Recognizing Families In the Wild (RFIW)

Data Challenge Workshop in conjunction with ACM MM 2017

Joseph P. Robinson
Northeastern University
Boston, Massachusetts
jrobins1@ece.neu.edu

Ming Shao
University of Massachusetts
Dartmouth, Massachusetts
mshao@umassd.edu

Handong Zhao
Northeastern University
Boston, Massachusetts
hdzhao@ece.neu.edu

Yue Wu
Northeastern University
Boston, Massachusetts
yuewu@ece.neu.edu

Timothy Gillis
Northeastern University
Boston, Massachusetts
gillis.ti@husky.neu.edu

Yun Fu
Northeastern University
Boston, Massachusetts
yunfu@ece.neu.edu

Figure 1: The Families In the Wild (FIW) dataset [33, 40] used to support RFIW. To the left of tree icon is a subset of the Bruce Lee family gallery chosen at random, and to the right are family photos of different families randomly selected.

ABSTRACT

Recognizing Families In the Wild (RFIW), the first large-scale automatic kinship recognition challenge, consisted of two tracks, kinship verification and family classification. The challenge was hosted in preparation for a Data Challenge Workshop during the 2017 ACM MM conference. In the end, 10 teams completed the challenge, several then submitted papers on the techniques and algorithms used. This paper reviews the evaluation phase of RFIW by introducing task protocols, reporting results, and summarizing various submissions. A discussions about kinship recognition that spans past, present, and future is then provided. Finally, remarks about the upcoming workshop are given in terms of the plans, goals, and hopes for RFIW2017.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
RFIW'17, October 27, 2017, Mountain View, CA, USA
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5511-7/17/10…$15.00
https://doi.org/10.1145/3134421.3134424

1 INTRODUCTION

Automatic kinship recognition is relevant in an abundance of use cases. For instance, it could be used for forensic investigations, automatic photo library management, historic lineage and genealogical studies, social-media analysis, cases of missing children and human trafficking, contemporary issues involving immigrants, their families and border patrol, and problems involving refugees, both homeland and abroad.

If so relevant and useful, then why has kinship recognition technology not yet transitioned from research-to-reality? For starters, researchers have lacked a data supply large enough to sufficiently mimic real-world data that by capturing the distributions of families found worldwide. Additional to limited data resources, kin-based tasks tend to be challenging even when compared to other vision problems (*e.g.*, conventional facial recognition or object classification). Thus, modeling kinship poses many challenges. Provided larger amounts of data with rich label information both reasons would be satisified,*i.e.*, a data collection truly reflective of the visual nature of kinship that can support the more complex, data-driven learning methods (*i.e.*, deep learning). This motivated the construction and recent release of Families In the Wild (FIW) [33, 40]– the first large-scale image collection capable of supporting multiple kin-based tasks.[1]

[1]Check out project page for on FIW, http://smile-fiw.weebly.com/.

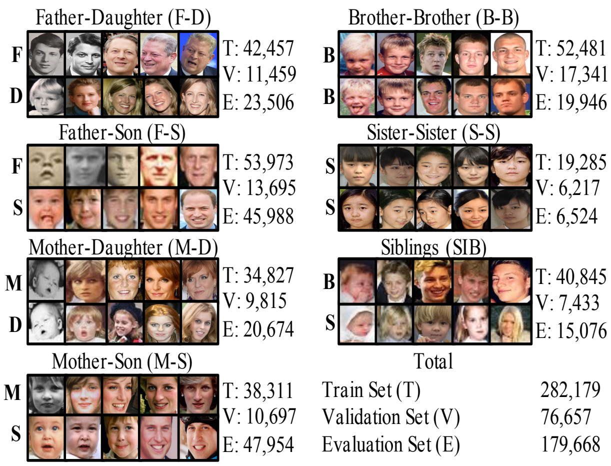

Figure 2: Pair counts per category and phase (*i.e., Train* T, *Validation* V, and *Test* Evaluation E).

FIW aims to bridge the gap between research and reality, for its size, diversity, rich labels, and image-level metadata packages much more than its predecessors. Now modern-day deep learning approaches can now be used in problems of recognizing kinship in imagery. In the end, FIW made this large-scale kinship recognition challenge, as part of a ACM MM 2017 Data Challenge Workshop, possible. Namely, Recognizing Families In the Wild (RFIW), a 2 task data challenge, supporting kinship verification (one-vs-one) and family classification (one-vs-many).

The rest of this paper is organized as follows. First, we explain overviews, data, and protocols for each task separately (Section 2). We then provide baselines, list results, and summarize submissions and report performance ratings (Section 3). Next, we review previous work related (Section 4) and discuss the impact of advancing automatic kinship technologies and some potential future directions this work may take (Section 5).

2 TASK EVALUATIONS AND PROTOCOL

In this section, we review the two tasks (*i.e.*, kinship verification and family classification). For each task, we provide an overview, highlight the intended use, and describe the experimental settings (*i.e.*, data splits and metrics).

2.1 Kinship Verification

The goal of kinship verification is to determine whether a pair of facial images of different subjects are kin of a specific type (*e.g.*, parent-child). This classic boolean problem formulates a *one-vs-one* view of automatic kinship recognition, with system responses being either KIN or NON-KIN (*i.e.*, true or false, respectively).

2.1.1 Intended Use. Prior research mainly considered parent-child types, *i.e.*, father-daughter (F-D), father-son (F-S), mother-daughter (M-D), mother-son (M-S). Although less, some attention was put on siblings, i.e., Sister-Sister (S-S), Brother-Brother (B-B), and siblings of opposite sex (SIBS). As studies in psychology and computer vision found, different types of kin share different familial features, so pair types are typically modeled and evaluated independent of all others. Thus, additional kinship types would further both our understanding and capabilities for kinship recognition. With FIW, the number of face pairs accessible for kinship verification is orders of magnitude greater than all existing kinship image collections. A subset of the kin types and face pairs were used for the challenge. Figure 2 lists the pair types and counts used for RFIW.

2.1.2 Data Splits. FIW contains 644, 000 pairs in total. From these, 538, 504 pairs of 7 different kin types were used for the challenge (Figure 2). Face pairs were split into 3 disjoint sets (*i.e.*, *Train*, *Validation*, and *Test*). Ground truth was provided for *Train* set and servers were open for scoring of *Validation* during **Phase 1**. Than, ground truth for *Validation* was made available during **Phase 2**. Finally, the "blind" *Test* set was available during **Phase 3**. No labels were provided for the *Test* set and teams were asked to only process the *Test* set to generate submissions and, hence, avoid any attempt to analyze or understand the *Test* set. Each set contained an equal number of positive and negative pairs. Note that there was no family or subject identity overlapping between sets.

2.1.3 Evaluation Settings & Metrics. Verification accuracy was the metric used and reported.

2.2 Family Classification

The goal of family classification is to determine which family a subject belongs to. Prior to testing, multiple face samples for several members are given for a set of known families (i.e., family classes). During testing, a face of an unseen subject is given as input, with the output being the family prediction. Hence, this task is formulated as a *one-to-many*, closed-form classification problem.

2.2.1 Intended Use. The goal of family classification is determine the family that an unseen subject belongs to. This is done by referencing faces, *i.e.*, families modeled using facial images of all but the held-out family members, then at test-time the held out members are used to evaluate on. Types of held out family members vary by type– from the youngest boy in a family tree that spans back several generations to a member whom assumes the role of being a mother, a sister, a daughter, and sits right in the middle of the family tree.

This is a challenging task that grows in difficulty with an increasing number of families. This is due to the large intra-class variations of families that are typically tricky to encode and model. Additional challenges come from the data being unconstrained (*i.e.*, faces *in the wild*). In other words, variations in pose, illumination, expression, etc., are present– developing algorithms to handle such data is critical for transitioning from research-to-reality.

In this evaluation, there is a set of faces for each of the several members from the 227 families. The goal is to identify the family labels for the family members held out during training split. Families with at least 5 members are included, and 1 family member is selected at random to be held out until test time.

2.2.2 Data Spits. FIW includes a total of 1, 001 families with multiple samples for each of the members, 227 of which are made

Figure 3: Photos of families sampled randomly from FIW (*i.e.*, 27 of 1, 001).

up of 6 or more members. These 227 families was the subset used for the challenge.

Similar to Track 1, the data was split into 3 disjoint sets referred to as Train, Validation, and Test sets. Ground truth for training was be provided during **Phase 1** and the validation data without labels (*i.e.*, scoring server was open to evaluate validation set). Then, labels for Validation were released at the start of **Phase 2**. Lastly, the "blind" Test set was released during **Phase 3**. No labels were provided for the Test set until after the competition was adjourned. Again, teams were asked to only process the Test set when generating submissions to prevent attempts to analyze and or understand the Test set. No member overlap existed between the different sets (*i.e.*, train, val, and test sets).

2.2.3 Evaluation Settings & Metrics. The results for this multi-class problem will be reported as top 1% error rates.

3 SUMMARY OF SUBMISSIONS

We next review results for each team, with kinship verification and family classification covered separately for each. First, however, we describe task-specific baselines.

3.1 Baseline Methods (1.A & 2.A-C)

Track 1– Kinship Verification

We used Centerface (CF) [42], which was proposed as a 22-layer residual convolutional neural network (CNN) trained on CASIA-Webface [49] for conventional facial recognition. CF introduced a new loss referred to a Center-loss. This loss focuses on reducing intra-class variations, from which state-of-the-art performance was achieved on several face benchmarks. We used the model provided by the authors to use as a feature extractor (*i.e.*, used *off-the-shelf* to extract face encodings). Face pairs were scored using cosine similarity.

For CF+FT, we fine-tune the CF model using the data that contains 400 families in training and validation set. The training used four Titan X GPUs with a batch size 256. The learning rate was initialized at 0.01, and decreased by a factor of 10 at 400 and then again at 600 iterations. Training was completed after 800 iterations.

Track 2– Family Classification

To classify the 228 families, we first used one-vs-all linear SVM [4] as the classifier and utilized two state-of-the-art deep face representations, *i.e.*, VGG and Centerface. Further, we fine-tuned a CF model by replacing the original loss layers with a softmax loss set to classify the 228 family classes. Same training strategy from Kinship Verification is used, including learning rate, batch size and iterations.

3.2 Institute of Computing Technology, China (1.B)

The authors proposed KinNet, a fine-to-coarse deep metric learning framework for kinship verification. In the framework, the authors transfered knowledge from the large-scale-data-driven face recognition task by pre-training the network with massive data for face recognition. Then, the network was fine-tuned to find a metric space where kin-related peoples are discriminant. This team obtained top performance in Track 1 and, thus, won 1^{st} place in the competition. See Figure 4 for the end-to-end framework.

Specifically, the authors adopted different variations of the residual network [17] as the basic architecture of KinNet. Pre-training was done on subset of MS-Celeb-1M dataset [16], with the output dimension of $fc - layer$ set to 41, 856. A softmax loss was added after the $fc - layer$ to decide between the 41, 856 subjects. While fine-tuning, the $fc - layer$ was replaced with a new $fc - layer$ of size $1024D$, which was followed by an L2-norm layer to normalize features to unit length. In the end, a soft triplet-loss was configured to force the KINs closer and NON-KINs farther apart. Since the first layer learns the most general representations (*i.e.*, lower-layers tend to learn more basic filters), the authors froze the bottommost 7×7 convolution layer and updated other layers to adapt the model for kinship verification. To increase the number of training images, as well as balance the number of images per member, the authors proposed an augmentation strategy that makes each member of the 300 families contain the same number of images. Finally, the reported kinship verification result is derived from the fusion of different models with different network depth and cropped sizes at the input.

Table 1: Verification accuracy scores (%) for Track 1 of RFIW2017. Top 3 scores are color coded: 1^{st} (blue), 2^{nd} (red), and 3rd (orange) highest scores. Reference strings under REF match text description with entry of table.

REF	USERNAME	F-D	F-S	M-D	M-S	SIBS	B-B	S-S	Avg.
1.A	BASELINE	68.53	69.02	73.11	71.54	68.37	69.97	75.09	70.81
1.B	mysee1989	70.79	71.46	77.87	78.62	79.91	74.77	70.57	74.86
1.C	ella	66.89	65.284	72.31	71.68	71.09	64.67	70.11	68.86
1.D	DQYCQU	65.196	63.38	70.86	65.22	72.10	64.00	66.51	66.58
1.E	viking	61.31	64.22	63.81	63.296	63.58	62.70	64.57	63.22
1.F	eranda	60.77	62.15	65.03	63.13	64.59	61.70	64.31	63.097
1.G	faud	61.31	61.98	62.30	57.69	63.83	60.60	61.62	61.33
1.H	Ji.L	59.95	59.72	61.12	55.69	59.40	57.68	58.30	58.83
1.I	BIU-team	56.16	60.594	60.00	54.33	61.205	59.70	58.02	58.21
1.J	oualid.laiadi	54.92	55.93	54.89	54.33	57.96	53.38	52.25	54.81

Table 2: Classification accuracy scores (%) for Track 2 of RFIW2017.

REF	USERNAME	Description	Acc.
2.A	BASELINE-1	VGG+SVM.	17.20
2.B	BASELINE-2	Center+SVM.	9.11
2.C	BASELINE-3	Center-Face Fine-tuned with softmax .	24.74
2.D	viking	Fine-tuned CNN using Center-loss, training FC5 and a new FC6 layer.	29.45
2.E	ella	Fine-tuned CNN with fc5 layer connected to a centerloss ($loss_{weight} = 0.001$).	29.36
2.F	mysee1989	Fine-tuned using new soft-max.	17.62

3.3 East China Normal University (1.C & 2.D)

Track 1– Kinship Verification

These participants fine-tuned a pre-trained VGG-Face CNN [31] (*i.e.*, VGG-16) for classification via the family labels. From which they used an embedding size of 512, instead of the last fully-connected layer used to identify 300 families. Then, the CNN was fine-tuned using triplet-loss function [35] on top, as defined in Eq. (1). The triplet-loss was found as the cosine similarity between two facial encodings, which slightly differs from [36]. To mine the triplet pairs, the authors used all positive pairs from each of the seven relationship types provided during training, then selecting negative samples using family labels to avoid conflicts (*i.e.*, ensure true negatives, opposed to be another sample or member from that same family). All models were trained using Adam [21], a batch size of 40, and iteration size of 3. In each batch, two images were used per class and all positive pairs were sampled. When mining negatives in one batch, experiments are carried out separately by three methods which are semi-hard (SH) [36], batch-hard (BH) [18] and batch-all (BA) [18].

$$L = (margin + CS_{an} - CSap)_+, \tag{1}$$

where CS represents the cosine similarity of two feature vectors; $margin = 0.2$ in their work.

Cosine similarity was computed for all pairs of each relationship type making up the validation set to evaluate the models. Each of the three negative mining methods were trained for 10 epochs with reported performance of the best epoch, rather than at the end of training. They achieved thresholds for each type of kin to use on evaluation set as ROC curves generated from the validation set. Since there was no family overlap in the training, validation, and test set, they selected 7 models for each negative mining method (*e.g.*, $L_{BA}(averagemodel)$) to average results of validation to improve thresholds for testing.

Track 2– Family Classification

Utilized the official pre-trained VGG model and employed a softmax loss as the loss function to solve this multi-class classification problem. They too used a Center-loss [43] as auxiliary loss– Center-loss simultaneously learns class-specific centroids while penalizing according to the distance between centroids of different classes. The loss weight of softmax was 1 and the loss weight of CF was 0.001.

3.4 Chongqing University, China (1.D)

Inspired by Maximum Mean Discrepancy (MMD) [2] and Generative Adversarial Net (GAN) [14], the authors proposed a family ID based Adversarial contrastive residual Network (AdvNet) for

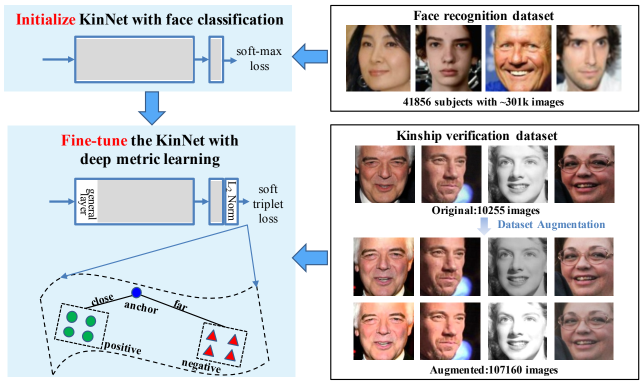

Figure 4: Proposed KinNet from Institute of Computer Technology in China (Ref 1.B). KinNet is first trained for face recognition on large amounts of data. It is then fine-tuned for coarser-grained recognition by finding kin-specific metric spaces with a soft triplet-loss.

large-scale (1 Million) kinship recognition. The proposed network interprets distribution differences between pairs of faces in the first fully-connected layer ($fc - layer$), which tends to minimize the inter-class discrepancy and maximize the intra-class discrepancy. In contrast, a contrastive loss was formulated to maximize the inter-class distance and minimize the intra-class distance in the second $fc - layer$. To integrate the family class, a softmax loss was added as the topmost layer to further improve the recognition performance. Finally, the feature augmentation was achieved by an ensemble of features from different deep models (see Figure 6).

3.5 Bar-Ilan University, Israel (1.F)

Proposed novel system that uses deep features to do kinship verification. As shown in Figure 7, the first step of the system used VGG-Face [31] to encode faces from the output of the last convolution layer. These outputs were $7x7x512$ feature maps, where each of the 512 features in the $7x7$ pixel-patch was derived from the spatial local domain in image space.

The next block in Figure 7 depicts the fusion of features from the two candidate images, *i.e.*, concatenated the two features maps from the first block to yield a single feature map. Then, they constructed a concept-convolution filter along each feature, simultaneously. Finally, a dense layer was added in order to do the binary classification (*i.e.*, KIN/ NON-KIN).

3.6 University of Biskra, University of Valenciennes (1.J)

Proposed a framework called LPQ-SIEDA, which was based on several well demonstrated components (see Figure 5). Specifically, *Side-information based Exponential Discriminant Analysis* (SIEDA) [30] was shown to be more effective than *Side-Information based Linear Discriminant analysis* (SILD) [20] in kinship verification. They also used cosine similarity as distance [29] measure; moreover, they experimentally proved it to perform better than did Euclidean distance when applied to this metric learning problem.

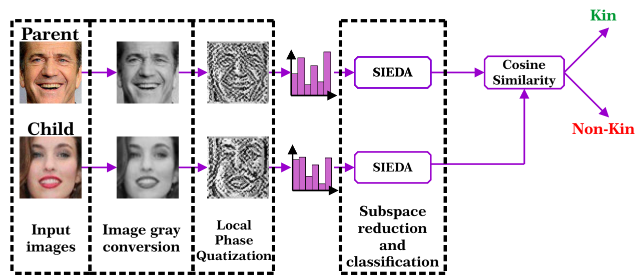

Figure 5: Verification framework of LPQ-SIEDA from University of Biskra & University of Valenciennes (Ref 1.J).

4 RELATED WORK

The problem of kinship recognition is fairly new, as research only dates back to 2010. Ever since, however, many experts have spent great efforts on this challenging problem [1, 3, 5–13, 15, 19, 22, 23, 25–28, 32–34, 37–41, 44–48, 50–55].

Looking back, a common trend emerged. With each new database for kinship recognition came a wave of attention and progress. Moreover, as the databases grew and evolved so did the technical approaches– the story of automatic kinship recognition follows that of its supportive data, where each database marked a major milestone in the journey from 2010 to now (*i.e.*, CornellKin [12] to FIW [33, 40]). Thus, to review the related work it is most natural to do so in reference to these milestones. We now review these critical moments in reverse chronological order.

Besides FIW, the last datasets released were KinWild I and II [28]. As done with FIW in RFIW, KinWild was used in a IEEE FG 2015 Data Challenge [26]. KinWild I and II support 4 pair types (*i.e.*, *parent-child*). KinWild I contains about 150 face pairs per each of the 4 categories, and KinWild II contains 250 face pairs per category. KinWild I and II differ in terms of the source of face images– KinWild I uses face pairs from the same source image, while KinWild II uses faces from different images). The larger of the two sets, KinWild II, has a total of 1,000 pairs and 4 different pair types– FIW has 644,000 pairs and 11 types, 4 provided to the vision community for the first time (*grandparent-grandchild*, which were left out of RFIW). In terms of diversity, [24] obtained state-of-art on KinWild using just simple similarity functions (*e.g.*, SSIM and color features) directly on the face pairs. From this, the authors made the supported claim that KinWild insufficiently represents the visual nature of kinship. Meaning it is unequipped to bridge the gap between research and reality– a hurdle overcome with FIW.

Nonetheless, KinWild has had a great impact, as many advances in kinship verification were made possible by it. Until FIW, KinWild was the largest dataset for verification.

One year prior, Family101 was released [11]. However, it was built to support a different task, family classification. Family101 consists of 101 non-overlapping families (*i.e.*, unique family trees) that include 607 different subjects. Like KinWild later did for verification, Family101 attracted much attention to what was then the new kin-based recognition problem (*i.e.*, family classification). The time has come to update our data resources to further advance automatic kinship recognition. This brings us to the goal of FIW:

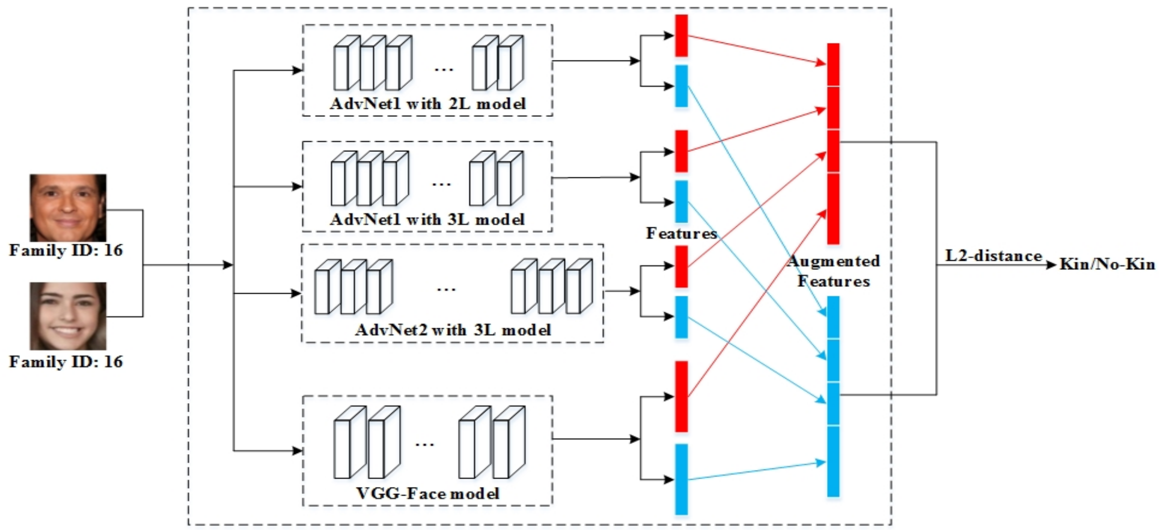

Figure 6: Framework used for kinship verification by team from Chongqing University (Ref 1.D).

collect family photos of 1, 000 unique family trees that span back a couple-to-several generations (*i.e.*, FIW consists of 1, 001 families). FIW contains nearly 6, 000 unique subjects from nearly 13, 000 family photos and, hence, each family averages around 6 members and 13 images each. A majority of these images are family photos– with over 30, 000 faces in total there is an average of about 3 relatives per photo. Note that only members of the respective family in the photo are counted, while all other faces are currently left-out of the image collection.

Stepping back another year, to 2012, UB KinFace [44] put emphasis on the effects of aging on kinship verification. The dataset includes 200 families, 400 people, and 600 images. The authors followed a transfer learning paradigm to transfer the knowledge learned from an easier task of "young parent – child" to the hard task "old parent – child". This work was later extended to the parsing of family photos using social cues found in photos.

Instead of pairs, TSKinFace [32] supported verification of kinship between two known parents and a single child. This was the first attempt to use faces of more than two subjects in kinship recognition. The goal then was to verify whether a child belonged to a pair of parents. An interesting perspective indeed, as several use-cases follow this setting. For instance, let us consider a scenario involving a child kidnapped and exploited online. If authorities stumbled upon this media, then vision technology could help identify the child. The problem here is it is unlikely that this young child is in any accessible database. However, chances are much higher that data on one or both parents are accessible. With the parents ID, the child's ID can then be inferred. In fact, this works with other family members as well. In essence, kinship serves as a powerful cue that greatly reduces the search space through family bonds. Now, this view to the problem can be explored, and on scales much larger than ever before, with FIW .

Finally, in 2010, the seminal work in [12] introduced CornellKin, which is made-up of parent-child pairs that were the first used to do kinship verification. Just like the others mentioned, [12] had great impact and was the first to publish on kinship recognition in the machine vision community.

Each major database released for kinship recognition proved critical to the advancement of this line of technology. Considering the incomparable size of FIW– in terms of the amount of data and types of labels– we expect great progress to had in the coming years. Not only will studies be conducted with sample sizes large enough to draw decisive conclusions, but the rich label information will pave the way to new ways of viewing, formulating and, thus, evaluating the problem of kinship recognition. In the end, FIW provides the data required to accurately reflect the true data distribution of families around the globe. Furthermore, vast collection of data enables deep learning models to be applied to kinship recognition like never before possible. To kick things off, we organized this data challenge as a somewhat constrained (*i.e.*, well defined) evaluation. We expect much more to come from FIW.

5 DISCUSSION

5.1 A Broader Impact

The size of FIW, along with the labels representing complex family structures of 1, 001 families, make it difficult to predict the exact future directions efforts on it will take (*i.e.*, many possibilities, each with high potential). For starters, improving upon existing benchmarks, as was the focus of RFIW. Also, adding new kin-based tasks and benchmarks could be next (*e.g.*, search & retrieval). Aside from evaluations, cross-discipline studies (e.g., nature-based) are too interesting directions. Even a more advanced analysis of the data making up FIW could also prove insightful, or, at the very least, help further characterize the data resource itself. The possible directions listed here are still far from exhausted. It is fair to say that– whether experimental, analytical, algorithmic, biological, practical– it is exciting to ponder on the many different directions others will imagine and propose once FIW is released.

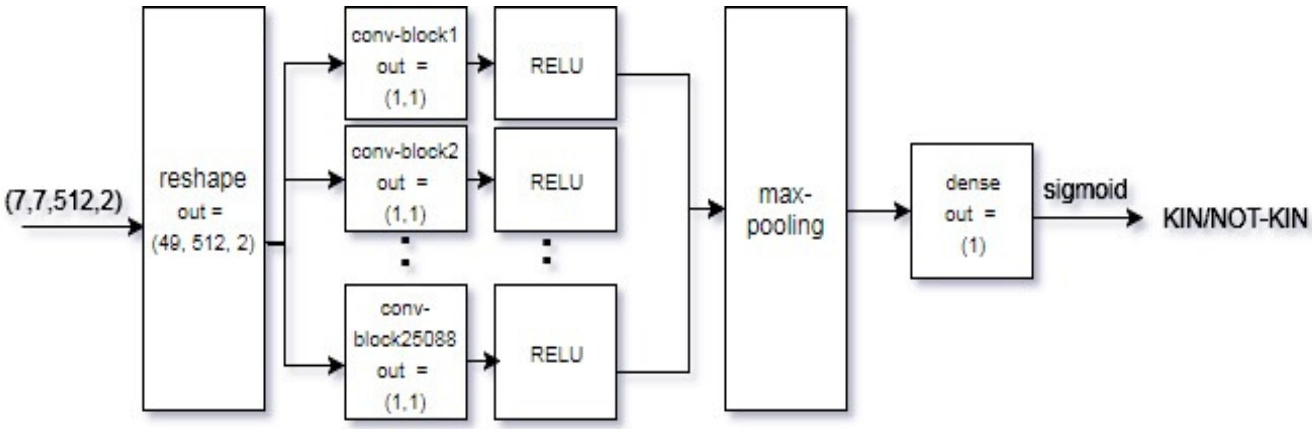

Figure 7: Modified convolution layer proposed by Team from Bar-Ilan University (Ref 1.F).

We anticipate that the Data Challenge Workshop itself will serve as a perfect opportunity to reflect on the current state of kinship recognition technologies, while also addressing questions concerning future directions and the next steps. We expect the following to spring up from the upcoming workshop during ACM MM 2017:

I. Cover the current state in kinship recognition;
II. Review data challenge submissions, hand out awards, and learn from top teams;
III. Discuss future directions for FIW and review aspects of it yet to be exploited (i.e., text metadata, new task evaluations, and more);
IV. Extend our list of practical uses of automatic kinship recognition technologies, and spread word to those focused on kinship-based technologies of this rich resource we have;
V. Attract newcomers to the problem, and also capture the interest of experts;
VI. Identify shortcomings of the evaluation to improve, along with what is working well;
VII. Establish cross-discipline studies between life sciences and machine vision communities;
VIII. Lay framework for the open-source community to join in the continued effort made possible with FIW and supporting SW tools that will be made public.

5.2 Conclusion

RFIW 2017 served as an ice-breaker for more big things to come for FIW and kinship recognition as a whole. All-in-all, the data challenge was a success. We established competitive benchmarks for both tasks. Additionally, source code for several of the methods reported in this paper will be available for download as part of the upcoming FIW API, with the hope such resources (to be introduced and demoed during workshop) will lessen the learning curve for newcomers, while still challenging the experts with the FIW database (*i.e.*, challenge the experts, while enabling newcomers). Thus, the story of FIW, as we see it, has only begun.

ACKNOWLEDGMENTS

We would like to thank the participants, PC Members, and Workshop Chairs for making RFIW possible by the great efforts of experts and rising scholars alike. Also, we would like to thank the organizers of ACM MM 2017 for both the opportunity to host RFIW Data Challenge Workshop.

This material is based upon work supported by the U.S. Department of Homeland Security, Science and Technology Directorate, Office of University Programs, under Grant Award 2013-ST-061-ED0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

REFERENCES

[1] Pendar Alirezazadeh, Abdolhossein Fathi, and Fardin Abdali-Mohammadi. 2015. A Genetic Algorithm-Based Feature Selection for Kinship Verification. *IEEE Signal Processing Letters* 22, 12 (2015), 2459–2463.
[2] Karsten M Borgwardt, Arthur Gretton, Malte J Rasch, Hans-Peter Kriegel, Bernhard Schölkopf, and Alex J Smola. 2006. Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics* 22, 14 (2006), e49–e57.
[3] Andrea Bottinok, Ihtesham Ul Islam, and Tiago Figueiredo Vieira. 2015. A multi-perspective holistic approach to Kinship Verification in the Wild. In *11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Vol. 2. 1–6.
[4] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)* 2, 3 (2011), 27.
[5] Xiaojing Chen, Le An, Songfan Yang, and Weimin Wu. 2015. Kinship verification in multi-linear coherent spaces. *Multimedia Tools and Applications* (2015), 1–18.
[6] Yunfei Chen, Hongwei Hu, Shujuan Cao, and Bo Ma. 2015. Sparse coding based kinship recognition. In *Multimedia Technology IV: Proceedings of the 4th International Conference on Multimedia Technology, Sydney, Australia, 28-30 March 2015*. CRC Press, 115.
[7] Afshin Dehghan, Enrique Ortiz, Ruben Villegas, and Mubarak Shah. 2014. Who do i look like? determining parent-offspring resemblance via gated autoencoders. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1757–1764.
[8] Hamdi Dibeklioglu, Albert Ali Salah, and Theo Gevers. 2013. Like father, like son: Facial expression dynamics for kinship verification. In *Proceedings of the IEEE International Conference on Computer Vision*. 1497–1504.
[9] Xiaodong Duan and Zheng-Hua Tan. 2015. A feature subtraction method for image based kinship verification under uncontrolled environments. In *IEEE International Conference on Image Processing (ICIP)*. 1573–1577.
[10] Xiaodong Duan and Zheng-Hua Tan. 2015. Neighbors Based Discriminative Feature Difference Learning for Kinship Verification. In *International Symposium*

on Visual Computing. Springer, 258–267.
[11] Ruogu Fang, Andrew C Gallagher, Tsuhan Chen, and Alexander Loui. 2013. Kinship classification by modeling facial feature heredity. In *2013 IEEE International Conference on Image Processing*. IEEE, 2983–2987.
[12] Ruogu Fang, Kevin D Tang, Noah Snavely, and Tsuhan Chen. 2010. Towards computational models of kinship verification. In *2010 IEEE International Conference on Image Processing*. IEEE, 1577–1580.
[13] Mohammad Ghahramani, Wei-Yun Yau, and Eam Khwang Teoh. 2014. Family verification based on similarity of individual family member's facial segments. *Machine Vision and Applications* 25, 4 (2014), 919–930.
[14] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.
[15] Yuanhao Guo, Hamdi Dibeklioglu, and Laurens van der Maaten. 2014. Graph-Based Kinship Recognition.. In *International Conference on Pattern Recognition*. Citeseer, 4287–4292.
[16] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. 2016. MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition. In *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III*. 87–102.
[17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*. 770–778.
[18] Alexander Hermans, Lucas Beyer, and Bastian Leibe. 2017. In Defense of the Triplet Loss for Person Re-Identification. *CoRR* abs/1703.07737 (2017).
[19] Junlin Hu, Jiwen Lu, Junsong Yuan, and Yap-Peng Tan. 2014. Large margin multi-metric learning for face and kinship verification in the wild. In *Asian Conference on Computer Vision*. Springer, 252–267.
[20] Meina Kan, Shiguang Shan, Dong Xu, and Xilin Chen. 2011. Side-Information based Linear Discriminant Analysis for Face Recognition. In *British Machine Vision Conference, BMVC 2011, Dundee, UK, August 29 - September 2, 2011. Proceedings*. 1–12.
[21] Diederik P. Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *CoRR* abs/1412.6980 (2014).
[22] Lu Kou, Xiuzhuang Zhou, Min Xu, and Yuanyuan Shang. 2015. Learning a Genetic Measure for Kinship Verification Using Facial Images. *Mathematical Problems in Engineering* 2015 (2015).
[23] Qingfeng Liu, Ajit Puthenputhussery, and Chengjun Liu. 2015. Inheritable Fisher vector feature for kinship verification. In *7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. IEEE, 1–6.
[24] Miguel Bordallo López, Elhocine Boutellaa, and Abdenour Hadid. 2016. Comments on the "kinship face in the wild" data sets. *IEEE transactions on pattern analysis and machine intelligence* 38, 11 (2016), 2342–2344.
[25] Jiwen Lu. 2012. Neighborhood Repulsed Metric Learning for Kinship Verification. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (CVPR '12)*. IEEE Computer Society, Washington, DC, USA, 2594–2601. http://dl.acm.org/citation.cfm?id=2354409.2354945
[26] Jiwen Lu, Junlin Hu, Venice Erin Liong, Xiuzhuang Zhou, Andrea Bottino, Ihtesham Ul Islam, Tiago Figueiredo Vieira, Xiaoqian Qin, Xiaoyang Tan, Songcan Chen, et al. 2015. The FG 2015 kinship verification in the Wild Evaluation. In *11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Vol. 1.
[27] Jiwen Lu, Junlin Hu, Xiuzhuang Zhou, Jie Zhou, Modesto Castrillón-Santana, Javier Lorenzo-Navarro, Lu Kou, Yuanyuan Shang, Andrea Bottino, and Tiago Figuieiredo Vieira. 2014. Kinship verification in the wild: The first kinship verification competition. In *International Joint Conference on Biometrics (IJCB)*. IEEE, 1–6.
[28] Jiwen Lu, Xiuzhuang Zhou, Yap-Pen Tan, Yuanyuan Shang, and Jie Zhou. 2014. Neighborhood repulsed metric learning for kinship verification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36, 2 (2014), 331–345.
[29] Hieu V. Nguyen and Li Bai. 2010. Cosine Similarity Metric Learning for Face Verification.. In *ACCV (2)*, Vol. 6493. Springer, 709–720.
[30] Abdelmalik Ouamane, Messaoud Bengherabi, Abdenour Hadid, and Mohamed Cheriet. 2015. Side-Information based Exponential Discriminant Analysis for face verification in the wild. In *11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2015, Ljubljana, Slovenia, May 4-8, 2015*. 1–6.
[31] Omkar M. Parkhi, Andrea Vedaldi, and Andrew Zisserman. 2015. Deep Face Recognition. In *Proceedings of the British Machine Vision Conference 2015, BMVC 2015, Swansea, UK, September 7-10, 2015*. 41.1–41.12.
[32] Xiaoqian Qin, Xiaoyang Tan, and Songcan Chen. 2015. Tri-Subject Kinship Verification: Understanding the Core of A Family. *CoRR* abs/1501.02555 (2015). http://arxiv.org/abs/1501.02555
[33] Joseph P Robinson, Ming Shao, Yue Wu, and Yun Fu. 2016. Families in the Wild (FIW): Large-Scale Kinship Image Database and Benchmarks. In *Proceedings of the 24th ACM international conference on Multimedia*. ACM, 242–246.
[34] J. Luo S. Xia, M. Shao and Y. Fu. 2012. Understanding Kin Relationships in a Photo. *IEEE Transactions on Multimedia* 14, 4 (2012), 1046–1056.
[35] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 815–823.
[36] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. FaceNet: A Unified Embedding for Face Recognition and Clustering. *CoRR* abs/1503.03832 (2015).
[37] Gowri Somanath and Chandra Kambhamettu. 2012. Can faces verify blood-relations?. In *IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 105–112.
[38] Tiago F Vieira, Andrea Bottino, and Ihtesham Ul Islam. 2013. Automatic verification of parent-child pairs from face images. In *Iberoamerican Congress on Pattern Recognition*. Springer, 326–333.
[39] Tiago F Vieira, Andrea Bottino, Aldo Laurentini, and Matteo De Simone. 2014. Detecting siblings in image pairs. *The Visual Computer* 30, 12 (2014), 1333–1345.
[40] Shuyang Wang, Joseph P Robinson, and Yun Fu. [n. d.]. Kinship Verification on Families in the Wild with Marginalized Denoising Metric Learning. In *12th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*.
[41] Xiaolong Wang and Chandra Kambhamettu. 2014. Leveraging appearance and geometry for kinship verification. In *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 5017–5021.
[42] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. 2016. *A Discriminative Feature Learning Approach for Deep Face Recognition*. Springer International Publishing, Cham, 499–515.
[43] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. 2016. A Discriminative Feature Learning Approach for Deep Face Recognition. In *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VII*. 499–515.
[44] Siyu Xia, Ming Shao, and Yun Fu. 2011. Kinship verification through transfer learning. In *IJCAI Proceedings-international joint conference on artificial intelligence*, Vol. 22. 2539.
[45] Siyu Xia, Ming Shao, Jiebo Luo, and Yun Fu. 2012. Understanding kin relationships in a photo. *IEEE Transactions on Multimedia* 14, 4 (2012).
[46] Min Xu and Yuanyuan Shang. 2016. Kinship Verification Using Facial Images by Robust Similarity Learning. *Mathematical Problems in Engineering* 2016 (2016).
[47] Haibin Yan, Jiwen Lu, Weihong Deng, and Xiuzhuang Zhou. 2014. Discriminative multimetric learning for kinship verification. *IEEE Transactions on Information forensics and security* 9, 7 (2014).
[48] Haibin Yan, Jiwen Lu, and Xiuzhuang Zhou. 2015. Prototype-based discriminative feature learning for kinship verification. *IEEE transactions on cybernetics* 45, 11 (2015), 2535–2545.
[49] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. 2014. Learning face representation from scratch. *arXiv preprint arXiv:1411.7923* (2014).
[50] Li Zhang, KengTeck Ma, Hossein Nejati, Lewis Foo, Terence Sim, and Dong Guo. 2014. A talking profile to distinguish identical twins. *Image and Vision Computing* 32, 10 (2014), 771–778.
[51] Ziming Zhang, Yuting Chen, and Venkatesh Saligrama. 2015. Group membership prediction. In *Proceedings of the IEEE International Conference on Computer Vision*. 3916–3924.
[52] Kaihao Zhang12, Yongzhen Huang, Chunfeng Song, Hong Wu, Liang Wang, and Statistical Machine Intelligence. [n. d.]. Kinship Verification with Deep Convolutional Neural Networks. ([n. d.]).
[53] Xiuzhuang Zhou, Junlin Hu, Jiwen Lu, Yuanyuan Shang, and Yong Guan. 2011. Kinship verification from facial images under uncontrolled conditions. In *Proceedings of the 19th ACM international conference on Multimedia*. ACM, 953–956.
[54] Xiuzhuang Zhou, Jiwen Lu, Junlin Hu, and Yuanyuan Shang. 2012. Gabor-based gradient orientation pyramid for kinship verification under uncontrolled environments. In *Proceedings of the 20th ACM international conference on Multimedia*. ACM, 725–728.
[55] Xiuzhuang Zhou, Yuanyuan Shang, Haibin Yan, and Guodong Guo. 2016. Ensemble similarity learning for kinship verification from facial images in the wild. *Information Fusion* 32 (2016), 40–48.