

doi: 10.11835/j.issn.1000-582X.2025.06.009

引用格式:冉光伟,何祺,王楠,等.基于非线性深度子空间学习的微表情识别方法研究[J].重庆大学学报,2025,48(6): 98-111.



基于非线性深度子空间学习的微表情识别方法研究

冉光伟¹,何祺²,王楠³,冯为嘉³,姜立标^{4,5}

(1. 星河智联汽车科技有限公司,广州 510335; 2. 广汽丰田汽车有限公司,广州 511455; 3. 天津师范大学计算机与信息工程学院,天津 300382; 4. 华南理工大学 机械与汽车工程学院,广州 510641; 5. 广州城市理工学院机械工程学院与机器人工程学院,广州 510850)

摘要:针对微表情识别中深度子空间网络鲁棒性差和泛化能力弱的问题,提出了基于非线性深度子空间学习和光流计算的微表情识别方法。该方法通过引入核变换充分挖掘微表情中的情感信息,同时使用光流特征捕捉微表情的运动信息,提高识别的鲁棒性。在 SMIC、SAMM、CASME 和 CASME II 4 个广泛使用的自发微表情数据集和 3DB-combined 复合数据集上的实验表明,所提方法识别性能优于 MACNN、Micro-Attention 等深度学习方法,在复合数据集上的准确率达到 0.834 6。此外,在 SMIC 数据集上添加 10%、20%、30% 和 40% 的随机噪声块后,在不同噪声水平下的未加权 F_1 分数均优于其他算法,表明该方法在微表情识别任务中的有效性和鲁棒性。

关键词:深度子空间;微表情识别;光流特征;主成分分析

中图分类号:TP391

文献标志码:A

文章编号:1000-582X(2025)06-098-14

Micro-expression recognition based on nonlinear deep subspace learning

RAN Guangwei¹, HE Qi², WANG Nan³, FENG Weijia³, JIANG Libiao^{4,5}

(1. Syncore Autotech Co., Ltd., Guangzhou 510335, P. R. China; 2. GAC Toyota Motor Co., Ltd., Guangzhou 511455, P. R. China; 3. College of Computer and Information Engineering, Tianjin Normal University, Tianjin 300382, P. R. China; 4. School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou 510641, P. R. China; 5. School of Mechanical Engineering and Robotics Engineering, Guangzhou City University of Technology, Guangzhou 510850, P. R. China)

Abstract: To address the issues of poor robustness and weak generalisation in deep subspace network-based micro-expression recognition, this paper proposes a novel method that integrates nonlinear deep subspace learning with optical flow computation. The method employs kernel transformation to comprehensively extract emotional features from micro-expressions while simultaneously utilizing optical flow characteristics to capture subtle motion dynamics, thereby enhancing recognition robustness. Experimental validation is performed on 4 widely

收稿日期:2024-08-12

基金项目:国家自然科学基金资助项目(61602345, 62002263)。

Supported by National Natural Science Foundation of China (61602345, 62002263).

作者简介:冉光伟(1980—),男,高级工程师,主要从事智能座舱系统的开发和应用方向研究,(E-mail)gwr@syncore.space。

通信作者:姜立标(1965—),男,副教授,(E-mail)jlb620620@163.com。

adopted spontaneous micro-expression datasets (SMIC, SAMM, CASME and CASME II) as well as a composite dataset 3DB-combined samples. Results demonstrate that the proposed method outperforms existing deep learning algorithms, including MACNN and Micro-Attention, achieving a recognition accuracy of 0.834 6 on the composite dataset. Furthermore, after adding 10%, 20%, 30%, and 40% random noise blocks to the SMIC dataset, the method consistently maintains superior unweighted F_1 scores compared to other algorithms. These findings substantiate its effectiveness and robustness in real-world micro-expression recognition scenarios.

Keywords: deep subspace; micro-expression recognition; optical flow features; principal component analysis

微表情作为一种瞬时而微小的面部表情^[1],具有丰富的情感信息。虽然持续时间短暂^[2](通常在1/25~1/5 s之间),但包含了个体的真实情感状态,常常反映出人内部情感冲突或试图隐藏情感的特质。对微表情识别算法进行研究,旨在精准揭示隐藏在人们面部微小肌肉变化中的真实情感,在判断陈述的真实性上提供可靠依据,因此,微表情识别已成为情感计算、心理学研究以及实际应用中的重要课题。微表情识别技术在多个领域具有广泛的应用前景,包括:金融欺诈检测、医疗诊断中的心理评估、人机交互的情感反馈系统,以及安全监控中的谎言检测等。这些应用场景都依赖于对人类情感状态的准确感知,因而微表情识别技术的研究显得尤为重要^[3-4]。

尽管微表情识别在理论和实践中具有重要价值,但其实现过程却面临着诸多挑战。微表情识别算法随着计算机视觉和机器学习技术的发展经历了多个阶段,目前主要分为以下2类。首先是早期的传统手工特征提取方法,通过提取图像中的某些关键特征,结合传统机器学习算法进行分类。例如:局部二值模式^[5](local binary patterns, LBP)、光流法^[6](optical flow, OF)、方向梯度直方图^[7](histogram of oriented gradients, HOG)、定向光流直方图^[8](histogram of oriented optical flow, HOOF)等方法,这类方法捕捉微小的面部运动变化大多依赖手工设计的提取器,需要专业的知识和复杂的参数调整过程,这些方法在真实场景实现微表情识别时存在泛化能力和稳健性差的问题。另一类是基于深度学习的微表情识别方法。通过卷积神经网络^[9](convolutional neural network, CNN)、循环神经网络^[10](recurrent neural network, RNN)等深度神经网络模型自动提取微表情特征实现识别分类,与传统手工特征提取方法相比,这些方法显著提高了微表情识别的精度,摆脱了手工设计特征的局限性。例如,Reddy等^[11]提出了一种基于3D时空卷积神经网络(3D spatiotemporal convolutional neural networks, 3D-CNN)的方法,能够学习并提取微表情在时间轴上的演变信息,例如面部肌肉在几个连续帧中的微小运动,提高微表情的识别精度。Khor等^[12]设计了1个长短期记忆卷积神经网络(long short-term memory convolutional neural network, CNN-LSTM),将卷积网络提取图像特征和LSTM捕捉动态变化信息相结合,将图像和3个光流特征分量输入到CNN-LSTM模型架构中,取得较好效果。除此之外,Liong等^[13]提出了一种浅三流3D卷积神经网络(shallow triple stream 3D convolutional neural networks, S-3D-CNN)的方法,该模型通过减少卷积核的数量和稀疏化处理,降低计算复杂度,在微表情识别任务中取得优异性能。Zhang等^[14]提出一种基于多尺度卷积神经网络(multi-scale convolutional neural network, MSCNN)的微表情识别方法,通过融合不同尺度的特征图,捕捉微表情的多层次信息,提升识别精度。然而,深度学习模型需要大量样本进行训练^[15],而微表情持续时间短暂且难以捕捉,自然情境下发生频率低,且准确标注需要专业知识,导致高质量数据十分稀缺,属于典型的小样本问题^[16],因此,深度学习方法面临如何在小样本下实现分类识别的问题。

为解决上述问题,Peng等^[17]基于迁移学习方法使用微调的预训练ResNet-10模型在微表情识别任务中取得较好的分类性能,准确率提高10%左右,验证了利用预训练模型来提取通用特征,是一种有效解决微表情识别小样本问题的方法。Gan等^[18]研究了一种基于生成对抗网络(generative adversarial network, GAN)的微表情识别方法,通过生成与判别器的对抗学习,GAN能够生成具有高度真实性的微表情样本,进一步扩展训练集的数据量,缓解小样本问题,提高模型泛化能力。Happy等^[19]提出了使用时序数据增强的方法,通过对微表情序列的时序结构进行处理,使模型能够更好地学习微表情的动态特征。除此之外,研究者们还采用子空

间学习的方法进行微表情识别研究。Chengeta等^[20]结合主成分分析(principal component analysis, PCA)技术与支持向量机(support vector machines, SVM)方法进行微表情识别,PCA和特征选择技术有效减少特征维度,提高计算效率和分类性能。Yu等^[21]提出鲁棒主成分分析(robust principal component analysis, RPCA)和局部时空方向特征(local spatiotemporal directional features, LSTD)结合的方法,通过有效处理噪声和捕捉微表情的动态信息,提高微表情识别的准确性。此外,Wang等^[22]结合光流特征和改进的主成分分析网络(principal component analysis network plus, PCANet+)进行微表情识别,PCANet+通过卷积操作提取面部图像的局部特征,并利用PCA进一步对特征进行降维和增强,这种组合使网络能有效学习面部特征的空间和结构信息。然而,由于微表情变化的复杂性,并且容易受到光照、拍摄角度和个体差异等多种因素影响,使数据呈现非线性分布的特点,属于非线性问题^[23]。而PCANet的PCA卷积核对异常值较为敏感,导致模型的鲁棒性较差且泛化能力较弱。

为了解决深度子空间网络PCANet在微表情识别中鲁棒性差和泛化能力弱的问题,提出了一种基于非线性深度子空间学习和光流计算的微表情识别方法,通过引入核变换充分挖掘微表情中的情感信息,使用光流特征来捕捉微表情的运动信息。核变换将微表情的高级语义特征进行有效提取,进一步减轻小样本问题和过拟合困扰。另外,光流特征的引入不仅能够捕捉微表情的动态信息,还能提高微表情的鲁棒性,使其对光线变化和遮挡更具抵抗力。最后,实验结果验证了本方法在改善模型泛化能力和鲁棒性方面的有效性,同时在微表情识别任务中取得较好的识别效果。

1 深度子空间学习算法

深度子空间学习算法详细描述为:假设大小为 $m \times n$ 的 N 幅输入图像 $\{\mathbf{M}_i\}$ ($i=1, 2, \dots, N$),令所有阶段的采样块大小为 $k_1 \times k_2$,其中, $0 < k_1 < m$, $0 < k_2 < n$ 。首先,对第 i 张图像 \mathbf{M}_i 进行 $k_1 \times k_2$ 大小的块采样,将得到的块数据展开为列向量并去均值后级联起来,得到输入图片 \mathbf{M}_i 的数据矩阵如下

$$\bar{\mathbf{X}}_i = [\bar{\mathbf{x}}_{i,1}, \bar{\mathbf{x}}_{i,2}, \dots, \bar{\mathbf{x}}_{i,\tilde{m}\tilde{n}}] \in \mathbf{R}^{k_1 \times k_2 \times \tilde{m}\tilde{n}}, \quad (1)$$

其中: $\tilde{m}=m-k_1+1$, $\tilde{n}=n-k_2+1$; $\bar{\mathbf{x}}_{i,\tilde{m}\tilde{n}}$ 表示第 i 张图片 \mathbf{M}_i 中第 k 块的采样结果。对所有输入图像进行相同操作可得

$$\mathbf{X} = [\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_N] \in \mathbf{R}^{k_1 \times k_2 \times N\tilde{m}\tilde{n}}. \quad (2)$$

1.1 PCANet卷积层

假设第1层卷积滤波器数量为 L_1 ,根据最小化重构误差原理,PCA滤波器的目标函数如下所示

$$\min_{V \in \mathbf{R}^{k_1 \times k_2 \times L_1}} \| \mathbf{X} - \mathbf{V}\mathbf{V}^T \mathbf{X} \|_{\text{F}}^2, \text{ s.t. } \mathbf{V}^T \mathbf{V} = \mathbf{E}, \quad (3)$$

其中: \mathbf{E} 表示大小为 $L_1 \times L_1$ 的单位矩阵;矩阵 \mathbf{V} 由 $\mathbf{X}\mathbf{X}^T$ 的 L_1 个特征向量组成。由此,PCANet第1层的滤波器组可以表示为

$$\mathbf{W}_l^1 = \text{mat}_{k_1, k_2}(\mathbf{q}_l(\mathbf{X}\mathbf{X}^T)) \in \mathbf{R}^{k_1 \times k_2}, l = 1, 2, \dots, L_1, \quad (4)$$

其中: $\text{mat}_{k_1, k_2}(\mathbf{q}_l(\mathbf{X}\mathbf{X}^T))$ 是将 $\mathbf{X}\mathbf{X}^T$ 的第 l 个主要特征向量 $\mathbf{q}_l(\mathbf{X}\mathbf{X}^T)$ 重塑为 $k_1 \times k_2$ 大小矩阵的函数,主成分分析用来提取去均值后采样块的主要信息。对图像 \mathbf{M}_i 进行卷积滤波,得到 L_1 个输出映射

$$\mathbf{M}_{i,l}^1 = \mathbf{M}_i * \mathbf{W}_l^1, i = 1, 2, \dots, N, l = 1, 2, \dots, L_1, \quad (5)$$

其中: $*$ 表示2D卷积运算。对 N 张图像进行相同操作,得到PCANet第1层的卷积输出。此外,在进行卷积滤波前要对图像进行补零操作,使卷积后的图像与输入图像尺寸相同。

第2层的卷积过程与第1层相似,不同点是第2层的输入是第1层的映射输出图像 $\mathbf{M}_{i,l}^1$ 。假设第2层的滤波器数量为 L_2 ,最后得到PCANet的卷积层输出为

$$\mathbf{O}_{i,l,b}^2 = \{ \mathbf{M}_{i,l}^1 * \mathbf{W}_b^2 \}, l = 1, 2, \dots, L_1, b = 1, 2, \dots, L_2. \quad (6)$$

1.2 PCANet输出层

在输出阶段进行二进制散列并计算块状直方图以特征编码。第2阶段卷积得到的输出矩阵 $\mathbf{O}_{i,l,b}^2$ 包含实数。为了将它们转换为整数,首先,使用Heaviside阶跃函数对第2阶段的输出进行二进制化,得到 $H(\mathbf{O}_{i,l,b}^2)$,

$H(\cdot)$ 为Heaviside阶跃函数,将正值映射为1反之为0。对于每个像素,把 L_2 个二进制位的向量看作是1个十进制数。接下来利用二进制散列法将图像 $\mathbf{O}_{i,l,b}^2$ 的 L_2 个输出转换为1个整数,如下所示

$$\mathbf{F}_i^l = \sum_{b=1}^{L_2} 2^{b-1} H(\mathbf{O}_{i,l,b}^2), \quad (7)$$

其中,每个像素都是 $[0, 2^{L_2-1}]$ 之间的整数。

在二进制散列步骤之后,计算块状的直方图。对于每个 L_1 图像 \mathbf{F}_i^l ($i = 1, 2, \dots, L_1$),将其划分为 P 块,每块大小为 $p_1 \times p_2$,这些块可能有重叠。然后统计每块十进制的直方图,并将所有 P 个直方图串联成1个向量,表示为 $\text{Bhist}(\mathbf{F}_i^l)$ 。经过编码后,输入图像 $\text{Bhist}(\mathbf{F}_i^l)$ 的特征被定义为块直方图的集合

$$\mathbf{F}_i = [\text{Bhist}(\mathbf{F}_i^1), \text{Bhist}(\mathbf{F}_i^2), \dots, \text{Bhist}(\mathbf{F}_i^{L_1})]^T \in \mathbb{R}^{(2^{L_2})L_1P}. \quad (8)$$

2 基于非线性深度子空间学习的微表情识别方法

基于PCANet提出一种基于光流计算和非线性深度子空间学习的微表情识别方法。首先,从输入的微表情视频片段中计算光流以获取面部运动信息;然后,将得到的光流特征图输送到2层核主成分分析网络(kernel principal component analysis network, KPCANet)中学习更丰富的时空特征;最后,采用线性支持向量机(support vector machine, SVM)实现微表情的分类。图1展示了基于KPCANet的微表情识别方法框架,该方法包括:数据预处理、光流特征计算、核主成分分析网络模型特征提取、分类,共4个步骤,接下来进行详细介绍。

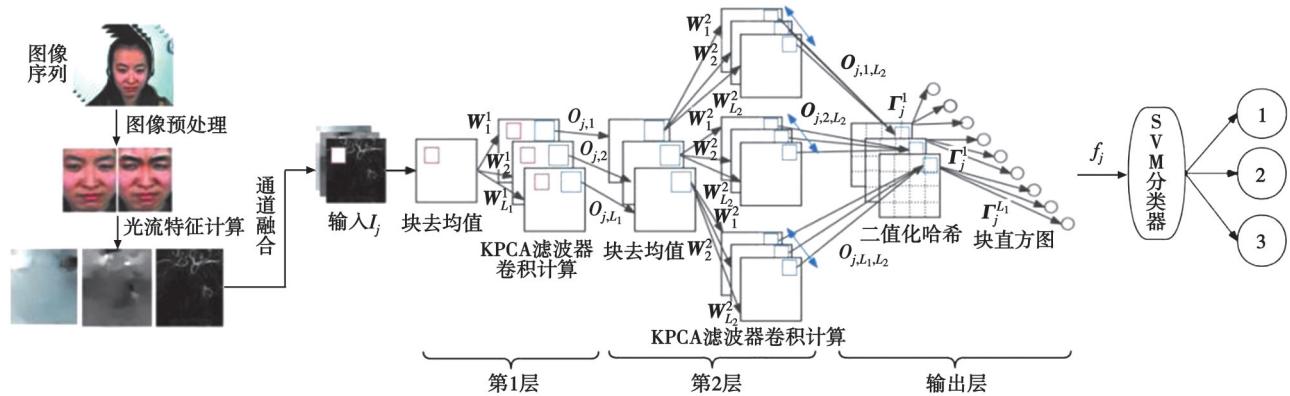


图1 基于KPCANet的微表情识别方法框架

Fig. 1 The framework of the micro-expression recognition method based on KPCANet

2.1 数据预处理

首先对微表情视频帧序列进行图像数据预处理,分为空间和时序2个方面。在空间上完成从原始数据到裁剪后的人脸图像序列,包括面部关键点检测、面部对齐和裁剪、尺度归一化等;其次,对视频帧进行面部运动放大处理,凸显和放大微表情的脸部动作^[24],有利于后续提取面部的微表情信息。

2.2 光流特征计算

光流可以很好地反映脸部的运动区域,也能在一定程度上抑制人脸身份信息带来的影响。通过光流计算获取面部运动信息,得到微表情动态的时空特征。考虑微表情图像序列中脸部的动作变化极度细微,且2个相邻帧之间的图像变化一般很难感知。如果以所有图像作为输入,会包含大量多余特征。因此,仅使用顶点帧^[25]和起始帧计算光流特征,可有效降低输入特征的维度。

假设对于第*i*个微表情视频序列 $\mathbf{S}_i = \{s_{i,a} | i = 1, 2, \dots, N, a = 1, 2, \dots, n\}$,共包括 n_i 帧图像,起始帧和顶点帧分别表示为 $s_{i,1}$ 和 $s_{i,n}$ 。令起始帧 $s_{i,1}$ 的亮度表示为 $B_{i,1}(x, y)$,顶点帧 $s_{i,n}$ 的亮度表示为 $B_{i,n}(x, y)$ 。根据亮度恒定性约束可得

$$B_{i,1}(x, y) = B_{i,n}\left(x + \frac{dx}{dt}, y + \frac{dy}{dt}\right), \quad (9)$$

对上面等式右侧部分采用泰勒级数展开,得到

$$B_{i,a}\left(x+\frac{dx}{dt}, y+\frac{dy}{dt}\right) \approx B_{i,1}(x, y) + \frac{\partial B}{\partial x} dx + \frac{\partial B}{\partial y} dy + \frac{\partial B}{\partial t} dt, \quad (10)$$

结合公式(9)(10),光流方程可以表示为

$$B_{i,1}(x, y) = B_{i,1}(x, y) + \frac{\partial B}{\partial x} dx + \frac{\partial B}{\partial y} dy + \frac{\partial B}{\partial t} dt, \quad (11)$$

令 $u(x, y) = \frac{dx}{dt}$ 和 $v(x, y) = \frac{dy}{dt}$ 分别表示光流向量场的水平分量和垂直分量。整理得到

$$u(x, y) \frac{\partial B}{\partial x} + v(x, y) \frac{\partial B}{\partial y} + \frac{\partial B}{\partial t} = 0, \quad (12)$$

使用经典的TV-L₁方法计算得到光流的水平分量 u 和垂直分量 v 。这一方法能更好地保持光流场的不连续性,并对照明变化、遮挡和噪声具有鲁棒性。此外,采用光学应变 ε 估算脸部动作变化的幅度大小,其定义如下

$$\varepsilon = \frac{1}{2} \left[\nabla \omega + (\nabla \omega)^T \right] = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \\ \frac{1}{2} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) & \frac{\partial v}{\partial y} \end{bmatrix}, \quad (13)$$

式中: $\omega = [u, v]^T$ 表示位移矢量。

从而,每个像素的光学应变的强度可以通过取法向应变分量和剪切应变分量的平方和来计算,得到计算公式如下

$$|\varepsilon(x, y)| = \sqrt{\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 + \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2}. \quad (14)$$

因此,每个视频 S_i 经过光流计算可以得到 1 个三元组的光流特征表示为

$$\mathbf{R}_i = \{ (u(x, y), v(x, y), \varepsilon(x, y)) | x = 1, 2, \dots, H, y = 1, 2, \dots, W \}, \quad (15)$$

式中: W 和 H 分别表示视频图像帧的宽度和高度,而 $u(x, y)$ 、 $v(x, y)$ 和 $\varepsilon(x, y)$ 分别表示光流场的水平分量、垂直分量和光学应变强度。经过第 1 阶段的光流计算,每个视频 S_i 生成了 1 张大小为 112×112 、通道数为 3 的光流特征图 \mathbf{F}_i 。

2.3 核主成分分析网络模型

KPCANet 与 PCANet 的整体结构相同,即图像样本在经过 2 层相连的卷积运算后进行二值哈希处理以及直方图统计,得到微表情图像特征的最终表示。不同之处在于学习卷积核之前,对样本数据进行核变换。KPCANet 利用核函数将输入数据映射到 1 个高维的特征空间。在这个高维空间中,原本线性不可分的数据可能变得线性可分,提升特征表达能力与识别效果。

对于大小为 $m \times n$ 的 N 幅输入光流图像 $\{ \mathbf{F}_i \} (i=1, 2, \dots, N)$,仍设 KPCANet 的采样块大小为 $k_1 \times k_2$,依次将其按卷积核大小进行取块、排列、去均值,可以得到重排后的训练样本

$$\bar{\mathbf{X}}_i = [\bar{\mathbf{x}}_{i,1}, \bar{\mathbf{x}}_{i,2}, \dots, \bar{\mathbf{x}}_{i,\tilde{m}\tilde{n}}] \in \mathbf{R}^{k_1 k_2 \times \tilde{m}\tilde{n}}, \quad (16)$$

然后,对每个 $\bar{\mathbf{X}}_i$ 的行向量进行核变换处理,这里引入多项式核函数作映射,计算公式如下

$$K(\mathbf{v}_p, \mathbf{v}_q) = \left((\mathbf{v}_p)^T \mathbf{v}_q + 1 \right)^d \in \mathbf{R}^{k_1 k_2 \times k_1 k_2}, \quad (17)$$

式中: $\mathbf{v}_p, \mathbf{v}_q \in \mathbf{R}^{\tilde{m}\tilde{n}}$ 是矩阵 $\bar{\mathbf{X}}_i$ 的第 p 和第 q 个行向量; d 为实数参数。令 \mathbf{K}_i 表示 $\bar{\mathbf{X}}_i$ 经过核变换后的矩阵,然后将所有的图像经过相同的处理得到对应的核矩阵,级联所有核矩阵得到

$$\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2, \dots, \mathbf{K}_N] \in \mathbf{R}^{k_1 k_2 \times N k_1 k_2}, \quad (18)$$

至此已完成了所有数据的核变换,将原始空间的数据投影到高维特征空间。假设第 1 层滤波器的数量为 L_1 ,下一步是在高维特征空间使用 PCA 算法进行降维,与 PCANet 相同,通过使用 PCA 目标函数学习投影矩阵,即

$$\min_{V \in \mathbb{R}^{k_1 \times L_1}} \|K - VV^T K\|_F^2, \text{ s.t. } V^T V = E, \quad (19)$$

式中: E 表示大小为 $L_1 \times L_1$ 的单位矩阵,矩阵 V 由 KK^T 的 L_1 个主特征向量组成。因此,第1层的KPCA滤波器可以表示为

$$W_a^1 = \text{mat}_{k_1, k_2}(q_a(KK^T)) \in \mathbb{R}^{k_1 \times k_2}, a = 1, 2, \dots, L_1, \quad (20)$$

对图像 F_i 进行卷积滤波,得到 L_1 个输出映射

$$F_{i,a}^1 = F_i * W_a^1, i = 1, 2, \dots, N, a = 1, 2, \dots, L_1. \quad (21)$$

与PCANet同理,将第1层的映射输出图像作为第2层的输入,得到KPCANet的卷积层输出,最后进行二进制散列与哈希编码后得到输入图像的特征表示。

2.4 分类

利用线性支持向量机(SVM)对经过核主成分分析网络处理的特征进行微表情分类识别。将学习到的特征和对应标签输入到SVM分类器中,选择线性核函数,在高维特征空间中构建有效的决策边界,通过调整超参数(如正则化参数),优化模型的性能和泛化能力。通过对训练数据集的学习,找到1个最优的超平面,将不同类别的微表情样本正确分离。在分类时,训练好的模型能够对新样本进行准确分类,最小化分类误差,提高系统对微表情的识别精度和鲁棒性。

3 实验结果与分析

为验证提出算法的性能,在SMIC^[26]、CASME^[27]、CASME II^[28]和SAMM^[29]4个广泛使用的单一微表情数据集和3DB-combined^[30]复合数据集上进行测试,并将实验结果与现有的2大类微表情识别算法(包括传统手工特征提取方法中的LBP-TOP^[5]、3DHOG^[7]、HOG^[8]与深度学习方法中的OFF-ApexNet^[31]、STSTNet^[13]、Dual-Inception^[32]、MACNN^[33]、Micro-Attention^[34]和Mini-AORCNN^[35],以及深度子空间学习方法PCANet)进行对比。实验软件环境为Windows10、Microsoft Visual Studio2015、python3.8.16和OpenCV4.5.5,采用Intel(R)Xeon(R)Platinum8 260 M(2.3 GHz)处理器,32 GB内存的硬件环境。

3.1 数据集介绍

SMIC数据集中图像样本的大小为320 pix×240 pix,包含16个受试者的164个样本,这些样本被分为3类:negative(70)、positive(51)和surprise(43)。CASME数据集中图像样本的大小为640 pix×480 pix,一共包含19个受试者的195个微表情样本,但类间样本分布不均衡。研究选用了154个样本,分4类:disgust(40)、repression(33)、surprise(18)和tense(63)。CASME II数据集中图像样本的大小为640 pix×480 pix,该数据集是为补充样本容量而改进CASME数据集的版本,其实验条件和CASME相近,但CASME II具有更高的时间和空间分辨率。这个数据集包含26个受试者的248个样本,共分为5类:disgust(62)、happiness(32)、repression(27)、surprise(28)和others(99)。SAMM数据集中图像样本的大小为256 pix×256 pix,包含32个受试者的159个样本,其中受试者来自13个不同的种族,男女受试者的比例为1:1。剔除部分样本数量小于10的小类别(如fear, sadness等),筛选出SAMM数据集中可用的134个样本,分5类:anger(56)、contempt(11)、happiness(26)、surprise(15)和other(26)。3DB-combined复合数据集是由SMIC、CASME II和SAMM 3个单一数据集组合而成,根据原始情感类别分为3大类:负面(即repression、anger、contempt、disgust、fear和sadness)、正面(happiness)、惊讶(surprise)。该数据集共包含68个受试者的442个样本,negative(250)、positive(109)、surprise(83)。

3.2 实验设置

为避免分类过程中的受试者依赖问题,在所有实验中采用留一受试者交叉验证^[36](leave-one-subject-out cross validation, LOSOCV)协议来计算,即每次选出1个受试者数据作为实验的测试集,其余受试者数据作为训练集,重复此过程,直至所有受试者都轮过测试集为止。这种验证方式能更好地模拟现实场景,评估对种族、性别、情绪敏感性等可变因素的稳健性。对单一数据集,使用平均识别准确率(accuracy)和 F_1 分数(F_1 -score)2个指标来报告微表情识别方法的性能。而复合数据集3DB-combined是通过合并SMIC、CASME II和SAMM 3个单一数据集获得,在类别分布方面明显不平衡,即惊讶:正面:负面类别的分布比例为1:1.3:3。为了正确处理这种类别不平衡,研究将使用未加权 F_1 分数(UF_1),也称为宏观 F_1 分数,和未加权平均召回率

(unweighted average recall, UAR)来评估各个算法的性能。准确率、 F_1 分数、未加权 F_1 分数和未加权平均召回率的计算公式如下

$$\text{Accuracy} = \sum_{i=1}^c \sum_{j=1}^s \frac{(TP_i^j + TN_i^j)}{(TP_i^j + TN_i^j + FP_i^j + FN_i^j)}, \quad (22)$$

$$\text{Recall} = \sum_{i=1}^c \sum_{j=1}^s \frac{TP_i^j}{(TP_i^j + FN_i^j)}, \quad (23)$$

$$\text{Precision} = \sum_{i=1}^c \sum_{j=1}^s \frac{TP_i^j}{(TP_i^j + FP_i^j)}, \quad (24)$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (25)$$

$$UF_1 = \frac{1}{c} \sum_{i=1}^c \frac{2 \sum_{j=1}^s TP_i^j}{2 \sum_{j=1}^s TP_i^j + \sum_{j=1}^s FP_i^j + \sum_{j=1}^s FN_i^j}, \quad (26)$$

$$\text{UAR} = \frac{1}{c} \sum_{i=1}^c \frac{\sum_{j=1}^s TP_i^j}{\sum_{j=1}^s TP_i^j + \sum_{j=1}^s FP_i^j + \sum_{j=1}^s FN_i^j}, \quad (27)$$

式中: c 表示数据集中情绪类别的数量; s 表示受试者的数量; TP_i^j , FP_i^j , TN_i^j 和 FN_i^j 分别代表第 i 类情绪中第 j 位受试者的真阳性(true positives)、假阳性(false positives)、真阴性(true negatives)和假阴性(false negatives)的数量。

3.3 分类率实验

为保证结果的公平性与一致性,所有方法的实验设置都是相同的,包括相同的样本数量和类标签,以及统一使用LOSO CV验证方法。表1~4报告了所有方法在4个单一微表情数据集上的性能比较结果。从结果可以看出,所提出的方法显然超过了3种微表情识别的基线方法,在SMIC、CASME、CASME II和SAMM数据库中的 F_1 分数分别比LBP-TOP基线平均提高约30%、26%、29%、17%。与最先进的方法相比,所提出的KPCANet算法取得了较好性能,并且比大多数现有算法的性能更好,验证了非线性深度子空间网络算法在微表情识别任务上的有效性。对比4个数据集的分类结果可以发现,SMIC、CASME和CASME II的最佳识别准确率均达到70%以上,而SAMM数据集却只有63.83%。这可能是因为SAMM数据集存在较为明显的类别不平衡问题,比如contempt和surprise的样本量分别只有11和15,只占总数的8%和11%左右,类别分布不均导致识别性能明显降低。

表1 SMIC数据集上不同方法的平均识别精度和 F_1

Table 1 Average recognition accuracy and F_1 of different methods on the SMIC dataset

方法	精度	F_1
LBP-TOP	0.414 1	0.404 9
3DHOG	0.328 0	0.344 1
HOOF	0.508 5	0.501 1
OFF-ApexNet	0.616 7	0.577 3
STSTNet	0.607 6	0.635 1
Dual-Inception	0.613 9	0.577 6
MACNN	0.710 6	0.644 9
Micro-Attention	0.690 7	0.624 8
Mini-AORCNN	0.674 8	0.601 3
PCANet	0.623 4	0.609 3
KPCANet	0.717 4	0.710 0

注:加粗数据表示该方法在准确率和 F_1 指标上表现最佳。

表2 CASME数据集上不同方法的平均识别精度和 F_1 Table 2 Average recognition accuracy and F_1 of different methods on the CASME dataset

方法	精度	F_1
LBP-TOP	0.472 8	0.363 8
3DHOG	0.478 9	0.342 9
HOOF	0.490 1	0.345 9
OFF-ApexNet	0.628 9	0.537 9
STSTNet	0.647 7	0.520 3
Dual-Inception	0.617 5	0.511 8
MACNN	0.724 4	0.671 4
Micro-Attention	0.752 2	0.665 6
Mini-AORCNN	0.739 7	0.669 8
PCANet	0.725 9	0.588 2
KPCANet	0.757 9	0.624 3

注:加粗数据表示该方法在准确率和 F_1 指标上表现最佳。

表3 CASME II数据集上不同方法的平均识别精度和 F_1 Table 3 Average recognition accuracy and F_1 of different methods on the CASME II dataset

方法	精度	F_1
LBP-TOP	0.460 1	0.361 8
3DHOG	0.401 0	0.300 9
HOOF	0.428 2	0.337 0
OFF-ApexNet	0.537 0	0.462 2
STSTNet	0.598 7	0.538 5
Dual-Inception	0.538 0	0.475 1
MACNN	0.656 7	0.658 8
Micro-Attention	0.651 9	0.624 3
Mini-AORCNN	0.661 0	0.626 2
PCANet	0.654 1	0.577 8
KPCANet	0.704 6	0.659 2

注:加粗数据表示该方法在准确率和 F_1 指标上表现最佳。

表4 SAMM数据集上不同方法的平均识别精度和 F_1 Table 4 Average recognition accuracy and F_1 of different methods on the SAMM dataset

方法	精度	F_1
LBP-TOP	0.440 3	0.341 8
3DHOG	0.450 8	0.343 7
HOOF	0.468 1	0.368 5
OFF-ApexNet	0.422 6	0.314 2
STSTNet	0.502 5	0.353 6
Dual-Inception	0.468 1	0.373 8
MACNN	0.557 0	0.448 9
Micro-Attention	0.563 4	0.466 5
Mini-AORCNN	0.567 1	0.472 3
PCANet	0.584 4	0.460 5
KPCANet	0.638 3	0.521 5

注:加粗数据表示该方法在准确率和 F_1 指标上表现最佳。

为了进一步探讨KPCANet算法在4个微表情数据集上的识别性能,图2展示了相关的混淆矩阵。对比分析显示,SMIC数据集的分类精度比较稳定,可能由于其类别较少且在整体上分布均衡。然而,在SAMM数据集中,由于类别分布不均衡,导致整体分类效果较差,尤其是在样本数量非常少的contempt类和surprise类中表现不佳。此外,在CASME和CASME II数据集中,样本数量较多的类别表现出色,例如,CASME中的tense类和CASME II中的others类,分别达到了81%和82%的准确率。相反,样本数量不足的小类别表现较差,例如,CASME II中的repression类只有33%的准确率。由此可见,微表情数据集中样本数量及类别分布的平衡性仍是影响模型分类性能的主要因素。

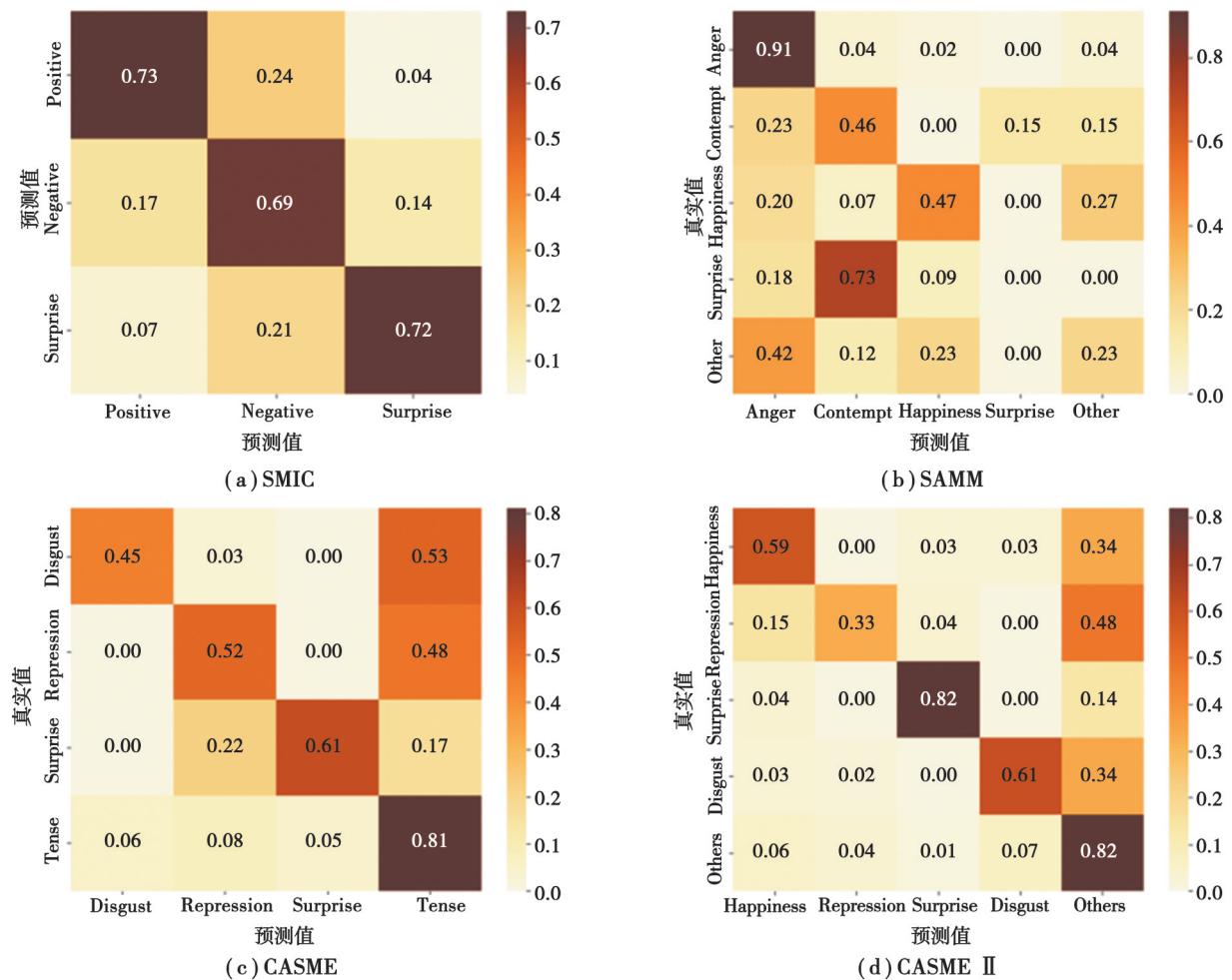


图2 4个微表情数据集的混淆矩阵
Fig. 2 Confusion matrices of four ME datasets

3.4 泛化能力和鲁棒性实验

在3DB-combined复合数据集上将所提出方法与其他微表情识别方法进行比较。该模式模拟了1个真实的情景,即来自不同背景(种族、性别、情感敏感性)的人分别在不同的环境和设置中注册到1个单一的识别系统中,实验结果如表5所示。图3展示了3DB-combined数据集上不同方法的表现,从结果看出,KPCANet算法在单一数据集上和组合而成的复合数据集上均展现出稳定的未加权 F_1 分数,说明其受数据集的多样性影响较小,具有较好的泛化能力,能够应用于真实世界中的各种复杂场景。

为验证所提方法的鲁棒性,通过引入噪声,在原始微表情数据集上进行遮挡实验。在SMIC数据集的微表情序列上合成了随机区域噪声,噪声块大小比例从0~40%不等。考虑到噪声区域的随机性,实验在噪声遮挡条件下重复进行5次,取平均值作为最终评估结果。表6报告了在SMIC数据集中不同噪声遮挡区域条件下各种方法的性能比较结果。图4显示了在不同噪声条件下,通过重复5次实验所得到的8种算法的平均分

类率。结果表明,随着噪声块的增加,KPCNet算法的识别性能较稳定,与其他微表情识别方法相比,提出的方法可有效减轻噪声遮挡对微表情识别的影响,具有较好的鲁棒性。

表5 3DB-combined数据集上不同方法的表现

Table 5 Performance of different methods on the 3DB-combined dataset

方法	Composite			SMIC		CASME II		SAMM	
	Acc	UF_1	UAR	UF_1	UAR	UF_1	UAR	UF_1	UAR
LBP-TOP	0.647	0.588 2	0.578 5	0.200 0	0.528 0	0.702 6	0.742 9	0.395 4	0.410 2
Bi-WooF	0.683 3	0.629 6	0.622 7	0.572 7	0.582 9	0.780 5	0.802 6	0.521 1	0.513 9
OFF-ApexNet	0.666	0.719 6	0.709 6	0.681 7	0.669 5	0.876 4	0.868 1	0.540 9	0.539 2
STSTNet	0.723 6	0.735 3	0.760 5	0.680 1	0.701 3	0.838 2	0.868 6	0.658 8	0.681 0
Dual-Inception	0.617 1	0.732 2	0.727 8	0.664 5	0.672 6	0.862 1	0.856 0	0.586 8	0.566 3
MACNN	0.761 3	0.702 8	0.703 4	0.635 2	0.638 6	0.843 8	0.852 2	0.585 9	0.574 4
Micro-Attention	0.790 7	0.732 7	0.727 1	0.691 2	0.697	0.883 1	0.880 6	0.523 6	0.508 5
Mini-AORCNN	0.754 2	0.706 4	0.701 3	0.646 5	0.653	0.831 4	0.818 1	0.555 2	0.535 9
PCANet	0.765 9	0.701 6	0.692 1	0.678 4	0.675 5	0.801 1	0.791 3	0.548 5	0.535 6
KPCANet	0.834 6	0.763 2	0.743 6	0.700 5	0.704 1	0.886 8	0.869 0	0.666 0	0.675 3

注:加粗数据表示该方法在对应性能指标上表现最佳。

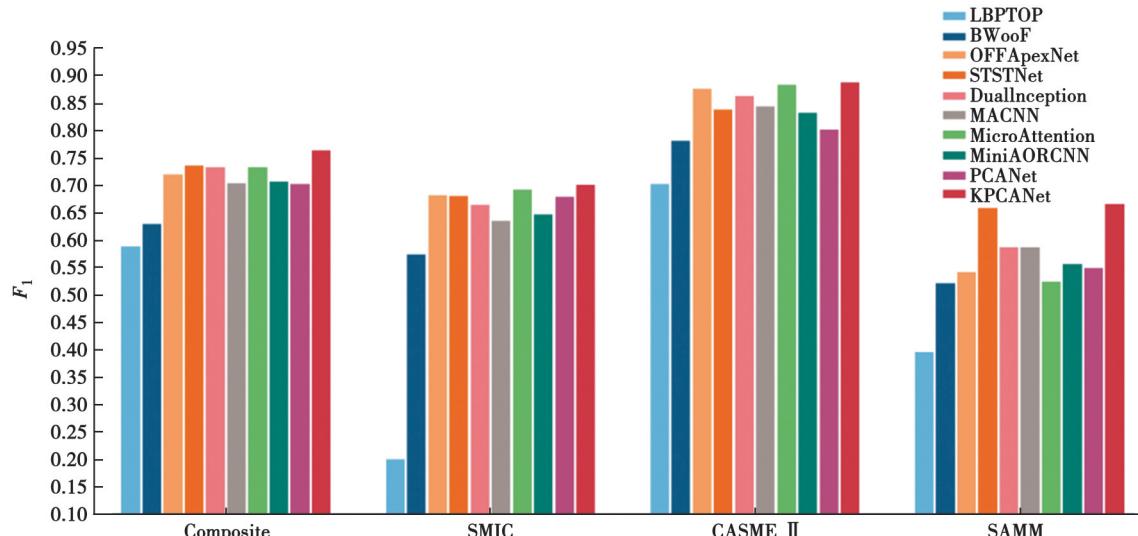
图3 3DB-combined数据集上不同算法的 F_1 分数Fig. 3 UF_1 scores of different algorithms on the 3DB-combined dataset

表6 SMIC数据集上不同噪声遮挡条件下各类方法的表现

Table 6 Performance of various methods under different noise occlusion conditions on the SMIC dataset

Noise	算法	Accuracy	UF_1	UAR	Recall
Un-occ	OFF-ApexNet	0.616 7	0.577 3	0.579 8	0.579 3
	STSTNet	0.581	0.604	0.603 6	0.603 7
	Dual-Inception	0.613 9	0.577 6	0.573 8	0.579 3
	MACNN	0.710 6	0.644 9	0.650 6	0.640 2
	Micro-Attention	0.690 7	0.624 8	0.628 5	0.622
	Mini-AORCNN	0.674 8	0.601 3	0.610 1	0.591 5
	PCANet	0.623 4	0.609 3	0.605 4	0.603 7

续表6

Noise	算法	Accuracy	UF_1	UAR	Recall
	KPCANet	0.717 4	0.710 0	0.710 7	0.707 3
Occ-10%	OFF-ApexNet	0.544 4	0.540 1	0.544 3	0.541 5
	STSTNet	0.520 8	0.545 2	0.543 9	0.550 0
	Dual-Inception	0.613 7	0.587 1	0.595 0	0.581 7
	MACNN	0.559 6	0.557 4	0.561 0	0.557 7
	Micro-Attention	0.651 2	0.603 0	0.611 7	0.597 6
	Mini-AORCNN	0.630 4	0.589 5	0.597 4	0.583 0
	PCANet	0.622 5	0.611 6	0.604 2	0.608 5
	KPCANet	0.648 9	0.649 8	0.649 2	0.646 3
Occ-20%	OFF-ApexNet	0.521 2	0.532 1	0.533 7	0.536 6
	STSTNet	0.497 3	0.537 5	0.536 2	0.543 9
	Dual-Inception	0.491 7	0.522 8	0.525 0	0.524 4
	MACNN	0.609 7	0.580 4	0.583 7	0.575 6
	Micro-Attention	0.625 1	0.598 4	0.606 6	0.591 5
	Mini-AORCNN	0.624 5	0.587 9	0.597 1	0.581 7
	PCANet	0.614 3	0.610 5	0.604 9	0.607 3
	KPCANet	0.643 1	0.648 6	0.644 3	0.641 4
Occ-30%	OFF-ApexNet	0.509 3	0.510 5	0.510 8	0.515 9
	STSTNet	0.497 0	0.535 5	0.543 2	0.537 8
	Dual-Inception	0.490 4	0.507 7	0.509 7	0.511 0
	MACNN	0.579 3	0.554 8	0.562 4	0.551 2
	Micro-Attention	0.630 7	0.589 6	0.598 7	0.584 1
	Mini-AORCNN	0.603 5	0.571 5	0.577 0	0.563 4
	PCANet	0.608 3	0.611 7	0.604 2	0.609 8
	KPCANet	0.627 6	0.642 1	0.636 8	0.639 0
Occ-40%	OFF-ApexNet	0.491 3	0.528 4	0.529 1	0.531 7
	STSTNet	0.480 5	0.521 3	0.520 5	0.526 8
	Dual-Inception	0.463 1	0.512 6	0.514 4	0.511 0
	MACNN	0.578 9	0.545 5	0.552 3	0.539 0
	Micro-Attention	0.606 1	0.587 0	0.592 7	0.582 9
	Mini-AORCNN	0.597 2	0.559 2	0.567 9	0.551 2
	PCANet	0.604 5	0.585 8	0.581 5	0.587 8
	KPCANet	0.619 1	0.612 6	0.607 7	0.611 0

注:加粗数据表示该方法在对应性能指标上表现最佳。

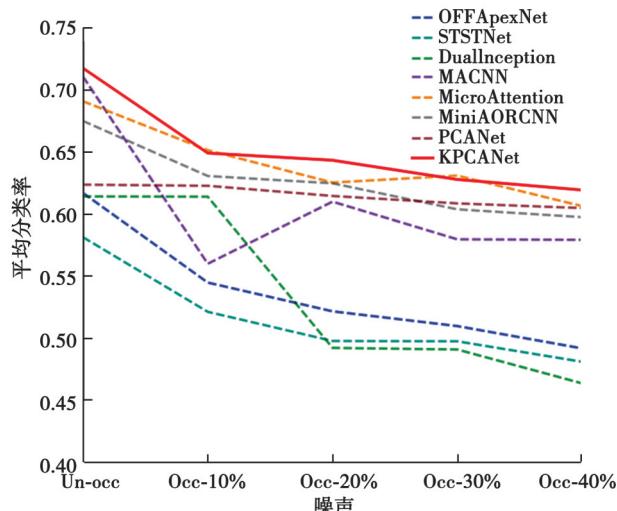


图4 不同噪声下的平均分类率

Fig. 4 Average classification rate under different noise levels

4 结 论

针对微表情识别任务中处理小样本困难以及模型鲁棒性差的问题,通过将传统手工特征提取方法和非线性深度子空间网络模型相结合,提出一种基于非线性深度子空间网络的微表情识别方法:首先,利用光流计算提取面部运动信息;再将得到的光流特征图输入KPCANet中,进一步增强和丰富微表情的表征特征;最后,采用线性支持向量机进行分类。实验结果表明所提出方法的识别性能优于MACNN、Micro-Attention等深度学习算法,并且取得较好的泛化能力和鲁棒性。尽管KPCANet在多样性数据集上表现出较好的鲁棒性,但在处理类不平衡情况时的性能不佳。在今后的研究工作中,将尝试在内核PCA方法中使用平衡样本集,即对小样本类别进行过采样和对大样本类别进行欠采样,提升微表情识别精度。

参考文献

- [1] Li J T, Wang T, Wang S J. Facial micro-expression recognition based on deep local-holistic network[J]. Applied Sciences, 2022, 12(9): 4643.
- [2] Li Y T, Wei J S, Liu Y, et al. Deep learning for micro-expression recognition: a survey[J]. IEEE Transactions on Affective Computing, 2022, 13(4): 2028-2046.
- [3] Ben X Y, Ren Y, Zhang J P, et al. Video-based facial micro-expression analysis: a survey of datasets, features and algorithms[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(9): 5826-5846.
- [4] Zhou L, Mao Q R, Huang X H, et al. Feature refinement: an expression-specific feature learning and fusion method for micro-expression recognition[J]. Pattern Recognition, 2022, 122: 108275.
- [5] Zhao G Y, Pietikäinen M. Dynamic texture recognition using local binary patterns with an application to facial expressions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 915-928.
- [6] Liu B, Sun Y F, Yu A Z, et al. Hyperspectral meets optical flow: spectral flow extraction for hyperspectral image classification [J]. IEEE Transactions on Image Processing, 2023, 32: 5181-5196.
- [7] Khan S A, Bin Abdulrahman Alawairdh M, AL-Akhras M. Texture and orientation-based feature extraction for robust facial expression recognition[C]//2023 IEEE/ACIS 21st International Conference on Software Engineering Research, Management and Applications. Orlando, USA: IEEE, 2023: 428-433.
- [8] Mu X F, Liu J J, Li Y, et al. A dual-network micro-expression recognition model based on optical flow features[C]//2022 8th International Conference on Big Data Computing and Communications. Xiamen, China: IEEE, 2022: 199-205.
- [9] Jin H M, He N, Li Z L, et al. Micro-expression recognition based on multi-scale 3D residual convolutional neural network[J]. Mathematical Biosciences and Engineering, 2024, 21(4): 5007-5031.
- [10] Li J, Wang T, Wang S J. Facial Micro-Expression Recognition Based on Deep Local-Holistic[J]. Applied Sciences, 2022, 12(9):

- 4643.
- [11] Teja Reddy S P, Teja Karri S, Dubey S R, et al. Spontaneous facial micro-expression recognition using 3D spatiotemporal convolutional neural networks[C]//2019 International Joint Conference on Neural Networks (IJCNN). Budapest, Hungary: IEEE, 2019: 1-8.
- [12] Khor H Q, See J, Phan R C W, et al. Enriched long-term recurrent convolutional network for facial micro-expression recognition[C]//2018 13th IEEE International Conference on Automatic Face & Gesture Recognition. Xi'an, China: IEEE, 2018: 667-674.
- [13] Liang S T, Gan Y S, See J, et al. Shallow triple stream three-dimensional cnn (ststnet) for micro-expression recognition[EB/OL]. (2019-08-21)[2024-05-12]. [Https://ieeexplore.ieee.org/document/8756567](https://ieeexplore.ieee.org/document/8756567).
- [14] Zhang X F, Fu X K, Qi G B, et al. A multi-scale feature fusion convolutional neural network for facial expression recognition [J]. Expert Systems, 2024, 41(4): e13517.
- [15] 刘东, 王生生. 精细拓扑结构表示与深度特征融合在多目标图像检索中的应用[J]. 重庆大学学报, 2021, 44(3): 132-143.
Liu D, Wang S S. On the application of detailed topological structure representation and deep feature fusion to multi-object image retrieval[J]. Journal of Chongqing University, 2021, 44(3): 132-143. (in Chinese)
- [16] Li J T, Dong Z Z, Lu S Y, et al. CAS(ME)3: a third generation facial spontaneous micro-expression database with depth information and high ecological validity[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 2782-2800.
- [17] Peng M, Wu Z, Zhang Z H, et al. From macro to micro expression recognition: deep learning on small datasets using transfer learning[C]//2018 13th IEEE International Conference on Automatic Face&Gesture Recognition. Xi'an, China: IEEE, 2018: 657-661.
- [18] Gan Y, et al. Micro-expression recognition using a generative adversarial network with dual discriminators[J]. IEEE Transactions on Multimedia, 2019, 21(8): 1925-1938.
- [19] Happy S L, Routray A. Fuzzy histogram of optical flow orientations for micro-expression recognition[J]. IEEE Transactions on Affective Computing, 2019, 10(3): 394-406.
- [20] Chengeta K. Automated facial micro-expression recognition using local binary patterns on three orthogonal planes with boosted classifiers: a survey[M]//Advances in Computer Communication and Computational Sciences. Singapore: Springer Singapore, 2019: 671-686.
- [21] Yu X H, Xie Z H, Zong W J. Dual-cross patterns with RPCA of key frame for facial micro-expression recognition[M]//Image and Graphics. Cham: Springer International Publishing, 2019: 750-759.
- [22] Wang S Q, Guan S E, Lin H, et al. Micro-expression recognition based on optical flow and PCANet[J]. Sensors, 2022, 22(11): 4296.
- [23] 李赛宇, 鞠传香, 丁航奇. 基于Iterative映射和非线性拟合的鲸鱼优化算法[J]. 重庆大学学报, 2023, 46(8): 120-131.
Li S Y, Ju C X, Ding H Q. Whale optimization algorithm based on iterative mapping and nonlinear fitting[J]. Journal of Chongqing University, 2023, 46(8): 120-131. (in Chinese)
- [24] 张洋铭, 吴凯, 王艺凡, 等. 基于随机权重分配策略的面目表情识别[J]. 重庆大学学报, 2022, 45(9): 135-140.
Zhang Y M, Wu K, Wang Y F, et al. Facial expression recognition based on random weight assignment strategy[J]. Journal of Chongqing University, 2022, 45(9): 135-140. (in Chinese)
- [25] Li Y T, Huang X H, Zhao G Y. Joint local and global information learning with single apex frame detection for micro-expression recognition[J]. IEEE Transactions on Image Processing, 2021, 30: 249-263.
- [26] Li X B, Pfister T, Huang X H, et al. A spontaneous micro-expression database: inducement, collection and baseline[EB/OL]. (2013-07-15)[2024-05-12]. [Https://ieeexplore.ieee.org/abstract/document/6553717](https://ieeexplore.ieee.org/abstract/document/6553717).
- [27] Yan W J, Qi W, Liu Y J, et al. CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces [EB/OL]. (2013-07-15)[2024-05-12]. [Https://ieeexplore.ieee.org/abstract/document/6553799](https://ieeexplore.ieee.org/abstract/document/6553799).
- [28] Yan W J, Li X B, Wang S J, et al. CASME II: an improved spontaneous micro-expression database and the baseline evaluation [J]. PLoS One, 2014, 9(1): e86041.
- [29] Davison A K, Lansley C, Costen N, et al. SAMM: a spontaneous micro-facial movement dataset[J]. IEEE Transactions on Affective Computing, 2018, 9(1): 116-129.
- [30] See J, Yap M H, Li J T, et al. MEGC 2019-the second facial micro-expressions grand challenge[EB/OL]. (2019-07-11)[2024-

- 05-12]. [Https://ieeexplore.ieee.org/abstract/document/8756611](https://ieeexplore.ieee.org/abstract/document/8756611).
- [31] Gan Y S, Liang S T, Yau W C, et al. OFF-ApexNet on micro-expression recognition system[J]. *Signal Processing: Image Communication*, 2019, 74: 129-139.
- [32] Zhou L, Mao Q R, Xue L Y. Dual-inception network for cross-database micro-expression recognition[EB/OL].(2019-07-11)[2024-05-12]. [Https://ieeexplore.ieee.org/document/8756579](https://ieeexplore.ieee.org/document/8756579).
- [33] Lai Z Y, Chen R H, Jia J L, et al. Real-time micro-expression recognition based on ResNet and atrous convolutions[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(11): 15215-15226.
- [34] Wang C Y, Peng M, Bi T, et al. Micro-attention for micro-expression recognition[J]. *Neurocomputing*, 2020, 410: 354-362.
- [35] Zhang L F, Hong X P, Arandjelović O, et al. Short and long range relation based spatio-temporal transformer for micro-expression recognition[J]. *IEEE Transactions on Affective Computing*, 2022, 13(4): 1973-1985.
- [36] Shang Z Y, Wang P H, Li X F. Micro-expression recognition based on differential feature fusion[J]. *Multimedia Tools and Applications*, 2024, 83(4): 11111-11126.

(编辑 侯湘)