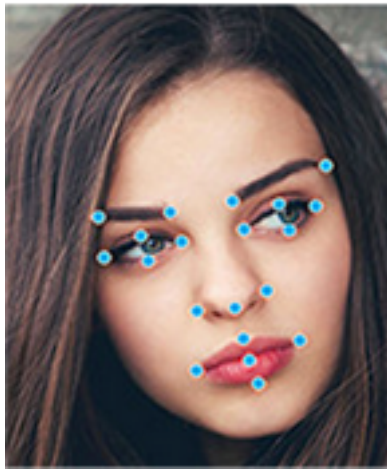
# Deep Learning For Face Analysis

Keke He

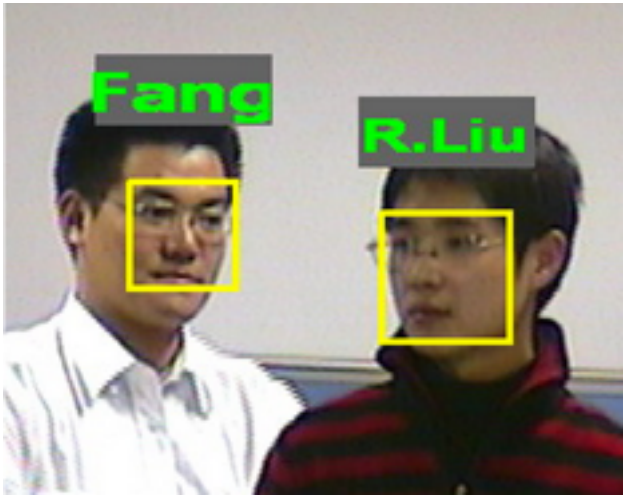
10.11 2017

# Contents



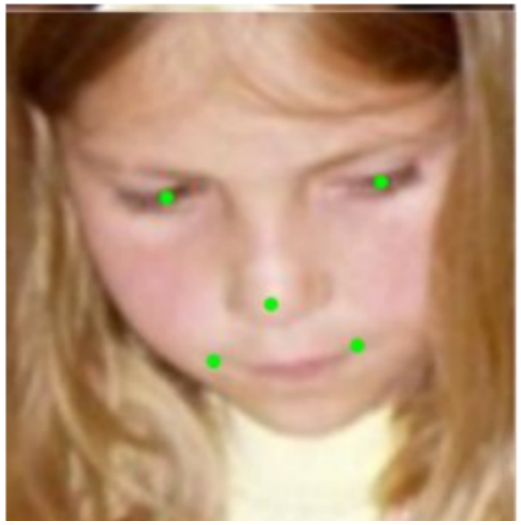
Facial Landmark Localization


Face Attribute Analysis

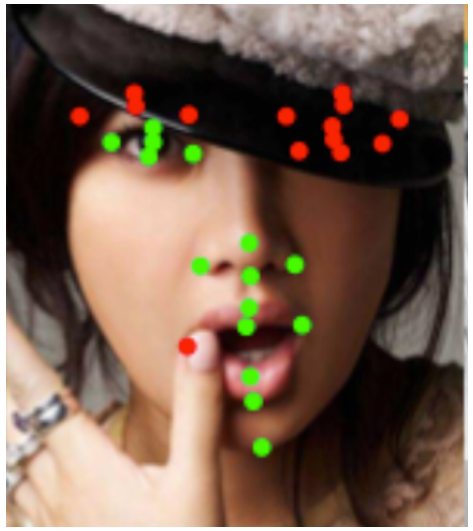



Face Recognition

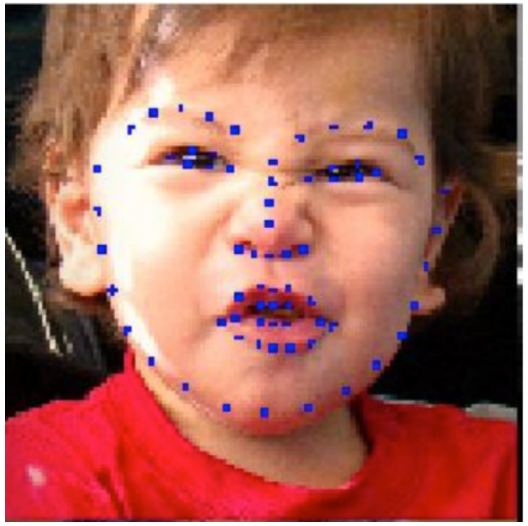
# Facial Landmark Localization



5 landmarks
cuhk-lfw-points
Total 13466



29 landmarks
COFW
Total 1007



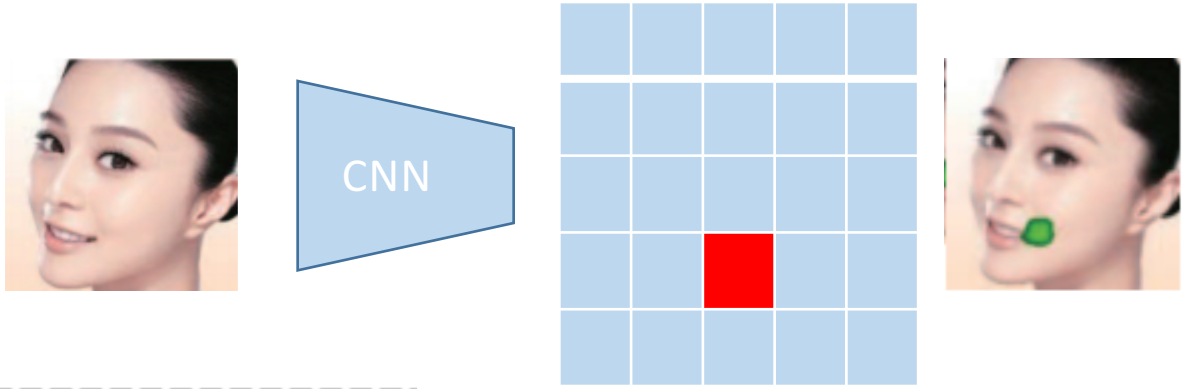
68 landmarks
AFW, LFPW, HELEN, IBUG
Total 3837

68 landmarks
LS3D-W
~230,000 images

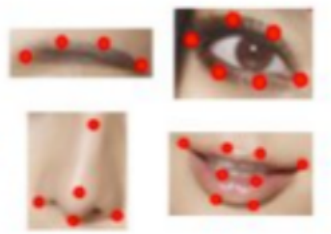
# Current Methods

- Regression vs Heat map


X1: 23
Y1: 12
…

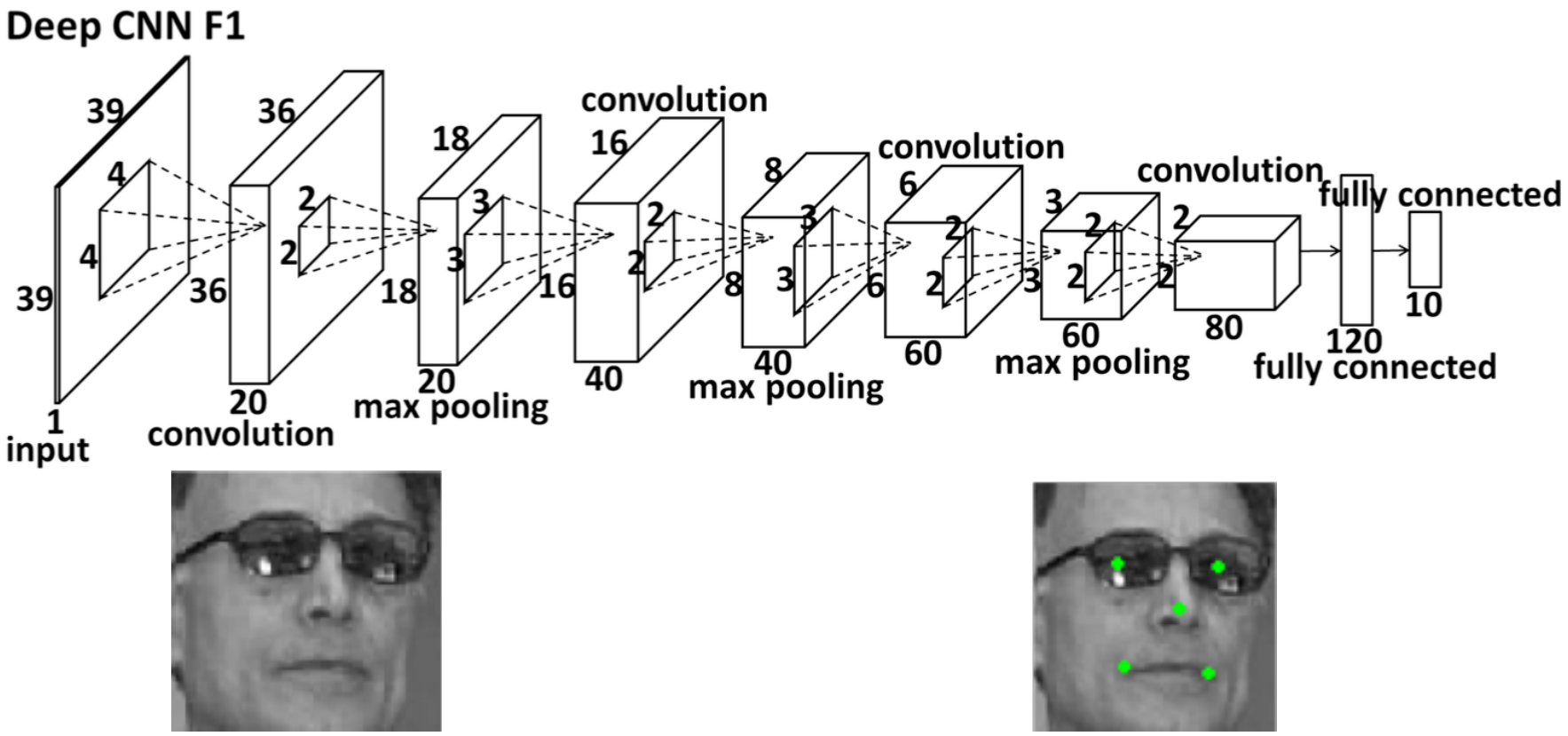



- Corse-to-fine vs Multi-task




| | |
|---|---|
| wearing glasses | × |
| smiling | × |
| gender | female |
| pose | right profile |

# Regression based



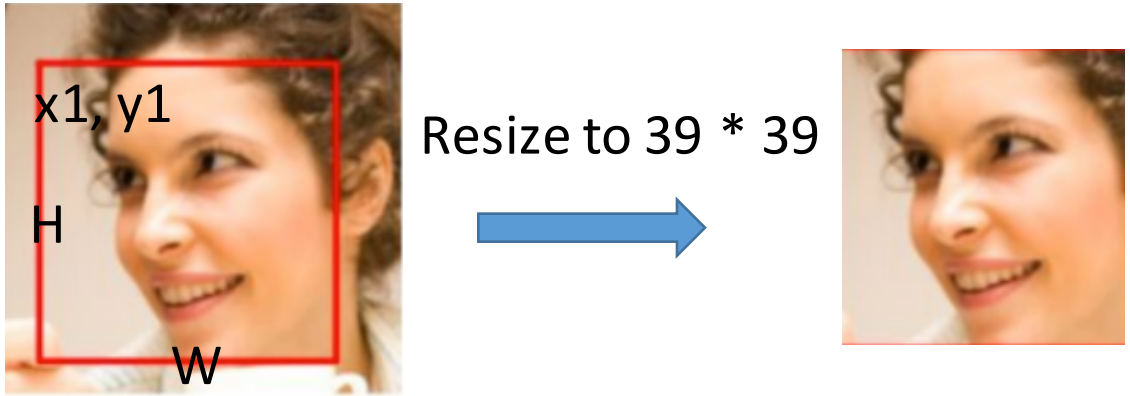





Yi Sun, Xiaogang Wang, Xiaoou Tang
Deep Convolutional Network Cascade for Facial Point Detection, CVPR 2013

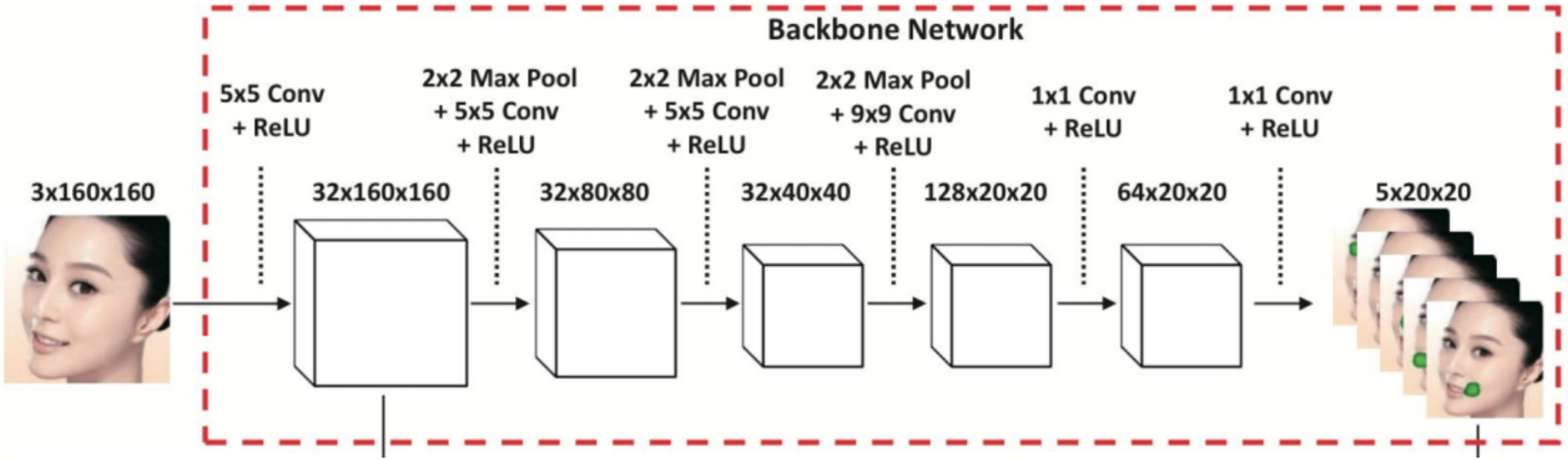
# Preprocess

- Preprocess of image




- Preprocess of landmarks
  - x = (x – x1)/W*39, y = (y – y1)/H*39 x~[0,39], y~[0,39]
  - x = (x – 19.5)/19.5 y = (y – 19.5)/19.5 x~[-1, 1], y~[-1,1]

# Heat map based



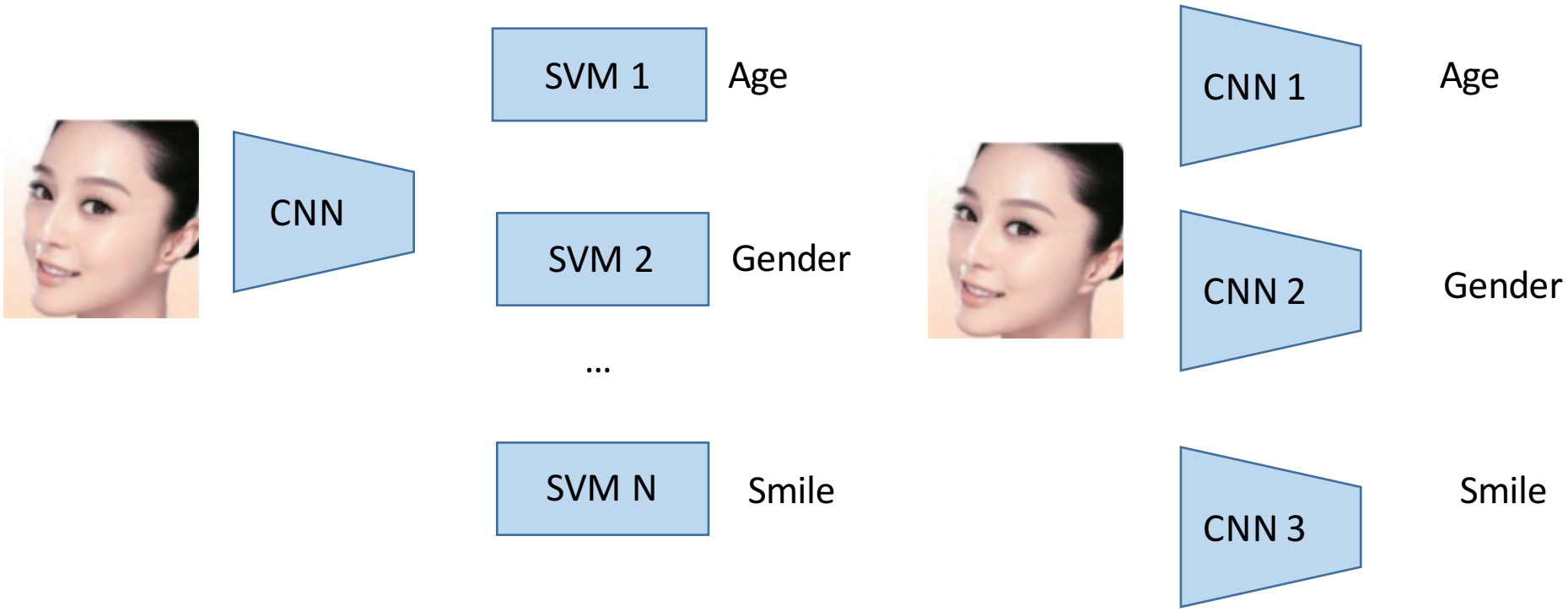

Zhujin Liang, Shengyong Ding, Liang Lin
Unconstrained Facial Landmark Localization with Backbone-Branches Fully-Convolutional Networks , arxiv 2016

# Face Attribute Analysis


CelebA
~200,000
40 attributes

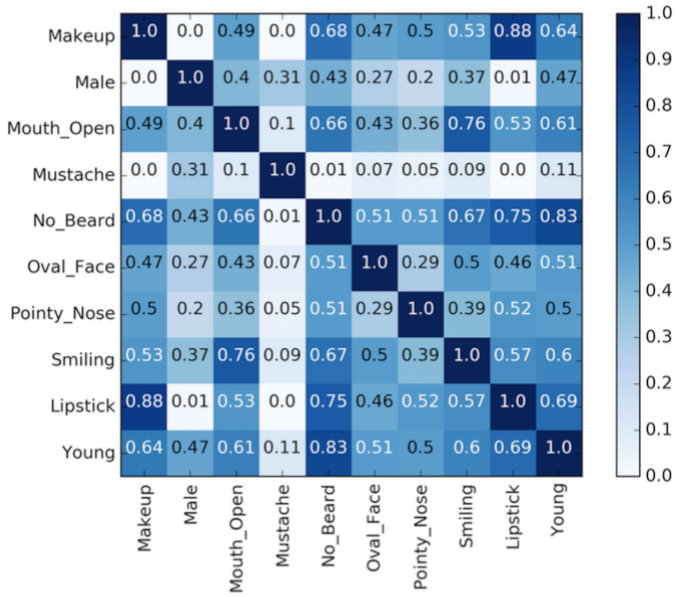
# Previous Methods

- Single model for one attribute

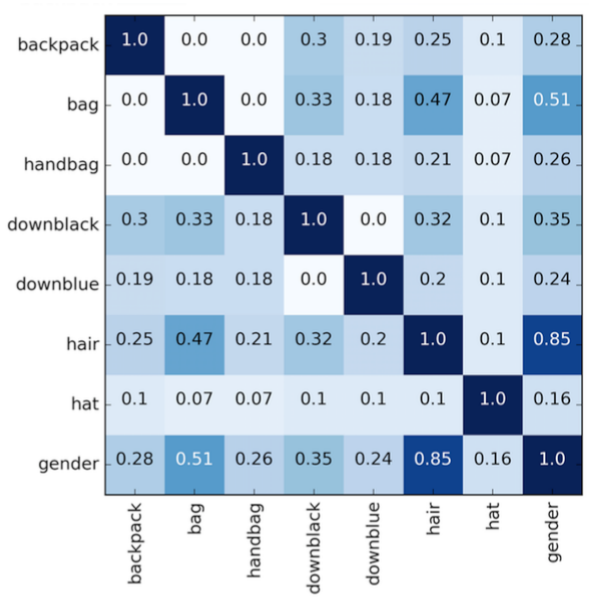



Ziwei Liu, Ping Luo, Xiaogang Wang, Xiaoou Tang
Deep learning face attributes in the wild. CVPR 2015

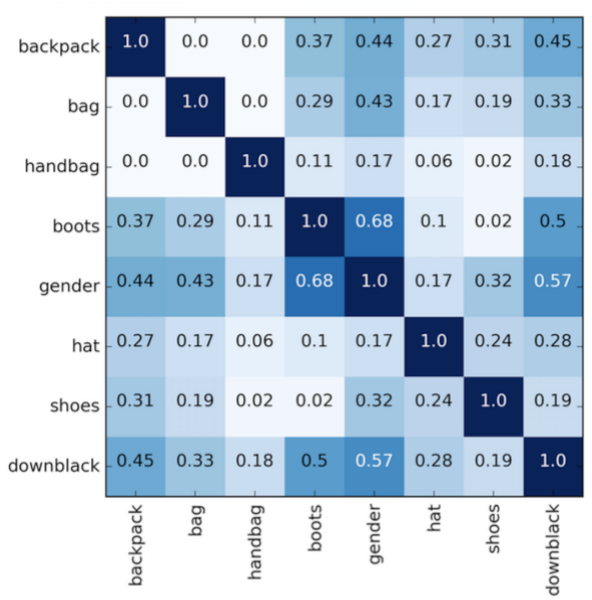
# Correlation in Attributes




(a) CelebA dataset

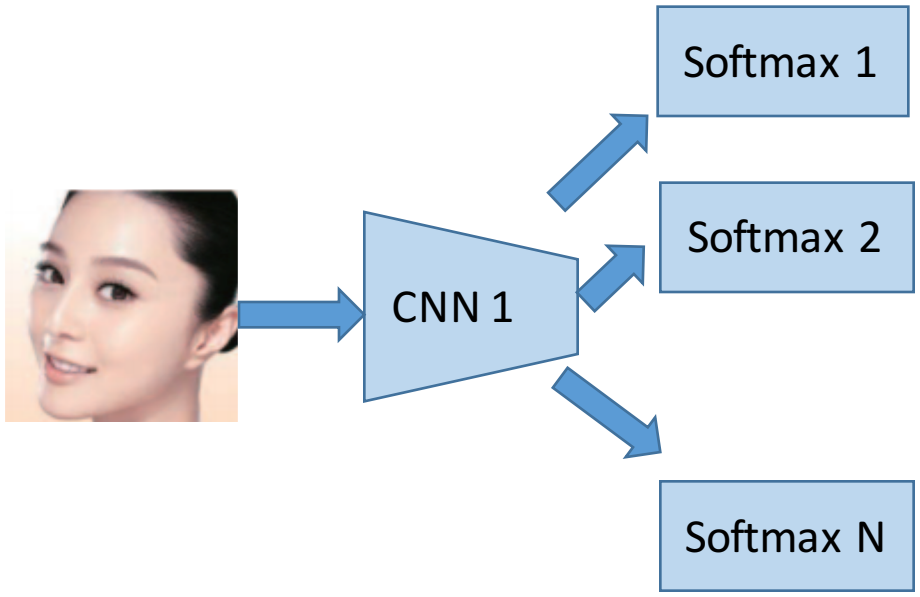



(b) Market-1501 attribute dataset



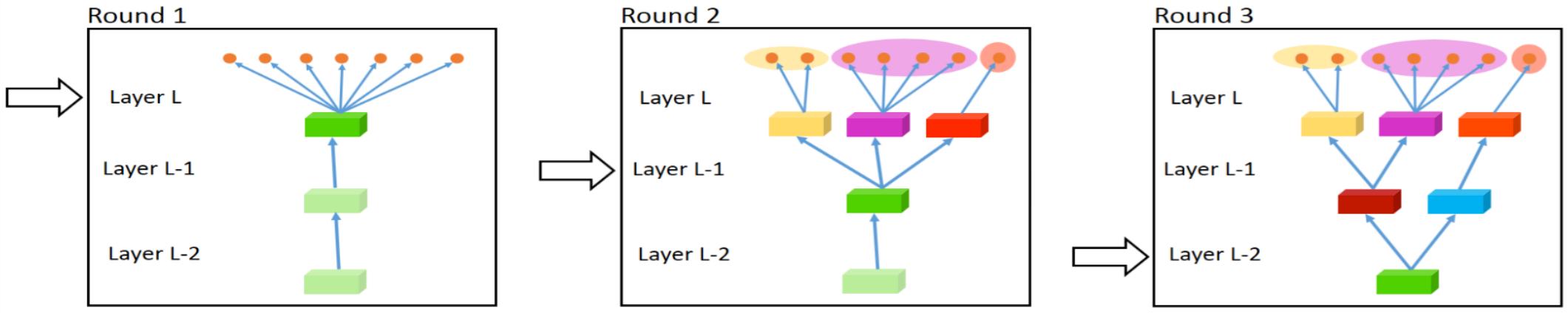

(c) Duke attribute dataset

# Previous Multi-task Methods




Ethan Rudd, Manuel Günther, Terrance Boult
Moon: A mixed objective optimization network for the recognition of facial attributes ECCV 2016

# Previous Multi-task Methods




Multi-task sharing structure

Yongxi Lu, Abhishek Kumar, Shuangfei Zhai, Yu Cheng, Tara Javidi, Rogerio Feris
Fully-adaptive Feature Sharing in Multi-Task Networks with Applications in Person Attribute Classification CVPR 2017

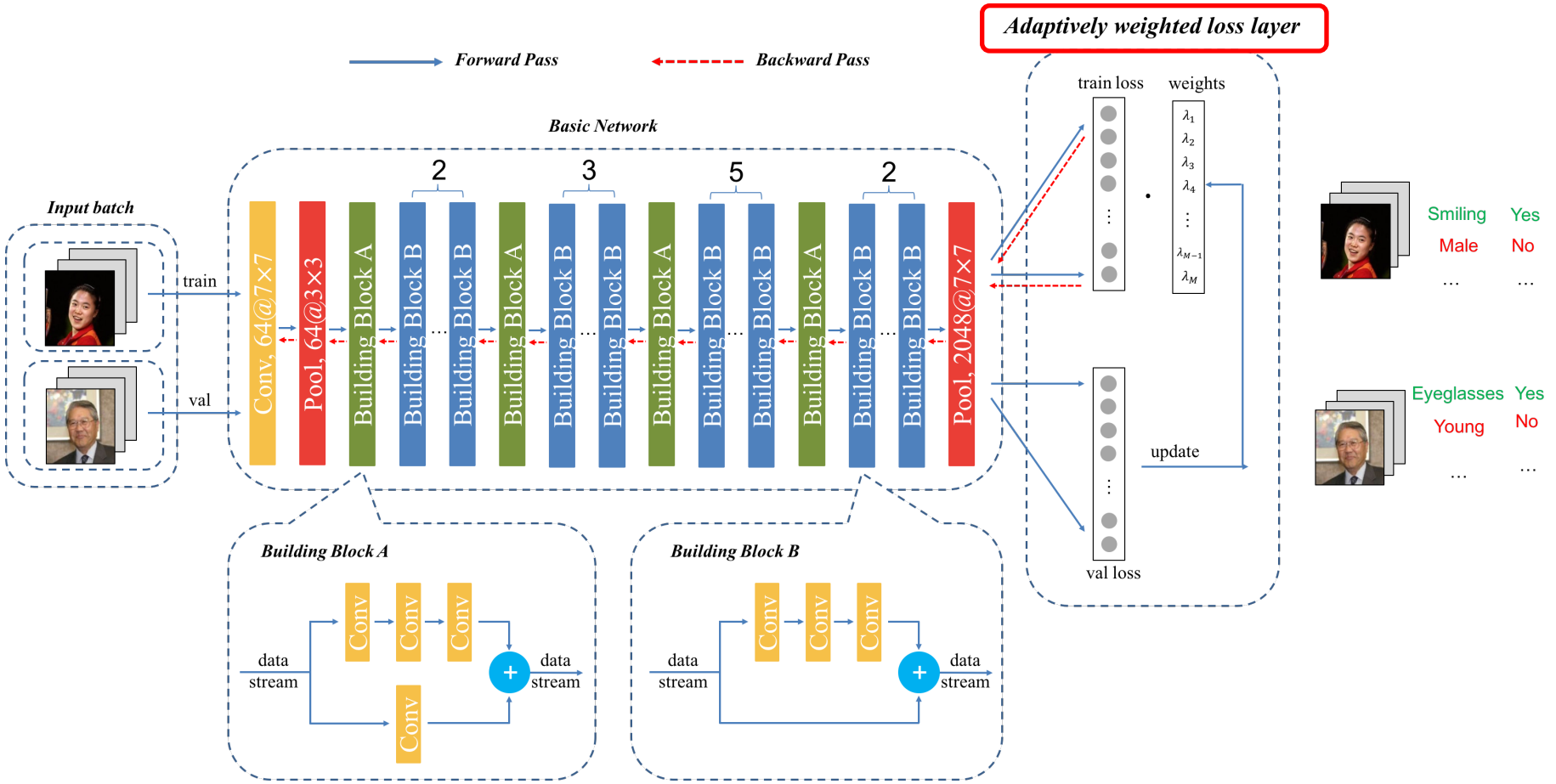
# Motivation

- When training multi-task attribute network
- Different tasks have various learning difficulties

Wearing Hat ?

Smiling ?

# Framework




Keke He, Zhanxiong Wang, Yanwei Fu, Rui Feng, Yu-Gang Jiang, Xiangyang Xue.
Adaptively Weighted Multi-task Deep Network for Person Attribute Classification. ACM MM 2017

# Approach-Training

**Algorithm 1** Network Training Procedure. $c$ is the current training iteration, $\lambda$ is the weights of all the tasks, $val\_loss\_list$ is a data structure to save the validation loss.

**Require:** $k$: the weight updating period.
**Initialize:** $c = 0$, $\lambda = \mathbf{1}$, $val\_loss\_list = []$.
**while** $c < max\_iter$ **do**
  $train\_loss, val\_loss \leftarrow net.feedward()$
  $val\_loss\_list.append(val\_loss)$
  **if** $c\%k = 0$ **then**
    $\lambda = update\_weights()$
  **end if**
  $weighted\_loss \leftarrow train\_loss * \lambda$
  $net.backward(weighted\_loss)$
  $c = c + 1$
**end while**

# Approach-Training

**Algorithm 2** Update weights of tasks: $update\_weights()$

**Require:** $k$ is the weight updating period, $c$ is the current training iteration, $val\_loss\_list$ is a data structure to save validation loss.
**Initialize:** $\lambda = 1$
**if** $c >= 2 * k$ **then**
  $pre\_mean \leftarrow mean(val\_loss\_list[c - 2 * k : c - k])$
  $cur\_mean \leftarrow mean(val\_loss\_list[c - k : c])$
  $trend \leftarrow abs(cur\_mean - pre\_man)/cur\_mean$
  $norm\_trend \leftarrow trend/mean(trend)$
  $norm\_loss \leftarrow cur\_mean/mean(cur\_mean)$
  $\lambda \leftarrow norm\_trend * norm\_loss$
  $\lambda \leftarrow \lambda/mean(\lambda)$
**end if**
return $\lambda$

# Datasets


Market-1501
Hat
Backpack
Handbag
Duke
Hat
Backpack
Handbag
CelebA
Hat
Eyeglasses
Smiling

# Experiments

| Methods | Accuracy (%) |
|---|---|
| FaceTracer [13] | 81 |
| PANDA-w [35] | 79 |
| PANDA-l [35] | 85 |
| LNets+ANet [20] | 87 |
| MT-RBM-PCA [6] | 87 |
| Off-the-Shelf CNN [42] | 86.6 |
| Walk-and-Learn [30] | 88 |
| Moon [26] | 90.94 |
| Adaptive Sharing [21] | 91.26 |
| Separate Models | 89.63 |
| Basic Model | 90.42 |
| Our Model | **91.80** |

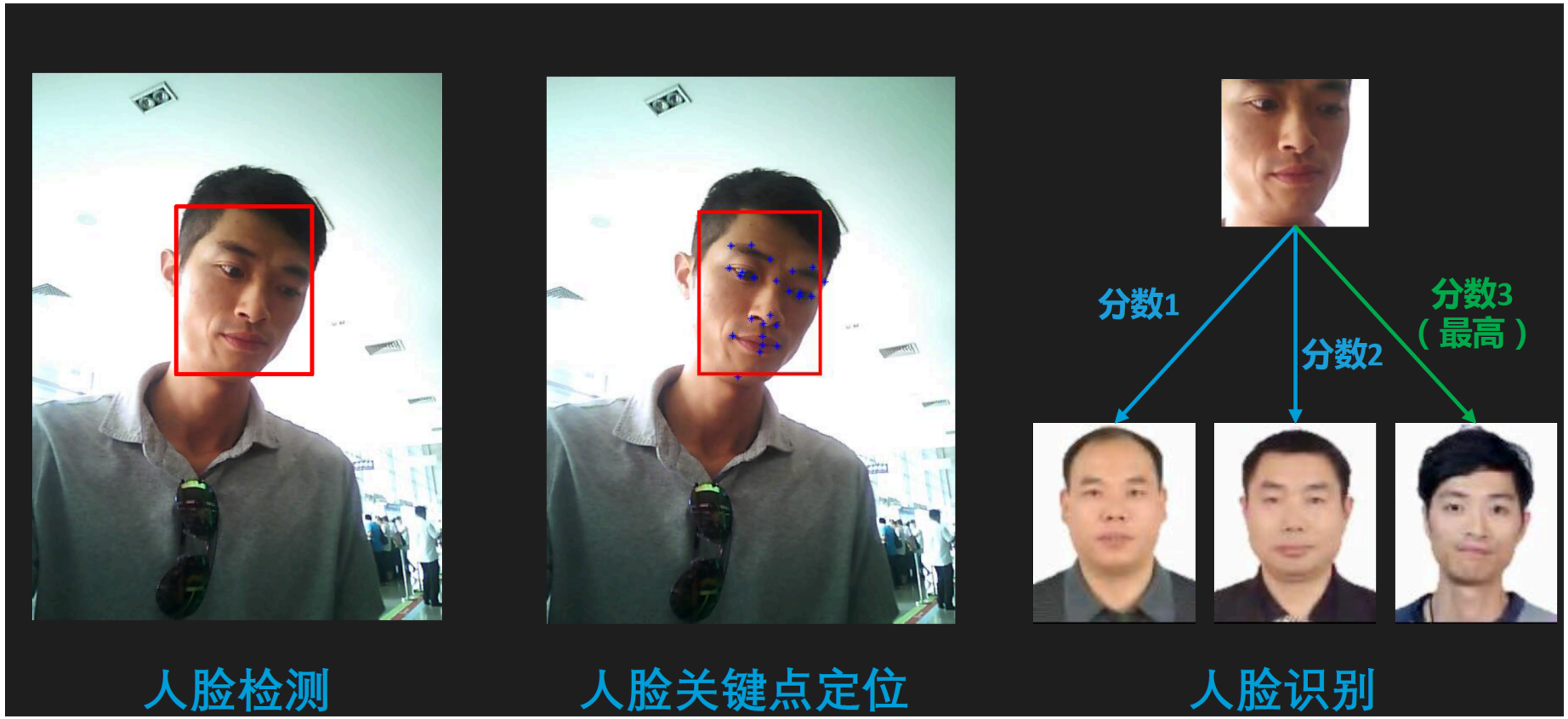
**Table 1: Comparison of mean accuracy on CelebA.**

# Experiments

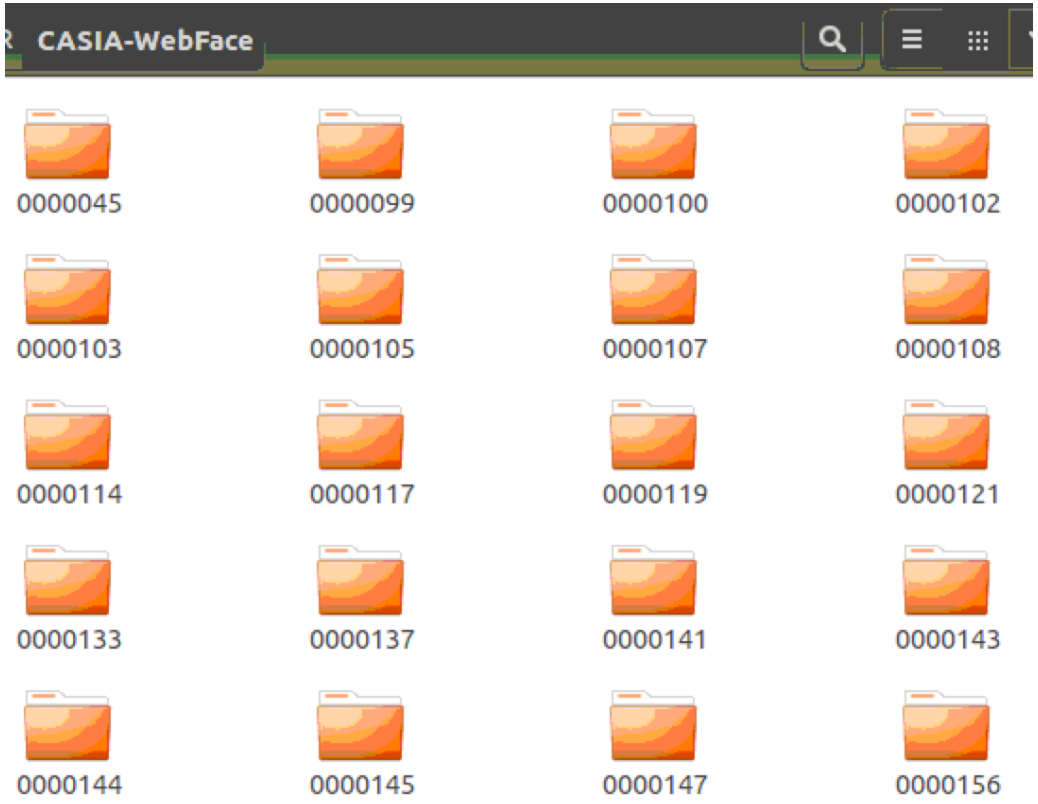
| Methods | Market-1501 (%) | Duke (%) |
|---|---|---|
| Ped_attrb_net [17] | 86.19 | 82.39 |
| APR [17] | 88.16 | 86.42 |
| Separate Models | 86.68 | 85.45 |
| Basic Model [35] | 86.84 | 85.91 |
| Our Model | **88.49** | **87.53** |

**Table 2: Comparison of mean recognition accuracy on Market-1501 and Duke attribute datasets.**

# Face Recognition



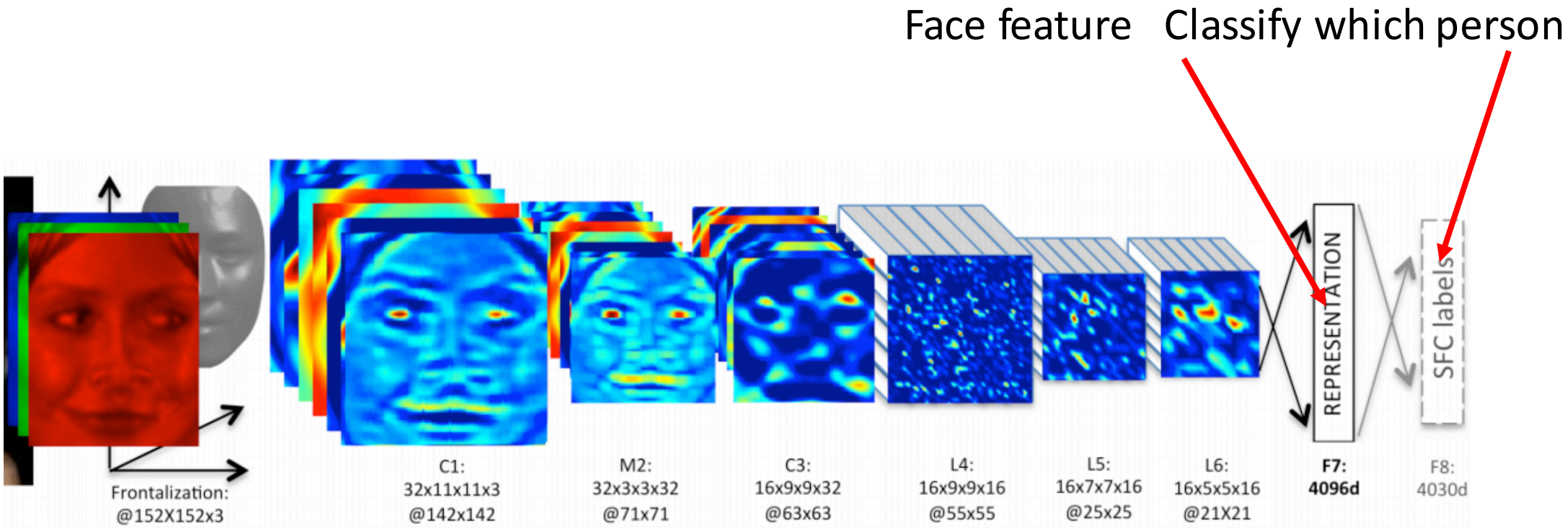
# Data





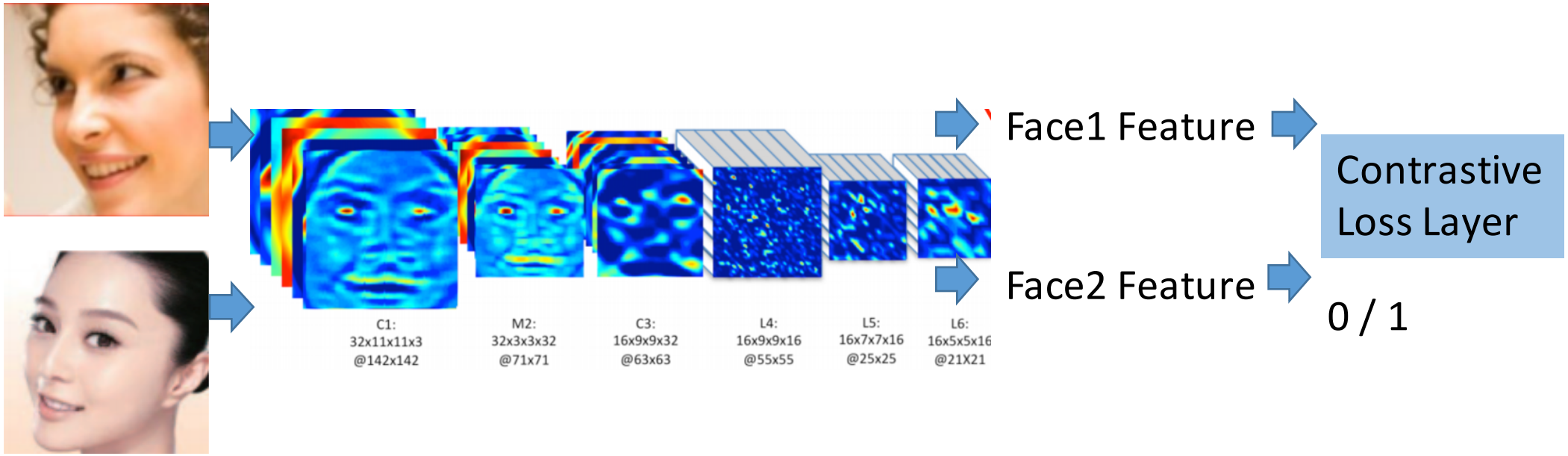
Casia-WebFace: 10575 person, total 494,414 images

# Softmax Loss

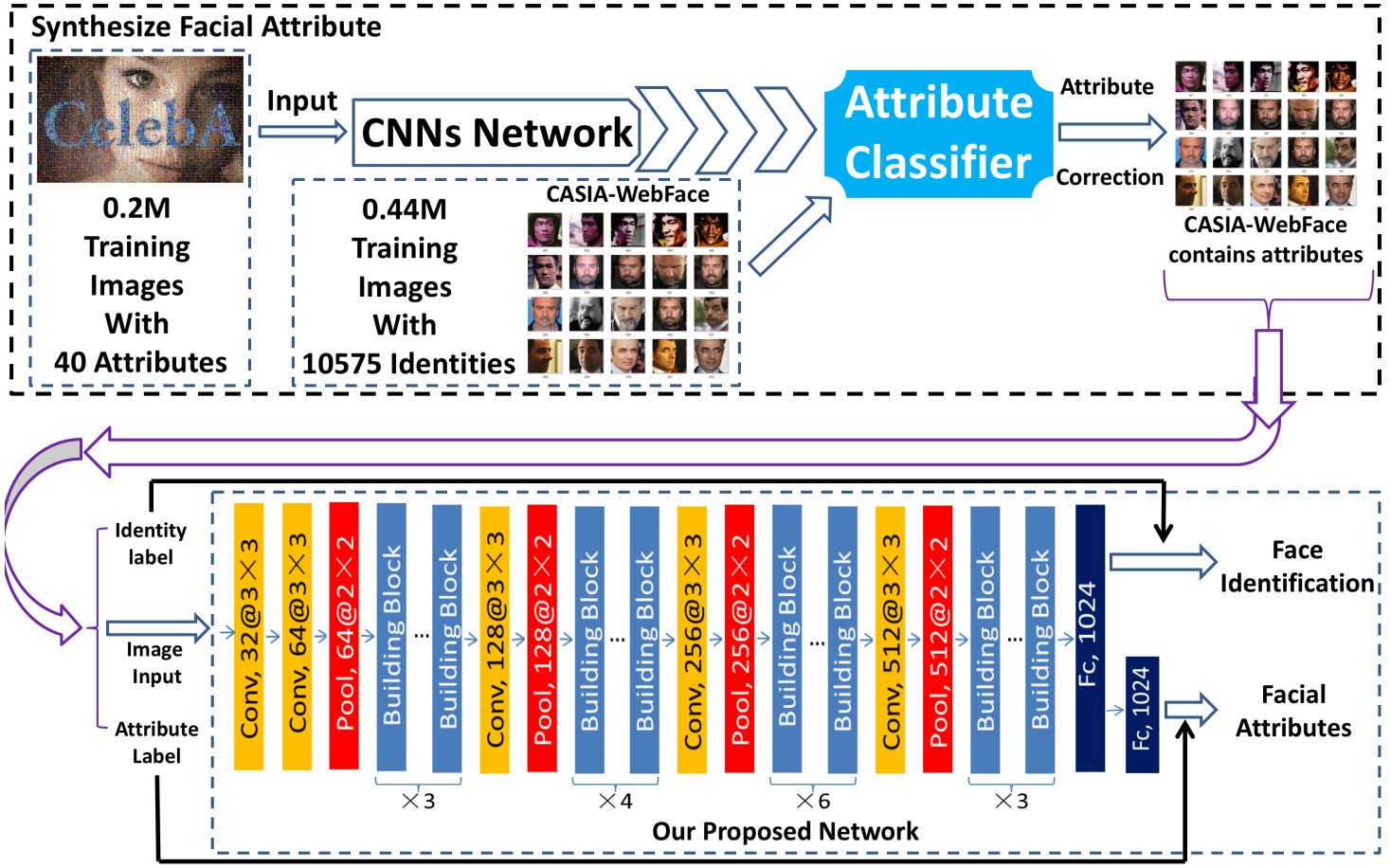



Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, Lior Wolf
DeepFace: Closing the Gap to Human-Level Performance in Face Verification. CVPR 2014

# Contrastive Loss



# This work focuses on Feature




Zhanxiong Wang*, Keke He*, Yanwei Fu, Rui Feng, Yu-Gang Jiang, Xiangyang Xue.
Multi-task Deep Neural Network for Joint Face Recognition and Facial Attribute Prediction. ICMR 2017

# Attributes

**Table 1: Selected attributes of the CelebA Dataset.**

| Group | Attributes |
|---|---|
| eye | Narrow Eyes |
| nose | Big Nose, Pointy Nose |
| global | Chubby, Double Chin, High Cheekbones, Male |

# MegaFace datasets



MegaFace is a very challenging dataset to evaluate the performance of face recognition algorithms. It includes gallery set and probe set. The gallery set consists of 690K different individuals with more than 1 million images.

# Experiments

**Table 2: Face identification Rank-1 results on MegaFace with 1M distractors. 'Release' indicates whether the details of this method are publicly released.**

| Method | Rels | Protocol | Acc. (%) |
|---|---|---|---|
| Google - FaceNet v8 [32] | ✓ | Large | 70.50 |
| NTechLAB - Large | × | Large | 73.30 |
| Faceall Co. - Norm_1600 | × | Large | 64.80 |
| Faceall Co. - FaceAll_1600 | × | Large | 63.98 |
| Lightened CNN [44] | ✓ | Small | 67.11 |
| Center Loss [43] | ✓ | Small | 65.23 |
| LBP [2] | ✓ | Small | 3.02 |
| Joint Bayes [5] | ✓ | Small | 2.33 |
| NTechLAB -Small | × | Small | 58.22 |
| 3DiVi Company | × | Small | 33.71 |
| SIAT_MMLAB | × | Small | 65.23 |
| Barebones FR | × | Small | 59.36 |
| Fast-Model (without Attribute) | ✓ | Small | 70.28 |
| Fast-Model | ✓ | Small | 73.07 |
| Full-Model (without Attribute) | ✓ | Small | 74.25 |
| Full-Model | ✓ | Small | **77.74** |

# On Mega-Face Challenge, our method can achieve the best Rank-1 identification performance under the small protocol :

| Algorithm | Date Submitted | Set 1 | Set 2 | Set 3 | Data Set Size (Large is >500K photos trained) |
|---|---|---|---|---|---|
| Vocord - deepVo V3 | 04/27/2017 | 91.763% | 91.711% | 91.704% | Large |
| YouTu Lab (Tencent Best-Image) | 04/08/2017 | 83.290% | 83.267% | 83.295% | Large |
| DeepSense V2 | 1/22/2017 | 81.298% | 81.298% | 81.298% | Large |
| Vocord-deepVo1.2 | 12/1/2016 | 80.258% | 80.195% | 80.241% | Large |
| Fudan University - FUDAN-CS_SDS | 1/29/2017 | 77.982% | 78.006% | 77.990% | Small |
| GRCCV | 12/1/2016 | 77.677% | 77.021% | 77.147% | Small |
| Beijing Faceall Co. - FaceAll V2 | 04/28/2017 | 76.661% | 76.643% | 76.607% | Small |
| SphereFace - Small | 12/1/2016 | 75.766% | 75.765% | 75.770% | Small |
| Vocord - DeepVo1 | 08/3/2016 | 75.127% | 75.093% | 75.125% | Large |
| DeepSense - Large | 07/31/2016 | 74.799% | 74.780% | 74.813% | Large |
| SIATMMLAB TencentVision | 12/1/2016 | 74.207% | 74.213% | 74.195% | Large |
| Shanghai Tech | 08/13/2016 | 74.049% | 74.032% | 74.020% | Large |
| NTechLAB - facenx_large | 10/20/2015 | 73.300% | 73.309% | 73.287% | Large |
| ForceInfo | 04/7/2017 | 72.11% | 72.084% | 72.121% | Large |
| 3DiVi Company - tdvm V2 | 04/15/2017 | 71.742% | 71.727% | 71.703% | Large |
| DeepSense - Small | 07/31/2016 | 70.983% | 70.948% | 70.962% | Small |
| Google - FaceNet v8 | 10/23/2015 | 70.496% | 70.492% | 70.551% | Large |

# Three factors

- Data
  - Collect and label more data
  - Synthesis more data [Masi et al ECCV 2016, Tran et al CVPR 2017]
- Structure
  - Network structure
- Metric Learning (Loss)
  - Softmax Loss [Sun et al CVPR 2014, Taigman et al CVPR 2014]
  - Contrastive Loss [Sun et al NIPS 2014]
  - Triplet loss [Schoroff et al CVPR 2015]
  - Center loss [Wen et al ECCV 2016]
  - Angular softmax loss [Liu et al CVPR 2017]

# What’s special in Face?

- Face Size  (Dense box)

- Locally Connected Convolution

# Locally Connected Convolution

2 * 2
Locally
convolution

# Outlook

- Scenario：
  - Pose
  - Light

- Data
  - How to leverage current millions of face data.

Thanks