**小组2第二次大作业**

**——长时微表情识别**

**0、环境配置**

本研究采用Python 3.8作为核心开发语言，基于其成熟的科学计算生态及深度学习框架支持特性。开发环境构建于PyCharm 2024.1社区版集成开发环境，通过其完善的代码管理、版本控制等功能提升开发效率。环境管理采用conda 25.1.1构建虚拟环境，有效隔离项目依赖并解决多版本库冲突问题。

硬件配置主体部署于本地计算平台，配备NVIDIA GeForce RTX 3060 Laptop显卡（6GB GDDR6X显存），配合CUDA 11.8及cuDNN 9.8加速库实现GPU加速计算。核心依赖库包括：OpenCV 4.5.2（图像预处理与特征提取）、PyTorch 2.3.1（深度学习模型构建与训练）、NumPy 1.24.3（张量运算与数值计算）及Pillow 10.4.0（图像I/O处理）。辅助工具链包含SciPy 1.10.1（科学计算）、Matplotlib 3.7.1（可视化分析）和TQDM 4.67.1（进度监控）。

当前环境配置及版本号为初步方案，可能根据实验需求动态调整。环境配置方案采用模块化设计原则，通过requirements.txt文件实现依赖项版本固化，确保实验可复现性。

1. **研究现状**

目前，我们对人类表情与情绪关系的理解还存在诸多未解之谜，尤其是在长时间维度下情绪的动态演变、不同情绪之间的相互转化以及表情与内心真实情感之间的复杂关联等方面。要完善人类情绪表达与认知的理论体系，对长时间的表情识别与分析进行深入研究具有十足的重要性。不仅如此，在许多场景中，掌握人类的情绪变化都具有至关重要的地位：在未来的人机交互中，长时间表情识别与分析可以使计算机从多角度接收人类情绪表达信号，推动智能家居、智能客服、智能驾驶等行业发展，对人类微表情的深刻理解可以帮助计算机自动感知和回应情绪，尤其在智能驾驶中，驾驶者的疲劳、愤怒、悲伤等情绪造成的许多危险驾驶行为都可通过车辆的主动介入避免；在心理医疗领域，机器对长时间表情的分析可以帮助心理医生捕捉患者的细微情绪变化，为病情评估提供更丰富的依据；在教育领域，对学生的情绪分析有助于老师明白学生实际的心理变化过程，依此为学生提供更有助于心理成长的教育方法，避免心理压力过大甚至产生心理问题；在公安领域，情绪分析能够帮助公安机关捕捉嫌疑人的微妙情绪变化，通过焦虑、紧张等情绪的积累缩小嫌疑范围……在当今这个信息爆炸、人际交往日益复杂而人工智能技术蓬勃发展的时代，基于深度学习研究长时间表情识别与分析具有深远且多维度的必要性。

在表情识别的特征提取方面，国内外已经有许多优秀的学者做了大量的工作。基于手工特征的方法，如局部二值模式（LBP）和光流特征，通过提取图像的纹理和运动信息，能够有效捕捉表情的动态特征。例如，G. Zhao 和 M. Pietikäinen 提出的 LBP-TOP（Local Binary Patterns on Three Orthogonal Planes） 通过在三个正交平面上提取LBP特征，能够有效建模微表情的时空特征，但其特征维度较高，计算复杂度较大。X. Ben 等人提出的 HWP-TOP（Hot Wheel Patterns on Three Orthogonal Planes） 和 DCP-TOP（Dual-Cross Patterns on Three Orthogonal Planes）在 LBP-TOP 的基础上引入了方向信息和双交叉模式，进一步提高了特征的判别能力。Y.-J. Liu 等人提出的 MDMO（Mean Directional Mean Optical Flow） 和 Sparse MDMO（Sparse Mean Directional Mean Optical Flow）通过计算主方向上的光流向量的均值和稀疏表示，有效地捕捉了微表情中的运动信息，并对光照变化和头部运动具有更好的鲁棒性。然而，现有的手工特征方法往往难以适应复杂的实际场景，泛化能力有限。为了克服这一局限，基于深度学习的方法逐渐成为表情识别的主流。H.-Q. Khor等人提出的 ELRCN（Enriched Long-term Recurrent Convolutional Network）结合CNN和LSTM，提取视频中的时空特征，能够有效处理视频中的时间信息，适用于动态表情识别。X. Jia等人提出的Dual-Inception Network\*\* 通过Inception模块提取多尺度特征，提高了模型的表征能力，进一步提升了识别的准确性。

**2、问题定义**

传统方法虽能有效捕捉微表情动态特征，但在处理小样本图像空间时，难以捕捉微表情的细微变化，特征提取多样性不足。现有的基于深度学习的算法在特征提取和模型优化上虽有显著进展，但对数据依赖性强，导致模型训练和泛化能力受限，且模型可解释性较差。同时，现有研究多集中在算法改进和性能提升上，对用户交互和结果呈现的用户友好度关注不够，缺乏对多维度数据的整合和关联展示，难以满足实际应用场景中用户对直观、易用软件界面的需求。

因此，本项目所研究的问题主要分为算法优化和软件设计两方面。首先在算法方面，本项目将致力于更有效地捕捉长时间视频序列中的微表情变化，提高识别准确率。将结合多种特征提取技术，充分利用微表情的时间和空间信息，设计更鲁棒的特征提取和识别算法。同时，针对微表情数据集样本稀缺问题，研究数据增强、迁移学习等方法，提升模型泛化能力。软件设计方面，本项目希望开发一款结合多种功能——包括图像情绪识别分析、视频情绪识别分析、实时情绪监控分析等——并且拥有对用户更加友好的软件界面的软件。除去优秀多样的功能，软件还将具备高效的数据处理与存储模块，以应对长时间视频产生的大量数据。采用多线程读取、批量处理等技术快速处理视频数据，并运用合适的数据存储结构和压缩技术，减少存储空间，提高读写效率。

**3、数据采集**

本小组阅读了近几年有关于微表情（Micro-Expression）识别的文章，了解到当前微表情识别领域较为主流的四个公开数据集：SMIC、CASME II、SAMM 和 MMEW。

SMIC 数据集最早发布于2013年，提供了三种模态（高速摄像、可见光、红外），情绪标签分为积极、消极和惊讶三类，适用于基础的微表情检测任务，但其标签较粗、样本数量有限，缺乏详细的动作单元标注，难以满足我们对图像和视频时序信息的建模需求；SAMM 数据集于2016年发布，具有更高的图像分辨率（2040×1088）和较丰富的实验者种族，支持跨人种的分析研究。其面部细节保留充分，适合高质量图像下的情绪建模，但样本数量偏少，情绪类别虽多但分布不均；MMEW 数据集是近年来发展较快，且较新的数据集之一，收集了微表情与宏表情视频共1200条，并提供了情绪标签与FACS动作单元的双重标注。然而，该数据集较新，尚缺乏统一的对比方法和评价标准，并且与我们的研究有出入。

综合考虑后，我们决定在现阶段使用CASME II数据集作为研究的数据集输入来源。CASME II由中科院心理研究所发布，采集过程中采用高帧率高速相机（200fps），在受控环境下诱导自发微表情，具有较高的数据质量。该数据集包含来自26位被试者的255个微表情片段，覆盖7种情绪类别，样本分布相对均衡，且每个样本均提供了起始帧、峰值帧、结束帧、情绪类别以及面部动作单元（AU）等详尽标注信息。同时，我们将使用新近的MMEW数据集来测试算法的普适性，增强对研究结果的验证。

对我们现阶段以图片作为输入、结合时序的微表情特征变化所反映的情绪变化的研究目标而言，CASME II数据集提供了较为优越的数据基础，既适用于静态帧的面部特征提取，也适用于短视频中情绪随时间演化的建模。同时，CASME II已被广泛应用于微表情识别研究领域中，结果评价手段成熟，在微表情特征提取的结果对比和重复实验的结果具有稳定性，为我们进一步开展研究有很大利好。

**图形用户界面, 文本, 应用程序, 电子邮件

AI 生成的内容可能不正确。图形用户界面, 文本, 应用程序, 电子邮件

AI 生成的内容可能不正确。**