pick a face region (manually labeled) and extract a feature-type and normalize it by either subtracting the mean and dividing it by std. deviation or perhaps by just dividing by the mean. Finally, aggregate or concatenate these values to construct the low level feature vector. Spatial regions are manually labelled different part of face images such as Eyes, nose, forehead etc. Thus, several attribute classifiers are trained with the positive and negative set of attributes as shown in fig.2.

Multi-Task CNN Architecture

Multi-task learning is a subield of ML where multiple learning tasks are solved at the same time. It aims to improve the performance of multiple classification tasks by learning them jointly. Multi-task learning has been successful in pose estimation, face detection and multiple other domains. We're applying MTL to attribute prediction since the relationship between the facial attributes is strong. The MCNN architecture as given in Fig 4 consists of three convolution layers followed by two fully connected layers resulting in 40 seperate attribute scores which is further thresholded to get binary outputs. In pre-processing of the images and the attributes, the image pixel value has been normalized from 0 to 1 from 0 to 255 and similarly, the attribute values are normalized from 0 to 1. Apart from this, data needs to be annotated 5 times, I am using annotations like Flip, Blur, Jittering etc.

The groups were based on spatial datapoints on the face image. I used Pytorch for my implementation. The attributes for each group with the corresponding names are as follows:

1. Gender
2. Nose
3. Eyes
4. Mouth
5. Face
6. AroundHead
7. FacialHair
8. Cheeks
9. Fat

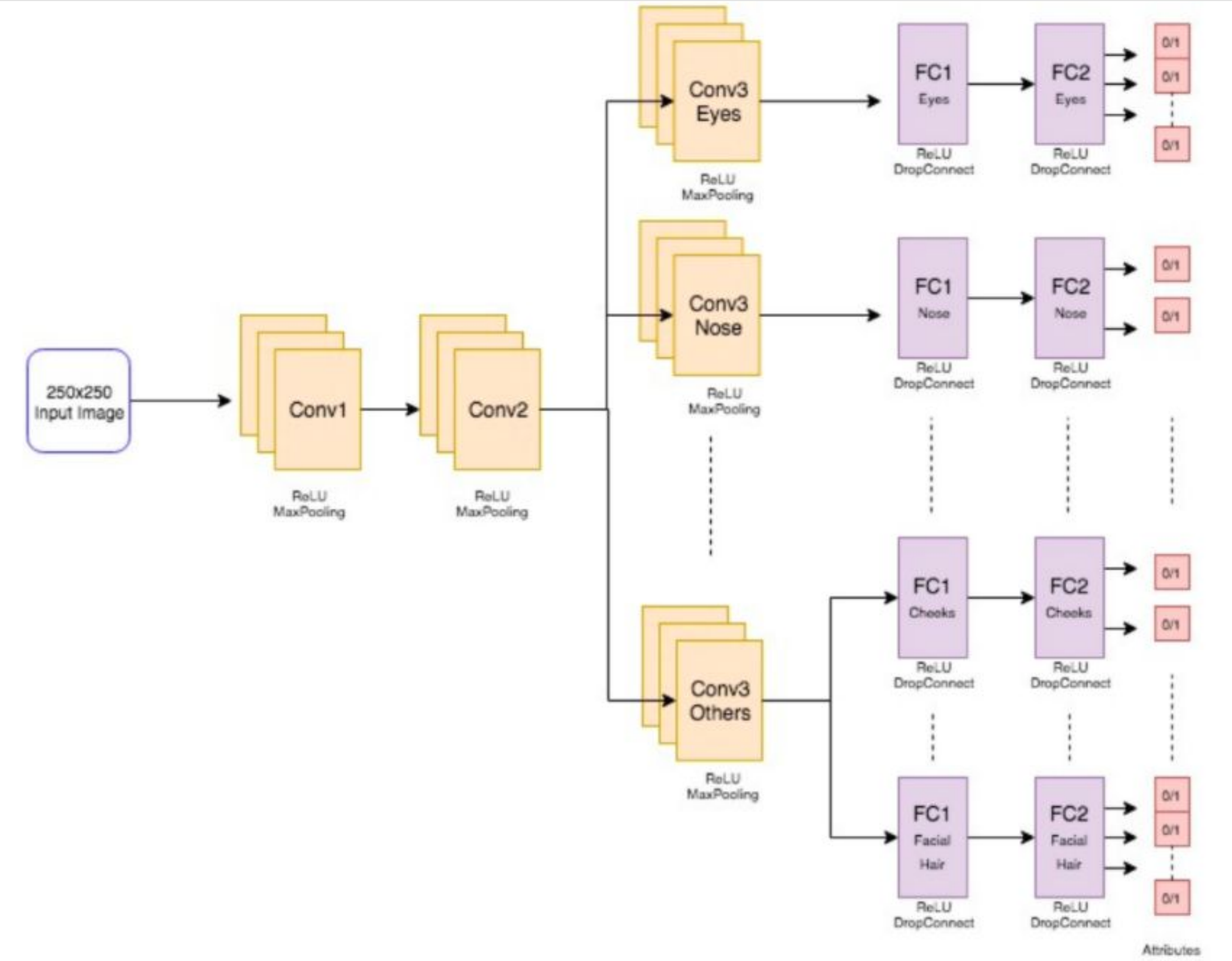


Fig. 4. Multi-Task CNN Architecture

Conclusion

In my first semester of Btech Project, I have done the following tasks:

- Deep understanding of the problem statement.
- Identify Challenges for the project.
- Thorough Literature Study.
- Implemented Baseline Models using SVM and MCNN.

Results of the two Algorithms

AResults for SVM based classification are pretty close to the baseline model. The difference arise since they used commercial systems to obtain the fiducial points which might be more accurate than an opensource model. And for the MCNN model, the results are yet not comparable since we need to augment the data to avoid overfitting. Apart from this, I am yet to test it with 50 epochs to make the results comparable to the paper.

No. of Epochs	Output	
	Average Accuracy	*Avg. Loss*
5	58.35%	0.02[a]
10	61.97%	0.018[a]
20	63.66%	0.01[a]

[a]The results are for unaugmented data.

Attribute Classification in MCNN)

Attributes	Paper	MyResults	Attributes	Paper	MyResults
Outdoor	84.8	64.38	Baby	90.45	99.54
White	91.48	89.59	Child	83.58	95.36
Senior	88.74	82.06	Middle Aged	78.39	90.07
No Beard	89.53	70.95	Round Jaw	66.99	86.98
Teeth Not Visible	91.64	77.38	Mouth Wide Open	89.63	90.52
No Eyewear	93.55	79.42	Wearing Hat	85.97	86.25
Eyeglasses	91.56	82.33	Male	81.22	76.96
Smiling	95.33	88.7	Brown Hair	72.42	62.2
Frowning	95.47	75.84	Mustache	91.88	87.16
Receding Hairline	84.15	75.79	Asian	92.32	92.17
Bushy Eyebrows	93.4	53.47	Chubby	77.24	63.84
Big Nose	87.5	68.94	Blurry	92.12	84.43
Mouth Closed	89.27	62.39	Harsh Lighting	78.74	67.94
Mouth Slightly Open	85.13	66.85	Curly Hair	68.88	59.83
Black Hair	80.32	86.34	Wavy Hair	64.49	57.38
Blond Hair	78.05	95.53	Straight Hair	76.81	63.21
Bald	83.22	89.61	Bangs	88.7	83.6
Sunglasses	94.91	98.63	Sideburns	71.07	69.03
Flash	72.33	81.15	Visble Forehead	77.02	68.4
Soft Lighting	67.81	70.12	Arched Eyebrows	80.9	73.96
Part.Visible Forehead	77.02	94.17	Eyes Open	92.52	86.25
Obstructed Forehead	79.11	85.98	Pointy Nose	86.87	69.77
Black	88.65	95.45	Big Lips		63.11
Goatee	80.35	75.44	Youth	85.79	82.79

Attribute Classification using SVM

References

[1] Emily M. Hand and Rama Chellappa: Attributes for Improved Attributes: A Multi-Task Network for Attribute Classification
[2] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. Attribute and Simile Classifiers for Face Verification. In ICCV, 2009
[3] Liu, Z., Luo, P., Wang, X., Tang, X.: Deep learning face attributes in the wild. ICCV (2015)
[4] Lubomir Bourdev, Subhransu Maji and Jitendra Malik: Describing People: A Poselet-Based Approach to Attribute Classification. ICCV (2011)
[5] Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Describable visual attributes for face verification and image search. PAMI (2011)
[6] Hui Ding, Hao Zhou, Shaohua Kevin Zhou, Rama Chellappa : A Deep Cascade Network for Unaligned Face Attribute Classification
[7] Ali Moeini, Hossein Moeini, A.M.Safai,Karim Faez : Regression Facial Attribute Classification via simultaneous dictionary learning, Elsevier,2017
[8] Matteo Fabbri,Simone Calderara, Rita Cucchiara : Generative adversarial models for people attribute recognition in surveillance, IEEE, 2017