

混合卷积-长短期记忆网络模型在实验室违规行为检测中的应用



渠连恩^{1, 2, 3,*}, 李梦月¹, 陈新¹, 陈悦容¹

¹青岛科技大学国际学院(泰中国际橡胶学院)科技创新实验室, 山东青岛 266061

²青岛科技大学信息科学与技术学院, 山东青岛 266061

³青岛科技大学人工智能与计算机视觉实验室, 山东青岛 266061

摘要: 本研究针对实验室人员违规行为检测, 提出了一种基于 CLSTM 模型的方法。该模型结合了卷积神经网络 (CNN) 和长短时记忆网络 (LSTM)。模型首先利用 CNN 感知图像特征, 并通过卷积和池化操作进行特征提取。这些提取的特征随后传递至 LSTM 网络模型, 利用 LSTM 的记忆能力在时间序列中捕捉和理解违规行为的模式。通过这种结构, CLSTM 模型不仅能够有效处理和分析复杂的图像数据, 还能在长期记忆和短期变化之间找到平衡, 提高违规行为检测的准确性和鲁棒性。模型通过卷积网络增强了图像特征的表征能力, 进一步优化了 LSTM 模型在计算能力和训练效率上的表现。相较于传统的 LSTM 模型, CLSTM 在训练过程中展现出更快的收敛速度。为验证模型的有效性, 本文在两个不同的数据集上进行了实验。实验结果显示, 所提出的 CLSTM 模型在性能上显著优于传统的 LSTM 方法, 性能提升幅度分别约为 7% 至 11%。这一结果突显了 CLSTM 模型在实验室违规行为检测任务中的优越性。

关键词: CLSTM; 卷积网络; LSTM 网络; 不当行为检测

DOI: [10.57237/j.cst.2024.03.003](https://doi.org/10.57237/j.cst.2024.03.003)

Application of Convolutional Long Short-Term Memory Network Model in Laboratory Violation Behavior Detection

Lianen Qu^{1, 2, 3,*}, Mengyue Li¹, Xin Chen¹, Yuerong Chen¹

¹Technology Innovation Lab of International College (Sino-Thai International Rubber College), Qingdao University of Science and Technology, Qingdao 266061, P.R.C.

²College of information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, P.R.C.

³Artificial Intelligence and Computer Vision Lab (AICV), Qingdao University of Science and Technology, Qingdao 266061, P.R.C.

Abstract: This study proposes a method based on the CLSTM model for detecting misconduct among laboratory personnel. The model integrates Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. Initially, the CNN perceives image features and performs feature extraction through convolutional and pooling

*通信作者: 渠连恩, quyufan@163.com

operations. These extracted features are subsequently fed into the LSTM network, which leverages its memory capability to capture and understand patterns of misconduct in the temporal sequence. This structure enables the CLSTM model not only to effectively process and analyze complex image data but also to strike a balance between long-term memory and short-term variations, thereby enhancing the accuracy and robustness of misconduct detection. The model enhances the representation capability of image features through the convolutional network, further optimizing the LSTM model's performance in computational efficiency and training effectiveness. Compared to traditional LSTM models, CLSTM demonstrates faster convergence during training. To validate the model's effectiveness, experiments were conducted on two different datasets. The results indicate that the proposed CLSTM model significantly outperforms traditional LSTM methods, achieving performance improvements ranging from 7% to 11%. These findings underscore the superiority of the CLSTM model in laboratory personnel misconduct detection tasks.

Keywords: CLSTM; CNN; LSTM; Improper Behavior Detection

1 引言

在当前数字时代，违规行为的监测和识别对于维护秩序和保障安全至关重要。特别是在复杂、动态的环境中，传统的监测方法往往无法满足管理和治理领域的实时监测需求。随着深度学习技术的兴起，视频监控领域的深度学习方法为违规行为检测带来了新的可能性。卷积神经网络（CNN）[1]和长短期记忆网络（LSTM）[2]的结合尤为突出，为图像处理和时间序列分析提供了强大的解决方案。

本文提出了 CLSTM 的网络模型，目的是克服传统方法在复杂环境中的局限性。CLSTM 模型融合了卷积操作在图像处理中的卓越性能和 LSTM 在序列建模中的优势，显著提高了违规行为的识别准确率并有效提升了管理效率[3, 4]。本文的主要贡献包括：

1) 模型创新：提出了基于 CLSTM 的混合模型，有效整合了视频图像的空间特征和时间序列数据，实现了更为精准的违规行为检测。通过深入探讨该模型的设计原理、实验验证以及应用场景，为从业者提供了一种高效违规行为检测方法。

2) 实验验证：通过详细的实验和案例分析，验证了 CLSTM 模型的稳健性和可靠性，展示了其在实验室安全管理中的实际应用潜力。该模型的应用场景不仅局限于实验室环境内的违规行为检测。

通过本文的研究，我们致力于推动实验室安全监控系统向智能化、高效化方向迈进，为管理者提供更可靠的工具和方法，以确保实验室环境的安全与秩序。

2 相关文献

Khaleghi 等[5]人提出了一种基于深度学习的视频监控摄像头异常检测方法，强调其在事件预测、在线监控和目标驱动分析方面提供有价值的数据。他们的方法利用人工智能技术，尤其是深度学习中的卷积神经网络，显著提高了异常检测的准确性。Lea 等[6]任提出了一种用于视频动作分割的创新方法。该网络通过分层方式捕捉低、中、高级时间尺度上的关系，相较于传统的循环神经网络(RNN)[7]训练所需时间较少。Simonyan 等[8]人进行了对深度卷积网络（ConvNets）在视频动作识别方面的判别性训练架构的研究，并提出了一种创新的双流 ConvNet 架构，其中包含了空间和时间网络。通过在多帧稠密光流上进行训练，他们证明了该 ConvNet 在训练数据有限的情况下能够取得良好的性能。此外，他们还展示了多任务学习在两个不同的动作分类数据集上的应用，以增加训练数据量并提高性能。Yue 等[9]人提出了两种能够处理完整视频的方法。第一种方法探索了各种卷积时序特征池化架构，研究了在调整 CNN 进行这项任务时需要做出的各种设计选择。第二种方法明确地将视频建模为帧的有序序列。

Tee 等[10]人提出了一种基于深度学习架构的最新人类活动识别模型，该模型包含卷积神经网络(CNN)、长短时记忆(LSTM)或多种类型混合的层，形成混合系统。分析对这些模型的实施方式进行了概述，并介绍了它们可能面临的一些潜在限制。Nagrath [11]提出了一种深度学习方法，目的是检测学校或大学校园中的异常行为。该方法通过提取视频相邻帧来实现监控，模型在这些摄像头帧上执行以检测异常活动。发现异

常, 系统开始保存视频流并向相关当局发送警报消息。通过在实时摄像头流中计算特征, 分类器基于提取的特征进行异常预测。

Einy [12]提出了一种创新的深度学习方法, 用于在时间延迟系统中检测体外受精中胚胎的同时异常, 从而实现存活和非存活出生的区分。该方法采用了局部二值卷积神经网络(LBCNN)和长短时记忆(LSTM)。LBCNN 通过利用深度和局部特征集显著提高了分类准确性, 与标准卷积层相比, 具有更少的可学习参数。

Yang [13]提出了一种融合了长短时记忆和卷积神经网络的混合网络, 采用非线性数据处理方法, 并通过数值反演(NI)方法填补缺失值。这一方法结合了相关分析、敏感性分析和随机森林(RF) [14]算法。

现代视频监控系统主要依赖人工操作员实时监视和解读个体行为, Kalyta 等[15]人出了一种简单的模型, 用于使用监控摄像头的低分辨率照片和视频帧检测个体情感状态的突然变化。所提出的技术包括一种对面部区域进行几何解释以提取面部表情特征的方法, 使用超平面分类方法在特征向量空间中识别情感状态的方法, 以及利用“人在环”原则获取透明和可解释分类器的视觉分析方法。

3 模型架构

本文提出的深度学习模型包括卷积神经网络(CNN)和长短期记忆网络(LSTM), 用于实现对违规行为的精确识别和分类, 如图 1 所示。首先, 输入层接收图像数据, 经过卷积层进行特征提取, 这些卷积层通过不同的滤波器捕捉图像的局部特征。随后的最大池化层用于降低特征映射的空间维度, 增强特征的表达能力。卷积网络处理的特征序列化后传递给 LSTM 模型, LSTM 利用其记忆单元捕捉时间序列数据中的长期依赖关系和动态模式。最后, 通过全连接层实现最终的违规行为分类任务, 输出每个类别的概率分布。

整体模型设计的核心包括图像特征提取、特征显著性增强、时间序列特征建模以及高效的多类别分类能力, 为复杂环境下违规行为检测提供了强大的解决方案。

模型的核心是卷积神经网络, 其主要由卷积层和池化层交替构成。CNN 通过学习输入数据与输出数据之间复杂的非线性映射关系来实现图像处理任务。网

络结构包括输入层、中间层和输出层, 其中, 中间层用于特征提取和学习。输入层接收图像信息, 而输出层则输出分类标签或识别结果。

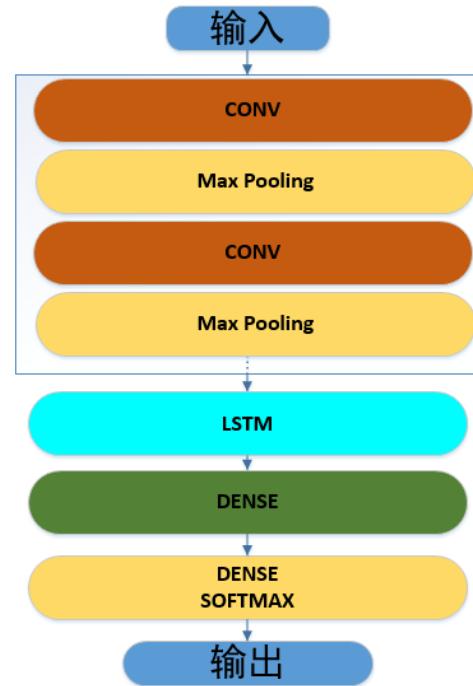


图 1 CLSTM 模型总体结构

卷积操作和池化操作是 CNN 的核心步骤, 用于感知图像中的关键特征并理解图像内容。卷积层利用不同尺寸和参数的卷积核来捕捉图像中的局部特征, 而池化层通常采用最大池化或平均池化等方法, 以降低数据的维度并增强特征的稳定性。

全连接层和激活函数位于卷积神经网络隐含层的最后部分, 全连接层的作用是将卷积层和池化层提取的特征信息进行融合和整合, 以进一步形成网络的输出。全连接层将特征向量连接到一起, 使得网络可以更好地理解和利用这些特征。

激活函数是全连接层中的一个关键组成部分, 该模型引入了非线性元素 ReLU 作为激活函数。该激活函数的作用在于使全连接层能够学习和表示复杂的非线性关系, 从而使卷积神经网络具备拟合任意非线性函数的能力。激活函数的引入有助于网络更好地捕捉和表示数据中的复杂模式和特征。

该模型的另一个重要组成部分是长短期记忆网络(LSTM), 它是一种改进的循环神经网络(RNN)结构。LSTM 通过引入细胞状态、隐藏状态和门结构来解决传统 RNN 中的梯度消失和梯度爆炸问题。这种设计

使得 LSTM 能够更有效地处理长序列信息，避免了传统 RNN 中容易出现的“顾此失彼”问题。

LSTM 的关键是其内部结构，它由三个门控制单元组成：遗忘门、输入门和输出门。遗忘门决定哪些信息需要被丢弃，使用一个激活函数来对当前输入和前一个时间步的隐藏状态进行处理，输出一个介于 0 到 1 之间的数值，其中 0 表示完全忘记，1 表示完全记住。遗忘门的引入使得 LSTM 可以有选择地保留重要信息，从而有效地处理梯度问题并学习长期依赖关系。

LSTM 单元的计算公式(1)所示，它通过门结构来更新和管理细胞状态和隐藏状态，这些状态允许网络更好地捕捉和利用时间序列数据中的长期依赖关系。

$$f_t = \sigma(w_f \bullet [h_{t-1}, x_t] + b_f) \quad (1)$$

其中： σ 是 sigmoid 函数，用于将输入的加权和转换为一个介于 0 和 1 之间的数值，表示门的遗忘程度。遗忘门决定了上一个细胞状态 c_{t-1} 中的信息有多少会被保留或遗忘。它输出一个介于 0 和 1 之间的数值 f_t ，其中 0 表示完全遗忘，1 表示完全保留。 w_f 是输入门的权重矩阵， h_{t-1} 是上一个时间步的隐藏状态， x_t 是当前时间步的输入， b_f 是输入门的偏置项。

输入门在循环神经网络中的作用是控制当前时间步的输入信息如何融合到细胞状态中。输入门的状态方程(2)：

$$i_t = \sigma(w_i \bullet [h_{t-1}, x_t] + b_i) \quad (2)$$

其中： i_t 是输入门的输出，是介于 0 和 1 之间的数值，表示在当前时间步 t 上，模型选择多少来更新细胞状态。 σ 是 sigmoid 函数，用于将输入的加权和转换为一个介于 0 和 1 之间的数值，表示门的打开程度。 w_i 是输入门的权重矩阵。 h_{t-1} 是上一个时间步的隐藏状态。 x_t 是当前时间步的输入， b_i 是输入门的偏置项。

输入门通过控制的大小，决定了在当前时间步的输入 x_t 对细胞状态的影响程度。输入门的作用是增强或减少当前输入对细胞状态的更新影响，从而帮助模型更好地捕捉时间序列数据中的长期依赖关系，更新候选值方程(3)：

$$\tilde{c}_t = \tanh(w_k \bullet [h_{t-1}, x_t] + b_c) \quad (3)$$

更新候选值计算了可以添加到细胞状态的新信息。通过使用 \tanh 函数将输入转换为介于 -1 和 1 之间的值，以便更好地控制细胞状态的更新。

细胞状态 c_t 在每个时间步根据遗忘门和输入门的决定进行更新。遗忘门决定了保留多少上一时刻的细胞状态信息，输入门决定了添加多少新的信息到细胞状态中，细胞状态更新方程(4)。

$$c_t = f_t \bullet c_{t-1} + i_t \bullet \tilde{c}_t \quad (4)$$

之后再由输出门计算出更新之后的隐藏状态和细胞状态，状态方程(5)。

$$o_t = \sigma(w_o \bullet [h_{t-1}, x_t] + b_o) \quad (5)$$

输出门决定了当前时间步的隐藏状态 h_t 是基于更新后的细胞状态的哪一部分。它输出一个介于 0 和 1 之间的数值 o_t ，其中 0 表示没有输出，1 表示完全输出。

隐藏状态 h_t 是基于当前时间步的细胞状态 c_t 和输出门的决定来计算得到的。它是 LSTM 单元最终输出的隐藏状态，用于传递给下一个时间步或者作为模型的输出，隐藏状态更新方程(6)。

$$h_t = o_t \tanh(c_t) \quad (6)$$

通过这些门控结构，LSTM 单元能够有效地处理和传递长期依赖关系，使得循环神经网络能够更好地应对时间序列数据中的复杂模式和长期记忆需求。

本文采用的 CLSTM 的模型来处理图像序列，其中每帧图像首先通过一个 3×3 的卷积核进行步长为 1 的卷积操作，将输入图像的通道数从 3 扩展到 64。接着，经过一个 2×2 的最大池化层，这一步操作旨在降低每个通道的空间维度，同时将特征图的尺寸减小为原来的一半，但通道数翻倍，从而有效地增强了每帧图像的特征信息。

随后，特征向量被送入一个全连接层，通过该层得到一个 256 维的特征向量。这些特征向量被看作是时间序列的一部分，然后被输入到一个包含 40 个 LSTM 细胞单元的 LSTM 模型中。每个 LSTM 单元负责学习和记忆输入序列中的时间相关性和长期依赖关系。最后一个 LSTM 细胞的输出被用作全连接层的输入，通过全连接层和非线性激活函数对 LSTM 计算的结果进行最终的分类操作。

4 实验配置

在研究综合考虑了硬件和软件配置, 选用了适当的深度学习框架, 以及详细的模型验证指标, 以确保模型的有效性和性能。

在软件配置方面, 为了确保系统的一致性和稳定性, 并设计本文的行为识别模型, 我们首先选择了 Torch 作为主要的模型构建工具。所有这些软件环境和工具都是在 Anaconda 中进行配置和构建的。我们使用了 CUDA 11.03、OpenCV-python 4.5.5、Torch 1.10 以及 Python 3.9.7 来支持视频的读取和处理任务。

硬件方面, 本研究使用了一台配置为第 12 代 Intel® Core™ i7-12700x20 CPU、64GB 内存和 Nvidia RTX A4000 GPU 的机器上进行的。

本文涉及使用 UCF-101 数据集和布朗大学发布的数据集, 分别包含不同动作类别, 对模型进行性能评估和比较。UCF-101 数据集以复杂的人际和人物交互动作作为特点, 比较具有挑战性。

模型的验证指标主要分为两部分。首先是训练过程中的损失值, 即训练标签真值与模型预测值之间的差异, 用来评估模型在训练过程中的收敛情况。其次是在测试集上的准确率, 用来衡量模型在新数据上的预测能力。为了进一步细化分析实验结果, 将整个测试集按照 8:2 的比例分为训练集和验证集, 分别用于训练模型和验证性能。

为了突出 CLSTM 模型的性能, 本文进行了与 LSTM 模型的比较, 主要从收敛速率和准确度两个方面进行对比。CLSTM 模型在架构上将卷积神经网络部分用于空间特征提取, 并将其结果输入 LSTM 进行时间特征提取, 以综合利用空间和时间信息。

在训练过程中, 采用了交叉熵损失函数和随机梯度下降优化器来优化模型。总共进行了 150 次训练迭代, 每 5 次迭代在测试集上进行一次性能评估。

在模型验证过程中, 重点关注了两个模型之间的

比较, 包括损失值的变化趋势和在测试集上的准确性表现, 以全面评估模型的预测性能。对不同动作类别的准确性进行了详细分析, 以更深入地理解实验结果的差异和模型的优劣势。

5 实验结果

(1) UCF-101 数据集

基于 UCF-101 数据集的实验结果显示, CLSTM 模型在性能和训练效率上优于 LSTM 网络, 如图 1 和 2 所示。在训练过程中, CLSTM 模型表现出更稳定的收敛性能和更低的训练损失下限。在经过 80 个 epoch 的训练后, CLSTM 模型的训练损失已接近 0.2, 优于 LSTM 模型。此外, CLSTM 模型的收敛速度更快, 且在损失值收敛后的波动幅度较小, 这表明其在学习过程中对数据的拟合能力更强。

图 1 展示了训练损失随时间的变化, 可以看出 CLSTM 模型相较于 LSTM 模型具有更快的损失下降速度和更稳定的收敛状态。图 2 则展示了在测试集上的准确率对比, 显示出 CLSTM 模型在相同的测试条件下能够获得更高的预测准确性。

在训练准确率对比中, LSTM 的训练准确率上升速度慢且最后约为 75.33%, 而 CLSTM 在整个训练过程中测试准确率全程高于 LSTM 模型且最后测试准确率约为 86.94%, 如表 1 所示。根据表格 1 中的数据, CLSTM 模型在整个测试集上的准确率比 LSTM 模型提升了超过 9%。CLSTM 与 LSTM 在对这 6 个动作进行识别分类时的综合准确率如表 2 所示。

总体而言, 这些结果表明 CLSTM 模型通过有效地整合卷积神经网络的空间特征提取能力和 LSTM 的时间序列建模能力, 显著改善了动作识别任务的性能。这种改进不仅提升了模型的训练效率和收敛速度, 还提高了在复杂动作分类中的准确性, 为动作识别领域的研究和应用带来了显著的进步。

表 1 LSTM 模型与 CLSTM 在准确率比较 (UCF-101)

模型	准确率 (UCF-101)
LSTM	75.33%
CLSTM	86.94%

表 2 LSTM 模型与 CLSTM 模型的准确率比较 (UCF-101)

模型	精确度 (%) (UCF-101)					
	刷牙	拳击	跳跃	刮胡子	太极	写字
LSTM	80.21	74.55	79.2	69.45	77.52	71.05
CLSTM	82.1	89.5	90.2	81.45	90.27	88.13

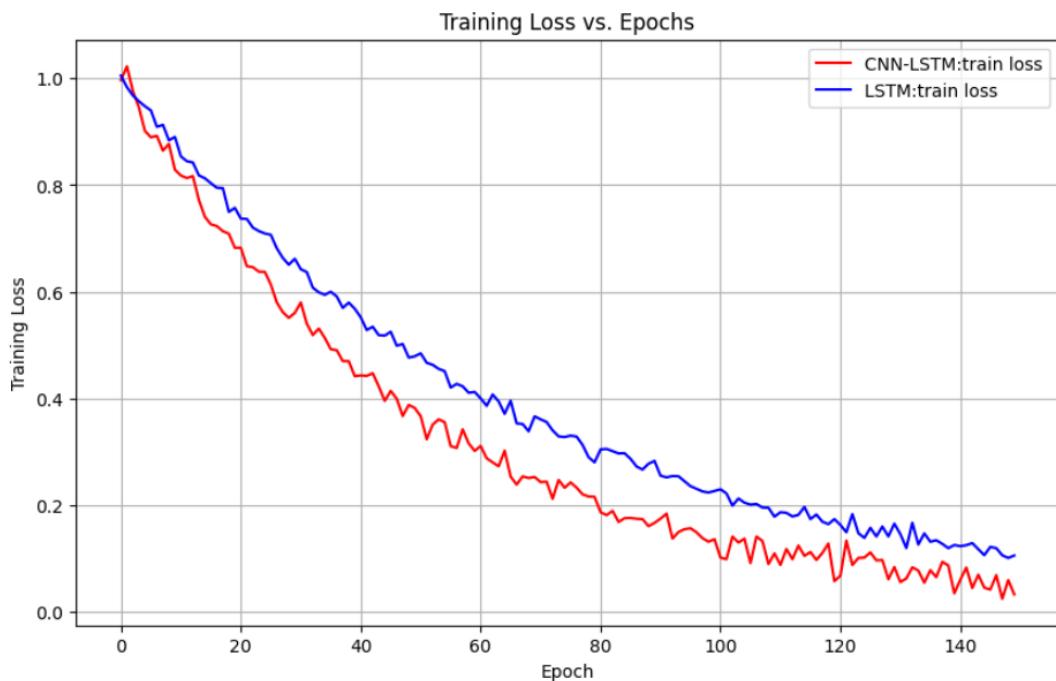


图 1 CLSTM 与 LSTM 模型训练 Loss 值 (UCF-101)

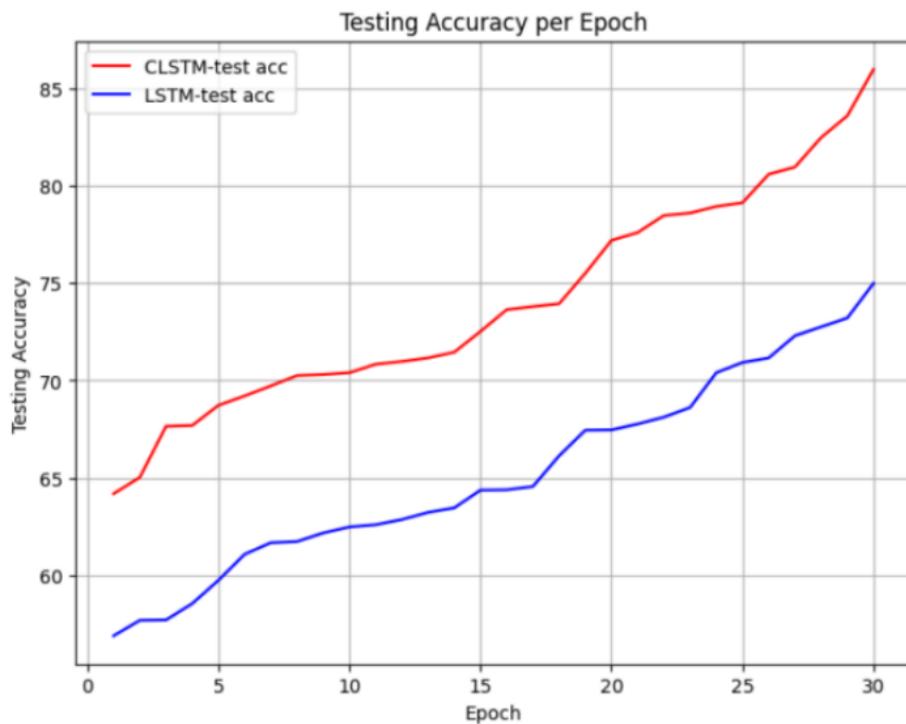


图 2 LSTM 与 CLSTM 模型测试正确率比较 (UCF-101)

根据表格数据显示, CLSTM 在需要同时考虑图像特征和时序建模的识别任务中表现突出。例如, 在跳跃和太极两种动作中, 其准确率达到了约 90%。而在刷牙和刮胡子这两个类别中, CLSTM 表现相似, 准确率约为 81%。分析显示, CLSTM 模型在具有明显空间

特征 (如跳跃) 或较大运动幅度 (如太极) 的类别中表现优异, 而在动作幅度较小的刷牙和刮胡子类别中则表现基本一致。

(2) 布朗大学数据集

根据布朗大学数据集上的比较结果, CLSTM 模型

展现了更高的训练效率、稳定的收敛速度以及优越的测试准确率, 如图 3 和 4 所示。然而, 在处理某些细

微动作或容易混淆的情况下, 其表现可能不如传统的 LSTM 模型, 如表 4 所示。

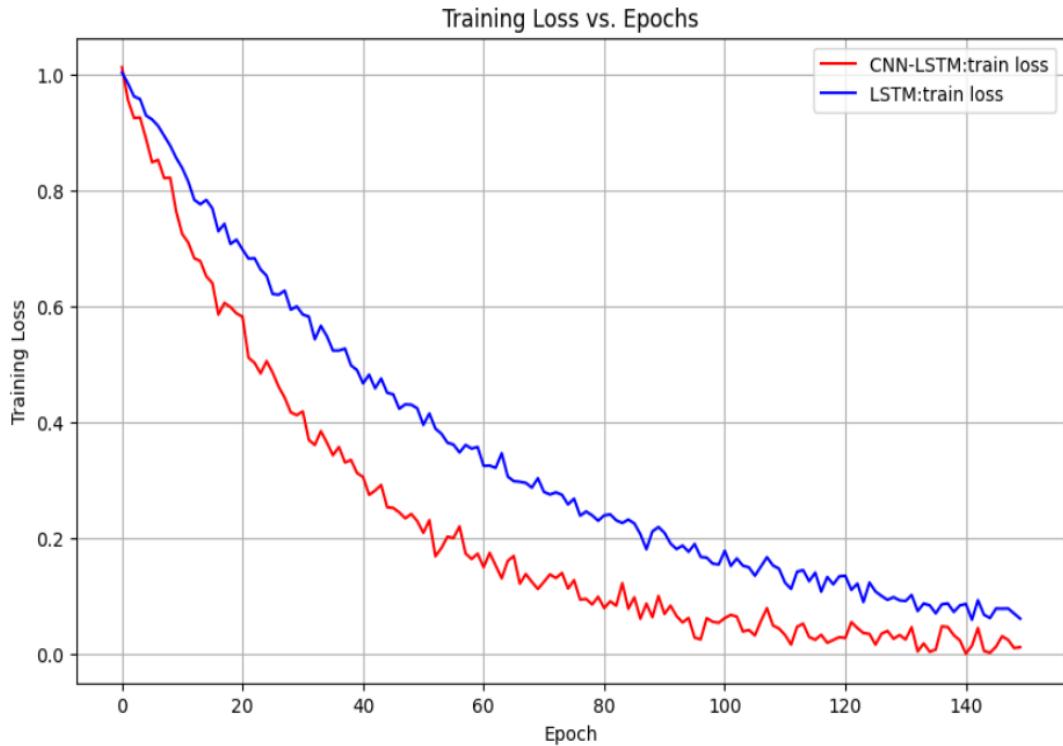


图 3 LSTM 与 CLSTM 模型训练 Loss 值比较

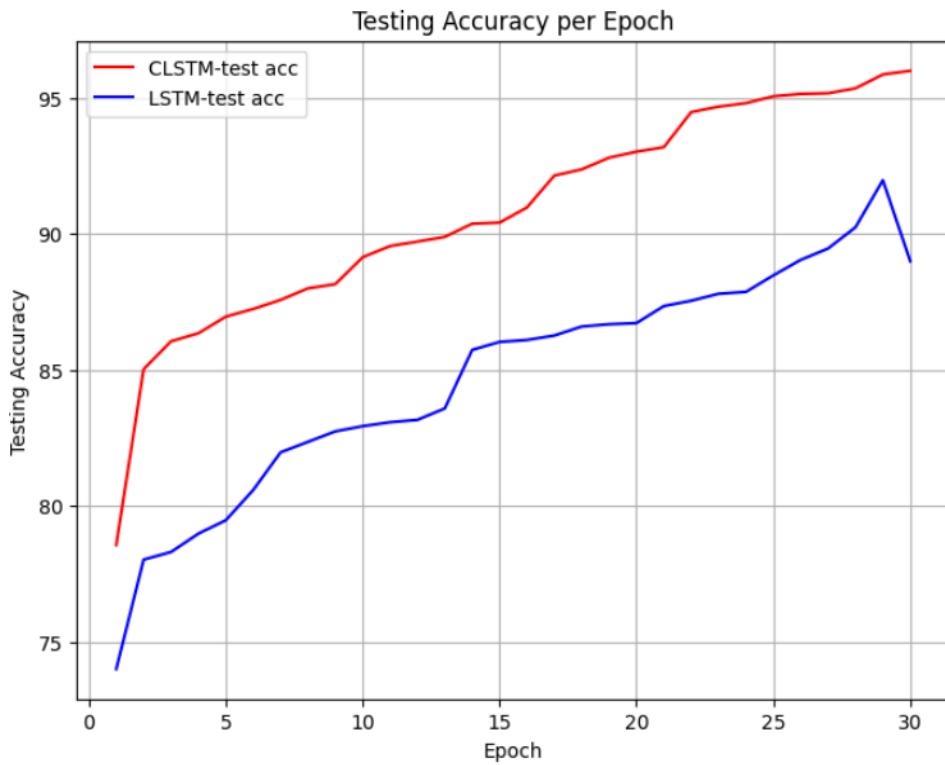


图 4 LSTM 与 CLSTM 模型测试正确率比较

表 3 LSTM 模型与 CLSTM 模型的准确率比较

模型	准确率 (布朗大学数据集)
LSTM	89.13%
CLSTM	96.1%

表 4 LSTM 模型与 CLSTM 模型准确率比较

模型	精确度 (%) (布朗大学数据集)					
	喝水	跳跃	快跑	抽烟	扔	进食
LSTM	90.71	91.47	90.1	78.59	93.54	90.43
CLSTM	99	100	98	99.45	89.43	84.92

在布朗大学数据集该数据集上的表现如下：

- 1) 训练效率和收敛速度: CLSTM 模型的训练 loss 收敛速度迅速且稳定, 约在 120 个 epoch 后达到了较低的 0.1 附近, 比 LSTM 模型提前了大约 40 个 epoch, 如图 3 所示。
- 2) 测试准确率: CLSTM 模型在整个训练过程中的测试准确率始终高于 LSTM 模型, 最终测试准确率约为 96.1%, 而 LSTM 模型的最终训练准确率仅约为 89.13%, 如图 4 所示。
- 3) 综合准确率对比: 根据表格 3 的数据, CLSTM 模型在整个测试集上相对于 LSTM 模型有了约 7% 以上的性能提升。
- 4) 类别特定表现: 根据表格 3 的数据, CLSTM 模型在跳跃、快跑和抽烟等动作类别中表现优异, 甚至在某些动作如跳跃中达到了 100% 的准确率。

CLSTM 与 LSTM 在对这 6 个动作进行识别分类时的综合准确率如表 4 所示。

6 结论

本文介绍了一种结合了卷积神经网络和长短期记忆网络的模型实现对化工实验室内人员违规行为的识别和分类检测。该模型充分利用了 CNN 在图像感知和学习方面的能力, 以及 LSTM 在时序信息学习方面的优势。通过这种结合, 提升了对复杂行为模式的识别和理解能力。

CLSTM 模型在多个主要性能指标上表现出色, 相较于 LSTM 神经网络, 取得了显著的提升。这表明了在行为识别任务中, 卷积神经网络与 LSTM 网络的有效融合能够显著改善模型的性能和准确性。

未来的研究方向应重点关注开发适应不同场景和需求的算法, 以实现特定行为的可靠检测。这包括对

模型的进一步优化, 适应复杂环境和动态变化的行为模式, 以及提升模型在实时性和鲁棒性方面的表现。

致谢

本文为青岛科技大学 2022 年教学改革研究项目《基于深度学习教与学行为的可视化研究》成果, 感谢青岛科技大学教学改革研究项目的支持。

参考文献

- [1] Hershey S, Chaudhuri S, Ellis D P W, et al. CNN architectures for large-scale audio classification [C] // 2017 ieee international conference on acoustics, speech and signal processing (icassp). IEEE, 2017: 131-135.
- [2] Graves, Alex, and Alex Graves. "Long short-term memory." Supervised sequence labelling with recurrent neural networks (2012): 37-45.
- [3] Huang, Shu-Chun L., and Wan-En Sun. "Exploration of social media for observing improper tourist behaviors in a national park." Sustainability 11.6 (2019): 1637.
- [4] Wang M, Zhao Y, Wu Q, et al. A YOLO-based Method for Improper Behavior Predictions [C] // 2023 IEEE International Conference on Contemporary Computing and Communications (InC4). IEEE, 2023, 1: 1-4.
- [5] Khaleghi, Ali, and Mohammad Shahram Moin. "Improved anomaly detection in surveillance videos based on a deep learning method." 2018 8th Conference of AI & Robotics and 10th RoboCup Iranopen International Symposium (IRANOPEN). IEEE, 2018.
- [6] Lea, Colin, et al. "Temporal convolutional networks: A unified approach to action segmentation." Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8–10 and 15–16, 2016, Proceedings, Part III 14. Springer International Publishing, 2016.

- [7] Grossberg, Stephen. "Recurrent neural networks." Scholarpedia 8.2 (2013): 1888.
- [8] Simonyan, Karen, and Andrew Zisserman. "Two-stream convolutional networks for action recognition in videos." *Advances in neural information processing systems* 27 (2014).
- [9] Yue-Hei Ng, J., Hausknecht, M., Vijayanarasimhan, S., Vinyals, O., Monga, R., & Toderici, G. (2015). Beyond short snippets: Deep networks for video classification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4694-4702).
- [10] Tee, Wei Zhong, et al. "A Close Look into Human Activity Recognition Models using Deep Learning." 2022 3rd International Conference on Computing, Networks and Internet of Things (CNIOT). IEEE, 2022.
- [11] Nagrath, P., Dharana, Dwivedi, S., Negi, R., & Singh, N. (2022). Real-Time Anomaly Detection Surveillance System.
- In Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1 (pp. 665-678). Springer Singapore.
- [12] Einy, S., Sen, E., Saygin, H., Hivechi, H., & Dorostkar Navaei, Y. (2023). Local Binary Convolutional Neural Networks' Long Short-Term Memory Model for Human Embryos' Anomaly Detection. *Scientific Programming*, 2023.
- [13] Yang, J., Qu, J., Mi, Q., & Li, Q. (2020). A CNN-LSTM model for tailings dam risk prediction. *IEEE Access*, 8, 206491-206502.
- [14] Rigatti, Steven J. "Random forest." *Journal of Insurance Medicine* 47.1 (2017): 31-39. (RF).
- [15] Kalyta, O., Barmak, O., Radiuk, P., & Krak, I. (2023). Facial Emotion Recognition for Photo and Video Surveillance Based on Machine Learning and Visual Analytics. *Applied Sciences*, 13(17), 9890.