Low-Resolution Face Recognition

Final-year Project Presentation

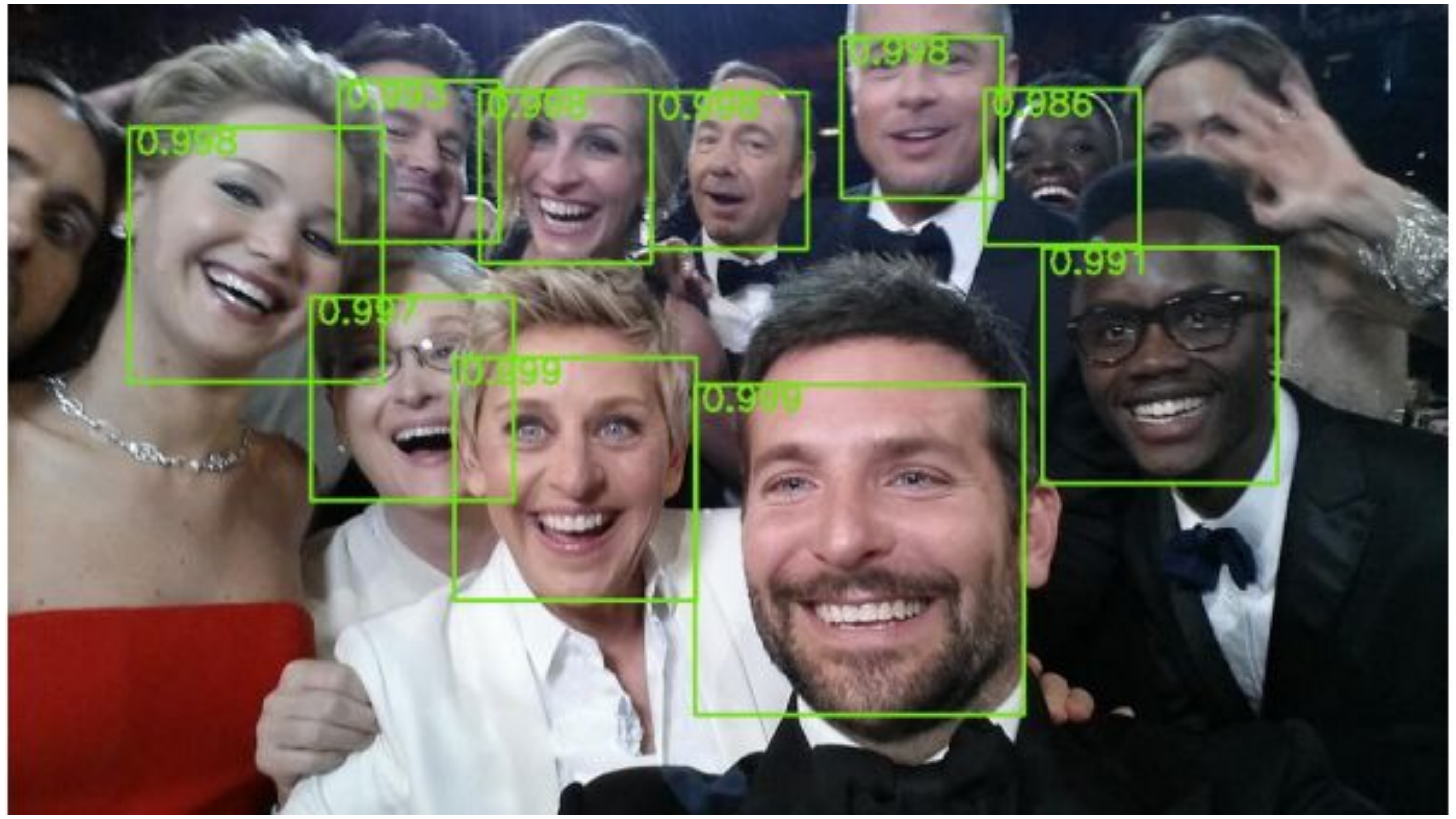
Department of Electronic and Information Engineering
Hong Kong Polytechnic University
Jiawei Tang
14109816d
Supervisor: Prof. Kenneth Lam
Assessor: Dr. Y. L. Chan

CONTENT

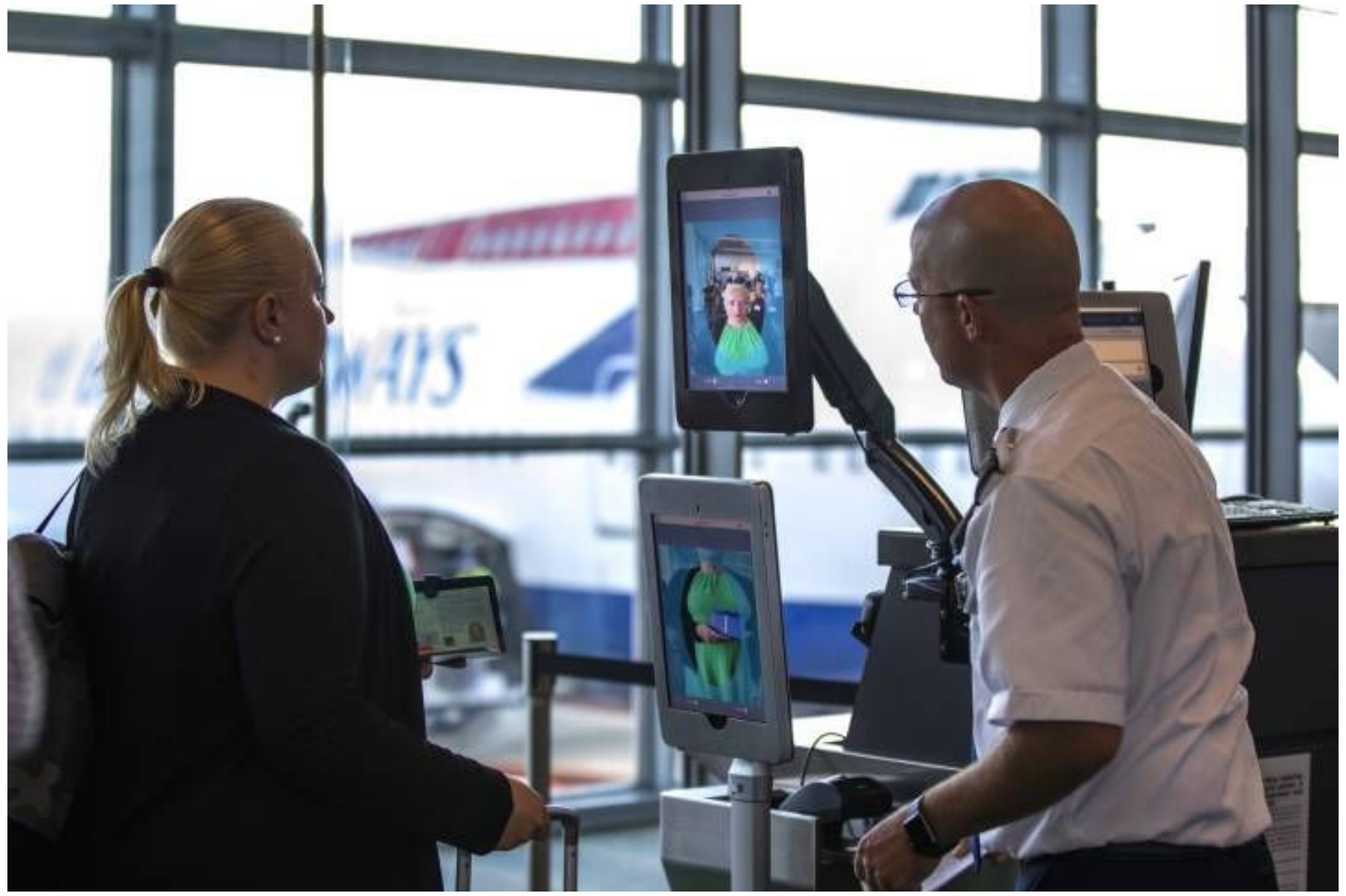
Introduction

Background

Face Recognition in our daily life



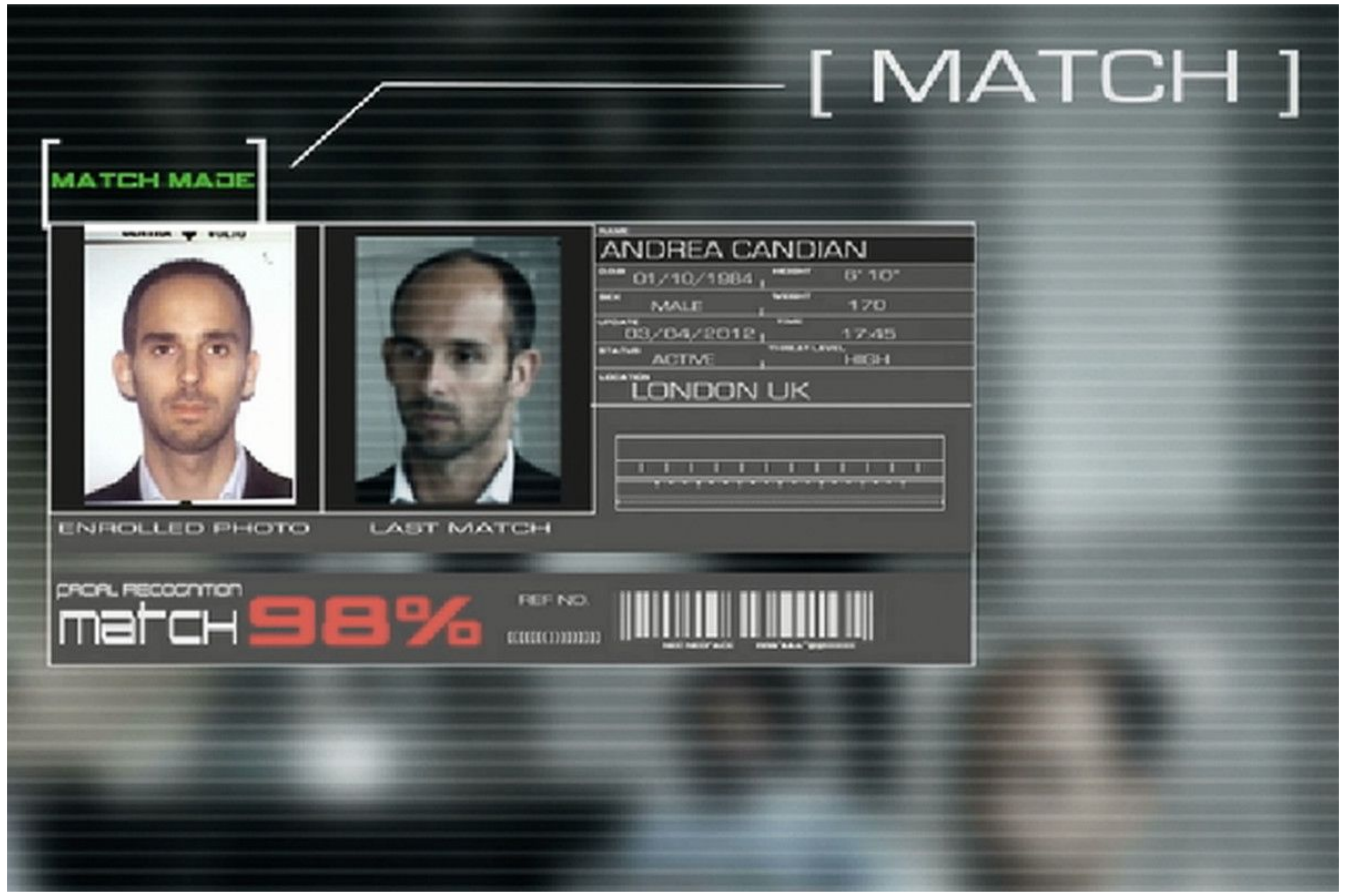
Face detection in digital camera



Access control

Background

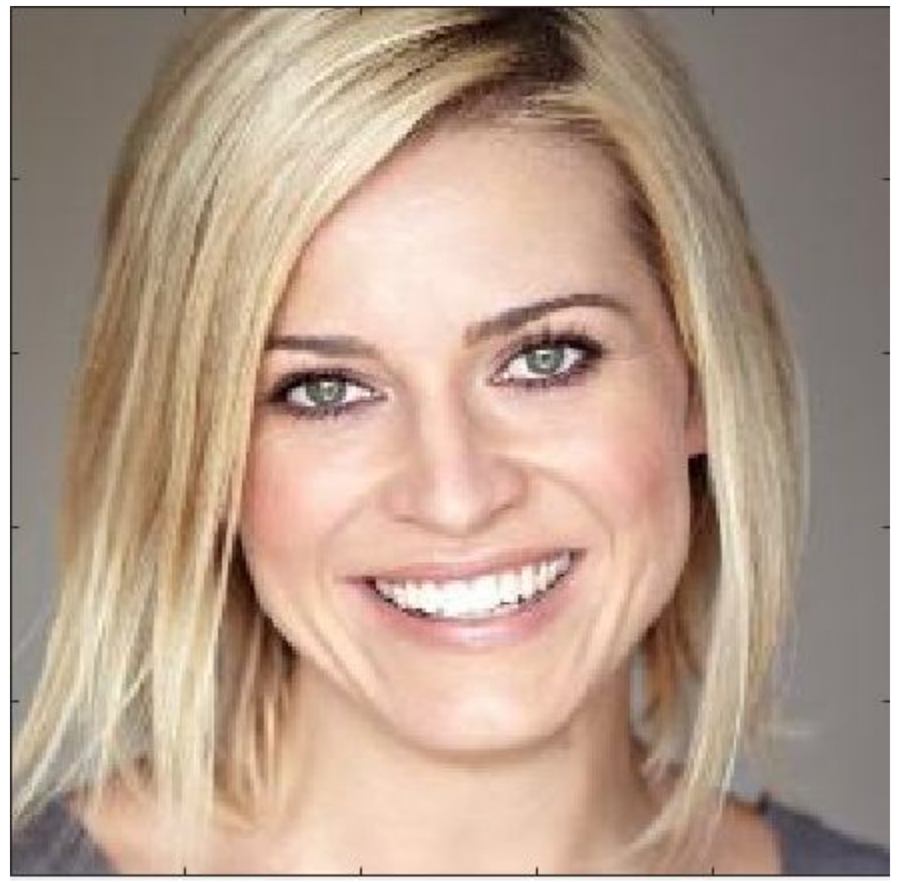
Face Recognition in our daily life



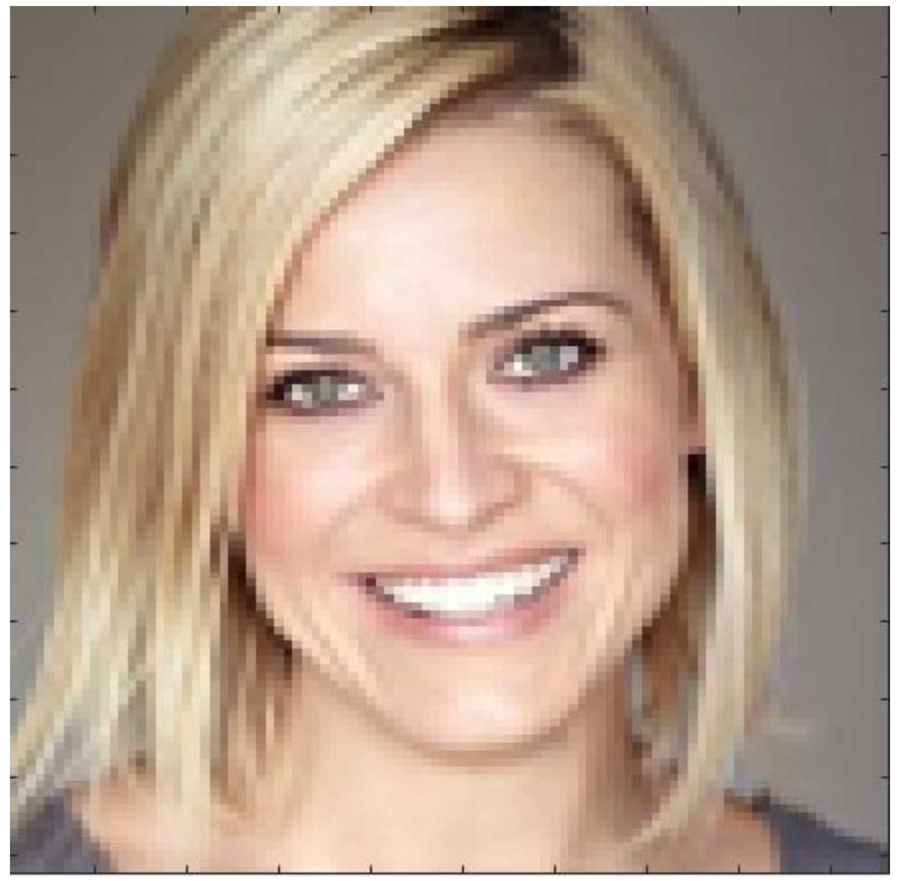
Payments

criminal identification

Background

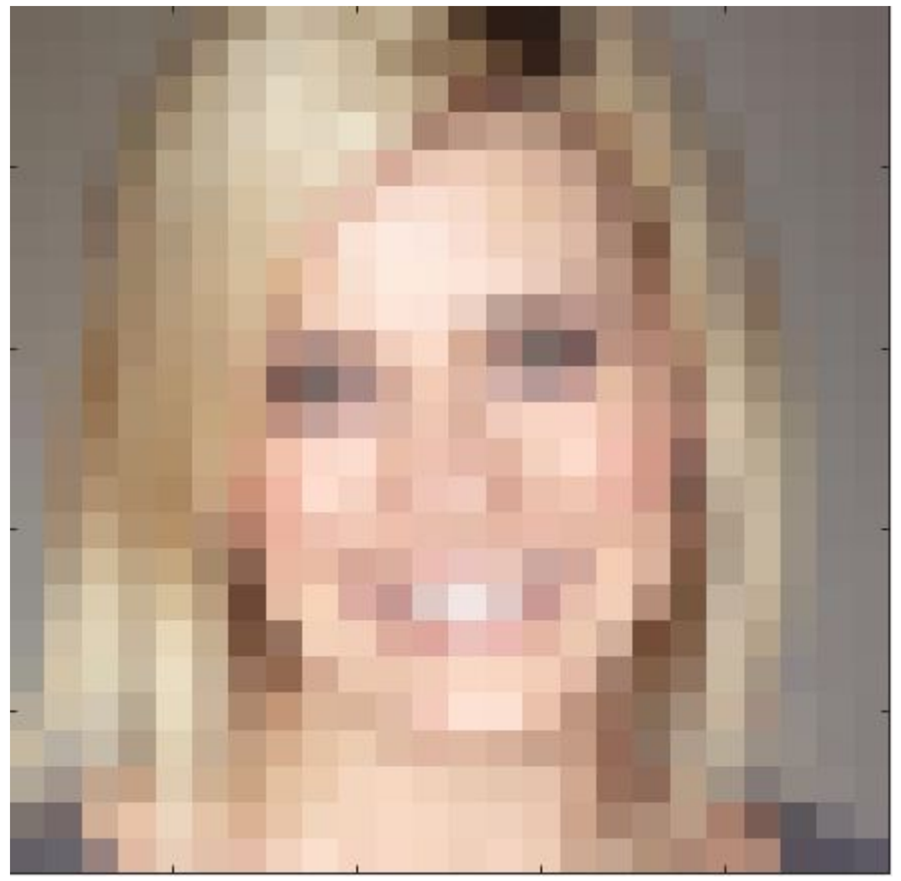


250x250

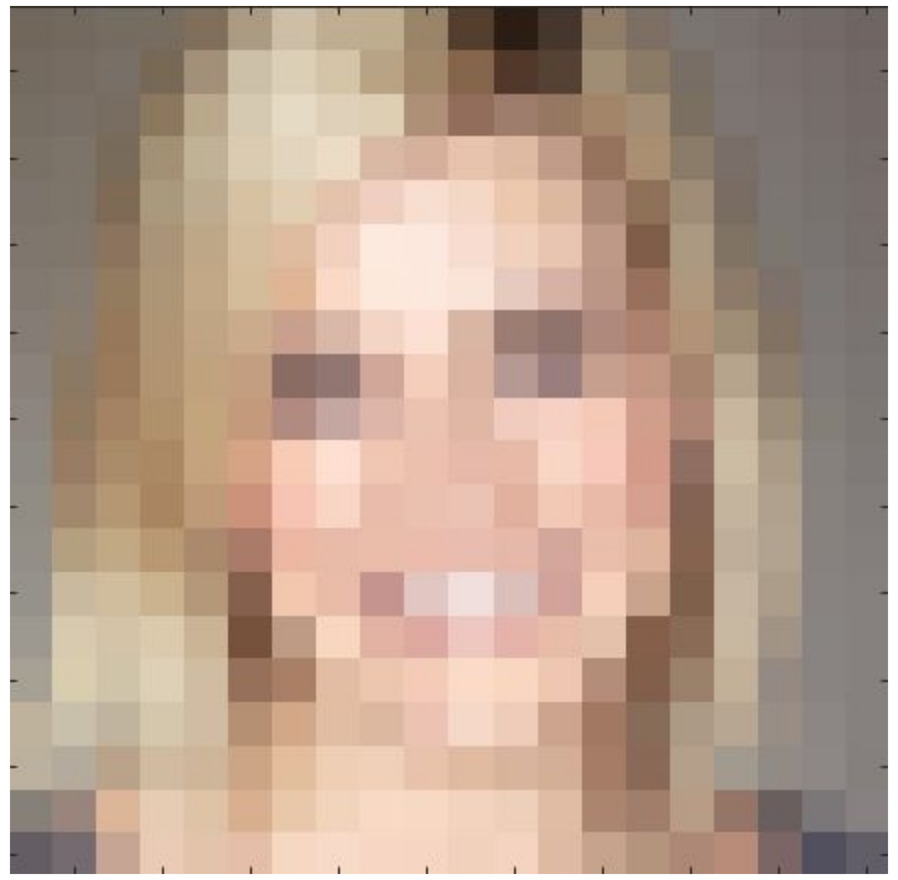


112x112

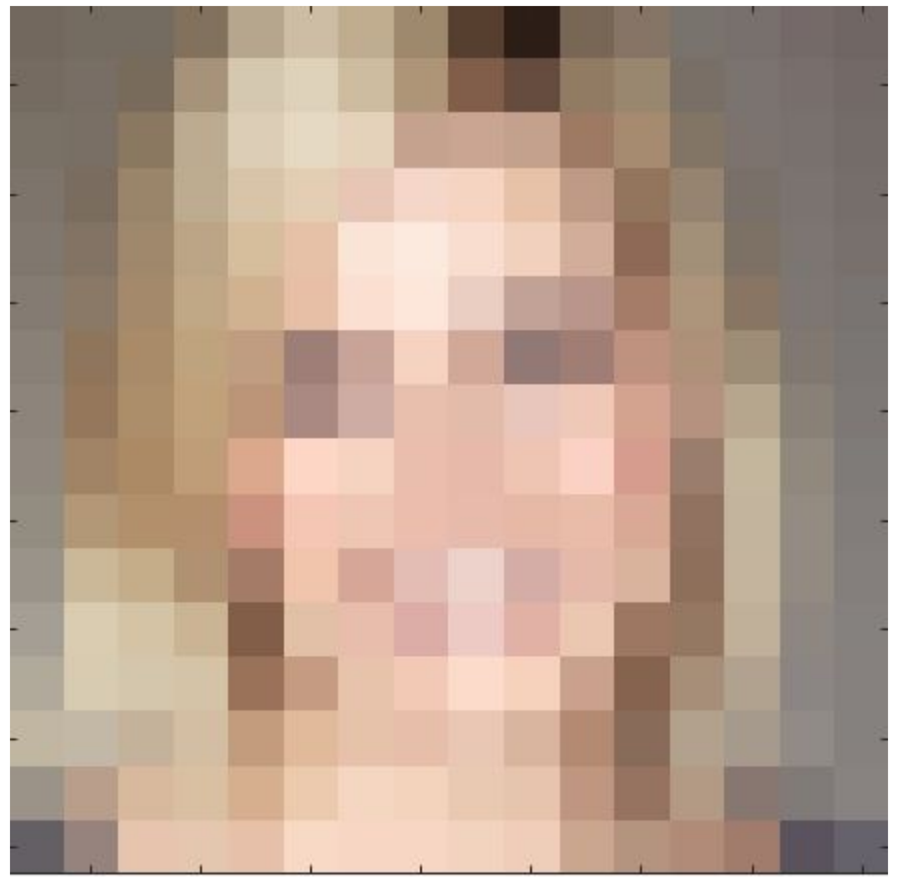
36x36



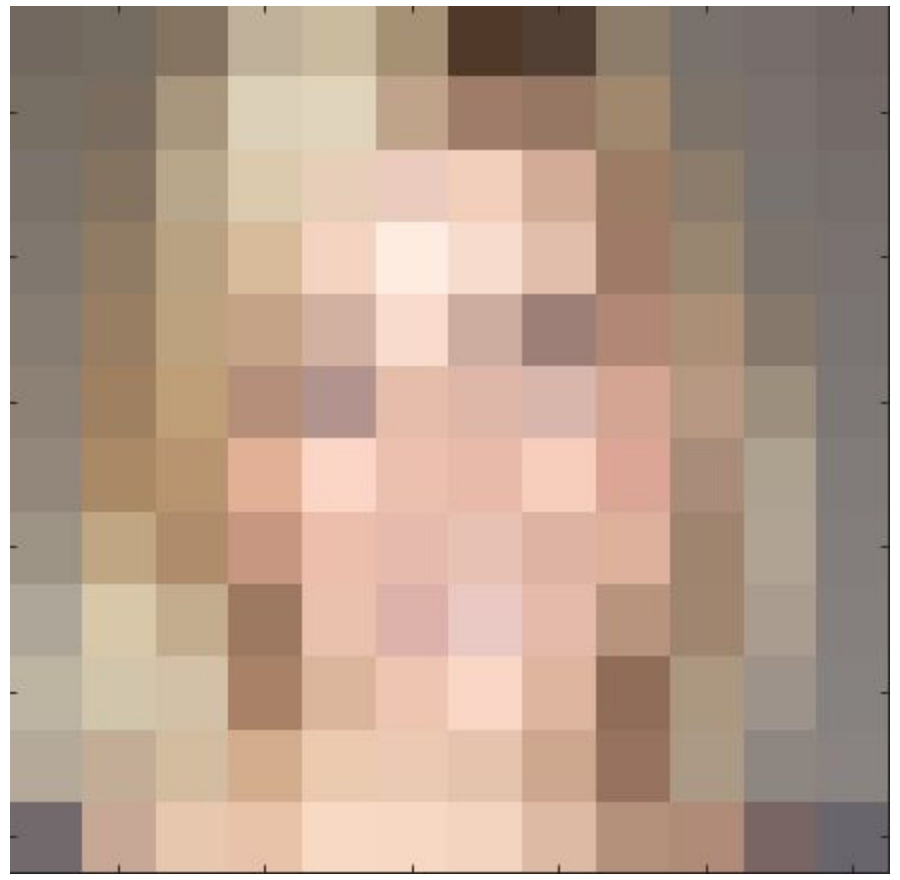
24x24



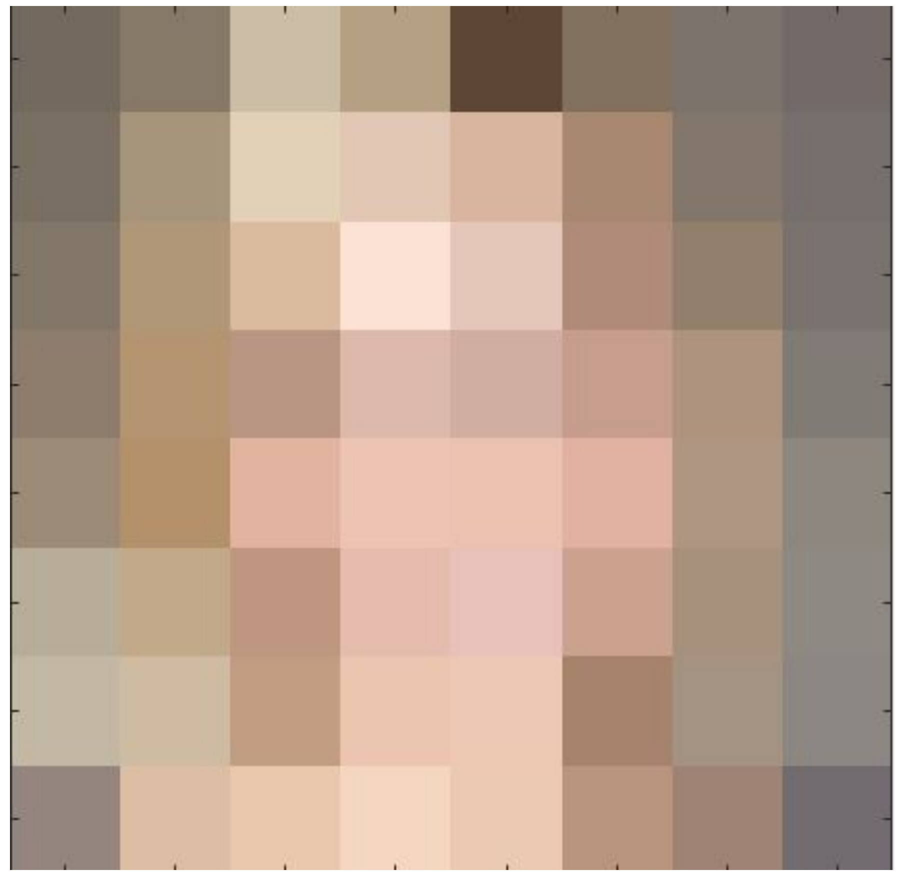
20x20



16x16



12x12



8x8

Background

Surveillance Camera

Objective

*Develop effective and robust face recognition algorithms that can achieve satisfactory performance in **low-resolution condition.***

Achievement

- Two algorithms that apply competently different techniques have been implemented.
  - The multidimensional scaling method for matching LR image
  - The deep-learning-based method with feature loss
- Experiments have been conducted around the built algorithms

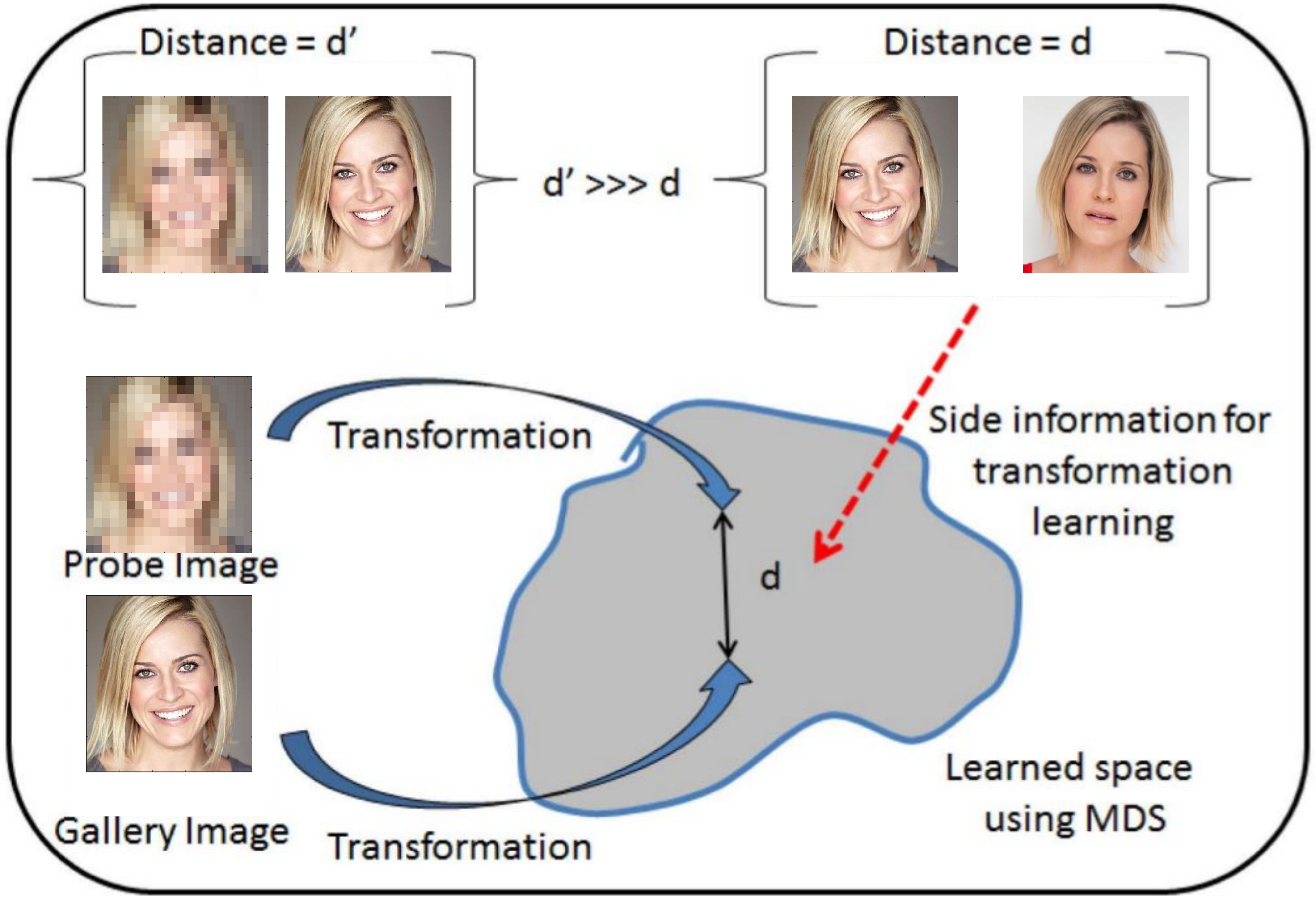
Conventional Method

Methodology

Multidimensional scaling (MDS) for Matching Low-resolution Images [1]

Key Ideas

- HR images and LR images are with different resolution, cannot be compared directly.
- Transform the HR image and LR image to a **common space** to minimize their distance.
- The result of LR images is close to the result of HR images.

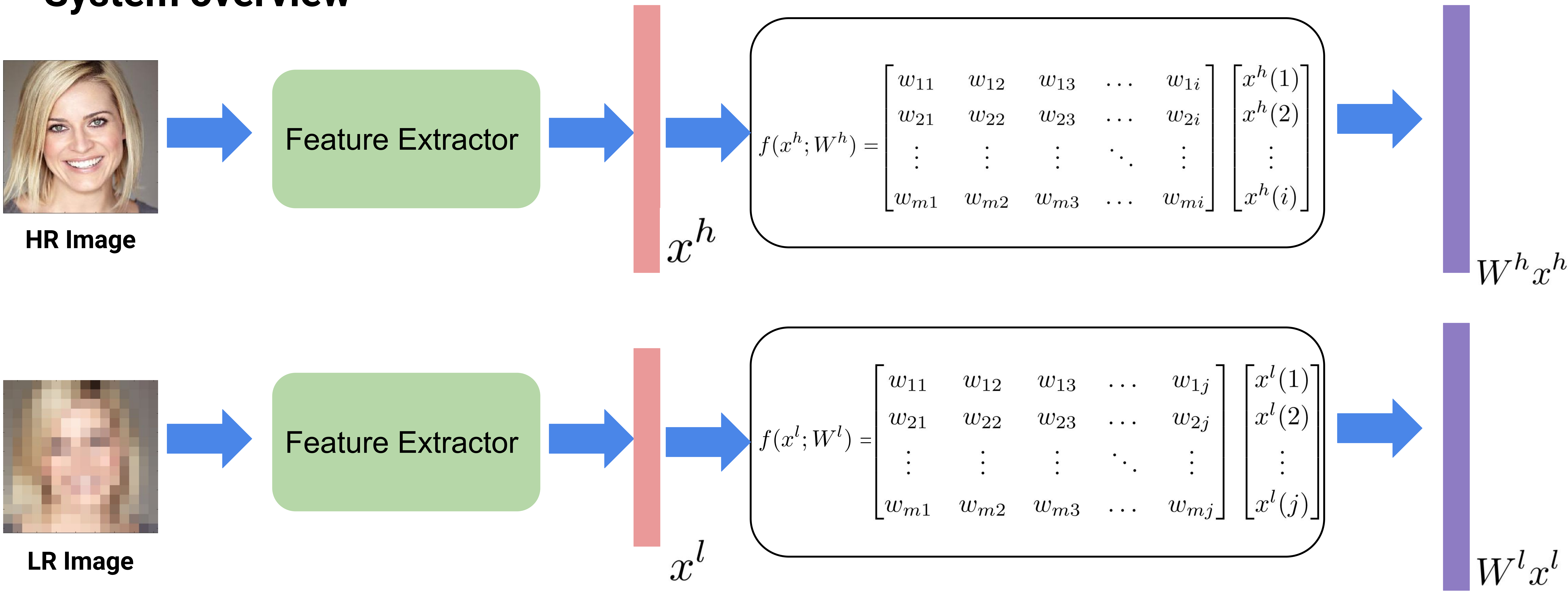


[1] Biswas, S., Bowyer, K. W., & Flynn, P. J. (2012). Multidimensional scaling for matching low-resolution face images. *IEEE transactions on pattern analysis and machine intelligence*, *34*(10), 2019-2030.

Methodology

Multidimensional scaling for Matching Low-resolution Images

System overview



PCA / LBP / SIFT

Methodology

Multidimensional scaling for Matching Low-resolution Images

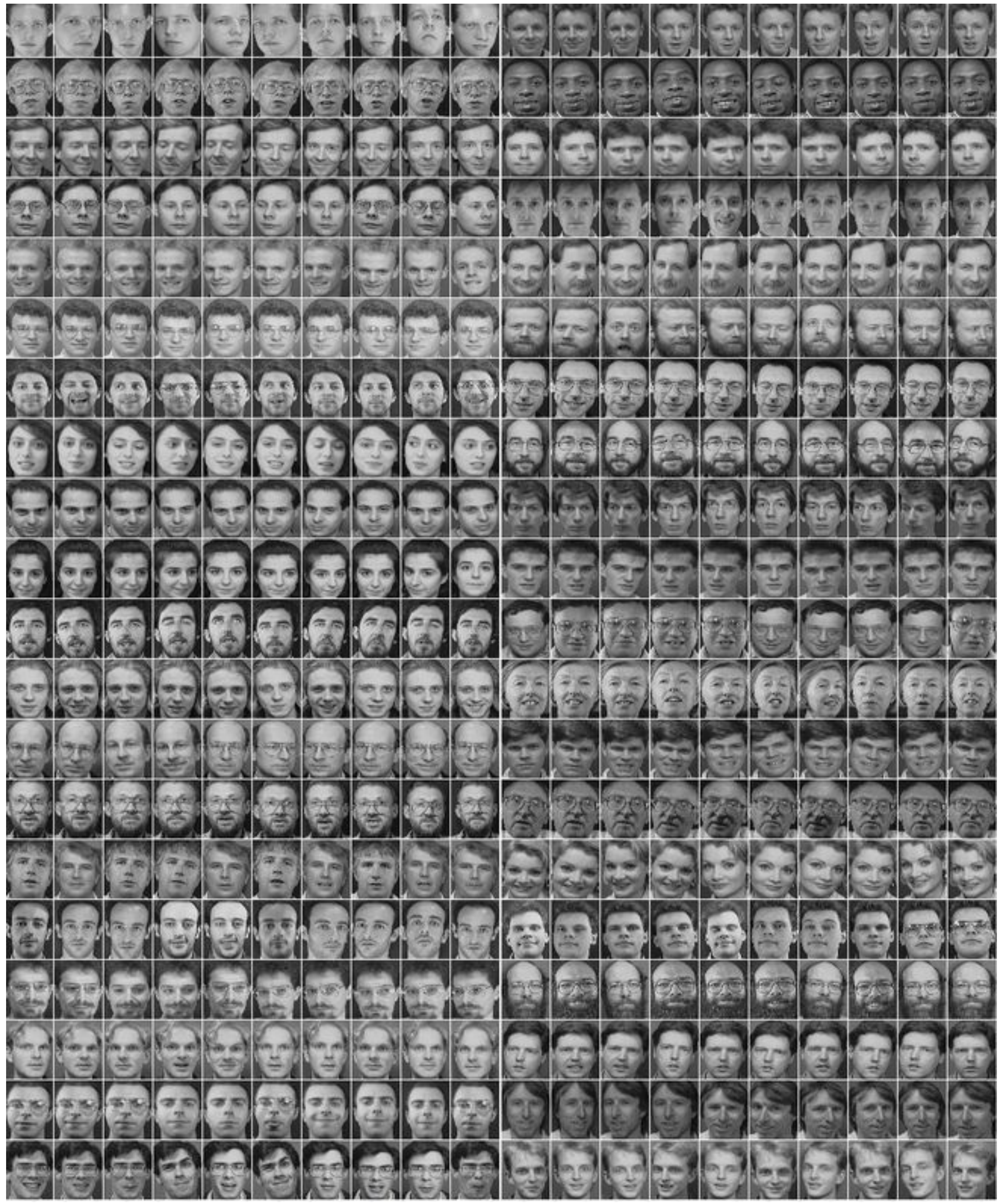
Objective function

$$J\left(W^{l}, W^{h}\right)=\sum_{i=1}^{N} \sum_{j=1}^{N}\left(\left|\left(W^{l}\right) x_{i}^{l}-\left(W^{h}\right) x_{i}^{h}\right|-d_{i j}^{h}\right)^{2}$$

Minimize the $\boldsymbol{J}(\boldsymbol{W}^{\boldsymbol{l}}, \boldsymbol{W}^{\boldsymbol{h}})$ by the **iterative majorization algorithm**

Experiment

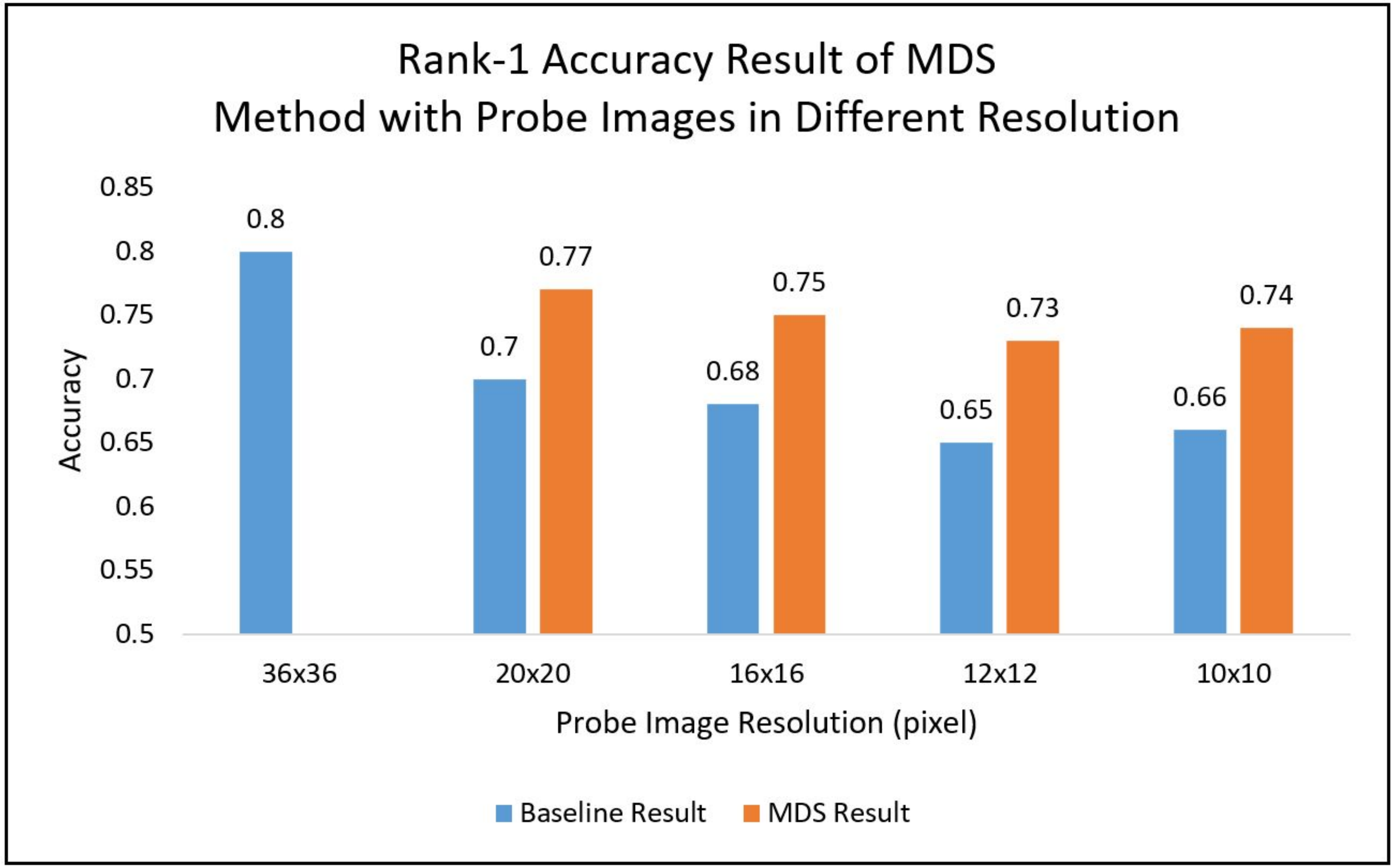
- ORL database: 400 face images of 40 subjects.
- For each subject: 2 for training, 8 for testing
- Feature: Principal Component Analysis (PCA)
- Baseline method : Eigenface for Recognition [2]



[2] Matthew Turk and Alex Pentland. Eigenfaces for recognition. Journal of cognitive neuroscience,3(1):71–86, 1991.

Conventional Method

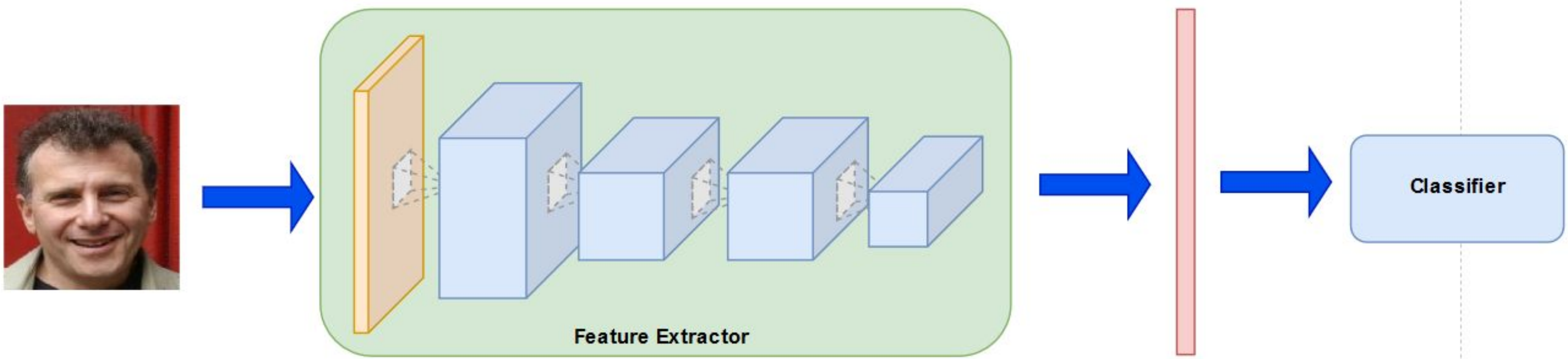
Result

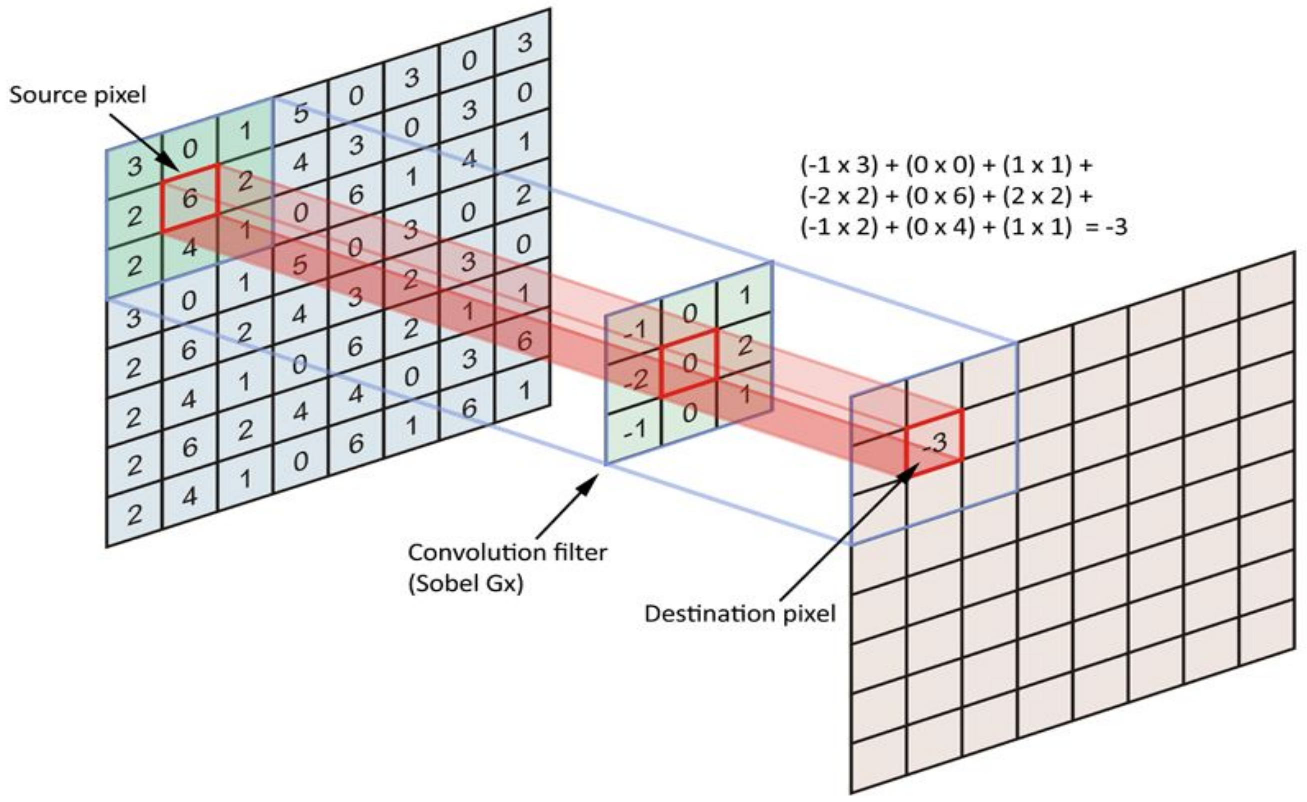


Deep-Learning Method

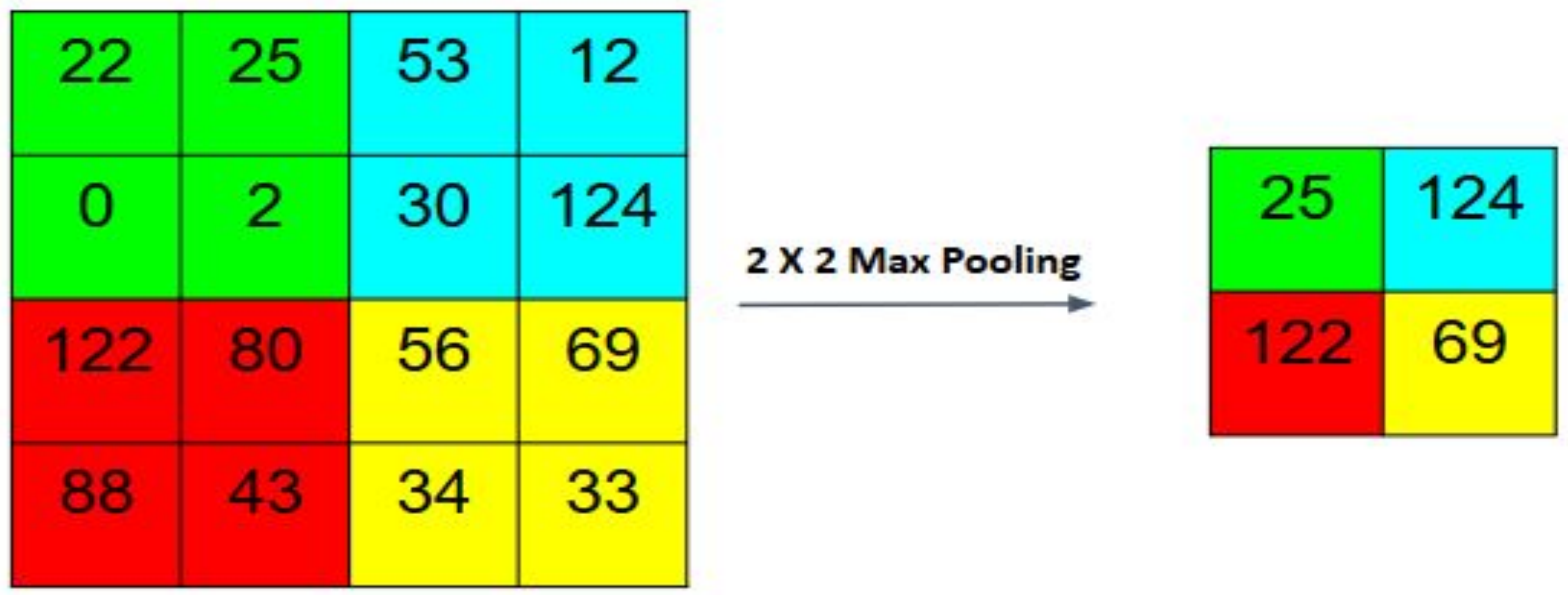
Methology

Convolutional Neural Network Model





Convolutional Layer

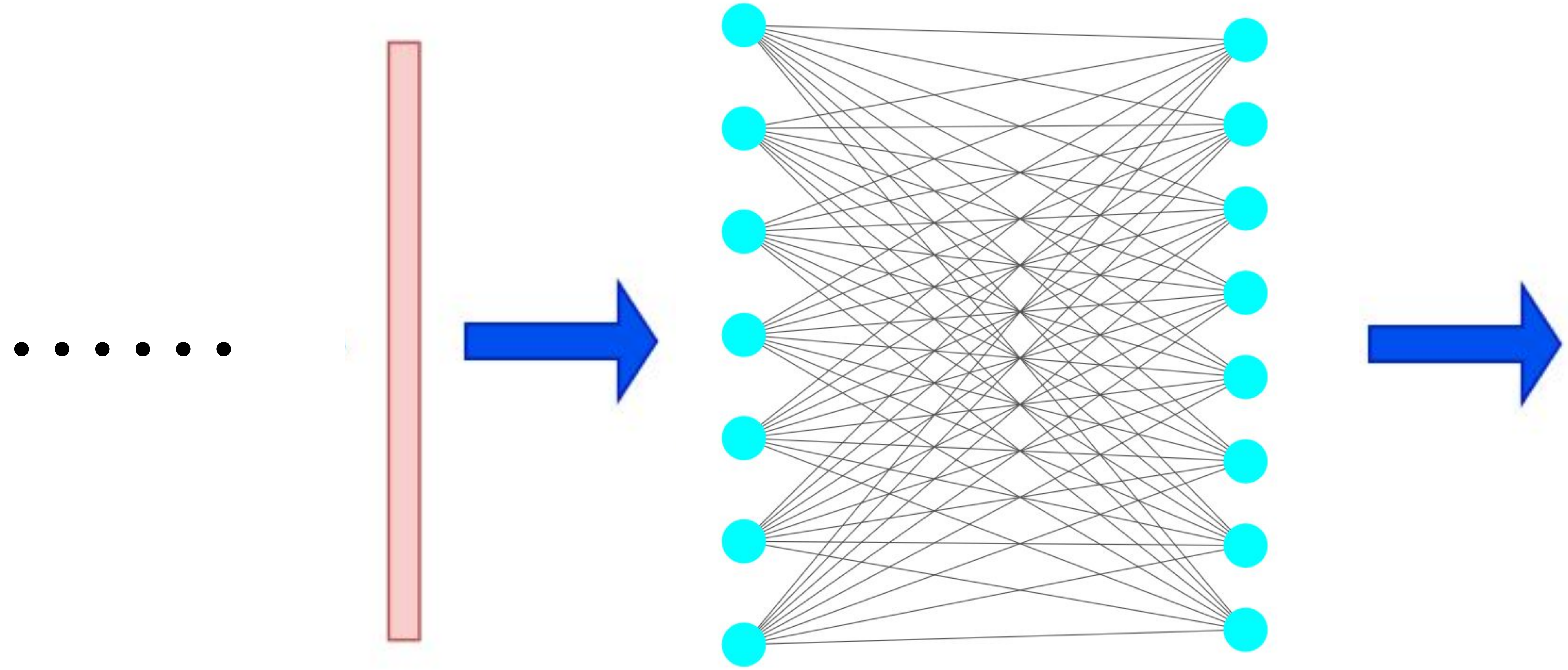


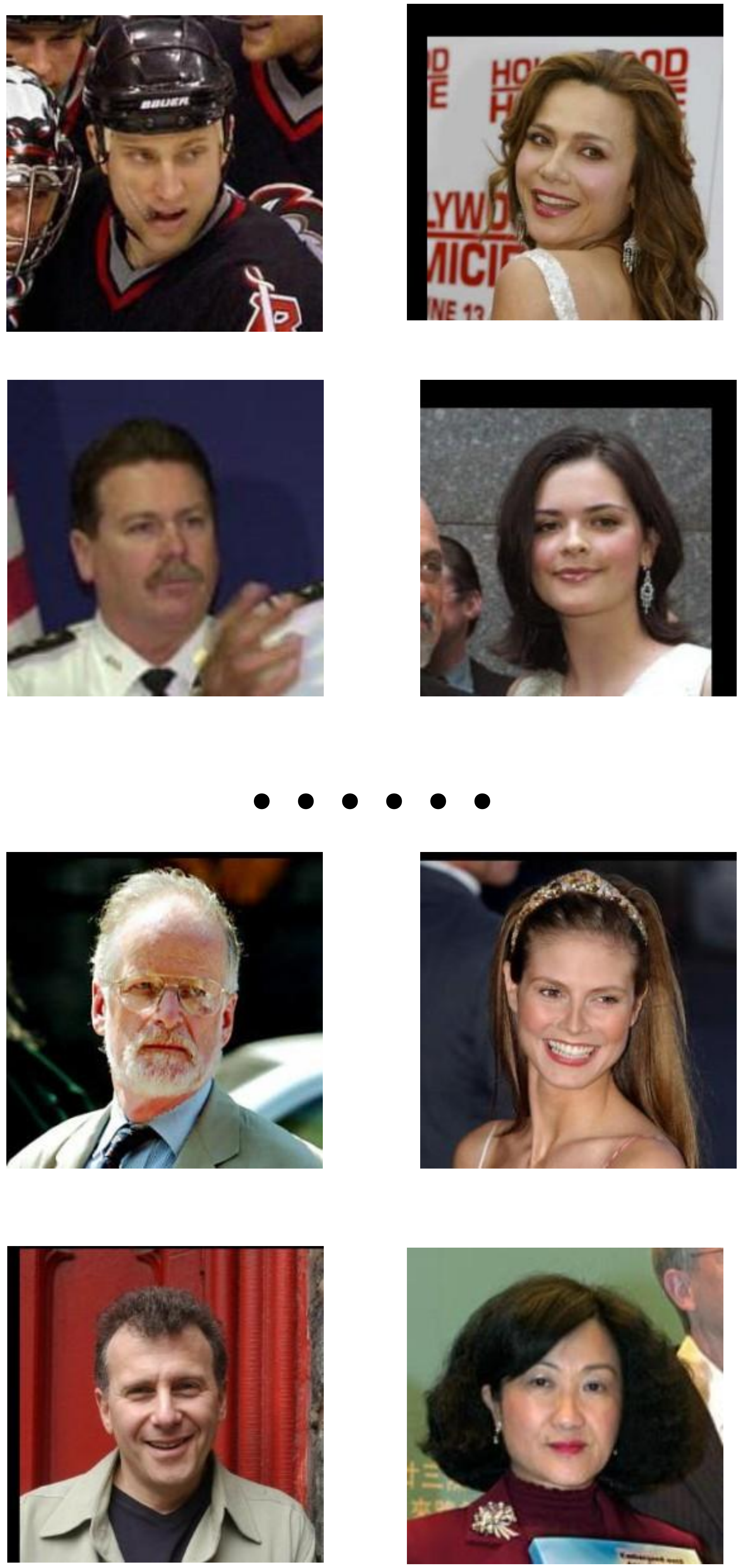
Pooling Layer

Deep-Learning Method

Methology

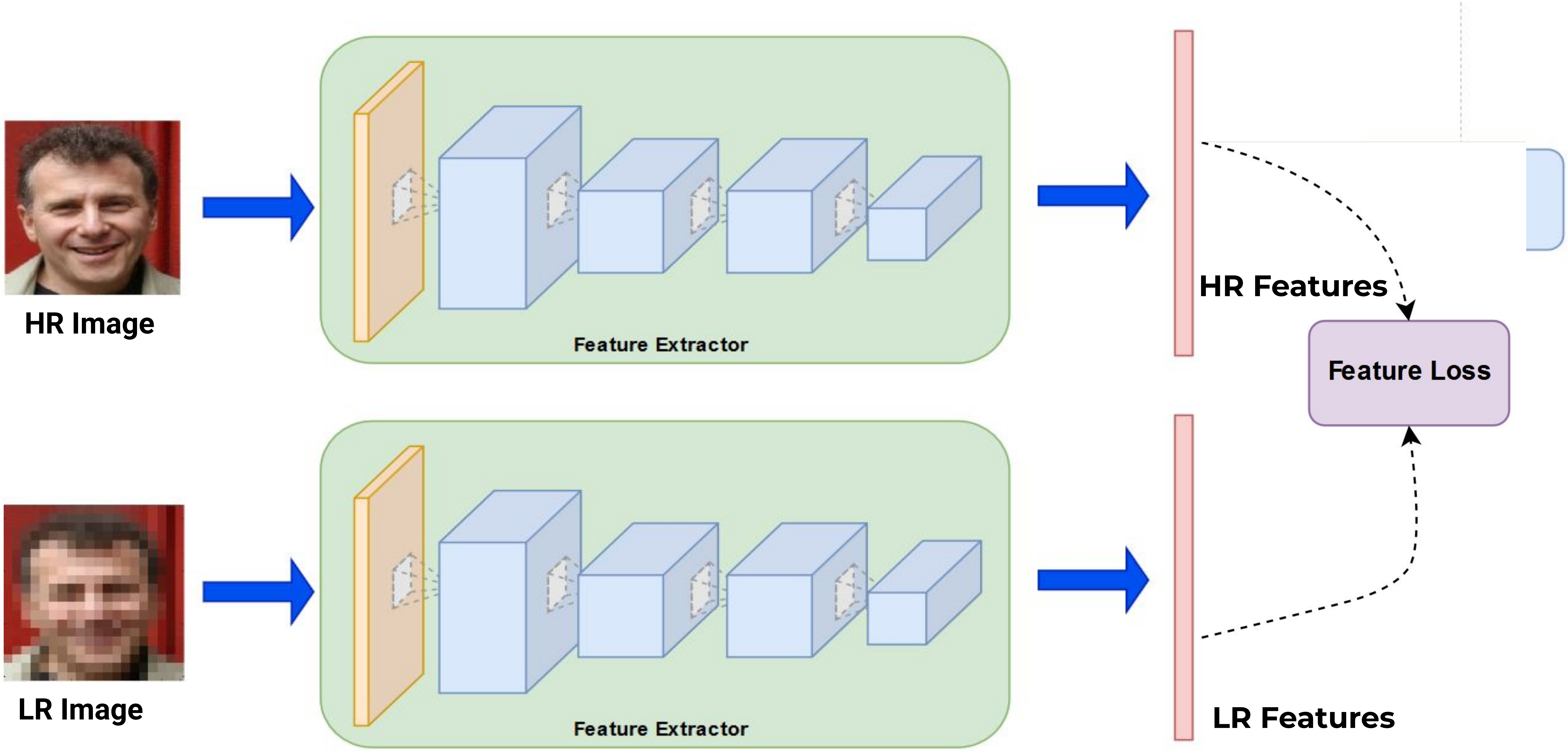
Convolutional Neural Network Model





Deep-learning Method

LR-CNN-Based Model Pipeline



Methodology

Deep-Learning Method

Loss function

- Mean Square Error

$$d(\mathbf{p},\mathbf{q}) = d(\mathbf{q},\mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

- L1 Distance

$$d_1(\mathbf{p},\mathbf{q}) = \|\mathbf{p} - \mathbf{q}\|_1 = \sum_{i=1}^{n} |p_i - q_i|$$

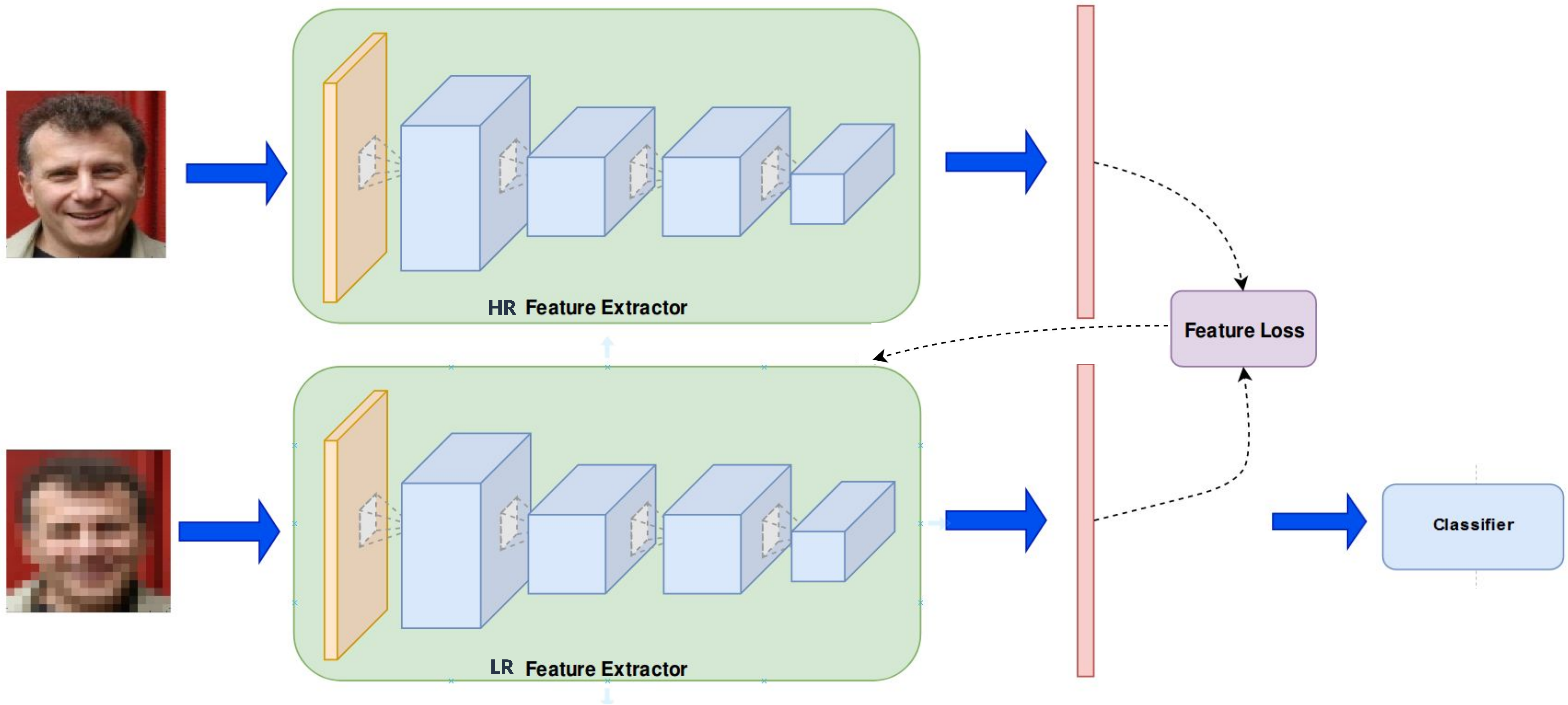
- Cosine Similarity

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum\limits_{i=1}^{n} A_i B_i}{\sqrt{\sum\limits_{i=1}^{n} A_i^2}\sqrt{\sum\limits_{i=1}^{n} B_i^2}}$$

Methodology

Deep-Learning Method

LR-CNN-Based Model Pipeline



Experiment

Database

Training set - Casia-webface [3]:
500k face images of 10k subjects

Testing set - LFW database [4]:
3k matching face images pairs
3k unmatching face images pairs

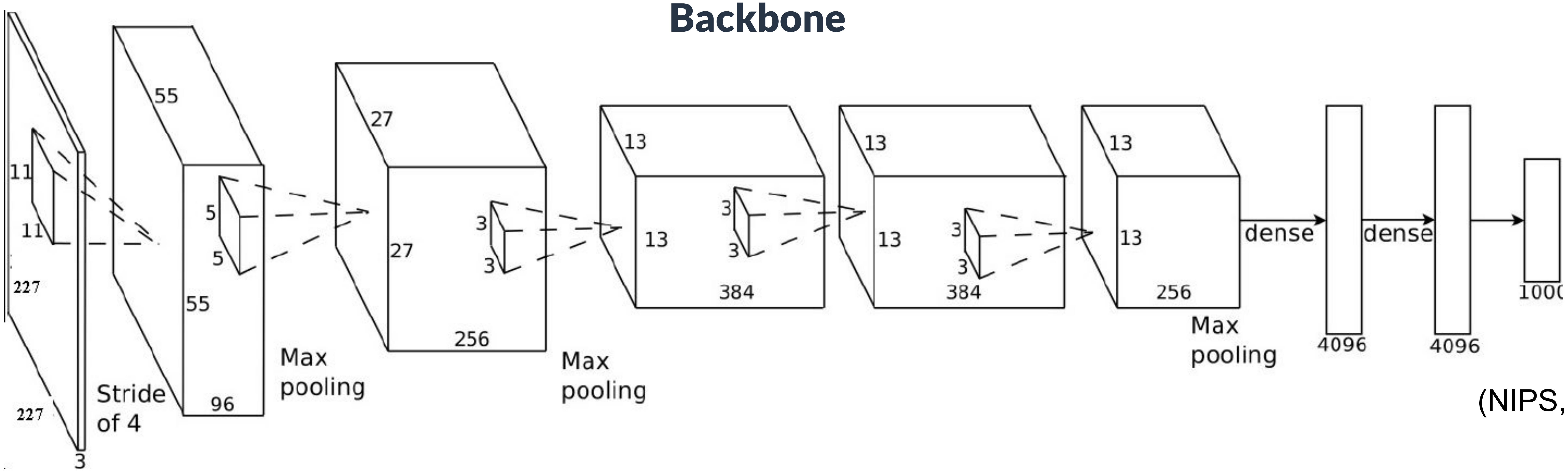
[3] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. Learning face representation from scratch.arXiv preprint arXiv:1411.7923, 2014.
[4] Gary B. Huang Erik Learned-Miller. Labeled faces in the wild: Updates and new reporting procedures. Technical Report UM-CS-2014-003, University of Massachusetts, Amherst, May 2014.

Deep-Learning Method

Experiment

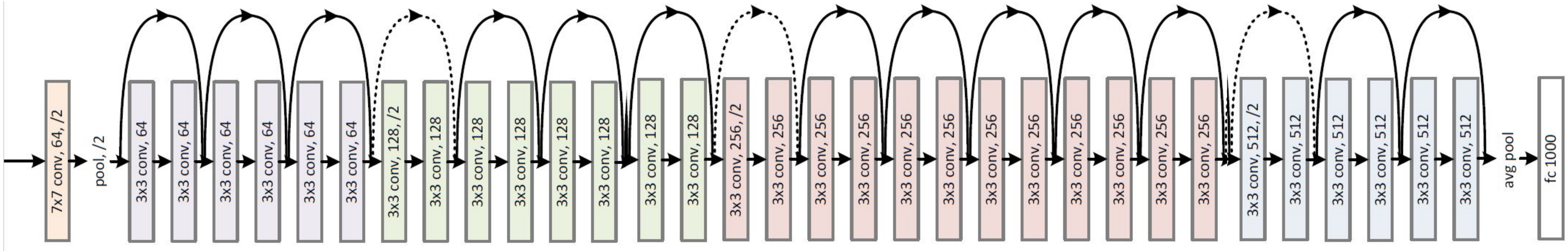
Alexnet

Backbone



(NIPS, 2012)

Resnet

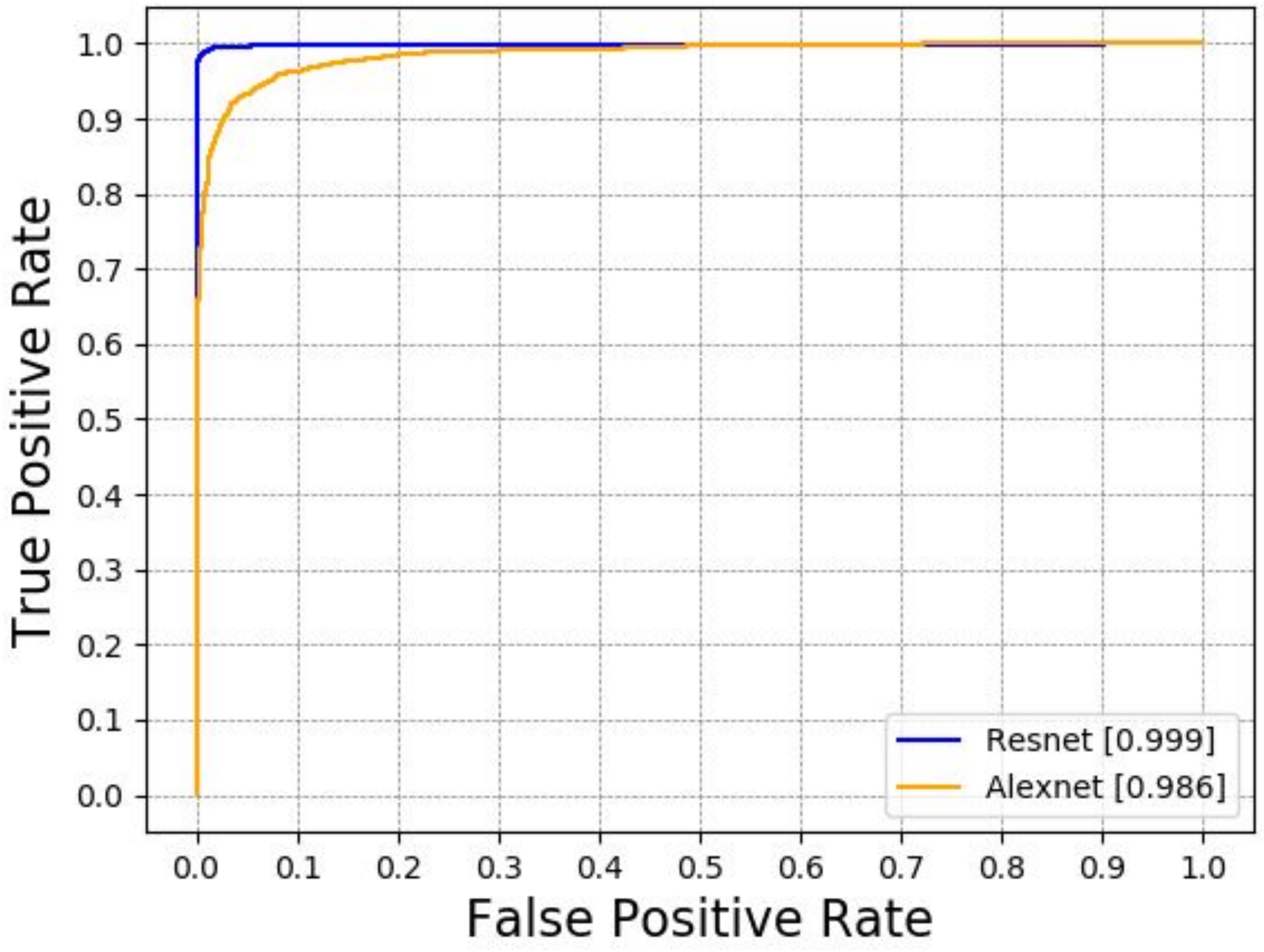


(CVPR, 2016)

Deep-Learning Method

Backbone Result

In high Resolution Condition



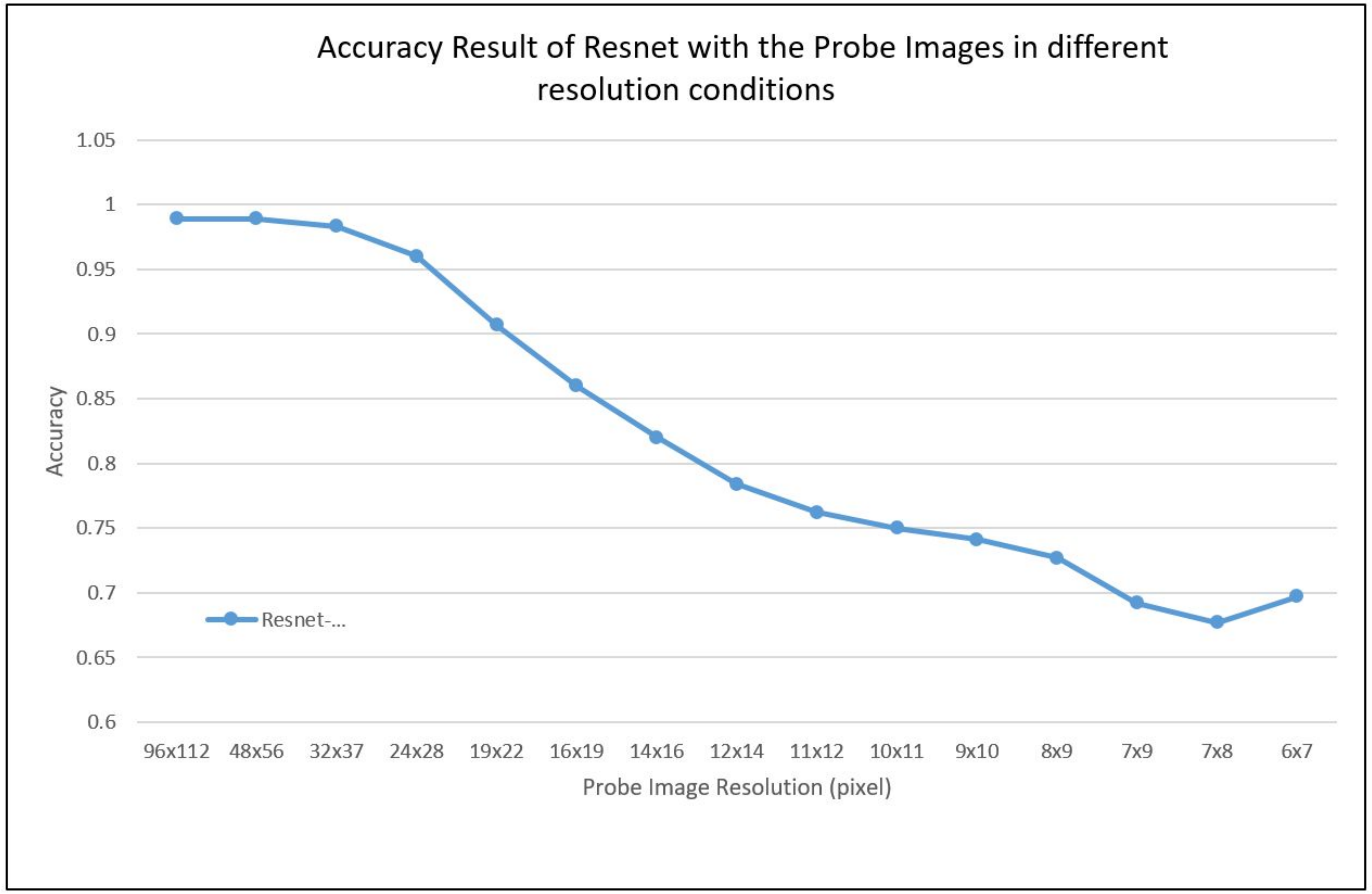
AUC of Resnet and Alexnet (HR Images)

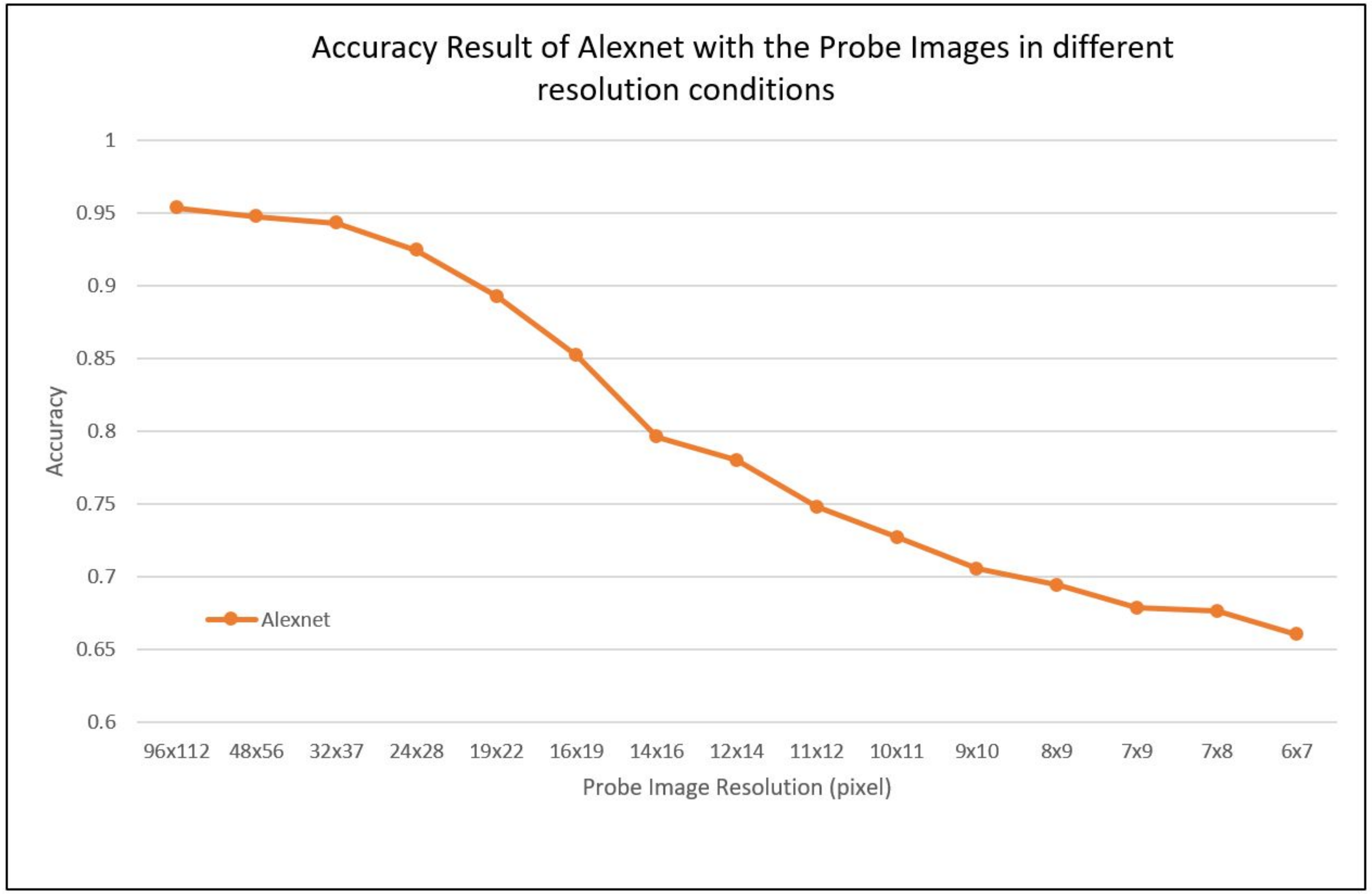
	Accuracy	AUC
Alexnet	0.953	0.986
Resnet	**0.989**	**0.999**

Deep-Learning Method

Backbone Result

In Low Resolution Condition



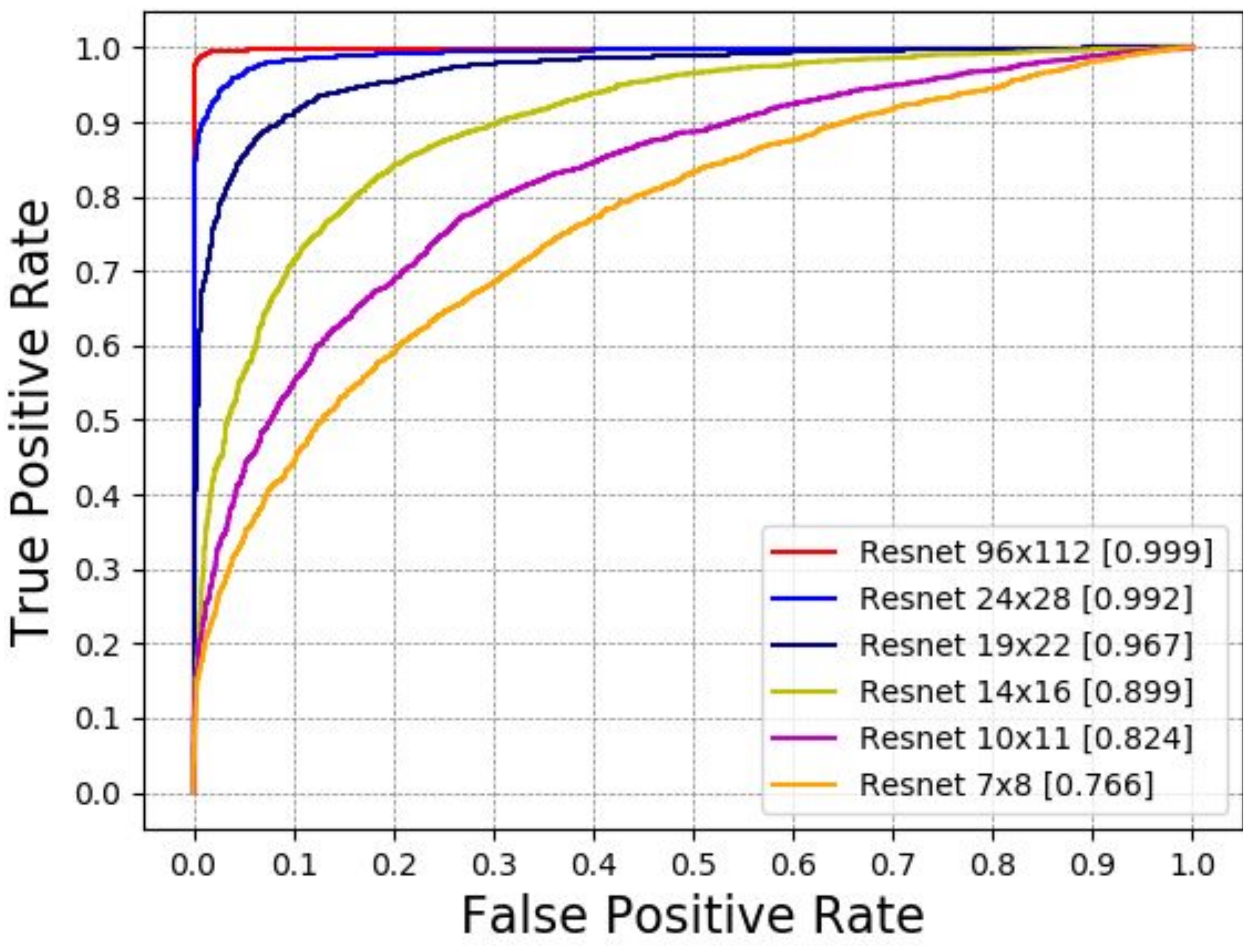


Accuracy of Resnet

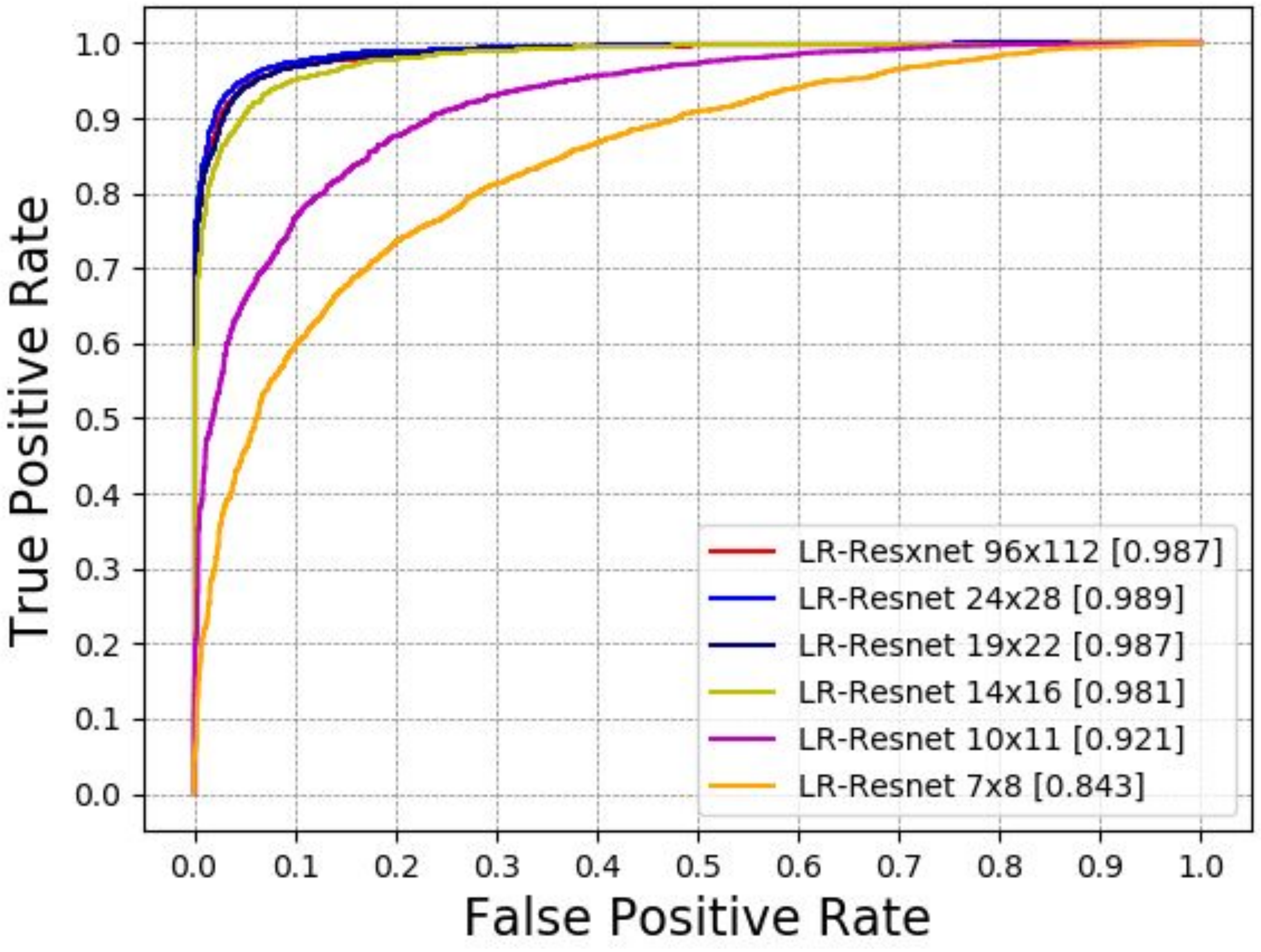
Accuracy of Alexnet

Deep-Learning Method

LR-Resnet Result



ROC Curve of Resnet

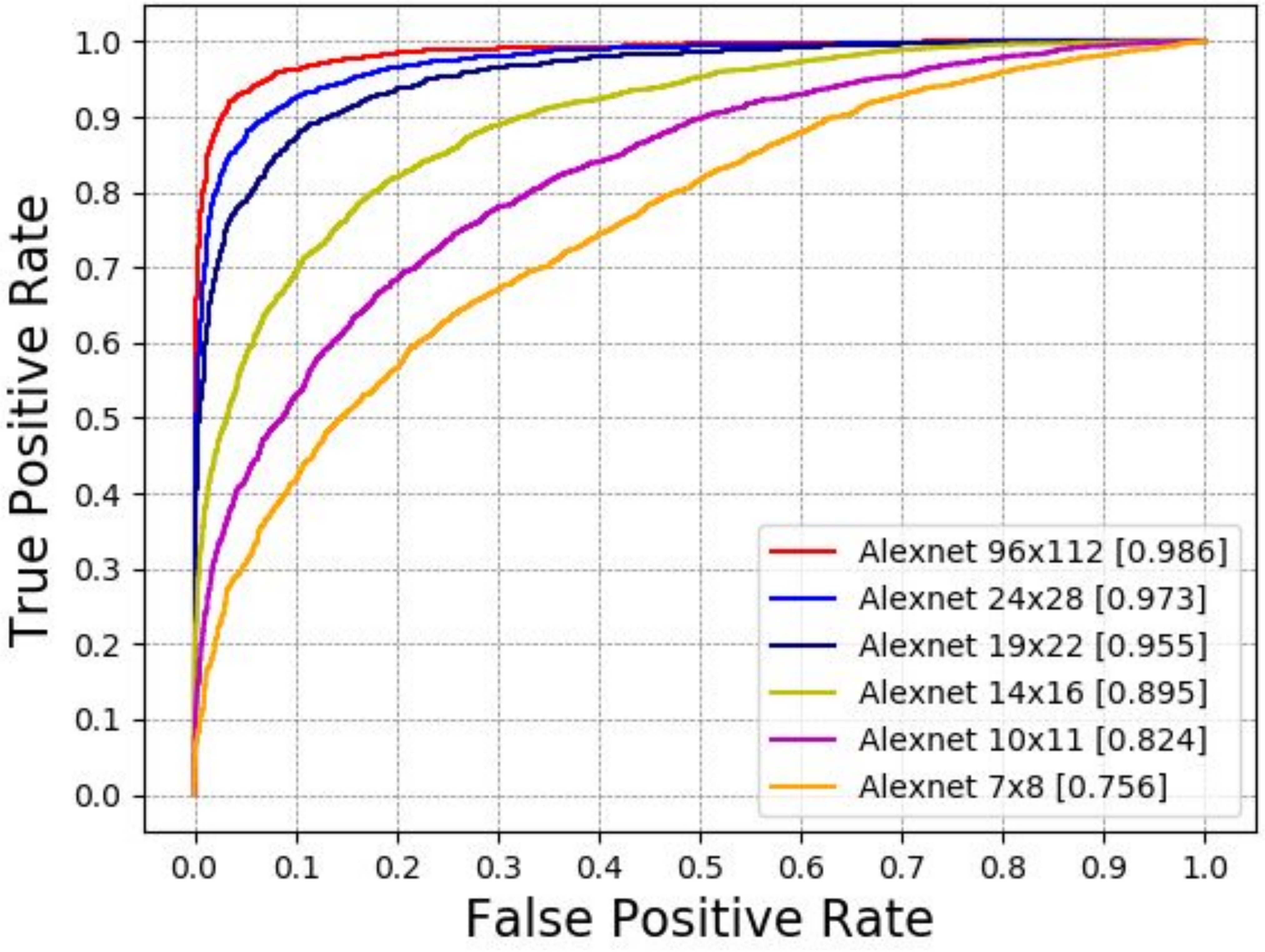


ROC Curve of LR-Resnet

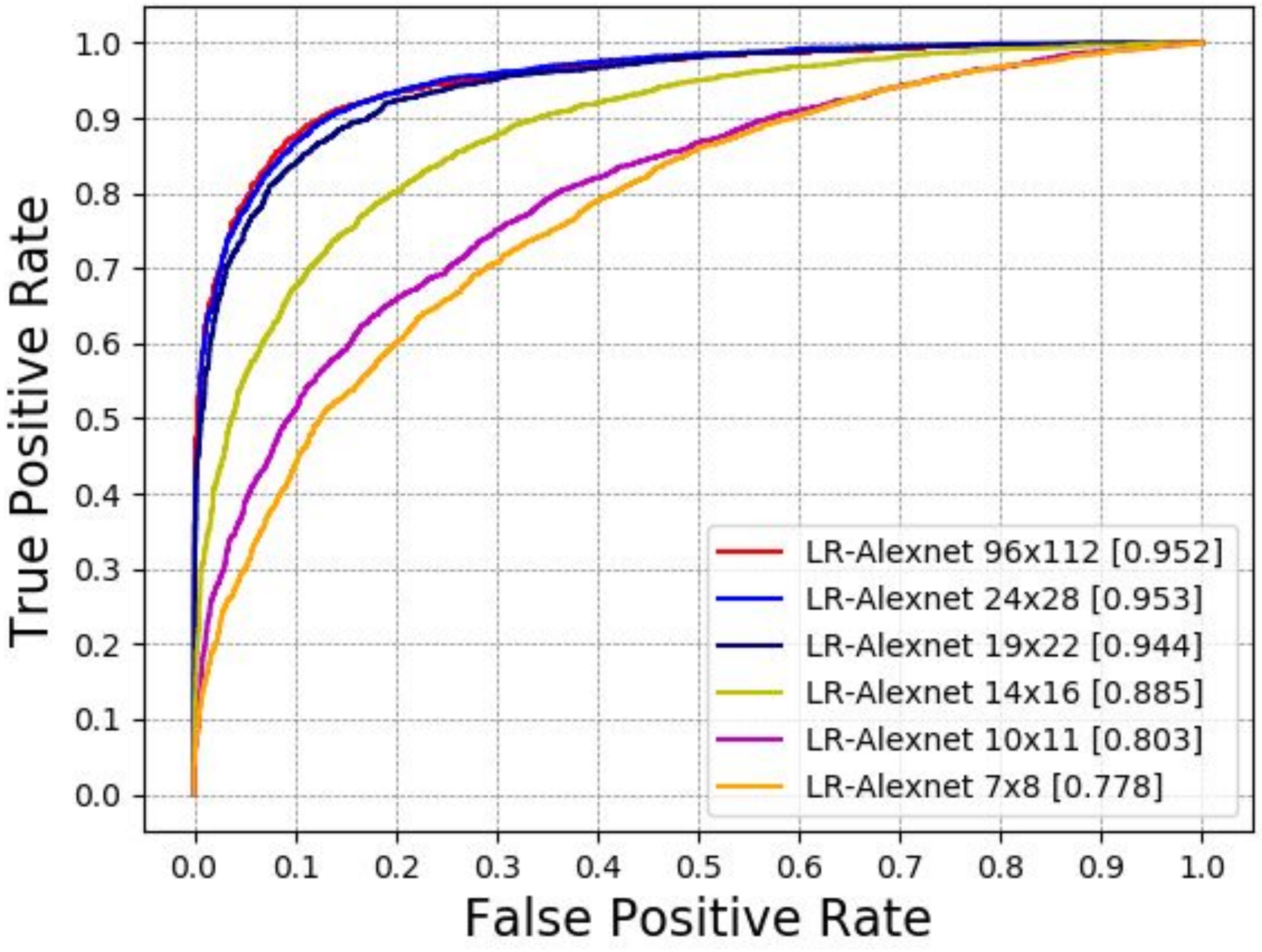
	Resnet	**LR-Resnet**
96x112	**0.989**	0.9462
24x28	0.9242	**0.9495**
19x22	0.8925	**0.9450**
14x16	0.7958	**0.9303**
10x11	0.7478	**0.8592**
7x8	0.6762	**0.7398**

Rank-1 Accuracy

LR-Alexnet Result



ROC of Alexnet



ROC of LR-Alexnet

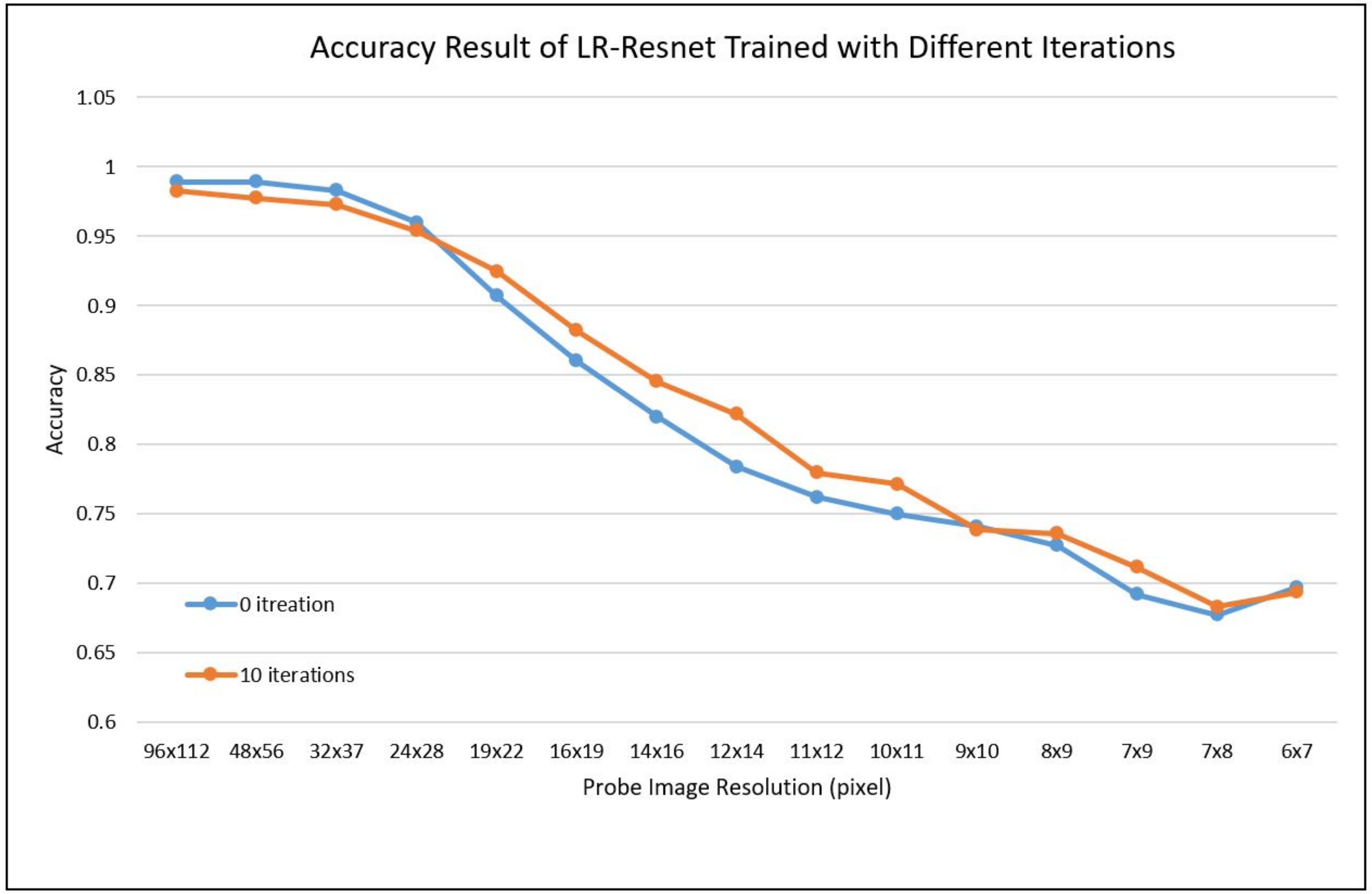
	Alexnet	**LR-Alexnet**
96x112	**0.9532**	0.8968
24x28	**0.9242**	0.9048
19x22	0.8925	**0.8988**
14x16	0.7958	**0.8582**
10x11	0.7478	**0.7808**
7x8	0.6762	**0.7205**

Rank-1 Accuracy

Deep-Learning Method

Interesting Finding

Training Interactions VS Accuracy



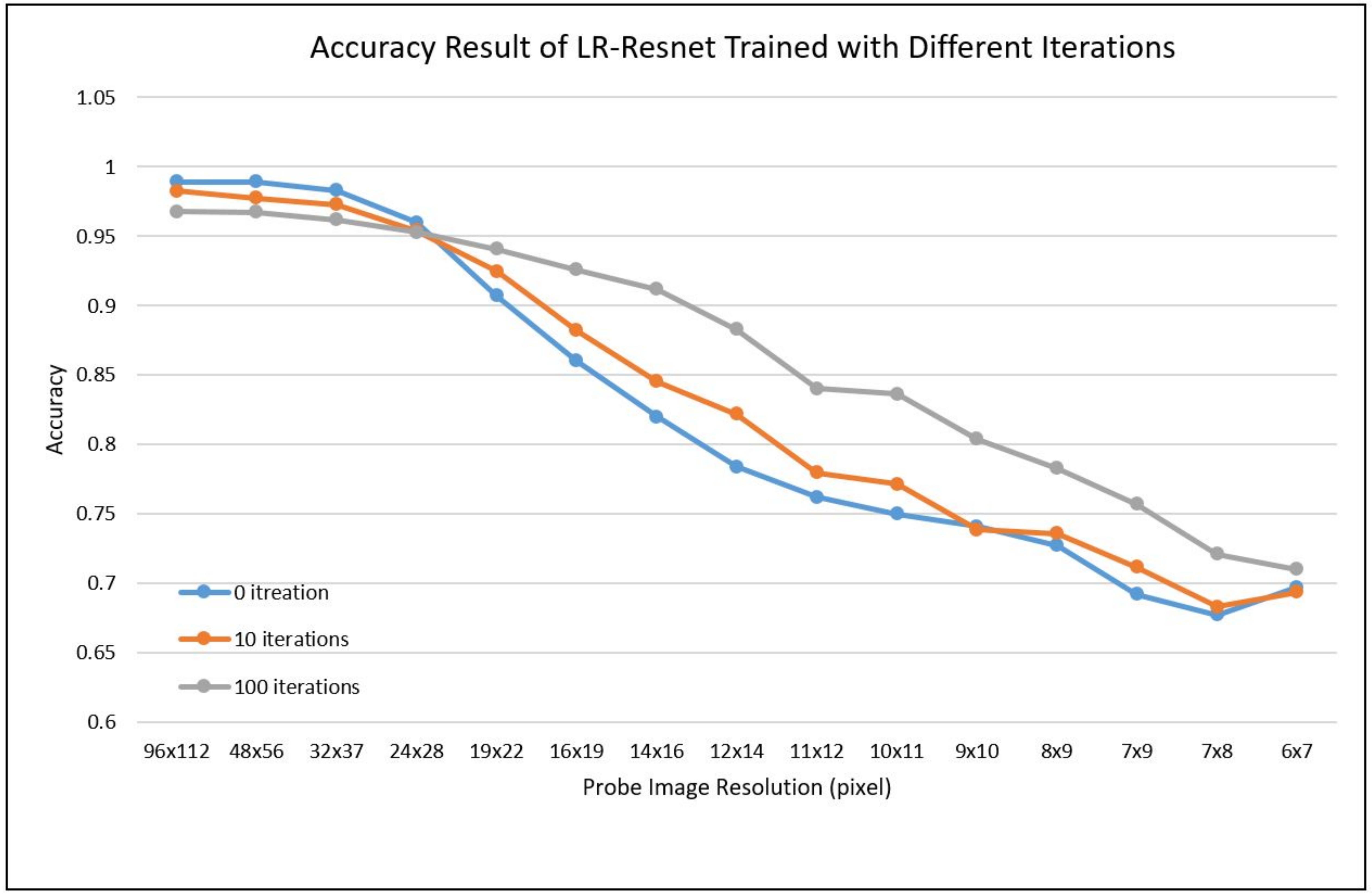
10 times

(LR-Resnet trained with 14x16 pixels LR images)

Deep-Learning Method

Interesting Finding

Training Interactions VS Accuracy



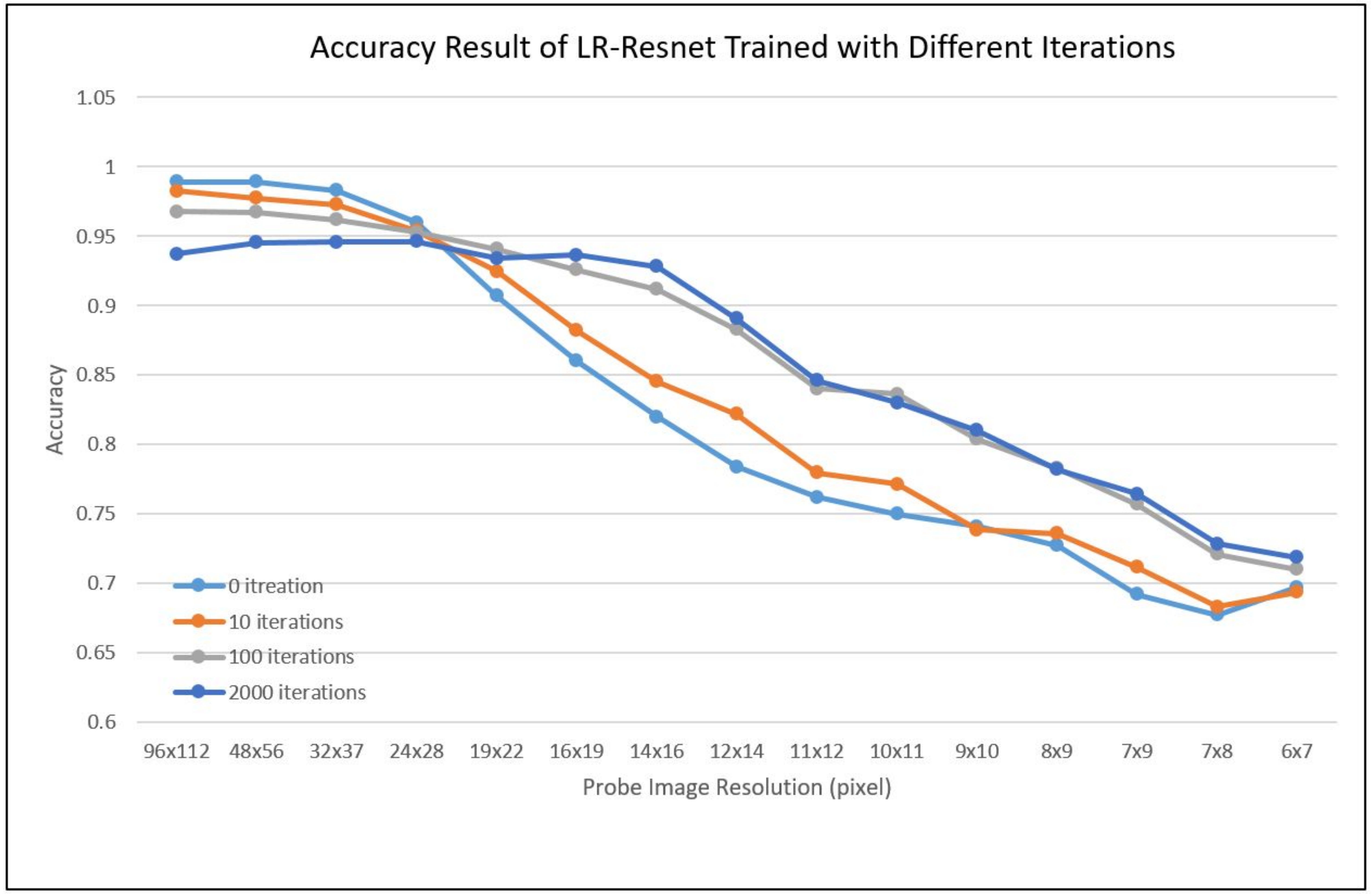
100 times

(LR-Resnet trained with 14x16 pixels LR images)

Deep-Learning Method

Interesting Finding

Training Interactions VS Accuracy



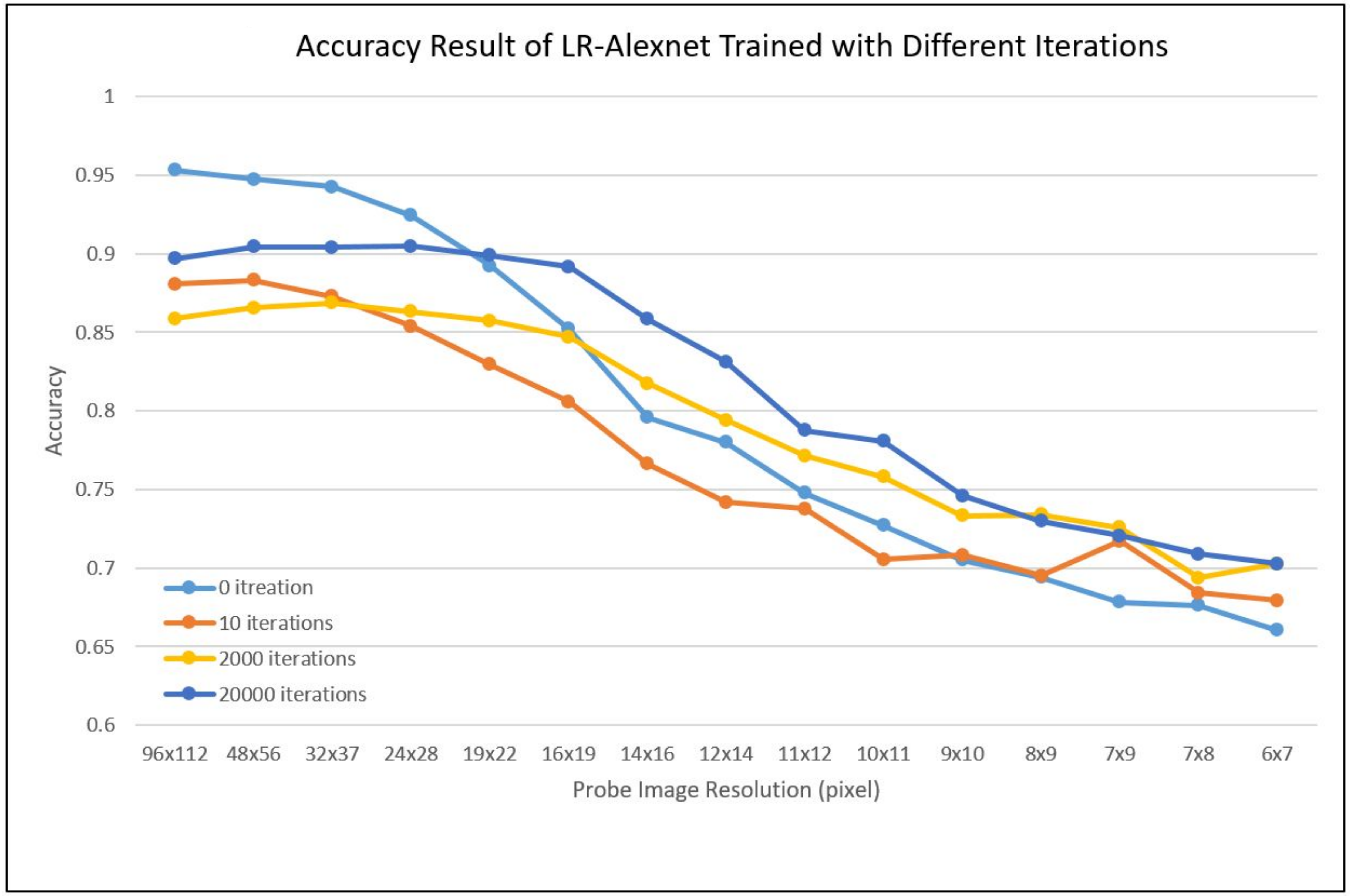
2000 times

(LR-Resnet trained with 14x16 pixels LR images)

Deep-Learning Method

Interesting Finding

Training Interactions VS Accuracy

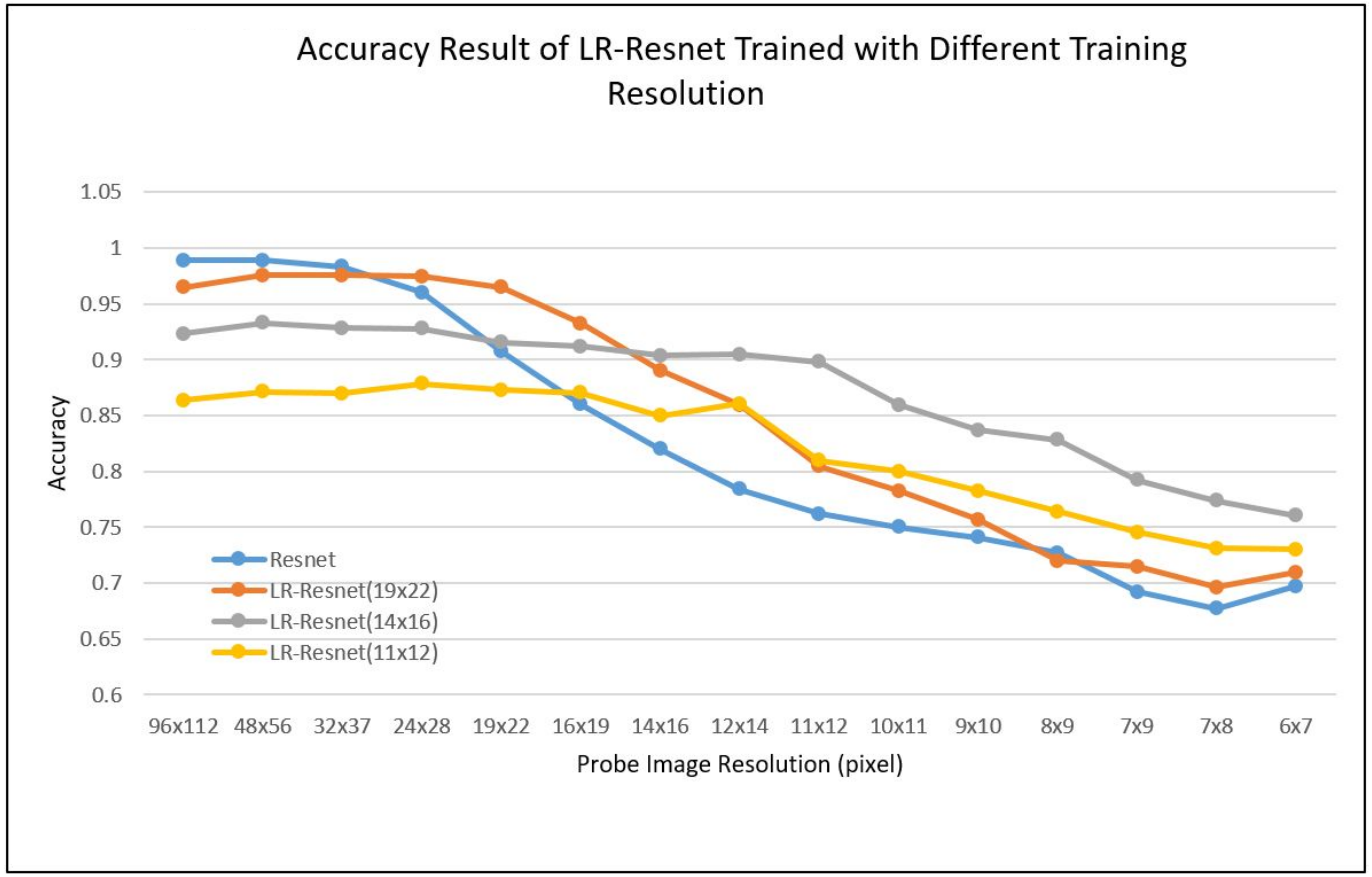


(LR-Alexnet trained with 11x12 pixels LR images)

Deep-Learning Method

Interesting Finding

Training Image Resolution VS Accuracy

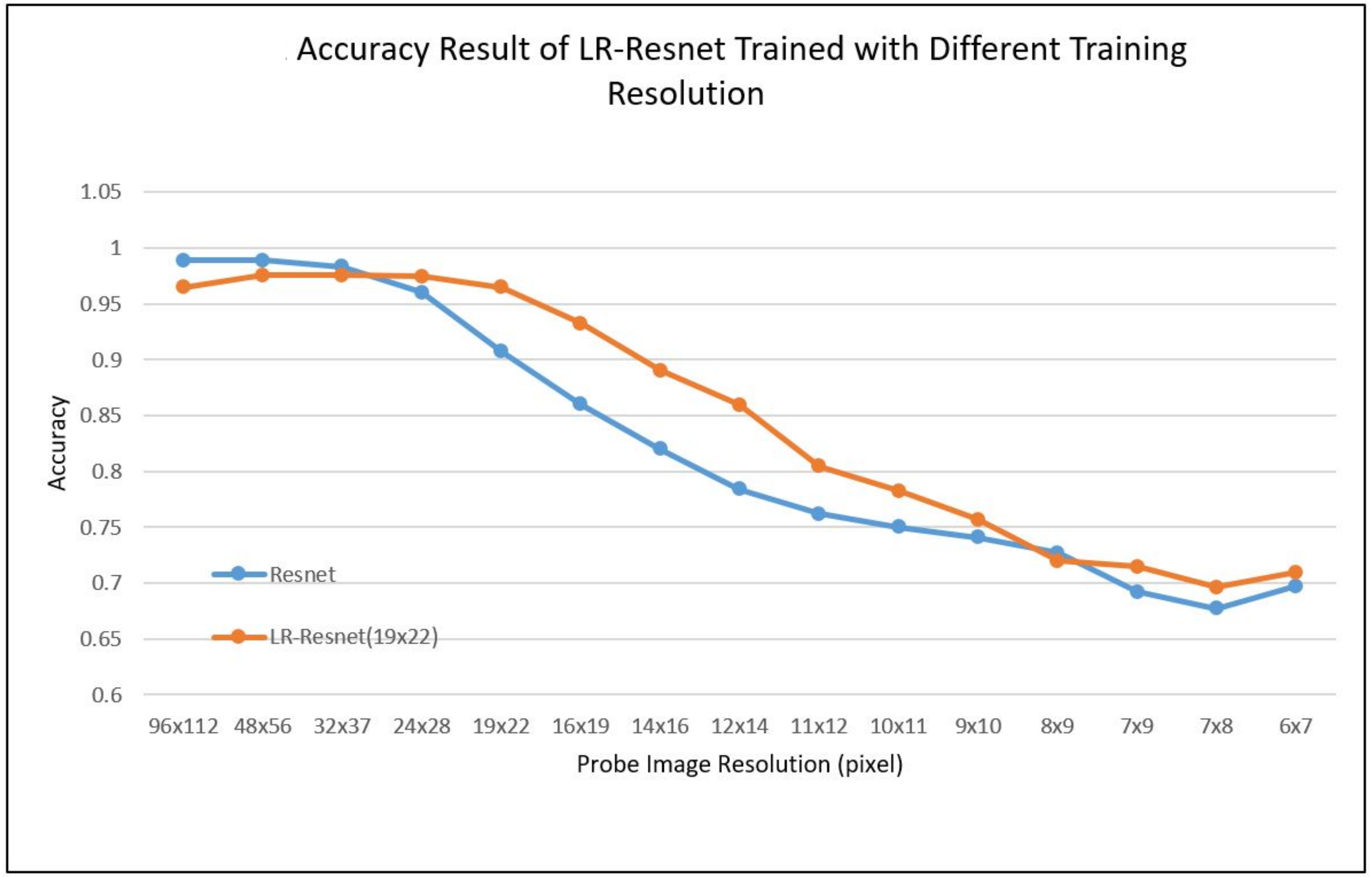


(LR-Resnet trained with 1000 iterations)

Deep-Learning Method

Interesting Finding

Training Image Resolution VS Accuracy



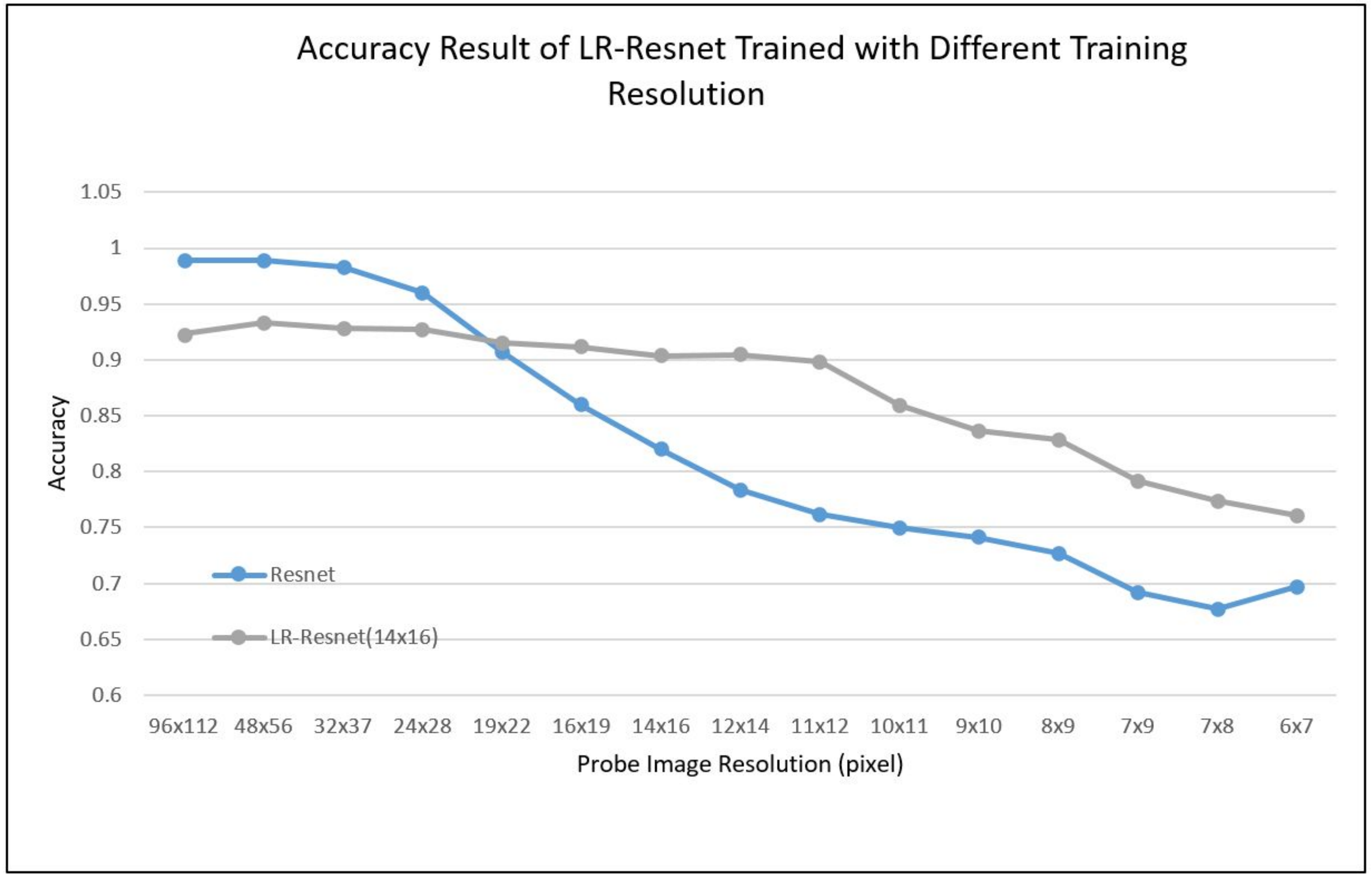
19x22

(LR-Resnet trained with 1000 iterations)

Deep-Learning Method

Interesting Finding

Training Image Resolution VS Accuracy



14x16

(LR-Resnet trained with 1000 iterations)

Conclusion

Conclusion

- In this project, both a conventional approach and a deep-learning-based approach are implemented and evaluated.

- Although the MDS method works, the result is not satisfactory.

- The feature loss can make the CNN-based model much suitable for low resolution condition and robust to image resolution, which realizes the objective of this project.

- As for the future work, we will try different loss functions to test the model’s performance and combine the image super-resolution with low-resolution face recognition.

Q & A

Thank You!