Similarity-Aware Deep Adversarial Learning for Facial Age Estimation

Penghui Sun, **Hao Liu***, Xing Wang, Zhenhua Yu and Suping Wu

Ningxia University, China

https://haoliuphd.github.io/paper/ICME2019Oral.pdf

Personal Introduction

Hao Liu

- Associate Professor, Ningxia University, China
- Ph.D supervised by Professor ***Jie Zhou*** and Professor ***Jiwen Lu*** of Tsinghua University
- Researches on facial analysis, particularly face alignment and facial age estimation

Selected Publications

- **Hao Liu**, Jiwen Lu*, Minghao Guo, Suping Wu and Jie Zhou. Learning Reasoning-Decision Networks for Robust Face Alignment, ***T-PAMI***, 2019
- **Hao Liu**, Jiwen Lu*, Jianjiang Feng and Jie Zhou. Two-Stream Transformer Networks for Video-based Face Alignment, ***T-PAMI***, 2018
- **Hao Liu**, Jiwen Lu*, Jianjiang Feng and Jie Zhou. Ordinal Deep Learning for Facial Age Estimation, ***T-CSVT***, 2019
- **Hao Liu,** Jiwen Lu*, Jianjiang Feng and Jie Zhou. Label-Sensitive Deep Metric Learning for Facial Age Estimation, ***T-IFS***, 2018

Age Estimation

- They are of nearly the same ages!

Zhi-ying Lin (**1974**)   De-gang Guo (**1973**)

Challenges

- Different ages look alike in appearance.

Brad Pitt (**1963**) Leonardo DiCaprio (**1974**)

Why Challenging?

- Apparent Age Estimation [ICCVW 15]

- Challenges
  - Large variances due to facial expressions and occlusions
  - Appearance changes with different facial make-up
  - Limited training samples/missing labels
  - **Label correlation for our human ages (real world)**

Conventional Methods

- Facial Age Estimation Framework

- Conventional Methods for Facial Age Estimation：
  - ✓ Feature Extraction ***(Requiring Much Strong Prior Knowledge)***
    - Hand-crafted Features: BIF, LBP, SIFT
    - Shallow Feature Learning: CS-LBFL [*Lu et al*, *T-IP 2015*]
  - ✓ Age Predictor ***(Imbalance of Class/Training Sample)***
    - LDL [*Geng et al*, *T-PAMI 2013*]
    - OHRANK [*Chen et al*, *CVPR 2012*]

Age Estimation by Deep Learning

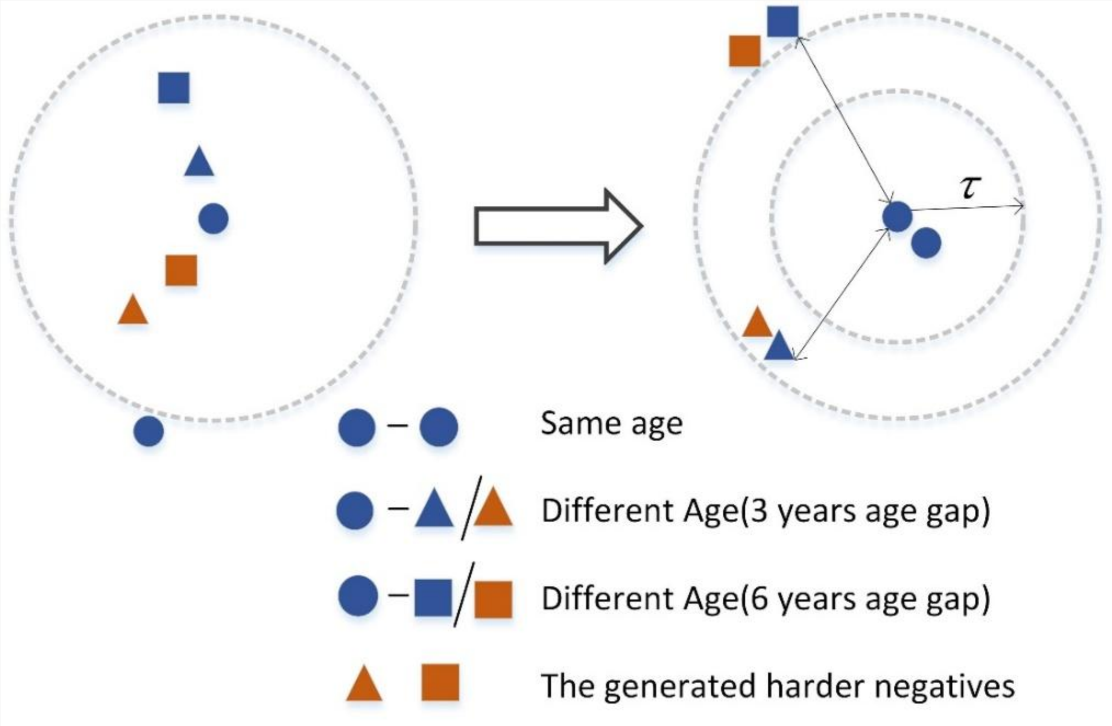
- Why Deep Feature Learning?
  - ✓ **Learning Features** directly from raw pixels
  - ✓ **Modeling Nonlinear Relationship** between Pixels and Labels
  - ✓ Transfer Learning (Fine-tuning)
- Label Correlation
  - ✓ Ordinal Regression [*Niu et al, CVPR 2016*]
  - ✓ Missing Labels [*Liu et al, PR 2017*]
- ***Goal: Jointly learning feature descriptors for face representation and exploiting the relationship of human age labels***
- ***ISSUES***
  - ***Hard/Semi-hard examples are meaningful (violates)***
  - ***Hard-Mining in unobserved space [Duan el al, CVPR 2018]***

Our Insight

- Two-fold criterions:

✓ The distance between each pair from different classes with a small age gap (circle and triangle) is smaller than that from a negative pair with a large age gap (circle and square).

✓ The distances of the pairs with same ages should become as smaller as possible.



Blue: **existing training samples**

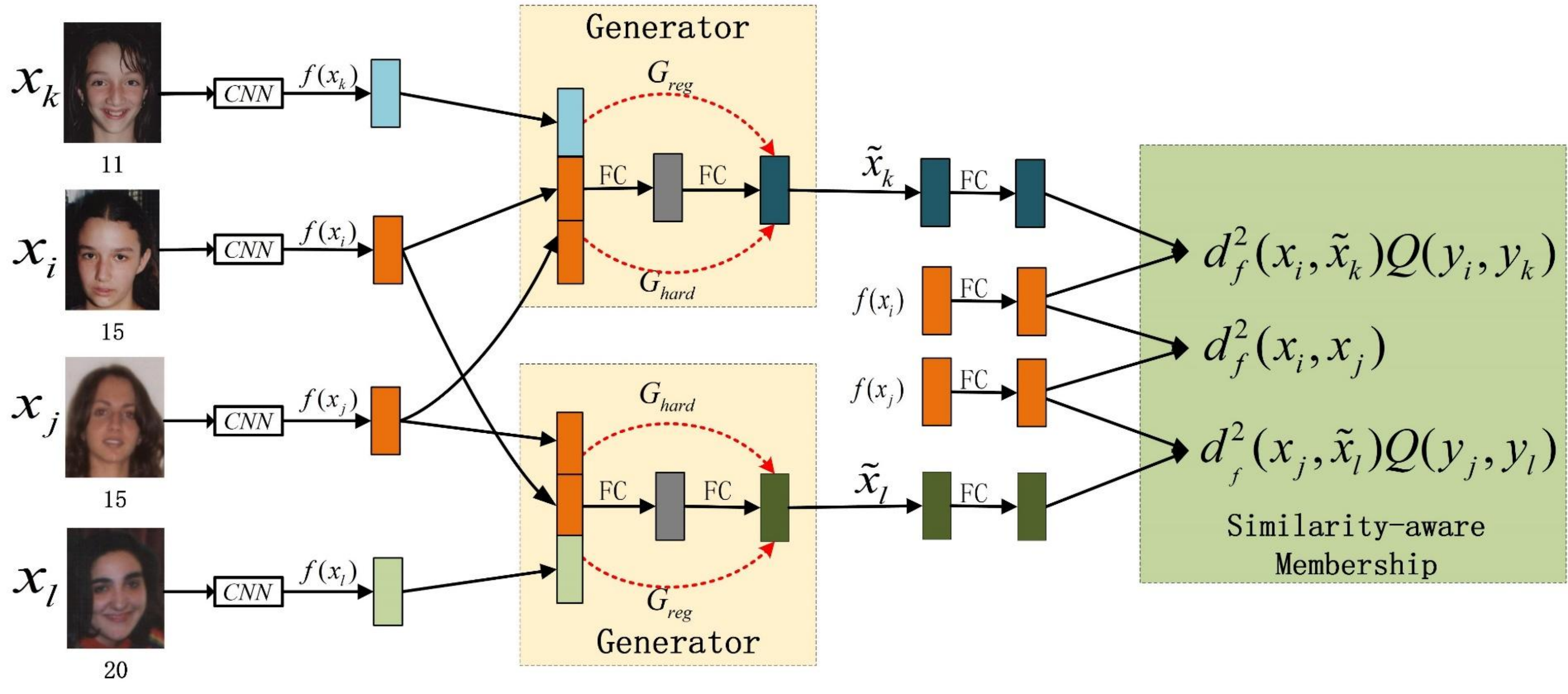
Orange: **generated hard-negatives**

Main Contributions

- Our method aims to seek batches of unobserved hard-negative samples based on existing training samples, which typically reinforces the discriminativeness of the learned feature representation for facial ages.

- Motivated by the fact that age labels are usually correlated in real-world scenarios, we carefully develop a similarity-aware function to well measure the distance of each face pair based on the age value gaps.

Proposed Framework



Objective Formulation

$$\min_{\theta_g,\theta_d} J = G_{(i,j,k,l)} + \lambda D_{(i,j,k,l)},$$

- The generator ***G*** aims to generate ***hard-negative*** samples in which the learned metric would misclassify.

- The discriminator ***D*** optimizes a discriminative distance metric, where ***the inter-class separability, intra-class compactness*** and label correlation of age classes are exploited to characterize the feature similarity simultaneously.

Hard-Negative Generation

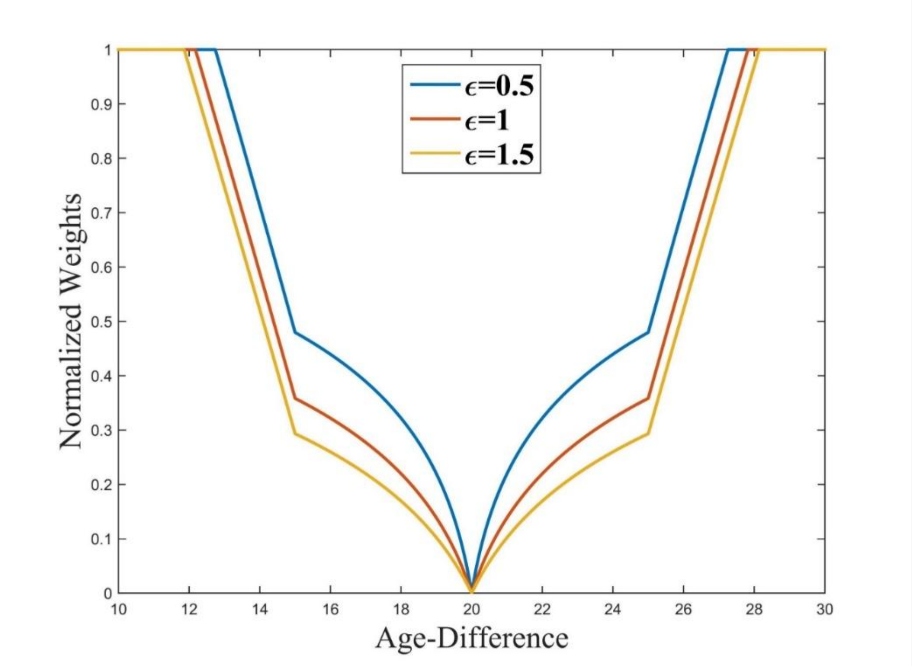
$$\min_{\theta_g,\theta_d} J = \boxed{G_{(i,j,k,l)}} + \lambda D_{(i,j,k,l)},$$

$$\min_{\theta_g} G_{(i,j,k,l)} = G_{hard} + G_{reg} + G_{adv},$$

Subject to

$$G_{(i,j,k,l)} = \begin{cases} \sum_{(i,j,k)} [\| \tilde{x}_k - x_i \|_2^2 + \| \tilde{x}_k - x_k \|_2^2 \\ + \max(0, d_f(x_i, \tilde{x}_k)[-Q(y_i, y_k)] - \tau)^2], \\ \\ \sum_{(i,j,l)} [\| \tilde{x}_l - x_j \|_2^2 + \| \tilde{x}_l - x_l \|_2^2 \\ + \max(0, d_f(x_j, \tilde{x}_l)[-Q(y_j, y_l)] - \tau)^2], \end{cases}$$

Similarity-Aware Function



$$Q(y_m, y_n) = \begin{cases} \frac{1}{L}\ln(1+\frac{|y_n - y_m|}{\varepsilon}), & |y_m - y_n| \le L \\ \frac{|y_n - y_m|}{L} - C, & otherwise \end{cases}$$

Age Smoothness

- With this function, the face pair with a larger age gap has a higher weight than that with a smaller age gap.
- At the same time, we amplify the differences between the pairs with interval L in the transformed feature space.

Deep Adversarial Learning

$$\min_{\theta_g,\theta_d} J = G_{(i,j,k,l)} + \lambda \boxed{D_{(i,j,k,l)},}$$

$$\min_{\theta_d} D_{(i,j,k,l)} = D_{(i,k)} + D_{(j,l)} + D_{(i,j)},$$

Subject to

$$D_{(i,j,k,l)} = \sum_{(i,j,k,l)} [\max_{(i,k)\in\hat{N}} (0, \tau - d_f(x_i, \tilde{x}_k) Q(y_i, y_k))^2$$

$$+ \max_{(j,l)\in\hat{N}} (0, \tau - d_f(x_j, \tilde{x}_l) Q(y_j, y_l))$$

$$+ \max_{(i,j)\in\hat{N}} (0, d_f(x_i, x_j))^2].$$

Experimental Results on Morph

Comparisons of MAEs with state-of-the-arts

Method	MAE	Year
BIF+KNN	9.64	-
OHRanker [6]	6.49	2011
LDL [7]	5.69	2013
CPNN [7]	5.67	2013
CA-SVR [22]	4.87	2013
CS-LBFL [9]	4.52	2015
CS-LBMFL [9]	4.37	2015
CSOHR [23]	3.74	2015
DeepRank [24]	3.57	2015
DeepRank+ [24]	3.49	2015
OR-CNN [13]	3.27	2016
ODFL [14]	3.12	2017
LSDML [10]	3.08	2018
M-LSDML [10]	2.89	2018
SADAL	**2.75**	-

Comparisons of MAEs with different deep learning approaches

Method	MAE
unsupervised VGG + KNN	7.21
unsupervised VGG + OHRanker	4.58
VGG + Single Label	3.63
VGG + Gaussian Label	3.44
ODFL [14]	3.12
SADAL	**2.75**

Experimental Results on FG-NET

Comparisons of MAEs compared with state-of-the-art approaches.

Method	MAE	Year
BIF+KNN	8.24	-
OHRanker [6]	4.48	2011
LDL [7]	5.77	2013
CPNN [7]	4.76	2013
CSOHR [23]	4.70	2015
CS-LBFL [9]	4.43	2015
CS-LBMFL [9]	4.36	2015
ODFL [14]	3.89	2017
LSDML [10]	3.92	2018
M-LSDML [10]	3.74	2018
SADAL	**3.67**	**-**

References

- **Hao Liu**, Jiwen Lu*, Jianjiang Feng and Jie Zhou: Label-Sensitive Deep Metric Learning For Facial Age Estimation. In IEEE Transactions on Information Forensics and Security (T-IFS), 13(2): 292-305 (2018).
- Yueqi Duan, Wenzhao Zheng, Xudong Lin, Jiwen Lu* and Jie Zhou: Deep Adversarial Metric Learning. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2780-2789 (2018).
- **Hao Liu**, Jiwen Lu*, Jianjiang Feng and Jie Zhou. Ordinal Deep Learning for Facial Age Estimation, IEEE Transactions on Circuits and Systems for Video Technology (T-CSVT), 2018.
- **Hao Liu**, Jiwen Lu*, Jianjiang Feng and Jie Zhou. Group-Aware Deep Feature Learning for Facial Age Estimation, Pattern Recognition (PR), 2017.
- **Hao Liu**, Penghui Sun, Jiaqiang Zhang, Suping Wu, Zhenhua Yu and Xuehong Sun: Similarity-Aware and Variational Deep Adversarial Learning for Robust Facial Age Estimation. In IEEE Transactions on Multimedia (T-MM), Under Review.

THANK YOU!   Q&A

https://haoliuphd.github.io/paper/ICME2019Oral.pdf