

学号 1606010313

年级 2016

河海大学

本科毕业论文

年龄不变的对抗特征学习及亲属关系验证

专 业 计算机科学与技术

姓 名 ***

指导教师 ***

评 阅 人

2020 年 6 月

中国 南京

**BACHELOR'S DEGREE THESIS
OF HOHAI UNIVERSITY**

**Age-Invariant Adversarial Feature Learning
with Kinship Verification**

College : College of Computer and Information

Subject : Computer Science and Technology

Name : ***

Directed by : ***

NANJING CHINA

郑 重 声 明

本人呈交的毕业论文，是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料真实可靠。尽我所知，除文中已经注明引用的内容外，本设计（论文）的研究成果不包含他人享有著作权的内容。对本设计（论文）所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确的方式标明。本设计（论文）的知识产权归属于培养单位。

本人签名：_____

日期：_____

摘 要

亲属关系验证的目标是确定两人是否具有血缘关系。该技术可用于许多现实世界的场景中，例如寻找失踪人员和跨学科研究。大多数现有方法通常是从给定的人脸图片中提取特征，然后将完整的人脸特征直接应用于亲属关系验证。然而，这些方法都存在两个问题：1) 它们很容易受到年龄差距的影响，当双方年龄差距大时准确率明显下降；2) 它们默认假设了人脸特征都具有相同的权重。针对这两个问题，本文提出了一种全新的年龄不变对抗特征学习方法。该方法包含年龄不变的对抗特征学习模块 (AIAF) 和基于注意力机制的特征加权模块 (AFW)。AIAF 能够将完整的面部特征分解为两个不相关的部分：身份特征 (即年龄不变特征) 和年龄特征。具体来说，我们利用一种对抗机制，使这两部分之间的相关性尽可能小。然后，只将提取的身份特征输入至 AFW 中，AFW 可以根据每个特征在亲属关系验证任务中的重要性自动地分配不同的权重。在两个公共数据集上的实验结果表明，我们提出的方法能够有效提取年龄不变的身份特征，从而在识别准确率上与其他最先进的方法相比得到较大提升。

关键词：亲属关系验证；年龄不变特征；对抗学习；多任务学习；注意力机制；深度学习

ABSTRACT

Kinship verification aims to determine whether two given persons are blood relatives. This technique can be leveraged in many real-world scenarios, such as finding missing people and interdisciplinary research. Most existing methods typically extract facial features from given images and then apply the full features to verify kinship directly. However, there exist two problems in these methods: 1) They are easily affected by the age gap. The accuracy is significantly reduced when the age gap between the two facial images is large. 2) They assume that the facial features have the same weight by default. To solve the two problems, this paper proposes a brand-new age-invariant adversarial feature learning approach. It consists of Age-Invariant Adversarial Feature Learning module (AIAF) and Attention-based Feature Weighted module (AFW). AIAF is capable of factorizing full facial features into two uncorrelated parts: identity-related features (i.e., age-invariant features) and age-related features. Specifically, we harness a type of adversarial mechanism to make the correlation between these two parts as small as possible. Then, only the purified identity features are fed into AFW, which can automatically assign different weights according to the importance of each feature in the kinship verification task. Experimental results on two public datasets show that our proposed approach can effectively capture useful age-invariant identity features, which achieves significant improvements in accuracy compared with other state-of-the-art methods.

Key words: Kinship verification, Age-invariant features, Adversarial learning, Multi-task learning, Attention mechanism, Deep learning

目 录

摘要	I
ABSTRACT	II
目 录	III
第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.2.1 传统的亲属关系验证方法	2
1.2.2 基于深度学习的亲属关系验证方法	2
1.3 存在的问题	4
1.4 本文研究内容及贡献	4
1.5 本文组织结构	6
第 2 章 相关技术介绍	7
2.1 引言	7
2.2 深度学习概述	7
2.2.1 深度学习的发展历史	7
2.2.2 深度学习的优势	8
2.2.3 激活函数	8
2.2.4 损失函数	10
2.2.5 优化器	11
2.3 卷积神经网络	12
2.3.1 卷积神经网络的发展历史	12
2.3.2 卷积神经网络的优势	13
2.3.3 卷积神经网络的构建	13
2.4 对抗学习	14
2.5 典型相关分析	16
2.6 多任务学习	17
2.7 注意力机制	19

2.8 本章小结	20
第 3 章 年龄不变的对抗特征学习.....	21
3.1 引言	21
3.2 年龄不变的对抗特征学习模块	21
3.2.1 特征分割器	22
3.2.2 身份特征损失函数	22
3.2.3 年龄特征损失函数	23
3.2.4 对抗性典型相关正则项损失函数	23
3.2.5 多任务学习	25
3.3 基于注意力机制的特征加权模块	27
3.4 本章小结	28
第 4 章 实验结果及分析.....	29
4.1 引言	29
4.2 实验设置	29
4.2.1 网络架构	29
4.2.2 数据预处理	29
4.3 评估指标	30
4.4 在 KinFaceW-I 上的实验	30
4.5 在 FIW 上的实验.....	32
4.6 消融实验	36
4.7 本章小结	36
第 5 章 总结与展望.....	37
5.1 本文工作总结	37
5.2 未来工作展望	38
致 谢.....	39
参考文献.....	40
附 录.....	50
A: 个人简介	50

B: 本科期间撰写的学术论文	50
C: 本科期间参加的科研项目	51
D: 本科期间申请的专利	51
E: 本科期间所获竞赛奖项.....	52
F: 本科期间所获荣誉	53
G: 本科期间社会工作	55

第 1 章 绪论

1.1 研究背景及意义

在世界范围内，每年失踪人口的数量十分巨大。根据调查，导致人口失踪的原因主要包括拐骗妇女、儿童和老人走失等。近年来，我国拐卖人口犯罪依然猖獗，为此，中国政府在 2013 年发布了《中国反对拐卖人口行动计划（2013-2020 年）》^[66]，旨在进一步打击贩卖人口犯罪，维护社会繁荣稳定。学术界和工业界广泛兴起了利用人工智能技术帮助失踪儿童的父母或其他亲属寻找失踪儿童。他们只需向公安部门提供孩子失踪前的照片，人工智能系统将这些照片与数据库信息以及全国各地的安全摄像头拍摄的图像进行比对整理，即可在茫茫人海中定位到失踪儿童。

上述问题的解决依赖于亲属关系验证技术的发展。亲属关系验证的目的是确定一对输入图像是否具有血缘关系。在多媒体领域和计算机视觉领域中，亲属关系验证是一个挑战性和趣味性并存的任务，在许多真实世界的场景中具有广泛的用途，比如寻找失踪人口、缓解人口贩卖问题，甚至跨学科研究，如历史人物的关系分析等。

从现代遗传学角度来看，亲属之间的面部特征有很多相似之处，面部变化过程除去生活环境、营养摄入等外界因素，主要取决于其父母的基因。基于这个观点，很多研究人员希望研究出一种系统^[74-76]，仅通过人脸面部图像就能判断其亲属关系。随着大数据时代的到来，越来越多的亲属照片涌现在互联网上，为亲属关系的研究提供了充足的数据保障。近年来，深度学习技术展现出了强大的挖掘数据潜在特征的能力和海量数据的能力，许多研究人脸借助深度学习算法进一步研究亲属关系验证问题。

此外，亲属关系验证属于人脸识别的一个子分支，同时也是人脸识别技术的一个典型的应用场景。比如，亲属关系验证问题通常更加注重于某些关键人脸特征点，需要研究人脸对这些特征进行充分挖掘，这些潜在特征可能在人脸识别中也非常重要。因此，亲属关系验证的研究不但有助于本研究自身的发展，也可促进人脸识别领域的发展。

综上，亲属关系验证问题能够帮助社会解决诸多重要问题，有利于跨学科研

究的发展,在大数据和人工智能技术的双重保证下,该研究具有重要的价值和意义。此外,亲属关系验证还可用于许多其他场景,如血缘关系鉴定、历史人物血缘关系研究等。目前,一些基于传统的机器学习方法^[1-7]或新兴的深度学习^[8-15]已经被提出,并用于亲属关系验证问题,取得了不错的效果。

1.2 国内外研究现状

1.2.1 传统的亲属关系验证方法

近十年来,人们提出了一些基于视觉图像的亲属关系确认方法。现有的方法大致可以分为两类:基于特征的方法^[4, 18-25]和基于相似性学习的方法^[2, 3, 7, 26-30]。

基于特征的方法通常会学习一个能从图像中提取鉴别性特征的模型,然后通过这些特征来判断是否具有亲属关系。Fang 等人^[18]在 2010 年首次提出利用计算机视觉技术解决亲属关系验证问题,他们从网络上收集了 150 对名人及其父母或子女的图片作为数据集,提取样本的局部特征,同时利用肤色、灰度值、梯度直方图来获取全局特征,并人工筛选最具有鉴别性的 22 个特征来进行亲属关系的验证。Zhou 等人^[19]提出了一种基于空间金字塔学习的人脸特征描述符,并将支持向量机应用于亲属关系验证。Guo 等人^[20]利用 DAISY 描述符^[54]提取显著特征用于亲属关系识别。Bottinok 等人^[21]利用 LPQ、WLD、TPLBP 和 FPLBP 纹理特征来解决亲属关系验证问题。Bottino 等人^[22]考虑到几何、纹理和整体特征,在 KinFaceW-II 数据集上取得了出色的性能。

此外,另一种基于相似学习的表明,通过度量学习或子空间学习来学习语义特征空间是解决亲属关系验证问题的另一种途径。Xia 等人^[31]认为在以前的模型中,儿童的脸与父母年轻时的脸更为相似,因此提出了一种基于转移子空间学习的算法来减少年龄带来的差异。此外,Lu 等人^[2]提出了一种邻域排斥度量学习方法,该方法将类内图像的距离拉近,将类间图像的距离推远。Zhao 等人^[32]提出了一种多核相似度度量方法,该方法通过加权几种不同局部特征来度量人脸相似度。为了解决不明确的关系对,Wei 等人^[33]使用一种对抗性相似度度量学习方法,迭代地学习到一种鲁棒的并且具有鉴别性的相似度度量。

1.2.2 基于深度学习的亲属关系验证方法

自 2012 年以来,越来越多的研究人员利用深度学习的方法来解决亲属关系验证问题。因为人工神经网络,特别是卷积神经网络(CNN)的发展,神经网络

成为一种自动化地提取面部特征的有力工具。因此，我们特别将这些基于深度学习的方法与上节所讲的传统的方法分开讨论。例如，Dehghan 等人^[34]提出了一种通过门控自动编码融合特征和度量的方法来学习遗传特征。Zhang 等人^[35]利用 CNN 提取人脸高维特征和面部关键点进行亲缘关系验证。Ding 等人^[36]提出了一种鉴别性的低秩度量学习方法（DLML），通过寻找低秩表示来发现全局数据结构。2016 年，Robinson 等人^[8]发布了迄今为止最大的亲属关系数据集，推动了亲属关系验证的发展。同时，他们通过实验证明了他们提出的 CNN 模型优于其他传统方法。Wang 等人^[10]提出了一种基于去噪度量学习（DML）及其边缘化版本（mDML）的去噪自动编码器，以保持数据的内在结构，同时将鉴别性的信息赋予学习的特征。此外，Robinson 等人^[12]利用 SphereFace 处理亲属关系验证问题。Wang 等人^[11]考虑到了年龄差距带来的影响，但是他们引入了 GAN 作为一个中间域，先生成某一年龄的人脸，再进行验证。此外，还有许多其他方法^[13, 15, 30, 37, 38]被提出。总的来说，与传统方法相比，深度模型能够提取出难以手工选择的具有代表性和综合性的特征。

尽管上述这些方法在亲属关系验证问题上取得了不错的效果，但是这几种方法^[8, 16]专注于学习可分离的特征而不是具有鉴别性的特征，所以很难将其直接推广到其他数据集上。因此，它们不能被直接应用于许多现实世界中的开集问题（open-set problem）。近年来，一些研究人员尝试利用区分特征做亲属关系识别问题。Liu 等人^[17]通过使用最大类内距离小于最小类间距离的策略，提出了 SphereFace。之后，Robinson 等人^[12]将其应用于亲属关系验证任务，获得了最先进的结果。尽管 SphereFace 采用了有效的策略来学习区分特征，但它仍然包含会对亲属关系验证结果产生干扰的年龄信息。众所周知，亲属关系不会随年龄变化，因此我们希望提出一种算法，无论年龄差距有多大，它都能够正确验证亲属关系。事实上，Wang 等人^[11]确实考虑到了年龄信息，但是他们是用了一种间接的方式。比如在验证母子关系时，照片中母亲的年龄明显大于儿子，他们会先引入一个中间域来生成年轻时母亲的图片，然后利用该生成的图片来执行亲属关系验证任务。具体来说，他们将问题分为了两个步骤，即先生成中间人脸图像以缩小两张人脸图像的年龄差距，然后再对其进行验证。然而，由于他们需要学习一个生成模型来生成年轻时的人脸图像，所以在训练中会引入额外的误差。此外，

由于他们使用了生成对抗模型^[65]，带来了训练的不稳定性。并且由于该方法是一种两阶段（two-stage）的方法，增加了计算成本。

1.3 存在的问题

尽管上述方法在基于图像的亲属关系验证问题上取得了诸多进展，但是为了使系统更加鲁棒，仍然有一些问题需要被解决。

（1）有效面部特征选取

在某些特定场景下（比如寻找失踪儿童），可能需要验证的双方年龄跨度较大，导致年龄因素对亲属关系的判断产生较大的影响。图 1.1（a）表明，由于年龄差异较大，两个有血缘关系的人之间的距离要比无血缘关系的人要大。具体来说，人脸的形状、皱纹等特征会随着时间推移逐渐改变，但尽管发生了这些变化，二者的血缘关系并不会因此而改变。与人脸识别中的光照变化、角度变化等外部因素变化不同，人脸老化属于人体内部因素变化，因此如何让算法学习到人脸随着年龄增长的变化规律是解决跨年龄亲属关系验证的重要问题之一，这也是朴素的人脸识别系统在亲属关系验证问题中效果并不理想的原因之一。尽管上节所述的方法在亲属关系验证问题上取得了很大进展，但很少考虑到年龄差距。

（2）面部特征权重设置

由于人脸具有较多的特征，在亲属关系验证问题中并不一定这些特征都具有相同的权重。比如当人类在判断两人是否具有亲属关系时，可能会将更多的注意力放在眼睛、鼻子、嘴巴等器官而非脸颊、额头等。鉴于此，我们从直觉上认为眼睛、鼻子等重要器官的权重可能比脸颊、额头等器官的权重要高，可能在亲属关系判断时起着更加决定性的作用。然而，大多数亲属关系验证方法都假设了每种特征的权重都相等，这可能并不一定成立。因此，本文认为应该为不同的特征赋予不同的权重。然而，尽管我们通过先验知识知道某些特征比较重要，但是如何将重要性进行量化、如何确定每种特征具体的权重仍然需要被仔细考虑。因此，如何根据每个特征在亲属关系验证中的重要性为其定量地分配不同的权重就显得尤为必要。

1.4 本文研究内容及贡献

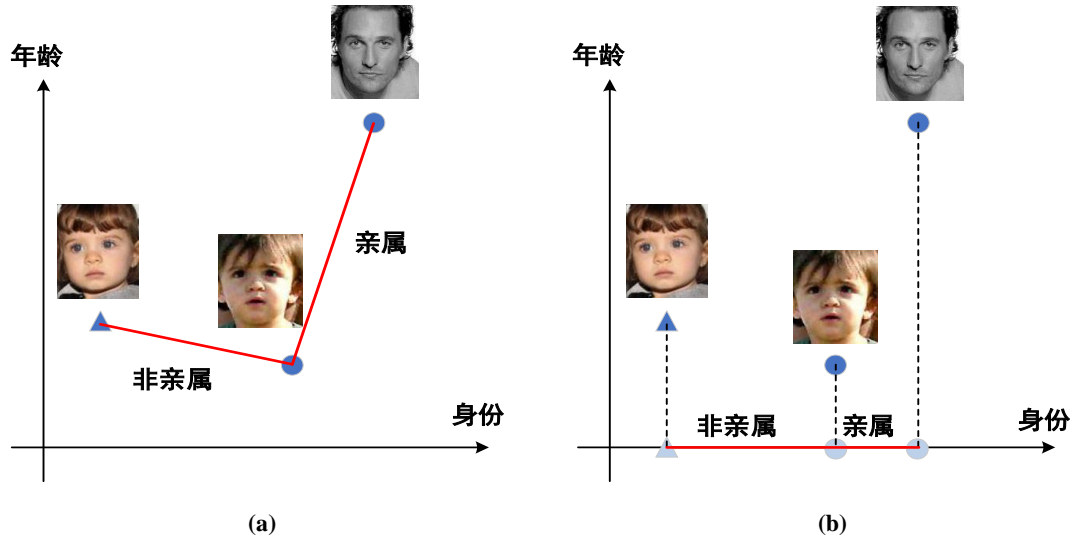


图 1.1 年龄对亲属关系验证的影响. (a) 未消除年龄因素时，亲属之间的距离可能会大于非亲属之间的距离; (b) 我们的方法能够提取身份特征来提高亲属关系验证的准确率

为了解决这些问题，我们提出了一种基于对抗机制的年龄不变的亲属关系验证方法。该方法由两个模块组成，分别是年龄不变的对抗特征学习（Age-Invariant Adversarial Feature Learning, AIAF）模块和基于注意力机制的特征加权（Attention-based Feature Weighted, AFW）模块。如图 1.1（b）所示，AIAF 能够真正剥离与年龄相关的面部特征，仅保留与身份相关的特征，以用于后续的亲缘关系验证。为了进一步提高模型的准确性，我们引入了 AFW 来根据身份特征在亲属关系验证中的重要性自动地为每个特征赋予不同的权重。我们的方法属于完全的端到端（end-to-end）方法，因此它将引入更少的误差，并且在保证了高准确性的同时也具有更好的训练稳定性。

总而言之，本文的贡献主要有以下三个方面：

（1）本文在对亲属关系验证问题进行分析时发现目前已有方法很少考虑到年龄因素，导致年龄相近的人更容易被判断为是亲属。针对该问题，我们设计了一个能有效提取年龄不变特征的年龄不变对抗特征学习模块（AIAF）。在 AIAF 中，一个特征分割器（FS）被用来将原始特征分割成两个特征。训练身份分类器和年龄分类器，使这两个特征分别代表身份信息 and 年龄信息。此外，我们提出了一种对抗机制来训练模型，引入了对抗性典型相关正则项（ACCR）来降低这两个特征之间的相关性。

（2）我们发现大多数亲属关系验证方法默认假设了面部特征都具有相同权

重，这导致算法对重要的特征和不重要的特征一视同仁，违背直觉。我们认为，与人脸识别不同的是，亲属关系验证问题更依赖于某些关键特征，比如眼睛的权重与额头的权重是不一样的。为此，我们精心设计了一个基于注意力机制的特征加权模块（AFW），用于根据不同的面部特征的重要性为其分配不同的权重。最后，利用加权特征对亲属关系做出最终决策。

（3）我们将我们提出的方法在两个公共数据集上进行了评估，数据集包含多达 11 种亲属关系类型。实验结果表明，与基线方法相比，我们的方法取得了最先进的结果。此外，消融实验为我们方法的有效性提供了更多的依据。

1.5 本文组织结构

根据不同内容，本论文主要分为五章，每章内容安排如下：

（1）第一章 绪论。首先简单介绍了研究背景及意义，分析了国内外研究人员针对亲属关系验证问题提出的方法。然后，本文根据现有的方法提出了一些存在的问题，并简要说明了本文的研究内容，提出了一些可能的解决方案。最后介绍了本文的组织结构。

（2）第二章 相关技术介绍。我们将介绍一些本文紧密相关的技术，包含深度学习概述、卷积神经网络、对抗学习、多任务学习和注意力机制，帮助读者能够更好地理解第三章中我们提出的方法。

（3）第三章 年龄不变的对抗特征学习。我们将详细介绍我们提出的方法。包含两个模块：年龄不变的对抗特征学习（AIAF）模块和基于注意力机制的特征加权（AFW）模块。其中 AIAF 能够有效提取年龄不变的身份特征，AFW 根据不同面部特征的重要性为其分配不同的权重。最后，利用加权特征对亲属关系做出最终决策。

（4）第四章 实验结果及分析。我们在两个公共数据集上评估了我们提出的方法，数据集包含多达 11 种亲属关系类型。实验结果表明，与基线方法相比，我们的方法取得了最先进的结果。此外，消融实验为我们方法的有效性提供了更多的依据。

（5）第五章 总结与展望。首先对本文提出的算法进行了回顾和总结，最后展望了今后的研究计划。

第 2 章 相关技术介绍

2.1 引言

本章将对本文提出的方法中使用到的相关技术进行介绍,分别从深度学习概述、卷积神经网络、对抗学习、典型相关分析、多任务学习、注意力机制这几个方面进行介绍,为第三章中提出的年龄不变的对抗特征学习方法奠定理论基础。

2.2 深度学习概述

自 AlphaGo 多次击败人类顶尖职业围棋选手后,人工智能重回人们的视野。得益于大数据时代的到来和 GPU 的飞速发展,人工智能迎来了第三个春天。深度学习(Deep learning)作为机器学习(Machine learning)的一个重要的子分支,在近些年被国内外学者广泛研究。深度学习能够通过学习海量数据样本的内在规律和表示层次,将其泛化至未见过的数据上,对外表现出机器能够像人一样具有举一反三的分析学习能力,在计算机视觉、自然语言处理等方面取得了不错的成绩,甚至在某些任务上已经超过了人类的表现。

2.2.1 深度学习的发展历史

为了帮助读者更加清楚的认识深度学习,本文首先对深度学习的发展历史进行简要介绍。

1943 年,McCulloch 和 Pitts^[77]提出了第一个神经元数学模型——MP 模型。在 20 世纪 50 年代末和 60 年代初,Rosenblatt^[78,79]通过在 MP 模型中加入学习能力,提出了一种单层感知机模型。然而,单层感知机不能处理线性不可分问题(如异或问题)。1986 年,Hinton 等人^[80]针对单层感知器无法解决的问题,提出了一种通过误差反向传播进行训练的多层前馈神经网络——反向传播网络(BP 网络)。由此,人工神经网络逐渐进入人们视野。然而好景不长,当神经网络层数增加时,传统的 BP 网络会遇到局部最优、过拟合、梯度消失和梯度爆炸等问题。此外,由于当时的计算机的计算能力不足以训练大规模的神经网络,并且数据也不够充足,随后的研究中心逐渐向支持向量机(Support Vector Machine)、最近邻分类器等具有坚实数学基础的传统机器学习算法倾斜,神经网络的相关研究被搁置。

2006 年,Hinton 等人^[87]在《科学》杂志上提出了以下两个观点:1)多层人工神经网络具有良好的特征学习能力;2)逐层训练能够有效克服深层神经网络训练的困难。从而带来了新一轮深度学习的研究。

需要强调的是，深度学习并非完全等同于神经网络。深度学习的核心思想是通过深度的结构来逼近复杂的非线性函数。神经网络是实现深度学习的一个强有力的工具，目前大多数深度学习模型都是基于神经网络进行实现的。

2.2.2 深度学习的优势

深度学习之所以具有强大的学习能力是因为其能够表达复杂的非线性变换。能够证明，浅层神经网络只能解决线性问题，而现实世界中的很多问题都是非线性的，比如突变等。很多现有方法在处理非线性问题时通常会采用近似的思想，将一个非线性问题分解为多段线性问题进行求解。虽然这些方法在某些问题上可能达到令人满意的结果，但是从本质上讲他们并没有解决复杂问题的能力。

从结构上讲，深度神经网络是具有层次性的。这决定了其能够表达由多个非线性函数组成的复合函数。从数学上可以证明，初等函数是由幂函数、指数函数、对数函数、三角函数、反三角函数与常数经过有限次的有理运算（加、减、乘、除、有理数次乘方、有理数次开方）及有限次函数复合所产生。由于神经网络中的每一层都可以在前一层的基础上进行变换，在每层之间的激活函数的帮助下具有非线性变换的能力，因此可以通过多层神经网络表示复杂函数。理论上，当每层神经元个数不变时，随着层数的增加，网络的非线性表达能力也随之增加，从而能够表达更加复杂的非线性函数。因此，层数的选取通常取决于待求解问题的复杂程度。

2.2.3 激活函数

在多层神经网络中，两层神经元之间存在一个函数，该函数称为激活函数。神经网络必须使用激活函数来表达非线性函数。与人脑神经元模型类似，神经网络中的每个神经元接受前一层神经元的输出值作为输入，并将处理后的值传递给下一层。这里的激活函数决定了哪些信息应该被传递给下一个神经元。

如果不使用激活函数或使用线性激活函数，则每一层的输入将是前一层输出的线性函数，即整个神经网络只能表示线性函数。例如，假设 $f(x) = W_1x + b_1$ 和 $g(x) = W_2x + b_2$ ，则：

$$\begin{aligned} f \odot g &= W_2(W_1x + b_1) + b_2 \\ &= W_2W_1x + W_2b_1 + b_2 \\ &= Wx + b \end{aligned} \tag{2.1}$$

其中 \odot 表示函数的复合运算。从上式可以看出,当没有非线性的激活函数时,尽管网络具有多层结构,本质上仍然是一种线性结构。理论上,具有非线性激活函数的深层神经网络可以逼近任意函数,极大地提高了神经网络的数据拟合能力。因此,非线性激活函数的引入是非常必要的。下文将简要介绍几种常用的激活函数。

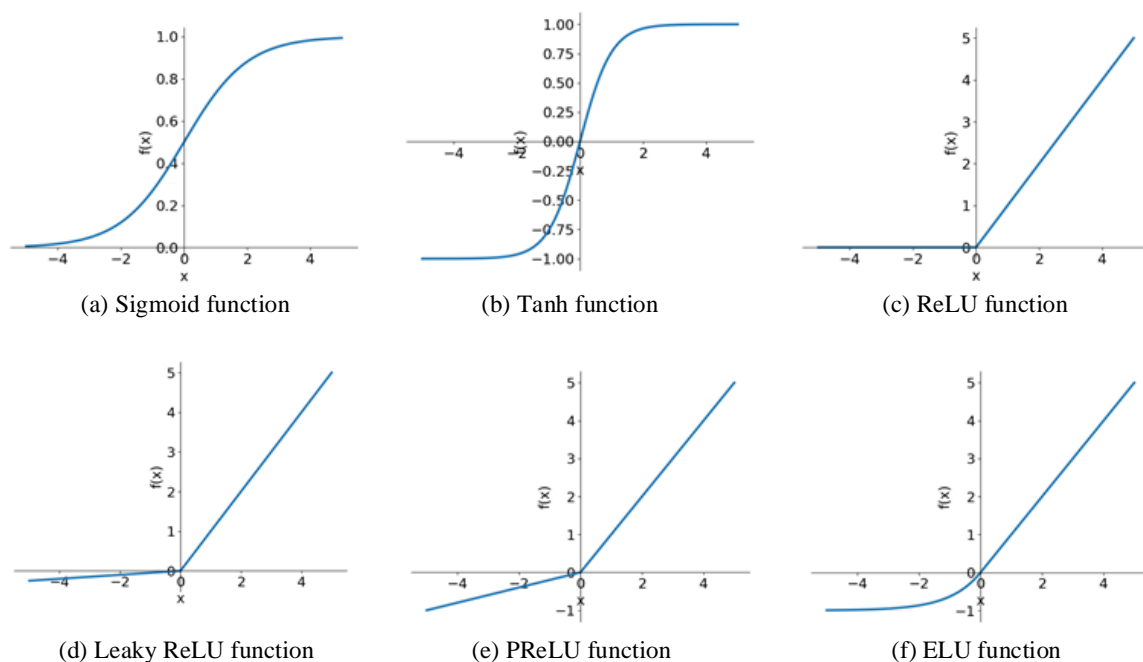


图 2.1 深度神经网络中 6 种常用的激活函数

(1) Sigmoid 激活函数。如图 2.1 (a) 所示, Sigmoid 从整体上看是一个 S 形函数, 当自变量接近 0 时, 梯度变陡, 远离 0 时, 梯度变缓。该函数可以将输入映射到 $(0, 1)$, 因此通常被用在神经网络的最后一层用来解决二分类问题。

(2) Tanh 激活函数。如图 2.1 (b) 所示, Tanh 函数与 Sigmoid 函数一样具有 S 形状, 但是 Tanh 可以将输入映射到 $(-1, 1)$ 。由于其均值为 0, 可以当做一种归一化来使用, 这缓解了梯度消失和梯度爆炸的问题, 使得后续的网络层更容易学习。

(3) ReLU 激活函数。如图 2.1 (c) 所示, ReLU 函数是另一种有效的激活函数。当输入值小于 0 时, 其输出值为 0, 相当于该部分信号被抑制; 当输入值大于或等于 0 时, 其输出值为 x 本身。与 Sigmoid 函数和 Tanh 函数相比, ReLU 函数的一个显著优点是可以加快学习速度, 因为 Sigmoid 和 Tanh 函数的计算涉及指数运算, 在计算导数时需要除法, 而 ReLU 的导数是常数。此外, 在 Sigmoid

函数和 Tanh 函数中，如果输入值太大或太小，函数的梯度很小，这会导致函数收敛缓慢。但是，在 ReLU 函数中，当输入值大于 0 时，其梯度为 1，可以更加快速地收敛。AlexNet 使用 ReLU 作为激活函数，缓解了当网络过深时的梯度消失问题，并验证了 ReLU 在深度网络中优于 Sigmoid 函数。

(4) ReLU 激活函数的变种。不难知道，ReLU 激活函数假设了当输入值小于 0 时都是无用的，会被直接抑制，输出为 0。然而，该假设可能并不总是成立。因此，Leaky ReLU、PReLU、ELU 等变种被提出。如图 2.1 (d) (e) (f) 所示，Leaky ReLU 在输入值小于 0 时增加了一个较小斜率的线性函数，使得输出不总为 0；PReLU 在 Leaky ReLU 基础上做了改进，让输入小于 0 时的斜率基于数据进行动态调整。ELU 函数让输出的平均值接近于 0，起到归一化的作用，使得收敛速度比 ReLU 快。

2.2.4 损失函数

通常，人们利用损失函数 (Loss function) 度量预测值与实际值之间的距离，从而通过训练使得预测值更加接近真实值。即，损失函数通常被用作优化问题的学习准则。损失函数可以与神经网络一起使用来处理回归问题或分类问题，其目标是使损失函数最小化。常见的损失函数包括平均绝对值误差 (MAE)、均方误差 (MSE)、交叉熵 (Cross Entropy) 等。

(1) 回归问题的损失函数

在处理回归问题时，我们通常使用 MAE 或 MSE 作为损失函数。MAE 计算预测值和实际值之间的绝对误差的平均值；MSE 计算它们之间的平方误差的平均值。与 MSE 相比，MAE 对异常值的鲁棒性更强，因为 MSE 会计算异常值的平方，从而增大了异常值的影响。因此，如果训练集中存在大量的异常值，并且这些异常值可能对模型产生负面影响，则更应该考虑使用 MAE 而非 MSE。否则，更应考虑 MSE。此外，因为 MSE 的结果是可导的，所以可以控制它的更新速度；MAE 的结果是不可导的，其更新速度在优化过程中无法确定。

(2) 分类问题的损失函数

在分类任务中，交叉熵损失是最典型的一种损失函数，用来评估当前训练得到的概率分布与实际分布之间的差异。此函数将预测概率与每一类的实际输出值进行比较，并根据它们之间的距离计算损失值。此外，对比损失 (contrastive loss)、

三重损失（triplet loss）等损失函数在人脸识别领域中也有较多的应用。

2.2.5 优化器

从每次优化使用数据量的角度来看，梯度下降的策略可以分为三类：批梯度下降（Batch Gradient Decent, BGD）、随机梯度下降（Stochastic Gradient Descent, SGD）和小批量梯度下降（Mini-Batch Gradient Descent, MBGD）。

若使用 BGD 进行梯度更新，每次更新都需要计算一整批数据以获得梯度，从而确保收敛到凸平面的全局最优和非凸平面的局部最优。但是，使用 BGD 进行更新的速度非常慢，因为需要计算整批样本的平均梯度才能进行一步更新。此外，当数据量非常大时，内存可能无法一次性存储这么多数据。因此，BGD 在实际训练中很少使用。

相反，SGD 每次更新只使用一个样本。很明显，每次更新的 SGD 时间大大少于 BGD，因为只需要计算一个样本的梯度。然而，由于 SGD 具有高方差的特性，会导致目标函数剧烈振荡。一方面，振荡会使模型更容易地跳出局部最优，最终到达一个较好的点；另一方面，由于无休止的振荡，SGD 可能永远无法收敛。

在 BGD 和 SGD 的基础上，MBGD 被提出了。它结合了 BGD 和 SGD 的优点，MBGD 每次更新都使用一小批样本，这样不仅可以比 BGD 更有效地进行梯度下降，而且可以减小方差，使收敛更加稳定。因此，大多数方法都采取 MBGD 作为更新策略。

在 MBGD 的基础上，研究人员们提出了一系列有效的优化算法来加速模型训练过程。本节将简单介绍几种常用的优化器。

在普通的优化器中，如果一个维度的梯度比另一个维度的梯度大得多，学习过程就会变得不平衡。为了解决这个问题，Qian 等人^[88]提出了动量优化算法。它模拟物理动量，使用梯度的指数加权平均值更新权重。动量优化算法可以防止一维的振荡，从而获得更快的收敛速度。一些经典的 CNN 模型，如 VGG^[89]、Inception v1^[90]和 ResNet^[91]等使用了动量优化算法。

然而，动量优化算法总是盲目地进行梯度下降，不对未来的梯度进行考虑。Nesterov 加速梯度（NAG）算法^[92]为动量优化算法提供了一种可预测性，使其在斜率快变为正之前减小学习率。它通过得到下一个位置的近似梯度，从而可以提

前调整步长。

Adagrad 算法^[93]是另一种基于梯度的优化算法。它可以根据参数调整学习率,对频繁的特征相关参数执行较小的更新(即以较小的学习率学习),对不频繁的参数执行较大的步骤更新(即以较大的学习率学习)。因此,它非常适合处理稀疏数据。Adagrad 的主要优点之一是不需要手动调整学习速率。在大多数情况下,我们只使用 0.01 作为默认学习率。FaceNet^[94]等网络在训练中就使用 Adagrad 作为优化器。

RMSprop 算法^[95]被设计用来解决 Adagrad 算法中学习率急剧下降的问题。此外,为了解决损失函数在更新过程中摆动幅度过大的问题,并进一步加快收敛速度, RMSProp 算法对权重和偏置的梯度使用了微分平方加权平均数。MobileNet、Inception v3 和 Inception v4 使用了 RMSprop 训练模型。

另一种常用的梯度优化算法是 Adam^[96]。本质上,它是一种由动量和 RMSprop 组合而成的算法。Adam 存储过去的平方梯度的指数衰减平均值和过去梯度的指数衰减平均值。实践证明, Adam 算法在许多问题上都有很好的应用,适用于许多不同的神经网络结构。

2.3 卷积神经网络

卷积神经网络(CNN)是一种能够通过卷积结构提取特征的前向神经网络。它能够提取输入图像的特征。与传统的特征提取方法^[67-69]不同, CNN 不需要手工提取或筛选特征,通过学习训练样本可以自动提取所需特征。

CNN 的结构来自人类大脑的视觉感知^[70],一个生物神经元对应于一个人工神经元; CNN 卷积核模拟检测不同特征的感受器,每种卷积核可以对某种特定的特征做出反应;激活函数模拟只有当神经电信号超过一定阈值时才能被传递到下一个神经元的功能。

2.3.1 卷积神经网络的发展历史

1987 年, Waibel 等人^[81]提出了一种用于语音识别的时滞神经网络(TDNN),它可以被看作是一个一维的卷积神经网络。之后, Zhang^[82]提出了第一个二维卷积神经网络——平移不变人工神经网络(SIANN)。LeCun 等人^[83]在 1989 年构造了一个用于手写体邮政编码识别的卷积神经网络,并首次使用了 LeNet 的原始版本卷积,通常其被看做是第一个现代的卷积神经网络。20 世纪 90 年代起,各种

浅层神经网络相继被提出，如混沌神经网络^[84]和广义回归神经网络^[85]等，最著名的是 LeNet-5^[86]卷积神经网络。2012 年，Alex 等人^[87]在 ImageNet 大规模视觉识别挑战赛（LSVRC）中使用深度卷积神经网络取得了当时最好的分类结果，引起了研究者的广泛关注^[97,98]，极大地推动了现代卷积神经网络的发展。

2.3.2 卷积神经网络的优势

与普通人工神经网络相比，CNN 具有诸多优点：

（1）局部连接。每个神经元不再与前一层的所有神经元相连，只与少数神经元相连，这有助于减少训练参数、加快收敛速度；

（2）权值共享。一组连接可以共享相同的权重，从而进一步减少参数。

（3）下采样降维。池化（Pooling）层利用图像局部相关的原理对图像进行降采样，在保留有用信息的同时减少数据量。此外，它还可以通过删除不重要的微小特性来减少参数量。上述三点使得 CNN 成为深度学习领域最具代表性的算法之一。

2.3.3 卷积神经网络的构建

具体来说，为了构建一个 CNN 模型，通常需要四个组件。首先，卷积（convolution）操作是特征提取中最重要的一步。卷积的输出通常被称为特征图（feature map）。当卷积核（convolution kernel）的大小固定后进行卷积操作时，由于卷积的特性，边界的信息很有可能会丢失。因此，为了减少信息的丢失，人们通常会在输入端引入了填充（padding）操作在特征图的四周用 0 填充，从而间接地调整输入端的大小。此外，为了控制卷积的密度，可以采用更大的步幅（stride）。步幅越大，卷积密度越低。经过卷积操作后，特征图中可能存在大量容易引起过拟合的特征，导致模型的泛化能力不足，因此，人们提出使用池化（pooling）操作来消除冗余。常用的池化操作包括最大池化（max pooling）和平均池化（average pooling）。一个二维的 CNN 的工作流程如下图所示，通常填充、卷积（以某一步长）、池化三个操作依次轮流进行。

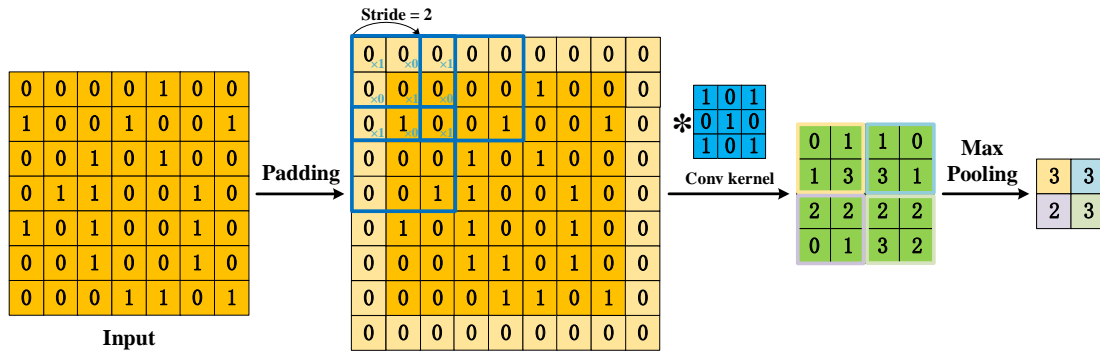


图 2.2 二维卷积工作流程示意图

下图展示了一种经典的卷积神经网络 LeNet-5 的结构图，通常被认为是第一个现代卷积神经网络。输入的二维图像先经过两组卷积层和池化层，再经过全连接层，最后使用 Softmax 作为输出层的激活函数。

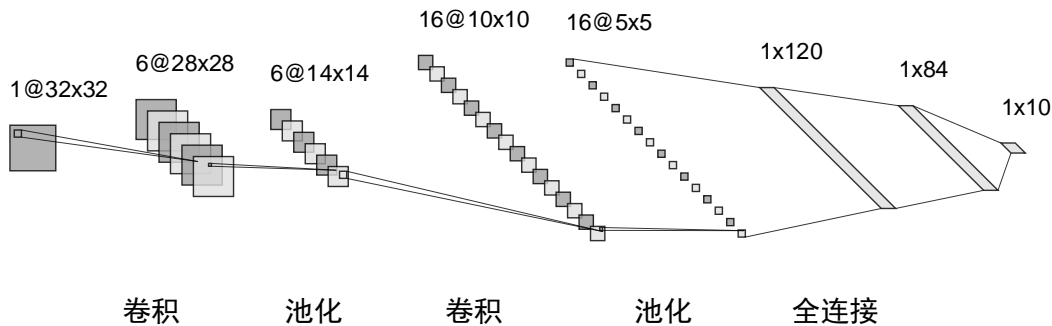


图 2.3 LeNet-5 网络结构图

2.4 对抗学习

对抗学习是一种新兴的机器学习方法，其思想是一种二人零和博弈思想，博弈双方的利益之和是一个常数，一方的利益增加意味着另一方利益的减少。换句话说，对抗学习是让两个“选手”同时玩“一场游戏”，但是双方的目的不同，相互竞争对抗，直到最后达到纳什均衡^[73]状态。

Ian Goodfellow 等人在 2014 年提出的生成对抗网络 (GAN)^[65]是最著名的对抗学习算法之一，该网络由生成模型 (Generator) 和判别模型 (Discriminator) 组成，生成模型通过学习训练数据的分布来生成样本，判别模型通过学习真实样本与生成样本的差异来尽可能区分生成样本。二者在相互博弈的过程中不断学习，直到判别模型无法区分输入的样本是真实的数据还是生成的数据。

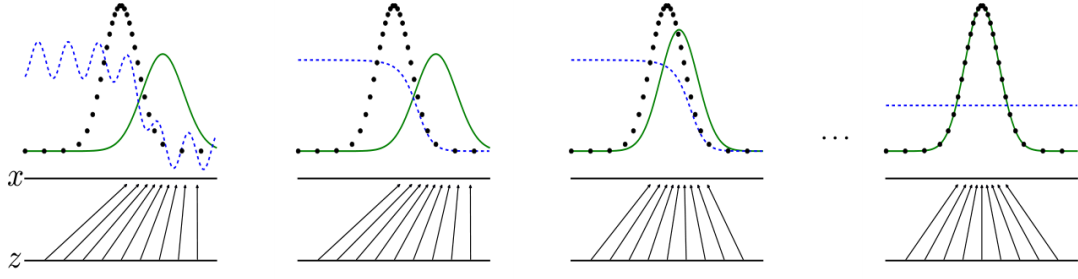


图 2.4 生成对抗网络训练过程示意图

如图 2.4 中上面一行图像所示，黑点表示真实数据样本，生成模型（绿线）逐渐学习拟合真实数据样本，从而能够生成更加真实的数据样本；判别模型（蓝色点线）用来区分数据是由生成模型生成的还是真实数据。可以看出，在训练过程中，生成模型越来越接近真实数据分布，判别模型起初可以正确判断，当生成模型完全学习到真实数据分布时，判别模型无法判断。在下面一行图像中，直线 x 表示真实数据的分布空间，直线 z 表示采样空间，生成模型能够逐渐学习到采样空间向真实数据分布空间的映射关系。用公式可以写作：

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log(D(x))] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2.2)$$

其中 $\log(D(x))$ 表示判别模型对真实数据的判断结果， $\log(1 - D(G(z)))$ 表示判别模型对生成数据的判断结果；通过极大极小博弈，循环交替地分别优化生成模型 G 和判别模型 D 来训练所需要的生成式网络与判别式网络，最终到达纳什均衡点。

在 GAN 被提出之后，很多研究人员借鉴类似的思想，提出了各种类型的生成对抗网络。2018 年的图灵奖得主 Yann LeCun 对其评价 “Adversarial training is the coolest thing since sliced bread.”。很显然 GAN 已经成为一个受欢迎的研究方向，越来越多关于 GAN 的论文发表^[103]。

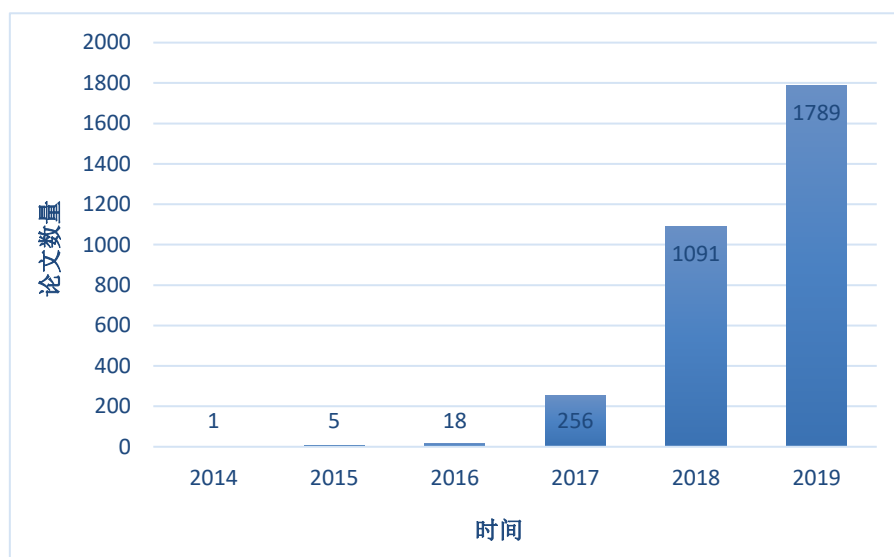


图 2.5 生成对抗网络相关论文在 Web of Science 上的发表情况

Radford 等人^[104]提出了深度卷积生成对抗网络(DCGAN)，他们将卷积神经网络应用到了生成对抗网络中，使得 GAN 能够更好地处理图像相关的任务。Ledig 等人^[105]提出 GAN—SRGAN，可以将低分辨率的图像转换为超分辨率图像。Karras 等人^[106]提出了 Progressive GAN，它可以用一种渐进的方式训练 GAN，同时，他们在训练过程中增加了一个能够处理高分辨率的网络层，使其能生成以假乱真的图像。由于传统的生成对抗网络是基于 Lipschitz 连续空间的，并且要求生成样本必须是可微的，这导致了对离散数据无法使用。因此，Hjelm 等人^[107]提出了边界寻找生成对抗网络 (BGAN)，使其能够训练离散数据。此外，对抗学习还可以应用在许多其他领域，比如在迁移学习中也有一些学者运用对抗学习的思想去解决相关问题。

2.5 典型相关分析

在概率论中，简单相关系数通常被用来研究两个变量之间的线性相关程度。它只能用来研究两个变量之间的关系，不适用于高维数据。1936 年，Hotelling 将简单相关系数进一步推广，提出了典型相关分析 (Canonical Correlation Analysis, CCA)^[64]，其可以用来衡量多个变量之间的线性相关程度，是一种多元统计方法。典型相关分析的基本思想是：从两组随机变量中选取若干个具有代表性的指标，用这些指标的相关系数来衡量两组变量之间的相关关系。在机器学习中，典型相关分析可以作为一种对高维数据进行降维的工具。

假设给定两组变量 $X = (x_1, x_2, \dots, x_m)$ 和 $Y = (y_1, y_2, \dots, y_n)$ ，我们希望衡量 X 和 Y 之间的相关程度，我们可以将 X 和 Y 分别投影到一维，即：

$$X' = A^T X \quad (2.3)$$

$$Y' = B^T Y \quad (2.4)$$

其中 A 和 B 表示投影向量，可以将高维向量投影到一维向量 X' 和 Y' 。此时我们可以计算二者之间的相关系数：

$$\rho(X', Y') = \frac{\text{cov}(X', Y')}{\sqrt{D(X')} \sqrt{D(Y')}} \quad (2.5)$$

典型相关分析的目标是优化投影向量 A 和 B ，使得相关系数 $\rho(X', Y')$ 最大，即：

$$\arg \max_{A, B} \rho(X', Y') \quad (2.6)$$

当 $\rho(X', Y')$ 最大时，表示找到了最优的投影向量 A^* 和 B^* 。

本质上，典型相关分析的目的就是找到两组变量 X 和 Y 最优的线性组合，即典型变量，来代表原变量。

2.6 多任务学习

在传统的机器学习算法中大多数都是单任务学习。如图 2.6 所示，在单任务学习中解决复杂问题时，通常会将其分解为简单且相互独立的子问题，分而治之，最后将结果合并，从而解决原问题。然而，由于现实世界中的很多问题无法被分解为独立的子问题，即使可以分解，各个子问题之间也可能相互存在某种关联。若将某个复杂问题强行拆分成若干个独立的单任务进行处理，这会忽略问题之间所蕴含的关联信息。

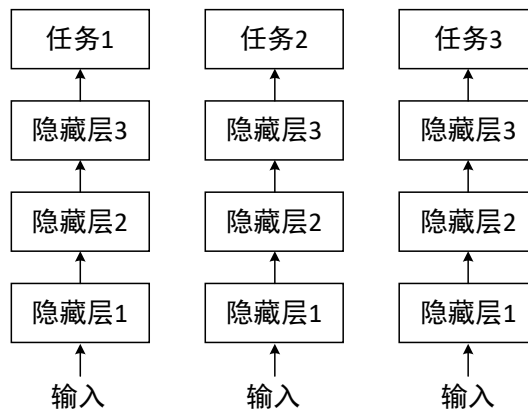


图 2.6 单任务学习模型示意图.

从生物学的角度，人类的大脑可被认为是一个巨大的神经网络，在学习某项任务时，我们经常通过学习相关任务获得的知识来帮助该任务的学习，即把多个相关的任务放在一起学习。例如，一些研究人员根据经验发现打壁球和打网球的技能可以互相帮助提升。由此，多任务学习（multi-task learning）诞生了。单任务学习中各个任务之间的模型是相互独立的，而多任务学习的目标是利用多个学习任务中所包含的有用信息来帮助每个任务学习得到更为准确的模型。因为多个任务之间可共享一些因素，它们可以在学习过程中共享它们所学到的知识，这是单任务学习不具备的，所以通常多任务学习比单任务学习能取得更好的泛化效果。

多任务学习通常可划分为硬参数共享（hard parameter sharing）和软参数共享（soft parameter sharing）。如图 2.7(a)所示，硬参数共享的多任务学习在某些层中所有任务共享一些参数，只有在特定的几层中，每个任务才拥有独有的参数。与硬参数共享不同，在软参数共享的多任务学习中，如图 2.7(b)所示，每个任务除了拥有特定的几层完全独立的参数外，其他网络层的参数也是较为独立的，只是这些参数之间被施加了相互的约束。

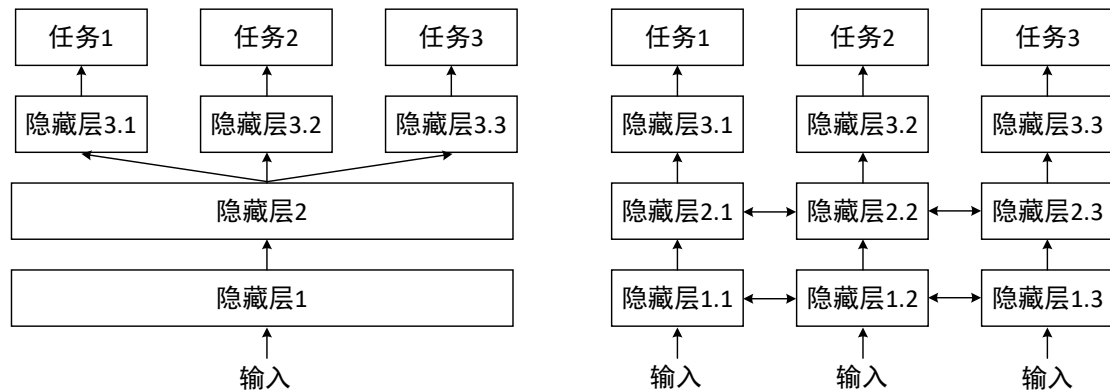


图 2.7 参数共享的多任务学习模型示意图. (a) 硬参数共享; (b) 软参数共享.

从机器学习的角度来看，多任务学习可以被视为一种归纳迁移（inductive transfer）。归纳迁移通过引入归纳偏置（inductive bias）来改进模型，使得模型更倾向于某些假设。比如，在神经网络中常见的一种归纳偏置是 L1 正则化，其能够使得模型产生较多的参数为 0，即 L1 正则化更倾向于稀疏解；而当归纳偏置是 L2 正则化时，模型更倾向于平滑解。在多任务学习场景中，归纳偏置是由辅助任务来提供的，这使得模型更倾向于选择那些可以同时解决多个任务的解，因

而模型具有更强的泛化能力。

2.7 注意力机制

注意力（Attention）机制是人类大脑在处理信号时的一种资源分配机制，是人类在长期进化中形成的一种生存机制。当人类在浏览图片时，大脑会快速扫描全局图像，获得需要重点关注的目标区域，也就是注意力焦点，然后对这一区域投入更多注意力资源，以获取更多所需要关注目标的细节信息，抑制其他无用信息，极大地提高了信息处理的效率。图 2.8 形象化地展示了人类在看到一张图像时是如何高效分配有限的注意力资源的，其中红色区域表明注意力更加集中的区域，可以看出人们会将更多的注意力投入到人的脸部，文本的标题以及文章首句等位置。

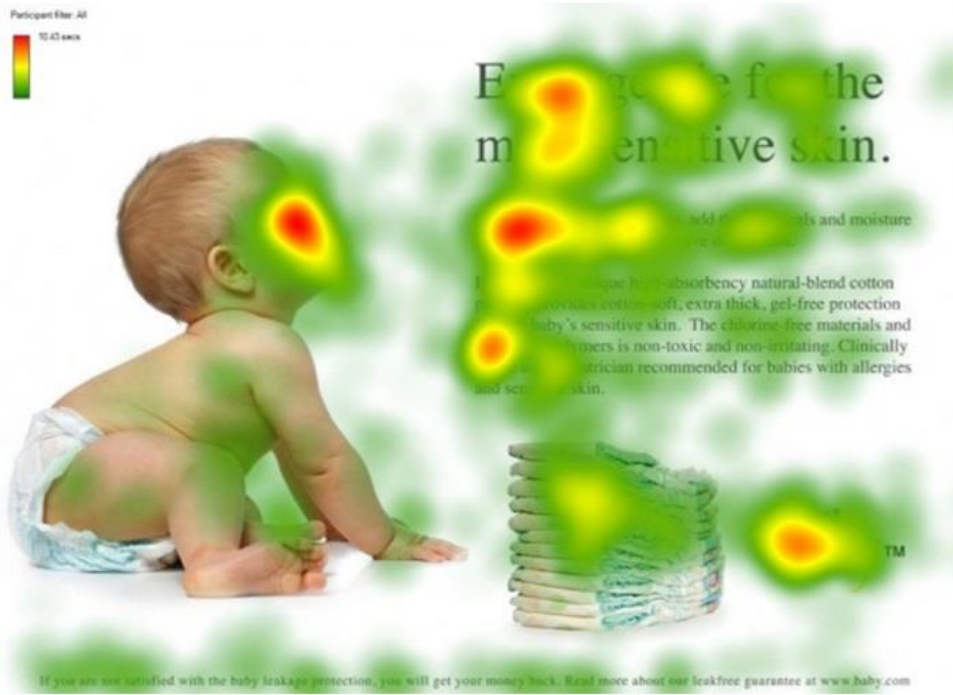


图 2.8 人类视觉注意力样例图

深度学习中的注意力机制借鉴了人类的注意力机制，是对人脑的一种仿生，其核心目标也是从众多信息中选择出对当前任务目标更关键的信息，从而提高信息处理的效率和准确性。目前，注意力机制已经被广泛应用在计算机视觉、自然语言处理和语音识别等各种类型的深度学习任务中，取得了显著的成效，具有极其广泛的研究价值。

2014 年，Google Mind 团队^[99]首先提出在 RNN 模型上使用注意力机制进行

图像分类，取得了很好的结果，随后越来越多的研究人员开始研究注意力机制。Bahdanau 等人^[100]第一个将注意力机制应用在 NLP 领域，提出将其应用在机器翻译任务上，使得翻译和对齐同时进行。之后，Xu 等人^[101]将注意力机制运用在计算机视觉的“看图说话”任务上。此后，注意力机制就被广泛应用在 RNN 模型的各种任务中。随后，如何在卷积神经网络中使用注意力机制也成为了研究的热点问题。2017 年，Google 研究人员^[102]提出在机器翻译任务上大量使用自注意力（self-attention）机制来学习文本表示。不难发现，注意力机制及其衍生算法在信息过载的今天具有广泛的研究意义和商业价值。

2.8 本章小结

在本章，我们介绍了本文方法中使用到的几种技术。具体来说，第一部分主要对深度学习做了简单的梳理，帮助读者了解深度学习领域中的基础知识，分别从深度学习的发展历史、与传统机器学习相比深度学习的优势、在构建神经网络中常用的激活函数、损失函数、优化器进行了讲解。第二部分重点介绍了深度学习中，尤其是计算机视觉与模式识别领域中常用的一种特征提取工具——卷积神经网络。分别从发展历史、与传统神经网络相比卷积神经网络的优势、如何从头构建一个卷积神经网络三个方面进行了简单介绍。第四部分介绍了对抗学习，详细描述了生成对抗网络的思想及其优化过程。第五部分介绍了多任务学习，并分析了多任务学习与单任务学习相比的优势。此外，我们从硬参数共享和软参数共享两个方面进行了介绍。第六部分主要对深度学习中的注意力机制进行了介绍，我们从人类大脑的注意力机制引入，说明为什么注意力机制能够有效提升模型的效率和准确率。本章工作为第三章中提出的年龄不变的对抗特征学习方法奠定了理论基础。

第 3 章 年龄不变的对抗特征学习

3.1 引言

现有的亲属关系验证方法大多利用传统的或深层的模型来提取不同的特征。然而,无论哪种方法,他们都忽略了年龄造成的差异。事实上,年龄与亲属关系是无关。因此,人脸特征向量中的年龄信息在亲属关系验证问题中不会起到积极的作用,甚至可能带来负面影响。如 1.2 节所述,尽管一些研究者考虑到了年龄,但还是存在一些不足。因此,本文的第一个动机是希望设计一个模型架构,能够仅保留与身份相关的特征,从数学角度消除与年龄相关的特征。此外,在亲属关系验证中,研究者们常利用两个人脸向量的余弦相似性来判断其是否具有亲属关系。余弦值越大,他们具有亲属关系的可能性就越大。然而,这里有一个默认的假设,即人脸特征向量中的每个特征都具有相同的权重。为此,我们设计了一个基于注意力的模块,根据特征在亲属关系验证中的重要性自动为每个特征分配不同的权重。

基于上述思想,我们提出了一种新的方法,以对抗的方式学习年龄不变的身份特征,同时,每个特征可以自动学习到不同的权重。该方法主要由两个模块组成:年龄不变的对抗性特征学习模块(AIAF)和基于注意力机制的特征加权模块(AFW)。在下面的小节中,我们将详细描述它们的结构,并解释如何训练每个模块。

3.2 年龄不变的对抗特征学习模块

如图 3.1 所示,本模块由五个组件构成,每个组件负责一个特定的功能。(1) 主干网络。它用于从输入图像中提取面部特征。(2) 特征分割器(FS)。它可以将一个完整的人脸特征向量分割为两个维数相等的特征向量。(3) 身份分类器和(4) 年龄分类器。这两个分类器可以保证被 FS 分割的两个向量分别包含身份信息和年龄信息。然后这两个向量就可以被视为身份相关特征和年龄相关特征。(5) 对抗性典型相关正则项(ACCR)。它负责通过一种对抗机制来最小化两个特征向量之间的相关性。因此,ACCR 保证了身份相关特征和年龄相关特征是不相关的。

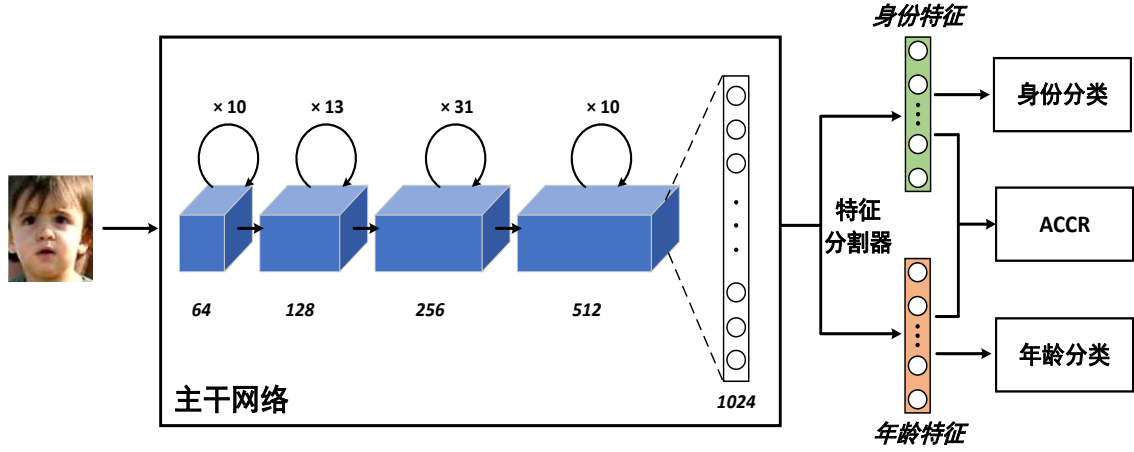


图 3.1 年龄不变的对抗特征学习模块结构图

3.2.1 特征分割器

如上所述，一般的人脸特征既包含身份信息又包含年龄信息。在亲属关系验证问题中，年龄特征是无用的，只有身份特征才是有用的。基于这个想法，我们提出了特征分割器用来将完整的人脸特征分割为两组特征。假设给定一个基于卷积神经网络的主干网络 \mathcal{B} 和一张人脸图片 p_i ，则完整的人脸特征向量 $x_i \in \mathbb{R}^{2L \times 1}$ 可以表示为 $x_i = \mathcal{B}(p_i)$ 。通过 FS， x_i 可被分解为：

$$x_i^1, x_i^2 = \mathcal{S}(x_i) \quad (3.1)$$

其中 \mathcal{S} 表示 FS 的操作， $x_i^1 \in \mathbb{R}^{L \times 1}$ 和 $x_i^2 \in \mathbb{R}^{L \times 1}$ 代表两组特征。需要注意的是 FS 仅仅是将特征向量从中间分割为了两个小的向量，此时没有代表任何的身份特征或年龄特征的含义。

3.2.2 身份特征损失函数

在经过特征分割器之后，两个向量可以被获得。正如 3.1 节所说，我们需要其中一个向量来表示身份特征，另一个来表示年龄特征。假设 x_i^1 表示身份特征，那么我们需要通过神经网络训练 x_i^1 去执行身份分类任务，当在一个巨大的数据集中不同的个体能够被正确分类时，我们认为 x_i^1 就能够提取鉴别性的特征。通常，在分类器中使用的传统的 Softmax 函数虽然能够正确的对人脸进行分类，但是它不够鲁棒，并且泛化性能很差。对于开集问题，我们希望分类器能够学习一条规则，尽可能地增大类间距离，缩小类内距离。目前，已经有许多方法^[39-43]被提出并且达到了令人满意的效果。因此，我们采用了 Additive Angular Margin 损失（ArcFace）^[43]来实现身份特征损失函数。

$$\mathcal{L}_{id} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i^{id}}+m))}}{e^{s(\cos(\theta_{y_i^{id}}+m))} + \sum_{j=1, j \neq y_i^{id}}^n e^{s(\cos \theta_j)}} \quad (3.2)$$

其中 N 是批次的大小， y_i^{id} 是第 i 个样例的身份标签， θ_j 是权重 W 的第 j 列与特征 x_i^1 的夹角， $\cos \theta_j = \frac{w_j^T x_i^1}{\|w_j\| \|x_i^1\|}$ 。此外， s 是一个放缩因子， m 是 $W_{y_i^{id}}$ 和 x_i^1 之间的一个惩罚项，为了减小类内距离并增大类间距离。在实际中， $\|w_j\|$ 通过 L2 归一化被固定为 1， $\|x_i^1\|$ 也通过 L2 归一化放缩到 s 。当 \mathcal{L}_{id} 接近 0 时，特征向量 x_i^1 此时能够被视为身份特征 x_i^{id} ，即 $x_i^{id} = x_i^1$ 。

3.2.3 年龄特征损失函数

在上一节中，我们利用特征向量 x_i^1 来表示与身份相关的特征 x_i^{id} 。在本节，剩余的特征向量 x_i^2 将用于表示与年龄相关的特征 x_i^{age} 。理论上，我们可以使用年龄标签来训练 x_i^2 向量，但是由于噪声的存在，年龄值很难精确地被标注。如果年龄标签按照所有整数进行分类，那么很容易导致错误分类。换句话说，即使是人类也很难识别一个人是 20 岁还是 21 岁。因此，我们的策略是将年龄分组，每个年龄标签都属于其中一个组。然后，我们采用 Softmax 和交叉熵损失对年龄分类任务进行优化。

$$\hat{y}_i^{age} = softmax(x_i^2) = \frac{e^{W_{y_i^{age}}^T x_i^2 + b_{y_i^{age}}}}{\sum_{j=1}^n e^{W_j^T x_i^2 + b_j}} \quad (3.3)$$

$$\mathcal{L}_{age} = -\frac{1}{N} \sum_{i=1}^N y_i^{age} \log(\hat{y}_i^{age}) \quad (3.4)$$

其中 x_i^2 指的是我们希望它来表示第 i 个样本的年龄特征， \hat{y}_i^{age} 是 x_i^2 经过 softmax 操作后的输出， y_i^{age} 是第 i 个样本的年龄标签， $W_{y_i^{age}}$ 和 $b_{y_i^{age}}$ 是 y_i^{age} 列的权重和偏置项。当 \mathcal{L}_{age} 接近 0 时，特征向量 x_i^2 此时能够被视为年龄特征 x_i^{age} ，即 $x_i^{age} = x_i^2$ 。

3.2.4 对抗性典型相关正则项损失函数

当一个完整的人脸特征向量 x_i 被 FS 分割成两个小的 x_i^1 和 x_i^2 向量时，他们可

能由于一些潜在因素导致是相关的。在对身份特征分类和年龄特征分类之后， x_i^1 和 x_i^2 能够代表身份特征 x_i^{id} 和年龄特征 x_i^{age} 。然而， x_i^{id} 和 x_i^{age} 的相关性并没有消除。因此，受到 Wang 等人^[44] 的启发，我们设计了对抗性典型相关正则项去最小化 x_i^{id} 和 x_i^{age} 之间的相关性。

在亲属关系验证问题中，由于身份相关的特征和年龄相关的特征是高维的，不能直接计算相关系数。相关系数只考虑两组特征之间的相关性，并未考虑两组内变量之间的相关性。因此，为了从整体上分析身份特征和年龄特征之间的相关性，我们利用典型相关分析^[64] 提取两个代表性变量（降维）来反映两组特征之间的总体相关性。具体来说，我们输入的是一批图像，而不是一个单独的图像。如图 3.2 所示，通过主干网络和 FS 模块，我们可以获得它们的身份特征和年龄特征。然后，由 ACCR 将它们分别投影到身份向量和年龄向量中。最后一步是输出两个向量之间的相关系数。

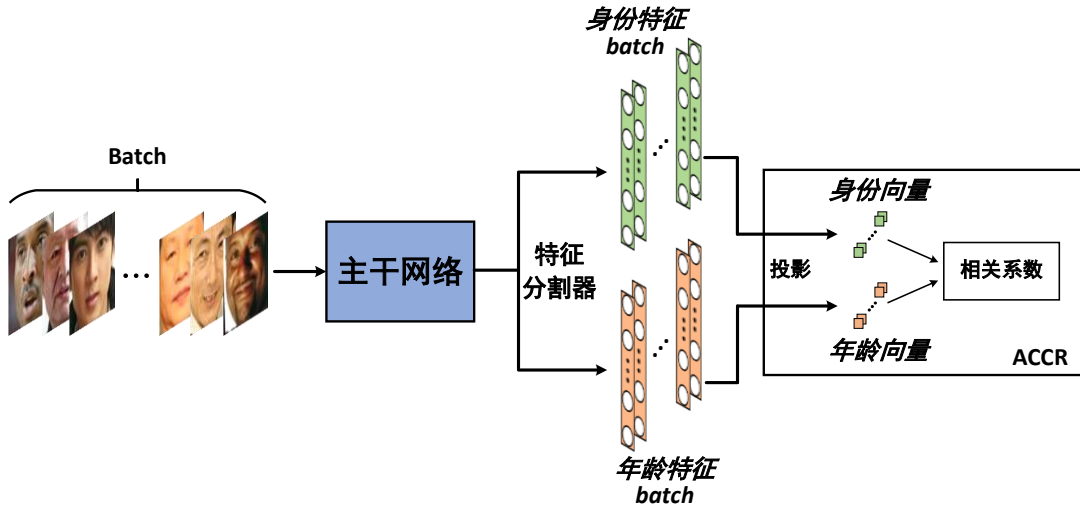


图 3.2 ACCR 训练流程图

假设给定两组由 FS 提取的特征 $\mathcal{X}_{id} \in \mathbb{R}^{L \times n}$ 和 $\mathcal{X}_{age} \in \mathbb{R}^{L \times n}$ ，其中 n 和 L 分别代表一批图像的大小和特征的维度。 \mathcal{X}_{id} 和 \mathcal{X}_{age} 的典型相关可以表示为：

$$\mathcal{X}'_{id} = W_{id}^T \mathcal{X}_{id} \quad (3.5)$$

$$\mathcal{X}'_{age} = W_{age}^T \mathcal{X}_{age} \quad (3.6)$$

其中 $W_{id} \in \mathbb{R}^{L \times 1}$ 和 $W_{age} \in \mathbb{R}^{L \times 1}$ 是可训练的向量， $\mathcal{X}'_{id} \in \mathbb{R}^{1 \times n}$ 和 $\mathcal{X}'_{age} \in \mathbb{R}^{1 \times n}$

是投影之后的向量。 W_{id} 和 W_{age} 能够将矩阵 \mathcal{X}_{id} 和 \mathcal{X}_{age} 分别投影到 \mathcal{X}'_{id} 和 \mathcal{X}'_{age} 。

为了优化 W_{id} 和 W_{age} ，我们定义相关系数为：

$$\rho(\mathcal{X}'_{id}, \mathcal{X}'_{age}) = \frac{cov(\mathcal{X}'_{id}, \mathcal{X}'_{age})}{\sqrt{D(\mathcal{X}'_{id})} \sqrt{D(\mathcal{X}'_{age})}} \quad (3.7)$$

基于该公式， \mathcal{X}'_{id} 和 \mathcal{X}'_{age} 之间的相关系数可以被计算得到。为了优化过程中的便捷性，我们使用 $\rho^2(\mathcal{X}'_{id}, \mathcal{X}'_{age})$ 作为目标函数。首先，我们冻结主干网络 \mathcal{B} 和特征分割器 \mathcal{S} ，然后训练 W_{id} 和 W_{age} 来最大化 ρ^2 。当 ρ^2 达到最大值后，我们解冻 \mathcal{B} 和 \mathcal{S} ，冻结 W_{id} 和 W_{age} 来最小化 ρ^2 。这两个操作交替进行直至 ρ^2 收敛至一个极小值。此时，向量 W_{id}^* 和 W_{age}^* 取得最优值。

这个训练过程可以被视为一个对抗游戏。一方想要最大化 ρ^2 ，另一方想要最小化它。换句话说，整个训练过程是反复地最小化最大的 ρ^2 ，这个过程可以表示为：

$$\mathcal{X}_{id}, \mathcal{X}_{age} = \mathcal{S}(\mathcal{B}(p_{1...n})) \quad (3.8)$$

$$\mathcal{L}_{ACCR}(\mathcal{X}_{id}, \mathcal{X}_{age}) = \min_{\mathcal{B}, \mathcal{S}} \max_{W_{id}, W_{age}} \left(\rho^2(W_{id}^T(\mathcal{X}_{id}), W_{age}^T(\mathcal{X}_{age})) \right) \quad (3.9)$$

其中 $p_{1...n}$ 表示一批包含 n 张图片的数据。总的来说，这个模块的目的是为了最小化 \mathcal{X}_{id} 和 \mathcal{X}_{age} 的典型相关。在典型相关分析的理论指导下，我们将原问题转化为了最小化 \mathcal{X}'_{id} 和 \mathcal{X}'_{age} 之间最大的相关系数。最终，经过足够的训练，由 FS 分割的两组向量 \mathcal{X}_{id} 和 \mathcal{X}_{age} 总是不相关的。

3.2.5 多任务学习

总的来说，为了训练不同的组件同时完成各自的工作，我们将上述三个损失组合成多任务学习损失，公式如下：

$$\mathcal{L} = \mathcal{L}_{id} + \lambda_1 \mathcal{L}_{age} + \lambda_2 \mathcal{L}_{ACCR} \quad (3.10)$$

其中 \mathcal{L}_{id} 、 \mathcal{L}_{age} 、和 \mathcal{L}_{ACCR} 分别是身份分类、年龄分类和 ACCR 的损失项。 λ_1 和 λ_2 是为了平衡三者权重的平衡因子。

为了更加清晰地阐述，AIAF 模块的训练过程可以用伪代码描述为：

算法 1 年龄不变的对抗特征学习算法

Input:

$T \in \mathbb{N}^+$: AIAF 的训练轮数
 $N \in \mathbb{N}^+$: AIAF 的训练样本总数
 $n \in \mathbb{N}^+$: 训练 AIAF 时每批数据的大小
 \mathcal{B} : 基于卷积神经网络的主干网络
 \mathcal{S} : 特征分割器
 W_{id}, W_{age} : 分别为身份和年龄的投影向量
 \mathcal{L}_{id} : 身份分类的损失函数
 \mathcal{L}_{age} : 年龄分类的损失函数
 \mathcal{L}_{ACCR} : 对抗性典型相关正则项损失
 $\lambda_1, \lambda_2 \in \mathbb{R}^+$: 平衡因子

Procedure:

```

1: for  $t = 1$  to  $T$  do
2:   for  $i = 1$  to  $\lfloor N/n \rfloor$  do
3:     Randomly sample unsampled data:
            $D_i = \{(x_1, y_1^{id}, y_1^{age}), \dots, (x_n, y_n^{id}, y_n^{age})\}$ 
4:      $\mathcal{X}_{id}, \mathcal{X}_{age} = \mathcal{S}(\mathcal{B}(D_i))$ 
5:      $\mathcal{X}'_{id} = W_{id}^T(\mathcal{X}_{id})$ 
6:      $\mathcal{X}'_{age} = W_{age}^T(\mathcal{X}_{age})$ 
7:      $\mathcal{L} = \mathcal{L}_{id}(\mathcal{X}_{id}, y_i^{id}) + \lambda_1 \mathcal{L}_{age}(\mathcal{X}_{age}, y_i^{age}) + \lambda_2 \mathcal{L}_{ACCR}(\mathcal{X}'_{id}, \mathcal{X}'_{age})$ 
8:     Jointly optimize  $\mathcal{L}$  via backpropagation to find optimal  $\mathcal{X}_{id}^*$ 
9:   end for
10: end for
    
```

3.3 基于注意力机制的特征加权模块

经过年龄不变的对抗特征学习模块 (AIAF) 的训练后, 我们可以提取出去除年龄的人脸特征。从人脸识别的角度来看, 这已经足够进行后续的任务了。但是从亲属关系验证的角度来看, 还需要考虑更多的细节。如第 3.1 节所述, 大多数现有算法利用两个面部特征的余弦相似性来确定它们是否是血缘关系。在这种情况下, 他们假设每个特征都具有相同的权重。事实上, 当人们决定亲属关系时, 我们通常更关注一些重要的器官, 比如眼睛、鼻子和嘴, 而不是脸颊和前额。虽然我们知道有些面部特征更为重要, 但每个特征的重要程度如何? 每个特征的重要性比例应该是多少? 此外, 基于深度学习方法提取出的人脸特征通常是高维的, 因此很难确定哪个点代表鼻子或眼睛。基于此, 我们设计了一个基于注意力机制的特征加权模块 (AFW), 它可以根据重要性自动为提取的身份特征分配不同的权重。

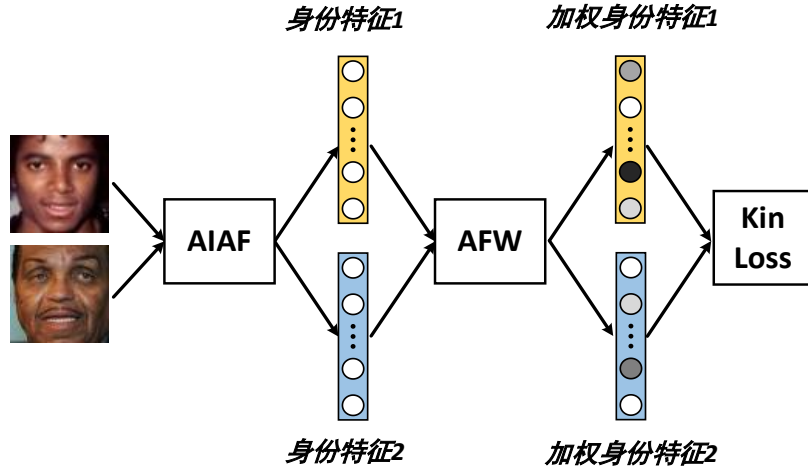


图 3.3 AFW 工作流程图

如图 3.3 所示, AFW 的输入是通过 AIAF 提取的两个身份特征。我们希望找到一个权重矩阵能够代表不同特征的权重。给定两个原始的身份特征 $x_1^{id} \in \mathbb{R}^{L \times 1}$, $x_2^{id} \in \mathbb{R}^{L \times 1}$ 和权重矩阵 $\mathcal{M} \in \mathbb{R}^{L \times L}$, 记加权之后两人的身份特征分别为 $\hat{x}_1^{id} \in \mathbb{R}^{L \times 1}$ 和 $\hat{x}_2^{id} \in \mathbb{R}^{L \times 1}$, 则可表示为:

$$\forall i \in \{1, 2\}: \hat{x}_i^{id} = \mathcal{M} \cdot x_i^{id} \quad (3.11)$$

需要注意的是我们将 \mathcal{M} 初始化为单位矩阵。一方面, 它代表在初始时为所有

特征分配了相同的权重；另一方面，因为 AIAF 提取的特征在大多数情况下都是可靠的，所以它能够加速训练。

之后， \tilde{x}_1^{id} 和 \tilde{x}_2^{id} 的余弦相似度可以通过下式进行计算：

$$\cos \theta = \frac{\tilde{x}_1^{id} \cdot \tilde{x}_2^{id}}{\max(\|\tilde{x}_1^{id}\|_2 \cdot \|\tilde{x}_2^{id}\|_2, \epsilon)} \quad (3.12)$$

其中 ϵ 是一个略大于 0 的值用来避免除 0 错误。 $\cos \theta$ 的范围是从 -1 到 1。我们使用如下的线性变换来改变它值域。

$$\tilde{y} = \frac{\cos \theta + 1}{2} \quad (3.13)$$

其中 $\tilde{y} \in [0, 1]$ ，它可以被视为具有亲属关系的概率。因此，它能够与二分类交叉熵（Binary Cross Entropy）共同使用来优化目标。给定真实标签 y ，亲属关系的损失函数 \mathcal{L}_{kin} 可被表示为：

$$\mathcal{L}_{kin} = -(y \cdot \log \tilde{y} + (1 - y) \cdot \log(1 - \tilde{y})) \quad (3.14)$$

在训练过程中， \mathcal{M} 可以通过最小化损失函数 \mathcal{L}_{kin} 来学习特征的重要性。最后，最优的特征加权矩阵 \mathcal{M}^* 可以学习到将更多的注意力放到重要的特征而非不重要的特征。

3.4 本章小结

本章针对如何选取有效的面部特征（剔除年龄的身份特征）、如何根据重要性为面部特征赋予不同权重两个问题提出了年龄不变的对抗特征学习方法，其中包含两个模块：年龄不变的对抗性特征学习模块（AIAF）和基于注意力机制的特征加权模块（AFW）。3.2 节和 3.3 节分别详细介绍了这两个模块的内部结构和训练方式。

第 4 章 实验结果及分析

4.1 引言

在本章，我们将对提出的算法进行测试，并与我们所知的当前最先进的一些方法进行比较。

本文采用的实验环境为：256GB 内存，Intel® Xeon® E5-2640 v4 (X2) 处理器，GPU 为 NVIDIA TITAN RTX (24GB)，实现语言为 Python 3.7.6、Pytorch 1.4.0。为加快模型的训练速度，本文使用 GPU 进行高速计算。同时使用 CUDA 9.0 和 cuDNN 加速库，进一步实现高性能并行计算。

4.2 实验设置

4.2.1 网络架构

主干网络。如图 3.1 所示，我们使用 65 层的卷积神经网络来构建主干网络。每个卷积层由批归一化、卷积核为 3 的卷积操作和 Leaky ReLU 构成。最后一层是具有 1024 个神经元的全连接层。

特征分割器。在本文的实现中，FS 是直接将 1024 维的输入特征向量从中间分割为两个 512 维特征向量的操作。

身份分类器。我们设置 angular margin 惩罚因子 m 为 0.3、放缩因子 s 为 64。

年龄分类器。它由三个全连接层构成，神经元个数分别为 512、512 和 8。前两层的输出用 Leaky ReLU 激活函数。此外，考虑到数据集中每类样本数量的平衡、处于青春期的儿童发育速度快、老化特征（如抬头纹、鱼尾纹）出现等问题，我们使用 Wang 等人^[44]的划分方式，将年龄标签分为八组： ≤ 12 、13-18、19-25、26-35、36-45、46-55、56-65、 ≥ 66 。

对抗性典型相关正则项。投影向量 W_{id} 和 W_{age} 的维度均为 512×1 。在迭代训练的过程中，我们先训练 20 步来最大化 ρ^2 ，再训练 50 步来最小化 ρ^2 。

多任务学习。根据经验，我们将 λ_1 和 λ_2 都设置为 0.1。一批数据的大小为 512，所有的组件都用 Adam 优化器进行优化，初始学习率为 $1e^{-4}$ ，训练大约 10 轮。

基于注意力机制的特征加权模块。特征权重矩阵 \mathcal{M} 被初始化为 512×512 的对角矩阵。 ϵ 被设置为 $1e^{-6}$ 。

4.2.2 数据预处理

在本文提出的方法中，我们利用多任务学习来训练 AIAF。人脸图像必须同时具有身份标签和年龄标签，并且这些图像应该具有较大的年龄差异。然而据我们所知，没有足够的数据集能够满足我们的需求。因此，我们的策略是先训练一个年龄估计模型来标记只包含身份标签的数据集。具体来说，首先根据这些工作^[45, 46]，我们在 IMDB-WIKI 数据集上训练了一个年龄估计模型，并依据这些方法^[47, 48]进行了数据增强。其次，将年龄估计模型应用于仅包含身份标签的 VGGFace2 数据集^[49]。然后，我们将包含两个标签的 CACD2000 数据集^[50]与已处理的 VGGFace2 数据集混合。最后，我们使用 MTCNN^[51]检测出面部的五个关键点（左眼、右眼、鼻子、左嘴角、右嘴角）。我们利用这些关键点来对齐人脸并将图像裁剪成 64×64 像素，作为 AIAF 模块的最终数据集。

4.3 评估指标

本文所解决的问题属于二分类问题，我们使用五折交叉验证进行实验，用平均准确率、接受者操作特性曲线（ROC）和 ROC 曲线下方的面积大小（AUC）作为算法的评估指标，平均准确率和 AUC 的值越大代表算法的性能越好。

4.4 在 KinFaceW-I 上的实验

KinFaceW-I^[2, 53]是一个被广泛使用的亲属关系数据集。该数据集由四种关系（父子、父女、母子和母女）组成，分别有 156、134、116、127 对亲属图片。下图展示了 KinFaceW-I 中部分图像对。



图 4.1 KinFaceW-I 数据集中部分图像对

在实验中，我们遵循大多数算法使用的 image-restricted 实验设置^[4]，该设置未使用 KinFaceW-I 数据集对特征提取模型进行微调，而是假设已经提取到特征之后进行的一系列操作，因此难以直接使用 AIAF 模块进行测试。我们使用预训练好的 AIAF 模块作为人脸身份特征提取器，然后使用 KinFaceW-I 的原论文^[4]中

预先划分好的训练集训练 AFW 模块，最后在预先划分好的测试集上进行测试。其中训练集和测试集是随机并且独立生成的。我们将本文提出的 AIAF+AFW 与当前最先进方法的结果^[4, 14]进行了比较，如表 4.1 和图 4.2 所示，我们的方法取得了最好的结果。

表 4.1 在 KinFaceW-I 数据集上 image-restricted^[4]条件下的准确率 (%)

方 法	亲属关系	父母-子女				平均
		F-S	F-D	M-S	M-D	
SILD (LBP ^[55])		78.22	69.40	66.81	70.10	71.13
SILD (HOG ^[56])		80.46	72.39	69.82	77.10	74.94
BIU ^[4]		86.90	76.48	73.89	79.75	79.25
NUAA ^[4]		86.25	80.64	81.03	83.93	82.96
ULPGC ^[4]		71.25	70.85	58.52	80.89	70.01
LIRIS ^[4]		83.04	80.63	82.30	84.98	82.74
IML ^[27]		70.5	67.5	65.5	72.0	68.9
MNRML ^[2]		72.5	66.5	66.2	72.0	69.3
MPDFL ^[25]		67.5	73.5	73.1	66.1	70.1
GA ^[34]		72.5	76.4	77.3	71.9	74.5
HumanA ^[2]		61.0	58.0	66.0	70.0	63.8
HumanB ^[2]		67.0	65.0	75.0	77.0	71.0
AIAF+AFW(Ours)		88.7	80.8	75.0	88.2	83.18

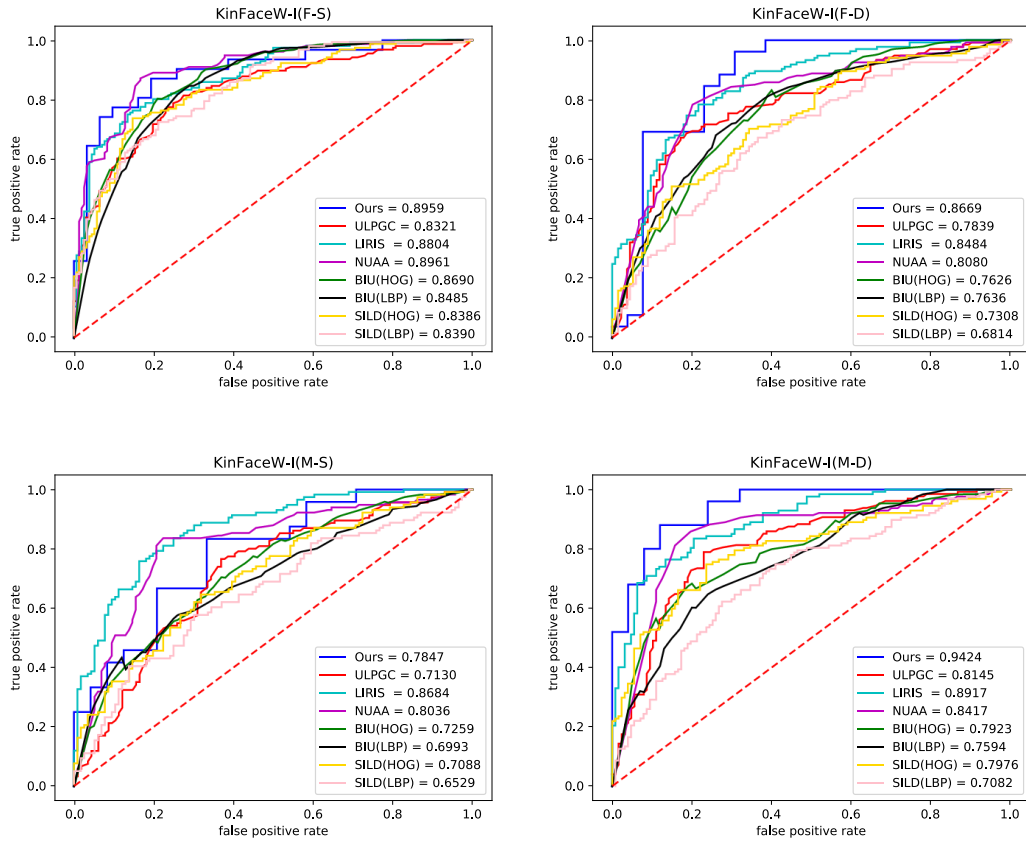


图 4.2 AIAP+AFW 和其他方法在 KinFaceW-I 数据集上的 ROC 曲线对比图

从实验结果中我们观察到：（1）我们的方法在大多数关系上和平均准确率上取得了最好的结果。（2）在 M-S 关系中，我们的方法并未超过 LIRIS、NUAA 等方法。通过分析，我们认为原因是 M-S 关系的数据量太少导致的。在小数据集上，基于深度学习的方法与传统方法相比失去了优势。

4.5 在 FIW 上的实验

为了证明我们的方法在大规模数据集上的优势，我们使用 Families In the Wild (FIW) [12] 数据集进行实验。据我们所知，该数据集是迄今为止最大、最全面的亲属关系数据集。它由 1,000 个家庭的成员，共 11,932 张图片组成，共包含 11 种关系（父女（F-D）、父子（F-S）、母女（M-D）、母子（M-S）、兄弟姐妹（B-B）、姐妹（S-S）、姐妹兄弟（SIBS）、祖父孙女（GF-GD）、祖父孙子（GF-GS）、祖母孙女（GM-GD）和祖母孙子（GM-GS））共有 656,954 个图像对。与 KinFaceW-I 数据集不同的是，FIW 中的人脸更加多样化，包含各种角度和光照，并且未对人脸进行对齐，下图展示了 FIW 中的部分图像。

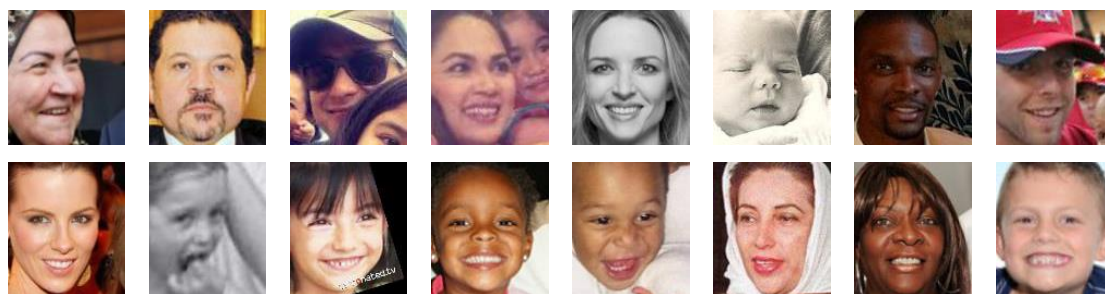


图 4.3 FIW 数据集中部分图像对

为了公平比较，我们遵循公认的实验方案^[12]进行五折交叉验证。我们将我们的方法与基于深度学习的方法 and 传统方法进行了比较，平均准确率、AUC 和 ROC 作为评价指标，实验结果^[11, 12]如表 4.2 和图 4.4 所示。需要注意的是，在表 4.2 中一些方法有两行结果，上面一行表示原始亲属关系对产生的结果，下面一行是通过引入这篇工作^[11]中提出的年轻父母/祖父母中间域来减小年龄差距的实验结果，他们未对兄弟姐妹关系进行测试。实验表明，我们的 AIAF+AFW 与其他方法相比取得了最好的结果，这证明了我们的方法在大规模数据集上的优越性。

表 4.2 在 FIW 数据集上五折交叉验证的亲属关系验证准确率 (%)
(每折之间的家庭无交叉)

方 法	亲属关系	兄弟姐妹			父母-子女				祖孙				平均
		B-B	S-S	SIBS	F-D	F-S	M-D	M-S	GF-GD	GF-GS	GM-GD	GM-GS	
LBP ^[55]		55.52	57.49	55.39	55.05	53.77	55.69	54.65	55.79	55.92	54.00	55.36	55.33
SIFT ^[56]		57.86	59.34	56.91	56.37	56.24	55.05	56.45	57.25	55.35	57.29	56.74	56.80
VGG-Face ^[57]		69.67	75.35	66.52	64.25	63.85	66.43	62.80	62.06	63.79	57.40	61.64	64.89
VGG+ITML ^[58]		57.15	61.61	56.98	58.07	54.73	57.26	59.09	62.52	59.60	62.08	59.92	59.00
VGG+LPP ^[59]		67.61	66.22	71.01	62.54	61.39	65.04	63.54	63.50	59.96	60.00	63.53	64.03
VGG+LMNN ^[60]		67.11	68.33	66.88	65.66	67.08	68.07	66.16	61.90	60.44	63.68	60.15	65.04
VGG+GmDAE ^[61]		68.05	68.55	67.33	66.53	68.30	68.15	66.71	62.10	63.93	63.84	63.10	66.05
VGG+DLML ^[36]		68.03	68.87	67.97	65.96	68.00	68.51	67.21	62.90	63.96	63.11	63.55	66.19
VGG+mDML ^[10]		69.10	70.15	68.11	67.90	66.24	70.39	67.40	65.20	66.78	63.11	63.45	67.07
SphereFace ^[17]		71.94	77.30	70.23	69.25	68.50	71.81	69.49	66.07	66.36	64.58	65.40	69.18
ResNet-22 ^[62]		65.57	69.65	60.12	59.45	60.27	61.45	59.37	55.37	58.15	59.74	59.70	61.34
ResNet+LPP ^[59]		69.04	75.84	67.41	65.86	65.12	68.64	64.82	58.38	59.73	62.29	62.63	65.47
		-	-	-	66.47	65.31	68.59	64.57	60.51	60.85	63.03	62.88	65.59
ResNet+LMNN ^[60]		70.38	75.18	65.29	63.53	64.46	67.01	63.64	60.31	57.57	60.81	56.01	64.06
		-	-	-	64.73	64.77	66.67	64.04	61.91	59.42	61.90	57.75	64.55
ResNet+NPE ^[63]		71.55	76.09	63.26	62.01	63.11	65.60	62.91	58.64	57.71	61.86	58.63	63.76
		-	-	-	63.99	63.46	65.90	63.37	60.15	59.42	62.82	61.43	64.67
ResNet+DLML ^[36]		71.41	75.66	68.61	64.31	64.56	67.63	65.67	61.36	59.99	60.60	59.78	65.42
		-	-	-	66.00	65.60	68.26	66.88	62.38	60.64	61.71	61.28	66.22
ResNet+mDML ^[10]		71.81	77.06	69.94	66.45	65.16	68.55	68.33	63.04	61.19	62.49	61.49	66.86
		-	-	-	67.58	66.48	69.53	68.45	64.11	62.58	63.37	62.73	67.60
ResNet+SDMLoss ^[11]		72.61	79.38	70.35	68.27	67.96	71.32	68.77	64.73	63.98	64.64	63.54	68.68
		-	-	-	69.02	68.60	72.28	69.59	65.89	65.12	66.41	64.90	69.47
AIAF (Ours)		70.05	84.18	76.89	75.86	76.63	73.93	78.20	63.61	62.03	70.26	64.86	72.41
AIAF+AFW (Ours)		73.86	85.57	77.60	79.13	78.22	76.16	86.59	69.37	69.36	70.53	78.38	76.80

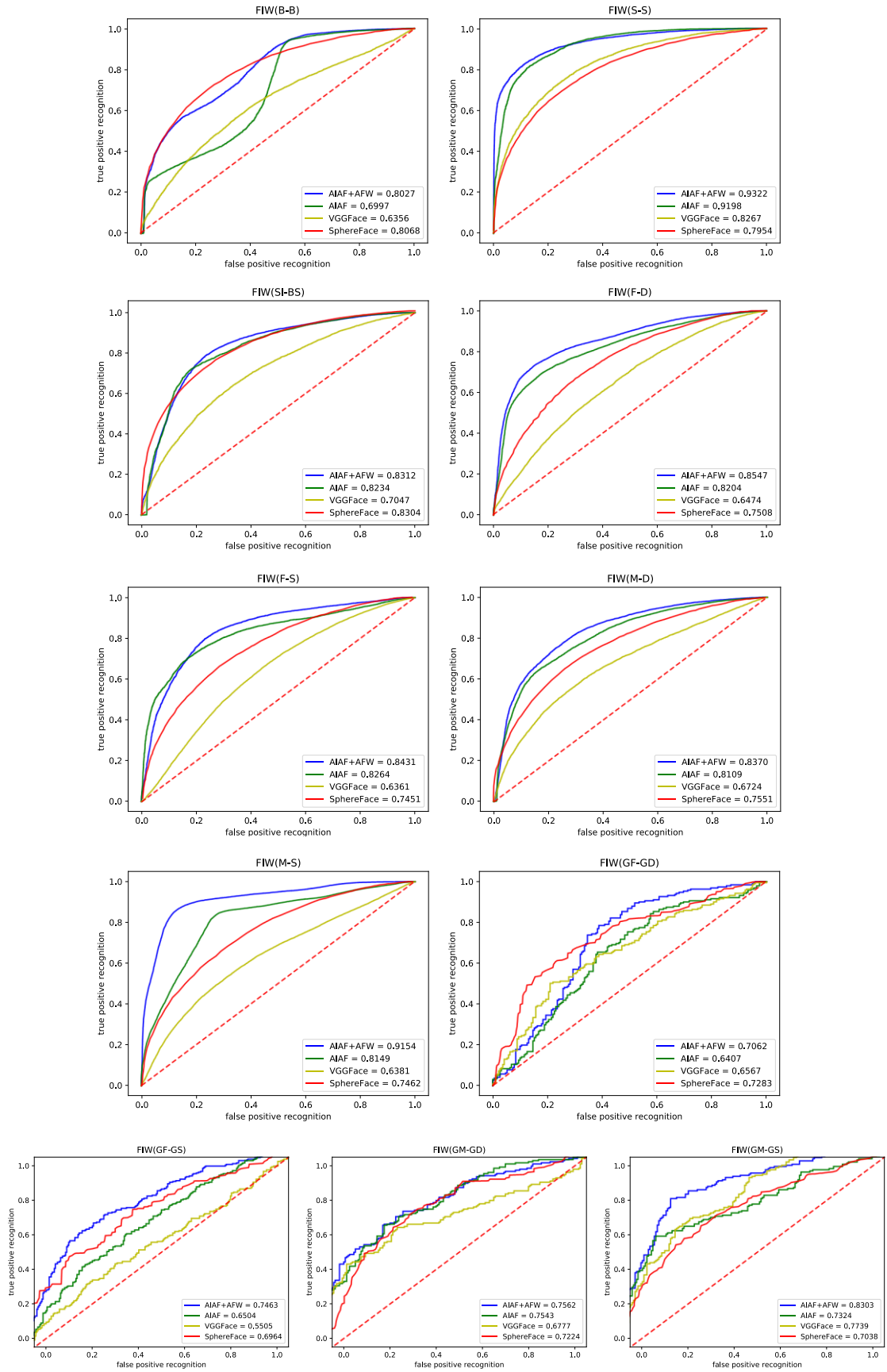


图 4.4 AIAF+AFW 和其他方法在 FIW 数据集上 11 种亲属关系验证的 ROC 曲线对比图

4.6 消融实验

如第 3.3 节所述，AFW 模块可以根据重要性为特征分配不同的权重。在本节中，我们将展示 AFW 的有效性。图 4.5 展示了人脸身份特征在通过 AFW 转换前后的部分权重，其中亮度代表特征的重要性。从图中可以看出，原始特征有很多较为模糊的部分（灰色区域），在经过 AFW 模块之后，部分特征得到了增强，另外一些被削弱。总的来说，经过计算，我们发现更多的特征被抑制了，这证实了我们在第 3.3 节中所讨论的，即一些面部特征在人脸识别中是重要的，但在亲属关系识别中并不是非常重要。

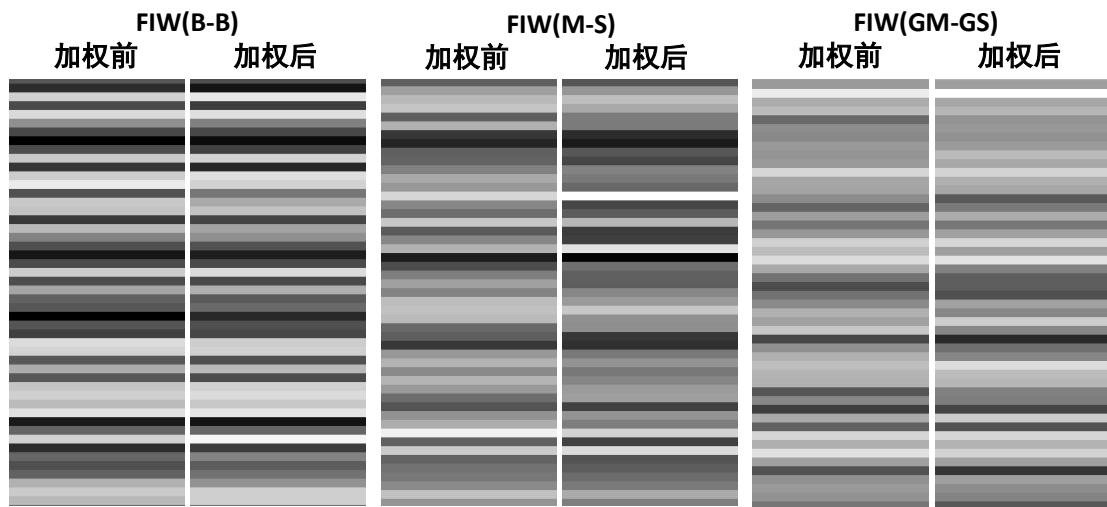


图 4.5 使用 AFW 模块前后人脸身份特征重要性的样例对比图（亮度表示特征的重要性）

4.7 本章小结

本章主要对本文提出的算法和其他一些最先进的算法在 KinFaceW-I 和 FIW 两个常用亲属关系数据集上进行了对比试验。实验结果表明，本文方法在小规模数据集上的优势不大，但总体来说还是取得了不错的效果。在大规模数据集上，本文方法取得了显著的进步，平均准确率比第二名高出 4.39%。从 ROC 和 AUC 也可看出我们方法的优势。此外，消融实验进一步证明了 AFW 模块的有效性。

第 5 章 总结与展望

5.1 本文工作总结

本文以机器学习相关理论为基础,研究了目前主流的一些亲属关系验证方法,包括传统的方法和基于深度学习的方法。针对这些方法中存在的问题,本文提出了一种基于年龄不变对抗特征学习的亲属关系验证方法,实验证明,本文方法在多个常用数据集上都达到了目前最先进的结果。具体地,本文工作主要包括以下几个方面:

(1) 本文首先对亲属关系验证问题的研究背景进行了深入的研究,并讨论了该问题的社会价值和科研意义。通过分析国内外在该领域的最新研究成果,本文指出了当前方法中存在的一些问题:1) 双方照片年龄差异大导致年龄相近的人更容易被判断为是亲属;2) 已有方法默认假设了面部特征都具有相同权重,这导致算法对重要的特征和不重要的特征一视同仁,违背直觉。这些问题都使得亲属关系验证的准确率不高。

(2) 本文对深度学习相关理论进行了介绍,包括其发展历史、与传统方法相比的优势,及其常用的激活函数、损失函数和优化器。此外,本文对我们提出的方法中使用到的一些技术也进行了简要介绍,包括卷积神经网络、对抗学习、多任务学习和注意力机制。

(3) 本文的重点是提出了一种基于年龄不变的对抗特征学习方法,该方法包括两个模块:年龄不变的对抗特征学习模块(AIAF)和基于注意力机制的特征加权模块(AFW)。具体来讲,通过 AIAF 对身份分类、年龄分类和对抗典型相关正则项使用多任务学习技术进行了联合优化,保证了身份特征与年龄特征的不相关性。然后,AFW 根据每个特征在亲属关系验证问题中的重要性为它们赋予不同的权重,越重要的特征会获得越大的权重。

(4) 通过大量实验表明,本文提出的方法优于其他最先进的方法。需要强调的是,我们提出的 AIAF 与 AFW 两个模块可以分别使用,具有松耦合性。AIAF 可以作为其他方法消除年龄因素影响的人脸特征描述的基础;AFW 可以作为其他方法在提取出人脸特征之后的特征权重赋值模块,具有极强的可拓展性。此外,本文方法并不局限于亲属关系验证问题,可以推广到亲属识别等其他相关任务中。

5.2 未来工作展望

本文针对亲属关系验证问题提出了一种基于年龄不变的对抗特征学习方法，实验结果表明，该方法与其他最先进方法相比平均准确率和 AUC 指标具有显著提升。尽管本文提出的方法取得了一定的进步，但部分工作仍有改进空间，今后我们将从以下几个方面对算法进行改进：

(1) 在实验中我们发现，本文方法在大规模数据集上远远优于在小规模数据集上的表现，在某些类别中甚至不如传统的机器学习方法。今后我们将研究在小样本情况下提升深度学习方法的能力。

(2) 理论上，本文方法中的 AIAF 与 AFW 两个模块可以与其他算法相结合，提高亲属关系验证的准确率，但是只是理论上的可能性，并未通过实验进行验证。今后我们将通过实验测试本文方法与其他方法相结合的效果。

(3) 本文方法目前在亲属关系验证任务上进行了测试，得到了不错的结果，我们认为在亲属关系识别等其他亲属相关问题上也能够取得较好的结果，所以今后我们将进一步探索该方法的可能性，促进亲属关系相关任务的研究和发展。

致 谢

父兮生我，母兮鞠我。抚我畜我，长我育我，顾我复我，出入腹我。养育之恩，无以为报！感谢父母在我求学期间的鼓励、帮助，在我人生转折点时竭尽所能给予我建议。在我放弃保研时，感谢你们百分之百地支持我所做的任何决定。

春满江山绿满园，桃李争春露笑颜。东西南北春常在，唯有师恩留心间。感谢***老师、***老师从大一下学期开始就指导我参加各种科技竞赛、撰写学术论文、专利，让我对人工智能产生了浓厚兴趣。千里马常有，而伯乐不常有。在***老师的帮助下，距本科毕业典礼不到一个月时，我终于实现了向顶会投稿的愿望。无论结果如何，终将是美好的回忆。此外，感谢***老师在我撰写毕业论文期间的悉心指导与帮助。

亦师亦友亦比邻。感谢***老师作为辅导员在这四年间的陪伴，您既像老师一样给我过来人的建议，又会像朋友一样倾听我的想法，从学习、生活等各方面都给予了我无微不至的关照，至今我还记得几年前在我遇到困难向您求助时的彻夜长谈，在此送给您一面云锦旗“最帅辅导员”。

没有人生来就是英雄，总有人用平凡成就伟大。2020 是不平凡的一年，新冠肺炎肆虐全球，牺牲了众多无辜的生命。病毒无情人有情，在此致敬白衣天使，致敬每一位奋战在一线的逆行者，是你们用生命保障了我们的安全。愿疫情早日结束，待春暖花开之时，你我无需口罩，共赴美好。

悄悄的我走了，正如我悄悄的来；我挥一挥衣袖，不带走一片云彩。

参考文献

- [1] Xia S, Shao M, Fu Y. Kinship verification through transfer learning[C]//Twenty-Second International Joint Conference on Artificial Intelligence. 2011.
- [2] Lu J, Zhou X, Tan Y P, et al. Neighborhood repulsed metric learning for kinship verification[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 36(2): 331-345.
- [3] Hu J, Lu J, Yuan J, et al. Large margin multi-metric learning for face and kinship verification in the wild[C]//Asian conference on computer vision. Springer, Cham, 2014: 252-267.
- [4] Lu J, Hu J, Liong V E, et al. The fg 2015 kinship verification in the wild evaluation[C]//2015 11th IEEE international conference and workshops on automatic face and gesture recognition (FG). IEEE, 2015, 1: 1-7.
- [5] Duan X, Tan Z H. A feature subtraction method for image based kinship verification under uncontrolled environments[C]//2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015: 1573-1577.
- [6] Kou L, Zhou X, Xu M, et al. Learning a genetic measure for kinship verification using facial images[J]. Mathematical Problems in Engineering, 2015.
- [7] Zhou X, Shang Y, Yan H, et al. Ensemble similarity learning for kinship verification from facial images in the wild[J]. Information Fusion, 2016, 32: 40-48.
- [8] Robinson J P, Shao M, Wu Y, et al. Families in the Wild (FIW) Large-Scale Kinship Image Database and Benchmarks[C]//Proceedings of the 24th ACM international conference on Multimedia. 2016: 242-246.
- [9] Zhou X, Yan H, Shang Y. Kinship verification from facial images by scalable similarity fusion[J]. Neurocomputing, 2016, 197: 136-142.
- [10] Wang S, Robinson J P, Fu Y. Kinship verification on families in the wild with marginalized denoising metric learning[C]//2017 12th IEEE International

- Conference on Automatic Face & Gesture Recognition (FG 2017). IEEE, 2017: 216-221.
- [11] Wang S, Ding Z, Fu Y. Cross-generation kinship verification with sparse discriminative metric[J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 41(11): 2783-2790.
- [12] Robinson J P, Shao M, Wu Y, et al. Visual kinship recognition of families in the wild[J]. IEEE Transactions on pattern analysis and machine intelligence, 2018, 40(11): 2624-2637.
- [13] Nandy A, Mondal S S. Kinship Verification using Deep Siamese Convolutional Neural Network[C]//2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019). IEEE, 2019: 1-5.
- [14] Yan H, Wang S. Learning part-aware attention networks for kinship verification[J]. Pattern Recognition Letters, 2019, 128: 169-175.
- [15] Jain A, Bhagat N, Srivastava V, et al. A Feature-Based Kinship Verification Technique Using Convolutional Neural Network[M]//Advances in Data Sciences, Security and Applications. Springer, Singapore, 2020: 353-362.
- [16] Chergui A, Ouchtati S, Mavromatis S, et al. Investigating Deep CNNs Models Applied in Kinship Verification through Facial Images[C]//2019 5th International Conference on Frontiers of Signal Processing (ICFSP). IEEE, 2019: 82-87.
- [17] Liu W, Wen Y, Yu Z, et al. Sphreface: Deep hypersphere embedding for face recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 212-220.
- [18] Fang R, Tang K D, Snavely N, et al. Towards computational models of kinship verification[C]//2010 IEEE International conference on image processing. IEEE, 2010: 1577-1580.
- [19] Zhou X, Hu J, Lu J, et al. Kinship verification from facial images under uncontrolled conditions[C]//Proceedings of the 19th ACM international conference on Multimedia. 2011: 953-956.
- [20] Guo G, Wang X. Kinship measurement on salient facial features[J]. IEEE

- Transactions on Instrumentation and Measurement, 2012, 61(8): 2322-2325.
- [21] Bottinok A, Islam I U, Vieira T F. A multi-perspective holistic approach to kinship verification in the wild[C]//2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). IEEE, 2015, 2: 1-6.
- [22] Bottino A, Vieira T F, Ul Islam I. Geometric and textural cues for automatic kinship verification[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2015, 29(03): 1556001.
- [23] Dornaika F, Bosaghzadeh A. Exponential local discriminant embedding and its application to face recognition[J]. IEEE transactions on cybernetics, 2013, 43(3): 921-934.
- [24] Kan M, Xu D, Shan S, et al. Learning prototype hyperplanes for face verification in the wild[J]. IEEE transactions on image processing, 2013, 22(8): 3310-3316.
- [25] Yan H, Lu J, Zhou X. Prototype-based discriminative feature learning for kinship verification[J]. IEEE Transactions on cybernetics, 2014, 45(11): 2535-2545.
- [26] Somanath G, Kambhamettu C. Can faces verify blood-relations?[C]//2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS). IEEE, 2012: 105-112.
- [27] Yan H, Lu J, Deng W, et al. Discriminative multimetric learning for kinship verification[J]. IEEE Transactions on Information forensics and security, 2014, 9(7): 1169-1178.
- [28] Xu M, Shang Y. Kinship measurement on face images by structured similarity fusion[J]. IEEE Access, 2016, 4: 10280-10287.
- [29] Qin X, Tan X, Chen S. Tri-subject kinship verification: Understanding the core of a family[J]. IEEE Transactions on Multimedia, 2015, 17(10): 1855-1867.
- [30] Zhou X, Jin K, Xu M, et al. Learning deep compact similarity metric for kinship verification from face images[J]. Information Fusion, 2019, 48: 84-94.

- [31] Xia S, Shao M, Luo J, et al. Understanding kin relationships in a photo[J]. IEEE Transactions on Multimedia, 2012, 14(4): 1046-1056.
- [32] Zhao Y G, Song Z, Zheng F, et al. Learning a multiple kernel similarity metric for kinship verification[J]. Information Sciences, 2018, 430: 247-260.
- [33] Wei Z, Xu M, Geng L, et al. Adversarial Similarity Metric Learning for Kinship Verification[J]. IEEE Access, 2019, 7: 100029-100035.
- [34] Dehghan A, Ortiz E G, Villegas R, et al. Who do i look like? determining parent-offspring resemblance via gated autoencoders[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 1757-1764.
- [35] Zhang¹² K, Huang Y, Song C, et al. Kinship verification with deep convolutional neural networks[J]. 2015.
- [36] Ding Z, Suh S, Han J J, et al. Discriminative low-rank metric learning for face recognition[C]//2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). IEEE, 2015, 1: 1-6.
- [37] Van T N, Hoang V T. Early and late features fusion for kinship verification based on constraint selection[C]//2019 25th Asia-Pacific Conference on Communications (APCC). IEEE, 2019: 116-121.
- [38] Dehshibi M M, Shanbehzadeh J. Cubic norm and kernel-based bi-directional PCA: toward age-aware facial kinship verification[J]. The Visual Computer, 2019, 35(1): 23-40.
- [39] Liu Y, Li H, Wang X. Rethinking feature discrimination and polymerization for large-scale recognition[J]. arXiv preprint arXiv:1710.00870, 2017.
- [40] Ranjan R, Castillo C D, Chellappa R. L2-constrained softmax loss for discriminative face verification[J]. arXiv preprint arXiv:1703.09507, 2017.
- [41] Wang F, Cheng J, Liu W, et al. Additive margin softmax for face verification[J]. IEEE Signal Processing Letters, 2018, 25(7): 926-930.
- [42] Wang H, Wang Y, Zhou Z, et al. Cosface: Large margin cosine loss for deep face recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 5265-5274.
- [43] Deng J, Guo J, Xue N, et al. Arcface: Additive angular margin loss for deep

- face recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019: 4690-4699.
- [44] Wang H, Gong D, Li Z, et al. Decorrelated adversarial learning for age-invariant face recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019: 3527-3536.
- [45] Rothe R, Timofte R, Van Gool L. Dex: Deep expectation of apparent age from a single image[C]//Proceedings of the IEEE international conference on computer vision workshops. 2015: 10-15.
- [46] Rothe R, Timofte R, Van Gool L. Deep expectation of real and apparent age from a single image without facial landmarks[J]. International Journal of Computer Vision, 2018, 126(2-4): 144-157.
- [47] Zhang H, Cisse M, Dauphin Y N, et al. mixup: Beyond Empirical Risk Minimization[J]. 2017.
- [48] Zhong Z, Zheng L, Kang G, et al. Random erasing data augmentation[J]. arXiv preprint arXiv:1708.04896, 2017.
- [49] Cao Q, Shen L, Xie W, et al. Vggface2: A dataset for recognising faces across pose and age[C]//2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). IEEE, 2018: 67-74.
- [50] Chen B C, Chen C S, Hsu W H. Cross-age reference coding for age-invariant face recognition and retrieval[C]//European conference on computer vision. Springer, Cham, 2014: 768-783.
- [51] Zhang K, Zhang Z, Li Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-1503.
- [52] Wang Y, Gong D, Zhou Z, et al. Orthogonal deep features decomposition for age-invariant face recognition[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 738-753.
- [53] Lu J, Hu J, Zhou X, et al. Neighborhood Repulsed Metric Learning for Kinship Verification[C]//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

- [54] E. Tola, V. Lepetit and P. Fua, Daisy: An efficient dense descriptor applied to wide-baseline stereo. In IEEE Trans. Pattern Anal. Mach. Intell. vol. 32, no. 5, pp. 815-830, May 2010.
- [55] Tola E, Lepetit V, Fua P. Daisy: An efficient dense descriptor applied to wide-baseline stereo[J]. IEEE transactions on pattern analysis and machine intelligence, 2009, 32(5): 815-830.
- [56] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). IEEE, 2005, 1: 886-893.
- [57] Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition[J]. 2015.
- [58] Davis J V, Kulis B, Jain P, et al. Information-theoretic metric learning[C]//Proceedings of the 24th international conference on Machine learning. 2007: 209-216.
- [59] He X, Niyogi P. Locality preserving projections[C]//Advances in neural information processing systems. 2004: 153-160.
- [60] Weinberger K Q, Saul L K. Distance metric learning for large margin nearest neighbor classification[J]. Journal of Machine Learning Research, 2009, 10(Feb): 207-244.
- [61] Peng Y, Wang S, Lu B L. Marginalized denoising autoencoder via graph regularization for domain adaptation[C]//International Conference on Neural Information Processing. Springer, Berlin, Heidelberg, 2013: 156-163.
- [62] Wen Y, Zhang K, Li Z, et al. A discriminative feature learning approach for deep face recognition[C]//European conference on computer vision. Springer, Cham, 2016: 499-515.
- [63] He X, Cai D, Yan S, et al. Neighborhood preserving embedding[C]//Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. IEEE, 2005, 2: 1208-1213.
- [64] Hotelling H. Relations between two sets of variates[M]//Breakthroughs in statistics. Springer, New York, NY, 1992: 162-190.
- [65] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial

- nets[C]//Advances in neural information processing systems. 2014: 2672-2680.
- [66] http://www.jiangsu.gov.cn/art/2014/4/11/art_46682_2585113.html 访问日期: 2020.5.20
- [67] Lindeberg T. Scale invariant feature transform[J]. 2012.
- [68] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). IEEE, 2005, 1: 886-893.
- [69] Ahonen T, Hadid A, Pietikainen M. Face description with local binary patterns: Application to face recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2006, 28(12): 2037-2041.
- [70] Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of physiology, 1962, 160(1): 106-154.
- [71] Hawkins D M. The problem of overfitting[J]. Journal of chemical information and computer sciences, 2004, 44(1): 1-12.
- [72] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological cybernetics, 1980, 36(4): 193-202.
- [73] https://en.wikipedia.org/wiki/Nash_equilibrium 访问日期: 2020.5.20
- [74] 秦晓倩. 基于 Web 图像的 Kinship 关系验证研究[D]. 南京航空航天大学, 2016.
- [75] 张艳丽. 基于图像的亲属识别关键问题研究[D]. 北京理工大学, 2016.
- [76] 袁琴. 基于图像的人物亲属关系分析与研究[D]. 基于图像的人物亲属关系分析与研究, 2019.
- [77] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. The bulletin of mathematical biophysics, 1943, 5(4): 115-133.
- [78] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. Psychological review, 1958, 65(6): 386.
- [79] Van Der Malsburg C. Frank Rosenblatt: principles of neurodynamics: perceptrons and the theory of brain mechanisms[M]//Brain theory. Springer,

- Berlin, Heidelberg, 1986: 245-248.
- [80] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. nature, 1986, 323(6088): 533-536.
- [81] Waibel A, Hanazawa T, Hinton G, et al. Phoneme recognition using time-delay neural networks[J]. IEEE transactions on acoustics, speech, and signal processing, 1989, 37(3): 328-339.
- [82] Zhang W. Shift-invariant pattern recognition neural network and its optical architecture[C]//Proceedings of annual conference of the Japan Society of Applied Physics. 1988.
- [83] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural computation, 1989, 1(4): 541-551.
- [84] Aihara K, Takabe T, Toyoda M. Chaotic neural networks[J]. Physics letters A, 1990, 144(6-7): 333-340.
- [85] Specht D F. A general regression neural network[J]. IEEE transactions on neural networks, 1991, 2(6): 568-576.
- [86] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [87] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.
- [88] Qian N. On the momentum term in gradient descent learning algorithms[J]. Neural networks, 1999, 12(1): 145-151.
- [89] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [90] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [91] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

- [92] Nesterov Y. A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$ [C]//Doklady an ussr. 1983, 269: 543-547.
- [93] Duchi J, Hazan E, Singer Y. Adaptive subgradient methods for online learning and stochastic optimization[J]. Journal of machine learning research, 2011, 12(Jul): 2121-2159.
- [94] Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 815-823.
- [95] Hinton G, Srivastava N, Swersky K. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent[J]. Cited on, 2012, 14(8).
- [96] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [97] Li Z, Yang W, Peng S, et al. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects[J]. arXiv preprint arXiv:2004.02806, 2020.
- [98] Khan A, Sohail A, Zahoor U, et al. A survey of the recent architectures of deep convolutional neural networks[J]. arXiv preprint arXiv:1901.06032, 2019.
- [99] Mnih V, Heess N, Graves A. Recurrent models of visual attention[C]//Advances in neural information processing systems. 2014: 2204-2212.
- [100] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv:1409.0473, 2014.
- [101] Xu K, Ba J, Kiros R, et al. Show, attend and tell: Neural image caption generation with visual attention[C]//International conference on machine learning. 2015: 2048-2057.
- [102] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in neural information processing systems. 2017: 5998-6008.

- [103] 邹秀芳, 朱定局. 生成对抗网络研究综述[J]. 计算机系统应用, 2019, 28(11): 1-9. <http://www.c-s-a.org.cn/1003-3254/7156.html>
- [104] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
- [105] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.
- [106] Karras T, Aila T, Laine S, et al. Progressive growing of gans for improved quality, stability, and variation[J]. arXiv preprint arXiv:1710.10196, 2017.
- [107] Hjelm R D, Jacob A P, Che T, et al. Boundary-seeking generative adversarial networks[J]. arXiv preprint arXiv:1702.08431, 2017.

附 录

A: 个人简介

***, 男, 中共党员, 河海大学信息学部计算机与信息学院计算机科学与技术专业 2016 级本科生, 曾担任计信院学生科协主席。曾获国家奖学金、严恺一等奖学金等奖学金共计 20 项, 江苏省“三好学生”、优秀学生标兵、优秀志愿者、优秀毕业生、科普工作先进个人等荣誉称号 20 项。在国内外核心期刊会议上已发表论文 4 篇(其中一篇 CCF B 类核心期刊第一作者)、在审论文 5 篇, 曾在 IEEE PIC 2018 国际会议上代表团队进行报告。获实用新型专利授权 1 项(第一发明人)、国家发明专利受理 8 项。曾获科技竞赛国家级奖项 2 项、省级奖项 4 项、校院级奖项 11 项, 主持国家级大学生创新训练项目 1 项, 优秀结题。

B: 本科期间撰写的学术论文

已发表的论文

- [J-1] 基于残差生成对抗网络的人脸图像复原[J]. 计算机科学. 2020. (CCF B 类), 第一作者.
- [J-2] 节水潜力预测研究综述[J]. 水利经济. 2018. (科技核心), 第三作者, 导师一作.
- [C-1] A Text-based CAPTCHA Cracking System with Generative Adversarial Networks[C]. IEEE International Symposium on Multimedia (ISM). 2018. (EI 检索), 第二作者, 导师一作.
- [C-2] An embedded approach for the CNN based real-time hard hat detection[C]. IEEE International Conference on Progress in Informatics and Computing (PIC). 2018. (EI 检索), 第三作者, 导师一作.

投稿在审的论文

- [J-1] A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects[J]. IEEE Transactions on Neural Networks and Learning Systems (TNNLS). 2020. (SCI 1 区, CCF B 类)
- 已预印至: <https://arxiv.org/abs/2004.02806>, 第一作者.

- [C-1] Age-Invariant Adversarial Feature Learning for Kinship Verification[C]. ACM Multimedia (MM). 2020. (CCF A 类, 多媒体领域顶级会议), 第一作者.
- [C-2] Deep Learning Based Single Sample Per Person Face Recognition: A Survey[C]. The 25th International Conference on Pattern Recognition (ICPR). 2020. (CCF C 类), 第三作者.
- [C-3] Supervised Auto-Encoder Based Sparse Representation for One-shot Face Recognition[C]. The 3rd Chinese Conference on Pattern Recognition and Computer Vision (PRCV). 2020. (EI), 第三作者.
- [C-4] A Review of Automatically Diagnosing COVID-19 Based on Scanning Image[C]. The 3rd Chinese Conference on Pattern Recognition and Computer Vision (PRCV). 2020. (EI), 第三作者.

C: 本科期间参加的科研项目

- 1、2018 年国家级创新训练计划:《AI 画质+:基于深度学习的视频增强技术及其应用》, 第一负责人, 项目验收结果为优秀.
- 2、国家自然科学基金委员会, 青年项目, 61602150, 视觉感知启发的单样本人脸特征表示与识别方法研究.
- 3、江苏省自然科学基金, 面上项目, BK20191298, 基于语义的单样本人脸识别关键技术研究.

D: 本科期间申请的专利

- [P-1] 一种智能旅行箱[P]. 中国实用新型专利. CN207707429U. 2018-08-10. (已授权), 第一作者.
- [P-2] 一种基于步态识别算法的智能跟随旅行箱 [P]. 中国发明专利. CN108813863A. 2018-11-16. (实审), 第一作者.
- [P-3] 一种基于残差生成对抗网络的人脸视频增强方法[P]. 中国发明专利. CN110276728A. 2019-09-24. (实审), 第一作者.
- [P-4] 一种智能电动旅行箱[P]. 中国发明专利. CN108851427A. 2018-11-23. (实审), 第一作者.

- [P-5] 一种基于电容式指纹解锁的智能电动旅行箱[P]. 中国发明专利. CN108937027A. 2018-12-07. (实审), 第二作者.
- [P-6] 一种基于边缘计算的智能安全帽佩戴监测系统[P]. 中国发明专利. CN109361902A. 2019-02-19. (实审), 第二作者.
- [P-7] 一种基于双目视觉的自动避障旅行箱[P]. 中国发明专利. CN108937026A. 2018-12-07. (实审), 第三作者.
- [P-8] 一种基于电动旅行箱的移动发电装置[P]. 中国发明专利. CN108879909A. 2018-11-23. (实审), 第四作者.
- [P-9] 一种基于 RFID 与无线通信技术的低功耗定位防丢旅行箱[P]. 中国发明专利. CN109820311A. 2019-05-31. (实审), 第五作者.

E: 本科期间所获竞赛奖项

- 1、2018 年第九届中国大学生服务外包创新创业大赛”国家二等奖, 第一负责人。
(中华人民共和国教育部、中华人民共和国商务部、无锡市人民政府)
- 2、2018 年第十二届 iCAN 国际创新创业大赛中国总决赛三等奖, 第一负责人。
(国际 iCAN 联盟、教育部创新方法教学指导分委员会、全球华人微纳米分子系统学会、北京大学)
- 3、2018 年全国大学生数学建模竞赛省赛二等奖, 队长。(中国工业与应用数学学会)
- 4、2018 年第十二届 iCAN 国际创新创业大赛江苏省赛二等奖, 第一负责人。(国际 iCAN 联盟、教育部创新方法教学指导分委员会、全球华人微纳米分子系统学会、北京大学)
- 5、2018 年“第九届蓝桥杯全国软件和信息技术专业人才大赛”江苏赛区个人二等奖(中国软件行业协会、中国电子商会、中国电子学会、中国半导体行业协会、工业和信息化部人才交流中心、蓝桥杯全国软件和信息技术专业人才大赛组织委员会)；
- 6、2017 年“第八届蓝桥杯全国软件和信息技术专业人才大赛”江苏赛区个人三等奖(中国软件行业协会、中国电子商会、中国电子学会、中国半导体行业协会、工业和信息化部人才交流中心、蓝桥杯全国软件和信息技术专业人才大赛组织委员会)；

- 7、2017 年河海大学创新创业大赛暨“建行杯”江苏省“互联网+”大学生创新创业大赛校内选拔赛初创组三等奖，第三负责人。（河海大学学生工作处）；
- 8、2018 年第四届中国“互联网+”大学生创新创业大赛校内选拔赛三等奖，第二负责人（河海大学学生工作处）；
- 9、“2018 溧智创业营”优秀奖，第一负责人。（江苏（溧水）人工智能创新创业大赛组委会）；
- 10、2016 年河海大学“闻道杯”辩论赛团体冠军（河海辩论联盟）
- 11、2016 河海大学计信院新生杯辩论赛“最佳辩手”；（计算机与信息学院科协）
- 12、2016 河海大学计信院新生杯辩论赛“最具风度辩手”；（计算机与信息学院科协）
- 13、2017 河海大学第二十九届校园科技节第七届 DIY 电脑装机大赛个人三等奖（河海大学党委）
- 14、2017 河海大学在线程序设计竞赛奋进奖（河海大学计算机与信息学院）
- 15、2016 河海大学第二届“HTML 网页制作大赛”优秀奖（河海大学计信院团委）
- 16、2019 第八届在宁高校计算机论坛“优秀团队”（中共河海大学委员会）
- 17、2019 第十六届“挑战杯”全国大学生课外学术科技作品竞赛校内选拔赛二等奖（河海大学）

F：本科期间所获荣誉

荣誉称号

- 1、2019 江苏省“三好学生”（江苏省教育厅、共青团江苏省委员会）
- 2、2016 年度暑期实践“先进个人”（中共河海大学委员会）
- 3、2016 年度《“智慧水务——计算机技术”在水资源处理与规划中的应用》“优秀调研报告”（中共河海大学委员会）
- 4、2016-2017 学年河海大学“优秀学生标兵”（河海大学）
- 5、2016-2017 学年河海大学“优秀学生干部”（河海大学）
- 6、2016-2017 学年河海大学“优秀学生”（河海大学）
- 7、2017-2018 学年河海大学“优秀学生标兵”（河海大学）
- 8、2017-2018 学年河海大学“优秀学生干部”（河海大学）

- 9、2017-2018 学年河海大学“优秀学生”（河海大学）
- 10、2018-2019 学年河海大学“优秀学生标兵”（河海大学）
- 11、2018-2019 学年河海大学“优秀学生干部”（河海大学）
- 12、2018-2019 学年河海大学“优秀学生”（河海大学）
- 13、2019-2020 学年河海大学“优秀学生标兵”（河海大学）
- 14、2019-2020 学年河海大学“优秀学生干部”（河海大学）
- 15、2019-2020 学年河海大学“优秀学生”（河海大学）
- 16、2017-2018 学年河海大学“优秀志愿者”（河海大学学生工作处）
- 17、2017 年度河海大学科普工作“先进个人”（河海大学科学技术协会）
- 18、2017 年河海大学计信院“十佳歌手”（河海大学计算机与信息学院）
- 19、河海大学 2020 届本科优秀毕业生（河海大学）
- 20、2016 级“1442 工程”优秀学员（河海大学）

奖学金

- 1、2017-2018 年国家奖学金（中华人民共和国教育部）
- 2、2016-2017 年国家励志奖学金（中华人民共和国教育部）
- 3、2019 年度严恺奖学金一等奖（严恺教育科技基金评奖委员会）
- 4、河海大学 2016-2017 学年学业优秀奖学金（河海大学学生处）
- 5、河海大学 2016-2017 学年科技创新奖学金（河海大学学生处）
- 6、河海大学 2016-2017 学年社会工作奖学金（河海大学学生处）
- 7、河海大学 2016-2017 学年精神文明奖学金（河海大学学生处）
- 8、河海大学 2016-2017 学年艺术体育优秀奖学金（河海大学学生处）
- 9、河海大学 2017-2018 学年学业优秀奖学金（河海大学学生处）
- 10、河海大学 2017-2018 学年科技创新奖学金（河海大学学生处）
- 11、河海大学 2017-2018 学年社会工作奖学金（河海大学学生处）
- 12、河海大学 2017-2018 学年学业进步奖学金（河海大学学生处）
- 13、河海大学 2018-2019 学年学业优秀奖学金（河海大学学生处）
- 14、河海大学 2018-2019 学年科技创新奖学金（河海大学学生处）
- 15、河海大学 2018-2019 学年社会工作奖学金（河海大学学生处）
- 16、河海大学 2018-2019 学年艺术体育优秀奖学金（河海大学学生处）

- 17、河海大学 2019-2020 学年学业优秀奖学金（河海大学学生处）
- 18、河海大学 2019-2020 学年科技创新奖学金（河海大学学生处）
- 19、河海大学 2019-2020 学年社会工作奖学金（河海大学学生处）
- 20、河海大学 2019-2020 学年精神文明奖学金（河海大学学生处）

G：本科期间社会工作

- 2017 年至 2020 年 6 月，河海大学“1442 工程”学员；
- 2016 年 9 月至 2020 年 6 月，计信院 16 级计算机 3 班副班长；
- 2016 年 9 月至 2017 年 9 月，计信院学生科协技术部干事；
- 2017 年 9 月至 2018 年 9 月，计信院学生科协技术部部长，考核为优秀；大学生职业发展联盟副队长；
- 2018 年 9 月至 2019 年 9 月，计信院学生科协主席。