

化学分析实验中的行为识别研究

李晓旭, 马兴录, 孙昊

青岛科技大学, 信息科学技术学院, 山东 青岛

收稿日期: 2022年3月21日; 录用日期: 2022年4月22日; 发布日期: 2022年4月29日

摘要

为了提高化学实验室的智能化水平, 本文对双流卷积网络进行研究, 提出一种应用在化学分析实验领域的行为识别网络。本文网络将输入视频分为RGB和光流, 通过改进的EfficientNetv2网络加深网络结构提取特征信息, 同时探究双流网络的融合位置, 改进损失函数, 确定网络的最佳策略。经验证在自制的数据集集中改进的双流卷积神经网络模型准确率可以达到92.4%, 相比常规的行为识别算法具有更高的识别率。

关键词

动作识别, 光流, 双流卷积, 融合

Research on Action Recognition in Chemical Analysis Experiments

Xiaoxu Li, Xinglu Ma, Hao Sun

Information Science and Technology, Qingdao University of Science and Technology, Qingdao Shandong

Received: Mar. 21st, 2022; accepted: Apr. 22nd, 2022; published: Apr. 29th, 2022

Abstract

In order to improve the intelligence level of the chemical laboratory, this paper studies the two-stream convolutional network and proposes an action recognition network applied in the field of chemical analysis experiments. In this paper, the network divides the input video into RGB and optical flow, and uses the improved EfficientNetv2 network to deepen the network structure to extract feature information. At the same time, it explores the fusion position of the dual-stream network, improves the loss function, and determines the best strategy for the network. It has been verified that the accuracy rate of the improved two-stream convolutional neural network model in the self-made dataset can reach 92.4%, which is higher than the conventional behavior recognition

algorithm.

Keywords

Action Recognition, Optical Flow, Two-Stream Convolution, Fusion

Copyright © 2022 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

进入智能时代，深度学习发展迅速，是国家科技大力发展的推动器，在各行业应用广泛，而随着我国人口老龄化的加剧，“三孩”政策全面实行，国家继续从高速发展转变为高质量发展，在此背景下，人工智能的广泛推广使用不仅可以改善劳动条件，提高劳动效率，完成预期各种活动工作。动作识别作为计算机视觉一大重要课题具有广泛的应用前景，而实验室的智能化应用却不足[1] [2]，对于枯燥频繁或有一定危险的化学分析实验过程需要与机械臂结合，减少人力浪费，保护人员安全，相关研究员已经将人工智能和机械臂技术结合在化学实验应用方面进行了研究[3] [4] [5]，化学实验室智能化也将迎来新的发展。

近年来，人体行为识别的算法主要有两类，传统的手工提取特征方法和基于神经网络的方法，将图像识别的方法应用在行为识别中也屡见不鲜。早期的动作识别研究算法在 RGB 图像的基础上进行，一些研究者是利用机器学习的方法手动提取视频动作信息，通过机器学习分类算法进行动作识别，方法费时费力然而准确率也不高。随着图像识别在深度学习上不断出现新的研究进展，许多研究者将深度学习引入动作识别中。大致分为三个方向：基于双流卷积神经网络，基于三位卷积神经网络，基于短时记忆网络。而动作识别与深度学习结合后的处理速度和准确率得到飞速提升的同时，在高准确率的实现背后是计算量的增大和运算速度的减小，加大了计算成本，限制了神经网络算法的应用以及最新深度学习的发展[6] [7] [8] [9] [10]。

动作识别经典算法双流卷积神经网络效仿人体视觉识别，在短视频识别领域具有较好的识别效果，在 UCF101 数据集上取得了 88.0% 的识别率，其应用也十分广泛，对其的改进在近年来也不断提高着双流卷积网络的深度和准确率[11]-[16]，陈颖和谢家龙等人将双流卷积网络和 3D 卷积神经网络结合，其方法增加了特征的提取程度，特征信息得到了利用的最大化，但处理的信息也增大了几倍。曾明如等人将双流卷积和循环神经网络结合，但网络所采用的算法也存在特征提取不充分和泛化能力低等问题。针对此问题，本文对双流卷积算法网络进行了实验研究，通过改进特征提取网络和损失函数，提高算法的识别准确率。

2. 改进的动作识别算法

2.1. 双流卷积

Simonyan 等人在 Karpathy 的基础上提出了视频动作识别模型双流卷积神经网络，该模型创新性地提出了通过两个独立网络提取特征，一路进行 RGB 图像的特征提取，一路进行光流特征提取，实现了提取时间序列特征。其模型如下图 1 所示。

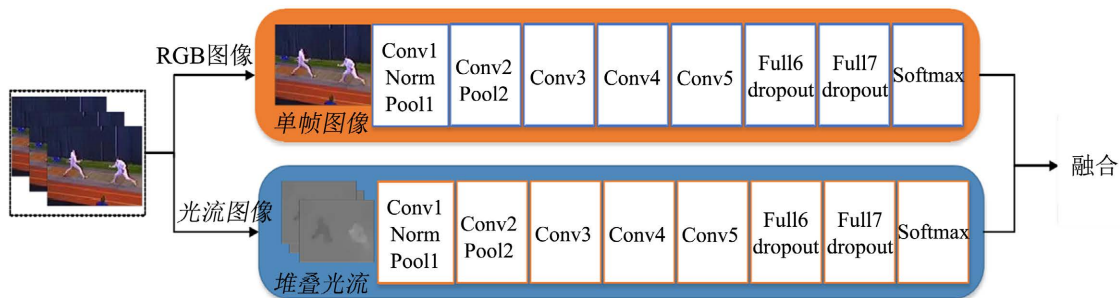


Figure 1. Two-stream convolutional neural network
图 1. 双流卷积神经网络

RGB 图像流处理的是单帧的视频图像，提取一段视频样本内视频帧的人体动作特征信息，通过静态图可以识别到动作视频内的大部分特征，视频帧的场景和人的位姿都对动作分类有较大贡献，而在双流卷积和 3D 卷积出现前，视频动作的分类依靠这个策略实现。

光流处理的单帧视频图像与上述不同，其输入的信息是通过叠加视频帧计算差值生成的灰色图像，通过对光流图像的特征提取可以学习到在静态图像中未学习到的时间变化，学习到动作变化之间的联系，有助于对视频信息更深的特征提取。

2.2. Efficient Netv2 网络

EfficientNet [17] [18] [19] [20] 兼顾了网络深度、网络宽度、图像分辨率，提出混合维度放大法，使用混合系数调整三个维度的放缩倍率，以此为基础调整网络网络结构实现参数少、性能优异。EfficientNetv2 网络结构如表 1 所示，由一个卷积层、三个 Fused-MBConv 模块、三个 MBConv 模块、一个卷积层池化层和全连接层组成，其核心为 MBConv 模块和 Fused-MBConv 模块，其组成如图 2 所示。MBConv 模块的组成为一个 1×1 普通卷积升维，然后是一个深度卷积和 SE 模块，最后是一个 1×1 普通卷积降维。Fused-MBConv 模块的组成不同于 MBConv 模块，将主分支中的普通卷积 1×1 和深度卷积 3×3 替换成一个普通卷积 3×3 ，表 2 所示对比 ResNetRS、EfficientNet 以及 EfficientNetv2 的参数量可知，本文通过 EfficientNetv2 网络进行特征提取可以有效减少计算压力。EfficientNetv2 网络结构可以在保证精度的同时，降低算法的计算量，具有较高的研究价值。

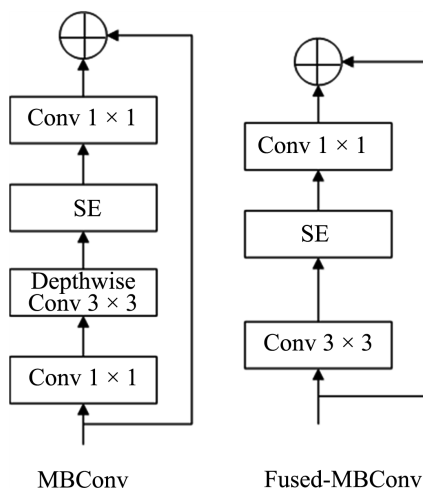


Figure 2. MBConv and Fused-MBConv structure
图 2. MBConv 和 Fused-MBConv 结构

Table 1. EfficientNetv2 network structure
表 1. EfficientNetv2 网络结构

	网络层	步长	堆叠
0	Conv3 × 3	2	1
1	Fused-MBConv1	1	2
2	Fused-MBConv4	2	4
3	Fused-MBConv4	2	4
4	MBConv4	2	6
5	MBConv6	1	9
6	MBConv6	2	15
7	Conv1 × 1 & Pooling & FC	-	1

Table 2. Parameter comparison
表 2. 参数量对比

实验模型	实验参数/M
ResNetRS	164
EfficientNet	43
EfficientNetv2	24

2.3. 空洞卷积

标准的卷积操作使用 3×3 或 5×5 的卷积核,较小尺寸的卷积核在模型训练期间可以减少训练参数,但同时也因为较小的卷积核对应的感受野也较小,对大分辨率图像特征提取十分不利。感受野指的是各层网络特征图的每一个像素点与输入图像之间的对应区域面积,感受野小就代表此次卷积操作提取特征小,感受野大代表提取的特征抽象,而空洞卷积核[21] (Dilated Convolution)结构的提出以扩大卷积核的感受野为目的,同时可以不增加模型的计算参数和计算量,其结构与普通卷积核不同的是在卷积滤波器插入零值,增加了超参数扩张率,而没有加入此超参数的标准的卷积核扩张率为 1。空洞卷积感受野变化如下图 3 所示。

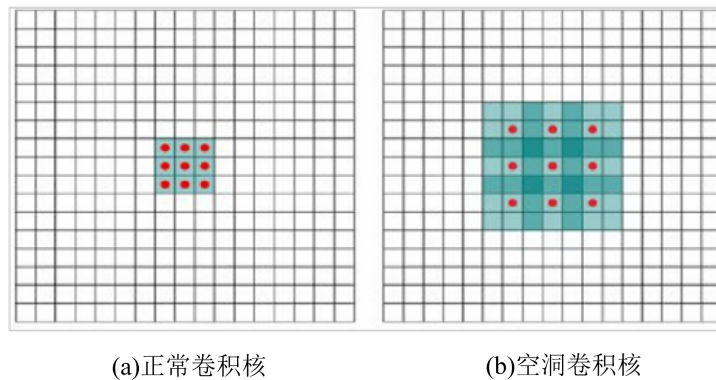


Figure 3. Receptive field changes
图 3. 感受野变化

2.4. 改进损失函数

本文为了提升模型的鲁棒性，本文联合两种损失函数，将 AM-softmax 损失函数[22]和中心损失函数[23]加权使用，其中 softmax 损失函数实现样本类别确定时只需判断 $\|x\| \cos \theta_1 > \|x\| \cos \theta_2$ ，AM-softmax 损失函数在其中增加权值因子 m ，属于同一类别的样本更佳紧密，不用样本之间差异更大，同时为了在训练时减少计算的压力的，引入另一个平衡因子 s ，通过间隔机制，减少计算量，其计算公式如下：

$$L_s = -\frac{1}{N} \sum \log \left(\frac{e^{m(\cos \theta_{yi} - s)}}{e^{m(\cos \theta_{yi} - s)} + \sum_{j \neq yi} e^{m \cos \theta_{yj}}} \right) \tag{1}$$

$$L_{center} = \frac{1}{2} \sum_{i=1}^N \|G^i - G_{yi}\|_2^2 \tag{2}$$

将中心损失函数引入本文，实现差异性小的样本训练优化，不断学习网络训练时各个类别深度特征，将类别之间中心的距离最小化。总损失值由式(1)和(2)联合组成，本文权值因子取值 0.1：

$$L = L_s + \theta L_{center} \tag{3}$$

2.5. 改进的动作识别算法结构

本文动作识别的底层算法是双流卷积神经网络，此算法在 2014 年提出后，研究员对加深网络结构方向进行了多方面研究，本文在改进动作识别算法时使用 EfficientNetv2 网络结构实现特征提取其网络结构如图 4 所示。其网络结构主要由多个 MBConv 和 Fused-MBConv 卷积块组成。

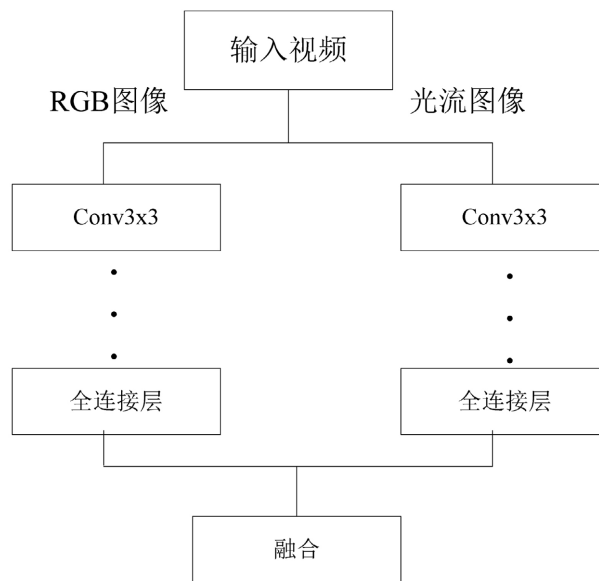


Figure 4. Action recognition model
图 4. 动作识别模型

3. 实验及分析

本文实验环境：Windows 操作系统，32 G 内存，Nvidia GeForce RTX 2080TI 显卡，搭载 pytorch 深度学习框架，在 Anaconda 编程环境下开发。在本节中，所有实验将初始学习率设置为 0.001，为保持损失下降稳定，设置衰减系数 0.1，bashsize 设置为 64。

3.1. 制作数据集

本文使用化学分析自制数据集，在实验室环境中，使用固定相机录制搅拌、拿起实验品、放下实验品、取用液体、混合液体等行为，由不同 20 人录制，每种行为录制 10 次。数据集视频时长控制在 10 秒内，录制视频在光照、背景、遮挡都有所变化。

3.2. 光流提取

在本文的网络结构中，预处理部分即是将视频流分为 RGB 帧图像和光流图像，其中 RGB 图像的获取非常简单，只需要截取视频帧即可；光流图像的提取(首先要满足基本假设，其次是光流法)。

光流图像提取按照不同实现方法可以在梯度、匹配、能量、相位的方法实现，本文采用 Farnback [24] [25]方法，基于梯度实现光流图像提取，如下图所示。

假定图像数列使用 $I(x, y, t)$ ，其中 $X = [x, y]$ ，视频中的前后视频帧取出之后就是一个图像数列，局部光流恒定时，对于任意 $y \in N(x)$ ， $d = \partial X / \partial t$ 不变，即

$$\frac{d}{dt} \nabla I(X, t) = \frac{\partial \nabla I}{\partial X} \frac{\partial X}{\partial t} + \frac{\partial \nabla I}{\partial t} = H(I) \cdot d + (\nabla I)t = 0 \quad (4)$$

将 $f_1(x)$ 多项式表示的空间信息进行位移估计，

$$f_1(x) = x^T A_1 x + b_1^T x + c_1 \quad (5)$$

像素移动后：

$$f_2(x) = f_1(x - d) = x^T A_2 x + b_2^T x + c_2 \quad (6)$$

假设像素值恒定，使 $f_1(x)$ 和 $f_2(x)$ 相等，即

$$A_1 = A_2 \quad (7)$$

$$b_2 = b_1 - 2A_1 d \quad (8)$$

$$c_2 = d^T A_1 d - b_1^T d + c_1 \quad (9)$$

由上述两个式子可以得到：

$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \quad (10)$$

3.3. 视频帧采集

实验室自制的化学分析数据集共有 1000 个化学分析动作视频，平均每个视频拥有 5 个剪辑视频。数据集拥有的视频帧数量巨大，且存在数据冗余、干扰等。对于训练和学习产生较大影响。

为了具体探究视频采样帧的数量对网络性能的影响，使用 10、15、20、25 和 30 共 5 种采样帧数作为动作数据样本制作输入数据，在网络模型训练 100 轮后统计识别准确率，其网络训练结果如下表 3 所示。

Table 3. Sample frame training results
表 3. 标准试验系统结果数据采样帧训练结果

采样帧	模型准确率/%
10	77.8
15	82.2
20	85.23
25	85.52
30	85.6

从上表中可以看到在采样帧为 10 时识别准确率最小，随着视频帧数量提高，准确率也有提升，而从帧数 25 和 30 两个数据看，准确率提升不大，而网络在特征提取时计算压力增大，为了在提升准确率的同时不增加额外的计算压力，本文在后续的视频处理中将 25 帧作为一个视频处理单元，减少计算消耗。

3.4. 网络训练实验

本文在自制化学分析数据集上进行训练，使用 ImageNet 训练初始化特征提取网络参数，进行如下操作：

- 1) 将视频进行处理，将视频分解为 RGB 图像，计算两帧之间的光流值；
- 2) 将图像信息输入第一个特征提取网络得到特征信息，光流信息输入第二个特征提取网络得到特征信息，通过全连接层融合；
- 3) 计算损失函数。

3.4.1. 改进特征识别网络实验

在对 EfficientNetv2 经过空洞卷积改进后，对加深网络深度对动作识别的影响进行实验研究，特征识别网络使用改进的 EfficientNetv2 模型替换，同时使用与双流卷积神经网络相同的平均融合方式。将 Resnet 网络、EfficientNet 网络、Inception 网络以及改进的 EfficientNetv2 网络制作对比实验，这些网络对比原版双流卷积具有更高的鲁棒性以及特征提取能力，其实验准确率对比如下表 4 所示：

Table 4. Accuracy of different feature extraction networks
表 4. 不同特征提取网络的准确率

网络模型	模型准确率/%
双流卷积	88.9
EfficientNet 替换	90.7
改进的 EfficientNetv2 替换	91.3
ResNet 替换	90.9
Inception 替换	90.5

从上述的准确率的对比可以看到，在模型在改进后平均准确率达到到了 91.3%，与未替换的双流卷积神经网络相比准确率得到了提升。对比上述表格中的算法也有准确率优势。

3.4.2. 融合位置对实验的影响

融合位置的选择对模型的性能也有较大影响，同时在融合位置的选择上需要保证在 RGB 图像流和光流保持特征输入尺寸的相同以及融合位置的相同。本文算法网络融合位置在训练的最终准确率数据的实验结果如表 5 所示。实验选取了不同模块进行实验对比，实验选取的融合方式使用卷积融合，其训练网络使用上文改进的迭代训练模型，其在 MBCConv4、MBCConv6、全连接层位置的融合准确率如下表所示。

表 5 不同融合位置的准确率对比。

Table 5. Accuracy comparison of different fusion positions
表 5. 不同融合位置的准确率对比

融合位置	准确率/%
MBCConv4	89.4
MBCConv6	90.2
全连接层	91.3

从上表中可以看出, 本算法在融合位置的选取在准确率方面有了较大影响, 融合位置在全连接层取得了最佳效果, 在两个通道分别进行充分的特征提取后再进行融合操作, 有助于获得更高的分类准确率。

3.4.3. 改进损失函数实验

最后对改进的损失函数进行实验探究, 经过上文的探索, 确定了本文动作识别模型, 确定了特征提取的模型以及融合位置。在前文介绍了本文对损失函数使用的是联合损失函数, 将 AM-softmax 损失函数和中心损失函数加权使用, 通过对原网络改进损失函数训练得到其准确率的对比如下表 6 所示。

Table 6. System resulting data of standard experiment

表 6. 不同损失函数的准确率对比

模型	准确率/%
原模型	88.9
改进损失函数的模型	90.7

从上表可以看到, 原网络模型的准确率为 88.9%, 在使用联合损失函数后, 准确率得到了提升, 达到了 90.7%, 可以推测, 对损失函数的改进可以对于整个模型的性能提升有较大帮助。

3.4.4. 本文的实验模型分析

由于本文使用的数据集为自制的化学分析实验数据集, 其数据量在较深卷积神经网络时会产生过拟合的问题, 为避免出现这类问题, 对实验搭建的卷积神经网络通过在 ImageNet 数据集使用迁移学习的方法, 在大型公共数据集提前预训练, 提前训练网络权重, 使用训练好的模型参数作为本文实验的初始参数。网络在 ImageNet 数据集上提前与训练后得到初始参数, 对于后续在自制数据集训练时可以采用更小的学习率, 模型迭代 200 轮的结果如下图所示, 横坐标表示迭代次数, 纵坐标表示准确率。从图中可以看到模型初始化准确率较低, 在经过 80 轮后准确率的值趋于平缓, 测试集可以达到 92.4% (图 5)。

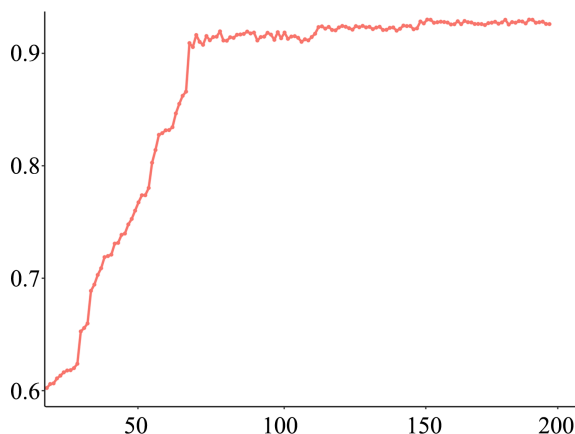


Figure 5. Improved action recognition model results

图 5. 改进动作识别模型结果图

在对所有影响网络准确率的因素进行影响分析后, 为验证在动作算法改进的效果, 将本文算法与先进的算法进行对比。本文的主要改进策略是在双流卷积神经网络架构下完成的, 改进的主要是特征提取的网络和损失函数, 为了保证实验的可对比性, 使用的设备、操作系统、数据集及预处理方式都保持一致, 取准确率的最大值作为参照。表 7 列出了与现有的几种动作识别相关模型的对比结果。

Table 7. Accuracy of different networks**表 7.** 不同网络的准确率

网络模型	模型准确率/%
Two-Stream	88.9
Two-Stream + LSTM	90.6
C3D	87.6
CNN KNN + SVM	94
Resnet-Two-Stream	90.9
本文	92.4

本文对比了 3 种双流卷积网络以及 C3D 算法, 从上表中可以看到, 双流卷积神经网络的准确率只有 88.9%, 本文在对其改进后准确率可以达到 92.4%, 在双流卷积神经网络基础上对特征提取网络及损失函数方向的改进对于模型的整体识别准确率比传统识别算法有一定提升, 证明本文在加深网络和改进损失函数的方法具有可行性。

4. 结束语

本文提出的应用于化学分析中的行为识别算法, 在双流卷积神经网络的基础上加深了网络结构, 引入空洞卷积改进特征识别, 并分析了融合位置对准确率的影响, 同时改进了损失函数, 将提取的 RGB 图像特征和光流图像特征进行融合分类, 识别效果明显。

网络中特征提取部分利用 EfficientNetv2 图像处理算法的处理速度快、参数少和准确率高的优势, 实现算法轻量级的同时将准确率降低的影响尽可能的缩小。网络与 CNN KNN + SVM 网络模型相比稍有逊色, 其较本文改进之前的双流网络准确率有所提升, 这个结果是可以接受的, 该研究方向对工业自动化、自动化具有重要意义。

基金项目

科技部, 国家重点研发计划子课题(2017YFB1400903)。

参考文献

- [1] 王春安, 危紫翼, 杨茜, 何宏斌, 徐艺珺, 陈艳萍, 曾坚朋. 国外先进实验室人员配置与经费情况对我国实验室建设运行的启示[J]. 实验技术与管理, 2021, 38(12): 243-248+282.
- [2] 冯云华. 智能化学化工实验室的研究与设计[J]. 石化技术, 2019, 26(3): 310+309.
- [3] 何爱欣. 桌面机械臂智能交互技术研究及应用[D]: [硕士学位论文]. 青岛: 青岛科技大学, 2019
- [4] 王涛. 基于 ROS 的多桌面型机械臂协作机制研究与应用[D]: [硕士学位论文]. 青岛: 青岛科技大学, 2021.
- [5] 张兴强. 桌面型机械臂在化学实验中的应用研究[D]: [硕士学位论文]. 青岛: 青岛科技大学, 2021.
- [6] 梁绪, 李文新, 张航宇. 人体行为识别方法研究综述[J]. 计算机应用研究, 2022, 39(3): 651-660.
- [7] 郝树新, 林锦州, 刘芳. 机器视觉在汽车质量检验领域的应用研究[J]. 时代汽车, 2021(21): 16-17.
- [8] 袁依文, 雷斌. 无人机视觉技术在道路裂缝分类中的应用[J/OL]. 机械设计与制造: 1-6. <https://doi.org/10.19356/j.cnki.1001-3997.20211106.011>, 2021-11-08.
- [9] 李柯泉, 陈燕, 刘佳晨, 牟向伟. 基于深度学习的目标检测算法综述[J/OL]. 计算机工程: 1-17. <https://doi.org/10.19678/j.issn.1000-3428.0062725>, 2021-11-17.
- [10] 施徐伟, 丁书艺, 储好, 郑思婕, 韦朴. 基于机器视觉的垃圾分拣系统[J]. 电子测试, 2021(21): 35-36+42.

-
- [11] 凌远坤, 仲志丹, 李栋, 金容鑫, 姚晨辉, 王继茂. 基于机器视觉图像识别及处理的方法[J]. 科技风, 2019(15): 95.
- [12] Zhao, R., Ali, H. and Patrick, V. (2017) Two-Stream RNN/CNN for Action Recognition in 3D Videos. 2017 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vancouver, 24-28 September 2017, 4260-4267. <https://doi.org/10.1109/IROS.2017.8206288>
- [13] 陈颖, 来兴雪, 周志全, 秦晓宏, 池亚平. 基于 3D 双流卷积神经网络和 GRU 网络的人体行为识别[J]. 计算机应用与软件, 2020, 37(5): 164-168+218.
- [14] 谢佳龙, 张波涛, 吕强. 一种基于双流融合 3D 卷积神经网络的动态头势识别方法[J]. 电子学报, 2021, 49(7): 1363-1369.
- [15] 曾明如, 郑子胜, 罗顺. 结合 LSTM 的双流卷积人体行为识别[J]. 现代电子技术, 2019, 42(19): 37-40.
- [16] Feichtenhofer, C., Pinz, A. and Zisserman, A. (2016) Convolutional Two-Stream Network Fusion for Video Action Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 1933-1941. <https://doi.org/10.1109/CVPR.2016.213>
- [17] Tan, M. and Le, Q.V. (2021) EfficientNetV2: Smaller Models and Faster Training. 2021 *International Conference on Machine Learning*, Online, 18-24 July 2021. <https://arxiv.org/abs/2104.00298>
- [18] Pakzad, F., Portmann, M., Tan, W.L. and Indulska, J. (2014) Efficient Topology Discovery in Software Defined Networks. 2014 *8th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, 15-17 December 2014, 1-8. <https://doi.org/10.1109/ICSPCS.2014.7021050>
- [19] Hong, L., Thanh, N.C. and Long, T.Q. (2020) Polyp Segmentation in Colonoscopy Images Using Ensembles of U-Nets with EfficientNet and Asymmetric Similarity Loss Function. 2020 *RIVF International Conference on Computing and Communication Technologies (RIVF)*, Chi Minh City, 14-15 October 2020, 1-6. <https://doi.org/10.1109/RIVF48685.2020.9140793>
- [20] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 2019 *International Conference on Machine Learning*, Long Beach, 10-15 June 2019, 6105-6114.
- [21] Yu, F. and Koltun, V. (2015) Multi-Scale Context Aggregation by Dilate Convolutions. arXivpreprint ar Xiv:1511.07122.
- [22] Wang, F., Cheng, J., Liu, W.Y. and Liu, H. (2018) Additive Margin Softmax for Face Verification. *IEEE Signal Processing Letters*, **25**, 926-930. <https://doi.org/10.1109/LSP.2018.2822810>
- [23] Wen, Y.D., Zhang, K.P., Li, Z.F. and Qiao, Y. (2016) A Discriminative Feature Learning Approach for Deep Face Recognition. *Proceedings of the European Conference on Computer Vision*, Amsterdam, 8-14 September 2018, 499-515. https://doi.org/10.1007/978-3-319-46478-7_31
- [24] Farnebeck, G. (2003) Two-Frame Motion Estimation Based on Polynomial Expansion. 13th *Scandinavian Conference on Image Analysis (SCIA 2003)*, Halmstad, 29 June-2 July 2003, 363-370. https://doi.org/10.1007/3-540-45103-X_50
- [25] Farnebeck, G. (2002) Polynomial Expansion for Orientation and Motion Estimation. Linköping University, Linköping.