

基于图像序列的微表情识别方法

李佳艳, 陈爽, 冯浩, 魏立臻, 张丽艳*

大连交通大学轨道智能工程学院, 辽宁 大连

收稿日期: 2026年2月3日; 录用日期: 2026年3月12日; 发布日期: 2026年3月24日

摘要

微表情因其持续时间极短、强度低及面部运动高度局部化的特点, 是情感计算与模式识别领域的重要研究方向, 在安全监控、临床诊断、人机交互等方面具有广泛应用价值。然而, 现有微表情识别方法仍面临两大挑战: 一是难以从时序与空间维度精准定位关键运动区域; 二是在小样本、类别不平衡数据下模型识别效果较弱。为此, 本文提出一种基于图像序列的微表情识别方法AD-Net。该方法以DenseNet121为主干网络, 引入坐标注意力机制, 在通道与空间维度上增强对微表情关键区域的特征提取能力; 进一步结合投影梯度下降对抗训练策略, 提升模型对噪声、光照等扰动的鲁棒性, 缓解小样本过拟合问题。此外, 通过在损失函数中引入类别权重矩阵, 改善类别不平衡对模型性能的影响。在CASME2和SAMM两个主流数据集上进行的消融实验与对比实验结果表明, AD-Net在关键性能指标(如准确率、召回率)上均显著优于现有基线方法, 充分验证了所提出模块与策略的有效性。

关键词

卷积神经网络, 对抗训练策略, 微表情识别, 坐标注意力

Micro-Expression Recognition Method Based on Image Sequences

Jiayan Li, Shuang Chen, Hao Feng, Lizhen Wei, Liyan Zhang*

School of Railway Intelligent Engineering, Dalian Jiao Tong University, Dalian Liaoning

Received: February 3, 2026; accepted: March 12, 2026; published: March 24, 2026

Abstract

Micro-expressions, characterized by their extremely short duration, low intensity, and highly localized facial movements, constitute a crucial research direction in affective computing and pattern recognition, with broad application value in fields such as security monitoring, clinical diagnosis,

*通讯作者。

文章引用: 李佳艳, 陈爽, 冯浩, 魏立臻, 张丽艳. 基于图像序列的微表情识别方法[J]. 人工智能与机器人研究, 2026, 15(2): 593-602. DOI: 10.12677/airr.2026.152057

and human-computer interaction. However, existing micro-expression recognition methods still face two major challenges: first, difficulty in precisely localizing key motion regions across both temporal and spatial dimensions; second, weak model recognition performance under conditions of small sample sizes and class imbalance. To address these issues, this paper proposes an image-sequence-based micro-expression recognition method named AD-Net. The method adopts DenseNet121 as the backbone network and incorporates a coordinate attention mechanism to enhance feature extraction of key micro-expression regions across channel and spatial dimensions. Furthermore, a projected gradient descent adversarial training strategy is integrated to improve the model's robustness against disturbances such as noise and illumination variation, thereby mitigating overfitting issues in small-sample scenarios. Additionally, a class weight matrix is introduced into the loss function to alleviate the negative impact of class imbalance on model performance. Ablation and comparative experiments conducted on two mainstream datasets, CASME2 and SAMM, demonstrate that AD-Net significantly outperforms existing baseline methods across key performance metrics (such as accuracy and recall), thereby fully validating the effectiveness of the proposed modules and strategies.

Keywords

Convolutional Neural Network, Adversarial Training Strategy, Micro-Expression Recognition, Coordinate Attention Mechanism

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

微表情因其极短的持续时间、极低的强度以及高度局部化的面部肌肉运动，构成了模式识别领域的重大挑战。微表情是了解人类真实情感的重要窗口，这使得它在医学诊断、国防安全、教育、刑事审判、刑事侦查、广告和消费等许多应用领域得到了广泛的应用[1]。

目前深度学习方法，特别是卷积神经网络(CNN)已成为微表情识别的主流技术。Kim 等人[2]首次将卷积神经网络(CNN)应用于微表情识别研究，但是微表情数据集大多是有限的并且是类别不平衡的，直接应用 CNN 时容易引起过拟合，Liong 等人[3]使用单一顶点帧进行微表情识别，大幅减少了输入帧的冗余信息，但是参数量相比于单分支结构更大，需要对全局和局部定义多重约束条件，同时需要引入多个损失函数来区分类内和类间特征，模型复杂度过高。任煜[4]提出使用 SE 模块中添加的改进残差网络 SE-RESNEXT-50 来提取微表情峰值帧的特征，解决了由于数据集数量少和不同类别 60 个分布不均而导致的识别困难和识别精度低的问题。虽然上述算法在表情识别方面取得了良好的效果，但微表情短而隐蔽，导致识别困难，识别准确率低。在微表情识别的研究中，目前仍存在以下两个问题：一是难以精准定位关键运动发生的时序片段(关键帧)与空间区域；二是数据样本少导致模型不稳定，识别效果不佳。为解决上述问题，本文提出了一种基于深度学习的微表情识别算法，以 densenet 为基础，通过增强对微表情序列通道上和空间上特征的提取，实现自适应地学习输入数据的权重和特征，增强微表情的可识别性，该算法融入坐标注意力(Coordinate Attention, CA)模块[5]，建模通道间依赖关系并沿水平和垂直方向编码精确的空间位置信息，同时在模型训练过程中采用对抗攻击训练策略，具体地，我们利用投影梯度下降(Projected Gradient Descent, PGD)方法[6]，在原始图像序列上生成对抗样本，并将这些扰动样本纳入训练过程，这种训练策略不仅有效提升了模型对光照波动、噪声污染等常见扰动的容忍度，其强大的正则化效应也显著缓解了小样本数据下的过拟合问题。通过这一技术，模型能够更好地识别图像序列之间的微

妙信息，从而提高微表情识别的准确性。

2. 相关工作

2.1. 人脸对齐和裁剪处理

在进行微表情识别之前，对原始的微表情数据集采取关键点检测和冗余部分的裁剪是很有必要的。本文中我们对数据的起始帧和顶点帧之间的序列进行等间隔抽取帧数，并去除距离顶点帧的间隔帧数大于 10 帧的冗余图像序列，组合起来后的新图像序列用来表示微表情的运动变化。

对于每个输入样本，利用 Retinaface [7]方法来检测面部关键点区域并将其统一裁剪为 128×128 的大小作为模型的输入，裁剪的流程如图 1 所示。

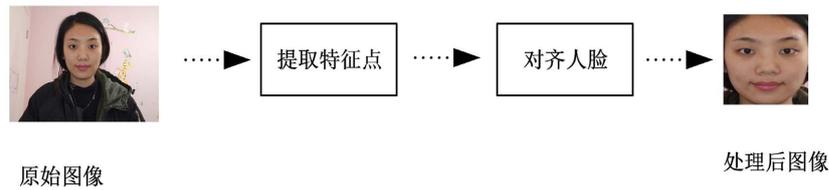


Figure 1. Original image alignment and cropping process diagram
图 1. 原始图像对齐裁剪过程图

2.2. 模型框架

本文的模型框架采用 AD-Net (Apex Distillation Network)，如图 2 所示，本文把顶点帧附近的图像作为网络的输入，大大减少了不必要信息，以 densenet121 为主干网络，每一层都可以直接获得所有先前层的特征图，从而实现特征重用和信息流，实现网络效率，图片在经过过渡层后，被送到坐标注意力模块 (CA) 中来强化关键时空区域定位和特征提取能力，并在分类之前我们引入特征细化来提取特定的微表情特征。此外，考虑到数据集稀缺的情况下，我们在训练过程中引入对抗攻击训练，旨在增强模型的稳定性和泛化能力。最后，为了应对微表情数据集中存在的分类不平衡现象，我们在损失函数中加入了权值矩阵，使得每个类别对损失值的贡献更加公平。这一改进有助于模型更加均衡地学习所有类别，尤其是在类别样本数量差异较大的数据集中，能够提高模型对少数类的识别能力，从而提升整体的性能。

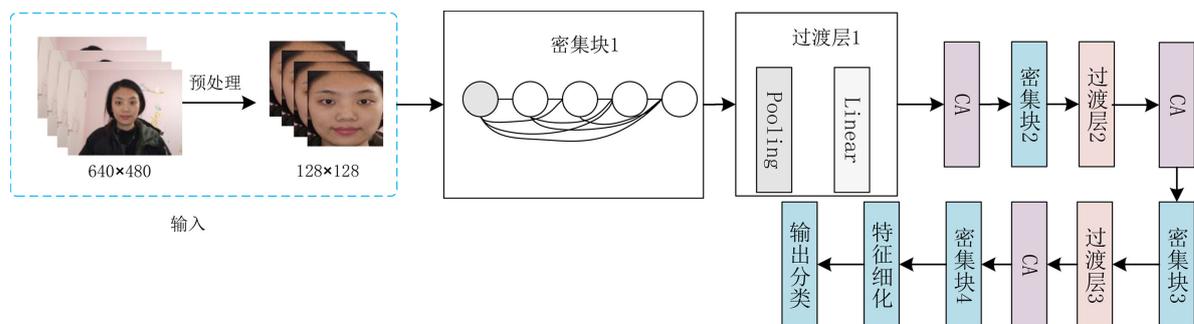


Figure 2. AD-Net structure diagram
图 2. AD-Net 结构图

2.3. 坐标注意力机制

坐标注意力模块的独特之处在于它能够识别和区分不同的空间方向，即坐标，从而生成具有坐标感知能力的特征图，坐标注意力机制结构图如图 3 所示。与传统的注意力机制相比，CA 在进行注意力分配

时，不仅深入考虑了输入数据的特征信息，而且还融入了每个像素点的空间位置信息[8]。这种结合特征和位置信息的方法，使得 CA 能够更加精准地捕捉到图像中的局部细节和全局结构，在空间关系的理解上具有更高的准确性和敏感度。本文引入了 CA 机制，使模型对面部特征表征能力好的通道和位置赋予更大权值，更加关注受试者的面部特征信息，降低光照、噪声等杂质的干扰能力，提高表情的识别精度。

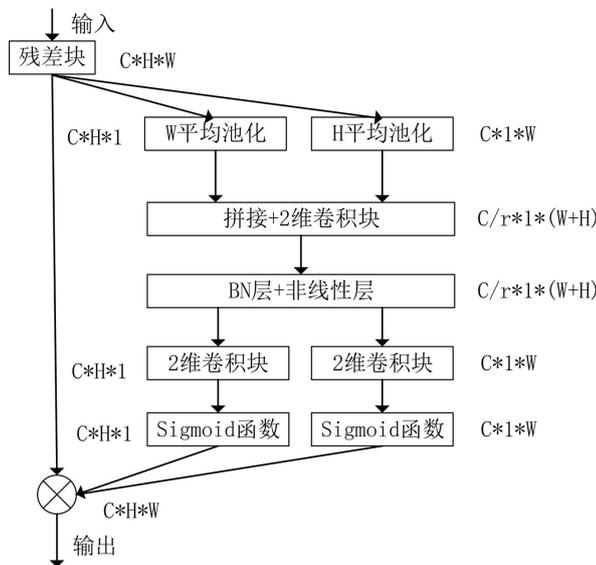


Figure 3. Coordinate attention mechanism structure diagram
图 3. 坐标注意力机制结构图

具体实现是将式(1)中所定义的全局池化分解为两个一维特征编码操作，如式(2)和式(3)所示。

$$Z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \tag{1}$$

$$Z_c^h(h) = \frac{1}{W} \sum_{i=1}^W x_c(h, i) \tag{2}$$

$$Z_c^w(w) = \frac{1}{H} \sum_{i=1}^H x_c(i, w) \tag{3}$$

式中， i, j 为坐标； W 为特征图宽度； H 为特征图高度； x_c 为第 C 通道输入； Z_c 为该通道输出。

具体过程为：输入特征图经过残差块处理，然后分别进行宽度和高度方向的平均池化，得到两个不同维度的特征图。这两个特征图被拼接并通过一个 2 维卷积块进行处理，接着通过批归一化层和非线性激活层进一步加工。之后，特征图通过另一个 2 维卷积块，并通过 Sigmoid 函数转换为注意力权重，这些权重表示不同位置特征的重要性。最后，原始输入特征图与这些注意力权重相乘，得到加权后的特征图，再与原始特征图进行逐元素相加，形成最终的输出。这一过程不仅保留了原始信息，还通过注意力机制增强了模型对关键特征的关注，从而提升了模型的性能和泛化能力。

2.4. PGD 对抗训练策略

PGD 对抗训练是一种有效的对抗训练方法，主要用于寻找损失函数的最小值。它通过在每个迭代步骤中，根据梯度下降的方向进行参数更新，以达到损失函数最小化的目的[6]。相对于其它优化算法，PGD 拥有许多优点，例如它对初始的选择并不敏感，可以处理带约束的问题，而且它的运行速度也非常快。

在正常的训练过程中,首先进行一次前向传播和反向传播,计算模型的损失和梯度。然后,使用 PGD 类进行对抗训练,通过多次迭代攻击来生成对抗样本,并计算对抗损失。将对抗损失的梯度累积到原始梯度上,然后进行一次梯度下降更新模型参数。在每次对抗训练迭代后,恢复嵌入层的原始参数,以便在下一次迭代中重新计算梯度。PGD 相关公式分别如式(4)、式(5)所示。

$$X_{adv,0} = X \quad (4)$$

$$X_{adv,n+1} = P_e \left(X_{adv,n} + a * \text{sign} \left(D_X J \left(X_{adv,n,y} \right) \right) \right) \quad (5)$$

其中, $X_{adv,0}$ 表示初始对抗数据等于原始数据; $X_{adv,n+1}$ 表示第 $(n+1)$ 次更新后的对抗数据; P 代表一种投影函数,可确保对抗数据保持在有效范围内(通常与原始数据保持一定距离); a 表示更新的学习率或步长; $\text{sign}(\cdot)$ 代表第 n 次迭代时损失函数相对于对抗数据的梯度符号,用于指导对抗更新; $J(X_{adv,n,y})$ 表示损失函数,指示模型性能; y 为数据的真实标签。

2.5. 损失函数改进

由于微表情数据集存在分类不平衡的现象,因此模型的损失函数在交叉熵函数的基础上,添加了权值矩阵,使得每个类别对损失值的贡献更加公平,其计算公式如式(6)所示。

$$Loss = -\frac{1}{J} \sum_{i=1}^J \sum_{k=1}^K v_k \cdot y(c_k, x_i) \log(p(c_k, x_i)) \quad (6)$$

其中, J 为模型计算中的一个批次的大小, $y(c_k, x_i)$ 表示样本 x_i 真实标签的编码在分类 c_k 上的值, v_k 为权值矩阵,计算公式如式(7)所示。

$$v_k = \frac{1}{N} \sum_{(x_n, y_n)} X(y_n, c_k), \quad n=1,2,3,\dots; \quad k=1,2,3,\dots,K \quad (7)$$

$$X(i, j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (8)$$

3. 实验及结果分析

3.1. 数据集

实验使用了 CASME2、SAMM 两个数据集。CASME2 是中国科学院心理研究所傅小兰课题组发布的自发微表情数据集,它使用了 200 FPS 帧率的高速摄像机录制微表情视频,采集了 26 个受试者的 247 条微表情样本,这些样本被分为五类。SAMM 是英国曼彻斯特大学 Davison 等人发布的,采用帧率为 200 FPS 的高速摄像机对 32 个受试者采集了 159 个微表情样本,并将其分为七类,同时标注顶点帧,提供 AU 信息。所有样本被分为三类微表情: Positive、Negative、Surprise。为了采用统一的情感标签系统来进行微表情识别任务,实验中对情绪类别划分如下:“积极”情绪类别包括“快乐”,而“消极”情绪类别包括“悲伤”、“厌恶”、“蔑视”、“恐惧”和“愤怒”,“惊喜”情感类别特指“惊喜”。

3.2. 实验配置

本实验环境为 windows11 搭载的操作系统平台,配备 16 GB 运行内存,具体的配置如表 1 所示。图片输入大小为 128×128 ,初始学习率为 1×10^{-4} , batch_size 设置为 16, epoch 为 100,数据集均使用的是 LOSO 交叉验证。

Table 1. Model training and testing environment configuration parameters**表 1.** 模型训练和测试环境配置参数

配置名称	参数或版本
GPU	NVIDIA GeForce RTX 4060 GPU
内存/显存	16 GB/8 GB
深度学习框架	Pytorch 2.4.1
Python/CUDA	3.9/12.4

3.3. 评价指标

本实验中我们使用了 CASME2 和 SAMM 数据集做实验，其中常用到的评价指标一般为 Accuracy 和 F1-score，具体涉及的计算公式如式(9)~(12)所示。

$$\text{Accuracy} = \frac{TP}{N} \times 100\% \quad (9)$$

$$\text{F1-Score} = \frac{2PR}{P+R} \quad (10)$$

$$\text{precision} = \frac{TP}{TP+FP} \quad (11)$$

$$\text{recall} = \frac{TP}{TP+FN} \quad (12)$$

其中， TP 表示正类预测正确的总数， FP 代表将负类判定为正类的总数， FN 代表正类判定为负类的总数， TN 代表负类判定正确的总数， N 表示测试样本的总数， P 代表 precision 精确率， R 代表 recall 召回率。Accuracy 从某种意义上可以评价一个分类器是否有效，但是对于样本不均衡的数据集，Accuracy 不再适合作为评判，此时，F1-score 更具有权威性，因为它是精确率与召回率的权衡。

3.4. 实验结果和分析

3.4.1. 消融实验

为了验证各个组件在模型中的重要性，本文进行了一系列的消融实验。消融实验的模型包括：仅 Densenet、Densenet 与 PGD 对抗训练策略、Densenet 与坐标注意力机制、Densenet 与坐标注意力机制和 PGD 方法结合的模型，需要说明的是，本文所有消融实验均基于相同的数据预处理策略，即在 apex 附近进行间隔抽帧，未引入显式的运动建模或区域级特征提取。在此基础上，通过引入 PGD 对抗训练和坐标注意力模块，对静态外观特征的判别能力进行增强，实验在 CASME2、SAMM 两个数据集上均进行了验证，如表 2、表 3 所示。

Table 2. Ablation results (%) of each component on the CASME2 database**表 2.** 各组件在 CASME2 数据库上的消融结果(%)

模型	Accuracy	F1-score
Dense	81.35	78.98
Dense + PGD	84.67	82.56
Dense + CA	87.98	84.30
Dense + PGD + CA (AD-Net)	90.45	84.94

Table 3. Ablation results (%) of each component on the SAMM database**表 3.** 各组件在 SAMM 数据库上的消融结果(%)

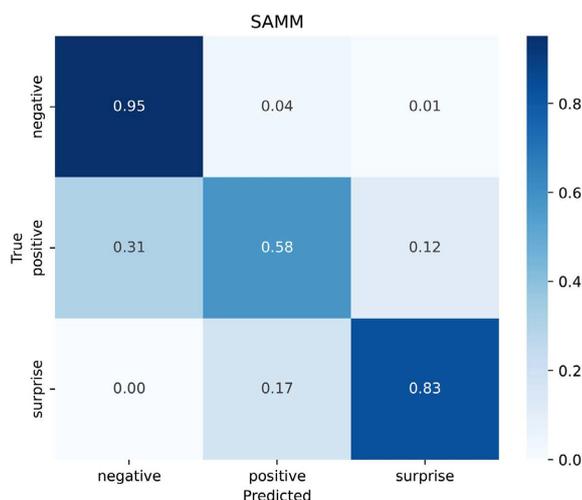
模型	Accuracy	F1-score
Dense	78.96	70.42
Dense + PGD	80.21	72.08
Dense + CA	82.10	74.10
Dense + PGD + CA (AD-Net)	87.46	80.03

由表 2 和表 3 可知, 所提出的 AD-Net 模型在 CASME2 和 SAMM 两个微表情数据库上均取得了优于单一模块改进模型的识别性能, 验证了各组成模块的有效性及其协同作用。

在 CASME2 数据集上, 基准 DenseNet 模型的准确率和 F1-score 分别为 81.35% 和 78.98%。当引入 PGD 对抗训练策略后, 模型的准确率提升至 84.67%, F1-score 提升至 82.56%, 分别提高了 3.32% 和 3.58%, 说明对抗训练能够在一定程度上增强模型对微表情细微变化的鲁棒性。进一步地, 在基准模型中引入坐标注意力(CA)模块后, 模型准确率和 F1-score 分别提升至 87.98% 和 84.30%, 相比基准模型均有明显提升, 表明 CA 模块有助于模型更有效地挖掘具有判别力的特征信息。最终, 当同时引入 PGD 和 CA 模块构建 AD-Net 模型时, 模型在 CASME2 数据集上取得了最优性能, 其准确率和 F1-score 分别达到 90.45% 和 84.94%, 进一步验证了两种改进策略在该数据集上的互补性和有效性。

在 SAMM 数据集上, 由于样本规模较小且微表情幅度更加细微, 整体识别难度更高。基准 DenseNet 模型在该数据集上的准确率和 F1-score 分别为 78.96% 和 70.42%。引入 PGD 对抗训练后, 模型性能有所提升, 准确率提高至 80.21%, F1-score 提高至 72.08%, 分别提升了 1.25% 和 1.66%。当加入 CA 模块后, 模型准确率和 F1-score 分别进一步提升至 82.10% 和 74.10%, 说明注意力机制在复杂微表情场景中同样具有一定的有效性。最终, 融合 PGD 与 CA 的 AD-Net 模型在 SAMM 数据集上取得了最优结果, 其准确率和 F1-score 分别达到 87.46% 和 80.03%, 表明所提出方法在不同微表情数据库上均具有良好的稳定性和泛化能力。

3.4.2. 基于混淆矩阵的性能分析

**Figure 4.** Confusion matrix on the CASME2 dataset**图 4.** CASME2 数据集混淆矩阵

为了进一步评估模型在不同情感类别上的分类性能, 本研究采用混淆矩阵对各类别的识别结果进行

细粒度分析。基于数据集的统一分类标准，本文将情感分为三个类别(positive, negative, surprise)进行评估。图 4 和图 5 分别展示了模型在 CASME2 与 SAMM 数据集上的混淆矩阵。从混淆矩阵中可以观察到，模型在 negative 类别的识别上表现稳定且准确率较高，说明模型能够较好地地区分该类别特征。然而，在 positive 类别上的识别效果相对较弱，存在一定的误判情况，这提示该类别的特征表达可能与其他类别存在相似性，或训练样本的特征区分度不足，后续可通过引入注意力机制或融合时序特征以提升其判别能力。此外，surprise 类别的识别虽然总体表现较好，但仍存在一定比例的误分，尤其是与 positive 类别之间的混淆，说明模型在该类别的边界特征学习上仍有优化空间。总之，本文方法在情感分类任务中具有一定有效性，尤其在 negative 类别上表现突出，而在 positive 和 surprise 类别上仍有提升潜力。

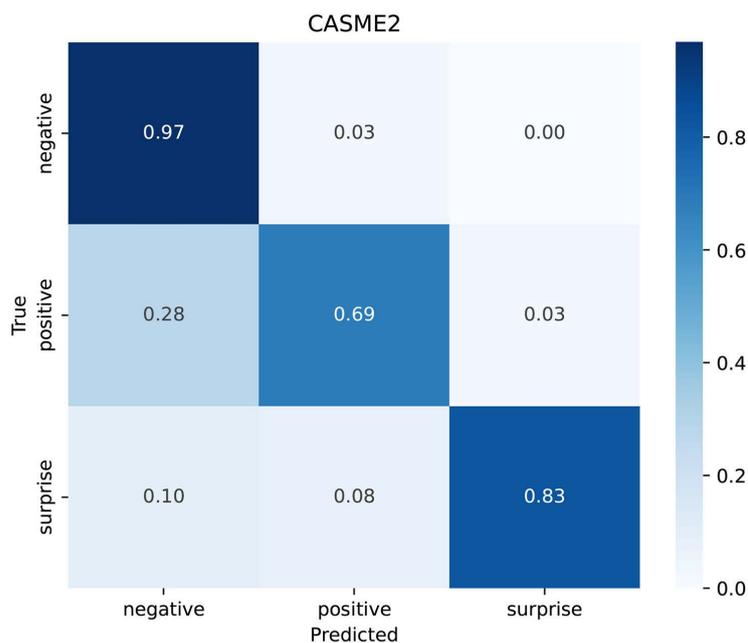


Figure 5. Confusion matrix on the SAMM dataset
图 5. SAMM 数据集混淆矩阵

3.4.3. 与其他模型的对比实验分析

为验证所提出 AD-Net 模型的有效性，本文将其与多种代表性微表情识别方法在 CASME2 和 SAMM 数据集上进行了对比实验，结果如表 4 和表 5 所示。

Table 4. Performance comparison (%) with other models on the CASME2 database
表 4. 与其他模型在 CASME2 数据库上的性能对比(%)

方法	类别	Accuary	F1-score
OFF-ApexNet [9]	3	88.28	86.97
STSTNet [10]	3	86.86	83.82
SLSTT [11]	3	75.81	75.30
AM3F-FlowNet [12]	3	84.52	82.88
MMNet [13]	3	88.60	86.90
AD-Net (ours)	3	90.45	84.94

Table 5. Performance comparison (%) with other models on the SAMM database
表 5. 与其他模型在 SAMM 数据库上的性能对比(%)

方法	类别	Accuracy	F1-score
OFF-ApexNet [9]	3	68.18	54.23
STSTNet [10]	3	68.10	65.88
SLSTT [11]	3	72.38	64.00
AM3F-FlowNet [12]	3	66.18	54.10
MMNet [13]	3	80.10	72.90
AD-Net (ours)	3	87.46	80.03

在 CASME2 数据集上, AD-Net 取得了 90.45% 的准确率, 高于多数对比方法, 相比 OFF-ApexNet 和 MMNet 分别提升约 2.17% 和 1.85%。在 F1-score 指标上, AD-Net 达到 84.94%, 整体处于较高水平, 表明该方法在保证较高分类准确率的同时, 仍能维持较为稳定的类别判别能力。

在 SAMM 数据集上, AD-Net 的性能提升更为明显, 其准确率达到 87.46%, 相较于性能较优的 MMNet 提升约 7.36%, F1-score 也提升至 80.03%。该结果表明, 在样本规模更小、类别分布更不均衡的复杂场景下, AD-Net 仍具有较好的鲁棒性和泛化能力。

4. 结论

本文提出了一种融合坐标注意力机制与对抗训练策略的深度学习模型 AD-Net。通过坐标注意力模块增强模型对微表情时空关键特征的感知与提取能力, 并结合 PGD 对抗训练提升模型在扰动环境下的鲁棒性, 有效抑制了过拟合现象。在 CASME2 和 SAMM 数据集上的消融实验与对比实验表明, AD-Net 在各项评价指标上均优于现有方法, 特别是在样本规模较小、类别分布不均衡的 SAMM 数据集上仍保持较高的识别性能, 说明该方法具有良好的适应性与泛化能力。然而, 实验也发现模型在“积极”类别上的识别效果相对较弱, 提示未来可在多模态特征融合、时序建模优化等方面进一步改进。总体而言, 本文提出的方法为微表情识别提供了一种有效的技术路径, 对推动其在真实场景中的应用具有一定的理论与实用价值。

参考文献

- [1] Zeng, X., Zhao, X., Zhong, X. and Liu, G. (2023) A Survey of Micro-Expression Recognition Methods Based on LBP, Optical Flow and Deep Learning. *Neural Processing Letters*, **55**, 5995-6026. <https://doi.org/10.1007/s11063-022-11123-x>
- [2] Kim, D.H., Baddar, W.J. and Ro, Y.M. (2016) Micro-Expression Recognition with Expression-State Constrained Spatio-Temporal Feature Representations. *Proceedings of the 24th ACM international conference on Multimedia*, Amsterdam, 15-19 October 2016, 382-386. <https://doi.org/10.1145/2964284.2967247>
- [3] Liong, S., See, J., Wong, K. and Phan, R.C. (2018) Less Is More: Micro-Expression Recognition from Video Using Apex Frame. *Signal Processing: Image Communication*, **62**, 82-92. <https://doi.org/10.1016/j.image.2017.11.006>
- [4] Ren, Y., Chen, X.Q., Wang, D.R. and Chen, X.Y. (2024) Improved Residual Network and Apex Frame for Micro-Expression Recognition. *Journal of Chongqing Technology and Business University (Natural Sciences Edition)*, **41**, 21-29.
- [5] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville 20-25 June 2021, 13708-13717. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [6] Madry, A., Makelov, A., Schmidt, L., et al. (2018) Towards Deep Learning Models Resistant to Adversarial Attacks. arXiv: 1706.06083.

-
- [7] Deng, J., Guo, J., Zhou, Y., et al. (2019) RetinaFace: Single-Stage Dense Face Localisation in the Wild. arXiv: 1905.00641.
- [8] Liu, M.L., Wang, S., Yang, Z.Q., et al. (2026) Change Detection in High-Resolution Satellite Imagery Based on Improved Attention and Convolution. *Information Recording Materials*, **27**, 193-196.
- [9] Gan, Y.S., Liong, S., Yau, W., Huang, Y. and Tan, L. (2019) Off-ApexNet on Micro-Expression Recognition System. *Signal Processing: Image Communication*, **74**, 129-139. <https://doi.org/10.1016/j.image.2019.02.005>
- [10] Liong, S., Gan, Y.S., See, J., Khor, H. and Huang, Y. (2019) Shallow Triple Stream Three-Dimensional CNN (STSTNet) for Micro-Expression Recognition. 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), Lille, 14-18 May 2019, 1-5. <https://doi.org/10.1109/fg.2019.8756567>
- [11] Zhang, L., Hong, X., Arandjelovic, O. and Zhao, G. (2022) Short and Long Range Relation Based Spatio-Temporal Transformer for Micro-Expression Recognition. *IEEE Transactions on Affective Computing*, **13**, 1973-1985. <https://doi.org/10.1109/taffc.2022.3213509>
- [12] Fu, C., Yang, W., Chen, D. and Wei, F. (2023) AM3F-FlowNet: Attention-Based Multi-Scale Multi-Branch Flow Network. *Entropy*, **25**, 1064. <https://doi.org/10.3390/e25071064>
- [13] Li, H., Sui, M., Zhu, Z. and Zhao, F. (2022) MMNet: Muscle Motion-Guided Network for Micro-Expression Recognition. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, Vienna, 23-29 July 2022, 1074-1080. <https://doi.org/10.24963/ijcai.2022/150>