Interpretation of Feature Space using Multi-Channel Attentional Sub-Network

Masanari Kimura[1] and Masayuki Tanaka[12]

[1]National Institute of Advanced Industrial Science and Technology, [2]Tokyo Institute of Technology

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Motivation

- Convolutional Neural Networks (CNNs) are very powerful tool of image recognition task.
- However, interpreting the internal mechanism is a challenging problem.
- We aim to obtain highly interpretable neural networks.

Multi-Channel Attention Mechanism

- We analyze the multi-attributes of the CelebA dataset.
- We propose a novel framework to visualize important features for each attribute by using an attention mechanism.

Our framework consists of the following components:

- Feature Extractor
- Binary Classifiers with Attention Mechanism
- Multi-Label Classifier
- Reconstructor

- Multi-Label Classifier and Reconstructor aim to obtain better feature representation.
- **Attention masks have the same dimensions as feature space.**

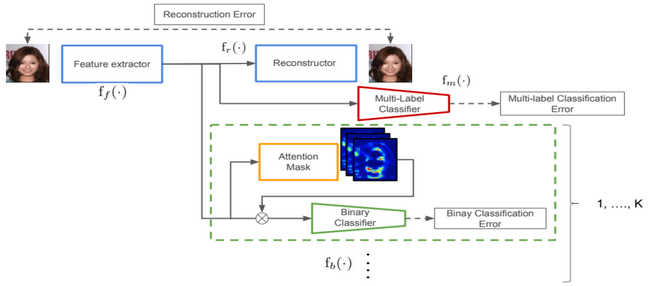

Figure 1. Overview of proposed network architecture. Here, there are K binary classification components, where K is the number of attributes.

Analysis of Relationships between Attributes

Visualization of Attention Masks

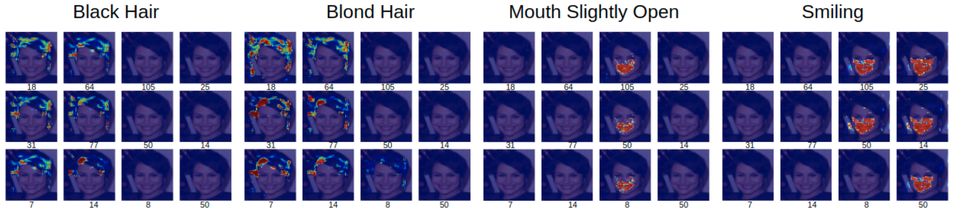

Figure 2. Visualizing attention masks on multiple facial attributes recognition. One element is one channel of the attention mask.

Top 5 Highly Correlated Attributes

Target Attribute	Top 5 Highly Correlated Attributes
Black Hair	Blond Hair, Brown Hair, Bald, Wearing Hat, Gray Hair
Heavy Makeup	Wearing Lipstick, Male, Rosy Cheeks, Attractive, Young
Bushy Eyebrows	Bags Under Eyes, Eyeglasses, Arched Eyebrows, Heavy Makeup, Attractive

Table 2. Correlation among the attributes. It lists the target attributes and the top five attributes that are highly correlated with the target.

Analysis of Relationships between Features

Our multi-channel attention can obtain correlations among each channel of the feature map.

This suggests that:

- Which features are important?
- Which features are similar?

and leads to various applications,
e.g. network pruning, image generation etc.

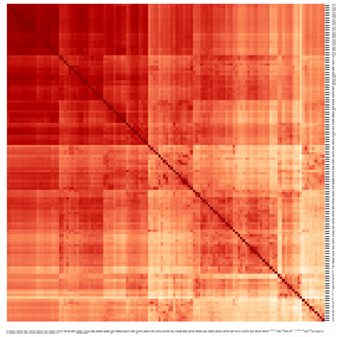

Figure 3. Visualization of correlation in feature space.

Experimental Results of the Classification Task

Classification accuracy on the celebA dataset.

Dataset Overview

- 40 facial attributes
- 182,637 training images
- 19,962 testing images

Comparison Methods

- MT-RBM PCA [1]
- LNets+ANet [2]
- FaceTracer [3]

The proposed method achieves good performance with many attributes and overall average accuracy.

Attribute	Ours	[1]	[7]	[5]
5 Shadow	**92.85**	90	91	85
Arched Eyebrows	**81.37**	77	79	76
Attractive	80.71	76	**81**	78
Bags Under Eyes	**83.79**	81	79	76
Bald	**98.30**	98	98	89
Bangs	94.10	88	**95**	88
Big Lips	**70.14**	69	68	64
Big Nose	**83.67**	81	78	74
Black Hair	**88.39**	76	88	70
Blond Hair	**95.10**	91	95	80
Blurry	**95.33**	95	84	81
Brown Hair	**86.55**	83	80	60
Bushy Eyebrows	**91.87**	88	90	80
Chubby	**96.02**	95	91	86
Double Chin	**96.68**	96	92	88
Eyeglasses	98.67	96	**99**	98
Goatee	**96.72**	96	95	93
Gray Hair	**97.89**	97	97	90
Heavy Makeup	89.49	85	**90**	85
High Cheekbone	86.77	83	**87**	84
Male	97.38	90	**98**	91
Mouth Open	**93.67**	82	92	87
Mustache	96.60	**97**	95	91
Narrow Eyes	**86.38**	86	81	82
No Beard	94.87	90	**95**	90
Oval Face	**73.33**	73	66	64
Pale Skin	**97.67**	96	91	83
Pointy Nose	**75.62**	73	72	68
Recede Hair	93.44	**96**	89	76
Rosy Cheeks	**94.67**	94	90	84
Sideburns	**97.65**	96	96	94
Smiling	**92.28**	88	92	89
Straight Hair	**81.60**	80	73	63
Wavy Hair	**81.64**	72	80	73
Earring	**84.61**	81	82	73
Hat	98.92	97	**99**	89
Lipstick	92.52	89	**93**	89
Necklace	86.37	**87**	71	68
Necktie	**96.30**	94	93	86
Young	**87.00**	81	**87**	80
Average	**92.05**	87	87	81

References

[1] Ehrlich, Max, et al. "Facial attributes classification using multi-task representation learning." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2016.
[2] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributesin the wild. InProceedings of the IEEE international conference on computer vision,pages 3730–3738, 2015.
[3] Neeraj Kumar, Peter Belhumeur, and Shree Nayar. Facetracer: A search engine forlarge collections of images with faces. InEuropean conference on computer vision,pages 340–353. Springer, 2008.

Project Page

http://www.ok.sc.e.titech.ac.jp/~mtanaka/proj/mam/