FG-2017

Northeastern University

Kinship Verification on Families In the Wild with Marginalized Denoising Metric Learning

Shuyang Wang[1*], Joseph P. Robinson[1*], and Yun Fu[1,2]

[1]Department of Electrical and Computer Engineering, Northeastern University, USA
[2]College of Computer and Information Science, Northeastern University, USA
* Equal Contribution

Synergetic Media Learning Lab

Motivation

- Automatic kinship recognition is a challenging task.
- Kinship verification is a practical problem with an abundance of applications.
- Much attention since 2010 but has yet to transition from research-to-reality. Reasoning is 2-fold:
  - Existing kinship recognition datasets are insufficient in terms of size, diversity and, thus are under represents true family distributions worldwide.
  - There are many hidden factors affecting the facial appearances amongst family members with less discriminative power when compared to more conventional problems (e.g. facial recognition or object classification).
- **Goal:** Build large-scale kinship dataset to best support the task of kinship recognition.

Table 1 Comparison of FIW with related datasets.

Dataset	No. Fam	No. People	No. Faces	Age Varies	Full Fam	Highlights
CornellKin [6]	150	300	300	No	No	Parent-child pairs.
UB KinFace-I [9]	90	180	270	Yes	No	Parent-child pairs. Parents' 139 images at various ages.
UB KinFace-II [9]	200	400	600	Yes	No	Parent-child pairs. Parents' 139 images at various ages.
KFW-I [7]	—	1,066	1,000	No	No	Parent-child pairs.
KFW-II [7]	—	2,000	2,000	No	No	Parent-child pairs.
TSKinFace [10]	787	2,589	—	Yes	Yes	2 parents-child pairs, tri-verification.
Family101 [8]	101	607	14,000	Yes	Yes	Family structured.
FIW(Ours) [1]	996	10,700	378,300	Yes	Yes	Both depth & breadth and multi-task evaluation offerings.

Table 2 Verification scores (%). No family overlap between 5-folds

Methods	F-D	F-S	M-D	M-S	GF-GD	GF-GS	GM-GD	GM-GS	SIB	Average
LBP	54.76	54.69	55.80	55.29	56.40	56.37	54.32	56.85	57.18	55.74
SIFT	56.13	56.34	56.30	55.36	56.90	56.07	60.32	57.95	58.80	57.13
VGG	63.92	64.02	65.99	63.70	60.80	63.11	59.89	61.85	73.21	64.05
VGG+LPP [21]	65.03	69.09	67.87	69.37	63.70	62.74	66.11	63.50	73.46	66.76
VGG+NPE [15]	64.25	63.78	64.75	64.74	59.90	61.93	64.95	61.60	73.68	64.40
VGG+LMNN [31]	65.66	67.08	68.07	67.16	63.90	60.44	63.68	60.15	73.88	65.56
VGG+GmDAE [23]	66.53	68.30	68.15	67.71	64.10	63.93	64.84	63.10	74.33	66.78
VGG+DLML [11]	65.96	69.00	68.51	69.21	62.90	62.96	64.11	64.55	74.97	66.90
VGG+ours(mDML)	67.90	70.24	**70.39**	70.40	63.20	63.78	66.11	66.45	75.11	68.18
VGG+ours(DML)	**68.08**	**71.03**	70.36	**70.76**	**64.90**	**64.81**	**67.37**	**66.50**	**75.27**	**68.79**

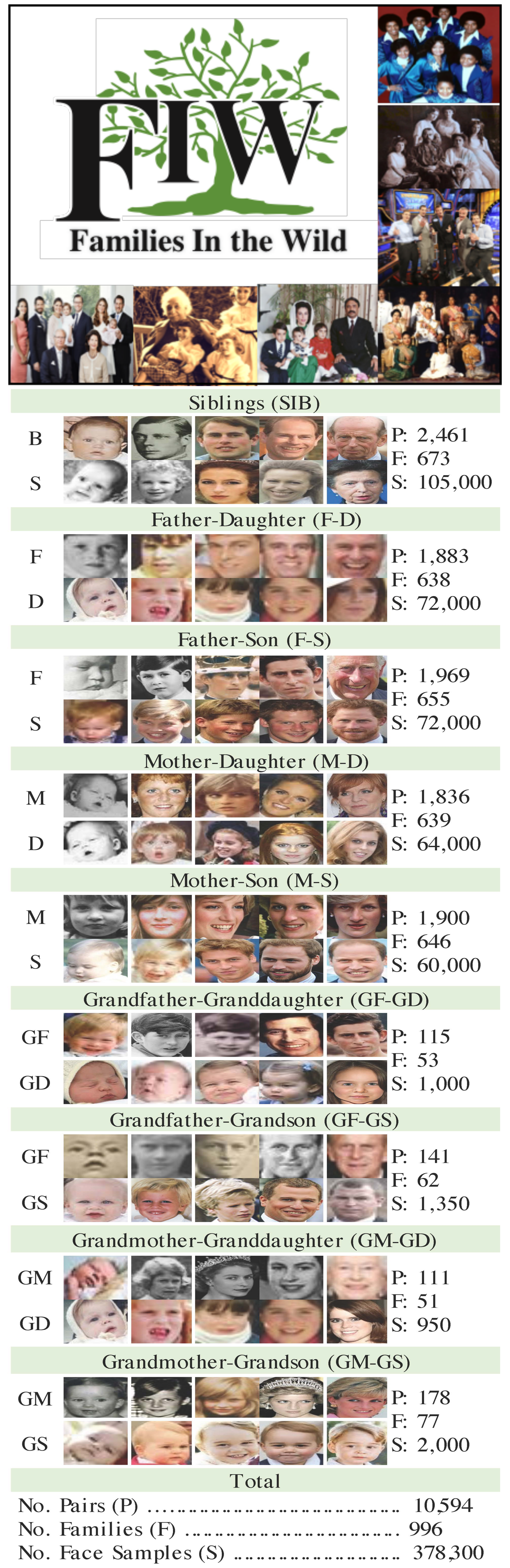
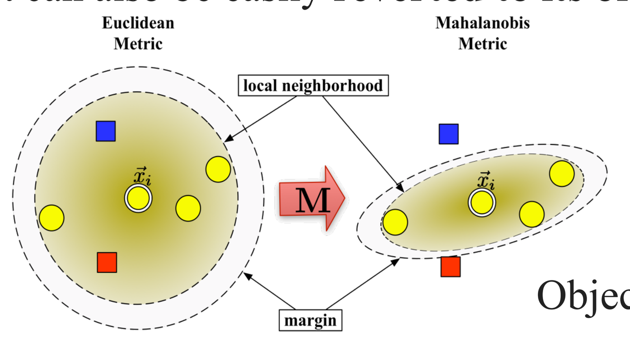


Fig 1 9 relationship types sampled from British Royal Family.

Methodology

- We propose a denoising AE (DAE) based metric learning (ML) method for kinship verification:
  - Incorporate metric learning on hidden layer units of AE, so that the projection matrix **W** in our model works both as the **encoder** in DAE and the **linear projection** in metric learning **simultaneously**.
  - Resulting data are not only be projected to a **maximum margin feature space**, but can also be easily **reverted to its original form**.



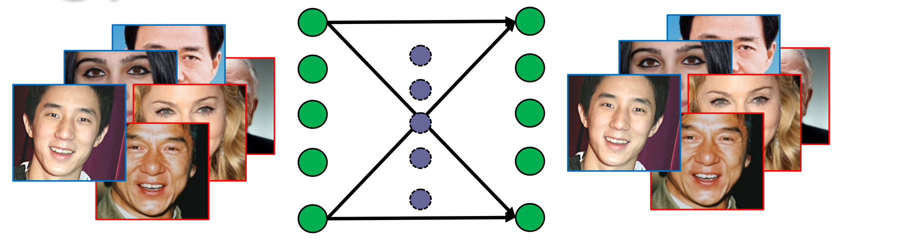
ML finds a Mahalanobis-like distance matrix from training data. We decompose M into W^TW, thus, the Mahalanobis distance converts to

$$d_M(x_i, x_j) = \|Wx_i - Wx_j\|_2$$

Formulation for discriminative ML is written as

$$\min_{WW^T=I} \frac{tr(WXL_wX^TW^T)}{tr(WXL_bX^TW^T)}$$

Where L_w and L_b are the Laplacian matrices.



Auto-encoder (AE) learns identity preserved representations in the hidden layer by setting the input and target as the same

$$z_i = \sigma(W_1x_i + b_1); \quad h(x_i) = \sigma(W_2z_i + b_2)$$

$$\min_{W_1,b_1,W_2,b_2} L(x) = \min_{W_1,b_1,W_2,b_2} \frac{1}{2n}\sum_i \|\hat{x}_i - h(x_i)\|_2^2$$

Marginalized Denoising Metric Learning

Objective Function:

$$\mathcal{L} = \min_{W_1,W_2,b_1,b_2} \frac{1}{2}\|X - \hat{X}\|_F^2 + \frac{\lambda}{2} tr\left(\frac{HL_wH^T}{HL_bH^T}\right)$$

where $H = \sigma(W_1X + B_1), \hat{X} = \sigma(W_2X + B_2)$, B_1, B_2 are the n-repeated column copy of b_1, b_2, respectfully. λ is the balance parameter between AE and ML.

Also, we have marginalized version:

$$\mathcal{L} = \min_W \frac{1}{2mn}\|\bar{X} - W^TW\tilde{X}\|_F^2 + \frac{\lambda}{2} tr\left(\frac{WXL_wX^TW^T}{WXL_bX^TW^T}\right)$$

Where $X = [x_1, \dots, x_n], \bar{X}$ is its m-times repeated version and $\tilde{X}$ is corrupted version

Experiments

- We report benchmarks using several different visual features and comparative methods. Note that our model could be based on both DAE or mDAE, which is denoted as DML and mDML, respectively.
- The comparative results show our approaches consistently outperform all the competitors, with the nonlinear version (DML) generally more accurate than the marginalized linear version (mDML). Table II illustrates the improvement of each method on all three features, along with training times. From the table, we can see that both our DML and marginalized version outperform all others. The marginalized version achieves a slightly smaller accuracy score, but at a **much faster training speed**, which is due to **the closed-form solution of mDAE**.
- The average verification rates for the 5-folds are shown in **Table 2**.
  - VGG-Face features produce better verification results than hand-craft features.
  - ROC curves are shown in **Fig 2**.

Table 3 Accuracy scores (%) and training time (sec) for father-son pairs.

Methods	LBP		SIFT		VGG	
	ACC	Time	ACC	Time	ACC	Time
baseline	54.69	-	56.34	-	64.02	-
LPP [21]	55.78	20.9	59.41	91.3	69.09	11.4
NPE [15]	54.57	20.2	56.51	58.7	58.3708	8.8
LMNN [31]	55.32	1052	58.93	681	67.08	363
GmDAE [23]	55.16	2.7	59.20	2.7	68.03	2.8
DLML [11]	55.25	131	58.62	112	69.00	97
mDML[Ours]	55.54	0.84	60.10	0.98	70.24	0.87
DML[Ours]	55.52	31.2	62.33	22.4	71.03	43.4

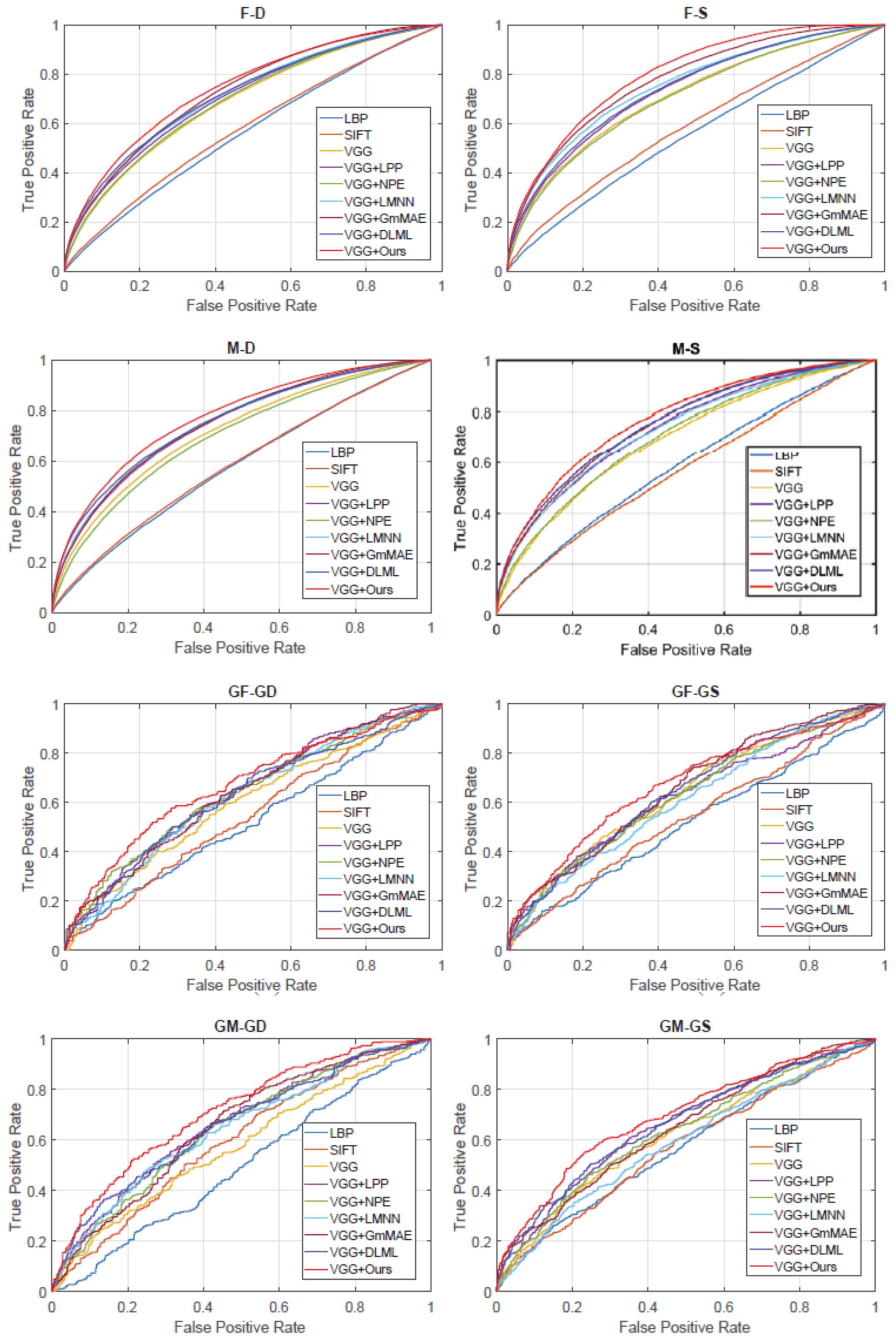


Fig 2 Relationship Specific ROC Curves.

References

[1] Siyu Xia, Ming Shao, and Yun Fu. Kinship verification through transfer learning. In IJCAI, 2011.

[2] J. P. Robinson, M. Shao, Y. Wu, H. Liu, T. Gillis, and Y. Fu, "Visual Kinship Recognition of Families In The Wild (FIW)," in IEEE Transactions on Pattern Analysis and Machine Intelligence Special Issue: The Computational Face (2017). (Under review).

[3] J. P. Robinson, M. Shao, Y. Wu, and Y. Fu, "Families in the wild (fiw): Large-scale kinship image database and benchmarks," in Proceedings of the 2016 ACM on Multimedia

[4] F.Schroff, D.Kalenichenko, J.Philbin. Facenet: A unified embedding for face recognition & clustering CVPR

[5] A. Bottino, M. De Simone, A. Laurentini, T. Vieira, "A new problem in face image analysis: Finding kinship clues for siblings pairs, "in Proceedings of International Conference on Pattern Recognition Application and Methods, Vilamoura, Algarve, Portugal, 6-8 February 2012. pp. 153-162.

[6] Y. Guo, H. Dibeklioglu, L. Maaten. Graph-based Kinship Recognition. ICPR (2014).

[7] Chen, Minmin, et al. "Marginalized Denoising Autoencoders for Domain Adaptation." Proceedings of the 29th International Conference on Machine Learning (ICML-12). 2012.