UNIVERSITY OF AMSTERDAM

# UvA-DARE (Digital Academic Repository)

## Kinship similarity for open sets

Wang, W.; You, S.; Karaoglu, S.; Gevers, T.

**DOI**
10.1016/j.patcog.2023.110123

**Publication date**
2024

**Document Version**
Final published version

**Published in**
Pattern Recognition

**License**
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)

Link to publication

**Citation for published version (APA):**
Wang, W., You, S., Karaoglu, S., & Gevers, T. (2024). Kinship similarity for open sets. *Pattern Recognition*, *148*, Article 110123. https://doi.org/10.1016/j.patcog.2023.110123

**General rights**
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)*

Download date:14 Nov 2026

Pattern Recognition 148 (2024) 110123

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

# Kinship similarity for open sets

Wei Wang [a,*], Shaodi You [a], Sezer Karaoglu [a,b], Theo Gevers [a,b]

[a] Computer Vision Lab, University of Amsterdam, Science Park 904, Amsterdam, 1098XH, The Netherlands
[b] 3DUniversum, Amsterdam, The Netherlands

ARTICLE INFO

*Keywords:*
Kinship similarity measurement
Kinship recognition
Open set

ABSTRACT

The aim of image-based kinship recognition methods is to determine the genetic relationship between people from their face images. Kinship recognition and related methods have many applications in computer vision. Most of the methods are very meaningful but are focused on closed sets. However, for real-life scenarios, kinship recognition is an open set problem.

In contrast to previous methods, this paper considers kinship recognition for open sets. Further, the aim is to determine family relationships and their corresponding degrees of kinship. Our method is pairwise-based and is able to exploit mutual information from positive pairs.

Large scale experiments and ablation studies show that our method (1) reaches SOTA performance on the FIW dataset, (2) is able to properly separate kinship categories using pairwise similarity and (3) generates uniform similarity distributions.

## 1. Introduction

Image-based kinship recognition [1] aims at determining the genetic relationship between people by analyzing images of their faces. In recent years, more and more attention has been focused on kinship related tasks such as kinship verification, family recognition, and kinship identification [2] due to many applications such as searching missing children [3], family album organization, forensic investigation [4], crime scene investigation [5], social media [6] and behavior analysis [7], historical and genealogical research [8], and automatic image annotation [9].

The recognition of kinship relationships between people as a hierarchical (tree) structure is an important task because people may share origins at different degrees. The genetic (kinship) relationship between two people is the amount of DNA they have in common because they are related. According to the average percent of DNA shared between relatives,[1] some kinship relations may have similar gene overlap. As illustrated in Fig. 1, family members have various degrees of kinship relationships. For example, the boy has 50% genetic origin with his (full) father and mother. In addition, the genetic relationship between the boy and his (full) brother is 0.5, as they share an average of 50% origins of their DNA. The boy and his grandmother share 25% DNA. Hence, genetic relationships between family members are determined by a family tree. Obviously, outside the family, there exists a large set of people who have no genetic origin at all (or very far) with the family members. Hence, kinship recognition is by definition an open set problem for real-life scenarios. However, the recognition of different degrees of kinship in open set scenarios has largely been ignored so far.

Therefore, this paper focuses on kinship at various degrees for open collections (*i.e.* including kin and non-kin related people). The aim is to determine family relationships and their corresponding degrees of kinship hierarchically. To this end, we propose a novel and more general task called the Open-set Kinship Similarity Measurement (OKSM). As illustrated in Fig. 1, similarities between images of faces of people are derived to measure different degrees of kinship relationship *i.e.* the first-degree relationship between parent and child, the second degree between grandparent and grandchild, *etc*. The more genetically related people are, the higher the similarity. Furthermore, current open set classification methods are mainly focused on distinguishable classes (not pairs), and most of the benchmark datasets are image-based [10]. In contrast, our method is pairwise-based and exploits mutual information from positive pairs by merely re-matching them. Moreover, we propose a data-driven strategy to compute hierarchical kinship degrees. Firstly, our method assigns images (Primary-Kin) into self-self (ego-self), kin-related (containing 50% genetics), and known-negative (unrelated) pairs. Then, a hierarchical kinship triplet loss is proposed to learn the similarity of the different pairs. Image pair features are projected onto a distance-based feature space. The distance measure is used to differentiate between primary kin, secondary kin, and non-kin relationships.

The main contributions of this paper are summarized as follows:

* Corresponding author.
*E-mail address:* w.wang@uva.nl (W. Wang).

[1] https://isogg.org/wiki/Autosomal_DNA_statistics.

https://doi.org/10.1016/j.patcog.2023.110123
Received 9 January 2023; Received in revised form 31 October 2023; Accepted 14 November 2023
Available online 17 November 2023
0031-3203/© 2023 Published by Elsevier Ltd.

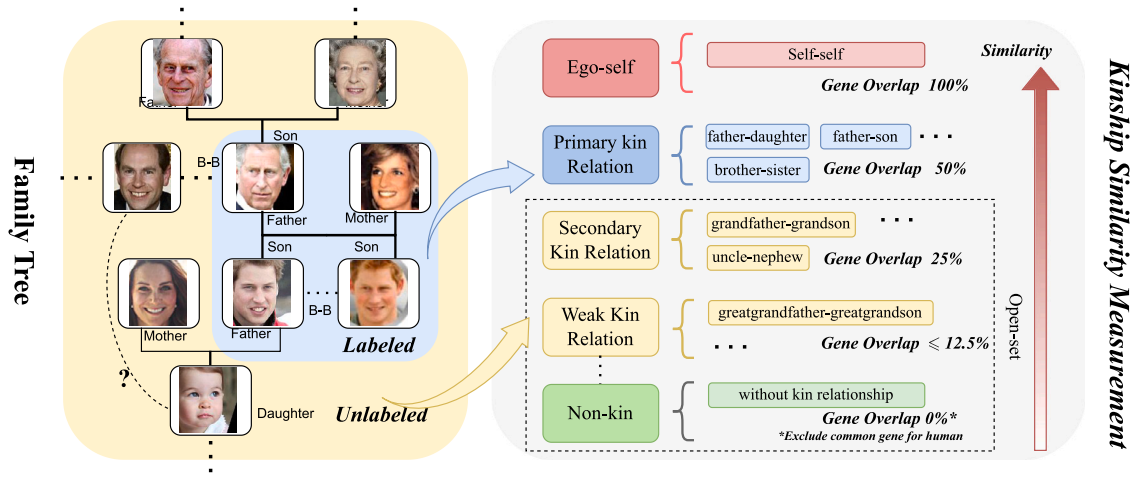



**Fig. 1.** The proposed open set kinship recognition task. The open set contains both (1) family members (having various degrees of kinship) and (2) non-kin persons. Kinship recognition is by definition an open set problem for real-life scenarios.

1. Kinship similarity is introduced for open sets.
2. A method is proposed to derive pairwise information by re-matching known (positive) pairs.
3. A hierarchical kinship network is proposed to distinguish primary kin, secondary kin, and non-kin relationships for open-set kinship scenarios.
4. The proposed method outperforms state-of-the-art methods on the FIW dataset.

## 2. Related work

### 2.1. Kinship recognition and related tasks

Image-based kinship recognition receives more and more attention in the computer vision community [1,11,12]. Kinship recognition typically includes kinship verification, kinship identification, and family recognition [2]. Kinship verification is a binary classification problem, determining whether two persons, represented by image pairs of their faces, are kin related or not [13]. Different methods are proposed in the field of kinship verification [9,14–24]. Kinship verification has limited applicability because it is not able to distinguish different kin-types. In contrast, kinship identification aims to determine the different kinship types [2]. For example, Guo et al. [16] propose a graph-based method recognizing the type of kinship using pairwise kinship information. Wang et al. [25] propose a deep kinship recognition framework to predict kin relationships using family trees. Different feature modalities are used such as kin-or-not, gender, and relative age attributes. Recently, Wang et al. propose a kinship identification method based on joint learning [26]. The method is specifically designed for pairwise kinship identification. However, the above kinship verification and identification methods are used for closed sets. Kinship recognition is an open set problem in real-life scenarios as many people are not (directly) kin related. Therefore, in this paper, we propose a new and more general open-set kinship similarity framework.

### 2.2. Open-set recognition

Open-set recognition (OSR) is the problem of properly handling unknown samples during the classification of the known ones [27–30]. Based on traditional machine learning methods, previous OSR methods utilize SVM [31], extreme value machines (EVM) [32], and sparse representation-based methods [33]. Lately, deep neural network-based methods are proposed. An OpenMax layer is proposed by Bendale and Boult [34] to circumvent the problem of the Softmax cross-entropy loss [10]. The method incorporates the likelihood of the recognition failures and adopts the concept of Meta-Recognition [35]. In addition, generative models are proposed [36,37]. For example, the G-OpenMax algorithm [36] uses synthesized (fake) unknown samples. Neal et al. [38] generate known-negative images as counterfactual images to improve the robustness of their model. However, the performance can be negatively affected by the (poor) quality of the generated (fake) images. Recently, the class anchor clustering (CAC) loss is proposed [39] by clustering known classes into predefined clusters. Unlike the OpenMax method, the CAC loss predefines anchors of known classes without updating them during training. More recently, Chen et al. [30] propose Reciprocal Point Learning (RPL) utilizing reciprocal points representing extra-classes for one specific class. The method is extended to Adversarial Reciprocal Point Learning (ARPL) by adding an instantiated adversarial enhancement process [29]. The method obtains state-of-the-art performance for different datasets [40]. None of the open-set methods is used for kinship recognition. Since our new task is an open-set problem, we propose a new pairwise open-set method for kinship similarity.

## 3. Problem formulation and comparison

### 3.1. Problem formulation

We formulate the open-set kinship problem as a classification problem. Face images are pre-processed in a pairwise form $[(I_A, I_B), y]$, where $I_A$ and $I_B$ denote two different images with kinship label $y$. In this way, a dataset $\{[(I_{A_i}, I_{B_i}), y_i] \mid i = 1, 2, \dots, n\}$ is obtained where $i$ is the index to the paired sample and $y$ indicates the type of kinship. Further, a feature embedding neural network $\phi$ is used to transform image pairs to vector pairs, *i.e.*, $(\phi(I_A), \phi(I_B)) = (\mathbf{x}_A, \mathbf{x}_B)$. $D = \{[(\mathbf{x}_{A_i}, \mathbf{x}_{B_i}), y_i] \mid i = 1, 2, \dots, n\}$ is the resulting dataset.

Based on kinship similarity, the samples are categorized into three classes $\{P, S, N\}$, where $P$, $S$ and $N$ represent the Primary-kin ($P$), Secondary-kin ($S$), and Non-kin ($N$) relationships respectively. These three classes contain different kinship types with label values $Y = \{\underbrace{1, 2, \dots, 7}_{P}, \underbrace{8, \dots, 11}_{S}, \underbrace{12}_{N}\}$. Specifically, $1, 2, \dots, 12$ in $Y$ represent Father–Daughter (F–D), Father–Son (F–S), Mother–Daughter (M–D), Mother–Son (M–S), Brother–Brother (B–B), Brother–Sister (B–S), and Sister–Sister (S–S), Grandfather–Granddaughter (GF–GD), Grandfather–Grandson (GF–GS), Grandmother–Granddaughter (GM–GD), Grandmother– Grandson (GM–GS), and pairs without kinship respectively. Since most existing kinship datasets only contain $1-7$ different kinship types, only samples with label $P$ are used for training, and samples with labels

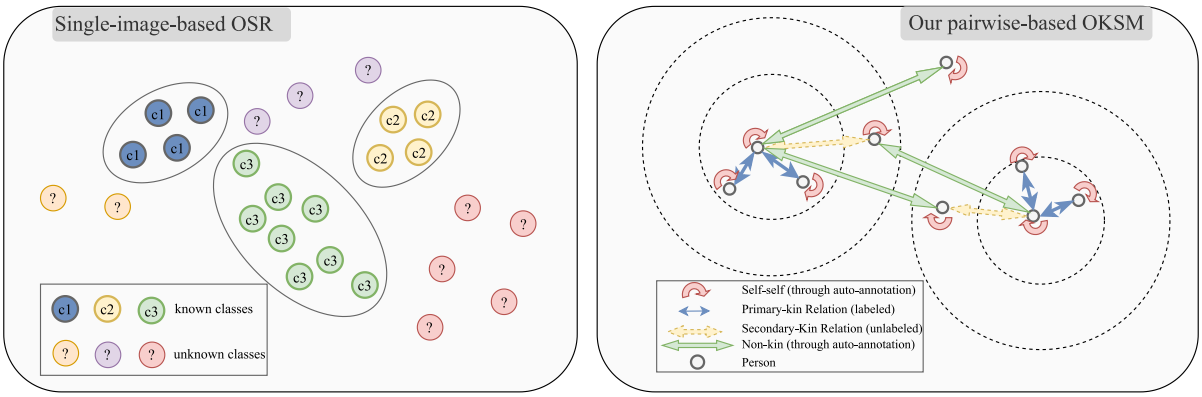



**Fig. 2.** *Left*: General OSR methods [10] are single-based and known samples are independent of each other. *Right*: Our OKSM task is pairwise-based. Labels denote the relationships between image pairs.

$P, S, N$ are used for testing. In our settings, a measurement model is used to compute the kinship similarity of three classes. Similarity distance $d$ is used to distinguish $P$, $S$, and $N$.

### 3.2. Comparison with kinship recognition

Our problem is to determine three categories $P$, $S$, and $N$ by computing kinship-related information from pairs. Different from existing kinship recognition, our approach yields kinship similarity types. The aim of existing kinship recognition is to bring positive pairs together and negative pairs apart. However, this strategy may have a negative influence on the computation of hierarchical kinship relationships as it may lead to incorrect predictions of similarity types of the test samples. Therefore, our approach focuses on measuring the hierarchical kin similarities of pairs depending on their genetic sharing.

The Softmax Cross-Entropy is used as the standard loss in deep learning-based kinship recognition or related kinship tasks. Denoting $\mathbf{z}$ as the extracted logits from a neural network, then the Softmax Cross-Entropy loss is defined by:

$$L = -\frac{1}{N}\sum_{i}^{N}\log\left(\frac{e^{\mathbf{z}_i}}{\sum_j e^{\mathbf{z}_j}}\right), \tag{1}$$

where $\mathbf{z}_j$ represents the $j$th element of the logits, and $\mathbf{z}_i$ is the target logit [41] of the ground truth. $N$ is the number of training samples. However, the close-set Softmax is not able to properly handle unknown samples because:

1. The Softmax Cross Entropy loss is not injective [42]. It cannot guarantee a proper clustering behavior [39].
2. The Softmax Cross Entropy [10] inherently has a closed set nature and can easily be misled by unknown samples [34].

On the contrary, our approach is specifically designed for open-set scenarios.

### 3.3. Comparison with open-set recognition

Open-set methods are specifically designed to deal with unknown samples. However, current open set methods mostly focus on the classification of (single) classes [10]. As illustrated in Fig. 2, the current open-set recognition approach mainly focuses on learning representative features for each known class and is single-image-based. The label information mainly lies in the feature of the single input. However, our open-set kinship approach is pairwise based. The classification is based on the mutual relation between pairs. Hence, our model computes corresponding feature relationships among pairs. Moreover, our pairwise kinship model inherently obtains extra information by merely re-matching the known positive pairs. The right side of Fig. 2 shows that after our data-driven strategy, the self-self and non-kin relationships can be auto-annotated and utilized for hierarchical feature representation. Standard open-set recognition methods consider two separate classes *i.e.* known $K$ and unknown $U$. Let the open set be denoted by $\mathcal{O}_k = S^o - S_k$; $S^o$ is the overall measure space [10], and $S_k$ is the embedding space [29] of $K$, then $\mathcal{O}_k^U$ is the unknown space of $U$, the open space risk can be quantitatively described by:

$$\mathcal{R}_o\left(\psi_k, \mathcal{O}_k^U\right) = \frac{\int_{\mathcal{O}_k^U}\psi_k(x)dx}{\int_{S_k\cup\mathcal{O}_k}\psi_k(x)dx}, \tag{2}$$

where $\psi_k(x)$ is a binary measurable function. However, our open-set kinship problem is to separate $P$, $S$ and $N$. Hence, following the previous notion [10,29], we re-define our open-set kinship measurement risk $\mathcal{R}_{ko}$ as follows:

$$\mathcal{R}_{ko}\left(\psi, \mathcal{O}_k^{U_N}\right) = \frac{\int_{\mathcal{O}_k^U}\psi_k(x)dx}{\int_{S_k\cup\mathcal{O}_k}\psi_k(x)dx} + \frac{\int_{\mathcal{O}_k^{U_N}}\psi_{u_k}(x)dx}{\int_{S_k\cup\mathcal{O}_k}\psi_{u_k}(x)dx}, \tag{3}$$

where $\psi_{u_k}(x) = 1$ indicates that $U_N$ is regarded as $U_k$. $U_N$ is the measure space for $N$ samples and $U_k$ is the measure space for $S$ samples. Hereby, the definition of our proposed Open-set Kin- ship Similarity Measurement (OKSM) is as follows:

**Definition.** Let $D$ be the training data, and let $\mathcal{R}_{ko}$, $\mathcal{R}_\varepsilon$ denote the open-set kinship risk and the empirical rick. The goal of Open-set Kinship Similarity Measurement is to find a suitable recognition function $\psi$, and $\psi$ is defined by minimizing the following risk:

$$\arg\min_{\psi}\left\{\mathcal{R}_{ko}(\psi) + \alpha R_\varepsilon(\psi(D))\right\}, \tag{4}$$

where $\alpha$ is a regularization parameter.

## 4. Methodology

The aim is to determine family relationships and their corresponding degrees of kinship hierarchically. To this end, we now introduce the architecture of the Hierarchical Kinship Measurement Network (HKMNet). The proposed hierarchical kinship network is a new 6-branch architecture integrating a modified *fc* layer, *tanh* layers and similarity values $\mathbf{c}$. As illustrated in Fig. 3, a data-driven strategy is adopted to extract hierarchical similarity from positive samples. Then, feature similarities of different hierarchical pairs are learned by our hierarchical kinship triplet loss. The final classification of test samples is dependent on the similarity measurements.

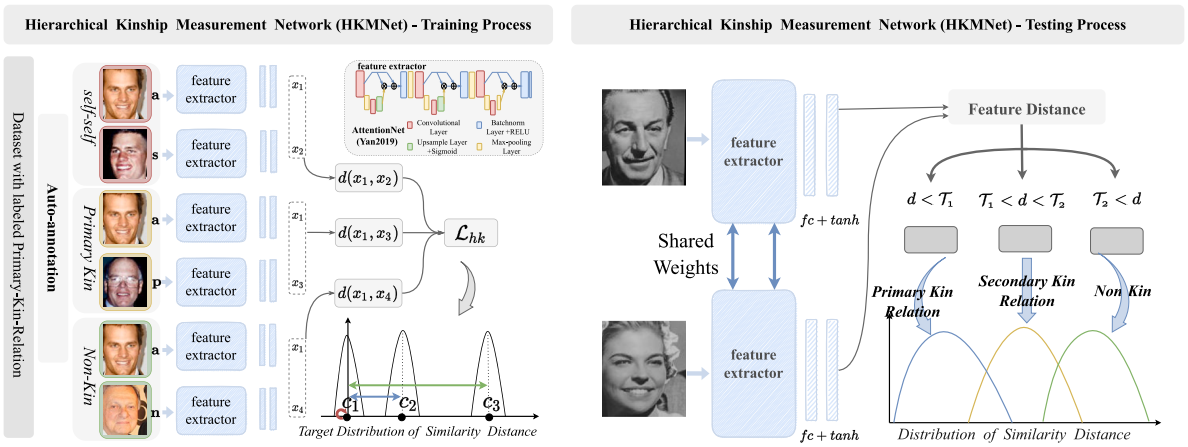



**Fig. 3.** Our Hierarchical Kinship Measurement Network (HKMNet).

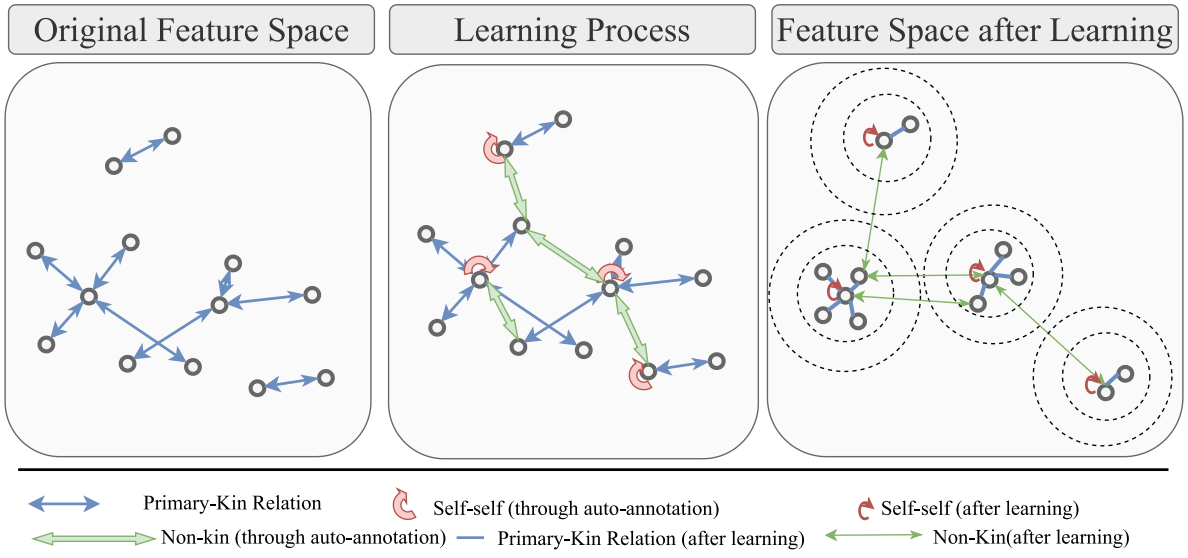



**Fig. 4.** Feature learning pipeline of our proposed Hierarchical Kin-ship Measurement Network (HKMNet).

### 4.1. Hierarchical information

Different from person re-identification or face retrieval, our new task predicts multiple classes. We only use Primary-Kin pairs as training samples. Secondary-Kin pairs are used as testing subsets for the open-set environment. Pairwise images inherently contain known-negative pairs of information. For example, as depicted in Fig. 4, during training, known positive pairs are separated and re-matched into new pairs. Pairs containing two images of the same person form self-self pairs. Pairs containing known-positive images are denoted by known-positive (Primary-Kin relation) pairs, and pairs containing two images from two different families correspond to known-negative (Non-kin relation) pairs. Based on genetic similarity, self-self (ego-self) pairs correspond to the highest similarity (100% genetic overlap). Known-positive samples (Primary-Kin containing father–son, mother–son, father–daughter, mother–daughter) relates to medium similarity (around 50% genetic overlap), and known-negative (Non-kin) pairs correspond to the lowest similarity. In this paper, the auto-annotation process $M(\cdot)$ is formulated as follows:

$$\{(\mathbf{a}, \mathbf{s}, \mathbf{p}, \mathbf{n})_i | i = 1, 2, \ldots, n\} = M(\{(I_{A_i}, I_{B_i}) | i = 1, 2, \ldots, n\}), \tag{5}$$

where the $i$th tuple $(\mathbf{a}, \mathbf{s}, \mathbf{p}, \mathbf{n})_i$ represent the anchor image of the person, the image of the same person, the image of a person with a known positive class, and the image from a different family, respectively. The total number $n$ of re-generated sets is the same as the previous known positive sets.

### 4.2. Distance-based losses

After generating new sets from known-positive (Primary-Kin) pairs, the images are given to the feature extractor. The extracted features are then projected into a kinship-feature space. To learn hierarchical similarities instead of separating "positive" and "negative" samples, the Hierarchical Kinship Triplet Loss is used.

#### 4.2.1. Hierarchical kinship triplet loss

Since the same kinship category (Primary-Kin, Secondary-Kin, *etc.*) has the same gene overlap,[2] the kinship similarity is the same for these kinship categories. We first pre-set the similarity value $\mathbf{c} = \{c_1, c_2, c_3\}$ for each training pair. Obviously, the feature distance of self-self pairs ($c_1$) should be smaller than that of known-positive pairs ($c_2$). And the feature distance of known-positive pairs ($c_2$) should be smaller than

[2] https://customercare.23andme.com/hc/en-us/articles/212170668-Average-Percent-DNA-Shared-Between-Relatives.

**Algorithm 1:** Pipeline of the Hierarchical Kinship Measurement Network.

**Input:** Image pairs $(I_{A_i}, I_{B_i}) = \{(I_{A_i}, I_{B_i}) \mid i = 1, 2, \ldots, n\}$
**Output:** *Primary-Kin, Secondary-Kin, Non-Kin*
**Training** *(only using Primary-Kin)*
**while** $\mathcal{L}_{hk} > \Delta$ **do**
  batch images = $\{(\mathbf{a}, \mathbf{s}, \mathbf{p}, \mathbf{n})_i | i = 1, 2, \ldots, N\}$, N = batch size;
  **if** do hard example mining ($\mathcal{H}_\mathcal{E}$) **then**
    *Concatenate* (batch images, hard images);
    calculate $\mathcal{L}_{hk}$;
    select hard images with smaller values of $d(x_1, x_4) - d(x_1, x_3)$;
  **else**
    calculate $\mathcal{L}_{hk}$;
  **end**
**end**
**Testing**
**if** $d(I_{A_i}, I_{B_i}) > \mathcal{T}_1$ **then**
  **if** $d(I_{A_i}, I_{B_i}) < \mathcal{T}_2$ **then**
    **Return** Secondary-Kin ($S$)
  **else**
    **Return** Non-Kin ($N$)
  **end**
**else**
  **Return** Primary-Kin ($P$)
**end**

the known-negative (Non-kin) pairs ($c_3$). To learn such hierarchical information, the hierarchical kinship triplet loss is defined as follows:

$$\mathcal{L}_{hk} = \sum_{i=1}^{3} \|d(x_1, x_{i+1}) - c_i\|^2 + \sum_{i=1}^{2} \mathcal{L}_t(x_1, x_{i+1}, x_{i+2}), \tag{6}$$

where the distance function is given by $d(\cdot)$, and $x_1, x_2, x_3, x_4$ representing the features of $\mathbf{a}, \mathbf{s}, \mathbf{p}, \mathbf{n}$ respectively. $\mathcal{L}_t$ is the triplet loss represented by:

$$\mathcal{L}_t(x_a, x_p, x_n) = max\{d(x_a, x_p) - d(x_a, x_n) + m, 0\}, \tag{7}$$

where $m$ is a margin, and $d(\cdot)$ is the Euclidean distance. $x_a, x_p, x_n$ represent the anchor, positive, and negative feature of a specific person. Different from previous triplet pairs (anchor, pos, neg) [43], our triplet loss uses the pairs with kin-based hierarchical information (self-self, known-positive, known-negative). The full pipeline is given in Algorithm 1.

## 5. Experiment

### 5.1. Datasets

In order to construct hierarchical similarity levels that effectively represent the open-set scenario, kinship datasets with various similarity levels (*e.g.*11 kin types) should be used. The Families In the Wild (FIW) [2] dataset is used as the training and testing dataset. FIW is the largest publicly available image-based dataset for the kinship-related tasks. For our task, we downloaded the newest version (*rfiw2020*) from the official website. The images of FIW are collected from the Internet. The dataset contains 11 193 unconstrained family photos of 1000 families. These images are processed into 418 060 pairs with 11 kinship types. Since our task is new, the settings of training and testing datasets are different from previous settings. We follow the 5-split protocol and fit the dataset into our task.[3] We only labeled the Primary-Kin images. The training set contains one subset:

- Primary-Kin set, which only contains direct kinship types.

Note that the Secondary-Kin images are not labeled and hence the problem becomes an open set problem. One of our tasks is to recall unlabeled data containing kinship information more effectively. Consequently, the Secondary-Kin set is not used for training and the training set corresponds to Primary-Kin label containing seven or four known positive kinship types: F–D, F–S, M–D, M–S, B–B, B–S, and S–S. These seven types are the mostly used kinship types in the literature. The testing set contains three subsets:

- Primary-Kin set, which contains known kinship types.
- Secondary-Kin (Unlabeled but kin-related) set, containing four kin type pairs (GF–GS, GF–GD, GM–GS, GM–GD) as unknown and unlabeled samples in real scenarios.
- Non-Kin set, which contains unknown pairs without kinship types.

More information (images and code) is anonymously available at our Github.

### 5.2. Experimental settings

During data prepossessing, the images are resized to $64 \times 64$. Only horizontal flipping is applied for the purpose of data augmentation. Our proposed HKMNet is a plug-and-play network which is suited for different existing neural networks. In this section, AttentionNet [44] is used as the backbone. Experimental results with other feature extractors are provided in the supplementary material. The $fc$ layer is shared-weighted and the *tanh* activation function is added after the $fc$ layer. The length of the features $(x_1, x_2, x_3, x_4)$ are $1 \times 64$. Our HKMNet model is trained with SGD. The learning rate is set to be 0.001 and the epoch to be 40. During training, the hierarchical kinship triplet loss is calculated and hard examples ($\mathcal{H}_\mathcal{E}$) are reused. Margin $m$ in $\mathcal{L}_t(x_1, x_2, x_3)$ is set to be 0.5. Further, $m$ in $\mathcal{L}_t(x_1, x_3, x_4)$ is set to be 3. The similarity values $c_1, c_2, c_3$ are set to be 0, 1, and 4 respectively.

#### 5.2.1. Comparison

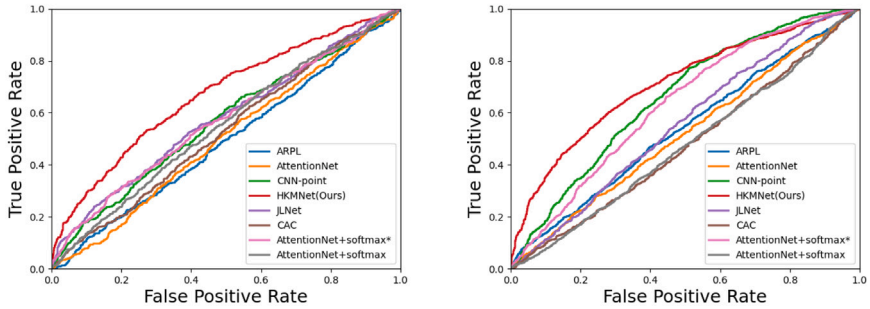
Unfortunately, publicly available codes are very limited in kinship recognition. Note that none of the existing methods can directly be applied to our new task. Therefore, all the compared methods are adapted to the new task at hand by changing their outputs accordingly.

To evaluate the performance of our proposed method, we compare our method with state-of-the-art methods in two different domains: (1) open-set recognition and (2) kinship-recognition. The CAC method [39] is one of the best performing distance-based open-set recognition methods. It explicitly trains known classes to form clusters around anchored class-dependent centers in logit space. ARPL [29] is the state-of-the-art approach of open-set recognition tasks for many datasets. RPL [30] and GCPL [45] are the well known open set recognition methods. It models the unexploited extra-class space with the concept of Reciprocal Point and uses an instantiated adversarial enhancement process.

For a fair comparison, networks are used with the same backbone [44]. AttentionNet + $Softmax$ is created following [39]. It uses kin types in the training set directly as classes. Moreover, two variations are added: AttentionNet+$Softmax$* and AttentionNet*.AttentionNet +$Softmax$* utilizes non-kin samples during training. AttentionNet* uses the same auto-annotated label as our HKMNet model.

We also compare our method with related kinship recognition methods. CNN-point [7] is selected as a representative of deep learning methods for kinship verification. JLNet is a pairwise-based kinship identification model. Ustc-nelslip [46] is a deep siamese neural network designed to fuse the features of two inputs for kinship verification. It utilizes kinship verification ensembles to enhance kinship identification

[3] We signed and strictly follow the 'Social Safety Support Guide and Rules of Code of Conduct' for using the human data.




(a) ROC curves (Secondary-Kin vs Non-kin) trained on 7 kin types

(b) ROC curves (Primary-Kin vs Unlabeled) trained on 7 kin types

**Fig. 5.** ROC curves of the different methods for FIW.

**Table 1**
Kinship similarity performance of 5-cross-validation on the FIW dataset. Here 7 kin-types (F–D, F–S, M–D, M–S, B–B, B–S, S–S) in Primary-Kin category are used for training. The Unlabeled includes both Secondary-Kin and Non-Kin.

| Methods | | Accuracy | AUROC Primary-Kin vs. Unlabeled | AUROC Secondary-Kin vs. Non-Kin | F1 |
|---|---|---|---|---|---|
| Same-backbone | AttentionNet+$Softmax^*$ [44] | 0.3752 | 0.6110 | 0.5780 | 0.3488 |
| | AttentionNet+$Softmax$ [44] | 0.3500 | 0.4842 | 0.5156 | 0.2835 |
| Open-set recognition | CAC [39] | 0.3084 | 0.4893 | 0.5195 | 0.2264 |
| | ARPL [29] | 0.3636 | 0.5388 | 0.4717 | 0.2590 |
| | RPL [30] | 0.3173 | 0.5127 | 0.4633 | 0.2639 |
| | GCPL [45] | 0.3510 | 0.4829 | 0.5143 | 0.2581 |
| Kinship related methods | CNN-point [7] | 0.4266 | 0.6413 | 0.5865 | 0.3832 |
| | AttentionNet* [44] | 0.2926 | 0.5182 | 0.4802 | 0.2242 |
| | JLNet [26] | 0.3807 | 0.5689 | 0.5539 | 0.3494 |
| | Ustc-nelslip [46] | 0.3417 | 0.6157 | 0.5523 | 0.2771 |
| **Ours** | HKMNet | 0.4298 | 0.6518 | 0.6043 | 0.4008 |
| | HKMNet + $\mathcal{H}_\varepsilon$ | **0.4706** | **0.6684** | **0.6433** | **0.4379** |

**Table 2**
Precision, Recall and F1 of 5-cross-validation for each testing category of different methods on the FIW dataset. Here, 7 kin types (F–D, F–S, M–D, M–S, B–B, B–S, S–S) within the Primary-Kin category are used for training.

| Methods | | Primary-Kin | | | Secondary-Kin | | | Non-kin | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Precision | Recall | F1 | Precision | Recall | F1 | Precision | Recall | F1 | F1-Macro |
| Same-backbone | AttentionNet+$Softmax^*$ [44] | 0.4490 | 0.4826 | 0.4538 | 0.2609 | 0.2824 | 0.2243 | 0.3711 | 0.4565 | 0.3682 | 0.3488 |
| | AttentionNet+$Softmax$ [44] | 0.3373 | 0.3992 | 0.3344 | 0.2401 | 0.1546 | 0.1783 | 0.2966 | 0.4390 | 0.3379 | 0.2835 |
| Open set | CAC [39] | 0.3486 | 0.2696 | 0.2537 | 0.0667 | 0.0171 | 0.0272 | 0.2941 | 0.7118 | 0.3982 | 0.2264 |
| | ARPL [29] | 0.5757 | 0.3362 | 0.4245 | 0.3206 | 0.7335 | 0.4462 | 0.2642 | 0.0745 | 0.1162 | 0.2590 |
| | RPL [30] | 0.5733 | 0.2192 | 0.3171 | 0.3562 | 0.4435 | 0.3951 | 0.2391 | 0.4814 | 0.3195 | 0.2639 |
| | GCPL [45] | 0.6333 | 0.0946 | 0.1647 | 0.3605 | 0.3113 | 0.3341 | 0.2369 | 0.7074 | 0.3549 | 0.2581 |
| Kinship related methods | CNN-point [7] | 0.4777 | 0.5162 | 0.4878 | 0.3399 | 0.1905 | 0.2215 | 0.3850 | 0.5654 | 0.4403 | 0.3832 |
| | AttentionNet* [44] | 0.3436 | 0.2778 | 0.2397 | 0.1411 | 0.2264 | 0.1393 | 0.2612 | 0.5037 | 0.2935 | 0.2242 |
| | JLNet [26] | 0.4142 | 0.4695 | 0.4163 | 0.3775 | 0.2961 | 0.2743 | 0.3594 | 0.3859 | 0.3576 | 0.3134 |
| | Ustc-nelslip [46] | 0.3865 | 0.6013 | 0.4215 | 0.1533 | 0.0564 | 0.0842 | 0.3019 | 0.5439 | 0.3273 | 0.2771 |
| **Ours** | HKMNet | 0.4617 | 0.5883 | 0.4904 | 0.3790 | 0.2465 | 0.2703 | 0.4193 | 0.4742 | 0.4418 | 0.4008 |
| | HKMNet + $\mathcal{H}_\varepsilon$ | 0.4802 | 0.5717 | 0.5122 | 0.3934 | 0.3627 | 0.3689 | 0.4918 | 0.4119 | 0.4327 | **0.4379** |

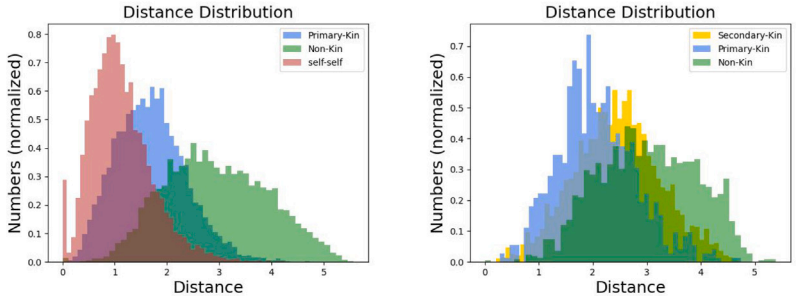
performance. CAC, ARPL, and JLNet use kin-types as their training classes following their training set in the experiments. CNN-point uses the same auto-annotated label as our HKMNet.

#### 5.2.2. Results and discussion

Accuracy measures the matching of correct categories. It corresponds to the ability in positive samples among different models. AUROC (Primary-Kin vs. unlabeled) corresponds to the ability to differentiate known labeled and unlabeled pairs where unknown samples contain both unknown-kin-related and unknown-non-kin pairs. AUROC (Secondary-Kin vs. Non-Kin) measures the ability to retrieve unlabeled but kin-related pairs from all unlabeled samples. F1 score computes the average F1 scores for three categories. Furthermore, precision, recall, and F1 scores are calculated for each category. We conduct 5-cross validation for the final results. Our HKMNet outperforms other methods for all metrics (e.g. distribution of distances and ROC curves).

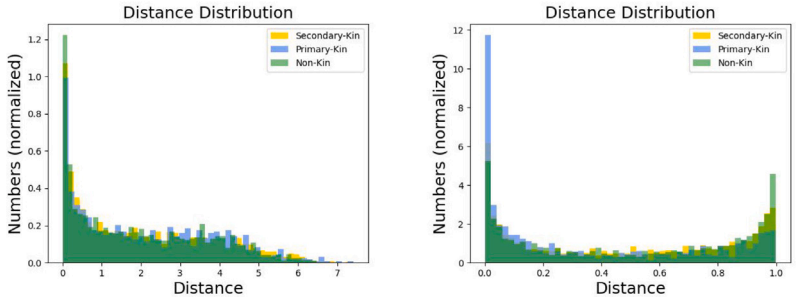
*Comparison with SOTA open-set methods.* Table 1 and Fig. 5 show comparisons of our method with SOTAs in OSR. Our approach, HKMNet, outperforms ARPL [29] and CAC [39] in all the four metrics. Note that $\mathcal{T}_1$ and $\mathcal{T}_2$ are derived from two ROC curves based on the Youden index [47]. If $\mathcal{T}_1 > \mathcal{T}_2$, there will be no optimal threshold for $S$. Then, Precision, Recall, and F1 will be set to 0. According to the results in Table 2, CAC and ARLP perform poorly in separating Non-kin type and unlabeled but kin-related pairs. ARLP obtains the lowest recall and F1 for Non-Kin CAC obtains the lowest recall and F1 for Secondary-Kin, which means it does not have the capability to distinguish and mine the unlabeled but kin-related pairs. Instead, our proposed model obtains more balanced scores.

*Comparison with SOTA kinship recognition methods.* CNN-Point, AttentionNet*, and JLNet are kinship recognition related methods and are specifically designed to extract kinship related features. The results show that CNN-Point and JLNet perform, in general, better than CAC




(a) Distribution of distances of HKMNet on the training dataset.

(b) Distribution of distances of HKMNet on the testing dataset.




(c) Distribution of distances of CAC on the testing dataset.

(d) Distribution of distances of JLNet on the testing dataset.

**Fig. 6.** Distribution of distances of HKMNet (ours), CAC and JLNet.

**Table 3**
Ablation study of HKMNet on FIW dataset.

| *tanh* | $fc$ share weights | $\mathcal{H}_\mathcal{E}$ | Accuracy | AUROC Primary-Kin vs. Unlabeled | AUROC Secondary-Kin vs. Non-Kin | F1 |
|---|---|---|---|---|---|---|
| | ✓ | | 0.4193 | 0.6486 | 0.6066 | 0.3772 |
| ✓ | | | 0.4444 | 0.6472 | 0.6098 | 0.3885 |
| ✓ | ✓ | | 0.4298 | 0.6518 | 0.6043 | 0.4008 |
| ✓ | ✓ | ✓ | **0.4706** | **0.6684** | **0.6433** | **0.4379** |

and ARPL. AttentionNet* utilizes the same feature backbone and label as HKMNet(ours). The results between HKMNet and AttentionNet* indicate the feasibility of our proposed methods. Fig. 6 illustrates the distributions of distances for HKMNet (ours), CAC and JLNet on the testing dataset. For CAC, the distance distributions of three testing categories are overlapping. As for JLNet, it can distinguish the known Primary-Kin and Non-kin categories. However, the Secondary-Kin category does not follow a uniform distribution. In contrast, our method is able to properly separate the three categories. Although the Secondary-Kin is unknown to our HKMNet, our model can still form a unified similarity distribution for the Secondary-Kin category.

*Comparison with the same backbone CNNs.* The results are shown in Table 1. Our model outperforms all (same) backbone models. As for precision, recall, and F1 among Secondary-Kin and Non-kin in Table 2, AttentionNet+ $Softmax$ performs in an unbalanced way. When comparing metric AUROC between AttentionNet+ $Softmax$ and AttentionNet+ $Softmax$*, it is shown that adding non-kin label information improves the capability of the AttentionNet+ $Softmax$* to distinguish different categories. Even though using the same auto-annotated labels, our HKMNet still outperforms AttentionNet*. This means that our proposed Hierarchical Kinship Triplet Loss is more beneficial than the Softmax+Cross-Entropy loss.

### 5.2.3. Ablation study

Ablation study is conducted. The performances of our network with different parameters are shown in Tables 3 and 4. The $\lambda$ represents the ratio between $\sum_{i=1}^{2}\mathcal{L}_t(x_1, x_{i+1}, x_{i+2})$ and $\sum_{i=1}^{3}\|d(x_1, x_{i+1}) - c_i\|^2$. In Table 3, when removing the *tanh* activation function, the performance drops. *tanh* activation function enforces the feature values to a limited scale, which helps the model to obtain a better representation. When the $fc$ layers do not share weights, it is difficult for the method to learn a uniform projection for the different extracted features. Table 4 shows that the model gets higher results when the c1, c2, c3, m1, m2, $\lambda$ are $0, 1, 4, 0.5, 3, 1$ respectively. When mining hard samples during the training process, the AUROC score increases. The Accuracy and average F1 are also increased. The hard sample mining process can be a good process to improve the performance.

### 5.2.4. Qualitative results

Qualitative results are shown in Fig. 7. "Secondary-Kin" pairs are unknown and unlabeled in the open set. They are easy to be taken as "Primary-Kin" or "Non-Kin". Compared to other methods, our model (HKMNet) is able to correctly recognize the testing pairs in the "Primary-Kin", "Secondary-Kin", and "Non-Kin" categories.

## 5.3. Further studies

Our experiments and results affirm the feasibility of our proposed approach for kinship recognition in open-set scenarios. Nonetheless, there is room for improvement in our current work. Firstly, we take a general approach to addressing the open-set problem, with a primary focus on similarity levels. Our experiments illustrate the efficacy of this method in open-set scenarios. In the future, more refined subtasks can be studied, such as recognizing specific kin types in an open-set environment. Secondly, our model is plug-and-play, not tied to a specific feature extractor. This allows our model to use various feature extraction methods. In this paper, we conduct experiments using different feature extractors [44,46]. In upcoming research, more feature extractors will be studied, including different cross-model architectures and Large Language Models. Additionally, we will investigate the influence of aging, gender, and ethical factors. Lastly, there remains a demand for multi-modal kinship datasets and additional kin-type

**Table 4**
5-cross-validation results of HKMNet (ours) with different **c**, $m$, and $\lambda$.

| Methods | c1 | c2 | c3 | $m_1$ | $m_2$ | $\lambda$ | $\mathcal{H}_\mathcal{E}$ | Accuracy | AUROC Primary-Kin vs. Unlabeled | AUROC Secondary-Kin vs. Non-Kin | F1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| HKMNet (Ours) | 0 | 1 | 2 | 0.3 | 3 | 1 | | 0.4499 | 0.6578 | 0.6185 | 0.4107 |
| HKMNet (Ours) | 0 | 1 | 4 | 1 | 1 | 1 | | 0.4261 | 0.6624 | 0.6124 | 0.3778 |
| HKMNet (Ours) | 0 | 1 | 4 | 0.5 | 3 | 2 | | 0.4170 | 0.6560 | 0.6403 | 0.3734 |
| HKMNet (Ours) | 0 | 1 | 4 | 0.5 | 3 | 2 | ✓ | 0.4573 | **0.6734** | 0.6293 | 0.4058 |
| HKMNet (Ours) | 0 | 1 | 4 | 0.5 | 3 | 1 | | 0.4298 | 0.6518 | 0.6043 | 0.4008 |
| HKMNet (Ours) | 0 | 1 | 4 | 0.5 | 3 | 1 | ✓ | **0.4706** | 0.6684 | **0.6433** | **0.4379** |

**Fig. 7.** Qualitative results of different methods. The testing set consists of three kinship categories: "Primary-Kin", "Secondary-Kin" and "Non-Kin". Here, "Secondary-Kin" and "Non-Kin" are unknown and unlabeled pairs in the open set.

datasets to broaden the scope of applicability for this emerging task. More datasets with secondary-kin and weak-kin relations should be collected in the future. In conclusion, our approach shows promise in real-world scenarios, and ample progress remains to be made.

## 6. Conclusion

A method has been proposed to determine family relationships and their corresponding degrees of kinship in an open set environment. It is pairwise-based and is able to exploit mutual information from positive pairs in a hierarchical way. Experiments and an ablation study show that our method outperforms the compared methods. Our model is able to properly separate kinship categories, and generates uniform similarity distributions.

Our approach has the following benefits for related kinship research: (1) pairs may exist (i) without any genetic relationship or (ii) with unlabeled kin-relationships in open-set collections. This may lead to (hard) negative samples possibly affecting the performance of close-set trained models in a negative way. Our method determines general kinship similarities of potential pairs to identify (hard) non-kin samples. This may yield cleaner close-set collections for kinship-related tasks, (2) current publicly available datasets may contain kin-related pairs without labels (*i.e.* kinship degree). Our approach is able to measure similarities for these unlabeled but kin-related samples yielding enhanced and larger kinship datasets.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.patcog.2023.110123.

## References


[1] J.P. Robinson, M. Shao, Y. Fu, Visual kinship recognition: A decade in the making, 2020, arXiv e-prints, arXiv–2006.
[2] J.P. Robinson, M. Shao, Y. Wu, Y. Fu, Families in the wild (fiw) large-scale kinship image database and benchmarks, in: ACM, 2016, pp. 242–246.
[3] Y. Wu, Z. Ding, H. Liu, J. Robinson, Y. Fu, Kinship classification through latent adaptive subspace, in: FG, IEEE, 2018, pp. 143–149.
[4] J.P. Robinson, M. Shao, H. Zhao, Y. Wu, T. Gillis, Y. Fu, Recognizing families in the wild (rfiw) data challenge workshop in conjunction with acm mm 2017, in: WRFW, 2017, pp. 5–12.
[5] N. Kohli, D. Yadav, M. Vatsa, R. Singh, A. Noore, Supervised mixed norm autoencoder for kinship verification in unconstrained videos, TIP 28 (3) (2018) 1329–1341.
[6] L.A. Zebrowitz, J.M. Montepare, Social psychological face perception: Why appearance matters, SPPC 2 (3) (2008) 1497–1517.
[7] K. Zhang, Y. Huang, C. Song, H. Wu, L. Wang, Kinship verification with deep convolutional neural networks, in: BMVC, 2015, pp. 148.1–148.12.
[8] M.J. Khoury, B.H. Cohen, C.A. Newill, W.B. Bias, V.A. McKusick, Inbreeding and prereproductive mortality in the Old Order Amish. II. Genealogic epidemiology of prereproductive mortality, AJE 125 3 (1987) 462–472.
[9] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, J. Zhou, Neighborhood repulsed metric learning for kinship verification, PAMI 36 (2) (2013) 331–345.
[10] C. Geng, S.-j. Huang, S. Chen, Recent advances in open set recognition: A survey, PAMI (2020) 3614–3631.
[11] R. Fang, K.D. Tang, N. Snavely, T. Chen, Towards computational models of kinship verification, in: ICIP, IEEE, 2010, pp. 1577–1580.
[12] M. Wang, X. Shu, J. Feng, X. Wang, J. Tang, Deep multi-person kinship matching and recognition for family photos, Pattern Recognit. (2020) 107342.
[13] J. Robinson, M. Shao, Y. Fu, Survey on the analysis and modeling of visual kinship: A decade in the making, PAMI (ISSN: 1939-3539) (2020) 1.
[14] S. Xia, M. Shao, Y. Fu, Kinship verification through transfer learning, in: IJCAI, 2011, pp. 2539–2544.
[15] X. Zhou, J. Lu, J. Hu, Y. Shang, Gabor-based gradient orientation pyramid for kinship verification under uncontrolled environments, in: ICM, ACM, 2012, pp. 725–728.
[16] Y. Guo, H. Dibeklioglu, L. Van der Maaten, Graph-based kinship recognition, in: ICPR, IEEE, 2014, pp. 4287–4292.
[17] H. Dibeklioglu, A. Ali Salah, T. Gevers, Like father, like son: Facial expression dynamics for kinship verification, in: ICCV, 2013, pp. 1497–1504.
[18] H. Yan, J. Lu, W. Deng, X. Zhou, Discriminative multimetric learning for kinship verification, TIFS 9 (7) (2014) 1169–1178.
[19] W. Li, Y. Zhang, K. Lv, J. Lu, J. Feng, J. Zhou, Graph-based kinship reasoning network, in: ICME, IEEE, 2020, pp. 1–6.
[20] W. Li, J. Lu, A. Wuerkaixi, J. Feng, J. Zhou, Reasoning graph networks for kinship verification: from star-shaped to hierarchical, TIP 30 (2021) 4947–4961.
[21] W. Li, S. Wang, J. Lu, J. Feng, J. Zhou, Meta-mining discriminative samples for kinship verification, in: CVPR, IEEE, 2021, pp. 16135–16144.
[22] X. Chen, C. Li, X. Zhu, L. Zheng, Y. Chen, S. Zheng, C. Yuan, Deep discriminant generation-shared feature learning for image-based kinship verification, SPIC 101 (2022) 116543.
[23] H. Yan, C. Song, Multi-scale deep relational reasoning for facial kinship verification, PR 110 (2021) 107541.
[24] H. Yan, J. Hu, Video-based kinship verification using distance metric learning, PR 75 (2018) 15–24.
[25] M. Wang, J. Feng, X. Shu, Z. Jie, J. Tang, Photo to family tree: Deep kinship understanding for nuclear family photos, in: WASMC, 2018, pp. 41–46.
[26] W. Wang, S. You, T. Gevers, Kinship identification through joint learning using kinship verification ensembles, in: ECCV, Springer, 2020, pp. 613–628.
[27] W.J. Scheirer, A. de Rezende Rocha, A. Sapkota, T.E. Boult, Toward open set recognition, PAMI 35 (7) (2012) 1757–1772.
[28] Z. Liu, Z. Miao, X. Zhan, J. Wang, B. Gong, S.X. Yu, Large-scale long-tailed recognition in an open world, in: CVPR, 2019, pp. 2537–2546.
[29] G. Chen, P. Peng, X. Wang, Y. Tian, Adversarial reciprocal points learning for open set recognition, PAMI (2021) 1.
[30] G. Chen, L. Qiao, Y. Shi, P. Peng, J. Li, T. Huang, S. Pu, Y. Tian, Learning open set network with discriminative reciprocal points, in: ECCV, Springer, 2020, pp. 507–522.
[31] W.J. Scheirer, L.P. Jain, T.E. Boult, Probability models for open set recognition, PAMI 36 (11) (2014) 2317–2324.
[32] M. Gunther, S. Cruz, E.M. Rudd, T.E. Boult, Toward open-set face recognition, in: CVPRW, IEEE, 2017, pp. 71–80.
[33] H. Zhang, V.M. Patel, Sparse representation-based open set recognition, PAMI 39 (8) (2016) 1690–1696.
[34] A. Bendale, T.E. Boult, Towards open set deep networks, in: CVPR, IEEE, 2016, pp. 1563–1572.
[35] P. Zhang, J. Wang, A. Farhadi, M. Hebert, D. Parikh, Predicting failures of vision systems, in: CVPR, IEEE, 2014, pp. 3566–3573.
[36] Z. Ge, S. Demyanov, Z. Chen, R. Garnavi, Generative openmax for multi-class open set classification, 2017, arXiv preprint arXiv:1707.07418.
[37] I. Jo, J. Kim, H. Kang, Y.-D. Kim, S. Choi, Open set recognition by regularising classifier with fake data generated by generative adversarial networks, in: ICASSP, IEEE, 2018, pp. 2686–2690.
[38] L. Neal, M. Olson, X. Fern, W.-K. Wong, F. Li, Open set learning with counterfactual images, in: ECCV, 2018, pp. 613–628.
[39] D. Miller, N. Sunderhauf, M. Milford, F. Dayoub, Class anchor clustering: A loss for distance-based open set recognition, in: WCACV, 2021, pp. 3570–3578.
[40] A. Krizhevsky, G. Hinton, et al., Learning Multiple Layers of Features from Tiny Images, Technical Report, Citeseer, 2009.
[41] G. Pereyra, G. Tucker, J. Chorowski, Ł. Kaiser, G. Hinton, Regularizing neural networks by penalizing confident output distributions, 2017, arXiv preprint arXiv:1701.06548.
[42] B. Gao, L. Pavel, On the properties of the softmax function with application in game theory and reinforcement learning, 2017, arXiv preprint arXiv:1704.00805.
[43] T.-D.H. Nguyen, H.-N.H. Nguyen, H. Dao, Recognizing families through images with pretrained encoder, in: 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), IEEE, 2020, pp. 887–891.
[44] H. Yan, S. Wang, Learning part-aware attention networks for kinship verification, PRL 128 (2019) 169–175.
[45] H.-M. Yang, X.-Y. Zhang, F. Yin, C.-L. Liu, Robust classification with convolutional prototype learning, in: CVPR, 2018, pp. 3474–3482.
[46] J. Yu, M. Li, X. Hao, G. Xie, Deep fusion siamese network for automatic kinship verification, in: FG, IEEE, 2020, pp. 892–899.
[47] W.J. Youden, Index for rating diagnostic tests, Cancer 3 (1) (1950) 32–35.

**Wei Wang** received the B.E. degree in Electronic Information Engineering from the Shandong University, Jinan, China, in 2015 and received M.Sc. degree in Media Technology from KTH Royal Institute of Technology, Stockholm, Sweden, in 2018. He is a Ph.D candidate in University of Amsterdam, Amsterdam, Netherlands. His research interests include facial recognition, kinship verification, kinship recognition and open-set recognition in computer vision.

**Shaodi You** is an Assistant Professor at University of Amsterdam (UvA), Netherlands. He receives his Ph.D. and M.E. degrees from The University of Tokyo, Japan in 2015 and 2012 and his bachelor's degree from Tsinghua University, P. R. China in 2009. He is best known for visual enhancement in bad weather for autonomous driving. He is an area chair of CVPR2022, ECCV2022 and ECCV2020. He is the program chair for ICCV2017, ICCV2019, ICCV2021 Workshop on Physics Based Vision meets Deep Learning.

**Sezer Karaoglu** received the double M.Sc. degrees in optics image vision from the Université Jean Monnet, Saint-Etienne, France, and in multimedia technologies from Gjvik University College, Gjvik, Norway, in 2011, and the Ph.D. degree from the Computer Vision Group, Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands, in 2016. He is the CTO and Co-Founder of 3DUniversum, a spin-off of the Informatics Institute, UvA. His research interests include 2D-3D computer vision and machine learning.

**Theo Gevers** is a Professor of computer vision with the University of Amsterdam. He is the Director of the Computer Vision Laboratory and Co-Director of the Atlas Laboratory and Delta Lab in Amsterdam. He is the Co-Founder of AI technology companies 3DUniversum and Sightcorp. His research area focuses on artificial intelligence with the focus on computer vision and deep learning, and in particular image processing, 3D (object) understanding and human-behavior analysis. He has authored or co-authored more than 230 papers and three books.