

基于深度学习的渐冻症患者实时监测系统

赵丛丛¹, 刘静超¹, 宋宇², 张杰^{1*}

¹西京学院计算机学院, 陕西 西安

²西京学院机械工程学院, 陕西 西安

收稿日期: 2025年2月2日; 录用日期: 2025年2月28日; 发布日期: 2025年3月12日

摘要

渐冻症(ALS)是一种严重影响运动神经元的神经退行性疾病,其早期症状通常表现为肌肉无力、抽搐和运动障碍。为了提高对渐冻症患者健康状况的实时监测能力,对病人进行24小时无接触看护。本研究采用深度卷积网络技术对图片进行跟踪运算,设计了实时监测系统,捕捉并甄别患者微表情、微动作来进行判断并报警,及时通知陪护人员前来处理危急情况,避免被看护人员发生危险。研究表明:该方法具有较高的检测精度和良好的实时性,可有效识别渐冻症患者的异常状态,显著提升健康监测的准确性与可靠性。本研究为渐冻症患者提供了一种便捷高效的监测手段,为个性化健康管理与早期干预提供了新的技术支持。

关键词

渐冻症, 微表情, 深度学习, 实时监测

A Real-Time Monitoring System for ALS Based on Deep Learning

Congcong Zhao¹, Jingchao Liu¹, Yu Song², Jie Zhang^{1*}

¹School of Computer Science, Xijing University, Xi'an Shaanxi

²School of Mechanical Engineering, Xijing University, Xi'an Shaanxi

Received: Feb. 2nd, 2025; accepted: Feb. 28th, 2025; published: Mar. 12th, 2025

Abstract

ALS (Amyotrophic Lateral Sclerosis) is a neurodegenerative disease that seriously affects motor neurons, and its early symptoms usually manifest as muscle weakness, twitches, and movement disorders. In order to improve the real-time monitoring ability of the health condition of ALS patients,

*通讯作者。

文章引用: 赵丛丛, 刘静超, 宋宇, 张杰. 基于深度学习的渐冻症患者实时监测系统[J]. 计算机科学与应用, 2025, 15(3): 20-28. DOI: 10.12677/csa.2025.153054

patients are monitored without physical contact for 24 hours. This study uses deep convolutional neural network technology to perform tracking calculations on images and designs a real-time monitoring system that captures and identifies patients' micro-expressions and micro-motions to make judgments and issue alarms. It notifies caregivers to handle emergencies within 3 seconds, thereby avoiding danger to the monitored person. The results show that the method has high detection accuracy and good real-time performance, can effectively identify abnormal states of ALS patients, and significantly improves the accuracy and reliability of health monitoring. This study provides a convenient and efficient monitoring method for ALS patients and provides new technical support for personalized health management and early intervention.

Keywords

Amyotrophic Lateral Sclerosis, Microexpressions, Deep Learning, Real-Time Monitoring

Copyright © 2025 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

渐冻症(肌萎缩侧索硬化症, ALS)是一种严重的神经退行性疾病,导致运动神经元逐渐退化,从而影响患者的运动能力和表情控制。早期症状包括微小的面部表情变化、肌肉抽搐或细微的肢体动作异常。患者的身体逐渐像被冰冻一般,逐步丧失语言表达和行动能力。对于重症渐冻症患者来说,即便是维持正常的呼吸、吞咽以及咳嗽功能,也变得异常困难。在无人看护的情况下,患者可能因无法及时咳嗽或呼救而面临生命危险。因此,持续的细致看护和对患者需求的迅速响应,成为渐冻症护理工作中的关键环节。本文通过计算机视觉技术,捕捉到被看护人的痛苦微表情和微动作来判定是否危险来报警,及时通知陪护人员处理紧急情况,避免被看护人发生生命危险。然而,这些细微变化通常难以通过传统的临床观察及时检测。因此,如何准确捕捉患者的微表情和微动作变化,成为实现早期监测和病情动态评估的重要挑战。

计算机视觉技术近年来取得了显著进展,尤其是在面部表情识别与动作捕捉领域,基于深度学习的视觉分析算法能够有效提取和识别微小的面部肌肉变化与肢体动作。近几年,卷积神经网络为代表的深度学习模型被广泛用于图像数据的识别分类任务,在人脸痛苦表情分类的领域更是逐渐取得有效成果。Egede 等[1]通过痛苦表情数据集训练了卷积神经网络,利用迁移学习的方法进行表情特征提取、痛苦识别。刘芳等[2]使用改进的 VGG-16 网络提取微表情特征。Zhou 等[3]训练了循环卷积神经网络识别痛苦表情,循环卷积神经网络的优势是能对视频序列进行识别。郝强[4]将 VGG-16 网络模型迁移至老人面部痛苦表情识别中,对老人这一特定人群更具有针对性,周婕等[5]提出一种改进 ResNet 网络进行表情识别的方法,罗明刚[6]等使用 Adaboost 算法进行了人脸检测研究,取得较佳的结果相比传统的监测方法,使用计算机视觉技术来进行检测,具有不用接触身体、成本低、实时性强的优势,特别适用于渐冻症患者的持续监测。通过高精度摄像头和深度学习模型,系统能够自动识别面部微表情(如皱眉、眼睑微颤)和肢体微动作的细微变化,为渐冻症患者提供连续、客观的病情监测支持。

本文为增强系统的检测准确性和鲁棒性,结合了基于时间序列的深度学习模型 3D ResNet。通过对缓冲区内 16 帧面部图像进行建模,3D ResNet 能够捕捉表情的动态变化并输出痛苦表情的预测结果。最终,系统综合 AU 分析与 3D ResNet 输出,通过融合多维信息提高痛苦表情的判定精度。当检测到痛苦

表情时，系统会触发报警机制，在界面上提示紧急情况，并通过语音合成播放相应的报警信息，提示相关床号已出现异常情况。这种基于动作单元和深度学习的双重方法，有效提升了痛苦检测的灵敏度和可靠性。

2. 方法

由于微表情是一种自发、无意识的面部表情，通常其持续时间短，动作幅度小，视频片段的相邻帧之间反映出的光流运动特征并不明显。本文采用预训练的 ResNet 网络来捕捉表情的动态变化进行检测，利用知识蒸馏技术，将一个较大、性能较好的 3D ResNet 模型作为教师网络，训练一个轻量化的学生网络，使其在参数更少的情况下尽量保留教师网络的性能。这样不仅可以降低模型复杂度，还能在推理时获得更快速度。选取脸部的关键动作单元，这些动作单元通过计算面部特征点之间的几何距离进行量化，进而生成对应的 AU 分数。随后，根据每个动作单元的重要性，使用加权公式计算综合痛苦分数。当该分数超过预设阈值时，系统判定为痛苦表情。

2.1. SA-ResNet

ResNet 是一种深度残差网络，解决了深度神经网络在加深时容易出现的梯度消失和梯度爆炸问题，使得网络可以在不丢失梯度信息的情况下加深，从而大幅提升网络性能。

ResNet50 的网络结构由 50 层深度组成，包括一个初始的卷积层、多个残差模块、全局平均池化层以及全连接层。初始卷积层采用了 7×7 的卷积核，步长为 2，接下来是一个 3×3 的最大池化层，用于缩小特征图的尺寸。网络的核心在于残差模块，每个模块都包含多个卷积层，并通过跳跃连接(Skip Connection)将输入加到输出上，这样的结构使得梯度能够更稳定地传播，避免了深层网络中常见的梯度消失问题。

本文所使用的 SA-ResNet 网络模型如图 1 所示，以 ResNet50 网络作为主体结构，构建 3D 残差块，将原先的 2D 残差块替换为 3D 残差块，在 3D 残差块中，除了卷积层外，还需要对输入数据进行时间维度的下采样，在网络的主干网络模块中使用 SimAm 轻量级注意力机制，可以提高网络学习面部特征的能力。

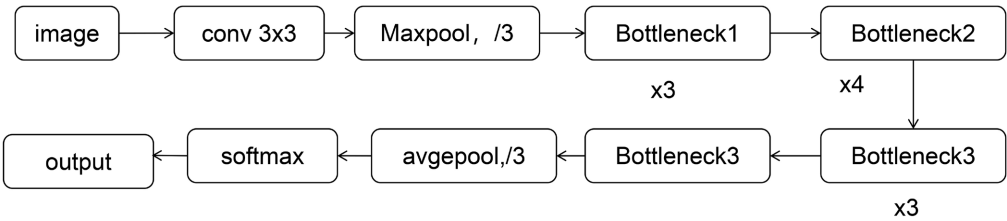


Figure 1. The SA-ResNet network architecture diagram
图 1. SA-ResNet 网络结构图

2.1.1. 3D 残差快

2D 卷积操作通常针对单一图像进行，能够有效地提取图像的细节特征，但它无法捕捉到图像在时间维度上的连续性特征，忽略了图像序列中的时序信息。微表情视频是由一系列相互依存的帧组成的，这些帧在时间轴上具有紧密的特征联系。因此，运用 3D 卷积能够更好地提取微表情视频中的时空特征，从而更全面地理解微表情的动态变化。3D 卷积通过将多个连续帧堆叠成一个立方体，并在立方体内应用 3D 卷积核来实现。在这种结构中，卷积层的每个特征图都会与上一层的多个相邻连续帧相连接，从而有效提取运动信息。

在 3D 卷积层中，每个特征图都与前一层的几个连续帧相连，从而能够捕获图像在时间维度上的信

息。微表情片段作为一个包含多帧图像的序列，其时间和空间特征紧密相关。本研究模型使用了 3D 卷积技术，以便更有效地提取关键帧序列中的时序信息，进而提升识别的精确度。首先需要定义 3D 卷积层卷积核为 $(3 \times 3 \times 3)$ ，RestNet 中的 2D 卷积核替换为 3D 卷积核。

网络中的 Bottleneck 模块如图 2 所示所示。

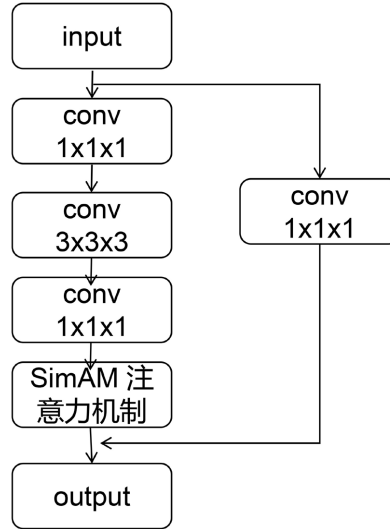


Figure 2. The diagram of the Bottleneck module

图 2. Bottleneck 模块结构图

2.1.2. 注意力机制

为使网络更好地聚焦于面部关键区域，提高检测精度，同时不使模型更加复杂，本文使用轻量级注意力机制 SimAM，SimAM 是一种轻量级、无参数的卷积神经网络注意力机制，它通过计算特征图的局部自相似性来生成注意力权重。SimAM 不需要引入任何额外参数，通过引导模型关注图像中的关键区域，显著提升了模型处理和理解图像的能力。

SimAm 注意力机制的工作原理可以分为以下几个步骤：

1) 特征图提取：通过卷积神经网络提取输入图像的特征图 $X \in R^{B \times C \times H \times W}$ ，其中 B 是批次大小， C 是通道数， H 和 W 分别是特征图的高度和宽度。

2) 计算局部自相似性：SimAM 针对特征图中的每个像素点 $x_{i,j}$ (其中 i, j 分别表示像素在图中的位置索引)，评估其与邻近像素点的相似程度。这种评估是通过比较像素间特征向量的差异来进行的，通常采用的是负的欧几里得距离平方。然而，SimAM 实际上是通过计算像素与其邻近区域像素之间差异平方的平均值(经过标准化处理)来间接测量这种相似性。具体操作为，对每个像素点，先计算其与周围像素点差异的平方，接着对这些平方值进行求和并实施标准化处理。

$$S_{i,j} = \frac{1}{N} \sum_{k \in \Omega_{i,j}} \|x_{i,j} - x_k\|_2^2 \quad (1)$$

其中 $\Omega_{i,j}$ 表示像素 $x_{i,j}$ 的邻域(不包括 $x_{i,j}$ 本身， N 是邻域内像素的数量)，在 SimAM 的实现过程中，通常采用特征图的整体均值进行中心化处理，然后从每个像素值中减去这个中心化后的均值，以此来计算差异的平方，这样做是为了简化计算过程。

3) 生成注意力权重：基于上面计算得到的 $S_{i,j}$ ，SimAM 通过以下公式生成注意力权重 $\omega_{i,j}$ ：

$$\omega_{i,j} = \frac{1}{1 + \exp\left(-\frac{1}{4}\left(\frac{s_{i,j}}{\sigma_{i,j}^2 + f} - 1\right)\right)} \quad (2)$$

其中, $\sigma_{i,j}^2$ 是 $s_{i,j}$ 的某种形式的归一化(在 SimAM 的实现中, 通常是通过整个特征图或局部区域的 $s_{i,j}$ 的平均值和标准差来近似), ϵ 是一个很小的常数。

4) 注意力图与特征图相乘: 将生成的注意力权重 $W \in R^{B \times 1 \times H \times W}$ 与原始特征图 X 相乘, 得到加权的特征图 $X' = W \odot X$ 。

2.1.3. 多尺度融合

在网络架构中增加多尺度特征提取模块, 将低层细节信息与高层语义信息进行融合, 确保系统对微小变化(例如局部肌肉运动)的捕捉更为敏感, 提高痛苦表情识别的准确性和鲁棒性。

2.2. 微表情的综合评估

动作单元是描述面部肌肉活动的标准化方法, 用于表征不同表情特征。选取的关键动作单元包括眼眶收紧(AU6)、上唇抬高(AU10)、嘴角上拉(AU12)、皱鼻子(AU4)、嘴角下拉(AU15)、下巴撇(AU17)、嘴巴闭合(AU23)、上眉毛上升(AU1 + AU2)以及嘴巴打开(AU27)。这些动作单元通过计算面部特征点之间的几何距离进行量化, 进而生成对应的 AU 分数。随后, 根据每个动作单元的重要性, 使用加权公式计算综合痛苦分数。当该分数超过预设阈值时, 系统判定为痛苦表情。计算公式如下:

$$\begin{aligned} \text{Pain Score} = & w_1 \cdot \text{AU1} + \text{AU2} + w_2 \cdot \text{AU4} + w_3 \cdot \text{AU6} + w_4 \cdot \text{AU10} + w_5 \cdot \text{AU12} \\ & + w_6 \cdot \text{AU15} + w_7 \cdot \text{AU17} + w_8 \cdot \text{AU23} + w_9 \cdot \text{AU27} \end{aligned} \quad (3)$$

其中 w 表示每个动作单元的权重, 反映其对痛苦表情的贡献度。系统会根据上述公式计算综合痛苦分数, 并与设定的阈值进行比较。如果综合痛苦分数超过阈值(例如 0.5), 则判定为痛苦表情, 触发报警机制。

3. 系统设计

3.1. 系统运行情况

采用 PC 作为处理终端, 作为摄像头人脸识别, 病房中安装的摄像头用于持续监测病人的面部状态, 并通过 WiFi 网络实时传输数据。系统配备的显示器能够显示具体的床号, 以便在需要时通知医护人员。若病人感到不适, 他们可以主动按下紧急按钮来发出警报。此外, 系统还能通过分析病人的面部表情来判断其痛苦程度, 当检测到痛苦表情时, 系统会自动触发报警机制, 在界面上提示紧急情况并通知相关监护人员, 并通过语音合成播放相应的报警信息, 提示相关床号已出现异常情况。这一功能在病人无法自行报警的情况下尤为重要。

3.2. 系统运行流程

本次研究的微表情, 通过人脸表情的活动单元: 眼眶收紧(AU6)、上唇抬高(AU10)、嘴角上拉(AU12)、皱鼻子(AU4)、嘴角下拉(AU15)、下巴撇(AU17)、嘴巴闭合(AU23)、上眉毛上升(AU1 + AU2)以及嘴巴打开(AU27)。这几部分来共同判断疼痛是否发生。系统运行流程图如图 3 所示。

1) 初始化冻结症患者实时监测系统, 将检测摄像头对准被看护人, 开始工作, 系统开始通过摄像头实时捕捉病人面部数据。

2) 面部数据采集。通过摄像头对病人进行实时监测, 系统调用摄像头实时获取视频流, 并逐帧处理,

使用 Mediapipe 的 FaceMesh 模块定位面部关键点，提取用于分析的 3D 坐标，每一帧的关键点数据被送入动作单元(AU)计算模块。

3) 系统并行收集视频帧，每 16 帧作为一个序列送入预训练的 3D ResNet 模型，动态分析结果结合 AU 结果，共同提高痛苦判定的准确性。

4) 将每个 AU 的数值按公式加权，计算综合痛苦分数，如果综合痛苦分数超过设定阈值，系统判定为发生紧急情况。

5) 当检测到痛苦表情，系统在报警界面显示报警信息，放大字体提示，并用语音提醒“床号 xx 有紧急情况”。并将紧急情况的信息发送给看护人员。

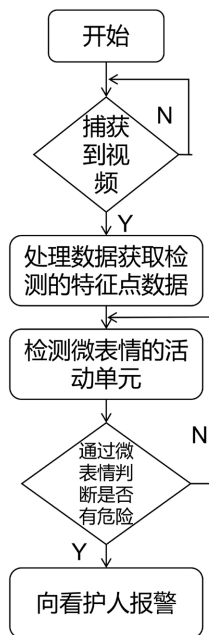


Figure 3. Flowchart of monitoring system operation

图 3. 监测系统运行流程图

3.3. 实验与分析

3.3.1. 数据集

由于目前缺乏专门针对渐冻症患者微表情监测的数据集，我们采取了其他方法进行研究。具体来说我们使用了一部分公开的面部表情数据集，并将其通过数据增强和仿真模拟进行扩展，还有一部分通过网络搜索到的渐冻症患者的面部表情，两者结合构成了本次使用的数据。

3.3.2. 实验结果

为了验证本文的渐冻症患者实时监测系统的检测精度、检测速度、分辨率等要素的准确可行性，在 pc 端上进行实验，本次用 pc 端作为显示端，通过自带摄像头，使用通过本文设计的监测系统，我们对本文的数据集进行了全面实验，经过实验测试表明，被看护人脸上出现痛苦的微表情时，系统会马上发出报警信息给看护人员，实验结果显示，系统在痛苦表情检测方面具有较高的准确性和较低的漏报率，同时在光照不足、噪声较多和运动模糊等复杂情况下，依然保持了良好的鲁棒性，在 NVIDIA GeForce RTX4060 显卡下，系统每秒能够处理 45 帧图像，在进行渐冻症患者面部微表情的实时监测时，处理速度的保持性和稳定性就是体现系统实时性的关键，系统的平均响应时间约为 1.2 秒，快速触发报警，有助于

实时反映患者的异常情况。与 YOLOv4 相比, 本系统的检测速度提高了大约 15%, 提供更多的实时应用潜力。1280 × 720 分辨率下, 系统处理速度为 45 FPS, 在 640 × 480 分辨率下, 系统性能基本保持一致, 但可以在处理速度上稍微有优势, 达到 50 FPS, 而在 1920 × 1080 分辨率下, 处理速度降到了 30 FPS, 精度未受显著影响。证明本文设计的监测系统可以增强监护的安全性。检测结果如下图(图 4~6)所示。



Figure 4. Monitor system initial interface
图 4. 监测系统初始界面



Figure 5. Monitoring system monitoring interface
图 5. 监测系统监测界面



Figure 6. Monitoring system alarm interface
图 6. 监测系统报警界面

4. 总结

本文设计了一种渐冻症患者实时监测系统。该系统将微表情识别技术融入了监护系统，除了现有的自动按铃和生理参数异常触发的报警机制外，新增了基于微表情的报警功能。本系统的价值主要体现在以下几个方面：首先，它满足了病人在遭遇痛苦时对医护人员及时响应的最高需求；其次，为病人构建了额外的安全保障——在特定情境下，当病人无法通过动作报警或生理监控出现故障时，监护系统能够自动发起报警；最为关键的是，这一系统显著提升了生命救援的宝贵时间，因为人体感知系统可以迅速捕捉到身体的变化，并通过痛苦表情表达出求助的信号，这在许多紧急场合具有决定性意义。

在隐私方面，首先，用户面部数据属于高度敏感信息，因此对数据存储和传输的安全性提出了极高要求，必须防止数据滥用或泄露，并严格遵守相关法律法规。其次，技术的普及过程中，用户对隐私保护、安全性和技术可靠性仍存顾虑。推广时应充分考虑用户心理和社会文化因素，通过提高技术透明度、赋予用户更大的控制权，从而增强公众对该技术的接受度。这些措施均为确保技术广泛应用和建立信任奠定了关键基础。

本文将致力于优化网络结构、改进算法，并探索新的模型设计思路，目标是在保持高精度的同时，实现实时的监测，后续会密切关注技术的前沿发展，并将其应用于相关的研究中。同时，隐私保护和技术规范的构建应始终以用户需求为核心。未来，透明化算法设计和用户可控的数据管理机制将成为增强用户信任的关键方向。同时，通过深入研究隐私保护算法(如差分隐私)和可信数据交互机制，可以更有效地保障用户数据在采集、存储与传输过程中的安全性。

基金项目

本研究得到西安市科技计划软科学研究一般项目(项目编号：24RKYJ0065)和陕西省大学生创新创业训练计划项目(项目编号：S202312715034)：渐冻症患者实时监测系统的支持。

参考文献

- [1] Egede, J., Valstar, M. and Martinez, B. (2017) Fusing Deep Learned and Hand-Crafted Features of Appearance, Shape,

- and Dynamics for Automatic Pain Estimation. 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). Washington, 30 May-3 June 2017, 689-696. <https://doi.org/10.1109/FG.2017.87>
- [2] 刘芳, 李俊吉. 融合 VGG 与注意力的学生微表情识别和情绪评估方法[J]. 现代计算机, 2024, 30(18): 28-33.
- [3] Zhou, J., Hong, X., Su, F., *et al.* (2016) Recurrent Convolutional Neural Network Regression for Continuous Pain Intensity Estimation in Video. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, 26 June-1 July 2016, 84-92. <https://doi.org/10.1109/CVPRW.2016.191>
- [4] 郝强. 基于双流特征深度学习的新生儿疼痛表情识别[D]: [硕士学位论文]. 南京: 南京邮电大学, 2019.
- [5] 周婕, 马明栋. 基于改进的 ResNet 网络的人脸表情识别[J]. 计算机技术与发展, 2022, 32(1): 25-29.
- [6] 罗明刚, 李一民, 曾素娣. 基于 Adaboost 算法的人脸检测研究[J]. 计算机与数字工程, 2007, 35(2): 7-8, 72.