

## 《文献检索与利用》课程综述报告

姓名	分数	备注
赵小钦		
秦思远		
张天宇		
刘天保		
彭留洋		
张弛		

# 人脸检测算法研究综述

201631107110 赵小钦

201631103228 秦思远

201631083416 张天宇

201631101227 刘天保

201631062622 彭留洋

201631107108 张弛

任课教师：贺映兰

评分项目	满分	得分
文本格式规范	15%	
应用检索原理检索方法情况	20%	
检索文献的范围和类型	20%	
参考文献的引用和著录规范	10%	
结论的准确性	25%	
各小组成员在课题完成中担负的责任大小	5%	

附件及其他	5%	
-------	----	--

阅卷老师签字：

完成日期：二〇一九 年 月 日

《文献检索与利用》综述选题表

拟选专题综述题目	人脸检测算法研究			
题目属性	(√) 教师指定      ( ) 自选			
课题组成员及分工情况(课题组成员 3-5 人)	姓名	身份	学号	课题内分工
	赵小钦	组长	201631107110	负责人, 选题分析, 搜集文献, 正文撰写
	秦思远	组员	201631103228	参考文献分析, 大纲编写、正文撰写、参考文献
	刘天保	组员	201631083416	资料收集、资料整理、摘要撰写, PPT 制作
	张天宇	组员	201631101227	资料收集、资料整理、摘要撰写, PPT 制作
	张弛	组员	201631107108	资料收集、资料整理、摘要撰写, PPT 制作
	彭留洋	组员	201631062622	资料收集、资料整理、摘要撰写, PPT 制作, PPT 汇报
联系人电话	17723487859			

**注：**课题内分工，指课题组成员拟承担的主要工作，如：分析课题要求、检索文献、获取原文、资料综合与分析、执笔写作、课题报告最后陈述人、课题总负责人，等。

# 1 目 录

## 目录

1	目 录.....	4
2	课题分析.....	5
3	检索策略.....	5
3.1	选择检索工具.....	5
3.2	选择检索词.....	5
4	检索步骤、检索结果及其评价.....	6
4.1	中国知网(CNKI).....	6
4.2	国家知识产权局.....	7
4.3	IEEE 数据库.....	8
4.4	WOS 数据库.....	9
5	文献综述.....	11
5.1	基于深度学习的人脸检测.....	11
5.1.1	图像预处理.....	12
5.1.2	人脸对齐.....	12
5.1.3	数据增强.....	12
5.1.4	面部检测.....	12
5.1.5	人脸归一化.....	12
5.1.6	面部特征提取.....	12
5.2	面部图像分类.....	14
5.2.1	KNN 算法进行面部图像分类.....	14
5.2.2	支持向量机算法进行面部图像分类.....	14
5.3	结论.....	15
6	总结与展望.....	15
7	参考文献.....	16
8	评价.....	17
9	附件.....	18

## 2 课题分析

随着当下科学技术的发展,一项新兴技术逐渐进入人们的视野,也慢慢渗透进人们生活的方方面面:生物特征识别。生物特征识别技术所研究的生物特征包括脸、指纹、手掌纹、虹膜、视网膜、声音(语音)、体形、个人习惯(例如敲击键盘的力度和频率、签字)等,相应的识别技术就有人脸识别、指纹识别、掌纹识别、虹膜识别、视网膜识别、语音识别(用语音识别可以进行身份识别,也可以进行语音内容的识别,只有前者属于生物特征识别技术)、体形识别、键盘敲击识别、签字识别等。

本次课题研究的人脸识别技术是基于人的脸部特征,对输入的人脸图象或者视频流,首先判断其是否存在人脸,如果存在人脸,则进一步的给出每个脸的位置、大小和各个主要面部器官的位置信息。并依据这些信息,进一步提取每个人脸中所蕴涵的身份特征,并将其与已知的人脸进行对比,从而识别每个人脸的身份。

## 3 检索策略

### 3.1 选择检索工具

检索工具名称	文献深度	语种	检索年代	文献类型
中国知网(CNKI)	全文	中文	2015-2018	期刊; 学位论文
国家知识产权局	全文	中文	2015-2018	专利
IEEE 数据库	全文	英文	2015-2018	期刊; 文献
WOS 数据库	文摘	英文	2015-2018	期刊

### 3.2 选择检索词

从课题字面选	从课题内涵选(同义词、近义词、上下位词)
人脸	面部、脸部、面容、特征 face
检测	识别技术; 浅层学习; 深度学习; 卷积神经网络; 深度置信网络; recognition technique; identification techn; CNN; neural network

## 4 检索步骤、检索结果及其评价

### 4.1 中国知网(CNKI)

检索时间范围: 2015-01-01——2018-09-28

序号	检索策略	检索结果	检索评价
1	题名=人脸*检测	2,406 条	相关性好
2	题名=人脸*识别技术	889 条	相关性好
3	题名=(面部+脸部)*(特征) AND 关键字=识别	3 条	查准率高
4	题名=面部 AND 关键字=(识别技术+深度学习+浅层学习+卷积神经网络+深度置信网络)	39 条	查准率高
5	题名=(面部+脸部)*(特征) AND 关键字=(识别技术+深度学习+浅层学习+卷积神经网络+深度置信网络)	6 条	查准率高
6	题名=脸部 AND 关键字=(识别技术+深度学习+浅层学习+卷积神经网络+深度置信网络)	3 条	查准率低

从序号 1 中选出 2 篇

(1) SrcDatabase-来源库: CJFDPREP

Title-题名: 基于多任务级联卷积网络模型的人脸检测和识别

Author-作者: 刘其嘉 郭一娜 任晓文 李健宇;

Organ-单位: 太原科技大学电子与信息工程学院;

Source-文献来源: 太原科技大学学报

Summary-摘要: 深度学习的集成特征提取这一优点使得它广泛应用于人脸检测和识别。提出了一种多任务级联卷积网络模型(Multitask Cascaded Convolution Network,MTCNN)。基于 Tensor Flow 平台,基于改进的任务级联卷积网络模型检测到人脸,并且用 Face Net 算法对人脸进行特征提取,用 KNN 算法对人脸进行识别。实验结果表明,对不同光照下多人图像和遮挡图像的人脸进行检测和识别,具有良好的鲁棒性。

(2) SrcDatabase-来源库: CAPJ2018

Title-题名: 两层级联卷积神经网络的人脸检测

Author-作者: 张海涛 李美霖 董帅含;

Organ-单位: 辽宁工程技术大学软件学院;

Source-文献来源: 中国图象图形学报

Summary-摘要: 目的传统人脸检测方法因人脸多姿态变化和人脸面部特征不完整等问题,导致检测效果不佳。为解决上述问题,提出一种两层级联卷积神经网络(TC\_CNN)人脸检测方法。方法首先,构建两层卷积神经网络模型,利用前端卷积神经网络模型对人脸图像进行特征粗略提取,再利用最大值池化方法对粗提取得到的人脸特征进行降维操作,输出多个疑似人脸窗口;其次,将前端粗提取得到的人脸窗口作为后端卷积神经网络模型的输入进行特征精细提取,并通过池化操作得到新的特征图;最后,通过全连接层判别输出最佳检测窗口,完成人脸检测全过程。结果实验选取 Fddb 人脸检测数据集中包含人脸多姿态变化以及人脸面部特征信息不完整等情况的图像进行测试,TC\_CNN 方法人脸检测率达到 96.39%,误检率低至 3.78%,相比当前流行方法在保证算法效率的同时检测率均有提高。结论两层级联卷积神经网络人脸检测方法能够在人脸多姿态变化和面部特征信息不完整等情况下实现精准检测,保证较高的检测率,有效降低误检率,方法具有较好的鲁棒性和泛化能力。

从序号 3 中选出 1 篇

(3) SrcDatabase-来源库: CAPI2018

Title-题名: 基于关键点 LBP 特征提取的面部 AU 组合识别

Author-作者: 任艳;

Organ-单位: 东南大学学习科学研究中心;

Source-文献来源: 信息化研究

Summary-摘要: 提出以面部关键点为中心选取子区域,对各个子区域提取(Local binary pattern,LBP)特征,用最近邻分类器进行面部运动单元(Action unit,AU)组合识别。AU 的运动与面部特定肌肉的变化有关,基于关键点特征的提取方法可最大限度提取有用的信息,减少冗余信息,与提取整幅面部表情图像的 LBP 特征方法相比较,其识别结果有明显的优势。

从序号 4 中选出 1 篇

(4) SrcDatabase-来源库: CJFD2017

Title-题名: 基于深度学习的面部表情识别研究

Author-作者: 陆嘉慧 张树美 赵俊莉;

Organ-单位: 青岛大学数据科学与软件工程学院;

Source-文献来源: 计算机应用研究

Summary-摘要: 近几年来,深度学习在语音识别、图像理解等许多应用领域取得了突破性成果。针对基于深度学习的静态人脸图像表情识别方法进行研究,首先介绍了深度学习的原理,并归纳了目前公开且常用的面部表情数据集;然后介绍了基于深度学习的表情识别的三个步骤,归纳了图像预处理和表情分类的主要方法,重点总结了目前性能较好用来提取特征的深度学习框架以及这些方法的基本原理和优劣势比较;最后指出了目前面部表情识别存在的问题和未来可能的发展趋势。

从序号 5 中选出 1 篇

(5) SrcDatabase-来源库: CJFD2017

Title-题名: 基于单个卷积神经网络的面部多特征点定位

Author-作者: 朱虹 李千目 李德强;

Organ-单位: 南京理工大学计算机科学与工程学院;

Source-文献来源: 计算机科学

Summary-摘要: 深度学习在面部特征点定位领域取得了比较显著的效果。然而,由于姿态、光照、表情和遮挡等因素引起的面部图像的复杂多样性,数目较多的面部特征点定位仍然是一个具有挑战性的问题。现有的用于面部特征点定位的深度学习方法是基于级联网络或基于任务约束的深度卷积网络,其不仅复杂,且训练非常困难。为了解决这些问题,提出了一种新的基于单个卷积神经网络的面部多特征点定位方法。与级联网络不同,该网络包含了 3 组堆叠层,每组由两个卷积层和最大池化层组成。这种网络结构可以提取更多的全局高级特征,能更精确地表达面部特征点。大量的实验表明,所提方法在姿态、光照、表情和遮挡等变化复杂的条件下优于现有的方法。

## 4.2 国家知识产权局

申请日>=20150101

序号	检索策略	检索结果	检索评价
1	发明名称=(人脸 识别)	109091 条	相关性差
2	发明名称=(人脸 面部 面容 脸部) AND 摘要	8359 条	相关性中

	= (识别)		
3	发明名称=(人脸 面部 面容 脸部) AND 摘要 =(识别技术 浅层学习 深度学习 卷积神经网络 深度 置信网络)	1514 条	相关性好

从序号 3 中选出 3 篇

(6) 专利名称：一种人脸识别方法及识别装置

申请号：CN201811504957

申请日：2018.12.10

公开（公告）号：CN109508694A

公开（公告）日：2019.03.22

IPC 分类号：G06K9/00;

申请（专利权）人：上海众源网络有限公司；

发明人：胡居成；

(7) 专利名称：一种基于人脸识别的考勤设备及方法

申请号：CN201811276685

申请日：2018.10.30

公开（公告）号：CN109191611A

公开（公告）日：2019.01.11

IPC 分类号：G07C1/10; G06K9/00; G06K9/46;

申请（专利权）人：惠州学院；

发明人：卢铿栋;蔡沛潜;陈晓星;周玲;罗中良;

(8) 专利名称：融合多任务和多尺度卷积神经网络的人脸分析方法及系统

申请号：CN201811260674

申请日：2018.10.26

公开（公告）号：CN109508654A

公开（公告）日：2019.03.22

IPC 分类号：G06K9/00; G06K9/62

申请（专利权）人：中国地质大学（武汉）；

发明人：刘袁缘;周顺平;张香兰;方芳;郭明强;姚尧;彭济耀;

### 4.3 IEEE 数据库

Year: 2015 - 2018

序号	检索策略	检索结果	检索评价
1	Face Recognition	30,323 条	相关性较好

从序号 1 中选出 2 篇

(9) Title: Face Recognition Method Based on Adaptively Weighted Block-Two Dimensional Principal Component Analysis

Author: Zhao Lihong;Guo Zikui

From: 2011 Third International Conference on Computational Intelligence, Communication Systems and Networks

Abstract: Face recognition is one of the most important part of biometrical recognition. 2-Dimensional Principal Component Analysis (2DPCA) is a classic method in face recognition, which is proposed to reduce the computational cost of the



standard Principal Component Analysis (PCA) algorithm, but the performance of 2-Dimensional Principal Component Analysis in reducing computational complexity and recognition rate is not satisfying. This paper mainly focuses on the feature extraction method of adaptively weighted Block 2-Dimensional Principal Component Analysis. The block methods divide a large picture into several smaller sub-blocks to get the local discrimination information and reduce the computational complexity. Then, a weighted Euclidean distance classifying algorithm is proposed to extract features of face images, and the Euclidean distance classifier is used for classifying. The experiments show that the Adaptively Weighted Block 2-Dimensional Principal Component Analysis method has better performance than standard 2-Dimensional Principal Component Analysis.

DOI: 10.1109/CICSyN.2011.18

(10) Title: A DSP-based approach for the implementation of face recognition algorithms

Author: A.U.Batur; B.E.Flinchbaugh ;

From: 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).

Abstract: Face recognition is an important part of today's emerging biometrics and video surveillance markets. Recent years have witnessed an exploding interest in the development of face recognition algorithms and products. Currently, face recognition systems are usually implemented on general purpose processors. As face recognition algorithms move from research labs to the real world, power consumption and cost become critical issues. This motivates searching for implementations using a digital signal processor (DSP). Our goal is to explore the feasibility of implementing DSP-based face recognition systems. To achieve this goal, we implement a fully automatic face recognition system on Texas Instruments' TMS320C6416 DSP, profile performance, and analyze opportunities for optimization. The results of our experiments demonstrate that a generic C implementation with a modest C level optimization effort results in a face recognition software prototype that has low CPU and memory requirements. Hence, it appears that well-optimized face recognition implementations on DSPs can be an effective choice for embedded face recognition products.

DOI: 10.1109/ICASSP.2003.1202342

#### 4.4 WOS 数据库

时间跨度: 2015-2018

序号	检索策略	结果	检索评价
1	标题: (face *) AND 标题: (recognition technique or identification techn* or CNN or neural network*)	789 条	相关性较好
2	标题: (face*) AND 主题: (recognition technique or CNN)	11625 条	相关性差

从序号 1 中选出 2 篇:

(11) 作者: Dang, LM (Dang, L. Minh); Hassan, SI (Hassan, Syed

Ibrahim); Im, S (Im, Suhyeon); Lee, J (Lee, Jaecheol) ; Lee, S (Lee, Sujin); Moon, H(Moon, Hyeonjoon)

标题: Deep Learning Based Computer Generated Face Identification Using Convolutional Neural Network

来源出版物: APPLIED SCIENCES-BASEL

卷: 8

期: 12

DOI: 10.3390/app8122610

出版年: DEC 2018

(12) 作者: Ding, CX (Ding, Changxing); Tao, DC (Tao, Dacheng)

标题: Trunk-Branch Ensemble Convolutional Neural Networks for Video-Based Face Recognition

来源出版物: IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

卷: 40

期: 4

页: 1002-1014

DOI: 10.1109/TPAMI.2017.2700390

出版年: APR 2018

## 5 文献综述

### 人脸检测算法研究综述

赵小钦 秦思远 刘天保 张天宇 张弛 彭留洋  
软件工程专业 2016 级

**摘要:** 随着安全入口控制和金融贸易方面应用需要的快速增长,生物统计识别技术得到了新的重视,而人脸识别是所有生物识别方法中应用最广泛的技术之一。人脸识别因其在安全验证系统、信用卡验证、医学、档案管理、视频会议、人机交互、社会公安等方面的巨大应用前景而越来越成为人工智能领域的一个研究热点,而其关键的技术在于如何检测到人脸,并将其转化为计算机可识别的数据。

**关键词:** 人脸检测; 深度学习; 卷积神经网络;

**人脸检测:** 是一种在任意数字图像中找到人脸的位置和大小的计算机技术。它可以检测出面部特征,并忽略诸如建筑物、树木和身体等其他任何东西。被广泛应用于视频监控、人机交互、机器视觉等领域。

**深度学习:** 深度学习是指多层神经网络上运用各种机器学习算法解决图像、文本等各种问题的算法集合。深度学习允许由多个处理层组成的计算模型来学习具有多个抽象级别的数据表示,它通过使用反向传播算法来发现大型数据集中的复杂结构,以指示机器应如何更改其内部参数,根据前一图层的表示计算每个图层中的表示。其核心是特征学习,旨在通过分层网络获取分层次的特征信息,从而解决以往需要人工提取特征的重要难题。<sup>[1]</sup>

**卷积神经网络(CNN):** 卷积神经网络是深度学习模型中的一种多层人工神经网络,包括输入层、卷积层( convolution layer)、池化层( pooling layer)、全连接层( FC) 和输出层。CNN<sup>[4]</sup>包含提取输入图像特征的特征提取神经网络和基于图像特征进行运算的分类神经网络,特征提取神经网络接收输入图像,经过卷积层(  $C_1, \dots, C_n$ ) 的卷积操作后得到特征映射图,通过池化(  $P_1, \dots, P_n$ ) 对卷积后的特征图进行降维,将降维后的特征值输入至分类神经网络中的全连接层,最后输出判别结果。在卷积神经网络中,卷积层和池化层是其核心部分,一个 CNN 模型可以由多个卷积层和池化层交替组成。<sup>[2]</sup>

#### 5.1 基于深度学习的人脸检测

基于深度学习的人脸检测一般包括数据获取、预处理、特征提取、图形分类等环节。在深度学习框架下,学习算法直接从原始图像学习判别性的人脸特征。在海量人脸数据支撑下,基于深度学习的人脸识别在速度和精度方面已经远远超过人类,深度学习借助于图形处理器(GPU)组成的运算系统作大数据分析,人脸识别作为图像处理及人工智能的一个重要指标,证明深度学习模型有助于推动人工智能发展,将来甚至可能超越人类的智能水平。

其识别技术包括以下部分:

### 5.1.1 图像预处理

预处理是在提取特征之前排除与面部表情无关的一切干扰，如光照、头部姿势以及不同的背景等，目的是将面部对准到公共参考系，使得从每个面提取的特征对应于相同的语义位置。其主要方法有人脸检测、人脸对齐、数据增强、人脸归一化。

### 5.1.2 人脸对齐

虽然面部检测是实现特征学习的必要过程，但进一步的人脸对齐可以大大提高面部表情识别性能。人脸对齐可以看做在一张人脸图像搜索人脸预先定义的点，也称为人脸形状，通常从一个粗估计的形状开始，然后通过迭代来细化形状的估计，当下处理的主要方式为利用级联形状回归模型（CPR）<sup>[8]</sup>来进行脸部的对齐，这种方式对初始化形状的角度比较敏感，使用不同的初始化做多次测试并融合多次预测结果可以一定程度上缓解初始化对于算法的影响，但并不能完全解决该问题，且多次测试会带来额外的运算开销。当目标物体被遮挡时，性能也会变差。

### 5.1.3 数据增强

深度神经网络需要足够有效的训练数据以确保识别任务的普遍性，但是公开提供的 FER<sup>[1]</sup>数据集没有足够数量的图像用于训练，数据量少往往会导致过拟合现象。因此，数据增强是面部识别的关键步骤。常用的数据增强方法有旋转/反射变换、翻转变换、缩放变换、平移变换、尺度变换、对比度变换、噪声扰动、颜色变化等。

### 5.1.4 面部检测

其重点在于去除背景以及非面部区域，传统的人脸检测方法是利用人工提取特征来训练分类器进行人脸检测，例如 opencv<sup>[1]</sup>源码中自带的人脸检测器就是利用 Haar<sup>[1]</sup>特征进行的，但在环境变化强烈的时候检测效果不理想；而在深度学习领域，提出了性能更好的 Faster-RCNN、R-FCN<sup>[1]</sup>系列以及速度更快的 YOLO、SSD<sup>[1]</sup>系列来检测人脸，可以适应环境变化和人脸不全等问题

### 5.1.5 人脸归一化

照明和头部姿势的变化会很大程度影响面部表情识别性能，因此引入两种典型的人脸归一化方法来改善这些变化：灰度归一化和几何归一化。灰度归一化，是增加图像的亮度，使图像的细节更加清楚，以减弱光线和光照强度的影响；除了亮度调整外，还包含了对比度调整。常见的方法有直方图均衡化、基于各向同性扩散（IS）归一化、基于离散余弦变换（DCT）归一化、高斯（DoG）归一化<sup>[1]</sup>。

### 5.1.6 面部特征提取

面部识别中的关键步骤就是如何提取面部的特征，并将其转化为计算机可识别的信息，

表 1-1 列举了三种深度学习框架的优缺点。

深度学习方法	深度信念网络	自动编码器方法	深度卷积神经网络
学习方式	无监督学习	无监督学习	监督学习
优点	数据去噪	可视化降维	可直接输入原始图像
缺点	学习过程慢	没有全局优化	需要大量样本
网络组成	卷积 DBN	神经网络编码器	卷积层、池化层、全连接层

表 1-1<sup>[1]</sup>

### (1) 深度信念网络

2006 年, Hinton 等人提出深度信念网络及其高效的学习算法, 并发表于《Science》上, 成为其后深度学习算法的主要框架。训练过程包括预训练和调优过程, 其中预训练过程相当于逐层训练每一个 RBM(玻尔兹曼机)<sup>[1]</sup>, 经过预训练的 DBN 已经 可用于模拟训练数据。而为了进一步提高网络的判别性能, 微调过程利用标签数据通过后向传播 (back propagation, BP) 算法对网络参数进行微调。DBN 根据人脑分级信息处理抽象 出多种表征并学习对象的特征 (知识) 层次, 实现对知识的分层次理解。它是一种生成模型, 通过训练其神经元间的权 重, 可以让整个神经网络按照最大概率来生成训练数据。所以不仅可以 使用 DBN<sup>[7]</sup>识别特征和分类数据, 还可以用它来生成数据。

### (2) 自动编码器方法

为了观测到面部表情细微的变化, 1978 年 Ek-man 和 Friesen 开发了面部运动编码系统<sup>[3]</sup>, 这个系统包含 44 个运动单元, 其中 30 个与解剖学中特定的面部肌肉运动相关, 如 AU1 对应了眉毛内侧肌肉的上升, 其 余的 14 个在解剖学上没有特定的肌肉运动单元与之对应。采用 FACS<sup>[3]</sup>进行面部表情识别就是识别相应的运动单元, 其应用非常广泛, 是识别复杂面部运动单元的有效途径。

其中基于关键点 LBP<sup>[3]</sup>的面部特征提取方法被广为使用, LBP 特征可以很好描述局部纹理信息, 抗光照能力好并且易于计算, 因此受到许多研究者的青睐, 基本的 LBP 算子是一个固定大小为 3x3 的矩 形块, 这个矩形块中的 9 个子块对应图像中的 9 个 灰度值, 将中心点邻域中的 8 个灰度值与中心灰度 值相比较, 大于等于中心灰度值的子块赋予值 1, 反之, 则赋予值 0。根据顺时针方向读取 8 个新赋的值, 并把二进制值转化成十进制, 该值就是这个 3x3 模块中心点处的 LBP 描述值。对于图像中 给定的一点 (x<sub>0</sub>, y<sub>0</sub>), 其 LBP 的特征值

可以用公式表示为:

$$LBP(x_0, y_0) = \sum_{i=1}^8 s(g_i - g_0) * 2^{i-1} \quad [3], \text{ 其中 } g_i \text{ 表示给定点}$$

$$s(g) = \begin{cases} 1, & \text{if } g \geq 0 \\ 0, & \text{if } g < 0 \end{cases}$$

的 3x3 邻域中第 i 个点的灰度值, 函数 s(g)定义为<sup>[3]</sup>。计算出图像中 每个点的 LBP 描述值以后, 对其进行直方图统计, 最后就得到了整幅图像的 LBP 描述子。

### (3) 深度卷积神经网络

实际场景中的人脸图片, 不仅包含人脸, 还有复杂的背景信息, 并且受人脸多姿态变化以及人脸面部特征不完整等外在因素影响较大。在用单个 CNN 模型实现人脸检测时, 保证检测精度的前提是 建立在复杂的网络结构设计之上的, 带来的时耗影响难以克服。由此, 在进行人脸检测的过程中, 为了兼顾检测精度和检测效率, 需要一种多层次的深度卷积神经网络进行识别。

一种融合多任务联合多尺度卷积神经网络的人脸分析方法，包含如下步骤：<sup>[5]</sup>

- 1) 对一张尺寸大小为  $N \times N$  的待学习的图片，采用关键区域搜索算法从所述图片中提取出  $K$  个不同尺度的人脸感兴趣区域，将其作为多尺度 CNN 三个通道的输入；其中， $N$  表示像素大小， $K \geq 2$  且为整数；
- 2) 采用 CNN 分别对  $K$  个所述人脸感兴趣区域进行特征提取，得到不同尺度的人脸特征，采用级联的方式将提取到的不同尺度的人脸特征进行融合，得到融合后的特征表达；
- 3) 将多个任务的损失函数进行融合得到联合损失函数，得到关于所述联合损失函数的最优解，从而得到所述多个任务的处理结果

## 5.2 面部图像分类

### 5.2.1 KNN 算法进行面部图像分类

当我们通过特征提取获得了人脸的面部特征数据后，就需要对该数据进行分类或者叫识别，以确定该图像的身份，当下的分类算法中，性能较好的是 KNN<sup>[6]</sup>（最邻近规则法），KNN 算法原理是计算输入数据和训练好的模型之间的距离，在阈值  $K$  之内的样本进行分类。如图 1 所示，对输入数据 进行分类，然后计算 与任意 一点之间的距离从而进行分类。<sup>[6]</sup>

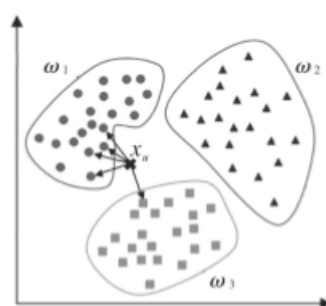


图 1<sup>[6]</sup>

算法主要步骤：

算距离：计算  $X$  样本和每个训练样本的距离或者相似度

找邻居：得到  $k$  个距离最近的样本；

做分类：统计  $k$  个距离最近样本中出现次数最多的归属类别，确定  $X$  样本的类别属性  
当我们通过跟数据库中的样本进行比对后，便可以估计出该面部图像的身份了。

### 5.2.2 支持向量机算法进行面部图像分类

常用的分类算法还有 SVM（支持向量机）算法<sup>[1]</sup>，支持向量机是针对二值分类问题提出的，并且成功地应用于解函数回归及一类分类问题，虽然支持向量机在解决二值分类问题时获得了巨大的成功，但实际应用中的大量多值分类问题也进一步要求如何将支持向量机推广到多分类问题上。它的基本模型是定义在特征空间上的间隔最大的线性分类器，间隔最大使它有别于感知机，它的学习策略就是间隔最大化，可形式化为一个求解凸二次规划的问题，也等价于正则化的合页损失函数的最小化问题，支持向量机的学习算法是求解凸二次规划的最优化算法。

SVM 算法是利用对目标函数的优化来找到分类间距最大的“决策面”，间隔的大小实际

上就是支持向量对应的样本点到决策面的距离的二倍，如图 2 所示。

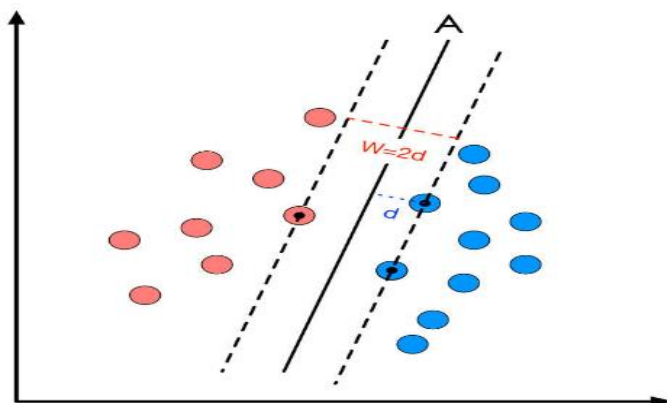


图 2

当然，这只是 SVM 的二元分类的展示，我们可以用它的多远分类来预测该人脸图像的身份。

### 5.3 结论

通过查阅相关资料后可以得出，人脸检测和识别的大致流程为：

- (1) 读取需要识别的图片；
- (2) 用数据集训练深度学习网络算法模型；
- (3) 检测出人脸后，用最邻近规则法(KNN)算法对人脸进行识别。

在进行人脸检测之前，需要对面部图像进行去噪、数据增强等预处理，以提高检测的准确度；之后便建立深度学习下的 CNN 神经网络，对采集到的大量人脸数据进行模型学习，提取其中的特征点；然后使用训练好的 CNN 神经网络对要识别的人脸图像进行回归检测，输出其特征点；最后使用 KNN 算法或者 SVM 算法在数据库中对提取出的特征数据进行比对，确定人脸图像的身份。

总之人脸识别的难点和关键点在于人脸检测，而传统的人工检测方法已经不再适用于新的时代，而随着机器性能的更新换代而越来越实用的深度学习算法则逐渐登上了舞台，它以其高效、快速的识别方式为人脸检测提供了更多的实现可能，而人脸检测的关键点在于特征点的提取，也就是轮廓的识别，在这方面普遍采用的方式是通过训练 CNN 卷积神经网络来识别某个面部图像的特征点，有了这样高效而又快捷的面部检测方式，相信以后的面部识别技术会越来越发达。

## 6 总结与展望

- (1) 本文针对基于卷积神经和深度学习的人脸检测前沿技术进行了综述与分析，阐述了各类方法的优点与不足，希望能给领域内的研究人员带来一些新的信息和研究思路。
- (2) 人脸检测问题一直以来是计算机视觉、模式识别领域的研究热点，尽管学习方法具有强大的特征学习能力，但应用于面部表情检测时仍存在问题，其鲁棒性有待进一步提高。

- (3) 人脸检测作为一种非常具有实用价值的身份识别技术,长期以来一直受到学术界和工业界的广泛关注,相关的理论研究和应用开发已经取得了丰硕的成果。在大量研究人员的共同努力下,现存的难题将会逐步被解决,各种成熟的应用产品也将会逐渐面世。我们可以期待,在不久的将来,人脸检测将会成为一种方便人们日常生活的常见技术。

## 7 参考文献

- [1] 陆嘉慧 张树美 赵俊莉 基于深度学习的面部表情识别研究[J]. 青岛大学数据科学与软件工程学院; 10.19734/j.issn.1001-3695.2018.10.0723
- [2] 张海涛 李美霖 董帅含 两层级联卷积神经网络的人脸检测[A] 辽宁工程技术大学软件学院; 1006-8961( 2019) 02-0203-12
- [3] 任艳 基于关键点 LBP 特征提取的面部 AU 组合识别[J] 东南大学学习科学研究中心; 210096
- [4] 朱虹 李千目 李德强 基于单个卷积神经网络的面部多特征点定位[J] 南京理工大学计算机科学与工程学院;210094
- [5] 刘袁缘;周顺平;张香兰;方芳;郭明强;姚尧;彭济耀; 融合多任务和多尺度卷积神经网络的人脸分析方法及系统[P] 中国地质大学(武汉); 201811260674. X
- [6] 刘其嘉 郭一娜 任晓文 李健宇 基于多任务级联卷积网络模型的人脸检测和识别[j] 太原科技大学电子与信息工程学院; 1673 —2057(2019)02 —0081 —05
- [7] Dang L,M Hassan S I Im S Lee JLee SMoon H Deep Learning Based Computer Generated Face Identification Using Convolutional Neural Network [A] APPLIED SCIENCES-BASEL
- [8] A.U.Batur; B.E.Flinchbaugh A DSP-based approach for the implementation of face recognition algorithms 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).



## 8 评价

1. 赵小钦: 经过本次文献检索调查活动, 我对文献检索的工具和方法有了更深的了解, 学习到了在查阅相关知识点时的检索方法和检索策略。在本次活动中, 我担任小组组长, 主要工作为协调小组成员间的工作, 以及收集资料和撰写综述报告, 依赖于强大的文献检索引擎, 我从本次活动中学习到了很多从未了解过的知识, 包括人脸识别的原理、相关的算法、以及对应的工具等, 同时也增加了自己的论文写作经验, 通过本次活动, 我对检索工具的掌握更加熟练了, 同时也了解到了除百度以外的其他检索工具, 即开拓了我的眼界, 又丰富了我的实践经验, 我从本次活动中学习到了很多。
2. 刘天保: 在文献检索课外报告合作的编写过程中, 对老师上课讲的一系列操作进行了实践, 巩固了之前理解的内容, 同时也发现了不少问题。我们的课题是面部检测技术的文献检索与总结, 在此之前也未曾了解过该方面的知识, 乘机机会也学习到了不少, 了解了平时常用的手机面部检测的一个大概原理; 在学习这门课程之前, 不懂的问题都是通过百度寻找答案, 而文献检索课程也正是为查阅专业知识方面问题提供了一种新的途径。
3. 彭留洋: 在这次的文献检索报告中, 本人主要负责资料的整理以及搜集, 摘要撰写, ppt 制作汇报等工作, 在这次的报告中, 遇到了不少的麻烦, 通过询问小组成员, 再加上网上查资料, 及时的得到了解决, 也加上了一些自己的想法, 感觉学到了不少新的知识, 总体来说, 《文献检索与利用》这门课让我知道了怎样查找有用的信息, 刚开始学这门课的觉得学来没什么用, 后来觉得对我们的工作还是有许多帮助的, 非常感谢老师的教导。
4. 张天宇: 学习了文献检索这门课程, 我才发现, 通过图书馆的电子资源, 我们可以查询到许许多多的有用文献, 对我们的学习具有相当大的作用, 另外, 还让我构成了借助这些数据库进行自主学习的习惯, 只要有需要, 我就会在这些数据库中查询自己感爱好的东西, 用来丰富自己的综合知识。合作报告的编写过程中, 我动手实践了这些理论知识并参与了 ppt 的制作, 整理和汇总了这次实践的收获, 提升了我极大的能力和知识面。
5. 张弛: 我觉得这门课程十分有用, 对学习十分有帮助, 老师课堂上能再多举一些例子就更好了。这次报告中, 我参与了查找文献、ppt 的制作以及答辩, 使我对文献的查找有了清晰的认识, 也增强了 ppt 的制作能力。
6. 秦思远: 通过对本次课程学习, 让我学会了如何进行文献检索的相关知识和技术, 首先能够更加精准的查找自己所需要的资料, 节省不必要的检索时间; 再者文献查找的途径渠道和方式上有了更加系统的认识; 并且了解了文献的各种含义。文献检索的作用不仅仅是为了完成课程要求, 更重要的是在实际使用中能够更方便、更准确、更专业进行检索。课程中老师的责任心和技术实力让我对文献检索课程产生了浓厚的兴趣, 不断的加强我对文献的认识和检索的能力, 在实验中, 通过亲自检索文献和完成试题, 使我更加清晰和明确文献检索的意义。

## 9 附件

### 1. 张海涛 李美霖 董帅含 两层级联卷积神经网络的人脸检测 辽宁工程技术大学软件学院:

E-mail: jlg@radl.ac.cn  
Website: www.cjlg.cn  
Tel: 010-58887035

中国图象图形学报  
JOURNAL OF IMAGE AND GRAPHICS  
©中国图象图形学报版权所有

203

中国法分类号: TP391 文献标识码: A 文章编号: 1006-8961 (2019) 02-0203-12

论文引用格式: Zhang H T, Li M L, Dong S H. Two-layer cascaded convolutional neural network for face detection [J]. Journal of Image and Graphics, 2019, 24 (02): 0203-0214. [张海涛, 李美霖, 董帅含. 两层级联卷积神经网络的人脸检测 [J]. 中国图象图形学报, 2019, 24 (02): 0203-0214.] [DOI: 10.11834/jig.180248]

## 两层级联卷积神经网络的人脸检测

张海涛, 李美霖, 董帅含

辽宁工程技术大学软件学院, 葫芦岛 125105

**摘要:** 目的 传统人脸检测方法因人脸多姿态变化和人脸面部特征不完整等问题, 导致检测效果不佳。为解决上述问题, 提出一种两层级联卷积神经网络 (TC-CNN) 人脸检测方法。方法 首先, 构建两层卷积神经网络模型, 利用前端卷积神经网络模型对人脸图像进行特征粗略提取, 再利用最大值池化方法对粗提取得到的人脸特征进行降维操作, 输出多个疑似人脸窗口; 其次, 将前端粗提取得到的人脸窗口作为后端卷积神经网络模型的输入进行特征精细提取, 并通过池化操作得到新的特征图; 最后, 通过全连接层判别输出最佳检测窗口, 完成人脸检测全过程。结果 实验选取 FDDB 人脸检测数据集中包含人脸多姿态变化以及人脸面部特征信息不完整等情况的图像进行测试, TC-CNN 方法人脸检测率达到 96.39%, 误检率低至 3.78%, 相比当前流行方法在保证算法效率的同时检测率均有提高。结论 两层级联卷积神经网络人脸检测方法能够在人脸多姿态变化和面部特征信息不完整等情况下实现精准检测, 保证较高的检测率, 有效降低误检率, 方法具有较好的鲁棒性和泛化能力。

**关键词:** 人脸检测; 卷积神经网络; 十折交叉验证; 两层级联卷积神经网络; 最大值池化

## Two-layer cascaded convolutional neural network for face detection

Zhang Haitao, Li Meilin, Dong Shuaihan

College of Software, Liaoning Technical University, Huludao 125105, China

**Abstract: Objective** As an important part of face recognition, face detection has attracted considerable attention in computer vision and has been widely investigated. Face detection determines the location and size of human faces in an image. Traditional face detection methods are limited by face multi-pose changes and incomplete facial features, which lead to their poor detection effect. Modern face detectors can easily detect near-frontal faces. Recent research in this area has focused on the uncontrolled face detection problem, where a number of factors, such as multi-pose changes and incomplete facial features, can lead to large visual variations in face appearance and can severely degrade the robustness of the face detector. A convolutional neural network can automatically select facial features, rapidly delete a large number of non-face background information, and can achieve good face detection results. However, a single convolutional neural network should possess three functions, namely, facial feature extraction, reduction of feature dimensions to decrease the computational complexity, and feature classification, which result in complex network structure, limited detection speed, and overfitting of the network. To solve these problems, this study presents a face detection method of two-layer cascaded convolutional neural network (TC-CNN). **Method** First, a two-layer convolutional neural network model is constructed. The first convolutional neural network model is used to extract the features of the face image, and a max pooling method is adopted to reduce the dimension of those features in which multiple suspected face windows are outputted. Second, the face windows are used as

收稿日期: 2018-04-12; 修回日期: 2018-08-31; 预印本日期: 2018-09-07

基金项目: 中国人民解放军总装备部装备预研基金项目 (61421070101162107002); 辽宁省自然科学基金面上项目 (20170540426)

**Supported by:** Natural Science Foundation of Liaoning Province, China (20170540426)

2. 刘其嘉 郭一娜 任晓文 李健宇 基于多任务级联卷积网络模型的人脸检测和识别  
太原科技大学电子与信息工程学院:

第 40 卷 第 2 期

太 原 科 技 大 学 学 报

Vol. 40 No. 2

2019 年 4 月

JOURNAL OF TAIYUAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Apr. 2019

文章编号: 1673 - 2057(2019)02 - 0081 - 05

## 基于多任务级联卷积网络模型的人脸检测和识别

刘其嘉, 郭一娜, 任晓文, 李健宇

(太原科技大学 电子与信息工程学院, 太原 030024)

**摘 要** 深度学习的集成特征提取这一优点使得它广泛应用于人脸检测和识别。提出了一种多任务级联卷积网络模型(Multitask Cascaded Convolution Network, MTCNN)。基于 Tensor Flow 平台, 基于改进的任务级联卷积网络模型检测到人脸, 并且用 Face Net 算法对人脸进行特征提取, 用 KNN 算法对人脸进行识别。实验结果表明, 对不同光照下多人图像和遮挡图像的人脸进行检测和识别, 具有良好的鲁棒性。

**关键词** 人脸检测和识别; 深度学习; Tensor Flow; 多任务级联卷积网络; Face Net

**中图分类号** TP3 **文献标志码** A **doi**: 10.3969/j.issn.1673-2057.2019.02.001

近年来, 互联网技术和计算机视觉的发展和兴盛使得全球化日益明显, 信息安全已经成为人类必不可少的环节。现在社会中数据传输量数以亿计的发生, 个人信息保密也将是很重要的技术, 生物识别技术作为个人信息保密的重要组成部分。其中人脸检测和识别技术比其他生物识别技术具有自然性, 非强制性, 非接触性, 并发性等优点。人脸检测和识别一直是机器语言, 计算机视觉等领域的热点。

人脸检测和识别的难点在于男女的分辨, 用户的配合度, 个体之间的相似性和易变性<sup>[1]</sup>。从非深度到深度的方法的发展, 使得人脸检测和识别的精度明显提高。1991 年, 麻省理工大学的 Turk<sup>[2]</sup> 等人提出了一种特征脸算法的人脸识别, 基于主成分分析, 通过 K-L 矩阵变换去识别人脸。1996 年 Sweets<sup>[3]</sup> 等人提出了一种线性判别方法, 通过对人脸图像降维去匹配人脸。这些非深度的方法简单并且速度快, 缺点是人脸检测效果不好, 鲁棒性不好。自从 Hinton 提出了深度信念网络<sup>[4]</sup>, 通过贪婪算法去训练模型, 然后反向传播优化参数。现在已

经成功应用到人脸识别。2001 年 Viola 等人提出基于 Ada Boost 算法<sup>[5]</sup> 的人脸检测器, 这个算法在人脸技术的发展史上有着里程碑的作用, 相比原来的算法明显的提高了人脸检测和识别的速度和准确率。2006 年 Hinton 等人在深度信念网(Deep Belief Networks, DBNs) 上革命性的发展使得深度学习进入了全新的领域。人脸技术和深度学习的结合随之而来。2014 年 Yaniv Taigman<sup>[6]</sup> 等人通过额外的 3D 模型, 改进了人脸对齐的方法, 基于大数据训练了 9 层的人工神经网络来进行人脸特征表达, 数据集上准确率打到了 0.9735。可见, 深度学习对于人脸检测和识别有着巨大的推动力。然而这些深度学习算法训练出来的模型对多目标和遮挡图像的识别精度不是很高, 迫切需要寻求快速识别和匹配人脸非刚性的算法和模型。

本文为解决 CNN 模型的识别精度问题, 在 CNN 的基础上对多任务级联卷积网络模型(Multitask Cascaded Convolution Network, MTCNN) 进行改进, 采用 Face Net 模型和 KNN 算法对人脸达到识别的目的, Face Net 算法可以实现人脸识别的结果。

收稿日期: 2017-10-10

基金项目: 国家自然科学基金(61301250); 山西省高等学校优秀青年学术带头人支持计划(晋教科[2015]3号); 山西省回国留学人员科研资助项目(2014-060)。

作者简介: 刘其嘉(1990-), 男, 硕士研究生, 研究领域为智能信号处理与模式识别。

3. 刘袁缘;周顺平;张香兰;方芳;郭明强;姚尧;彭济耀;融合多任务和多尺度卷积神经网络的人脸分析方法及系统 中国地质大学(武汉):

(19)中华人民共和国国家知识产权局



(12)发明专利申请



(10)申请公布号 CN 109508654 A

(43)申请公布日 2019.03.22

(21)申请号 201811260674.X

(22)申请日 2018.10.26

(71)申请人 中国地质大学(武汉)

地址 430000 湖北省武汉市洪山区鲁磨路  
388号

(72)发明人 刘袁缘 周顺平 张香兰 方芳  
郭明强 姚尧 彭济耀

(74)专利代理机构 武汉知产时代知识产权代理  
有限公司 42238

代理人 龚春来

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

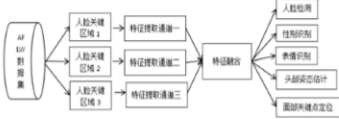
权利要求书2页 说明书6页 附图2页

(54)发明名称

融合多任务和多尺度卷积神经网络的人脸  
分析方法及系统

(57)摘要

本发明公开了一种融合多任务联合多尺度卷积神经网络的人脸分析方法及系统,先对一张尺寸大小为 $N \times N$ 的待学习的图片,采用关键区域搜索算法从所述图片中提取出 $K$ 个不同尺度的人脸感兴趣区域,将其作为多尺度CNN三个通道的输入;然后采用CNN分别对 $K$ 个所述人脸感兴趣区域进行特征提取,得到不同尺度的人脸特征,采用级联的方式将提取到的不同尺度的人脸特征进行融合,得到融合后的特征表达;再将多个任务的损失函数进行融合得到联合损失函数,以所述特征表达作为学习输入,得到关于所述联合损失函数的最优解,从而得到所述多个任务的处理结果。本发明利用任务间的相关性互相促进来,提高了单个任务的预测准确率。



CN 109508654 A