

参赛队员姓名： 蒋昕昀

中学： 杭州外国语学校

省份： 浙江省

国家/地区： 中国

指导教师姓名： 钱运涛

指导教师单位： 浙江大学计算机科学与技术学院

论文题目： 基于多步特征分解的跨年龄人脸身份识别深度神经网络方法

2020 S.-T. Yau High School Science Award

本参赛团队声明所提交的论文是在指导老师指导下进行的研究工作和取得的研究成果。尽本团队所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果。若有不实之处，本人愿意承担一切相关责任。

参赛队员：

蒋昕雨

指导老师：

魏江涛

2020年9月6日

基于多步特征分解的

跨年龄人脸身份识别深度神经网络方法

A Deep Neural Network Approach for Cross-Age Face
Recognition Based on Multi-Step Feature Decomposition

蒋昕昀

xinyunsunshine@126.com

杭州外国语学校

2020 S.-T. Yau High School Science Award

摘要

计算机人脸自动身份识别已经在各个领域都得到了广泛应用。在部分人脸识别应用场景中，同一人在人脸库中的图片和需要识别的图片是在不同年龄下拍摄的，它们之间差异很大，这就提出了一个更复杂和更具挑战性的跨年龄人脸身份识别问题。跨年龄人脸识别的关键是对人脸特征进行分解，把和身份相关的特征从整个特征中分离出来，特别是把年龄相关的特征和身份相关特征进行比较彻底的分离，这样跨年龄的人脸身份识别就会减少年龄的影响和干扰。现有的跨年龄识别模型存在特征分解不彻底、分解后的特征相互交叉、分解导致信息丢失等问题。为了实现更好的跨年龄人脸识别，本研究提出了一种人脸特征分步分解的跨年龄人脸身份识别深度神经网络模型 (SWNet)。该模型充分利用训练样本中的身份和年龄标记信息，构造身份识别和年龄估计协同多任务学习网络，同时在深度神经网络的每步特征分解中，身份特征和年龄特征都直接连接最后的损失函数进行监督学习，保证每层分离特征都直接对身份识别产生影响，降低分解带来的信息损失。在标准数据集上的实验表明，相比已有方法，SWNet 在跨年龄人脸身份识别上表现出良好的性能。

关键词：跨年龄人脸识别，特征分解，深度神经网络

Abstract

Compared with the conventional face recognition, cross-age face recognition has more extensive application scenarios and more complex technical challenges. It aims at solving the problem that the pictures of the same person in the face database and the pictures to be recognized are taken at different ages, in which the key technique is to decompose facial features into two components to represent age-related and identity-related features. Then, the cross-age face recognition is realized by using the identity-related features (features independent of age). The existing cross-age recognition models have some problems, such as incomplete feature decomposition, overlapping of the decomposed features and information loss caused by decomposition. In order to better achieve cross-age face recognition, this study proposes a deep neural network model (SWNet) for cross-age face identity recognition based on multi-step features decomposition. The model makes full use of the marked identity and age information in training samples to construct a collaborative multi-task learning network for identity recognition and age estimation. At the same time, in each step of the eigendecomposition of deep neural network, the identity and age features are directly connected with the final loss function for supervised learning, and the features of each layer are integrated with those of the previous layer to compensate for the information loss caused by the previous decomposition. Extensive experiments on CACD show the effectiveness of the proposed approach.

Key words: cross-age face recognition, feature decomposition, deep neural network

目录

1 引言	6
1.1 问题背景	6
1.2 相关工作	6
1.2.1 基于特征正交分解的跨年龄人脸识别模型 (OENet)	7
1.2.2 基于年龄估计的跨年龄人脸识别模型 (AENet)	8
2 提出的方法	9
2.1 动机	9
2.2 方法	9
2.2.1 网络结构	9
2.2.2 特征分解	10
2.2.3 Loss 函数选择	11
2.2.4 特征解耦	12
2.2.5 网络训练优化方法	13
3 实验评估	13
3.1 数据集	13
3.2 实验设计	13
3.3 实验结果	14
4 讨论	17
4.1 联系	17
4.2 区别	17
5 结论	18
参考文献	19
致谢	20
附件：参赛选手与指导老师简历	21

1 引言

1.1 问题背景

借助于深度学习技术，普通的基于人脸的身份识别和基于人脸年龄估计，在较良好的可控使用环境下，目前已经取得了很高的精度，特别是身份识别已经在考勤系统、门禁系统、远程认证、娱乐系统中得到广泛应用 [1]。相对一般条件下的人脸识别，跨年龄人脸识别目前更具挑战性，也具有更广泛的潜在应用场景 [2]。比如在寻找丢失儿童、抓捕罪犯、护照和身份证识别等特殊场景下具有重要的实际应用价值，要求比一般的人脸识别系统对年龄差异更具鲁棒性。图 1 是同一个人的不同年龄阶段的人脸图片展示，可以发现老化过程带来的同一个体外貌的明显变化，这就给跨年龄人脸识别任务带来了很大考验。

目前基于深度学习的跨年龄人脸识别在利用人脸数据库中的年龄标记信息时，往往采用单独学习人脸老化模型或单独提取年龄相关特征等方法，很少在同一深度神经网络模型中同时进行身份识别和年龄估计的学习。考虑到该人脸识别和年龄密切相关，我们希望能用一个模型同时训练跨年龄人脸识别和年龄估计两个任务。我们希望通过标注了身份和年龄的数据集联合学习跨年龄人脸识别模型和年龄估计模型，通过多任务学习机制，年龄估计模型可以有效监督人脸特征分解，在学习过程中将年龄相关特征和身份相关特征的分解更加彻底。同时充分发挥深度学习在特征分解中的深度表达能力，通过深度多步特征分解，使得到的特征同时满足特征分离和深度特征表示。

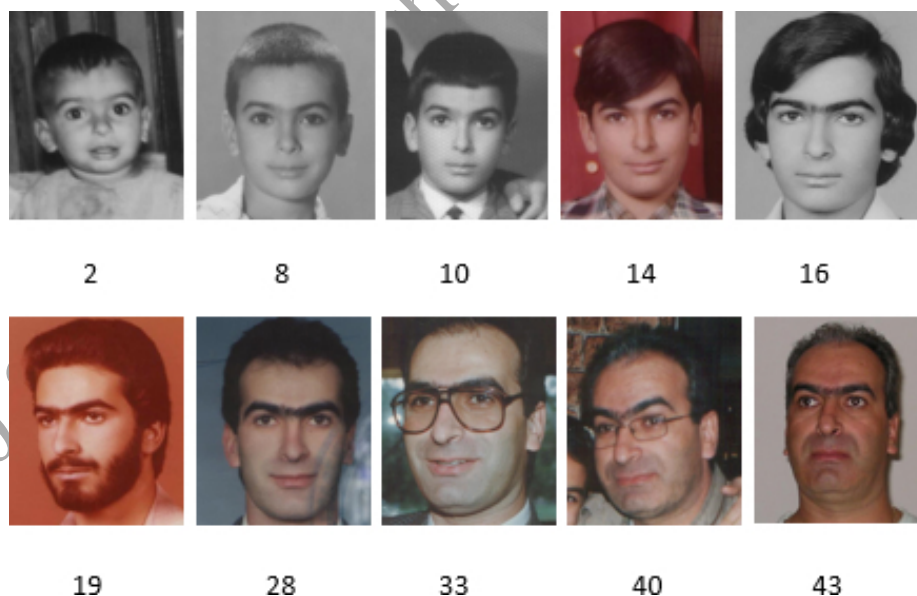


图 1 同一身份个体在不同年龄阶段的人脸图片

1.2 相关工作

现有的抗年龄干扰的人脸识别模型有两类：生成模型和判别模型。生成模型利用大量个体不同年龄的人脸图像，构建一个年龄影响下的人脸特征渐进变化模型，建模人脸

的形状和肌理变化。该模型利用人脸的形态和纹理等随时间变化的规律，利用同一人部分年龄段的人脸图像生成其他年龄段的人脸图像，这样就可以用基本相同年龄段的人脸图像进行人脸识别。判别模型是基于人脸图像中只和身份相关的特征进行身份判别，这些特征在不同年龄段都是比较稳定的，和年龄无关或者关联很小，这样进行人脸身份识别时就没有跨年龄的干扰。本论文我们采用判别模型方法。

判别模型方法主要流程包括特征提取、特征分解和特征匹配（分类），关键在于人脸特征提取和提取之后的分解。在特征提取方面，通过不断研究特征提取技术，希望能够挖掘出人脸内在的特征，从而避免光照、姿态、表情、遮挡等影响。在特征分解方面，即分解出身份和年龄特征，普遍采用基于潜在因素分析（HFA）的分解技术。随着深度学习的发展，近年来无论普通人脸识别还是跨年龄人脸识别都普遍采用深层卷积神经网络（CNN）提取人脸特征 [3]，卷积神经网络由多层次局部卷积滤波操作为主干，加入池化操作和非线性激励操作，可以有效提取图像包含的丰富内在特征信息，并且提取的深度特征对于图像中对象的平移、旋转、缩放有一定的不变性。同时借助反向传播和随机梯度等训练策略，可以实现大规模卷积网络模型的参数优化，具有很强的学习能力。基于这些优越的性能，相比全连接神经网络等其他经典神经网络模型，它的性能更加突出。本文提出模型的基本架构也是基于卷积网络。

在卷积神经网络框架下，跨年龄人脸识别的研究重点基本都集中在卷积特征的分解方法上。最早的是把 HFA 分解直接应用到 CNN 中用于分解卷积特征 [4]，之后又陆续出现了特征的残差分解、正交分解、年龄估计导向的减法分解等，目的都是为了从卷积后的人脸特征中分解出身份相关的特征（同时也分解出年龄相关的特征）用于人脸识别，使得分解出的身份特征尽可能只包含身份相关的信息以实现跨年龄的人脸匹配。本论文的主要贡献也集中在特征分解上，希望分解后的身份和年龄特征尽可能地解耦，通过构造身份识别和年龄估计两个任务的多任务学习模型，以及深度多步特征分解等创新方法，达到更好的跨年龄人脸身份识别效果。

下面我们将介绍两个和本文方法密切相关的工作，而且这两种方法在后续实验中将和我们提出的方法进行对比，用以说明提出方法的优点。

1.2.1 基于特征正交分解的跨年龄人脸识别模型（OENet）

OENet 是 2018 年在 ECCV 会议上提出的一种典型基于深度神经网络的跨年龄人脸身份识别的判别模型 [5]，如图 2 所示。OENet 的主干网络是一个类似 ResNet [6] 的多层卷积神经网络，之后添加一层全连接层做卷积特征的分解，分解的结果是表示年龄的特征向量和表示身份的特征向量。在分解过程中，对于年龄和身份这两类特征施加了正交约束，使得它们尽可能的正交去相关。年龄特征采用均方差的回归 Loss 函数监督，身份特征采用多分类的 Loss 函数，两者相加并用一个权重系数控制年龄 Loss 的比重，以实现用多任务学习的方式进行模型训练。

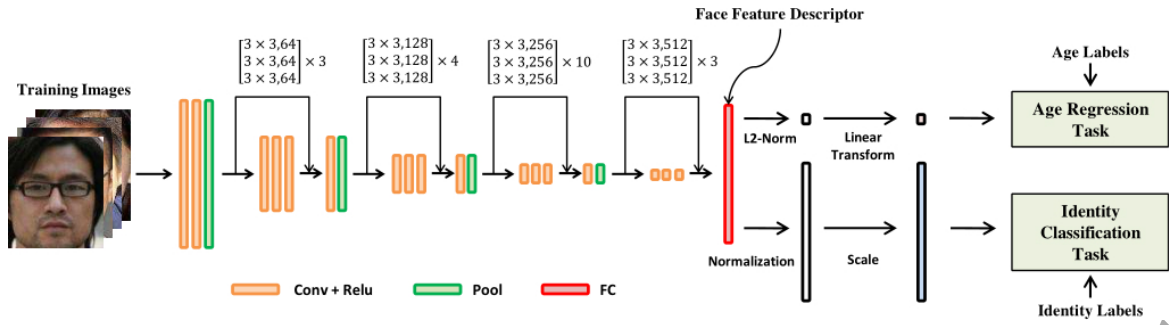


图 2 OENet: 基于正交特征分解的跨年龄人脸识别网络架构

1.2.2 基于年龄估计的跨年龄人脸识别模型 (AENet)

AENet 是 2017 年在 CVPR 上提出的另一种典型基于深度神经网络的跨年龄人脸身份识别的判别模型 [7], 典型如图 3 所示。AENet 是年龄估计驱动的跨年龄人脸识别。我们知道从跨年龄人脸数据提取的特征应该是混合了年龄特征和身份特征的, 为了让两张同一身份不同年龄的人脸成功匹配, 就要将人脸特征中的身份特征单独提取出来做匹配。所以, 基于年龄估计导向的方法就是用年龄估计任务尽可能提取准确的年龄信息并从混合特征中分离出去。分解方式是直接采用混合特征减去年龄特征 (残差分解), 整体网络架构也是先通过一个卷积神经网络提取人脸特征, 然后接一个全连接层得到混合人脸特征, 之后混合特征历经两个分支任务, 年龄估计分支经过一个全连接层作为年龄特征, 人脸身份识别分支用混合特征减去年龄特征的结果作为身份特征, 分别用年龄 Loss 函数和身份 Loss 函数同时训练两个任务, 年龄估计和人脸身份识别都有采用多分类 Softmax Loss 函数。

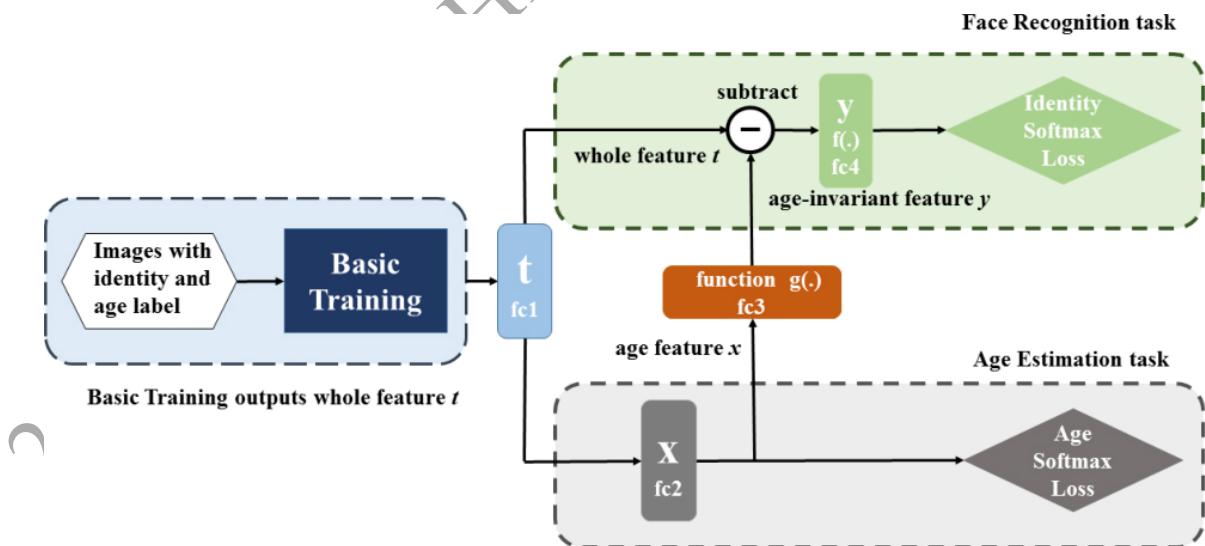


图 3 AENet: 基于年龄估计的跨年龄人脸识别网络架构。其中 t 表示人脸混合特征, x 表示在年龄估计分支任务中的年龄特征, y 表示与年龄无关的身份特征。

AENet 的特征提取模块采取了基于 LightCNN [8] 如图 4 所示的 LightCNN 网络架构。

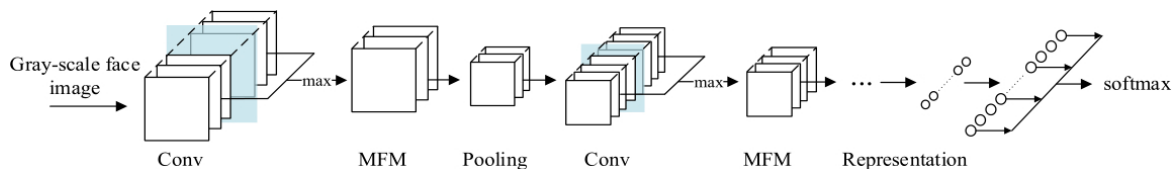


图 4 基于 LightCNN 的人脸深度特征提取网络

2 提出的方法

2.1 动机

目前针对跨年龄人脸身份识别任务的生成模型和判别模型两大类方法中，生成式方法较明显的两个缺点是强参数假设和算法复杂度，导致计算代价高昂，实际应用中结果不稳定，并且生成式方法不是端到端的，不能一步解决识别问题。所以目前判别式方法已经成为主流，并且由于卷积神经网络的发展和在视觉特征提取和特征分类方面的巨大优势，目前判别式方法基本都是基于深度卷积神经网络的。

判别式方法中普遍把提取后的人脸特征看作是代表身份和年龄等特征的混合，采用混合特征分解的方式来提取只和身份相关的特征。比如，论文 [7] 试图通过引入典型相关性分析 (CCA) 度量分解后的身份特征和年龄特征之间的相关性，然后对抗地先最大化这种相关性再最小化以进一步解耦分解后的两种特征。但无论什么分解方法，分解不彻底、分解后的特征相互混合、分解导致信息丢失仍然是主要和根本的问题。以往的方法都是在网络的最后引入一个全连接层做一次特征分解，我们猜想这可能是导致分解不彻底的原因。从这点出发，我们希望能够多步分解将混合特征分解得更彻底，并且确保分解后特征的信息损失尽量小。

本研究方法的核心思想是采用分步分解 (stage-wise) 的方式，每步采用相同的残差分解或者正交分解，每步分解的身份特征和年龄特征都直接连接最后的 Loss 函数进行训练，并且每个阶段分解之前的特征都是包含了之前所有阶段或前一个阶段的特征，以弥补之前分解带来的信息损失。我们将这一方法命名为 SWNet。

2.2 方法

2.2.1 网络结构

整个网络结构大致分为主体的卷积特征提取模块 (backbone)、分步特征分解模块、分类模块。其中主体模块一般选择类似 ResNet 的卷积结构来提取深层特征，这些特征很好挖掘了人脸的内在本质信息，它们同时含有身份信息 and 年龄信息，用于后面的分解。特征分解模块由 K 步构成 (这里我们取 $K=4$)，每步结构相同，分解方式有两种选择：正交分解 (图 5) 和残差分解 (图 6)。其中 X_n 表示网络中间的某一模块的输出特征，同时也是下一模块的输入特征， H_n 表示网络中的模块 (包含若干网络层)。我们对每个模块提取的混合特征进行分解，比如第一阶段的混合特征为 $H_1(X_0)$ 。每步分

解之后的身份和年龄特征都参与到最后的 identity loss 和 age loss 中以达到对每步分解的特征的监督训练。为了弥补分解中造成的特征损失，正交分解中，每步分解前的混合特征包含了之前所有阶段的输出特征（受 DenseNet [9] 启发）。残差分解中，每步分解前的混合特征包含了前一阶段的输出特征，即每次经过一个模块后提取的混合特征加上前一阶段的特征，例如 $X_2 = X_1 + H_2(X_1)$ ，其中 $H_2(X_1)$ 表示第二阶段的混合特征， X_1 表示第一阶段最终输入到第二阶段的特征。这种操作既可以解决神经网络反向传播梯度不稳定的问题，提高网络浅层学习能力，也可以为当前阶段获取前面的浅层特征，使得包含的信息更丰富（受 ResNet [6] 启发）。

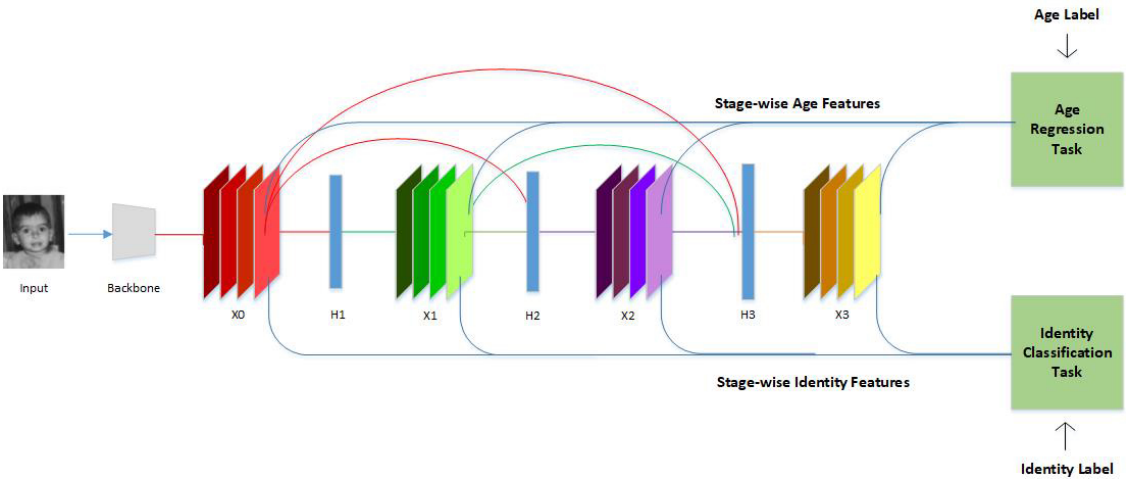


图 5 SWNet 正交分解模式

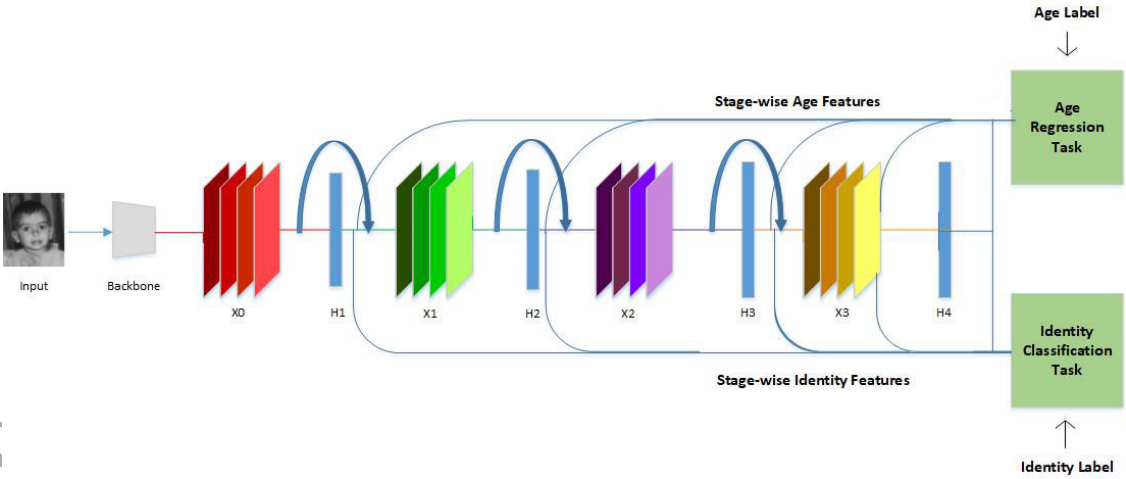


图 6 SWNet 残差分解模式

2.2.2 特征分解

(1) 正交分解

$$x = x_{age} \cdot x_{id}$$

其中 $x_{\text{age}} = \|x\|_2$, $x_{\text{id}} = \left\{ \frac{x_1}{\|x\|_2}, \frac{x_2}{\|x\|_2}, \dots, \frac{x_n}{\|x\|_2} \right\}$, 所以 $\|x_{\text{id}}\|_2 = 1$ 。 $x_{\text{age}} = \|x\|_2$ 表示 L2 范数, 即向量的模长。根据 [5], 每阶段的混合特征 $x_l = H_l([x_0, x_1, \dots, x_{l-1}])$, 然后每个 x_l 分解为 x_{age} 和 x_{id} 。之所以选取混合特征的 L2 范数作为年龄特征, 是因为相比残差分解把人脸混合特征看作是身份特征和年龄特征相加的思路, 正交分解把人脸混合特征看作是身份特征和年龄特征相乘的结果, 提供了另一种新的分解思路, 这样身份特征就是混合特征除以 L2 范数的归一化向量, 可以用于人脸识别多分类任务, 年龄特征就是一个实数值, 可以直接用于年龄回归任务。

(2) 残差分解

$$x = x_{\text{id}} + x_{\text{age}}, x_{\text{age}} = H(x), x_{\text{id}} = x - H(x)$$

根据图 6, 每阶段的混合特征 $x_l = H_l(x_{l-1}) + x_{l-1}$, 然后每个 $H_l(x_{l-1})$ 作为 x_{age} , $x_{l-1} - H_l(x_{l-1})$ 作为 x_{id} 。目前基于残差分解的方法 (图 3), 例如 [5] 和 [7] 等, 都是在卷积网络的最后引入全连接层, 让混合的人脸卷积特征通过全连接层, 将通过全连接层提取的特征设定为年龄特征, 然后用混合特征即通过全连接层之前的特征减去年龄特征得到的特征向量就设定为身份特征。

2.2.3 Loss 函数选择

年龄回归 Loss 函数采用简单的平均绝对误差 (MAE):

$$L_{\text{age}} = \frac{1}{2M} \sum_{i=1}^M \|f(x_{\text{age}}^i) - z_i\|_2^2$$

x_{age}^i 是第 i 个训练样本的混合特征的 L2 范数 (即对应分解后的年龄特征), z_i 是第 i 个样本的年龄标签, 使用线性多项式 $f(x) = k \cdot x + b$ 映射年龄标量。身份识别属于多类分类, Loss 函数可以有多种选择, 我们一般采用 A-Softmax Loss [11] 或 Large Margin Cosine Loss [12]。

(1) A-Softmax Loss

$$L_{\text{id-ang}} = \frac{1}{M} \sum_{i=1}^M -\log \left(\frac{e^{s \cdot \psi(\theta_{y_i, i})}}{e^{s \cdot \psi(\theta_{y_i, i})} + \sum_{j \neq y_i} e^{s \cdot \cos(\theta_j, i)}} \right)$$

其中 $\psi(\cdot)$ 定义为 $\psi(\theta_{y_i, i}) = (-1)^k \cos(m\theta_{y_i, i}) - 2k$, $\theta_{y_i, i}$ 是第 i 个样本身份特征和身份标签 y_i 对应的权重向量之间的角度, $\theta_{y_i, i} \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m} \right]$, $k \in [0, m-1]$, $m \geq 1$, $s > 0$, m 是控制角裕度大小的整型超参数, s 是可调整的尺度因子, 用来弥补 Softmax 学习的不足。

OENet 模型 [5] 的人脸身份识别任务的损失函数就是采用了 A-Softmax loss 函数。常用的 Softmax loss 函数单纯地将人脸识别看作是传统多分类任务, 取网络最终全连接层的输出得到一个概率分布。但是 A-Softmax loss 提供了一种新的构造多分类损失函数的思路, 从向量夹角的角度来学习具有判别性的特征, 损失函数优化的目标就是使得特

征向量和真实类别所在的那个权重向量之间的夹角尽可能小（即距离尽可能近），和其它权重向量之间的夹角尽可能大（即距离尽可能远）。

(2) Large Margin Cosine Loss

$$L_{\text{id-lmc}} = \frac{1}{M} \sum_{i=1}^M -\log \left(\frac{e^{s(\cos(\theta_{y_i,i})-m)}}{e^{s(\cos(\theta_{y_i,i})-m)} + \sum_{j \neq y_i} e^{s \cdot \cos(\theta_j,i)}} \right)$$

其中 $W = \frac{W^*}{\|W^*\|}$, $x = \frac{x^*}{\|x^*\|}$, $\cos(\theta_j, i) = W_j^T x_i$, x_i 是第 i 个样本特征向量，对应身份类标签 y_i , W_j 是第 j 类的权重向量， θ_j 是 W_j 和 x_i 之间的角度。从表达式可以看出，Large Margin Cosine Loss 其实和 A-Softmax loss 类似，只不过将控制角裕度大小的整型超参数 m 用来控制余弦值大小。

2.2.4 特征解耦

解耦是为了进一步减小分解后的身份特征和年龄特征之间的关联性和相似性，因为前期分解可能不够彻底。每一次分解后得到身份特征和年龄特征，分别来计算身份特征的损失和年龄特征的损失，通过加权求和得到总的损失，如果引入解耦的话，我们还要度量分解后的身份特征和年龄特征之间的相似性或关联性，还要在总损失中加入这个相似性度量一起来优化。

我们考虑在残差分解模式中引入解耦。在分解不够彻底时，解耦可以进一步减小分解后的身份特征和年龄特征之间的关联性和相似性。分解后的年龄特征和身份特征的解耦可以采用最小化 CMM (Canonical Mapping Module) 或最小化向量余弦。解耦也可以有两种方案：可以针对每阶段分解的特征进行，也可以只针对最后一步分解的特征进行。将要优化的解耦的度量加入最终的损失函数一同监督模型的训练。

CMM 度量表达如下：

$$\forall t \in \{\text{id, age}\} : v_t = C(x_t) = w_t^T x_t$$

$$\rho = \frac{\text{Cov}(v_{\text{id}}, v_{\text{age}})}{\sqrt{\text{Var}(v_{\text{id}})\text{Var}(v_{\text{age}})}}$$

CMM 是用来度量向量之间的线性相关性或关联性，[10] 引入这个是为了度量身份特征和年龄特征之间的关联性，并加入到损失函数一起进行最小化训练，以达到进一步解耦身份特征和年龄特征的目的。CMM 将身份特征和年龄特征通过权重参数映射到一个新的 Canonical 变量，变量的取值和权重参数有关，然后再定义 Canonical 相关性 ρ ，即 canonical correlation，通过优化这个相关性来实现年龄和身份特征的去相关。优化分为两个步骤：第一个优化学习步骤就是通过更新权重参数来最大化 canonical correlation 系数，此时身份变量和年龄变量之间的相关性达到最大值；第二个优化学习步骤就是将第一步学到的身份和年龄之间的最大相关性加入模型最终的损失函数进行最小化，以尽可能解耦身份特征和年龄特征。这种对抗学习的思路，使得模型每次优化都是减小身份特征和年龄特征之间的最大的 canonical correlation，这样可以显著地解耦身份和年龄，使得身份特征尽可能不包含年龄信息。

至于余弦解耦，则是优化年龄和身份特征向量之间的余弦值。即通过余弦值来度量年龄特征和身份特征之间的距离，然后加入最后的损失函数一起优化。本文采用 CMM 解耦。

2.2.5 网络训练优化方法

网络训练采用反向传播算法和随机梯度下降算法的结合 [3]。

3 实验评估

实验中我们采用 Dell 工作站和一块 NVIDIA 1080Ti 显卡作为硬件环境，采用 CentOS 系统、Python 3.6，深度学习框架 Pytorch 1.0 作为软件环境。

3.1 数据集

实验数据采用 CACD 跨年龄人脸数据集 [13]，该数据集有超过 16 万张图片，包含 2000 个身份人体，年龄范围跨越 16 到 62，我们依据论文 [4] 的实验设置，选取了身份 ID 为 1 到 120 的人脸图片作为后续的一种测试方案的测试集（我们命名为 CACD120），其余 1880 个带有身份标记和年龄标记的人脸图片全部作为训练集。关于数据的预处理，我们首先用 MTCNN 模型检测人脸并修正人脸，裁剪出的人脸图片统一为 144×144 ，在输入网络之前，再随机裁剪成 128×128 ，所有输入图片都被灰度化和归一化。这一人脸检测、修正和裁剪的预处理效果如图 7 所示。



图 7 第一行是原始图像，第二行是预处理后的图像

另一种测试方案采用 CACD_VS 数据集。CACD_VS 是 CACD 的一个子集，一般用作测试集，总共包含 2000 组正样本（同一个体的两张图片）和 2000 组负样本（不同个体的两张图片），一共 4000 组 8000 张人脸。这里用作我们第二种测试方案的测试集。

3.2 实验设计

训练过程中输入为批量标注了身份和年龄的人脸图像。输出为分解后的年龄特征 (age feature) 和身份特征 (identity feature)。训练中每步的年龄和身份 loss 函数都是

直接求和得到总的任务 loss 函数。如果引入解耦的话解耦度量也一并加入总 loss 函数。最终的多任务 loss 函数如下：

$$L = L_{id}(x_{id}) + \lambda_1 L_{age}(x_{age}) + \lambda_2 L_{\rho}(x_{id}, x_{age})$$

其中右边第一项是身份识别任务的 loss 函数，第二项是年估计的 loss 函数，第三项是特征解耦的 loss 函数。

为了便于和已有方法进行公平对比，我们的测试方案有两种，分别对应上文提到的 CACD120 和 CACD_VS，两种方案具体细节有所不同，但都是针对每组人脸的身份特征之间进行余弦相似度的计算。

基于 CACD120 的测试，我们遵循论文 [7] 的测试实验设置，我们选取身份 id 为 1-120 并且 rank 标记为 3-5 的人脸作为测试集，其中在 2013 年拍摄的图片作为查询图像 (query images)，剩下的作为数据库内的被查询图像 (gallery images)。在本测试集中，使用 MAP (mean average precision) 作为评估协议，Cosine 距离用于计算两张图片之间的相似度。具体来说，令 $q_i \in Q$ 作为 query images， Q 是 query images 集合。对于 q_i ，相关图片的数量是 m_i ，相关图像可以表示为 Y_1, Y_2, \dots, Y_{m_i} ， E_{ic} 是 q_i 从上到下的检索结果。 q_i 的平均精度 (AP) 可计算如下：

$$AP(q_i) = \frac{1}{m_i} \sum_{c=1}^{m_i} \text{Precision}(E_{ic})$$

其中 $\text{Precision}(E_{ic})$ 是指 E_{ic} 中检索得到图像和应该被检索得到图像的比率。然后， Q 的所有图像 (MAP) 的 AP 的均值可以计算如下：

$$MAP(Q) = \frac{1}{|Q|} \sum_{i=1}^{|Q|} AP(q_i)$$

基于 CACD_VS 数据集的测试中，我们遵循论文 [5] 中的实验设置，测试时计算所有 4000 组图片的余弦相似度，将 4000 组均分成 10 份，取其中 9 份计算基于余弦的相似度阈值，计算得到的阈值用于预测剩下的 1 份。重复以上过程 10 次，最后取均值作为测试准确率的结果。

3.3 实验结果

在上述实验数据和实验设置下，得到实验结果如表 1 和表 2 所示。实验除了实现我们自己提出的模型 (采用残差分解模式的 SWNet-Res 和正交分解模式的 SWNet-OE)，还复现了相关工作 [5, 7] 中提到的两种模型 (AENet 和 OENet) 以用来对比，其中 AENet 采用了 LightCNN 骨干网络，同时为了更清晰的对比，我们还对不引入特征分解的 AENet 和 OENet 模型进行了实验。为了区分，我们将不含有特征分解的模型用 Basic 表示，引入特征分解的模型用 Decomposition 表示。表中可以看到残差分解模式精准度最高，而正交分解模型无论采用单步还是多步都比基于残差的分解方式要差一些。可能是相比于残差分解，正交分解的约束条件太强，使得信息有所丢失。

另外我们在模型训练中设置 batch_size 大小为 64，数据加载开了 4 个线程，一个完整的 epoch（训练周期）大约 20 分钟，总共设置了 80 个 epoch，学习率设为 0.001。而测试一张图像的时间约为 2.5ms。

需要指出的是深度学习方法的性能受制于多种因素的影响，包括网络大小、初始化方法、训练优化方法、参数调节的技巧，以及图像预处理等。尽管我们尽最大可能采用比较算法在原始论文中的设置，但是部分相关内容并未在原始论文中得以清晰阐释，我们根据自己的理解和反复调试复现 AENet 和 OENet 算法。

表 1 CACD120 上各种算法比较测试结果

模型	MAP
AeNet_Basic	0.804
AeNet_Decomposition	0.802
OENet Basic	0.670
OENet Decomposition	0.645
SWNet-OE	0.786
SWNet-Res	0.815

表 2 CACD_VS 上各种算法比较测试结果

模型	相似度
AENet_Basic	0.897
AENet_Decomposition	0.904
OENet Basic	0.870
OENet Decomposition	0.712
SWNet-OE	0.837
SWNet-Res	0.930

图 8 是本文提出模型在 CACD_VS 中身份识别成功的几组样例结果，图 9 是本文方法在 CACD 数据集中识别失败的一些样例。可以发现，尽管识别失败，但是模型识别结果图片和查询图片肉眼看来是很相似的。从这些例子也可以看出跨年龄人脸身份识别是很具挑战性的，也激发了我继续进行这一问题研究的动力。



图 8 CACD_VS 数据集中识别成功的案例，每组图片第一张为查询图片，第二张为对应的同一身份的图片

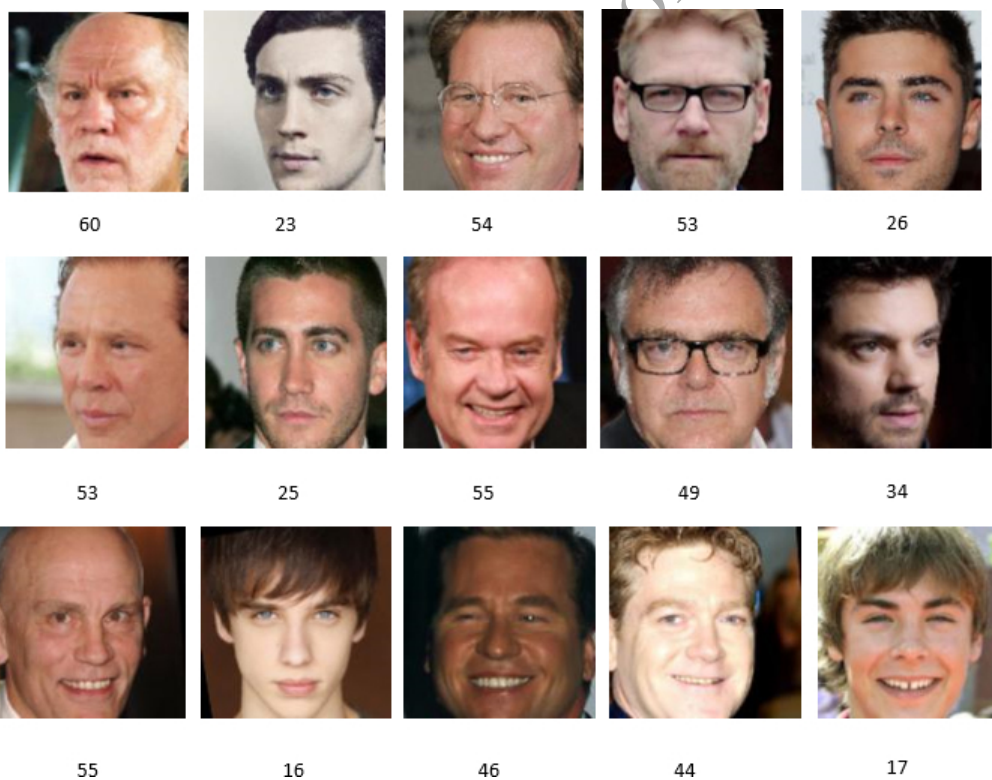


图 9 CACD 数据集中检索失败的案例，第一行是查询图片，第二行是不正确的查询结果图片，第三行是对应的正确的结果图片

4 讨论

随着神经网络和深度学习在计算机视觉领域的广泛应用，人脸识别和年龄识别各自都取得了显著进步，一般的基于神经网络的人脸识别或人脸年龄识别都是单个独立的任务，很难建立联系。但是随着跨年龄人脸识别的现实需求和任务的提出，需要对年龄有跨度的人脸进行识别，而解决跨年龄人脸识别任务需要的数据集不仅仅是具有身份信息的人脸图片，而是需要针对同一身份的人具有多个分布在不同年龄段的人脸图片。由于跨年龄身份识别任务的特殊性，目前针对此任务提出的基于深度神经网络的方法往往需要同时考虑年龄特征和身份特征，因此决定了这个任务相比一般的人脸识别任务，更需要利用年龄信息来帮助实现人脸身份识别。

以往的基于深度神经网络的判别式方法主要是通过卷积网络来提取人脸混合特征，然后通过一次分解提取与年龄无关的身份特征来进行人脸识别。我们的方法与之前方法的不同之处，主要在于我们的模型分解是多步的，而不是在网络的最后进行一次对混合特征进行分解，每次分解都会连接到最后的损失函数分别对分解后年龄特征和身份特征进行学习，并且每步分解的特征都是混合了之前所有层或者前一层特征。我们认为这既可以通过多次分解将卷积特征分解得更彻底，也可以在后面阶段中保留之前分解损失的信息。另外，在我们的模型中，每步分解都引入分解后特征的解耦，以进一步区分年龄相关特征和身份相关特征。

下面我们总结一下我们的方法和相关方法的联系与区别。

4.1 联系

我们的方法和之前基于神经网络的方法（代表的方法 [2, 4]）最大的相同之处在于我们也是基于混合特征分解的，就是将卷积网络提取的特征看作是混合了身份信息和年龄信息的特征，都是在进行人脸身份识别之前进行特征的分解获取只和身份相关的特征，都是多任务的，分别用年龄标记和身份标记来训练身份特征和年龄特征，端到端地一步到位地同时进行身份识别和年龄估计。年龄估计任务对于身份识别任务起到提供辅助引导信息的作用。我们的方法都在实验中尝试了已有的正交特征分解和残差分解策略，并且我们的方法也采用了分解之后特征做进一步解耦。

4.2 区别

(1) 我们的方法和以往方法最明显的不同之处在于我们不是在网络的最后针对混合人脸特征仅分解一次，而是在整个网络的过程中进行多步特征分解，即在提取特征的同时就不断进行分解，以达到将身份信息和年龄信息分解彻底的目标，同时我们的特征解耦也是伴随着每一次的分解，每一步的分解特征都会用训练的身份标签和年龄标签进行监督，同时每一步的特征解耦也会加入最终的损失函数进行优化。

(2) 我们采用的多步分解方法的每步都是采用残差分解或者都是正解分解，其中残差分解发生在每个 ResBlock 中 shortcut 加入特征图之前，混合特征依然会随着网络进

入下一个 ResBlock, 这样每个阶段还会包含上一阶段的信息 (如图 6)。正交分解则是每一阶段除了分解, 还会融合之前所有阶段的信息 (如图 5)。残差分解模式下年龄任务分支采用分类, 而正交分解则采用回归。

5 结论

在本论文中, 我们提出了一个基于分步特征分解的跨年龄人脸识别深度神经网络 (SWNet) 方法。考虑到利用现有方法身份相关和年龄相关的特征难以彻底分解, 我们对其在每步分解中都用年龄和身份标签对分解后的特征进行监督学习, 并且将前一阶段或之前所有阶段的特征加入提取的混合特征以弥补之前分解带来的信息损失, 利用解耦减小分解后的身份特征和年龄特征之间的关联性和相似性, 并在公开的跨年龄人脸数据集进行实验, 证明了 SWNet 方法的有效性。

2020 S.-T. Yau High School Science Award

参考文献

- [1] 李倩玉, 蒋建国, 齐美彬, 基于改进深层网络的人脸识别算法, 电子学报, 45 (3) : 619-625, 2017.
- [2] 吴长虹, 苏剑波, 陈叶飞, 抗年龄干扰的人脸识别, 电子学报, 46 (7) : 1593-1600, 2018.
- [3] 周飞燕, 金林鹏, 董军, 卷积神经网络研究综述, 计算机学报, 6: 1229-1251, 2017.
- [4] D. Gong, Z. Li, D. Lin, J. Liu, and X. Tang. Hidden factor analysis for age invariant face recognition. International Conference on Computer Vision (ICCV), 2013.
- [5] Y. Wang, D. Gong, Z. Zhou, X. Ji, , H. Wang, Z. Li, W. Liu, and T. Zhang. Orthogonal Deep Features Decomposition for Age-Invariant Face Recognition. European Conference on Computer Vision (ECCV), 2018.
- [6] K.He, X.Zhang, S.Ren, andJ.Sun. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [7] T. Zheng, W. Deng, and J. Hu. Age Estimation Guided Convolutional Neural Network for Age-Invariant Face Recognition. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017.
- [8] Wu X, He R, Sun Z. A lightened cnn for deep face representation. arXiv preprint arXiv:1511.02683, 2015.
- [9] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [10] 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. CVPR, 2019.
- [11] Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., Song, L.: SphereFace: Deep Hypersphere Embedding for Face Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [12] H. Wang, Y. Wang, Z. Zhou, X. Ji, Z. Li, D. Gong, J. Zhou, and W. Liu. Cosface: Large margin cosine loss for deep face recognition. Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [13] B.-C. Chen, C.-S. Chen, and W. H. Hsu. Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval, pages 768–783. Springer International Publishing, Cham, 2014.

致谢

我于 2019 年初入选了由中国科协和教育部共同组织实施的中学生科技创新后备人才培养计划（简称“英才计划”）。在指导老师——浙江大学计算机科学与技术学院钱运涛教授的引领下，我系统地参与了人工智能方向的理论学习、学术研讨和科研实践。在学习导师推荐的计算机视觉方向最新顶级会议论文（如 CVPR 2019）时，我发现跨年龄人脸识别问题受到了日益广泛的关注。然而，识别年龄间隔较大的人脸图像仍然具有较大挑战性，这主要是因为年龄变化会引起人脸图像呈现出较大的差异。我考虑到多步的特征分解也许可以更彻底地分离身份特征和年龄特征。因此，我提出了一个基于多步特征分解的跨年龄人脸识别深度神经网络（SWNet）方法。这个研究设想得到了导师的肯定，他鼓励我大胆地把研究做下去。

本研究的理论推演、代码编写、实验设计与实施、结果的分析与比较、论文撰写等工作的全部内容均由我本人独立完成。在整个研究进程中，钱老师多次为我指点迷津，当我遇到困难或研究止步不前时，都给予我积极的指导与帮助。钱老师是我最尊敬的学术导师，他严谨的治学态度、渊博的学术知识、诲人不倦的敬业精神将使我终身受益。

本课题的研究过程是颇为艰苦的，由于研究内容属于近年来迅速迭代发展的前沿领域，需要投入大量的时间和精力。这对一个同时还需要准备国内高考、国际升学的高中生而言是一个极大的挑战。很幸运的是，我完成了这项研究。这个成果的顺利完成还要归功于一直以来给予我无条件支持与陪伴的父母，予以悉心指导与精心培养的老师，为我提供学习和研究平台的学校。在此对他们表达深深的敬意和诚挚的谢意。

附件：参赛选手与指导老师简历

蒋昕昀同学，女，杭州外国语学校高三（6）班学生，勤奋自律，学业优异，在初高中阶段年级排名始终为前 1%，荣获校一等奖学金、综合素质优秀生、三好学生等荣誉，入选 2019 年浙江省“英才计划”、2020 年美国斯坦福大学数学夏令营（中国大陆地区仅两席）；知行合一，勇于创新，积极参与海内外各级各类学术竞赛和社会实践活动，表现突出，硕果累累，代表性成果如下：

- 2020 年浙江省青少年科技创新大赛一等奖，入围全国总决赛
- 2020 年世界青少年创新发明展金奖
- 2019 年美国高中生数学建模大赛特等奖提名奖
- 2019 年世界机器人大赛一等奖
- 2019 年英国物理奥林匹克竞赛超级金奖
- 2019 年美国数学竞赛（AMC）全球排名前 1%
- 2019 年 ASDAN 美式数学竞赛中国赛区团队第 1 名
- 2018 年中央电视台“希望之星”英语演讲比赛浙江省亚军
- 2019 年中国日报社“21 世纪杯”英语演讲比赛浙江省一等奖
- 申请专利《作业通：线上线下融合的学生作业智能管理系统》
- 受邀参加 2020 年伦敦国际青年科学论坛
- 受邀参加 2020 年西方七国首脑会议青年峰会，获得“杰出代表”称号
- 担任“Little Elite Program for Girls（女生小英才计划）”校际社团社长，鼓舞全国多地女生在 STEAM 领域深耕发展
- 原创微纪录片《00 后高中生的抗疫日记》浏览量超百万，受到新华社、中国日报、中国教育报等权威媒体与教育部官网、哈佛中国教育论坛的关注与报道
- 翻译国际合作学习顶级专家卡干·斯宾塞著作《小组建设》与《团队建设》（由广东教育出版社出版）

钱运涛老师，浙江大学计算机科学与技术学院教授、博士生导师、计算智能实验室主任。毕业于西安交通大学，分别于1989年、1992年获自动控制专业学士和硕士学位，1996年3月毕业于西安电子科技大学，获信号处理专业博士学位，1996-1998年在西北工业大学计算机系从事博士后研究，1998年进入浙江大学计算机学院，2002年晋升教授。中国海洋大学兼职教授、博导，浙江省151人才。现任“IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing”副主编，“信号处理”、“哈尔滨工业大学学报”、“计算机工程”、“信息融合学报”等杂志编委。发表学术论文近200篇。主要从事机器学习、信号处理、人工智能研究。

2020 S.-T. Yau High School Science Award