ICPR 2020

Identifying Missing Children: Face Age Progression via Deep Feature Aging

Debayan Deb
Michigan State University

Divyansh Aggarwal
Michigan State University

Anil K. Jain
Michigan State University

15th January 2021

- According to UNICEF and ICAT, **28% of the identified victims of human trafficking are children.**
- Around **8 million children go missing** around the world every year
- In 2019, there were **421,394 NCIC (National Crime Information Center) entries for missing children** in the US

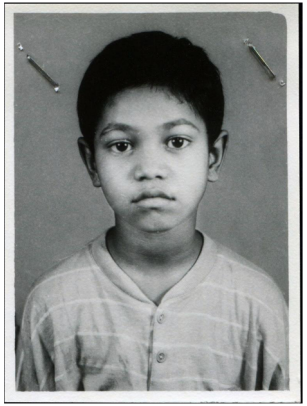
Child Trafficking

Missing Refugees/ Migrants

Abduction

1. https://www.fbi.gov/file-repository/2019-ncic-missing-person-and-unidentified-person-statistics.pdf/view

Identifying Missing Children



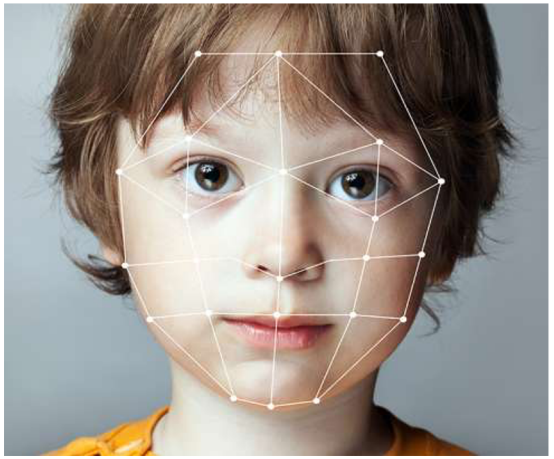
Saroo Brierley lost at the age of 5 (left) and later reunited with his family at the age of 30 (right)

Jaycee Dugard abducted at the age of 11 (left) and later retrieved at the age of 29 (right)

Which Biometric trait to use?

Iris

Fingerprints



Face

Parents or relatives are more likely to have the missing child's face images opposed to iris or fingerprints

Effect of Aging on AFR systems

26 years

6 years

0.33

Hannah Taylor Gordon

16 years

5 years

0.34

Max Burkholder

Cosine Similarity scores ϵ [-1,1] via CosFace. **Score > 0.35** is considered as match (Threshold @ 0.1% FAR)

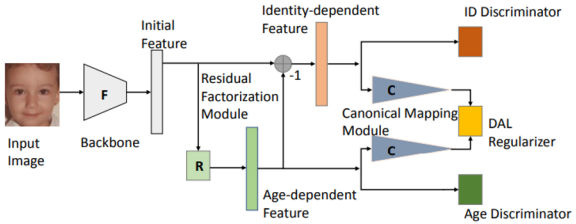
Requirements of an Age Progression Method

Identity Preservation

Visual Realism

Cosine Similarity scores ϵ [-1,1] via CosFace. **Score > 0.35** is considered as match (Threshold @ 0.1% FAR)

Prior Approaches

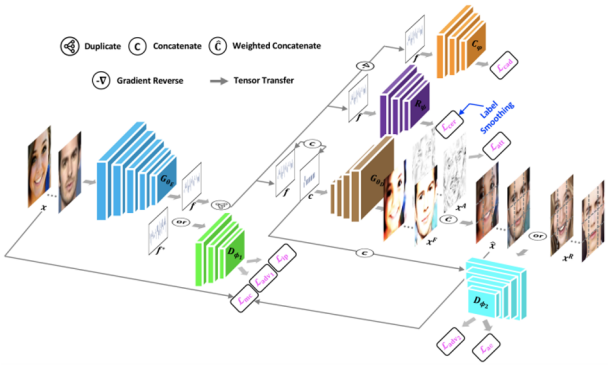


Discriminative Approaches [1]

+ Improve cross-age face recognition performance by discarding age information from face features

- Assume age and identity can be disentangled

- Assume identity specific features are adequate for face recognition



Generative Approaches [2]

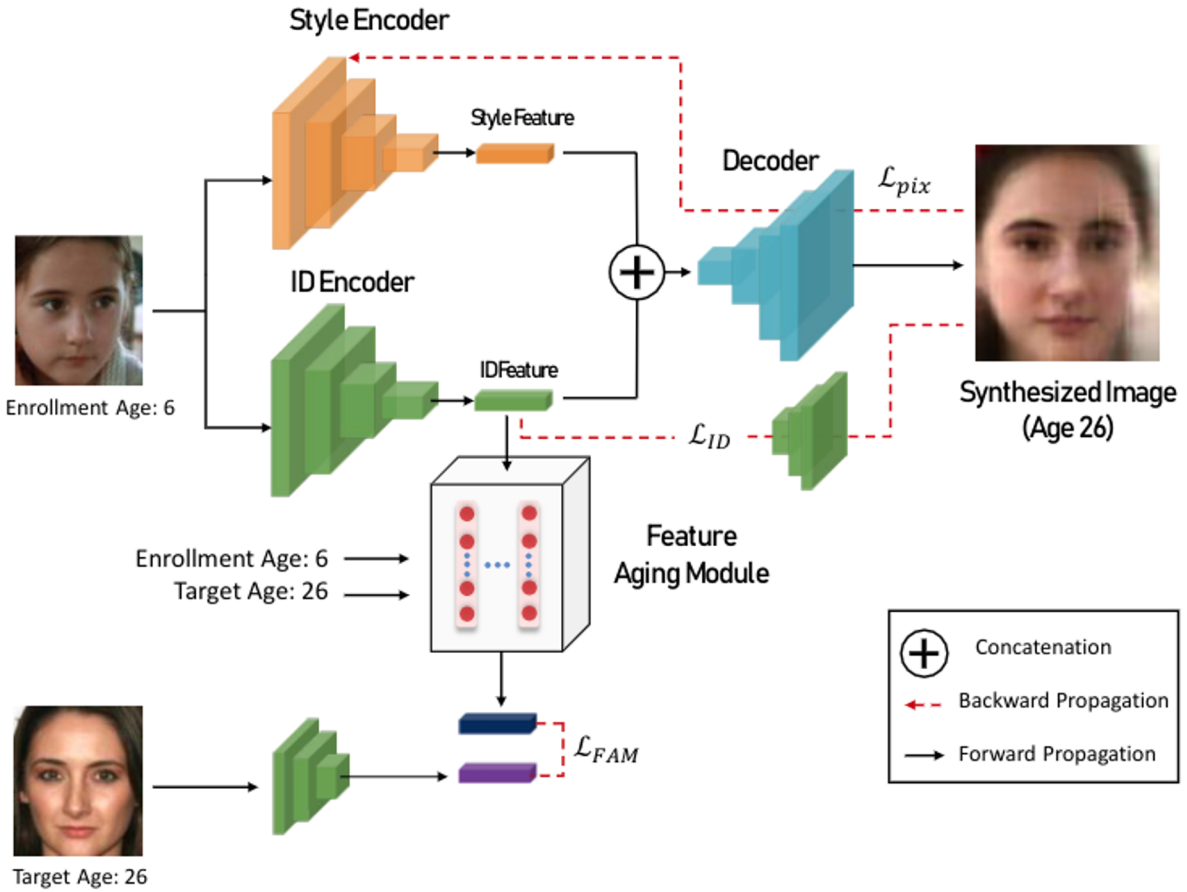
+ Synthesize realistic age-progressed faces by learning aging patterns from face aging datasets

- Identity is not preserved during the synthesis

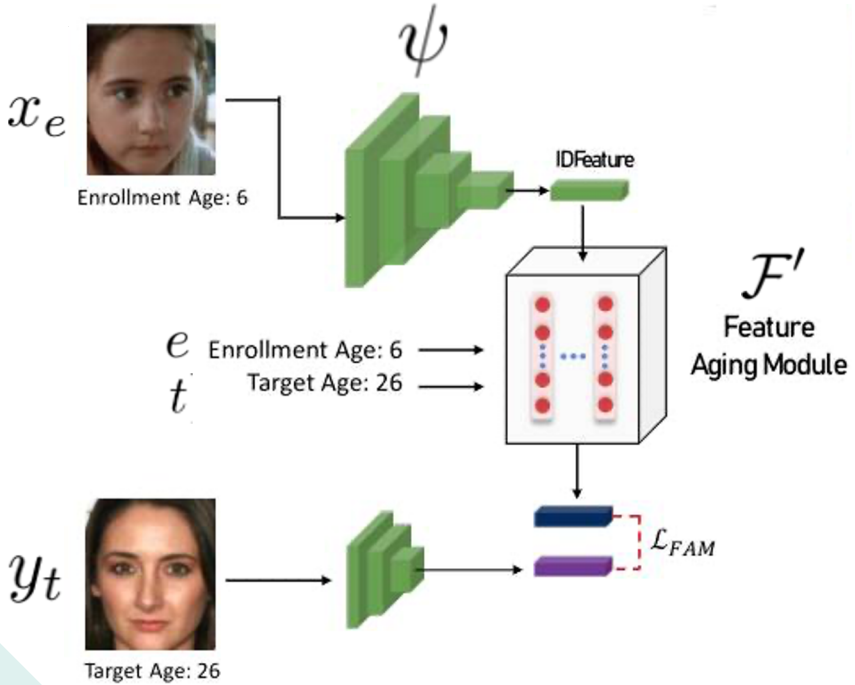
- Do not report cross-age recognition performance on the synthesized faces

1. H. Wang, D. Gong, Z. Li, and W. Liu, “Decorrelated adversarial learning for age-invariant face recognition,” in CVPR , 2019.
2. J. Zhao et al, “Look across elapse: Disentangled representation earning and photorealistic cross-age face synthesis for age-invariant face recognition,” in AAAI 2019

Proposed Approach



Feature Aging Module (FAM)



$$\mathcal{L}_{FAM} = \frac{1}{|\mathcal{P}|} \sum_{(i,j)\in\mathcal{P}} ||\mathcal{F}'(\psi(x_e), e, t) - \psi(y_t)||_2^2$$

Where, P is the Set of all genuine pairs

Image Generator

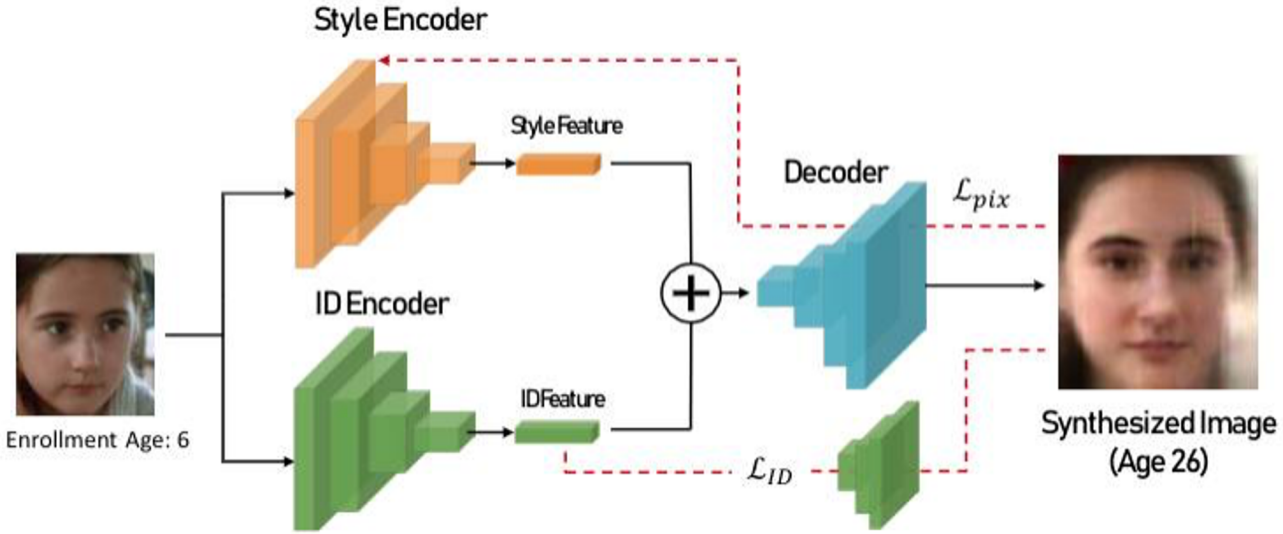


Image Generator

$$\mathcal{L}_{ID} = \sum_{i=0}^{n} ||\mathcal{E}_{ID}(\mathcal{D}(\mathcal{E}_{style}(x_i), \mathcal{E}_{ID}(x_i))) - \mathcal{E}_{ID}(x_i)||_2^2$$

Identity Preservation Loss
Preserves the identity in the synthesized image

$$\mathcal{L}_{pix} = \sum_{i=0}^{n} ||\mathcal{D}(\mathcal{E}_{style}(x_i), \mathcal{E}_{ID}(x_i)) - x_i||_1$$

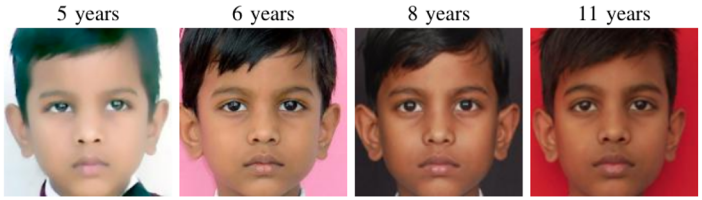
Pixel-level supervision loss
Ensures other details such as background etc. are preserved from the original image

$$\mathcal{L}_{TV} = \sum_{i=0}^{n} \left[\sum_{r,c}^{H,W} \left[\left(x_{i_{r+1,c}} - x_{i_{r,c}}\right)^2 + \left(x_{i_{r,c+1}} - x_{i_{r,c}}\right)^2 \right] \right]$$

Total variation Loss
To synthesize a smooth image

$$\mathcal{L}_{(\mathcal{E}_{style}, \mathcal{D})} = \lambda_{ID}\mathcal{L}_{ID} + \lambda_{pix}\mathcal{L}_{pix} + \lambda_{TV}\mathcal{L}_{TV}$$

Datasets



Child Face Aging (CFA)

25,180 images of 9,196 subjects
No. of Images/Subject : 2-6
Age Range (years) : 2-18
Avg. Age : 8 years

In the Wild Child Celebrity (ITWCC)

7,990 images of 745 subjects
No. of Images/Subject : 3-37
Age Range (years) : 0-32
Avg. Age : 13 years

Experimental Results

Quantitative Results

Method	CFA (Constrained)		ITWCC (Semi-Constrained)	
	Closed-set	Open-set	Closed-set	Open-set
	Rank-1	Rank-1 @ 1% FAR	Rank-1	Rank-1 @ 1% FAR
	P: 642 , G: 2213	P: 3290 , G: 2213	P: 611 , G: 2234	P: 2849 , G: 2234
COTS	91.74	91.58	53.35	16.20
FaceNet (w/o FAM)	38.16	36.76	16.53	16.04
FaceNet (with FAM)	**55.30**	**53.58**	**21.44**	**19.96**
CosFace (w/o FAM)	91.12	90.81	60.72	22.91
CosFace (with FAM)	**94.24**	**94.24**	**66.12**	**25.04**
CosFace (Image Aging)	**93.18**	**92.47**	**64.87**	**23.40**

Rank-1 identification accuracy on two child face datasets, CFA and ITWCC, when the time gap between a probe and its true mate in the gallery is larger than 5 years and 10 years, respectively. The proposed aging scheme (in both the feature space as well as the image space) improves the performance of FaceNet and CosFace on cross-age face matching. We also report the number of probes (P) and gallery sizes (G) for each experiment.

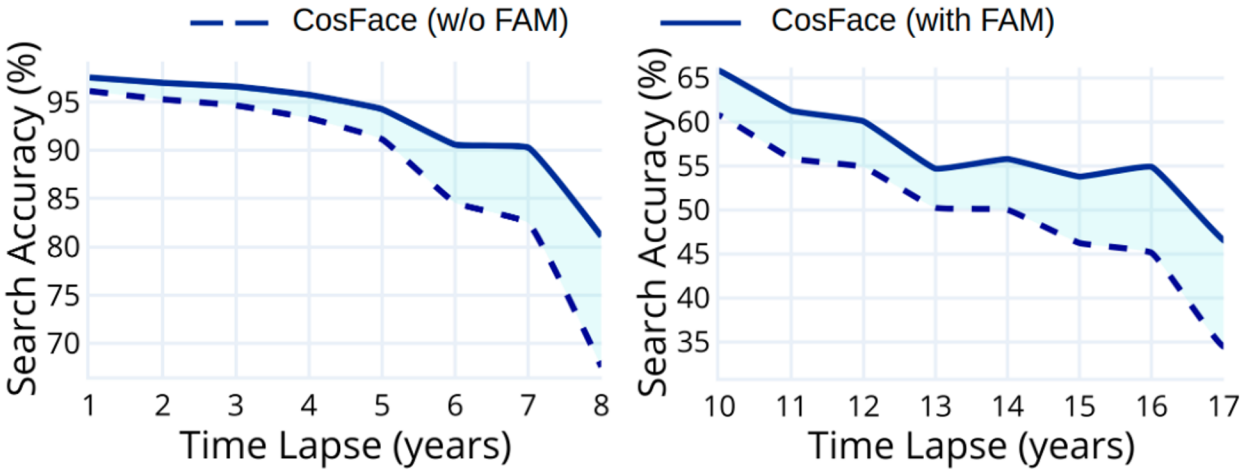
Quantitative Results

Face Recognition Performance on FG-NET and CACD-VS

Method	FG-NET	CACD-VS
	Rank-1 (%)	Accuracy (%)
HFA [1]	69.00	84.40
LF-CNN [2]	88.10	98.50
AIM [3]	93.20	99.38
Wang et al. [4]	94.50	99.40
COTS	93.61	99.32
CosFace (w/o FAM)	94.91	99.50
CosFace (Finetuned on children)	93.71	96.78
CosFace (with FAM)	**95.91**	**99.58**

[1] D. Gong et. al. "Hidden factor analysis for age invariant face recognition," in CVPR , 2013.
[2] C. Nhan Duong et. al. , "Temporal non-volume preserving approach to facial age-progression and age-invariant face recognition," in ICCV , 2017.
[3] J. Zhao et al, "Look across elapse: Disentangled representation earning and photorealistic cross-age face synthesis for age-invariant face recognition," in AAAI 2019
[4] H. Wang, D. Gong, Z. Li, and W. Liu, "Decorrelated adversarial learning for age-invariant face recognition," in CVPR , 2019.

Quantitative Results

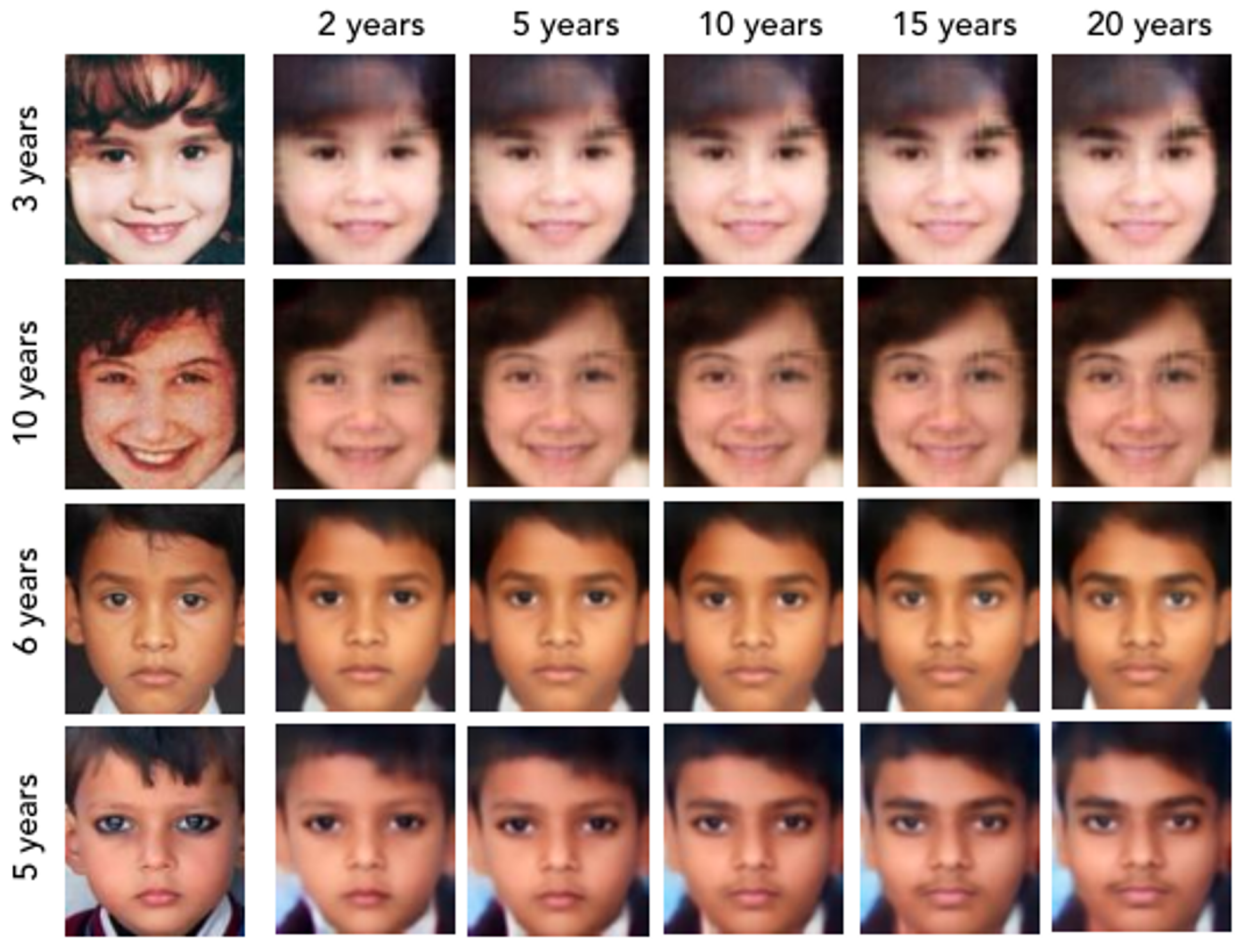


(a) CFA (b) ITWCC

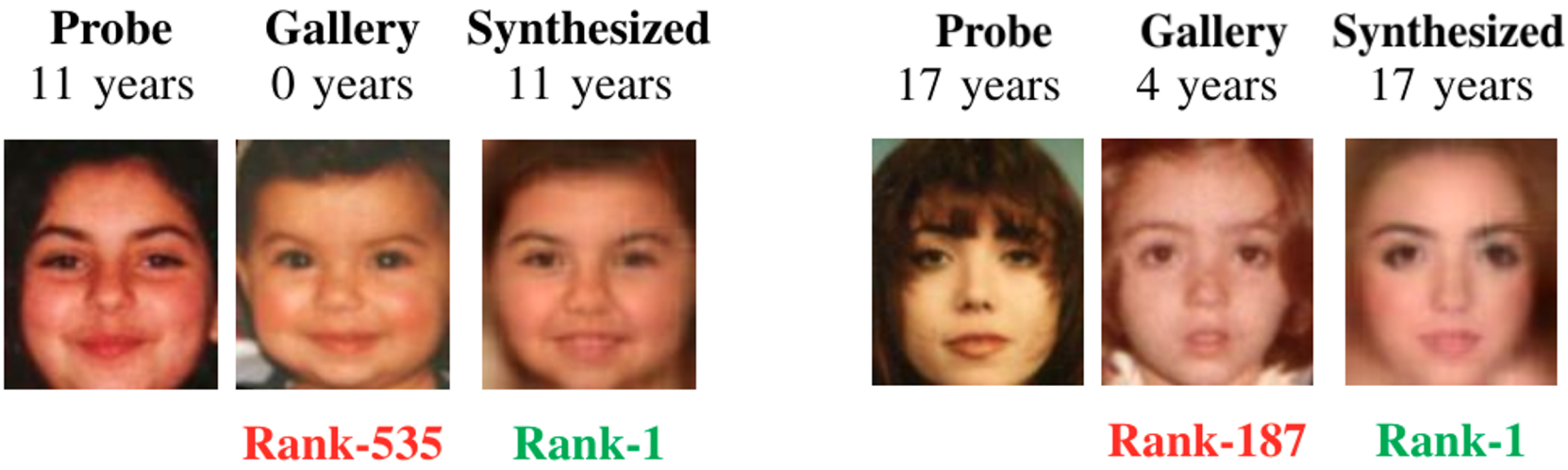
Fig. 4: Rank-1 search accuracy for CosFace [1] on (a) CFA and (b) ITWCC datasets with and without the proposed Feature Aging Module (FAM).

[1] H. Wang, Y. Wang, Z. Zhou, X. Ji, D. Gong, J. Zhou, Z. Li, and W. Liu, "Cosface: Large margin cosine loss for deep face recognition," in CVPR , 2018.

Qualitative Results

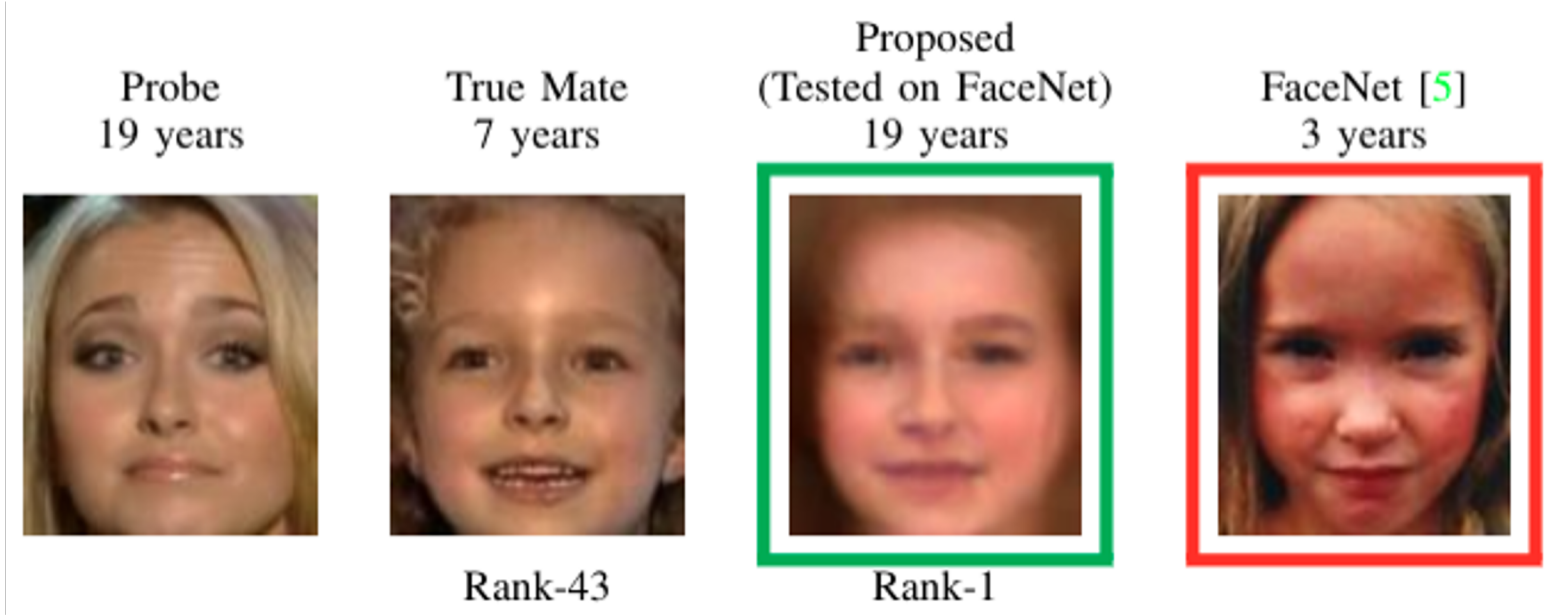


Retrieval Results



The true mate in the gallery was recovered at ranks highlighted in red. But with the proposed approach, CosFace is able to retrieve the true mate at Rank-1

Generalizability



FaceNet originally achieves 16.53% Rank-1 accuracy.

Via aged images (trained via CosFace), FaceNet achieves **21.11%** Rank-1 identification rate.

Case Studies

Probe
(24 years)

28 years

Incorrect Retrieval at Rank-1 by CosFace

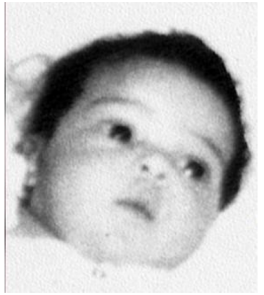
True Mate
(5 years)

With proposed, correctly retrieved at **Rank-1** by CosFace

Richard Wayne Landers abducted by his grandparents at age 5 in July 1994 in Indiana was later identified at age 24

Probe
(23 years)

True Mate
(19 days)



CosFace : Rank 3,069
COTS : Rank 1,242
Proposed : **Rank 268**

Carlina White was abducted from the Harlem hospital center in New York City when she was 19 days old and later reunited with her family at the age of 23 years

Thank You