

基于人脸状态识别的算法研究

罗 龙

江西理工大学 江西赣州 341000

摘 要: 采用计算机视觉实现对采集图像中人脸所处状态的识别技术, 通过对人脸状态的识别能够充分地反映出一个人的心理与生理方面的状态信息, 这种利用人脸识别方式来判断一个人的行为, 如今已经成为当前人脸状态检测的方向标。现阶段各个领域中的应用的主流人脸识别技术, 主要以人的面部问题, 器官特征以及骨架特征等静态情况来判断人脸所处的状态, 但无法根据脸部全局肌肉运动情况来判断人脸当前所处的状态。因此, 针对该问题, 本文提出了一种基于深度学习的时间系列人脸状态识别算法, 以此提高了对不同的人脸面部特征、复杂背景的鲁棒性。首先, 利用共享权值直连了双通道级联网络的设计, 并实现了对特征图的提取。然后, 采用 Liteflownet 从特征图当中完成了光流图的提取, 以及实现了 L-R 模型的设计, 解决了头部偏转矢量导致识别结果存在误差的问题; 其次, 完成了分类器的设计。最后经测试结果证明, 本文提出的人脸状态识别算法, 不仅可以促进人脸状态的识别率得到提高, 还拥有较高的鲁棒性。

关键词: 时间序列识别算法; 人脸状态识别; 面部表情; 图像处理; 深度学习

前言: 利用生物视觉系统实现对面脸面部状态的识别, 主要依据人脸面部纹理特征以及面部肌肉的变化等两个方面, 来反映一个人的状态(情绪、疲劳程度以及认知等)。在人脸状态识别过程中, 传统识别算法采用手工标注特征点方式实现特征的提取, 接着对提取的表情特征实施分类处理, 这样一来就导致算法非常依赖于人工的干预, 出现状态识别率低、鲁棒性差等问题^[1]。近几年, 随着深度学习(Deep Learning)的发展、完善, 使该算法在人脸状态识别方面得到了广泛的应用, 如疲劳监测、人脸状态替换、智能医疗等。但是在人脸状态识别的研究方面, 深度学习的应用一直停留在静态帧检测方面, 简单来说就是通过对提取的人脸面部纹理进行状态识别。因此, 本文提出了一种基于深度学习的时间序列检测人脸面部全局肌肉状态变化的深度网络 Liteflownet-R, 用于识别人脸状态, 并设计了实时状态检测算法, 以此来提高对人脸状态识别的效率, 同时该算法的应用进一步降低人脸状态的检测成本。

1、人脸状态识别概述

随着计算机技术及其硬件性能的提升, 推动了基于三维图像的人脸状态识别研究得到了进一步的发展。如结合稀疏编码(SRC)提出的人脸状态识别算法, 降低了识别过程中遮挡、噪音对人脸状态识别精度的影响; 同时结合 Gabor 特征进行了识别模型设计, 以此促进了人脸状态识别算法的鲁棒性得到进一步提高。近年来, 随着学者对人脸状态识别领域的不断研究, 逐渐形成静态图片流和视频流等两个流派。从静态图片的图像特征提取和分类器方面来看, 主要分为深度学习、支持向量机以及自适应增强等几种方法; 从视

频流方面来看, 其分类器构造、图像特征提取主要分为差分图像法、光流法、动态特征定位以及特征点定位等 4 种方法。为进一步提高人脸面部识别的效率和鲁棒性, 本文选择采用深度学习, 基于对图像深层特征的提取, 提取出人脸面部纹理以及器官的特征点, 然后借助大量的样本训练分类器, 实现了对人脸面部状态的识别^[2]。

然后, 结合 Liteflownet 提出了一种可以充分识别人脸面部状态的深度网络(L-R), 接着又对深度网络参数及其结构设计等进行了相应的测试以及调整, 完成了基于人脸面部肌肉运动状态特征提取的模块设计。最后, 利用卷积神经网络 CNN 和循环神经网络 RNN, 实现了时间序列状态分类器设计, 为最终分类结果的输出提供相应的工具。

2、深度学习及图像预处理方法选择

2.1 深度学习算法选择

深度学习主要通过对样本的内在规律进行挖掘和表现, 实现了学习过程中所获得信息的深层次解释, 如图像、文字以及声音等信号; 而训练完成的深度网络可以实现对文字数据、图像数据以及声音数据的识别。随着深度网络结构的不断发展, 从 BP 神经网络、卷积神经网络以及 RNN 等网络结构中延伸出来了 SSD、YOLO 以及 R-CNN 网络结构。因此, 在对人脸状态识别前, 本文先是对上述三种神经网络结构进行分析, 选择出适合本文所提出算法的图像特征提取器、分类器以及时间序列特征提取器的网络结构^[3]。如能够用于对序列数据处理的 RNN 基础网络结构, 该神经网络在层之间的神经元之间同样可以实现全连接。其中, 从其架构上来看, 网络输入端

代表了序列信息；单神经元实现了对数据信息的提取，多神经元实现了对数据信息的继承；并且，序列信息特征的输出非常适合应用在对时间序列数据特征提取和分类处理等方面上，也就是说本文的时间序列特征算法设计可以选择用 RNN 网络为基础的架构。

2.2 图像预处理方法选择

摄影器材设备以及摄影的背景环境等因素会对图像信号的获取造成一定程度的影响，摄影时摄影场景中存在的噪声问题、光照强弱问题等会导致人脸状态识别以及检测的最终结果受到一定程度的影响。图像预处理简单来说就是借助先进的操作技术来降低拍摄时存在的差异问题，这样一来就能够得到真实的图像信息。不过，由于针对图像预处理的技术类型繁多，所以本文在对人脸状态识别算法设计前，对于图像预处理的选择可以分为三个方面：图像模糊、图像均衡化和图像插值等。其中，从图像模糊方面分析来看，该部分的作用在图像预处理时，在充分保障原图像不失真，且还能够使得噪声问题对图像信号造成的影响减少。从图像均衡化方面设计分析来看，它在图像预处理当中主要起到了平衡图像像素信息间的差异，并利用均衡化的方法，使得光照强弱不均匀导致的图像受到的影响降低。从图像插值设计分析来看，图像插值在图像预处理时，当图像发生失真，图像插值就可以改变原图像的尺寸，不过由于采集的测试样本的尺寸大小不同，所以测试过程当中，还需要对所采集的测试样本尺寸实施处理，这样一来才能够保证每个图像的尺寸大小一致^[4]。

3、基于人脸状态识别的算法研究设计

3.1 人脸状态识别算法

现阶段，静态帧状态识别算法是人脸状态识别研究领域当中的主流算法。接着，通过对基于人脸面部器官特征的状态识别算法、二级投票算法及加权局部旋度算法等三种不同类型的识别算法研究分析发现，这三种识别算法的分类器大致一样，其差别主要在于特征提取方面。首先，从面部器官特征算法方面分析来看，在对人脸状态识别过程前需要先检测出人脸的准确位置，然后结合面部的集合特征分布规律，对人脸中的器官（眼镜、嘴巴）等位置进行检测，在利用直方图图像预处理方法从图像中将眼镜、嘴巴等器官的几何特征提取出来，再转化成直方图分布概率，将其和配置完成的数据库进行对比分析，并输出所识别的信息。其次，从由 CNN 和 STN 构成的 ROI 二级投票算法方面进行分析，将投票机制添加到判别系统当中，这样一来就能够降低识别时存在的误判可能。简单来说，该算法的原理就是将人脸识别图像划分成多个不同的区域，然

后通过加权之后，将其添加到 STN 网络分类当中进行识别，并且，该网络对人脸头部偏转识别具备较高的鲁棒性。最后，从加权局部旋度算法方面，此算法在应用时，借助 LCPs 实现对面部的旋转向量特征进行提取^[5]。

3.2 深度网络 LiteflowNet-R 算法设计

本文提出了一种基于时间序列信息的人脸状态识别算法，该算法主要基于人脸面部肌肉移动变化的时间序列构建二次，可以快速解决特征丢失以及环境扰动等缺点。

3.2.1 级联网络特征设计

由于深度网络算法在进行图像识别时，输入的图像为双通道人脸图像，所以本文基于共享权值提出并设计了一个双通道级联网络。首先，该网络的双通道输入设计，分别对应了时间序列的两幅连续的图像，接着通过对图像中人脸特征网络提取，并输出不同类型空间维度的特征图，而特征提取网络设计主要由 CNN 网络组成。其次，级联网络特征的设计具有 A1 和 A2 两个输入端，且共用一个网络权值，所以该网络可以进一步降低网络参数的数据，使得训练过程得到加快，从而在一定程度上解决了拟合和梯度爆炸的问题。最后，由两个输入端所提取、标识图像的面部单元特征之后，可以将放入相应的模块中进行匹配，并输出人脸面部肌肉运动变化的光流图^[6]。

3.2.2 深度网络模块设计

输入图像中的每一个图像像素点的速度矢量，只有利用运动单元提取模块才能够将速度矢量提取出来，但无法实现对图像中人脸面部头部偏转的速度矢量进行识别。如当对两帧图像上的头部偏转特征提取时，最后输出的人脸面部肌肉运动变化光流图，则是由人脸面部运动速度以及头部偏转速度两者叠加的。所以，针对上述问题，本文结合深度学习，提出了一种 L-R 头部偏转特征维稳模块，以此来解决头部偏转速度矢量求解的问题^[7]。

4、测试结果与分析

为了进一步验证本文算法对人脸状态识别的可靠性，本文选择对采用弱光照的场景下，人脸面部均匀光照的扰动影响进行测试，并将其和上述提到的三种识别算法之间进行交叉对比分析，同时对这三种识别算法模型分别配置了 96、85、530 组时间序列测试图像，以此来验证本文算法经图像预处理之后，对人脸状态识别时的可靠性进行验证。

4.1 运动特征提取模型效果的评估分析

针对弱光场景下人脸面部均匀光照扰动的人脸面部肌肉运动特征提取模型分析，本文选择 20 幅时间序列图像为此次测试样本，其

中该图像主要分别包含了 20 张人脸序列图像以及面部肌肉运动特征图，具体选取的检测图像见图 1 所示。

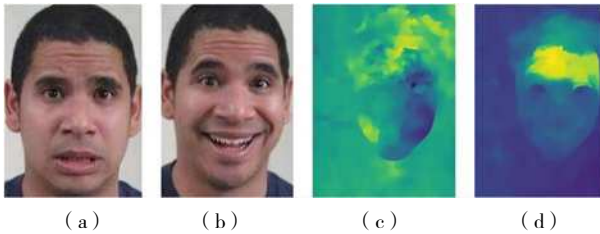


图 1 人脸面部肌肉运动特征提取效果的示意图

从图 1 上来看，在弱光场景条件下，该模型可以实现对人脸的眼部、脸颊肌肉运动以及嘴巴的运动提取。并且，在人脸状态识别算法中，其特征表现出了较强的鲁棒性。

4.2 L-R 分类器效果评估分析

为了验证分类器的可行性，本文选择 100 幅修正过的头部偏转的人脸面部肌肉运动特征图序列为此次分类器测试的样本，以及 100 条对应的状态标签。并将人脸状态分为 5 种类型：兴奋、悲伤、疲劳、平淡以及愤怒，且每种状态各 20 幅时间序列图像，然后通过交叉对比对本文提出的分类器效果进行评估，具体结果见表 1、表 2 所示。

表 1 交叉评估 mAP 结果

	兴奋	悲伤	疲劳	平淡	愤怒	均值
1-mAP	0.87	0.83	0.80	0.84	0.80	0.822
2-mAP	0.31	0.39	0.35	0.41	0.45	0.382
3-mAP	0.62	0.76	0.62	0.76	0.70	0.692
4-mAP	0.82	0.72	0.82	0.74	0.80	0.78

从上表可知，抵挡在弱光场景下，对人脸脸部均匀光照的测试，L-R 分类器的测试样本 mAP 为 82%。这就说明本文算法比上述三种算法更加优秀，且在多维度的扰动干扰之下，此模型能够精准地实现人脸面部状态的识别。

结语：

综上所述，针对人脸状态的识别，本文提出了一种基于深度学习的时间序列特征状态识别算法。首先，该算法的设计主要分为基于共享权值的双通道级联网络，该网络的设计不仅提高了收敛速度，还降低了权值数量；其次，设计了运动特征提取模块，提高了对人脸面部运动特征的提取速度；设计了深度网络 L-R，用于降低人脸头部转动所造成的扰动问题；最后，人脸识别器的设计，输出了最终人脸状态识别的结果。同时，还对 L-R 模型的训练及测试过程进行了阐述，并对该识别算法的可靠性进行了验证。

参考文献：

- [1]徐武,陈盈君,汤弘毅,等. 局部特征映射与融合网络的人脸识别优化算法[J]. 河南科技大学学报(自然科学版),2023,44(2):59-64,72.
- [2]曾曦,辛月兰,谢琪琦. 基于性别约束的多分支网络人脸表情识别[J]. 计算机工程与应用,2023,59(9):245-254.
- [3]焦阳阳,黄润才. 基于改进纹理特征与迁移学习的人脸表情识别[J]. 智能计算机与应用,2023,13(5):181-186.
- [4]张妍. 基于动态人脸表情相似度的情绪状态识别研究[D]. 上海:东华大学,2022.
- [5]胡佳玲,施一萍,谢思雅,等. 基于轻量级卷积神经网络人脸识别算法的研究与应用[J]. 传感器与微系统,2022,41(1):153-156.
- [6]郑焱. 深度学习在人脸表情识别算法中的研究与应用[D]. 江西:江西理工大学,2022.
- [7]吕鹏,单剑锋. 基于多特征融合的人脸表情识别算法[J]. 计算机技术与发展,2022,32(10):151-155,181.

作者简介：姓名：罗龙，性别：男，籍贯：四川射洪，民族：汉，出生年月：1987.01.29，学位：学士，职称：实验师，研究方向：嵌入系统。

本论文受江西省教育厅科技项目资助，项目编号：GJJ210838