

学校编码: 10384

分类号\_\_\_\_密级

学 号: 24320141152399

UDC

厦 门 大 学

硕 士 学 位 论 文

基于级联卷积神经网络的人脸检测算法的  
研究

Research on Face Detection Algorithm Based on Cascaded  
Convolutional Neural Networks

刘伟强

指导教师姓名: 王备战 教授

专业名称: 计算机科学与技术

论文提交日期: 2017 年 4 月

论文答辩日期: 2017 年 5 月

学位授予日期: 2017 年 月

指 导 教 师:

答 辩 委 员 会 主 席:

2017 年 4 月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（）课题（组）的研究成果，获得（）课题（组）经费或实验室的资助，在（）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年 月 日

# 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1.经厦门大学保密委员会审查核定的保密学位论文，  
于 年 月 日解密，解密后适用上述授权。

2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

## 摘 要

在人脸识别领域中,人脸检测是其中的一个重要环节,2001年 Viola 和 Jones 发明了 VJ 人脸检测算法,该算法的诞生使人脸检测能够得到实际应用,但对于人脸多角度、多光照、多姿态等问题,VJ 人脸检测器表现不佳。DPM 算法的引入较好解决了复杂情况下的人脸检测问题,但是该算法存在计算量大,前期人工标记成本大等问题。神经网络特别是卷积神经网络的发展和 GPU 编程的实现给图像处理、音频处理、视频处理等领域注入了新的活力。卷积神经网络(CNN)在人脸检测领域有着无可比拟的优点,比如自学习特征、较好适应复杂情况下人脸、节省人力成本等。CNN 用于人脸检测效果良好,但也存在着解释性较差的缺点。虽然人脸检测算法取得了巨大的进步,但是其中的很多工作并没有非常完美地解决,依然存在着巨大的挑战,如图像质量问题、人脸遮挡、多姿态、多表情、多光照、实时检测等问题,这些问题亟需解决。

目前,通常采用传统方法和深度学习方法结合来解决人脸检测的问题。本文设计了一种基于级联卷积神经网络的人脸检测算法,级联结构策略权衡了算法准确度和运行时间,在保证较高准确度的情况下,运行时间尽量短。算法通过第一阶段的全卷积神经网络(FCN)提取人脸候选区域,比 Selective Search、Edge Box 等算法更高效。在第二、第三阶段特征学习能力更强的网络中,结合边框回归、NMS 等算法进一步筛选人脸候选框,校准人脸框位置,最终输出头像中是否含有人脸,如果有,则输出人脸位置。为了提高人脸框的准确度和增强算法辨别人脸的能力,本文在训练方法上进行改良,优化了训练数据集,并设计了多任务学习网络。

实验表明,该算法在 FDDB 上的人脸检测上准确度较高,检测时间较短。

**关键词:** 人脸检测; 深度学习; 多任务学习

## Abstract

Face detection is an important step in face recognition. In 2001, Viola and Jones invented the VJ face detector which made it possible in practical use. However, the VJ face detector has poor performance while detecting the face in the wild. The introduction of DPM algorithm solves the problem of face detection in complex cases, but it has large cost in computation and manual marking. Thanks to the development of neural network, especially the convolutional neural network, and the realization of GPU programming, they provide a more effective technical solution for image processing, audio processing, video processing and so on. Convolutional neural network (CNN) has incomparable advantages in the field of face detection, such as self-learning characteristics, better adapting to the complex situation of the face, saving manual costs, etc. CNN is good for face detection, but it also has the disadvantage of poor interpretability. Although the face detection algorithm has made great progress, but many of them didn't work perfectly, there are still huge challenges, such as image quality, face pose and occlusion, facial expressions and illumination, real-time detection and other issues, these problems need to be solved.

At present, the combination of the deep learning and the traditional method can solve the problem of face detection. In this dissertation, we design a face detection algorithm based on cascaded convolutional neural network, and the cascade structure is used to balance the accuracy and running time cost. We use full convolutional neural network ( FCN ) to exact candidate regions of human face in the first stage, which is more efficient than selective search, edge box and other algorithms. Combining with the NMS algorithm and bounding box regression during the whole process, we can get the more accurate of face position. In order to improve the accuracy and enhance the ability of the algorithm to distinguish the face, we try to improve the training method and optimized the training set, finally we realized the multi-task learning network.

The experimental results show that the algorithm has high accuracy in Fddb

data set, and the detection time is short.

**Key Words:** Face Detection; Deep Learning; Multi-task Learning

厦门大学博硕士学位论文摘要库

## 目 录

<b>第一章 绪论</b> .....	<b>1</b>
1.1 研究背景与意义 .....	1
1.2 国内外研究现状 .....	3
1.3 论文主要研究内容及组织结构 .....	6
<b>第二章 相关理论介绍</b> .....	<b>8</b>
2.1 人脸检测概述 .....	8
2.1.1 算法框架 .....	8
2.1.2 人脸库及测评评价标准 .....	9
2.2 人工神经网络 .....	11
2.3 卷积神经网络 .....	13
2.3.1 卷积层 .....	14
2.3.2 ReLU 层 .....	15
2.3.3 池化层 .....	16
2.3.4 全连接层 .....	17
2.4 人脸检测相关算法介绍 .....	17
2.4.1 VJ 人脸检测器 .....	17
2.4.2 级联网络人脸检测算法 .....	19
2.4.3 区域推荐 (Region Proposal) 算法 .....	21
2.5 本章小结 .....	24
<b>第三章 基于级联卷积神经网络的人脸检测算法设计</b> .....	<b>25</b>
3.1 总体设计 .....	25
3.2 网络数据结构 .....	26
3.3 级联网络设计 .....	30
3.3.1 人脸候选区域网络设计 .....	30
3.3.2 强化网络设计 .....	33
3.3.3 输出网络设计 .....	35



3.4 损失函数.....	38
3.4.1 损失函数概述.....	38
3.4.2 损失函数.....	38
3.5 本章小结.....	41
<b>第四章 难分负样本挖掘及多任务学习实现.....</b>	<b>42</b>
4.1 难分负样本挖掘实现.....	42
4.1.1 难分负样本挖掘概述.....	42
4.1.2 难分负样本挖掘实现.....	43
4.2 多任务学习实现.....	45
4.2.1 多任务学习概述.....	45
4.2.1 多任务学习实现.....	46
4.3 本章小结.....	50
<b>第五章 实验结果 .....</b>	<b>51</b>
5.1 实验环境搭建.....	51
5.2 实验数据处理.....	52
5.3 级联网络训练.....	57
5.4 实验结果.....	59
5.5 本章小结.....	65
<b>第六章 总结与展望 .....</b>	<b>66</b>
6.1 总结.....	66
6.2 展望.....	67
<b>参考文献.....</b>	<b>69</b>
<b>硕士期间发表学术论文情况 .....</b>	<b>73</b>
<b>致谢.....</b>	<b>74</b>

## Contents

<b>Chapter 1 Introduction .....</b>	<b>1</b>
<b>1.1 Research Background and Significance.....</b>	<b>1</b>
<b>1.2 Overseas and Domestic Research Status.....</b>	<b>3</b>
<b>1.3 Main Research Contents and Structure Arrangements .....</b>	<b>6</b>
<b>Chapter 2 Related Theories Overview.....</b>	<b>8</b>
<b>2.1 Overview .....</b>	<b>8</b>
2.1.1 Algorithm Frameworks .....	8
2.1.2 Face Database and Evaluation Standards.....	9
<b>2.2 Artificial Neural Networks.....</b>	<b>11</b>
<b>2.3 Converlutional Neural Networks.....</b>	<b>13</b>
2.3.1 Convolutional Layer .....	14
2.3.2 ReLU Layer .....	15
2.3.3 Pooling Layer.....	16
2.3.4 Fully Connected Layer.....	17
<b>2.4 Face Detection Algorithm.....</b>	<b>17</b>
2.4.1 VJ Face Detector .....	17
2.4.2 Cascade CNN.....	19
2.4.3 Region Proposals Algorithm .....	21
<b>2.5 Summary.....</b>	<b>24</b>
<b>Chapter 3 Design of Face Detection based on Cascade CNN.....</b>	<b>25</b>
<b>3.1 Overall Design .....</b>	<b>25</b>
<b>3.2 Network Data Structure .....</b>	<b>26</b>
<b>3.3 Cascade Network Design .....</b>	<b>30</b>
3.3.1 Region Proposals Network Design .....	30
3.3.2 Strengthen Network Design.....	33
3.3.3 Output Network Design.....	35

---

<b>3.3 Loss Function</b> .....	<b>38</b>
3.3.1 Overview .....	38
3.3.2 Loss Function.....	38
<b>3.4 Summary</b> .....	<b>41</b>
<b>Chapter 4 Hard Negative Mining and Multi-task Learning</b> .....	<b>42</b>
<b>4.1 Hard Negative Mining</b> .....	<b>42</b>
4.1.1 Overview .....	42
4.1.2 Realization of Hard Negative Mining .....	43
<b>4.2 Multi-task Learning</b> .....	<b>45</b>
4.2.1 Overview .....	45
4.2.1 Realization of Multi-task Learning .....	46
<b>4.3 Summary</b> .....	<b>50</b>
<b>Chapter 5 Experimental Results</b> .....	<b>51</b>
<b>5.1 Experimental Environment</b> .....	<b>51</b>
<b>5.2 Data Processing</b> .....	<b>52</b>
<b>5.3 Training of Cascade CNN</b> .....	<b>57</b>
<b>5.4 Results</b> .....	<b>59</b>
<b>5.5 Summary</b> .....	<b>65</b>
<b>Chapter 6 Conclusions and Prospects</b> .....	<b>66</b>
<b>6.1 Conclusions</b> .....	<b>66</b>
<b>6.2 Prospects</b> .....	<b>67</b>
<b>References</b> .....	<b>69</b>
<b>Publications</b> .....	<b>73</b>
<b>Acknowledgements</b> .....	<b>74</b>

厦门大学博硕士学位论文摘要库

## 第一章 绪论

### 1.1 研究背景与意义

人脸识别指的是基于人脸的特征信息进行生物识别，是计算机视觉中的一个重要研究领域。人脸识别技术应用的场景很多，大都在自然情况下发生，且不需要强制执行，也不需要让人为接触某些设备，并且可以同一时间段内并发检测，该技术发展到目前为止，精度高并且安装使用简易，稳定性高。人脸识别技术也是一种生物识别技术，具有难以伪造的特性，具有极其广阔的市场应用前景。

自 2015 年来，人脸识别瞬间爆发，成为了科技圈中的热门对象，有关人脸识别的应用迅速崛起，成为了生物识别领域的新宠。人脸识别技术不仅在理论研究上大放异彩，也在应用层面得到了快速发展。旷世科技在支付宝、园区门禁、智能分析等商业应用中“稳坐宝座”，云从科技在公安、银行、机场、火车站等行业应用“深耕细作”，还有商汤科技主要提供 SDK、API 等服务。总结起来，人脸识别可运用到安防、金融、平安城市、教育、交通、监狱、办公楼、住宅等领域。

人脸识别在最近几年发展迅速，其热门程度早已超越了指纹识别技术，是什么原因让人脸识别技术迅速成为火爆热门的技术？首先是技术层面的原因，人工智能、深度学习和大数据等热门技术的发展与成熟给了人脸识别新的解决思路。特别是近年来，卷积神经网络的技术发展为人脸识别问题提供了良好的技术基础，让原本费时费力，难以解决的问题都有了较好的解决方案，技术的进步给了人脸识别的实际应用提供了强有力的支持。第二个方面主要是社会需求的增长，随着我国城镇化进程的加快，人员流动大，各种突发情况越来越频繁。特别是在安防领域，传统的一些方法已经不再适用，必须通过技术来提高执法人员的效率。除此之外，金融安全、出行安全等领域均需要人脸识别产品的支持。

人脸识别主要经过两个阶段，第一阶段是主要是对图片进行预处理，包括

人脸检测、特征点定位与人脸对齐。第二阶段是识别阶段，该阶段的主要任务是认证、鉴别与属性分类。认证指的是判断两张人脸图片是否属于同一个人，鉴别指的是计算输入人脸图片的身份，而属性分类指的是分析人脸图片中的一些人脸部特征。人脸识别框架图如图 1-1 所示。论文主要针对第一阶段的人脸检测进行研究。

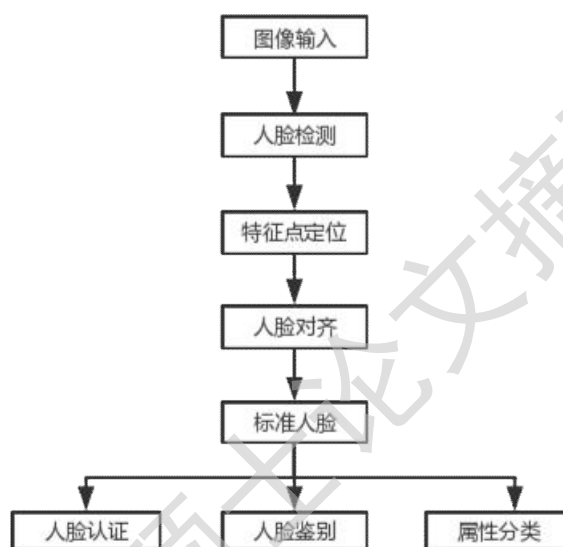


图 1-1: 人脸识别框架

人脸检测(face detection)的主要任务是对于输入的一张图片，判断该图片中是否包含人脸，如果包含人脸则需要确认人脸的位置以及人脸中的眼睛、鼻子、嘴巴等重要部位的位置。人脸检测在实际应用中具有重要意义，该项研究早在上个世纪 70 年代就已经开始。但是由于受到技术条件的限制和需求不大的影响，人脸检测直到上个世纪 90 年代才开始加快向前发展的脚步。

到目前为止，人脸检测取得了非常多的成果，但是人脸检测依然还有许多问题没有很好的解决，人脸检测仍然是一个富有挑战性的任务。主要问题体现在以下几个方面：

- 1、人对事物的感知过程仍然不清楚，对于人脸也一样，人们对自身的研究还不够。
- 2、图像的质量，不同光照下，人脸的变化比较大。
- 3、人脸的姿态，VJ 人脸检测器较好解决了正脸的检测，但是实际情况下的

人脸姿态千变万化：低头、抬头、正脸旋转、侧脸旋转等。

4、人脸的表情，喜怒哀乐是人传达情感的一种方式，这些情感提现在人脸会上会让人脸产生一定的变化，面部特征将会改变。

5、人脸的遮挡，如戴上墨镜、眼镜，头发遮挡脸部，一些装饰品等的遮挡都会对人脸检测产生较大的影响。

人脸识别应用广泛，具有巨大的商业价值。深度学习在人脸识别等领域具有非常大的技术优势，值得大家去深究。目前，人脸检测仍然还是个富有挑战的工作等待着大家去探索。

## 1.2 国内外研究现状

早期的人脸检测关注的都是人的正脸，通过图像边缘信息和灰度值对人脸图像进行有限的分析，引入模式识别<sup>[1]</sup>方法，结合先验知识来设计人脸模型和算法。模板匹配主要是通过人脸特征模板与目标图像相同位置的像素灰度值或者颜色值进行比较，通过计算灰度值之和或者颜色之和的差来表示偏离程度。Adaboost<sup>[2]</sup>算法训练级联的分类器，效果良好。在 2001 年，人脸检测技术得到了突破，由 Viola 和 Jones<sup>[3]</sup>发明的 VJ 人脸检测器极大提高了实时的人脸检测性能，Viola-Jones (VJ) 人脸检测器优化了 Haar<sup>[4]</sup>特征，引入了积分图。但是在现实世界中，人脸图像千变万化，受到各种因素的影响，如光线明暗不一、脸部动作千奇百怪、人脸不完整、背景复杂、人脸角度不一致等，如图 1-2 表示复杂情况下的人脸图像。在这种复杂情况下，VJ 人脸检测器的检测结果并不理想。



图 1-2: 不同光照、表情下的人脸

对于这个问题，人脸检测引入了 Pedro Felzenszwalb 的 DPM<sup>[5]</sup>(Deformable Parts Model)目标检测算法，因为 DPM 目标检测算法是基于部件检测，所以对于复杂情况下的人脸、扭曲、角度、姿态等问题，检测效果较好。但是该算法计算代价大，训练阶段的样本标记工作量大，成本高。

到了中期，由于目标检测的理论研究得到了发展，如诞生了主成份分析(PCA)<sup>[6]</sup>、统计学习、神经网络等算法，因此人脸检测算法也得到了改进。统计学习主要是从大量的正反样本中学习分类规则，使得分类规则能最大程度上区分出正样本，排除反例样本，在人脸检测中就是通过学习人脸与非人脸样本得到分类算法。

近几年，人脸检测的算法大放异彩，由于前期的积累使得研究方向更多，不同的方法互相结合以达到最佳效果。目前的趋势是根据不同的应用采用不同的方法，结合人脸特征模板匹配方法、神经网络、卷积神经网络、主成分分析、DPM、颜色信息等方法，以达到某个环境下的最优的人脸检测结果。

深度学习的引入为计算机视觉的各个领域如图像分类、人脸识别等带来了突破。基于卷积神经网络的人脸识别算法渐渐流行起来，并取得了一些成果。FaceBook 的 Deep Face<sup>[7]</sup>通过 CNNs (Convolutional neural networks) 提取特征，采用 3D 面部绘图功能，对面部不同区域建模，提升面部识别的精确度，目前 Deep Face 的人脸识别准确率已经达到 97.25%，此外还有 Google 的 FaceNet<sup>[8]</sup>



Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库