

人脸亲子关系验证研究综述

吴晓婷^{1,2} 冯晓毅¹ 黄安³ 张雪毅³ 董晶⁴ 刘丽^{2,3}

摘要 人脸亲子关系验证即通过给定的不同人的两幅人脸图像判断其是否具有亲子关系,是计算机视觉和机器学习领域中一个重要的研究问题,在丢失儿童寻找、社交媒体分析、图像自动标注等领域具有广泛的应用价值.随着人脸亲子关系验证问题受到越来越多的关注,其在多个方面都得到了相应的发展,本文对人脸亲子关系验证方法做了综述整理.首先,简要介绍了人脸亲子关系验证在近十年的研究现状,随后对问题进行了定义并讨论其面临的挑战.接下来,汇总了常用的亲子关系数据库,对数据库属性做了详细的总结和对比.然后,对人脸亲子关系验证方法进行了分类总结、对比,以及不同方法的性能表现.最后,展望了人脸亲子关系验证今后可能的研究方向.

关键词 人脸亲子关系验证,人脸分析,相似度计算,亲子数据集,深度学习

引用格式 吴晓婷,冯晓毅,黄安,张雪毅,董晶,刘丽.人脸亲子关系验证研究综述.自动化学报,2022,48(12):2886-2910

DOI 10.16383/j.aas.c201023



开放科学(资源服务)标识码(OSID):

Survey on Facial Kinship Verification

WU Xiao-Ting^{1,2} FENG Xiao-Yi¹ HUANG An³ ZHANG Xue-Yi³ DONG Jing⁴ LIU Li^{2,3}

Abstract Facial kinship verification aims to verify whether two facial images have kin relations or not. It is an important research topic in computer vision and machine learning domain. Kinship verification has a wide range of applications in fields of missing children finding, social media analysis and image automatic labelling. It receives more and more attention and has been developed in many directions. In this paper, we do a survey on the kinship verification. First, we briefly introduce the kinship verification's state of research within last ten years. Then we define the problem and discuss the challenges. Then, we summarize and compare the attributes of public kinship databases. Kinship verification algorithms are categorized and reviewed as well. We also analysis different kinship verification methods and compare their performance. Finally, we discuss the future possible research directions of kinship verification.

Key words Facial kinship verification, face analysis, similarity computation, kinship databases, deep learning

Citation Wu Xiao-Ting, Feng Xiao-Yi, Huang An, Zhang Xue-Yi, Dong Jing, Liu Li. Survey on facial kinship verification. *Acta Automatica Sinica*, 2022, 48(12): 2886-2910

人脸亲子关系验证(Facial kinship verifica-

收稿日期 2020-12-09 录用日期 2021-03-02

Manuscript received December 9, 2020; accepted March 2, 2021
国家自然科学基金(61872379, 91846301), 湖南省重点领域研发计划(2019GK2131), 陕西省科技计划重点项目(2020GY-050)资助

Supported by National Natural Science Foundation of China (61872379, 91846301), Key Research and Development Program of Hunan Province (2019GK2131), and Key Research and Development Program of Shaanxi (2020GY-050)

本文责任编辑 张敏灵

Recommended by Associate Editor ZHANG Min-Ling

文中涉及的人脸图像均已获得使用授权,仅限于本文使用

The facial images involved in this article have been authorized to use, only for this article.

1. 西北工业大学电子信息学院 西安 710129 中国 2. 奥卢大学机器视觉与信号分析中心 奥卢 90570 芬兰 3. 国防科技大学系统工程学院 长沙 410073 中国 4. 中国科学院自动化研究所智能感知与计算研究中心 北京 100190 中国

1. School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China 2. Center for Machine Vision and Signal Analysis, University of Oulu, Oulu 90570 Finland 3. College of System Engineering, National University of Defense Technology, Changsha 410073, China 4. Center for Research on Intelligent Perception and Computing, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

tion),是指应用计算机视觉和机器学习方法,分析来自不同人的两幅人脸图像的面部特征相似性来自动判断其是否具有某种亲子关系,如父子、父女、母子、母女关系等.人脸亲子关系验证是在生物遗传学理论上,受心理认知学启发^[1-5]衍生出来的一个计算机视觉中的重要研究方向,因为心理学研究表明人们可以通过面部推断亲子关系.人脸亲子关系验证技术在诸多领域具有广泛的应用价值.在人类学、基因学领域,有助于分析社会关系中近源亲属的遗传特性;在社会公共安全领域,可以应用于丢失儿童寻找、国家边境控制、刑事侦查等,如可以快速验证家庭团聚签证申请人的身份信息以及服务快速侦察儿童乞讨、拐卖儿童等违法行为;在社交媒体领域,可以用于家庭信息分析、家庭相簿整理、自动图像标注等;在生活娱乐方面,可以用于子女人脸图像合成、年老人脸合成等;此外,亲子关系验证在智能家居、物联网等领域也有潜在应用价值,

如亲子行为理解、家庭信息互联等。该问题由 Fang 等^[6]于 2010 年提出, 至今仅有十余年的研究历史, 是一个近几年新兴的、具有挑战的一个研究方向, 鉴于人脸亲子关系验证研究潜在的学术价值和应用价值, 近期国内外研究者对该问题的关注呈上升趋势。

国际上, 开展人脸亲子关系验证研究的主要有美国康奈尔大学、东北大学和西弗吉尼亚大学、荷兰阿姆斯特丹大学、芬兰奥卢大学等; 国内开展该方面研究的主要有清华大学、北京邮电大学、西北工业大学等。近年来, 国际上多次举办了人脸亲子关系验证比赛和专题研讨会。第 1 届国际亲子关系验证比赛 (Kinship Verification in the Wild, KVW) 于 2014 年在国际生物特征识别会议 (International Joint Conference on Biometrics, IJCB) 上举办^[7]; 第 2 届 KVW 比赛举办于 2015 年国际人脸与手势自动识别会议 (IEEE Conference on Automatic Face and Gesture Recognition, FG)^[8]; KVW 系列比赛均为单任务, 基于人脸亲子关系验证展开。随后, 美国东北大学 SMILE 实验室^[9-10]自 2017 年起每年举办家庭识别竞赛 (Recognizing Families in the Wild, RFIW), 该竞赛于 2017 年在 ACM 国际多媒体会议 (ACM International Conference on Multimedia, ACM MM)^[10]上举办, 之后每年在 FG 会议上^[9]举办, 2019 年在 Kaggle 平台上举办了 RFIW 竞赛^{2, 3}。与以往不同, RFIW 系列比赛基于大规模亲子数据库^[11]展开的多任务人脸亲子识别比赛, 包括人脸亲子关系验证、三人组人脸亲子关系验证以及家庭识别、家庭检索。

综上所述, 人脸亲子关系验证的研究获得计算机视觉和模式识别领域研究者的广泛关注, 研究者提出很多方法来解决该问题^[6, 12-20]。近期, 国外已经有一些学者对人脸亲子关系验证问题进行了综述, 表 1 对比了这些综述论文, 具体代表性工作总结如下。2014 年, Dandekar 等^[21]对早期人脸亲子关系验证方法进行了总结, 包括研究目标、应用价值、方法框架和早期代表性方法。2016 年, Wu 等^[22]对早期人脸亲子关系验证方法进行了归类总结, 将已有方法分为 4 类: 基于特征的方法、基于度量学习的方法、基于深度学习的方法和其他方法。Almuashi 等^[23]在 2017 年的综述则侧重于阐述人脸亲子关系验证问题的衍生、定义、研究意义以及难点与挑战, 但较少讨论解决方法。Georgopoulos 等^[24]在 2018 年将人脸亲子关系验证任务和人脸衰老问题结合, 分析了其之间的相互关系, 介绍了年龄因素对于亲子关系验证问题的影响, 以及祖辈人脸图像在人脸衰老

变化中的引导作用。2020 年, Qin 等^[25]对人脸亲子关系验证方法进行归类总结, 将其分为基于亲子特征的方法和基于数学度量分析的方法, 但是较少涉及近期基于深度学习的人脸亲子关系验证方法。2020 年, Robinson 等^[26]侧重于总结 RFIW 系列比赛, 包括比赛任务介绍、现阶段性能对比及今后可能的研究方向, 但忽略了其他人脸亲子验证方法的汇总。目前, 国内尚没有关于人脸亲子关系验证方法的综述。

总而言之, 已有的综述对现有人脸亲子关系验证问题的研究方法总结还很不全面, 尤其缺乏对近期基于深度学习的人脸亲子关系验证方法的总结以及近几年以来出现的重要数据集的总结。为吸引更多的研究学者参与到该领域的研究, 促进其发展, 本文尝试首次对人脸亲子关系验证的相关研究成果进行全面综述。首先从人脸亲子关系验证问题的基本定义出发, 对该问题研究中存在的困难与挑战进行了深入剖析; 接下来, 本文对人脸亲子关系验证领域的主流数据库的发展脉络进行了梳理与总结; 随后系统地对比亲子关系验证的方法进行了梳理、归纳、评述和总结; 最后对人脸亲子关系验证未来的研究方向进行了探讨。

1 问题定义及其挑战

心理学研究表明, 人们具有仅通过面部图像推断其是否具有亲子关系的能力^[1-5]。进一步, 相关研究学者探究了影响人们亲子推断的可能因素, 包括面部旋转性^[27]、面部表情^[28]、人脸的对称性^[2]、性别和年龄因素^[3], 以及不同面部区域对亲子验证的贡献分布^[1, 5]。其中, 面部的旋转并不会降低亲子识别的准确率^[27]; 而对比于无表情人脸图像, 无约束的面部表情会对亲子验证产生负面影响, 从而降低识别准确率^[28]; 面部左右脸部分对亲子验证的贡献相当, 仅稍次于整个面部; DeBruine 等^[3]指出性别和年龄因素会大大降低亲子验证的准确率; 就面部区域影响上, 上半脸包含的亲子信息量最大, 而嘴部区域由于形态多样, 容易对亲子识别产生噪声影响, Alvergne 等^[5]指出亲子线索取决于有效面部区域, 而非整个面部区域。心理学领域中的研究成果, 对基于计算机视觉和机器学习方法的自动人脸亲子关系验证问题具有借鉴价值。

图 1 中展示了部分亲子图像对和非亲子图像对, 亲子的推断主要从亲子间的相似性与非亲子间的不相似性得出。然而, 人眼通过感官知觉上判断图像的相似程度, 很难对来自不同人的两幅图像进行相似度量, 如眼睛间距离、五官形状与大小等。此外, 人眼对颜色的区分感知能力较低, 因此在进行亲子关系验证时, 人脑处于模糊判断, 从而造成

¹ <https://web.northeastern.edu/smilelab/RFIW2017/>

² Kaggle 是一个数据建模和数据分析竞赛平台, 信息来自维基百科。

³ <https://www.kaggle.com/c/Recognizing-Faces-in-the-Wild>

表 1 现有脸亲子关系验证综述论文总结
Table 1 The summary of the existing facial kinship verification survey papers

综述	问题定义	难点与挑战	数据集总结	人脸亲子关系验证算法	性能分析	未来的研究方向	侧重点
[21]	×	×	✓	对早期的人脸亲子关系验证算法进行了总结.	×	×	对人脸亲子关系验证问题的基础介绍.
[22]	×	✓	✓	对早期的人脸亲子关系验证算法进行了归类、总结.	×	✓	对人脸亲子关系验证问题做了较为全面的总结.
[23]	✓	✓	✓	主要就亲子特征提取方法进行了对比.	✓	✓	侧重阐述人脸亲子关系验证问题的衍生、定义、研究意义以及难点与挑战.
[24]	×	×	✓	对人脸亲子关系验证算法作了分类、汇总.	✓	×	将人脸亲子关系验证任务和人脸衰老问题结合, 分析了其之间的相互关系.
[25]	✓	✓	✓	对人脸亲子关系验证算法作了分类、汇总.	✓	✓	对亲子关系验证问题作了全面总结, 但是缺乏近期深度学习相关方法的汇总.
[26]	✓	✓	✓	对 RFIW 系列比赛中的方法进行了归纳.	✓	✓	侧重于总结 RFIW 系列比赛, 包括比赛任务介绍、现阶段性能对比及今后可能的研究方向.
本文	✓	✓	✓	对近十年来的亲子验证算法进行了全面的归纳、总结与分析.	✓	✓	文章从问题的数学定义出发, 对人脸亲子关系验证问题进行了全方位的总结、剖析与展望.

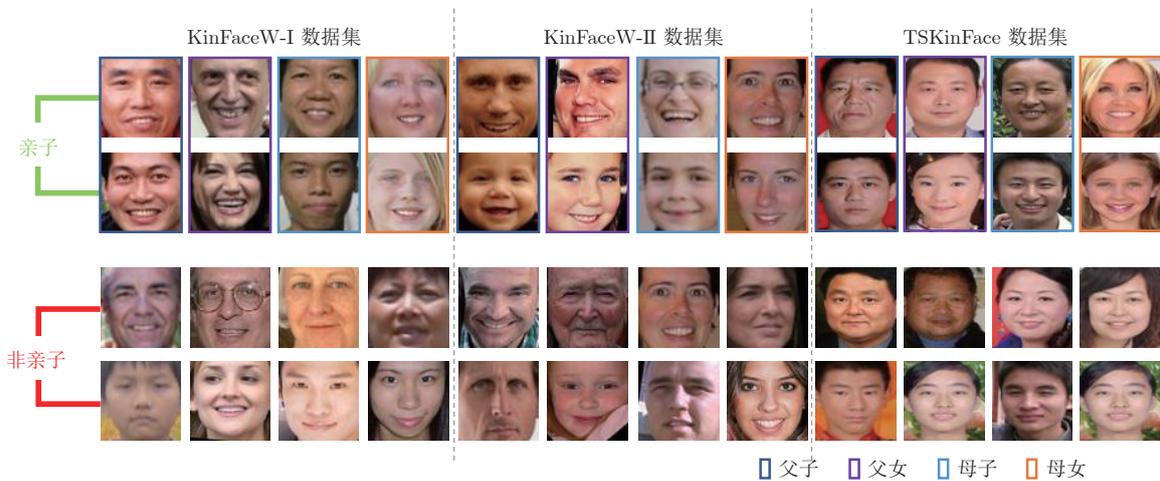


图 1 正样本对 (具有亲子关系) 和负样本对 (不具有亲子关系) 示意图 (© (2018) IEEE^[29] 授权修改版)

Fig.1 The illustration of positive pairs (with kin relations) and negative pairs (without kin relations) (© (2018) IEEE. Modified, with permission, from [29])

识别准确率低下. 不同于人眼对亲子关系验证的判断, 计算机视觉能够在形状和颜色等方面, 准确捕捉亲子间的遗传相似度, 以及判断非亲子图像间的不相似程度, 在数字化基础上有据可依, 通过优化特征表示和分类器, 使得机器学习方法能够更加准确的推断两幅人脸图像是否具有亲子关系.

自动人脸亲子关系验证, 旨在通过图像特征提取和机器学习方法, 分析两幅不同的人脸图像的面部特征相似性来判断其是否具有某种亲子关系, 如父子、父女、母子、母女关系等. 这是一个极具挑战的问题, 因为有的时候具有亲子关系的两幅人脸图像仅存在细微的相似性. 正式地, 给定输入的两幅人脸图像 (或两段视频), 用 I_p 和 I_c 表示, 分别对图像 (或视频) 提取特征表示, 记为 $\phi(I_p)$ 与 $\phi(I_c)$, 两输入之间的距离用 $d(\phi(I_p), \phi(I_c))$ 表示, 通过阈值 θ

判断输入间是否具有亲子关系, 则 I_p 与 I_c

$$\begin{cases} \text{具有亲子关系,} & \text{若 } d(\phi(I_p), \phi(I_c)) < \theta \\ \text{不具有亲子关系,} & \text{若 } d(\phi(I_p), \phi(I_c)) \geq \theta \end{cases} \quad (1)$$

人脸亲子关系验证系统的关键环节在于两个方面, 一是亲子特征的有效表示; 二是人脸图像间距离的准确度量. 人脸验证与人脸亲子关系验证问题两者具有相似性, 却又有很大不同. 两者都需要基于两幅人脸图像的面部特征相似性来进行判断, 前者旨在判断两幅人脸图像是否属于同一个人, 后者旨在判断其是否具有某种亲子关系. 两个问题的正负样本空间是不同的, 前者正样本对相似度高, 负样本对存在明显差异性; 而后的正样本对相似程度较低, 容易与负样本对混淆, 鉴于部分情况下具有亲子关系的两幅人脸图像仅在细微之处存在相似

性,增加了问题的难度系数,人脸亲子关系验证是一个极具挑战的问题。

除了目前研究最热的人脸亲子关系验证问题,研究学者也指出由于孩子的遗传基因来源于父母双方,因此引出了另一个亲子研究问题,即三人组亲子关系验证 (Tri-subject kinship verification), 输入包括父母二人的人脸图像,以及子女人脸图像,分别用 I_f 和 I_m 表示父亲、母亲的人脸图像,用 I_c 表示孩子人脸图像,来判断其是否具有亲子关系。父母、孩子的特征表示为 $\phi(I_f)$ 、 $\phi(I_m)$ 和 $\phi(I_c)$ 。孩子与父母间的距离表示为 $d((\phi(I_f), \phi(I_m)), \phi(I_c))$, 与人脸亲子关系验证问题类似,通过阈值对距离进行分类。三人组亲子关系验证与人脸亲子关系验证均为二分类问题 (Binary classification)。

在人脸亲子关系验证基础上,研究也延伸到家庭分析上,包括家庭分类^[30]、家庭成员搜索和检索^[11]。家庭分类问题是一个多类别的分类问题 (Multi-class classification), 即分类任务中包含多个类别,每个类别为一个家庭,给定一幅待定人脸图像判断其属于哪个家庭。用 $\chi = \{X_1, X_2, \dots, X_k\}$ 表示包含 k 个家庭的集合,对应的多分类标签为 $\{y_1, y_2, \dots, y_k\}$, 训练一个分类器,对输入的人脸图像 x_i 输出其家庭标签。家庭分类问题与人脸亲子关系验证的区别在于,人脸亲子关系验证的任务是分析两输入间的相似性,而家庭分类需要分析一个输入与多个家庭集间的相似与不相似性,从而增加了问题的难度。同其他图像分类问题类似,随着家庭数目的增加,家庭分类的难度也相应增大,因此分类性能也会跟着下降。例如目前最大的亲子数据集 FIW, 家庭分类任务中包含 564 个家庭,分类准确率仅有 16.18%。家庭成员检索问题旨在对输入的人脸图像匹配家庭成员,其中搜索集由所有的家庭成员组成,输入的人脸图像为待查询样本,输出给出最为匹配的前 K 个家庭成员。家庭识别和家庭成员检索问题的区别在于,家庭识别侧重于家庭分类模型的训练,而家庭成员检索则通过相似性度量学习,

检索出与待查询图像较相似的人脸图像,从而找到输入的父母及其他亲属成员。

图 2 总结了 4 种亲子关系相关的任务示意图。由于人脸亲子关系分析仍处于研究初级阶段,且人脸亲子关系验证方法是亲子关系研究中的关键和核心,因此受关注度最高的是人脸亲子关系验证问题,表 2 对现有的代表性人脸亲子关系验证算法进行了分类汇总,分析了其所针对的问题,对比了不同方法的优势与不足,本文第 3 节与第 4 节对具体的算法进行了详细的分析。

人脸亲子关系验证是一个具有挑战的问题,其主要的挑战和难点总结为以下几个方面。

1) 同一人不同形态差异 (Inner-person variations) 的影响。这一点继承了人脸验证 (Face verification) 问题存在的挑战。一方面,由于成像环境导致的各种变化,受到多种因素带来的干扰,如光照变化、成像距离、成像角度、相机参数等;另一方面,由于主体自身的各种变化,例如姿态变化、表情变化、年龄变化、眼镜等配饰引起的遮挡等。这些因素,会导致同一个人的人脸图像表现特征产生很大差异,影响图像特征的提取。

2) 亲子长相的差异性 (Inter-person variations)。待判定是否具有亲子关系的两幅不同人脸图像可能存在很大的年龄差异,也可能来自不同性别,或者仅存在细微的相似性,这些都增加了人脸亲子关系验证的难度。然而性别、年龄等差异性,会导致具有亲子关系的人脸图像差异性增大。可以看出相比于年轻时父母的人脸图像,子女与年老父母图像差异更大。而性别的差异性,间接性地影响了判断是否具有亲子关系的准确性。此外,隔辈人脸亲子关系 (如爷爷-孙子、爷爷-孙女等) 验证问题,由于年龄跨度大、共同遗传信息减少使得判断难度更大。

3) 亲子/非亲子分类边界模糊。具有亲子关系的两个不同人可能看起来有很大差异,仅存在细微的相似性,这种情况被称为困难正样本 (Hard positive sample)。相反,不具有亲子关系的不同人或许

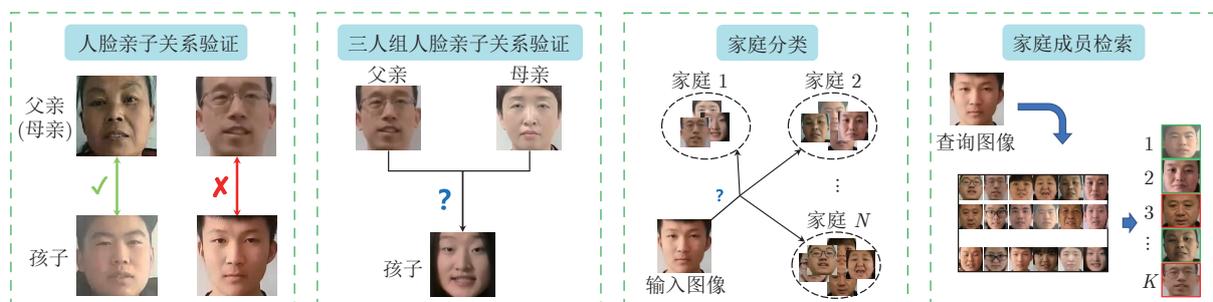


图 2 亲子关系任务示意图

Fig.2 The illustration of kinship related tasks

表 2 人脸亲子关系验证方法的优缺点总结

Table 2 The summary of advantages and disadvantages of the facial kinship verification methods

人脸亲子关系验证方法	代表性方法	针对的问题	优势	不足
基于特征表示	基于显著特征 ^[31-35] , 手工特征 ^[14, 36-39] , 基于颜色的特征变换 ^[13, 29, 40-41] , 特征选择 ^[42-45] , ...	为了解决亲子特征的有效表达问题, 从传统手工设计规则出发, 依据形状、颜色、纹理等特性描述亲子的相关特征.	算法复杂度低, 特征提取速度快.	对光照、人脸角度等因素敏感, 在自然条件下性能较差.
传统方法	邻域驳斥度量学习 ^[15, 46-47] , 基于双线性相似度 ^[48-49] , 图构造 ^[50-51] , 迁移子空间学习 ^[52-54] , 余弦相似度 ^[55-56] , ...	就一般距离度量不能很好地描述类内聚敛性和类间分离性问题, 提出度量学习优化亲子图像对间的距离衡量.	优化亲子图像的距离度量方式, 使得亲子间距离缩小, 并且拉大非亲子间的距离.	依赖于特征的有效表示, 距离优化程度有限.
基于深度学习的方法	基本卷积神经网络 ^[57] , 深度度量学习 ^[17, 58-62] , 基于自编码器 ^[18-19, 63-64] , 基于注意力机制 ^[65] , ...	传统方法固化于特定的规则, 很难适应于复杂情况, 深度神经网络针对问题自适应的学习数据的有效特征表示.	依据损失函数训练网络参数, 特征表达能力强.	需要大量的训练数据, 算法模型复杂, 训练时间长.

在长相上具有一定的相似性, 称为困难负样本 (Hard negative sample). 亲子人脸图像对相似度低、非亲子人脸图像相似度高, 就造成距离度量往往难以分类, 分类边界难以界定.

4) 图像高维特征提取与数据样本量少之间的矛盾. 可见人脸亲子关系验证问题是一个难度很高的研究课题, 人脸亲子特征的准确表示就是技术需攻破的关键壁垒之一. 随着深度学习在计算机视觉领域的不断渗透和发展, 在学习提取高维可判别的图像特征中, 对训练数据的要求非常高, 然而现有的亲子数据集往往存在数据量少、数据质量低下等问题, 给深度模型的训练带来巨大的挑战.

5) 大规模亲子数据库的建立. 现有的亲子数据库往往存在数据量少、数据质量低、涉及情景少等缺点. 大规模全面数据库对机器学习算法的研究推进起着重要的作用, 然而亲子数据库的建立依赖于已知人物间的亲子关系, 由于亲子关系属于个人的家庭信息, 有时候很难直接获得该方面信息, 这就对大规模的亲子数据库的建立产生了巨大的困难.

2 数据集和评估准则

2.1 数据集

数据集在人脸亲子关系验证研究中发挥着重要的作用. 在如今的大数据时代, 大量数据的收集显得越来越重要. 一方面, 公开的标准数据集为研究者提供了实验数据和统一的评价标准, 另一方面, 数据集的构建与发展也进一步促进研究的问题趋于完善. 在亲子关系验证研究出现以前, 没有相关的亲子人脸数据集, 因此, 康奈尔大学的 Fang 等^[6]于 2010 年建立了第一个亲子关系数据集 Cornell KinFace, 该数据集由 300 幅人脸图像 (即 150 个亲子对) 组成, 包含了父亲-儿子, 父亲-女儿, 母亲-儿子, 母亲-女儿 4 种主要的亲子关系. 此后十年间,

许多学者建立了多种亲子数据集. 相较于最初的数据集, 新的数据集在数据集大小、结构、亲子关系类型以及数据的模态上各有丰富和发展.

目前常用的公共人脸亲子关系数据集有 12 个, 分别为: Cornell KinFace^[6], UB KinFace^[52-53], KinFaceW^[15], Family 101^[30], UVA-NEMO Smile^[66-67], TSKinFace^[68], FIW^[11], WVU^[69], KFWW^[70], FFWW^[71], KIVI^[72] 和 TALKIN^[73] 数据集. 表 3 对现有亲子数据集参数进行了总结和对比. 下面我们将对这些数据集展开介绍其特性.

1) Cornell KinFace 数据集⁴. Cornell KinFace^[6]是第一个公共人脸亲子关系数据集, 其由康奈尔大学搜集. 其包含 150 对名人以及名人的父母或者子女的人脸图像, 每幅图像的大小是 100×100 像素, 这些图像均来源于网络, 图像采集环境为自然条件, 即对光照、姿态、背景、表情、遮挡等因素上没有要求. 该数据集的提出具有开创性意义, 为之后的数据集构建提供了借鉴. Cornell KinFace 数据集没有对父母-子女亲子关系进一步细分, 根据名人的身份信息推断亲子关系类型发现, 亲子关系类型分布数量不均衡, 其中母亲-儿子这一亲子关系类型数量最少, 父亲-儿子关系对最多.

2) UB KinFace 数据集⁵. 由美国东北大学的 Shao 等采集的 UB KinFace^[52-53], 是第一个包括有子女、年轻父母和年老父母的人脸图像数据集. 与其他的亲子关系数据集不同, 它收集了父母年轻时期的照片, 这是由于研究人员发现年轻父母与其子女的人脸图像之间具有更高的相似性. 它包括 400 个人的 600 幅图像, 图像的分辨率为 89×96 像素. 这 600 幅图像分为 200 组, 每一组中包括了子女、年轻时父母及年老时父母的人脸图像各一幅. UB KinFace 涉及到的人种分为两部分: 亚洲人和非亚洲

⁴ <http://chenlab.ece.cornell.edu/projects/KinshipVerification/>

⁵ <http://www1.ece.neu.edu/yunfu/research/Kinface/Kinface.htm>

表 3 亲子关系数据库属性总结
Table 3 Characteristics of kinship databases

数据集, 年份	数据集大小	属性	分辨率 (像素)	有无多幅图像/视频	年龄变化	有无家庭结构	录制环境	数据集特点	针对的问题
CornellKin ^[6] , 2010	300 幅图像 150 对父子	图像	100×100	无	无	无	自然环境	第一个亲子数据集	1V1
UB Kinface ^[52-53] , 2011	600 幅图像 200 组父子	图像	89×96	无	有	无	自然环境	包含年轻与年老的父母图像的数据集	1V1
UvA-NEMO Smile ^[66-67] , 2012	1240 段视频	视频	1920×1080	有	无	无	约束环境	第一个亲子视频数据集	1V1
Family 101 ^[30] , 2013	14816 幅图像 101 个家庭树	图像	120×150	有	有	有	自然环境	第一个有家庭结构的数据集	1V1, 家庭识别
KinFaceW-I ^[15] , 2014	1066 幅图像 533 对父子	图像	64×64	无	无	无	自然环境	常用的数据集, 图像来自不同照片	1V1
KinFaceW-II ^[15] , 2014	2000 幅图像 1000 对父子	图像	64×64	无	无	无	自然环境	常用的数据集, 图像来自同一照片	1V1
TSKinFace ^[68] , 2015	787 幅图像 1015 组父子	图像	64×64	无	无	有	自然环境	同时包含父亲和母亲的人脸图像	2V1
FIW ^[18] , 2016	11 193 幅图像 1000 个家庭	图像	108×124	有	有	有	自然环境	当前规模最大的亲子数据集	1V1, 2V1 家庭识别, 检索
WVU ^[69] , 2017	904 幅图像 113 对父子	图像	32×32	有	未知	无	自然环境	每人均有4幅图像不同的图片	1V1
KFVW ^[70] , 2018	836 段视频 418 对父子	视频	900×500	无	无	无	自然环境	第一个无表情约束视频数据集	1V1
FFVW ^[71] , 2018	300 段视频 100 组父子	视频	未知	无	无	有	自然环境	可用于三人组视频亲子关系验证	2V1
KIVI ^[72] , 2019	503 段视频 211 个家庭	视频	未知	无	无	有	自然环境	无约束的亲子视频数据集	1V1
TALKIN ^[73] , 2019	800 段视频 400 对父子	视频音频	1920×1080	无	无	无	自然环境	第一个多模态数据集	1V1

表中缩写含义: 1V1表示两输入的亲子关系验证, 2V1表示三人组, 即父亲-母亲-孩子的亲子关系验证。

人, 各有 100 组. 该数据集对具体的亲子关系类型进行了区分, 分为 4 种亲子关系, 其中父亲-儿子, 父亲-女儿, 母亲-儿子, 母亲-女儿分别有 91 组, 79 组, 21 组以及 15 组.

3) KinFaceW 数据集⁶. KinFaceW 数据集由两个子数据集组成: KinFaceW-I 和 KinFaceW-II, 数据是从网络上收集的自然环境下的人脸图像, 由 Lu 等提出^[15], 所有图像的大小为 64×64 像素. 两个子数据集的数据结构相同, 不同之处在于, KinFaceW-I 中的人脸图像来自不同的照片, 而 KinFaceW-II 中的人脸图像对是从同一幅图像中分割得到的. KinFaceW-I 中, 父亲-儿子, 父亲-女儿, 母亲-儿子, 母亲-女儿分别有 134 对, 156 对, 127 对和 116 对; KinFaceW-II 中, 每一种亲子关系分别有 250 对, 由于其数据规模适中, 数据分布合理, 便于进行训练, 绝大多数基于人脸亲子关系验证研究中都用到了该数据集, 但随着深度学习算法在人脸亲子关系验证中的运用, 该数据集也暴露出其在数据量上的短板.

4) Family 101 数据集⁷. Family 101^[30] 数据集是第一个包含有家庭树结构的数据集, 该数据集是由康奈尔大学先进多媒体处理 (Advanced multimedia processing, AMP) 实验室建立. Family 101 数据集包含了 206 个核心家庭, 涉及 607 个人, 共计 14816 幅人脸图像. 每一棵家庭树有 1~7 个核心家庭, 而每一个核心家庭包括 3~9 个家庭成员. 数据集中的图像大部分是灰度图, 图像的分辨率各有不同, 并没有进行统一. Family 101 数据集中包含了白种人、亚洲人和非裔美国人等多个人种. Family 101 数据集中有多种亲子关系类型, 父亲-儿子, 父亲-女儿, 母亲-儿子, 母亲-女儿分别有 213 对, 147 对, 184 对和 148 对. Family 101 数据集以家庭树结构组织数据, 高效地表示了不同样本间的亲子关系. 此外, Family101 中的多数人具有多幅人脸图像, 部分图像拍摄于不同的年龄时期.

5) UVA-NEMO Smile 数据集⁸. UVA-NEMO

⁷ <http://chenlab.ece.cornell.edu/projects/KinshipClassification/index.html>

⁸ <http://www.uva-nemo.org/>

⁶ <http://www.kinfacew.com/>

Smile 数据集由 Dibeklioglu 等^[66-67] 提出, 最初建立该数据集的目的是为了检测参与者是在“真笑”还是“假笑”, 由于该数据集的参与者是以家庭为单位的, 其也被视为第一个由人脸视频组成的亲子数据集. UVA-NEMO Smile 数据集包括有 1240 段微笑视频 (其中 597 段“真笑”视频, 643 段“假笑”视频), 所有视频的拍摄背景均为黑色, 光照条件固定. 参与者的年龄在 8~76 岁之间, 视频帧大小为 1920×1080 像素. 该数据集包含了父母-子女的 4 种主要亲子关系类型以及兄弟姐妹 3 种亲子关系类型 (根据性别不同分成的姐妹关系, 兄弟关系, 姐弟/兄妹关系). 该数据集的亲子关系对数较少, 仅有 95 对, 且视频中包含的人种相对单一, 绝大多数为白种人.

6) TSKinFace 数据集⁹. TSKinFace 数据集是由淮阴师范大学的 Qin 等^[68] 提出的人脸图像数据集, 与之前的数据集不同, 它主要用于研究三人组亲子关系验证问题. TSKinFace 包含两种组合亲子关系类型: 父亲-母亲-儿子和父亲-母亲-女儿, 分别有 513 组和 502 组, 人脸图像来自于网络并且对拍摄环境没有约束, 每幅人脸图像大小归一化为 64×64 像素, 数据集中涉及了多个人种.

7) FIW 数据集¹⁰. FIW (Families in the Wild)^[11] 数据集是目前规模最大, 综合性最强的亲子图像数据集, 由 Robinson 等建立. FIW 数据集以家庭结构组织数据, 包含 1000 个家庭的超过 13000 幅人脸图像, 每幅人脸图像大小为 224×224 像素. FIW 涉及了家庭成员在不同时间的多张照片. 在亲子关系类型上, 不仅包含了在 UVA-NEMO Smile 中涉及的 7 种亲子关系类型, 而且还涉及了四种跨代亲子关系类型 (祖父-孙子, 祖父-孙女, 祖母-孙子, 祖母-孙女). FIW 不仅在数据集的规模上远远超过了 Family 101, 而且在家庭结构方面也比 Family 101 更为复杂. 它的数据量也远超其他的数据集.

8) WVU 数据集¹¹. WVU 数据集由西弗吉尼亚大学 (West Virginia University, WVU) 的 Kohli 等^[69] 建立, 数据集包括 113 对亲子关系对, 包含了 UVA-NEMO Smile 中涉及的 7 种亲子关系类型, 其中 22 对兄弟, 9 对姐弟/兄妹, 13 对姐妹, 14 对父女, 34 对父子, 13 对母女和 8 对母子. 为了增加人脸图像多样性, 数据集每个参与者包含 4 幅人脸图像. 缺点是数据集规模小, 亲子对数量有限.

9) KFVW 数据集¹². Yan 等^[70] 提出了 KFVW

(Kinship Face Videos in the Wild) 这一人脸视频亲子关系数据集, 与 UVA-NEMO Smile 的固定背景不同, 它采集于自然环境. 视频在光照、姿势、遮挡情况、背景、表情、化妆情况、年龄等方面没有约束. KFVW 数据集包含了 418 对人脸视频, 每段视频包含 100~500 帧图像, 视频帧的大小为 900×500 像素. 这些视频数据均来源于网络电视节目. KFVW 包含了 4 种主要的亲子关系类型, 该数据集在亲子关系类型的分布上比较均衡, 数据量较 UVA-NEMO Smile 也有了很大的提升, 而不足的是 KFVW 数据集没有以家庭结构组织数据, 每个个体没有多段视频.

10) FFVW 数据集. FFVW 数据集是由 Sun 等^[71] 建立的一个人脸视频亲子关系数据集, 与 TSKinFace 数据集相似, 该数据集主要用于基于人脸视频的三人组亲子关系验证. FFVW 包含了 100 组对象的 300 段视频, 数据来源于互联网, 视频中的环境均为自然环境. 与 TSKinFace 不同, FFVW 将父母-子女关系统一作为研究对象, 没有对子女性别进行细分.

11) KIVI 数据集¹³. KIVI 数据集由 Kohli 等^[72] 建立, 数据以家庭为结构进行组织. 包含了 503 个个体的脸视频, 这 503 个个体来自于 211 个家庭, 共计组成了 355 对亲子对. 该数据集从互联网中收集, 视频的平均时长为 18.78 s, 平均帧频率为 26.79 帧/s, 静态总帧数超过了 25 万. 与先前提出的视频亲子关系数据集不同, KIVI 数据集包括了加上兄弟姐妹关系在内的 7 种亲属关系, 在光照、姿势、遮挡情况、种族和表情方面没有约束, 且数据集中包括了多个人种.

12) TALKIN 数据集. TALKIN^[73] 数据集是第一个多模态的亲子关系数据集, 由人脸视频和语音两个模态的数据组成. TALKIN 数据集包含了从互联网中收集的 800 段视频, 包括了 4 种亲子关系, 父亲-儿子、父亲-女儿、母亲-儿子和母亲-女儿. 每种关系有一百对视频. 视频均为说话者视频, 包含人脸视频和声音信号. 视频中人物为名人或者来源于家庭综艺中的普通人. TALKIN 数据集包含了来自 9 个国家的人种, 其中白种人占有很大的比例. 视频在背景环境、人脸姿态、有无遮挡以及说话的文本内容等均无限制.

2.2 性能评估准则

实验中数据分为正样本对和负样本对, 正样本对是由数据集当中所有具有亲子关系的样本组成,

⁹ <http://parsec.nuaa.edu.cn/xtan/data/TSKinFace.html>

¹⁰ <https://web.northeastern.edu/smilelab/fiw/>

¹¹ <https://sites.google.com/a/mix.wvu.edu/namankohli/resources>

¹² <https://www.kinfacew.com/datasets.html>

¹³ <http://iab-rubric.org/resources/KIVI.html>

负样本对为随机产生的不具有亲子关系的样本. 通常正负样本对数量相同, 但也有研究工作为了生成更多的学习样本, 负样本会多于正样本. 由于亲子关系数据量往往比较小, 为了避免过拟合, 实验中采取五折验证方法. 其中, 四折作为训练集, 剩下的一折作为测试集, 通过五轮交替后求平均得出最终的结果. 需要特别注意的是, 每一折数据的正负样本对均在本折中产生, 不可以产生不同折之间的数据交叉.

识别准确率是亲子关系验证常用的评价指标, 准确率可以描述一个分类器在测试数据集上的分类能力^[74], 分类器的分类结果可以分为两部分: 正确分类部分和不正确分类部分. 正确分类部分包括正确分类的正样本 (True positive, TP) 与正确分类的负样本 (True negative, TN); 错误分类部分包括错误分类的正样本 (False positive, FP) 和错误分类的负样本 (False negative, FN). 准确率 A 的计算方式为

$$A = \frac{TP + TN}{P + N} \quad (2)$$

3 传统人脸亲子关系验证方法

短短十年的研究历史, 涌现了不少人脸亲子关系验证的研究工作, 经历了传统手工设计工程和以深度学习为代表的学习训练时代. 为方便总结, 我们将人脸亲子关系验证方法分为两个阶段进行总结: 传统的手工设计方法 (本节) 和基于深度学习的方法 (第 4 节). 在本节, 我们先介绍了传统人脸亲子关系验证的一般过程, 之后对传统方法分为两个大类介绍, 分别为传统人脸亲子特征提取方法和亲子度量学习方法.

3.1 人脸亲子关系验证方法框架

亲子关系验证的一般流程如图 3 所示, 主要由 4 个关键环节组成: 人脸检测、矫正和分割、亲子特征提取、距离计算和分类得出结果.

1) 人脸检测、矫正和分割. 这一步骤是对输入

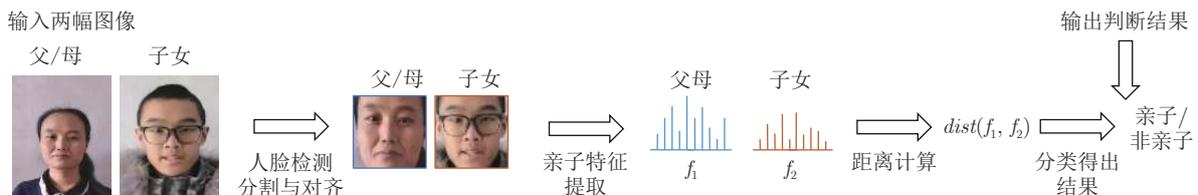


图 3 亲子关系验证的一般流程

Fig.3 The general framework of kinship verification

进行定位后, 通常以眼睛位置作为关键比对点, 对人脸进行对齐和矫正. 对齐和矫正的目的是减少由于尺度、人脸角度的影响, 常用的人脸分割和对齐方法有 MTCNN (Multi-task convolutional neural network)^[75] 和 ERT (Ensemble of regression trees)^[76]. 本文将不再对此进行详细介绍, 读者可参阅相关文献.

2) 亲子特征提取. 对给定的两幅输入人脸图像 I_1 , I_2 分别进行特征提取, 将每幅人脸图像表示成一个紧致的特征矢量, 即 $\phi(I_1)$ 和 $\phi(I_2)$, 用于后续的距离度量 and 分类. 这一环节属于亲子关系验证的一个核心研究内容, 也是决定其性能的关键所在. 在深度学习出现以前的传统特征工程时代, 常用的特征描述子有局部二值模式 (Local binary pattern, LBP)、二进制统计图像特征 (Binarized statistical image feature, BSIF)、尺度不变特征变换 (Scale-Invariant feature transform, SIFT)、方向梯度直方图 (Histogram of oriented gradient, HOG) 以及局部相位量化 (Local phase quantization, LPQ) 等. 深度学习出现以后, 传统的特征描述子已经逐渐被深度特征表达方法取代. 本文在第 3.2 节和第 4 节中按照传统特征工程方法和深度学习方法两类对代表性的人脸亲子特征表达方法进行详细的归类、分析和总结.

3) 距离度量. 经过亲子特征提取, 两幅人脸图像分别表示成两个固定维度的特征向量, 接下来需要采用合适的距离度量来计算两个特征矢量在特征空间中的距离, 来评估两幅人脸图像之间的相似性. 常用的距离度量方法有欧氏距离、马氏距离、余弦相似度等, 在基本距离度量上, 研究者尝试引入度量学习 (Metric learning) 算法, 度量学习由 Xing 等^[77], 在 NIPS 2002 (Conference and Workshop on Neural Information Processing Systems 2002) 第一次提出, 度量学习旨在学习能使具有亲子关系样本对 (即正样本对) 间的距离变小, 而非亲子关系样本对 (即负样本对) 间的距离变大的变换矩阵, 将提取的特征向量通过学习到的变换矩阵嵌入到新的特征空间, 以提高人脸亲子关系验证的性能^[15, 46, 78]. 这

部分将在第 3.3 节中进行详细介绍。

4) 分类. 经过上述步骤, 已经得到人脸图像样本对之间的距离, 接下来就是学习一个二元分类器对输入人脸图像样本对间的距离进行分类. 常用的分类器有 K 近邻, 支持向量机和阈值判别法。

在人脸亲子关系验证问题中, 亲子特征提取和距离度量起着更为重要的作用, 也是研究关注的焦点, 为此也是本文讨论的重点。

3.2 传统人脸亲子特征提取方法

解决人脸亲子关系验证问题首先要对人脸图像中的关键部位进行有效的特征表示, 特征提取方法研究的目的是在面部图像中提取包含亲子遗传属性的特征表示, 依据算法侧重不同及其拟解决的问题, 可以分为描述特征、面部显著特征、手工特征、基于颜色的特征变换和特征选择方法等. 首先, 描述特征由 Fang 等^[6]学者提出, 他们着手于面部底层特征, 从不同维度表示面部特征, 包括面部颜色属性: 眼睛颜色、肤色、发色; 五官间距离: 用欧氏距离测量不同部位间的距离以及其梯度特性; 脸型: 五官器官的大小. 之后 Xia 等^[79-80]对面部更多的信息进行描述, 包括图像人的年龄、性别、人种等, 这些属性用二值化的特征表示, 即 -1 和 $+1$. 最后将 20 种属性特征级联作为特征向量. 然而, 描述特征大多需要人为介入进行标注, 来定义特征描述的属性, 特征向量维度低, 信息表示片面, 不能够全面挖掘图像中的亲子特性. 随着人脸亲子关系验证逐渐获得更高的关注度, 研究员们提出了性能佳、效率高的自动特征提取方法, 接下来将对这些方法进行分类和汇总。

3.2.1 基于显著特征的亲子关系验证

这类方法旨在通过对面部的显著部位, 如眼睛、鼻子、嘴巴比较亲子间的相似度^[31-33, 81], 因此首先需要对人脸进行关键点定位. 给定一幅人脸图像, 对其提取显著性部位, Guo 等^[31]提出使用眼睛、嘴巴和鼻子作为面部显著性区域, 如图 4 所示, 再用 DAISY 描述子^[82]对每个区域提取相似性特征, 而后计算一对图像的相似性. Kohli 等^[81]提出使用角

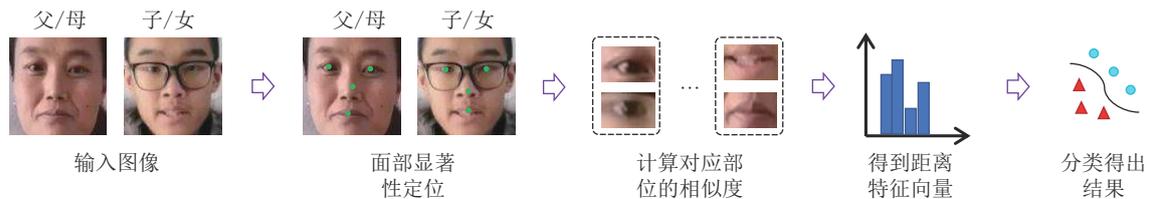


图 4 基于显著性区域的亲子关系验证算法框架图

Fig. 4 The architecture of saliency based kinship verification

点检测 (Differences of Gaussians, DoG) 方法定位面部关键区域, 之后 Wang 等^[32]在 2014 年将广泛应用的 68 个人脸特征点检测方法应用到亲子关系验证中来^[34], 并提出使用基于 Grassmann 流形的几何度量方法, 计算两幅人脸图像的相似性. 除了对人脸的关键点、关键部位检测和提取, Goyal 等^[33]提出基于边缘 (Edge) 检测的亲子特征提取方法, 作者首先用 Canny 边缘检测算子对两幅人脸图像提取边缘特征. Canny 算子具有低错误率的边缘检测, 最优的边缘定位以及对任意边缘仅标记一次的优良特性^[35], 在获得图像边缘信息后, 分析边缘条纹包裹的闭环面部区域作为最终获得的显著性区域。

3.2.2 手工特征方法

前面介绍的面部几何形状特征是在图像的像素级上检测和提取的, 受检测准确度、面部表情形变、噪声以及倾斜角度影响较大, 从而导致识别准确率, 在复杂条件下对噪声容忍度低. 针对这类问题, 研究学者提出使用手工特征提取方法^[14, 36-39, 83]. 其中局部二值模式 (LBP) 是应用最为广泛的手工特征表示提取方法, LBP 是一种用来描述图像局部纹理特征的算子, 它具有旋转不变性和灰度不变性等显著的优点, 最早由 Ojala 等^[84-85]在 1996 年提出, LBP 采用圆形邻域编码方法, 如图 5 所示. 在 3×3 大小的窗口内, 以中心像素作为阈值, 分别与邻域的 8 个像素进行比较, 生成 8 个二值码. 若周围像素大于阈值, 则标记为 1, 否则标记为 0.

由以上方法得到的 8 位无符号数即为该窗口的 LBP 编码, 这个编码可以反映区域的纹理特征. LBP 编码可以通过改变邻域半径和取值点, 得到不同维度下的特征表示. 设窗口半径为 R , 取 P 个邻域相似点, 则 (x_c, y_c) 点的 LBP 值计算式可以表示为

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3)$$

其中, g_c 为中心像素点, g_p 为邻域像素点. $s(x)$ 为符号函数. 经过 LBP 算子后, 特征采用 10 进制方

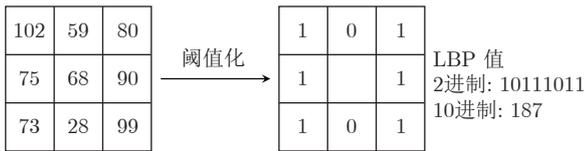


图 5 LBP 算法

Fig.5 The algorithm of LBP

式表达该点的二值模式, 共有 2^P 种, 在实际应用中可以对 P 和 R 值进行改变获得多尺度下的 LBP 表达^[38]. 通常, 在 LBP 特征提取前会将图像进行分块, 计算每个子块中每个像素的 LBP 值, 对子块进行直方图统计得到该子图像块的 LBP 值, 利用子图像块的 LBP 值即可以描述该图像的纹理特征. 图 6 给出了一个灰度图及其对应的 LBP 图谱, LBP 图谱中的每个像素值代表了该点的 LBP 值.

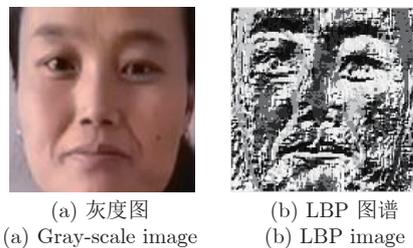


图 6 LBP 可视化特征

Fig.6 The visualization of LBP feature

在基础的手工特征基础上, 现有的方法也对其进行了不同程度的发展研究. 金字塔多尺度协方差描述子 (Pyramid multi-level covariance descriptor, PML-COV)^[35] 建立了金字塔手工特征, 将不同分辨率下的 HOG 和 LBP 两种特征与协方差描述子结合, 取得了传统特征提取方法中最好的识别率, 在 KinFaceW-I 与 KinFaceW-II 数据集上的识别率均达到 88.2%.

3.2.3 基于颜色的特征变换方法

传统手工特征往往在灰度图像上提取, 仅从光照强弱 (Luminance) 程度层面分析图像的特征表示, 而忽略了图像的色度信息 (Chrominance). 面部中的颜色包含了亲子间的遗传信息, 如: 眼睛颜色、肤色和发色等, 为了能够充分提取面部图像颜色信息, Wu 等^[12] 提出颜色纹理特征提取方法. 颜色纹理特征结合了图像中颜色和纹理两个方面信息, 使得表征具有更高的可判别性. 常用的颜色空间包括 RGB, HSV, YCbCr 等, 该方法首先将输入图像转换为目标颜色空间, 在目标颜色空间中对每个通道上的图像分别提取纹理特征, 再将各个通道的特征进行级联得到最终的颜色纹理特征, 如图 7 所示. 实验表明 HSV 颜色空间相较于其他颜色空间可以

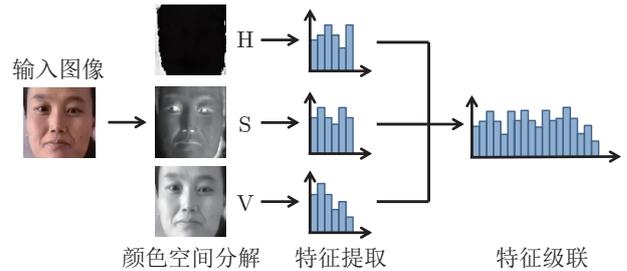


图 7 彩色纹理特征提取方法

Fig.7 The method of color-texture feature extraction

提供更为丰富的亲子信息, 并在 TSKinFace 数据集上取得了较好的识别准确率, 亲子平均识别率为 81.2%. 并且之后的相关研究均表明了颜色纹理特征在亲子关系验证问题上的有效性^[29, 40].

除了在已知的颜色空间提取特征表示, Liu 等^[13, 41] 提出遗传颜色空间 (Inheritable color space, InCS), 将亲子距离度量作为目标函数, 从而学习一个映射矩阵 W , 映射矩阵可以将原图像所在的颜色空间映射到新的颜色空间. 假设原图像对可以用两个矩阵表示, m_p 和 $m_c \in \mathbf{R}^{3 \times n}$ ($n = h \times w$), 在新颜色空间下的图像为 $x, y \in \mathbf{R}^{3 \times n}$. InCS 通过一个变换矩阵 W 将原图像变换得到新的颜色空间

$$x = W^T m_p$$

$$y = W^T m_c \quad (4)$$

其中, W 矩阵在目标函数下训练得到. InCS 颜色空间的各个颜色通道信息不相关, 减小了信息冗余度. 研究表明 InCS 对光照变化, (包括光强变化和光照颜色的变化) 鲁棒, 在多个数据集上的实验结果也表明了 InCS 相较于常用颜色空间的优越性.

3.2.4 特征选择方法及其他方法

与单一特征提取方法相比, 特征选择 (Feature selection) 方法旨在研究多个特征间的融合和选择方法, 从而可以丰富特征表示, 并去除冗余信息^[42-45]. 特征选择方法往往输入为多个特征, 在目标函数或者识别准确率的约束下, 从多个特征中选择出其中有效的表示方法. Alirezazadeh 等^[42] 在 2015 年首先提出融合局部特征与全局特征, 并从其中选取有效特征的方法. Bottinok 等^[43] 提取图像的多个特征, 包括频域局部相位量化 (Local phase quantization, LPQ)、韦伯局部描述符 (Weber's local descriptor, WLD) 和 LBP. 在进行特征分类之前, 为了提高识别准确率, 用最大相关和最小冗余算法 (Max-relevance and min-redundancy, mRMR) 从特征候选集选出一个最优子集作为最终的分器输入.

除此之外, 研究者也有用分类器做导向来进行

特征选择. Cui 等^[44]提出自适应的特征选择方法, 用 W 对不同特征进行选取, 对于特征 f_j , $j = 1, \dots, N$ 训练一个弱分类器 h_j , 对 N 个弱分类器的选择和优化通过以下目标函数实现:

$$\begin{aligned} \min_{\xi, \omega_j, b_j} & \|\omega_j\|^2 + C \sum_{i:y_i=1}^{N_i} W_i \xi_i \\ \text{s.t.} & y_i(\omega_j^T \Phi(x_{ij}) + b_j) \geq 1 - \xi_i, \quad i = 1, \dots, N, \\ & \xi_i \geq 0, \quad i = 1, \dots, N \end{aligned} \quad (5)$$

其中, W_i 为正则化参数, ξ_i 是每个特征的松弛变量, ω_j , b_j 是第 i 个支持向量机 (Support vector machine, SVM) 的超参数, C 是 SVM 的惩罚参数, x_{ij} 为 i 样本对的第 j 特征表示, $\Phi(x_{ij})$ 是输入空间的特征图, 有效的弱分类器组成一个强分类器. 与文献 [44] 类似, Chen 等^[45]采用典型相关分析 (Canonical correlation analysis, CCA) 方法, 来寻找一个多特征映射选择矩阵, 从而增大亲子关系对的特征相关性.

除了上述提到的特征提取方法, 还有一些学者从其他角度解决特征提取的问题. Fang 等^[30]选取多个不同人脸部分组成一个字典, 当输入新的面部图像时, 在对应器官图像字典中寻找其编码方式作为该图像的特征向量. Duan 等^[86]提出特征消减法去除特征中的噪声信息, 保留有效信息. Bessaoudi 等^[87]提取了特征的高阶表示, 而 Laiadi 等^[88]提出张量交叉视图二次判别分析 (Tensor cross-view quadratic discriminant analysis, TXQDA) 方法, 通过特征映射方法, 学习低维张量, 来减小由年龄、性别等多种因素带来的影响.

本小节对基于特征表示的人脸亲子关系验证方法进行了全面的总结, 从多个分类角度分析了方法的切入点和创新点. 特征提取的好坏直接影响到最终识别性能的高低, 传统的特征提取方法有计算复

杂度低、特征提取速度快的优点, 然而, 由于特征固化于特定的提取规则, 往往有不适用于复杂情况、抗干扰能力差等缺点.

3.3 亲子度量学习方法

在特征提取方法基础上, 我们期望找到一个最佳的度量方法去计算图像对间的距离, 使得具有亲子关系的图像对 (正样本对) 距离小, 而不具有亲子关系的图像对 (负样本对) 距离大. 这就引入了本小节将要介绍的研究内容, 基于度量学习的人脸亲子关系验证算法. 度量学习 (Metric learning) 旨在根据不同的任务自动学习出一个距离度量函数, 将距离度量空间映射到一个新的度量空间中. 这是由于在常用的度量空间中, 同一类样本往往不能够很好地聚类, 度量学习就是为了解决该类问题、更好地描述样本间的相似问题, 这也是人脸亲子关系验证研究的核心问题之一.

度量学习通常在距离度量/相似性度量的基础上增加一个映射矩阵, 依据任务不同构造损失函数, 通过迭代优化或最优解的方法得到映射矩阵, 从而达到映射度量空间的目的. 其中涉及到的距离度量包括欧氏距离^[46]、马氏距离^[15, 54, 78, 89-92]、双线性相似度^[47-49, 93-95]、图构造^[50-51]、余弦相似度^[55-56]、CCA^[96]及其他度量模型^[97-101], 本小节将对这些方法进行总结概述.

3.3.1 邻域驳斥度量学习算法

Lu 等^[15]2014年提出邻域驳斥度量学习 (Neighborhood repulsed metric learning, NRML) 算法, 这是度量学习在人脸亲子关系验证中的第一次尝试, 为度量学习在人脸亲子关系验证研究中的研究奠定了理论和实践基础. NRML 算法的思想是, 正样本相邻的负样本会对分类器产生干扰, 因此方法对正样本 k 个邻域中的负样本进行排斥, 同时使正样本互相吸引, 从而分离正负样本, 算法示意图如图 8 所示.

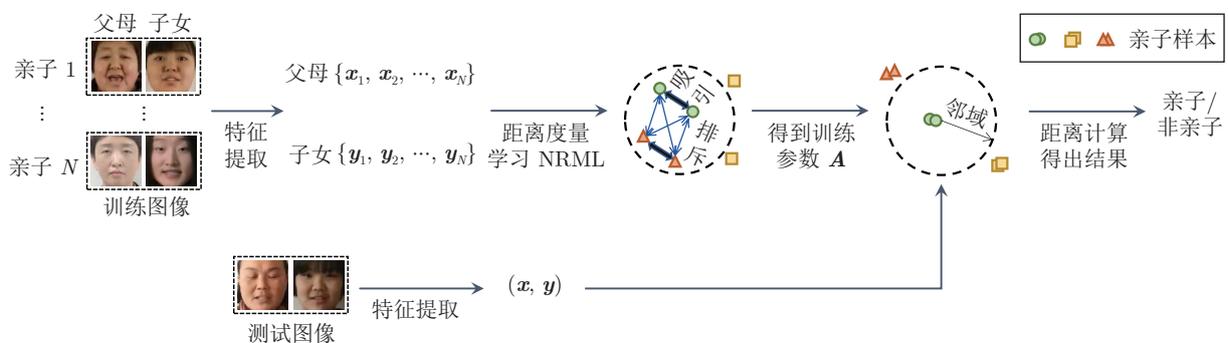


图 8 NRML 算法框架图

Fig.8 The architecture of NRML algorithm

我们用 $S = \{(\mathbf{x}_i, \mathbf{y}_i) \mid i = 1, 2, \dots, N\}$ 表示训练集, 其中, 包含了 N 对具有亲子关系的图像对, $\mathbf{x}_i, \mathbf{y}_i$ 为第 i 对亲子图像的特征表示, 用 m 维的列向量表示. \mathbf{x}_i 和 \mathbf{y}_j 之间的距离用马氏距离 (Mahalanobis distance) 表示为

$$d(\mathbf{x}_i, \mathbf{y}_j) = \sqrt{(\mathbf{x}_i - \mathbf{y}_j)^\top \mathbf{A} (\mathbf{x}_i - \mathbf{y}_j)} \quad (6)$$

其中, \mathbf{A} 是一个 $m \times m$ 的对称正定矩阵, 通过找到一个合适的 \mathbf{A} , 使得当 $i = j$ 时, 距离尽可能小; 而 $i \neq j$ 时, 距离尽可能大. 对 \mathbf{A} 构造目标函数如下:

$$\begin{aligned} \max_{\mathbf{A}} J(\mathbf{A}) &= J_1(\mathbf{A}) + J_2(\mathbf{A}) - J_3(\mathbf{A}) = \\ &= \frac{1}{Nk} \sum_{i=1}^N \sum_{t_1=1}^k d^2(\mathbf{x}_i, \mathbf{y}_{it_1}) + \\ &= \frac{1}{Nk} \sum_{i=1}^N \sum_{t_2=1}^k d^2(\mathbf{x}_{it_2}, \mathbf{y}_i) - \\ &= \frac{1}{N} \sum_{i=1}^N d^2(\mathbf{x}_i, \mathbf{y}_i) \end{aligned} \quad (7)$$

其中, \mathbf{y}_{it_1} 表示 \mathbf{y}_i 的 k 邻域中第 t_1 个负样本, 而 \mathbf{x}_{it_2} 表示 \mathbf{x}_i 的 k 邻域中第 t_2 个负样本. 式 (7) 中的前两项是为了驳斥 \mathbf{x}_i 和 \mathbf{y}_i 邻域中的负样本, $J_1(\mathbf{A})$ 以 \mathbf{x}_i 为基准, 最大化其与 k 个邻域负样本的距离; 反过来, $J_2(\mathbf{A})$ 以 \mathbf{y}_i 为基准, 最大化其与 k 个邻域负样本的距离, 而第 3 项 $J_3(\mathbf{A})$ 则使正样本对 \mathbf{x}_i 和 \mathbf{y}_i 距离尽可能小. 这样就实现了正负样本对的分离, 文献 [15] 中用迭代优化的方法对 \mathbf{A} 进行求解, 本文将不再赘述, 感兴趣的读者可以参阅相关文献.

NRML 算法在当时研究水平上取得了最好的识别率, 分别在 KinFaceW-I 和 KinFaceW-II 数据集上的平均识别率为 73.8% 与 69.9%. 因此, NRML 的邻域驳斥思想也被多个研究学者应用于其他度量方法上, Yan 等 [46] 提出将特征向量映射到 SVM 超平面 (Hyperplane) 上, 再通过构造邻域驳斥思想优化系数矩阵. Xu 等 [47] 用多特征级联作为图像的特征向量, 提出用双线性相似度结合 NRML 算法的度量学习方法. 之后, Yan 等 [55] 和 Lei 等 [96] 分别将度量方法用余弦相似度和 CCA 代替, 也证明了 NRML 算法的有效性.

除了上文提到的 NRML 算法, 研究学者在马氏距离的基础上也提出了其他度量学习算法, 从不同角度切入优化距离度量方式. Yan 等 [78] 引入概率模型, 最大化亲子关系距离小于非亲子关系距离的概率, 即 $P(d(\mathbf{x}_i, \mathbf{y}_i) < d(\mathbf{x}_i, \mathbf{y}_j))$, 通过最小化概率 P 达到优化矩阵 \mathbf{A} 的目的. 之后, Hu 等 [89-90] 提出大

裕量多维度量学习算法 (Large-margin multi-metric learning, LM3L), 该算法融合多角度特征的度量学习算法, 同时在分离正负样本的基础上, 增加一个距离裕量, 可以进一步拉大正负样本的距离, 如图 9 所示.

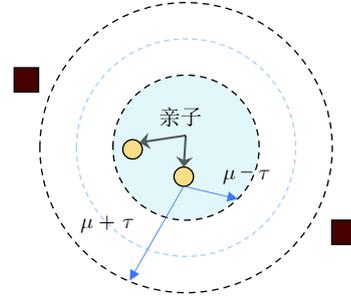


图 9 大裕量多维度量学习算法示意图

Fig.9 The illustration of LM3L

LM3L 在单一特征 (第 k 特征) 下的损失函数可以表示为

$$\min_{M_k} J_k = \sum_{i,j} h(\tau_k - y_{ij} (\mu_k - d_{M_k}^2(\mathbf{x}_i^k, \mathbf{x}_j^k))) \quad (8)$$

其中, $h(x) = \max(x, 0)$, J_k 为第 k 特征下的损失函数, $\mathbf{x}_i^k, \mathbf{x}_j^k$ 表示父/母和子/女的人脸在第 k 角度下的特征向量, $y_{i,j}$ 是标签, 表示是否具有亲子关系, 而 τ_k 和 μ_k 是阈值. 由图 9 可以直观地看出, 当图像对具有亲子关系时, $d_{M_k}^2(\mathbf{x}_i^k, \mathbf{x}_j^k) < \mu_k - \tau_k$; 不具有亲子关系时, $d_{M_k}^2(\mathbf{x}_i^k, \mathbf{x}_j^k) > \mu_k + \tau_k$. 在引入多角度的特征向量后, 在不同特征空间需要学习该特征空间下的度量矩阵参数, 即在 l 特征空间, 距离矩阵表示为 M_l . 作者为了进一步缩小由不同特征空间所带来的差距, 上述损失函数可以变为

$$\begin{aligned} \min_{\{M_k, w_k\}_{k=1}^K} J &= \sum_{k=1}^K w_k^p J_k + \\ &= \lambda \sum_{k,\ell=1, k<\ell}^K \sum_{i,j} (d_{M_k}(\mathbf{x}_i^k, \mathbf{x}_j^k) - \\ &= d_{M_\ell}(\mathbf{x}_i^\ell, \mathbf{x}_j^\ell))^2 \\ \text{s.t. } \sum_{k=1}^K w_k &= 1, w_k \geq 0, \lambda > 0 \end{aligned} \quad (9)$$

其中, w_k 是权值参数, 为了避免平凡解, 对权值取 p 次方. 第 1 项对不同特征空间学习权重函数; 第 2 项为了减少不同特征空间的差异性, 由此最终可以得到多特征的度量学习模型. 作者采用交替优化方法优化式 (9), 具体细节可参见文献 [78].

3.3.2 基于双线性相似度的度量学习方法

除了研究较多的马氏距离度量方法, 还有一种度量方式是双线性相似度(式(10)), M 是参数矩阵, $\mathbf{x}_i, \mathbf{x}_j$ 表示两个特征向量, 当 M 为单位阵 (Identity matrix) 时, 双线性相似度可以视为没有归一化的余弦相似度. 双线性相似度在图像检索上展现了较好的度量性能^[102-103], 并且对于稀疏特征向量, 其可以有效计算特征间相似度, 即

$$S_M(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T M \mathbf{x}_j \quad (10)$$

Zhou 等^[48-49] 提出四输入的相似度学习算法 (Ensemble similarity learning, ESL), 输入为 $\mathbf{x}_i, \mathbf{y}_i, \mathbf{x}_j, \mathbf{y}_j$, 其距离关系目标函数可以表示为

$$\begin{cases} S_M(\mathbf{x}_i, \mathbf{y}_i) \geq S_M(\mathbf{x}_i, \mathbf{y}_j) + 1 \\ S_M(\mathbf{x}_i, \mathbf{y}_i) \geq S_M(\mathbf{x}_j, \mathbf{y}_i) + 1 \end{cases} \quad (11)$$

其中, \mathbf{x}_i 和 \mathbf{y}_i 表示一对正样本对, 而 \mathbf{y}_j 和 \mathbf{x}_j 分别为对应负样本. 为了满足上述的约束条件, Zhou 等^[48-49] 提出如下损失函数:

$$\begin{aligned} \ell(S_M, \mathcal{D}) = & \frac{1}{nk} \sum_{\substack{(\mathbf{x}_i, \mathbf{y}_i) \in \mathcal{D} \\ \mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i)}} \ell_h(\mathbf{x}_i, \mathbf{y}_i, \mathbf{x}_j; S_M) + \\ & \frac{1}{nk} \sum_{\substack{(\mathbf{x}_i, \mathbf{y}_i) \in \mathcal{D} \\ \mathbf{y}_l \in \mathcal{N}_k(\mathbf{y}_i)}} \ell_h(\mathbf{x}_i, \mathbf{y}_i, \mathbf{y}_l; S_M) \end{aligned}$$

其中,

$$\ell_h(\mathbf{x}, \mathbf{y}, \mathbf{z}; S_M) = \max(0, 1 - S_M(\mathbf{x}, \mathbf{y}) + S_M(\mathbf{x}, \mathbf{z})) \quad (12)$$

其中, 1 是一个边缘阈值常量用于分离正负样本对. 作者提出使用梯度下降的迭代方法优化参数. ESL 算法在计算效率方面具有一定的优势, 因此其在实际应用中可以适应于高维特征表示. 其次, ESL 通过四输入的数据组, 满足每对图像相似度模型上的类内类间约束. 实验结果表明, ESL 在计算效率上优于现有的方法, 识别准确率在 KinFaceW-I 和 KinFaceW-II 上分别取得 74.1% 和 74.3%, 在当时获得最佳识别性能. Qin 等^[94] 提出多任务的双线性相似性学习方法, 将多种亲子关系任务结合, 用多任务 (Multi-task) 的思想进一步提高算法的鲁棒性. Fang 等^[95] 引入线性回归模型, 建立正样本的后验概率模型, 平滑目标函数, 进一步帮助最优解求解.

3.3.3 迁移子空间学习

基于迁移子空间学习的人脸亲子关系验证方法将子女图像与父母图像分别看作源域 (Source domain) 与目标域 (Target domain), 旨在通过学习一个矩阵映射, 缩小两个域之间的差异, 同时保留特征的可判别性^[52-53]. Xia 等^[52-53] 提出子女长相更相

像于年轻时父母, 为了保证源域到目标域的有效迁移, 将年轻父母的人脸图像作为源域到目标域的桥梁, 称为中间域 (Intermediate domain). Zhang 等^[54] 将该思想与 NRML 算法结合, 提出邻域驳斥迁移度量学习 (Neighborhood repulsed transfer metric learning, NRTML) 算法, 同时缩小子女与年轻父母之间的距离、子女与年老父母之间的距离以及年轻与年老父母之间的距离. 这类方法的缺点是同时需要年轻与年老时的父母人脸图像, 对数据搜集要求高, 难度大.

3.3.4 其他度量学习方法

除了上述应用中较为广泛的度量学习方法, 研究学者也从其他角度对亲子度量方法进行了研究. Zhang 等^[97] 提出三人组的亲子关系度量方法, 在高维特征空间中, 距离由子/女点到父母点连线的垂直距离表示. Liu 等^[98-99] 通过向量的对角向量表示距离关系, Wu 等^[100] 引入低秩的度量学习, 通过自适应的隐子空间学习方法挖掘更具有可判别特征表示. Zhao 等^[101] 提出多核的距离度量方法, 包括线性和非线性度量方法, 在不同核函数下计算距离后加权融合. 图学习方法也在亲子关系度量中展现了其优越性和有效性, Liang 等^[51] 依据数据节点间的距离关系, 将 NRML 思想和图理论结合, 通过构造本质图 (Intrinsic graph) 和惩罚图 (Penalty graph), 同时描述类内紧凑性和类间分离性, 在公开数据集上取得了较好的识别效果, 在 TSKinFace 数据集上的亲子平均识别率为 90.5%, 在 KinFaceW-I 和 KinFaceW-II 数据集上的平均识别率分别为 78.7% 和 82.8%.

度量学习可以看作是对特征向量的聚类操作, 通过不断优化学习, 拉近具有亲子关系的图像对, 而不具有亲子关系的图像对互斥. 本小节对现有的亲子度量学习算法进行了归纳总结, 传统的度量学习方法是在特征表示基础上提出的, 与特征提取方法是两个独立的模块. 深度度量学习将特征提取和距离度量结合, 用样本间的距离度量引导网络学习表征能力强的特征提取方法. 通常深度网络抽象特征向量, 相似性度量作为损失函数, 通过优化损失函数, 不断更新网络参数, 最后可以得到判别力强的特征向量.

4 基于深度学习的人脸亲子关系验证方法

传统手工特征的描述能力有限, 很难适应复杂场景下的大规模数据任务. 基于卷积神经网络 (Convolutional neural network, CNN) 的深度学习方具

有很好的特征描述能力, 通过对大规模的数据进行训练分析, 自适应地从原始数据中提取有效特征向量, 避免了手工特征单一、特定的特征提取规则。

随着深度学习在计算机视觉领域的不断渗透和大规模亲子数据集的提出, 2016 年以后研究人员的研究重心逐渐向深度学习方法迁移. 相较于传统方法, 深度学习方法通过设计神经网络, 自适应地挖掘图像的高层次特征, 一般通过端到端 (End-to-end) 的训练方法获得任务的训练模型 (Model). 深度学习方法在复杂情况下表现出了优越的性能, 对光照、遮挡等因素鲁棒性高. 人脸亲子关系验证算法中涉及到多个先进深度学习框架, 包括基本卷积神经网络、深度度量学习、基于自编码器的框架和注意力机制神经网络框架. 下面将对这些算法进行汇总和介绍。

近年来, 随着深度学习方法在计算机视觉和语音处理领域得到成功应用, 其在人脸亲子关系验证问题中也取得了不少进展^[16-20, 57-63, 65, 72-73, 104-109]. 深度学习网络包含了许多隐含层及通过损失函数的反向传播, 使得其具有相对于传统手工特征的强大的表达能力, 以及对特定问题的学习能力. 2015 年, Wang 等^[16] 首次提出用深度网络解决亲子关系验证问题, 该方法分为两个阶段, 特征提取与深度度量学习. 其中, 面部的非线性特征通过自编码器 (AutoEncoder) 提取, 而后, 基于优化马氏距离的深度度量学习, 可以更好地分离正负样本. 然而, 该方法的输入为 LBP 特征, 丢失了原始图像的细节特征. Zhang 等^[57] 开启了基于端到端 (End-to-end) 深度学习的人脸亲子关系验证方法, 框架图如图 10 所示, 是一个基本的卷积神经网络 (CNN). 网络的输入为两幅人脸图像, 经过多层的网络卷积最终输出判断结果. 该网络架构简洁明了, 仅用单一卷积网络对两输入分析其之间的关系情况. 尽管网络结构比较单一, CNN 在公开数据集上同样表现出了很好的识别性能, 在 KinFaceW-I 和 KinFaceW-II 数据集上的平均识别率分别为 77.5% 和 88.4%, 该项工作奠定了深度学习在人脸亲子关系验证研究中的研究基础。

4.1 深度度量学习方法

在卷积神经网络的基础上, 为了优化学习输入的亲子图像间的距离, 研究者引入距离度量参与网络训练, 即深度度量学习 (Deep metric learning) 方法^[17, 58-62]. 典型的网络结构为孪生网络 (Siamese network), 如图 11 所示. 不同于基于表征学习的深度网络, 度量学习旨在通过优化损失函数, 寻找到一个最优的特征空间, 使得具有亲子关系的图像对 (Positive pair) 距离缩小, 不具有亲子关系

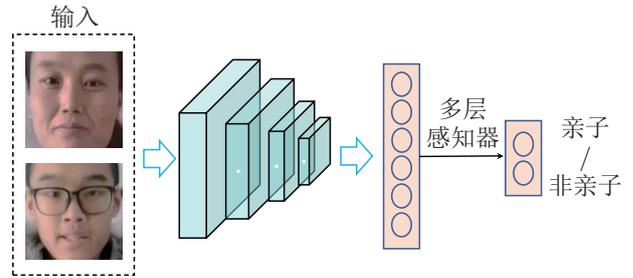


图 10 基于基本 CNN 的亲子关系验证框架图

Fig. 10 The architecture of the basic CNN based kinship verification

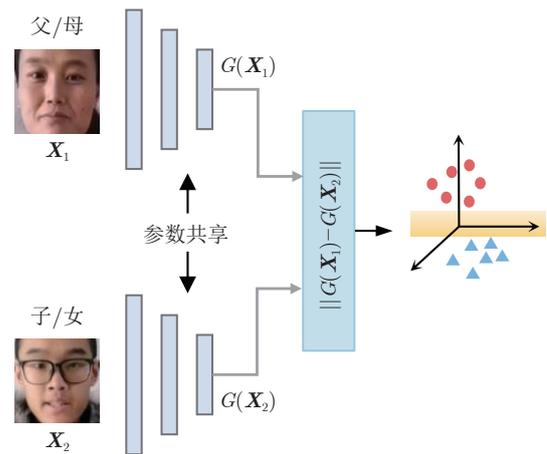


图 11 基于孪生网络的亲子关系验证框架

Fig. 11 The Siamese network based kinship verification

的图像对 (Negative pair) 距离变大, 而该特征空间即为我们训练所得到的网络。

Li 等^[17] 提出基于相似度量的卷积神经网络 (Similarity metric based convolutional neural networks, SMCNN), 输入为两幅人脸图像 X_1 和 X_2 . $G(\cdot)$ 表示网络的全连接层特征输出, 两幅图像经过参数共享的网络得到特征的嵌入表示, 特征间的距离用 L_1 范数表示, 如式 (13) 所示。

$$D(X_1, X_2) = \|G(X_1) - G(X_2)\|_1 \quad (13)$$

训练过程中, 为了区分正负样本对, 增加一个阈值参数 τ , 使得正样本对间距离大于 τ , 而负样本对间距离小于 τ . 正负样本的标签分别用 $y = 1$ 与 $y = -1$ 表示, 此时, 可以得到网络的损失函数

$$L_{SMCNN} = f(1 - y(\tau - D(X_1, X_2))) \quad (14)$$

其中, $f(\cdot)$ 表示归一化的逻辑回归. 通过不断地优化最小化损失函数, 可以得到最终的参数用于问题的求解. 此外, 常用的孪生网络损失函数还包括对比损失函数 (Contrastive loss) 和三元损失函数 (Triplet loss)^[17-18, 73], 这两个损失函数均基于距离

度量, 如欧氏距离和余弦相似度, 用 $d = D(\mathbf{X}_1, \mathbf{X}_2)$ 表示. 同样, y 表示输入图像对的标签. 对比损失函数可以表示为

$$L_c = yd^2 + (1 - y)(\alpha - d)_+^2 \quad (15)$$

其中, $(x)_+ = \max(x, 0)$, 参数 y 的 0 或 1 取值就使得损失函数在正负样本情况下保留特定项. 当 $y = 1$ 时 (正样本), 优化 d 使其逐渐缩小, 从而使得亲子关系对间距离变小. 反之, 当 $y = 0$ 时 (负样本), 负样本对会逐渐分离, 距离大于预设定的 α 阈值. 不同于对比损失函数, 三元损失函数的输入为三元组 $\langle a, p, n \rangle$, 其包括 3 个样本, 分别为, 固定图像 (Anchor) a 、正样本图像 (Positive) p 和负样本图像 (Negative) n , 正负样本均参照输入固定图像而言. 三元组损失函数表示为

$$L_t = (d(a, p) - d(a, n) + \alpha)_+ \quad (16)$$

三元损失函数以固定样本作为参考, 在拉近正样本的同时远离负样本. 在两个条件作用下, 达到聚类正样本对的目的, 实现正负样本对的分离, 如图 12 所示. 然而, 对于深度度量学习而言, 网络的训练样本对直接决定网络训练的效率和网络性能, 因此研究者提出一系列难样本挖掘 (Hard negative samples mining) 方法^[110]. 通常难样本挖掘方法致力于在训练批 (Batch) 中寻找距离较远的正样本对以及距离相近的负样本对, 这些样本会产生较大的误差回传, 从而有效训练网络. 由于深度度量学习在人脸亲子关系验证中的研究尚处于较初级阶段, 且人脸亲子关系验证由于环境等因素, 问题难度大, 至今难样本挖掘方法的研究在人脸亲子关系领域中仍处于空白.

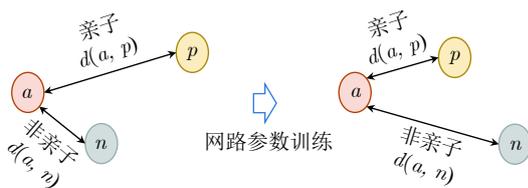


图 12 三元损失函数示意图
Fig.12 The illustration of triplet loss

4.2 基于自编码器的方法框架

另一个解决亲子关系验证问题的框架基于自编码器 (Auto-encoder, AE)^[16, 19, 72, 104, 106, 111-114], 早期自编码器在人脸亲子关系验证领域的应用主要针对特征的无监督训练, 如图 13 所示. 自编码器包含两个部分: 编码器和解码器, 输入为特征向量, 经过多层隐含层 (Stacked autoencoders) 后得到输出特征向量, 此时的输出近似于输入. 其中编码特征 (En-

coded feature) 可以作为输入的特征用于任务. 自编码器通过无监督的学习方式, 对输入特征进行降维和去噪.

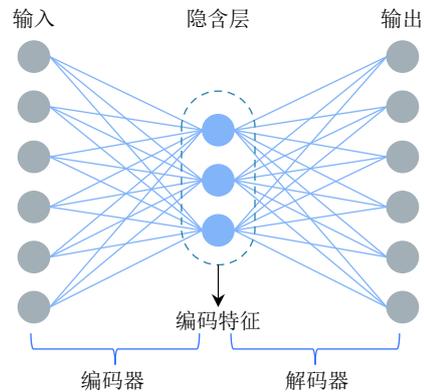


图 13 自编码器框架图
Fig.13 The architecture of auto-encoder

由于自编码器输入与输出相似的特点, 如图 14 所示, Dibeklioglu 等^[18] 将亲子对作为自编码器的输入, 自编码器的输出不仅与原图像相似, 又使其和自己的亲子相似, 从而得到具有亲子表示的特征向量. 用 x_1, x_2 表示两幅输入图像, 经过自编码器后的输出分别为 \hat{x}_1 和 \hat{x}_2 . 则亲子损失函数可以表示为

$$L_{kin} = C(x_1, \hat{x}_2) + C(x_2, \hat{x}_1) \quad (17)$$

其中, $C(\cdot, \cdot)$ 表示两幅图像的余弦相似度, 通常面部图像会用大数据预训练的深度学习提取特征向量, 如 VGG (Visual geometry group) 网络. 亲子损失函数描述了亲子对间的相互转化, 保留了亲子间的有效特征表示. 而非亲子损失函数是建立在亲子图像与其非亲子图像间的, 对于父母和子女的非亲子图像表示为 n_1 和 n_2 , 则非亲子损失函数为

$$L_{non-kin} = -C(\hat{x}_2, n_1) - C(\hat{x}_1, n_2) \quad (18)$$

结合亲子损失函数和非亲子损失函数, 可以得到

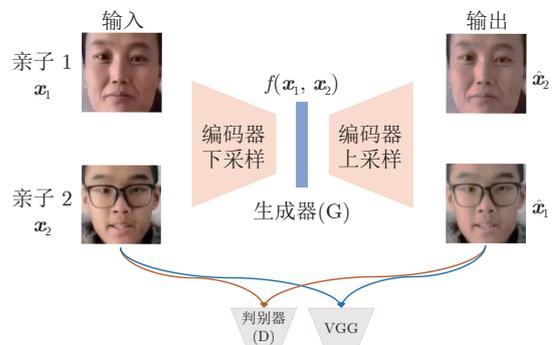


图 14 基于自编码器的亲子验证/图像合成框架图
Fig.14 The architecture of auto-encoder based kinship verification/kin face synthesis

$$L_{ac} = \lambda L_{kin} + (1 - \lambda) L_{non-kin} \quad (19)$$

其中, λ 是一个权重参数, 用来控制亲子损失函数和非亲子损失函数间的贡献程度. 通过对自编码器的训练, 可以得到编码特征 $f(\mathbf{x}_1, \mathbf{x}_2)$, 描述输入图像间的关系. 此外, 也有部分研究者利用其图像生成特性, 研究亲子图像合成技术^[19, 63, 64]. 生成对抗网络 (Generative adversarial networks, GAN) 是一个无监督的学习模型, 对子女图像的生成可以实现数据域之间的转换. GAN 包括两个部分, 分别为生成器 (Generator, G) 和判别器 (Discriminator, D), 生成器接受父母的图像而合成子女的图像, 缩小与真实子女图像间的距离. 同时判别器用对抗学习的方法, 判断生成的图像是否是真实图像. 因此在训练过程中, 生成器的目标是合成逼真的子女图像去欺诈判别器, 而判别器在尽力区分真实图像与合成图像. 通过这样的一个博弈过程, GAN 就可以完成对子女图像的合成.

4.3 基于注意力机制的框架

Yan 等^[65] 就亲子遗传线索往往存在于特定的局部面部区域, 而非整个人脸的问题, 提出基于注意力机制的亲子关系验证方法. 通过注意力机制, 学习面部局部特征表示方法, 对面部不同区域进行加权操作, 对比与传统的卷积神经网络, 该方法可以学习到局部区域的高维特征, 从而去除噪声和无效信息的干扰, 提高方法的性能, 算法如图 15 所示.

Yan 等^[65] 采用残差注意力机制方法, 对原特征图进行加权. 首先输入特征图 (Feature map), 经过一个池化和卷积, 为了使得权值和原特征图大小相同, 用上采样操作得到一个二维的特征图. 特征点通过 Sigmoid 函数, 将特征图上的每个点都归一化到 0 到 1 之间. 用 $C(\mathbf{X})$ 表示输入, 即待加权的特征图, $F(\mathbf{X})$ 代表得到的注意力图, 得到的经过注意力操作的特征图为

$$P(\mathbf{X}) = F(\mathbf{X}) * C(\mathbf{X}) \quad (20)$$

其中, $*$ 表示点乘. 但是, 由于提出的注意力权值通用于特征图所有通道, 而权值在 0 到 1 之间, 会造成丢弃深层特征值, 因此网络增加残差 (Residual) 结构, 可以得到最终的特征图表示为

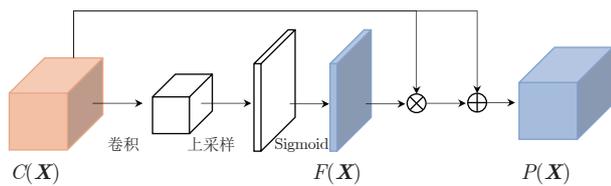


图 15 注意力机制示意图

Fig. 15 The illustration of attention mechanism

$$P(\mathbf{X}) = (1 + F(\mathbf{X})) * C(\mathbf{X}) \quad (21)$$

当 $F(\mathbf{X})$ 的值接近于 0 时, $(1 + F(\mathbf{X}))$ 的值接近于 1, 不会对原特征图造成特征消失的影响, 保留了网络的高维信息. 进一步, 为了更加准确地定位亲子特征部位, 指导网络对局部区域的学习作用, 作者对不同面部区域遮挡再学习注意力权值, 例如, 眼睛、鼻子、嘴巴部位. 包括全脸和部分遮挡的面部图像在内, 网络输入 6 幅面部图像, 可以得到一个合成的网络, 全面地描述面部信息. 注意力机制网络在 KinFaceW-I 和 KinFaceW-II 数据集上的平均识别率为 82.6% 和 92.0%, 是目前算法中性能比较好的.

注意力机制网络借鉴了人类观察时选择性的注意力特点, 可以根据任务的不同, 学习注意力方法. 在亲子关系验证的问题上尤为适用, 生物学家和心理学家发现对于长相遗传, 眼睛部位会存在大量的遗传信息. 相比于眼睛, 嘴巴部位由于状态多样化, 并不能够提供充足的遗传线索, 尤其对于静态图像而言. 基于注意力的亲子关系验证研究尚处于一个初级阶段, 而今后如何将度量结合注意力学习是研究的关键点之一.

4.4 其他框架

除了上述提到的几种典型的深度学习框架, 研究者们也在包括外观模型和几何模型分离、多模态融合、基于图网络的框架和数据扩充问题上提出了相关解决办法.

Zhang 等^[107] 提出分离外观模型与几何模型的方法, 结合全局面部外貌和脸型形状两个方面因素分析亲子特征. 外观特征基于人的身份信息得到, 脸型信息通过面部关键点得到. 然而脸型信息的获取受到角度和姿态的影响, 因此提出仿射不变性的形状特征提取方法, 详细的方法细节参见文献 [108]. 实验表明, 两个维度特征的融合使得识别准确率在单特征基础上提高了 10%.

Wu 等^[73] 就子女不仅长相与父母相似, 在说话声音上也具有一定的相似性特点, 提出融合语音特征和面部特征来解决亲子关系验证问题, 为亲子特征提供更加全面的信息. 网络结构分为两个部分: 声音网络和视觉网络, 分别在大规模的说话者数据集和人脸数据集上预训练后进行微调. 融合模块采用对比损失方法, 学习融合机制, 首先将语音和视觉特征级联, 经过全连接层输出融合后的特征表示. 实验结果表明, 多模态融合相较于单一模态可以进一步提高识别准确率.

Li 等^[109] 发现对比亲子关系时, 往往通过对比

对应的面部属性, 而 CNN 的输出特征的每一维可以看作是一种基因特征. 基于此, 作者提出一种基于图的关系推理方法. 两个输入图像的特征向量 $g(\mathbf{x}_i)$ 和 $g(\mathbf{y}_j)$ 的每一维级联后得到图的 N 个特征节点, $\mathbf{h}_d = [g_d(\mathbf{x}_i) || g_d(\mathbf{y}_j)]$. 为了融合不同维度上的关系, 在所有节点中选出一个节点作为超级节点, 剩余节点均和该节点连接, 由此建立了亲子关系推理网络. 通过多层图卷积后, 将二值交叉熵损失函数作为目标方程来训练整个图网络.

数据量不足是影响亲子关系验证研究在深度学习领域发展的一个重要因素, Song 等^[110] 提出 KINMIX 的亲子数据扩充方法. 该方法在特征级而不是原始图像上进行扩充. 作者提出两个亲子图像特征向量间的线性变化后的特征向量, 仍与原向量具有亲子关系, 通过改变线性变化的权值, 可以扩充出大量的特征用于训练. 实验表明, 扩充后相较于原始数据会对方法性能有一定的提升.

深度学习算法具有较好的学习能力, 在人脸亲子关系验证中具有很大的发展潜力, 也是未来的主要研究方向之一. 为了更加清楚直观地对比现有基于深度学习的人脸亲子关系验证算法, 本文在表 4 中对其在多个维度上进行了汇总比较, 并分析了算法所针对的问题.

5 现有方法性能比较

本节将对现有的人脸亲子关系验证算法进行性能比较, 由于人脸亲子关系验证的方法很多, 我们无法对所有的方法穷举. 因此, 本文选取了 3 个研究方向中具有代表性的研究工作, 以及发表在顶级会议和顶级期刊上性能好、具有里程碑意义的方法. 表 5 总结了 10 余种最具代表性的亲子关系验证算法在常用基准测试数据集上的性能, 分别从特征提取、度量学习和深度学习 3 个方面选取几种最具代表性的方法. 常用基准测试数据集包括亲子静态人脸数据集 KinFaceW-I、KinFaceW-II、TSKinFace、FIW 和 Cornell KinFace 等, 视频数据集包括 Smile、TALKIN 和 KIVI. 为了方便比较, 我们选取了应用较多的几种亲子关系, 包括: 父亲-儿子 (FS)、父亲-女儿 (FD), 母亲-儿子 (MS), 母亲-女儿 (MD).

传统人脸亲子关系包括两个方面, 分别为基于特征提取的方法和基于度量学习的方法. 人脸亲子关系验证研究初期以传统特征提取方法为主, 识别率相对较低. 以第一个人脸亲子关系验证方法“计算模型”为例, 平均识别准确率仅有 65.7%. 随着人脸特征表示研究的不断进步, 基于特征提取的方法取得了比较大的性能提升. 颜色特征方法^[12] 在 TSKinFace 数据集上的平均识别准确率为 81.2%, PML-

COV^[38] 方法在 KinFaceW-I 和 KinFaceW-II 数据集上的平均识别率均可以达到 88.2%. 在特征提取的基础上, 度量学习方法从距离度量的角度着手, 进一步优化亲子和非亲子图像对的距离度量方式, 提高正负样本的识别准确率. MNRML 算法作为度量学习在人脸亲子关系验证研究中的里程碑工作, 取得了当时最好的识别准确率, 分别在 KinFaceW-I 和 KinFaceW-II 数据集上的平均识别准确率为 73.8% 和 69.9%. WGEML 方法引入图结构, 得到了更高的识别准确率. 传统方法为人工设计的特定规则, 在复杂情景下鲁棒性低, 从表 3 中可以看出传统方法均针对于小规模数据集, 涉及的情景单一、数据收集环境也比较理想. 深度学习方法进一步实现了算法性能的提升, 其中 SCCAE 算法在 Smile 数据集上取得了 93.3% 的平均识别准确率, Attention 算法在 KinFaceW-I 和 KinFaceW-II 数据集上击败了传统方法, 识别准确率分别为 82.6% 和 92.0%. 随着小数据集上识别准确率趋近于饱和, 研究学者提出大规模的人脸亲子关系验证, 目前, FIW 是最大的人脸亲子数据集, 数据采集于无约束条件下, 问题复杂度高, 现有的识别准确率仅有 68.5%, 大规模自然环境下的人脸亲子关系验证问题难度大, 仍需进一步研究.

人脸亲子关系验证也受到多方面因素的影响, 从数据采集条件角度分析, 包括录制环境以及人脸图像来源等因素; 从人脸特性而言, 影响人脸亲子关系验证的因素包括面部表情、年龄和性别; 此外, 亲子数据集规模也是影响亲子验证识别准确率的因素之一, 下面我们将依据对应实验结果对这些因素进行分析.

1) 数据采集. 数据集采集的条件包括两种, 一种是在自然环境下无约束的数据采集, 一种是约束条件下录制. 约束条件下, 数据采集的各个参数均保持一致, 包括背景、光照、受验者面部角度、面部表情、面部是否遮挡以及采集数据的设备和拍摄位置. 以 Smile 亲子数据集为例, 由于 Smile 数据集是在约束条件下拍摄的, 受到来源于外部因素的干扰可以降到最低, 因此识别准确率处于较高的水平, SCCAE^[18] 方法取得了 93.3% 的平均识别准确率. KinFaceW 数据集研究了亲子人脸图像来源于同一张照片和不同照片的差异性, 从表 3 的实验结果可以看出, KinFaceW-I 数据集中亲子图像来自于不同照片, 其平均识别率相对低于亲子图像来自同一照片的 KinFaceW-II 数据集, 这是由于图像背景、光照等因素的相似性为亲子验证提供了参考, 从而简化了问题复杂度. 文献 [115] 对此进行了详细的分析, 感兴趣读者可以参阅相关文献.

表 4 基于深度学习的亲子关系验证方法总结
Table 4 The summary of deep learning based kinship verification

方法	特征	图学习	距离度量	自编码器	数据类型	针对的问题
DKV ^[16] (MMSP2015)	全局特征	否	马氏距离	是	图像	针对亲子验证中的非线性问题, 提出用自编码器及深度度量学习改进图像特征和距离度量.
CNN-Points ^[57] (BMVC2015)	局部特征	否	神经网络	否	图像	对比于传统手工特征, 第一个提出端到端的深度学习算法, 并将面部分块作为网络输入.
CMDAE ^[104] (IJCAI2016)	全局特征	否	欧氏距离	是	图像	将父母和子女看做两个不同的域, 年轻时父母作为中间桥梁, 减小由年龄带来的差异性.
SMCNN ^[17] (ICIAR2016)	全局特征	否	L1范数	否	图像	将深度学习网络和度量学习结合, 提出使用L1范数作为距离度量, 端到端的更新网络参数.
AdvNet ^[58] (RFIW2017)	全局特征	否	欧氏距离	否	图像	大数据下的亲子关系验证问题, 提出使用多个对抗对比残差网络提取图像的多角度特征.
DDML ^[59] (TIP2017)	全局特征	否	欧氏距离	否	图像	针对传统度量学习大多为线性变换, 提出多特征的深度度量学习, 将亲子面部特征映射到非线性平面.
CFT ^[105] (ICCVW2017)	全局特征	否	欧氏距离	否	图像	针对亲子关系数据量小时, CNN 网络容易过拟合问题, 提出人脸迁移学习方法.
KimNet ^[60] (RFIW2017)	全局特征	否	马氏距离	否	图像	引入 Triplet 损失函数训练网络, 并通过改变分辨率、加噪声等操作增加训练数据量.
DRF ^[106] (CCCV2017)	局部特征	否	神经网络	是	图像	对于大部分方法都基于相似度度量, 提出使用自编码器度量亲子关系距离.
SCCAE ^[18] (ICCV2017)	全局特征	否	后验概率	是	视频	对于亲子间具有相似的表情, 通过表情匹配方法, 分析亲子间表情的关联性.
Appearance+Shape ^[107] (ICIP2019)	全局特征	否	内积	否	图像	将面部的外观模型和几何模型进行分离, 从两个维度考虑亲子对的相似性.
KML ^[61] (Inf Fusion2019)	全局特征	否	余弦相似度	否	图像	针对隔代因素间的差异性, 提出紧凑型深度度量学习方法. 方法使用四输入比较其相似性与非相似性.
Attention ^[65] (PRL2019)	局部特征	否	神经网络	否	图像	针对亲子特征往往存在于人脸区域中, 并非全脸, 提出注意力机制的亲子关系验证方法.
SMNAE ^[72] (TIP2019)	全局特征	否	神经网络	是	视频	为解决亲子关系验证问题大都从静止图像分析, 提出空间-时间的亲子关系验证框架.
Fusion Kinship ^[67] (ICB2019)	全局特征	否	余弦相似度	否	视频+语音	针对现有亲子关系验证仅从视觉信息入手, 论文融合声音信号来提高模型的性能.
AdvKin ^[62] (TCYB2020)	局部特征	否	欧氏距离	否	图像	提出细粒度和更加严格的度量方法. 将人脸图像进行分块, 结合对比损失函数与对抗函数优化网络.
GKR ^[109] (ICME2020)	全局特征	是	神经网络	否	图像	针对如何进行特征融合及亲子关系推理. 对 CNN 输出的特征向量, 通过图结构分析两幅图像相似性.
KINMIX ^[110] (ICME2020)	全局特征	否	神经网络	否	图像	就数据量不足的问题, 在图像特征级上进行线性变换进行数据扩充, 数据可靠性高.

2) 面部表情. 心理学研究表明亲子关系验证准确率受到面部表情^[27]变化的影响, 相比于不同面部表情, 待验证的两幅人脸图像在无面部表情时识别率较高. 表 5 中, 文献 [18] 采用微笑的人脸亲子数据进行测试, 实验结果表明了具有亲子关系的两个人不仅在长相上相似, 也具有相似的微笑表情.

3) 年龄和性别. 人脸面部随着年龄的变化存在结构和纹理上的差异^[116], 而这些差异对人脸亲子关系类内距离产生影响, 从而降低识别性能. UBKin-Face 数据集包含子女、年轻父母以及年老父母的人脸图像, 在表 5 的 DMML^[78] 算法实验结果可以看到, 年轻和年老父母在亲子关系验证上性能相差 4.5%. 在亲子关系性别上的差异表现于父亲-女儿和母亲-儿子这两种亲子关系, 性别不同在人脸属

性上会产生一定的差异^[117], 从表 3 中统计的实验结果分析, FS、MD 关系的识别性能相对优于 FD、MS 关系.

4) 数据集规模. 自人脸亲子关系验证研究以来, 亲子数据集的建立经历了由小规模到大规模的发展过程. 目前大规模亲子数据库为 FIW, 其余数据库大都属于小规模数据库. 小规模亲子数据库涉及情景单一, 数据搜集条件比较理想, 受到来自遮挡、视角等因素的干扰少, 人脸图像提供了丰富的面部视觉信息. 小数据库中样本量少更容易学习类内的聚敛性和类间的分离性, 从表 5 的实验结果可以看出, 在小数据库上, 算法识别性能已接近饱和. 而当数据量剧增时, 以 FIW 数据库为例, 每个个体具有多幅不相同的人脸图像, 人脸形态在年龄维度、装

表 5 人脸亲子关系验证方法识别准确率对比 (%)
Table 5 The comparison of accuracies or kinship verification methods (%)

数据集	方法类别	方法	特征	度量/分类器	FS	FD	MS	MD	平均值	
Cornell KinFace	特征提取方法	计算模型 ^[6]	低层次特征	SVM	72.9	54.6	73.8	61.3	65.7	
	度量学习方法	PDFL ^[46]	LBP, SPLE, SIFT	SVM	74.8	69.1	77.5	66.1	71.9	
		DMML ^[78]	LBP, SPLE, SIFT	SVM	76.0	70.5	77.5	71.0	73.8	
KinFaceW-I	特征提取方法	D-CBFD ^[39]	LBP	阈值	79.6	73.6	76.1	81.5	77.6	
		特征选择 ^[86]	LBP, LPQ, SIFT	高斯概率	75.4	63.8	69.9	74.6	70.9	
		PML-COV ^[38]	HOG, LBP	SVM	91.0	84.3	87.1	90.2	88.2	
	度量学习方法	MNRML ^[15]	LBP, SIFT, LE, TPLBP	SVM	72.5	66.5	66.2	72.0	69.9	
		PDFL ^[46]	LBP, SPLE, SIFT	SVM	73.5	67.5	66.1	73.1	70.1	
		ESL ^[48]	LBP	双线性相似度	81.7	71.1	69.6	74.3	74.1	
	深度学习	DMML ^[78]	LBP, SPLE, SIFT	SVM	74.5	69.5	69.5	75.5	72.3	
		WGEML ^[51]	LBP, HOG, SIFT, CNN	KNN	78.5	73.9	80.6	81.9	78.7	
		CNN-points ^[57]	深度特征	神经网络	71.8	76.1	84.1	78.0	77.5	
		SMCNN ^[17]	深度特征	阈值	75.0	75.0	72.2	68.7	72.7	
		DDML ^[50]	深度特征	阈值	86.4	79.1	81.4	87.0	83.5	
		AdvKin ^[62]	深度特征	神经网络	76.6	77.3	78.4	86.2	79.6	
	Attention ^[65]	深度特征	神经网络	81.2	85.9	78.2	85.2	82.6		
	KinFaceW-II	特征提取方法	D-CBFD ^[39]	LBP	阈值	81.0	76.2	77.4	79.3	78.5
			特征选择 ^[86]	LBP, LPQ, SIFT	高斯概率	82.4	76.2	76.6	73.2	70.0
PML-COV ^[38]			HOG, LBP	SVM	88.6	85.8	87.2	91.0	88.2	
度量学习方法		MNRML ^[15]	LBP, SIFT, LE, TPLBP	SVM	76.9	74.3	77.4	77.6	76.5	
		PDFL ^[46]	LBP, SPLE, SIFT	SVM	77.3	74.7	77.8	78.0	77.0	
		LM3L ^[90]	LBP, SIFT, LE, TPLBP	阈值	82.4	78.2	78.8	80.4	80.0	
深度学习		ESL ^[48]	LBP	双线性相似度	80.5	72.2	72.8	71.6	74.3	
		DMML ^[78]	LBP, SPLE, SIFT	SVM	78.5	76.5	78.5	79.5	78.3	
		WGEML ^[51]	LBP, HOG, SIFT, CNN	KNN	88.6	77.4	83.4	81.6	82.8	
		CNN-points ^[57]	深度特征	神经网络	81.9	89.4	92.4	89.9	88.4	
		SMCNN ^[17]	深度特征	阈值	79.0	75.0	85.0	78.0	79.3	
		DDML ^[50]	深度特征	神经网络	87.4	83.8	83.2	83.0	84.3	
AdvKin ^[62]		深度特征	神经网络	91.6	85.2	90.2	92.4	89.9		
Attention ^[65]		深度特征	神经网络	91.8	89.8	92.8	93.4	92.0		
TSKinFace		特征提取方法	颜色特征 ^[12]	BSIF-HSV	阈值	81.5	81.4	79.9	82.0	81.2
	度量学习方法	WGEML ^[51]	LBP, HOG, SIFT, CNN	KNN	90.3	89.8	91.4	90.4	90.5	
	深度学习	DDML ^[50]	深度特征	神经网络	88.5	87.0	87.9	87.8	87.8	
UBKinFace	度量学习方法	DMML ^[78]	LBP, SPLE, SIFT	SVM	年轻父母-孩子 74.5 70.0		年老父母-孩子		72.3	
FIW		AdvKin ^[62]	深度特征	神经网络	68.8	67.8	67.3	69.9	68.5	
Smile	深度学习	SCCAE ^[18]	自编码器隐含层	神经网络	93.4	93.8	92.2	93.6	93.3	
KIVI		SMNAE ^[72]	自编码器隐含层	SVM	80.0	81.8	77.8	92.3	83.0	
TALKIN		模态融合 ^[73]	深度特征	神经网络	80.0	70.5	73.5	72.5	74.1	

束、拍摄条件发生变化时具有多样化, 干扰类内距离的度量, 问题复杂度增加, 传统方法已无法满足解决问题的需要. AdvKin^[62] 深度学习方法也仅取得了 68.5% 的平均识别准确率. 基于大规模数据集的人脸亲子关系验证研究更贴合于实际应用, 同时

也伴随着更大的挑战性, 值得今后投入更多研究.

5) 多模态信号. 现有的亲子数据集大多数仅包含人脸视觉信息, 人脸亲子关系验证算法基于视觉单模态实现. 文献 [73] 通过将语音信号和视觉信号结合, 在视觉模态上通过语音模态提供互补信息,

性能得到了进一步提升. 通过文献 [73] 的实验结果可以看到单一语音模态和视觉模态分别取得 65.8% 和 71.9% 的平均识别准确率, 融合两模态后可以达到 74.1% 的平均识别准确率, 因此所涉及的模态也是影响识别性能的因素之一.

整体而言, 人脸亲子关系验证的识别率总体仍比较低, 特别是在复杂情况下受多种因素干扰时性能鲁棒性较差, 距离实际应用尚存在很大差距, 仍需要进一步的研究.

6 对人脸亲子关系验证的思考

近年来人脸亲子关系验证领域出现大量新的研究工作, 在包括方法研究、数据集构建和应用拓展等方面均取得一定的研究进展, 这是对其应用价值的肯定. 未来会有更多的工作尝试进一步提高人脸亲子关系验证的准确率和易用性, 并逐渐将人脸亲子关系验证方法应用到现实问题中. 本文在回顾了人脸亲子关系验证问题的发展历程、分析其研究现状、归类和评述了现有主流方法的基础上, 下面总结现有人脸亲子关系验证研究的一些问题和未来可能的研究方向.

1) 大规模人脸亲子关系数据集构建问题. 深度学习在亲子关系验证中得到了成功应用, 为了训练性能更佳、使用性更为广泛的深度模型, 需要建立一个大的亲子数据库用于算法的训练. 目前最大的亲子数据库 FIW 的数据量为 1000 个家庭, 距离百万级的数据库还有很大的距离. 并且目前的亲子数据库所涉及的场景也不够丰富, 例如通常亲子图像来源于同一幅相片^[115]. 现有的亲子数据库所涉及的参与者信息不够全面, 大部分仅涉及到了参与者的视觉静态的面部长相, 而忽略了参与者时间-空间的面部视频、声音以及步态等信息. 除此之外, 亲子数据库的搜集是建立在已知的亲子关系基础上, 多个亲属参与的数据搜集工作, 需要家庭中两人以上作为单位进行数据搜集. 然而, 亲子关系属于相对隐私的家庭关系, 随着人们对个人隐私保护意识的不断提高, 无疑对亲子数据库的构建形成了一个阻碍. 构建大规模且信息源丰富的亲子数据集是推动亲子关系验证研究发展的一项重要工作.

2) 有效亲子特征表示的研究. 现有亲子关系验证方法中涉及到的面部特征提取大多基于整个面部的纹理特征、几何特征以及深度特征, 上述特征大多为人脸的一般特征表示, 包含了与亲子遗传信息无关的特征. 而亲子特征的有效表示直接影响了算法的识别准确率, 因此有效的亲子特征表示方法研究是今后的一个重要研究方向. 对亲子特征提取的研究包括, 面部亲子区域定位、亲子特征表示方法

等. 在 Dal Martello 等^[1, 5, 27] 的心理学研究中就表明了不同面部区域对亲子验证的贡献程度不同, 从实验结果看出嘴巴部分由于形态的多样性, 易对亲子关系验证产生负面的噪声影响, 而上半部脸包含了丰富的亲子特征线索, 今后一个值得关注的可能解决办法包括, 注意力机制网络在亲子关系验证问题的深入探究. 在区域定位基础上, 如何应用深度卷积网络从大数据中学习亲子特征表示也是今后值得研究的关键之一.

3) 如何降低年龄和性别因素对人脸亲子关系验证的影响. 人脸亲子关系验证本质上是对比两个输入间的相似程度, DeBruine 等^[8] 通过研究表明性别和年龄会大大影响亲子验证系统的准确性, 而如何降低这两个因素带来的影响是今后研究值得关注的一点. 其中一个思路是通过 GAN 网络在原输入图像的基础上修改图像的某种特性, 如性别、年龄, 并生成新的人脸图像. 另一个思路是, 将年龄、性别标签作为约束条件, 通过将年龄、性别特征与亲子特征解耦, 从而学习年龄和性别不变性的人脸特征表示.

4) 非限定环境下人脸亲子关系验证的鲁棒性问题. 本文在第 1 节中提到人脸亲子关系验证所面临的挑战包括人脸姿态变化、成像条件、图像分辨率、面部表情变化等, 目前为了解决自然条件下的人脸姿态差异性, 在实验过程中通过眼睛位置对人脸进行归一化对齐操作. 然而目前很少对光照、表情等因素进行考虑, 依旧是尚待解决的问题, 如我们考虑引入表情估计模型, 利用表情估计模型估计出人脸的表情, 再进行局部特征匹配, 从而解决表情不统一的问题^[118]. 再者对于光照不均匀和光线弱的情况, 是否可以引入近红外成像光电器件解决光线弱的问题, 引入本征图分解方法提取与周围环境光无关的本征反射图, 从而消除由光照不均匀和阴影带来的影响^[119].

5) 视听亲子关系验证 (Audio-visual kinship verification). 现有研究工作大多从静止人脸图像判断二者是否具有亲子关系, 而基因学^[120]、心理学^[121] 和声学^[122-125] 等研究均表明声音也具有遗传性, 子女具有与父母相似的说话声音. Sataloff^[120] 从基因学的角度研究人类发声规律, 他指出人说话音色取决于发声器官, 而发生器官往往由基因决定, 包括喉部软骨大小、声带的长度以及声门形状. 声学相关学者也通过语音特征: 基频 F_0 ^[122-124]、基频扰动 (Jitter)、振幅扰动 (shimmer) 以及谐噪比 (Harmonics-to-noise ratio, HNR)^[125] 证明了亲属间语音是具有相似性的. 此外, 心理学研究表明人们具有依据说话声音判断是否具有亲属关系的能力^[121], 而计算机是否也可以通过语音来进行亲子关系验证是

未来值得探究的一个研究方向. 声音模态可以为人脸视觉模态提供补充信息, 提高系统的识别准确率. 目前在亲子关系验证研究中, 视觉语音双模态融合研究还非常少, 处于起步阶段, 而如何利用模态间的互补性、提高系统鲁棒性是未来的研究热点之一.

6) 亲子图像生成相关研究. 生成对抗网络 (GAN)^[126] 是现阶段深度学习的一个重要发展分支, 在多个研究领域展现了其卓越的性能. 在亲子关系验证研究上, 其中一个难点就是亲子人脸数据的难获得性和亲子数据量不足、缺乏数据多样性, GAN 模型就可以通过亲子图像生成、数据扩充方法帮助解决这个问题. 此外, 就丢失儿童问题而言, 其中的一个重要解决思想就是通过父母的人脸图像生成其子女图像, 从而更好的匹配和寻找丢失的儿童, 基于 GAN 的子女人脸图像的生成研究也可以应用于手机终端娱乐. 基于 GAN 的亲子图像间的转换和分析研究是一个今后值得关注的研究方向.

7) 基于模型迁移学习的亲子关系验证研究. 我们在前面提到现有的亲子数据集规模大都较小, 训练样本不足, 除了对数据集进行数据扩充, 另一个思路就是用知识迁移 (Knowledge transfer)^[127] 方法, 是否可以借鉴人脸相关训练模型所学习到的知识, 如在大规模人脸识别数据库上预训练的深度学习模型, 从而更有效的解决亲子关系验证问题. 除此之外, 现有的人脸亲子关系验证研究将每种亲子关系看作独立的任务, 以“父亲-儿子 (FS)”关系为例, 其余人脸亲子数据, 如: 母亲-儿子, 是否有助于训练单任务模型, 比如将模型通过其他类型关系训练网络的初始化参数, 再就特定的亲子类型学习或微调网络. 如何充分应用亲属数据也将是解决数据量不足的一个思路.

8) 小样本亲子关系验证. 深度学习网路往往需要大量的标注数据训练模型, 然而就人脸亲子关系验证问题而言, 场景不同、问题背景不同时训练样本不足, 这种情境下就需要解决样本少时的亲子关系验证问题. 我们是否可以引入小样本学习 (Few-shot learning, FSL)^[128], FSL 是深度学习的一个重要分支, 在没有大量标注数据情况下, 能够快速泛化学习内容, 训练模型. 小样本亲子关系验证的一个解决思路是, 通过元学习 (Meta-learning) 使得模型学会学习. 具体为将有限的人脸亲子训练数据构造一系列训练任务, 让系统学会学习, 从而达到快速学习目的, 同时避免过拟合 (Overfitting) 问题. 小样本亲子关系验证问题研究目前处于空白, 具有很大的研究价值和意义.

9) 家庭检索和家庭分类. 家庭检索和家庭分类任务的训练数据集为所有家庭成员, 测试输入为丢

失儿童图像, 输出其所属的家庭. 家庭检索和家庭分类对于丢失儿童匹配目标家庭有重要的现实意义. 但是, 家庭分析研究由于其涉及范围广、隔辈差异性大等原因, 问题复杂度高、难度大, 因此目前的相关研究很少, 仍处于基准方法水平, 算法的性能具有非常大的提升空间. 比如, 应用数学模型, 对家庭构建图结构或树结构, 更好的优化家庭聚类. 家庭检索和家庭分类是今后极具潜力的一个研究方向.

7 结束语

人脸亲子关系验证是计算机视觉领域一个新兴的研究方向, 具有重要的理论意义和实际应用价值, 同时也存在诸多难点与挑战. 本文对人脸亲子关系验证问题的起源与发展做了详细的回顾, 对该领域主流数据集进行了总结, 对已有人脸亲子关系验证方法进行了详细地梳理、归类和评述, 对主流方法的优缺点进行了分析、性能进行了对比; 在此基础上, 对未来人脸亲子关系验证的可能研究方向进行了展望. 由于篇幅所限, 许多问题未及深究. 本文旨在使国内同行对该领域有一个较为全面的了解, 以期引起更多研究者对该领域的关注.

References

- 1 Dal Martello M F, Maloney L T. Where are kin recognition signals in the human face? *Journal of Vision*, 2006, **6**(12): 1356–1366
- 2 Dal Martello M F, Maloney L T. Lateralization of kin recognition signals in the human face. *Journal of Vision*, 2010, **10**(8): Article No. 9
- 3 DeBruine L M, Smith F G, Jones B C, Roberts S C, Petrie M, Spector T D. Kin recognition signals in adult faces. *Vision Research*, 2009, **49**(1): 38–43
- 4 Maloney L T, Dal Martello M F. Kin recognition and the perceived facial similarity of children. *Journal of Vision*, 2006, **6**(10): 1047–1056
- 5 Alvergne A, Perreau F, Mazur A, Mueller U, Raymond M. Identification of visual paternity cues in humans. *Biology Letters*, 2014, **10**(4): Article No. 20140063
- 6 Fang R G, Tang K D, Snavely N, Chen T. Towards computational models of kinship verification. In: Proceedings of the 2010 IEEE International Conference on Image Processing. Hong Kong, China: IEEE, 2010. 1577–1580
- 7 Lu J W, Hu J L, Zhou X Z, Zhou J, Castrillón-Santana M, Lorenzo-Navarro J, et al. Kinship verification in the wild: The first kinship verification competition. In: Proceedings of the 2014 IEEE International Joint Conference on Biometrics. Clearwater, USA: IEEE, 2014. 1–6
- 8 Lu J W, Hu J L, Liang V E, Zhou X Z, Bottino A, Islam I U, et al. The FG 2015 kinship verification in the wild evaluation. In: Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). Ljubljana, Slovenia: IEEE, 2015. 1–7
- 9 Robinson J P, Yin Y, Khan Z, Shan M, Xia S Y, Stopa M, et al. Recognizing families in the wild (RFIW): The 4th edition. In: Proceedings of the 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020). Buenos Aires, Argentina: IEEE, 2020.

- 10 Robinson J P, Shao M, Fu Y. To recognize families in the wild: A machine vision tutorial. In: Proceedings of the 26th ACM International Conference on Multimedia. Seoul, Republic of Korea: ACM, 2018. 2096–2097
- 11 Robinson J P, Shao M, Wu Y, Liu H F, Gillis T, Fu Y. Visual kinship recognition of families in the wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **40**(11): 2624–2637
- 12 Wu X T, Boutellaa E, López M B, Feng X Y, Hadid A. On the usefulness of color for kinship verification from face images. In: Proceedings of the 2016 IEEE International Workshop on Information Forensics and Security (WIFS). Abu Dhabi, United Arab Emirates: IEEE, 2016. 1–6
- 13 Liu Q F, Puthenputhussery A, Liu C J. A novel inheritable color space with application to kinship verification. In: Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Placid, USA: IEEE, 2016. 1–9
- 14 Puthenputhussery A, Liu Q F, Liu C J. SIFT flow based genetic fisher vector feature for kinship verification. In: Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP). Phoenix, USA: IEEE, 2016. 2921–2925
- 15 Lu J W, Zhou X Z, Tan Y P, Shang Y Y, Zhou J. Neighborhood repulsed metric learning for kinship verification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(2): 331–345
- 16 Wang M Y, Li Z C, Shu X B, Jingdong, Tang J H. Deep kinship verification. In: Proceedings of the 17th International Workshop on Multimedia Signal Processing (MMSP). Xiamen, China: IEEE, 2015. 1–6
- 17 Li L, Feng X Y, Wu X T, Xia Z Q, Hadid A. Kinship verification from faces via similarity metric based convolutional neural network. In: Proceedings of the 13th International Conference on Image Analysis and Recognition. Póvoa de Varzim, Portugal: Springer, 2016. 539–548
- 18 Dibeklioglu H. Visual transformation aided contrastive learning for video-based kinship verification. In: Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 2478–2487
- 19 Gao P Y, Xia S Y, Robinson J, Zhang J K, Xia C, Shao M, et al. What will your child look like? DNA-Net: Age and gender aware kin face synthesizer. arXiv preprint arXiv: 1911.07014, 2019.
- 20 Ghatas F S, Hemayed E E. GANKIN: Generating kin faces using disentangled GAN. *SN Applied Sciences*, 2020, **2**(2): Article No. 166
- 21 Dandekar A R, Nimbarte M S. A survey: Verification of family relationship from parents and child facial images. In: Proceedings of the 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science. Bhopal, India: IEEE, 2014. 1–6
- 22 Wu X T, Boutellaa E, Feng X Y, Hadid A. Kinship verification from faces: Methods, databases and challenges. In: Proceedings of the 2016 IEEE International Conference on Signal Processing (ICSPCC). Hong Kong, China: IEEE, 2016. 1–6
- 23 Almuashi M, Hashim S Z M, Mohamad D, Alkawaz M H, Ali A. Automated kinship verification and identification through human facial images: A survey. *Multimedia Tools and Applications*, 2017, **76**(1): 265–307
- 24 Georgopoulos M, Panagakis Y, Pantic M. Modeling of facial aging and kinship: A survey. *Image and Vision Computing*, 2018, **80**: 58–79
- 25 Qin X Q, Liu D K, Wang D. A literature survey on kinship verification through facial images. *Neurocomputing*, 2020, **377**: 213–224
- 26 Robinson J P, Shao M, Fu Y. Survey on the analysis and modeling of visual kinship: A decade in the making. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, **44**(8): 4432v4453
- 27 Dal Martello M F, DeBruine L M, Maloney L T. Allocentric kin recognition is not affected by facial inversion. *Journal of Vision*, 2015, **15**(13): Article No. 5
- 28 Fasolt V, Holzleitner I, Lee A J, O'Shea K J, DeBruine L M. Facial expressions influence kin recognition accuracy. *Human Ethology Bulletin*, 2018, **33**(4): 19–27
- 29 Wu X T, Feng X Y, Boutellaa E, Hadid A. Kinship verification using color features and extreme learning machine. In: Proceedings of the 3rd IEEE International Conference on Signal and Image Processing (ICSIP). Shenzhen, China: IEEE, 2018. 187–191
- 30 Fang R G, Gallagher A C, Chen T, Loui A. Kinship classification by modeling facial feature heredity. In: Proceedings of the 2013 IEEE International Conference on Image Processing. Melbourne, Australia: IEEE, 2013. 2983–2987
- 31 Guo G D, Wang X L. Kinship measurement on salient facial features. *IEEE Transactions on Instrumentation and Measurement*, 2012, **61**(8): 2322–2325
- 32 Wang X L, Kambhamettu C. Leveraging appearance and geometry for kinship verification. In: Proceedings of the 2014 International Conference on Image Processing (ICIP). Paris, France: IEEE, 2014. 5017–5021
- 33 Goyal A, Meenpal T. Detection of facial parts in kinship verification based on edge information. In: Proceedings of the 2018 Conference on Information and Communication Technology (CICT). Jabalpur, India: IEEE, 2018. 1–6
- 34 Asthana A, Zafeiriou S, Cheng S Y, Pantic M. Robust discriminative response map fitting with constrained local models. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 3444–3451
- 35 Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, **PAMI-8**(6): 679–698
- 36 Zhou X Z, Hu J L, Lu J W, Shang Y Y, Guan Y. Kinship verification from facial images under uncontrolled conditions. In: Proceedings of the 19th ACM International Conference on Multimedia. Scottsdale, USA: ACM, 2011. 953–956
- 37 Zhou X Z, Lu J W, Hu J L, Shang Y Y. Gabor-based gradient orientation pyramid for kinship verification under uncontrolled environments. In: Proceedings of the 20th ACM International Conference on Multimedia. Nara, Japan: ACM, 2012. 725–728
- 38 Moujahid A, Dornaika F. A pyramid multi-level face descriptor: Application to kinship verification. *Multimedia Tools and Applications*, 2019, **78**(7): 9335–9354
- 39 Yan H B. Learning discriminative compact binary face descriptor for kinship verification. *Pattern Recognition Letters*, 2019, **117**: 146–152
- 40 Laiadi O, Ouamane A, Boutellaa E, Benakcha A, Taleb-Ahmed A, Hadid A. Kinship verification from face images in discriminative subspaces of color components. *Multimedia Tools and Applications*, 2019, **78**(12): 16465–16487
- 41 Liu Q F, Puthenputhussery A, Liu C J. Inheritable fisher vector feature for kinship verification. In: Proceedings of the 7th International Conference on Biometrics Theory, Applications and Systems (BTAS). Arlington, USA: IEEE, 2015. 1–6
- 42 Alirezazadeh P, Fathi A, Abdali-Mohammadi F. A genetic algorithm-based feature selection for kinship verification. *IEEE Signal Processing Letters*, 2015, **22**(12): 2459–2463
- 43 Bottinok A, Islam I U, Vieira T F. A multi-perspective holistic approach to kinship verification in the wild. In: Proceedings of the 11th International Conference and Workshops on Automatic Face and Gesture Recognition (FG). Ljubljana, Slovenia: IEEE, 2015. 1–6
- 44 Cui L Y, Ma B. Adaptive feature selection for kinship verification

- tion. In: Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME). Hong Kong, China: IEEE, 2017. 751–756
- 45 Chen X J, An L, Yang S F, Wu W M. Kinship verification in multi-linear coherent spaces. *Multimedia Tools and Applications*, 2017, **76**(3): 4105–4122
- 46 Yan H B, Lu J W, Zhou X Z. Prototype-based discriminative feature learning for kinship verification. *IEEE Transactions on Cybernetics*, 2015, **45**(11): 2535–2545
- 47 Xu M, Shang Y Y. Kinship measurement on face images by structured similarity fusion. *IEEE Access*, 2016, **4**: 10280–10287
- 48 Zhou X Z, Shang Y Y, Yan H B, Guo G D. Ensemble similarity learning for kinship verification from facial images in the wild. *Information Fusion*, 2016, **32**: 40–48
- 49 Zhou X Z, Yan H B, Shang Y Y. Kinship verification from facial images by scalable similarity fusion. *Neurocomputing*, 2016, **197**: 136–142
- 50 Guo Y H, Dibeklioglu H, Van Der Maaten L. Graph-based kinship recognition. In: Proceedings of the 22nd International Conference on Pattern Recognition. Stockholm, Sweden: IEEE, 2014. 4287–4292
- 51 Liang J Q, Hu Q H, Dang C Y, Zuo W M. Weighted graph embedding-based metric learning for kinship verification. *IEEE Transactions on Image Processing*, 2019, **28**(3): 1149–1162
- 52 Shao M, Xia S Y, Fu Y. Genealogical face recognition based on UB KinFace database. In: Proceedings of the CVPR 2011 WORKSHOPS. Colorado Springs, USA: IEEE, 2011. 60–65
- 53 Xia S Y, Shao M, Fu Y. Kinship verification through transfer learning. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence. Barcelona, Spain: AAAI Press, 2011. 2539–2544
- 54 Zhang Y L, Ma B, Huang L H, Hu H W. Transfer metric learning for kinship verification with locality-constrained sparse features. In: Proceedings of the 22nd International Conference on Neural Information Processing. Istanbul, Turkey: Springer, 2015. 234–243
- 55 Yan H B, Zhou X Z, Ge Y X. Neighborhood repulsed correlation metric learning for kinship verification. In: Proceedings of the 2015 Visual Communications and Image Processing (VCIP). Singapore: IEEE, 2015. 1–4
- 56 Yan H B. Kinship verification using neighborhood repulsed correlation metric learning. *Image and Vision Computing*, 2017, **60**: 91–97
- 57 Zhang K H, Huang Y Z, Song C F, Wu H, Wang L. Kinship verification with deep convolutional neural networks. In: Proceedings of the 2015 British Machine Vision Conference. Swansea, UK: BMVA Press, 2015.
- 58 Duan Q Y, Zhang L. AdvNet: Adversarial contrastive residual net for 1 million kinship recognition. In: Proceedings of the 2017 Workshop on Recognizing Families in the Wild. Mountain View, USA: ACM, 2017.
- 59 Lu J W, Hu J L, Tan Y P. Discriminative deep metric learning for face and kinship verification. *IEEE Transactions on Image Processing*, 2017, **26**(9): 4269–4282
- 60 Yong L, Zeng J B, Zhang J, Dai A B, Kan M N, Shan S G, et al. KinNet: Fine-to-coarse deep metric learning for kinship verification. In: Proceedings of the 2017 Workshop on Recognizing Families in the Wild. Mountain View, USA: ACM, 2017.
- 61 Zhou X Z, Jin K, Xu M, Guo G D. Learning deep compact similarity metric for kinship verification from face images. *Information Fusion*, 2019, **48**: 84–94
- 62 Zhang L, Duan Q Y, Zhang D, Jia W, Wang X Z. AdvKin: Adversarial convolutional network for kinship verification. *IEEE Transactions on Cybernetics*, 2021, **51**(12): 5883–5896
- 63 Wang S Y, Ding Z M, Fu Y. Cross-generation kinship verification with sparse discriminative metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, **41**(11): 2783–2790
- 64 Ozkan S, Ozkan A. Kinshipgan: Synthesizing of kinship faces from family photos by regularizing a deep face network. In: Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece: IEEE, 2018. 2142–2146
- 65 Yan H B, Wang S W. Learning part-aware attention networks for kinship verification. *Pattern Recognition Letters*, 2019, **128**: 169–175
- 66 Dibeklioglu H, Salah A A, Gevers T. Are you really smiling at me? Spontaneous versus posed enjoyment smiles. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: Springer, 2012. 525–538
- 67 Dibeklioglu H, Salah A A, Gevers T. Like father, like son: Facial expression dynamics for kinship verification. In: Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013. 1497–1504
- 68 Qin X Q, Tan X Y, Chen S C. Tri-subject kinship verification: Understanding the core of a family. *IEEE Transactions on Multimedia*, 2015, **17**(10): 1855–1867
- 69 Kohli N, Vatsa M, Singh R, Noore A, Majumdar A. Hierarchical representation learning for kinship verification. *IEEE Transactions on Image Processing*, 2017, **26**(1): 289–302
- 70 Yan H B, Hu J L. Video-based kinship verification using distance metric learning. *Pattern Recognition*, 2018, **75**: 15–24
- 71 Sun Y, Li J C, Wei Y W, Yan H B. Video-based parent-child relationship prediction. In: Proceedings of the 2018 IEEE Visual Communications and Image Processing (VCIP). Taichung, China: IEEE, 2018. 1–4
- 72 Kohli N, Yadav D, Vatsa M, Singh R, Noore A. Supervised mixed norm autoencoder for kinship verification in unconstrained videos. *IEEE Transactions on Image Processing*, 2019, **28**(3): 1329–1341
- 73 Wu X T, Granger E, Kinnunen T H, Feng X Y, Hadid A. Audio-visual kinship verification in the wild. In: Proceedings of the 2019 International Conference on Biometrics (ICB). Crete, Greece: IEEE, 2019. 1–8
- 74 Wu Xiao-Ting. Kinship Verification Based on Color Texture Features and Metric Learning [Master dissertation], Northwestern Polytechnical University, China, 2016. (吴晓婷. 基于颜色纹理特征与度量学习的亲子关系验证研究 [硕士学位论文], 西北工业大学, 中国, 2016.)
- 75 Zhang K P, Zhang Z P, Li Z F, Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 2016, **23**(10): 1499–1503
- 76 Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014. 1867–1874
- 77 Xing E P, Ng A Y, Jordan M I, Russell S J. Distance metric learning, with application to clustering with side-information. In: Proceedings of the 15th International Conference on Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2002. 521–528
- 78 Yan H B, Lu J W, Deng W H, Zhou X Z. Discriminative multimetric learning for kinship verification. *IEEE Transactions on Information Forensics and Security*, 2014, **9**(7): 1169–1178
- 79 Xia S Y, Shao M, Fu Y. Toward kinship verification using visual attributes. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). Tsukuba, Japan: IEEE, 2012. 549–552
- 80 Xia S Y, Shao M, Luo J B, Fu Y. Understanding kin relationships in a photo. *IEEE Transactions on Multimedia*, 2012, **14**(4): 1046–1056
- 81 Kohli N, Singh R, Vatsa M. Self-similarity representation of Weber faces for kinship classification. In: Proceedings of the IEEE 5th International Conference on Biometrics: Theory, Ap-

- lications and Systems (BTAS). Arlington, USA: IEEE, 2012. 245–250
- 82 Tola E, Lepetit V, Fua P. DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(5): 815–830
- 83 Patel B, Maheshwari R P, Raman B. Evaluation of periocular features for kinship verification in the wild. *Computer Vision and Image Understanding*, 2017, **160**: 24–35
- 84 Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 1996, **29**(1): 51–59
- 85 Ojala T, Pietikäinen M, Maenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(7): 971–987
- 86 Duan X D, Tan Z H. A feature subtraction method for image based kinship verification under uncontrolled environments. In: Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP). Québec City, Canada: IEEE, 2015. 1573–1577
- 87 Bessaoudi M, Ouamane A, Belahcene M, Chouchane A, Boutellaa E, Bourennane S. Multilinear side-information based discriminant analysis for face and kinship verification in the wild. *Neurocomputing*, 2019, **329**: 267–278
- 88 Laiadi O, Ouamane A, Benakcha A, Taleb-Ahmed A, Hadid A. Tensor cross-view quadratic discriminant analysis for kinship verification in the wild. *Neurocomputing*, 2020, **377**: 286–300
- 89 Hu J L, Lu J W, Yuan J S, Zhou X P. Large margin multi-metric learning for face and kinship verification in the wild. In: Proceedings of the 12th Asian Conference on Computer Vision. Singapore: Springer, 2014. 252–267
- 90 Hu J L, Lu J W, Tan Y P, Yuan J S, Zhou J. Local large-margin multi-metric learning for face and kinship verification. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, **28**(8): 1875–1891
- 91 Kou L, Zhou X Z, Xu M, Shang Y Y. Learning a genetic measure for kinship verification using facial images. *Mathematical Problems in Engineering*, 2015, **2015**: Article No. 472473
- 92 Wei Z Q, Xu M, Geng L, Liu H M, Yin H. Adversarial similarity metric learning for kinship verification. *IEEE Access*, 2019, **7**: 100029–100035
- 93 Xu M, Shang Y Y. Kinship verification using facial images by robust similarity learning. *Mathematical Problems in Engineering*, 2016, **2016**: Article No. 4072323
- 94 Qin X Q, Tan X Y, Chen S C. Mixed bi-subject kinship verification via multi-view multi-task learning. *Neurocomputing*, 2016, **214**: 350–357
- 95 Fang Y, Yan Y, Chen S, Wang H Z, Shu C. Sparse similarity metric learning for kinship verification. In: Proceedings of the 2016 Visual Communications and Image Processing (VCIP). Chengdu, China: IEEE, 2016. 1–4
- 96 Lei X H, Li B, Xie J. Locality discriminative canonical correlation analysis for kinship verification. In: Proceedings of the 12th IEEE Conference on Industrial Electronics and Applications (ICIEA). Siem Reap, Cambodia: IEEE, 2017.
- 97 Zhang J K, Xia S Y, Pan H, Qin A K. A genetics-motivated unsupervised model for tri-subject kinship verification. In: Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP). Phoenix, USA: IEEE, 2016. 2916–2920
- 98 Liu H J, Cheng J, Wang F. Kinship verification based on status-aware projection learning. In: Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China: IEEE, 2017. 1072–1076
- 99 Liu H J, Zhu C. Status-aware projection metric learning for kinship verification. In: Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME). Hong Kong, China: IEEE, 2017. 319–324
- 100 Wu Y, Ding Z M, Liu H F, Robinson J, Yun F. Kinship classification through latent adaptive subspace. In: Proceedings of the 13th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2018). Xi'an, China: IEEE, 2018. 143–149
- 101 Zhao Y G, Song Z, Zheng F, Shao L. Learning a multiple kernel similarity metric for kinship verification. *Information Sciences*, 2018, **430–431**: 247–260
- 102 Deng J, Berg A C, Fei-Fei L. Hierarchical semantic indexing for large scale image retrieval. In: Proceedings of the 2011 Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE, 2011. 785–792
- 103 Gao X Y, Hoi S C H, Zhang Y D, Wan J, Li J T. SOML: Sparse online metric learning with application to image retrieval. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence. Québec City, Canada: ACM, 2014. 1206–1212
- 104 Wang S Y, Ding Z M, Fu Y. Coupled marginalized auto-encoders for cross-domain multi-view learning. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence. New York, USA: IJCAI/AAAI Press, 2016. 2125–2131
- 105 Duan Q Y, Zhang L, Zuo W M. From face recognition to kinship verification: An adaptation approach. In: Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCVW). Venice, Italy: IEEE, 2017. 1590–1598
- 106 Liang J Y, Guo J L, Lao S Y, Li J. Using deep relational features to verify kinship. In: Proceedings of the 2nd CCF Chinese Conference on Computer Vision. Tianjin, China: Springer, 2017. 563–573
- 107 Zhang H M, Wang X L, Kuo C C J. Deep kinship verification VIA appearance-shape joint prediction and adaptation-based approach. In: Proceedings of the 2019 IEEE International Conference on Image Processing. Taipei, China: IEEE, 2019. 3856–3860
- 108 Wu T, Turaga P, Chellappa R. Age estimation and face verification across aging using landmarks. *IEEE Transactions on Information Forensics and Security*, 2012, **7**(6): 1780–1788
- 109 Li W H, Zhang Y Q, Lv K C, Lu J W, Feng J J, Zhou J. Graph-based kinship reasoning network. In: Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME). London, UK: IEEE, 2020. 1–6
- 110 Song C H, Yan H B. KINMIX: A data augmentation approach for kinship verification. In: Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME). London, UK: IEEE, 2020. 1–6
- 111 Suh Y, Han B, Kim W, Lee K M. Stochastic class-based hard example mining for deep metric learning. In: Proceedings of the 2019 IEEE/CVF Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 7251–7259
- 112 Ertugrul I Ö, Dibeklioglu H. What will your future child look like? Modeling and synthesis of hereditary patterns of facial dynamics. In: Proceedings of the 12th International Conference on Automatic Face and Gesture Recognition. Washington, USA: IEEE, 2017. 33–40
- 113 Dehghan A, Ortiz E G, Villegas R, Shah M. Who do I look like? Determining parent-offspring resemblance via gated autoencoders. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014. 1757–1764
- 114 Wang S Y, Robinson J P, Fu Y. Kinship verification on families in the wild with marginalized denoising metric learning. In: Proceedings of the 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG2017). Washington, USA: IEEE, 2017. 216–221
- 115 López M B, Boutellaa E, Hadid A. Comments on the “kinship face in the wild” data sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(11): 2342–2344

- 116 He Xing-Chen, Guo Yong, Li Qi-Long, Gao Chang. Age invariant face recognition based on deep learning. *Acta Automatica Sinica*, 2022, **48**(3): 877–886
(何星辰, 郭勇, 李奇龙, 高唱. 基于深度学习的抗年龄干扰人脸识别. 自动化学报, 2022, **48**(3): 877–886)
- 117 Zhang Ting, Li Yu-Jian, Hu Hai-He, Zhang Ya-Hong. Gender classification model based on cross-connected convolutional neural networks. *Acta Automatica Sinica*, 2016, **42**(6): 858–865
(张婷, 李玉鑑, 胡海鹤, 张亚红. 基于跨连卷积神经网络的性别分类模型. 自动化学报, 2016, **42**(6): 858–865)
- 118 Luo Hao, Jiang Wei, Fan Xing, Zhang Si-Peng. A survey on deep learning based person re-identification. *Acta Automatica Sinica*, 2019, **45**(11): 2032–2049
(罗浩, 姜伟, 范星, 张思朋. 基于深度学习的行人重识别研究进展. 自动化学报, 2019, **45**(11): 2032–2049)
- 119 Li Chen. Facial Intrinsic Image Decomposition and its Application [Ph.D. dissertation], Zhejiang University, China, 2017.
(李琛. 人脸本征图像分解及其应用 [博士学位论文], 浙江大学, 中国, 2017.)
- 120 Sataloff R T. Genetics of the voice. *Journal of Voice*, 1995, **9**(1): 16–19
- 121 Van Gysel W D, Vercammen J, Debruyne F. Voice similarity in identical twins. *Acta Oto-Rhino-Laryngologica Belgica*, 2001, **55**(1): 49–55
- 122 Whiteside S P, Rixon E. Speech tempo and fundamental frequency patterns: A case study of male monozygotic twins and an age-and sex-matched sibling. *Logopedics Phoniatrics Vocology*, 2013, **38**(4): 173–181
- 123 Debruyne F, Decoster W, Van Gijzel A, Vercammen J. Speaking fundamental frequency in monozygotic and dizygotic twins. *Journal of Voice*, 2002, **16**(4): 466–471
- 124 Nolan F, McDougall K, Hudson T. Some acoustic correlates of perceived (Dis) similarity between same-accent voices. In: Proceedings of the 17th International Congress of Phonetic Sciences. Hong Kong, China: ICPhS, 2011. 1506–1509
- 125 Weirich M, Lancia L. Perceived auditory similarity and its acoustic correlates in twins and unrelated speakers. In: Proceedings of the 17th International Congress of Phonetic Sciences. Hong Kong, China: ICPhS, 2011. 2118–2121
- 126 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: ACM, 2014. 2672–2680
- 127 Pan S J, Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, **22**(10): 1345–1359
- 128 Wang Y Q, Yao Q M, Kwok J T, Ni L M. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys*, 2021, **53**(3): Article No. 63



吴晓婷 西北工业大学电子信息学院与芬兰奥卢大学机器视觉与信号分析中心博士研究生. 主要研究方向为人脸亲子关系验证, 计算机视觉, 深度学习方向.

E-mail: wuxt14@mail.nwpu.edu.cn

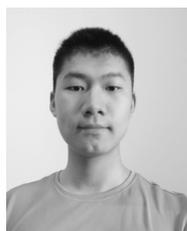
(WU Xiao-Ting Ph.D. candidate at the School of Electronics and Information, Northwestern Polytechnical University and Center for Machine Vision and Signal Analysis, University of Oulu, Oulu, Finland. Her research interest covers facial kinship verification, computer vision, and deep learning.)



冯晓毅 西北工业大学电子信息学院教授. 主要研究方向为计算机视觉, 图像处理, 雷达成像和识别.

E-mail: fengxiao@nwpu.edu.cn

(FENG Xiao-Yi Professor at the School of Electronics and Information, Northwestern Polytechnical University. Her research interest covers computer vision, image process, and radar imagery and recognition.)



黄安 国防科技大学系统工程学院硕士研究生. 主要研究方向为计算机视觉, 模式识别.

E-mail: huangan99@nudt.edu.cn

(HUANG An Master student at the College of Electronic Science, National University of Defense Technology. His research interest covers computer vision and pattern recognition.)



张雪毅 国防科技大学系统工程学院硕士研究生. 主要研究方向为计算机视觉, 模式识别.

E-mail: zhangxy1998@nudt.edu.cn

(ZHANG Xue-Yi Master student at the College of Electronic Science, National University of Defense Technology. His research interest covers computer vision and pattern recognition.)



董晶 中国科学院自动化研究所智能感知与计算研究中心研究员. 主要研究方向为模式识别与图像处理, 计算机视觉, 多媒体内容安全.

E-mail: jdong@nlpr.ia.ac.cn

(DONG Jing Research fellow at the Center for Research on Intelligent Perception and Computing, Institute of Automation, Chinese Academy of Sciences. Her research interest covers pattern recognition, image processing, computer vision, and digital image forensics.)



刘丽 国防科技大学系统工程学院教授. 主要研究方向为图像理解, 计算机视觉, 模式识别. 本文通信作者.

E-mail: liuli_nudt@nudt.edu.cn

(LIU Li Professor at the College of System Engineering, National University of Defense technology. Her research interest covers image understanding, computer vision, and pattern recognition. Corresponding author of this paper.)