

学校编码: 10384

分类号_____密级_____

学 号: 24320141152406

UDC_____

厦 门 大 学

硕 士 学 位 论 文

基于深度卷积神经网络的人脸特征点检测
的研究

Research on Facial Landmark Detection Based on Deep
Convolutional Neural Networks

陈凌宇

指导教师姓名: 王备战 教授

专业名称: 软 件 工 程

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

近年来,如何提高优化人脸特征点检测技术已经成为了人脸识别技术领域中的研究热点。许多研究学者相继提出了一些用于求解各种类型的人脸识别问题的人脸特征点检测算法。但在复杂环境下,例如姿势不同、光照条件、以及遮挡因素等,传统的人脸特征点检测算法的精度会出现大幅度下降。本文在全面介绍特征点检测的理论知识以及研究现状的基础上,主要研究如何应用卷积神经网络技术设计出有效的人脸特征点检测算法。本论文的主要研究工作如下:

(1) 针对传统卷积神经网络模型在处理人脸特征点检测问题的不足之处,提出基于小滤波器的深卷积神经网络(DCNNSF)。该算法引入小滤波器思想和以拓展“网络深度”为优先的深层卷积神经网络模型,针对人脸特征点检测重新设计训练,以提高算法的有效性与适用性。本文将所提到算法 DCNNSF 用于在 ALFW 和 AFW 人脸数据集上预测 5 点人脸特征点问题,并与其他多个经典算法进行对比分析,实验结果表明:基于小滤波器的深卷积神经网络 DCNNSF 在预测人脸 5 点特征点问题上有很好的准确性和鲁棒性。

(2) 针对单个卷积神经网络模型在处理多特征点检测问题上的不足,本文在原有的 DCNNSF 基础上,提出一种从整体到局部检测的级联卷积神经网络模型 DCNNSF-CFC。DCNNSF-CFC 算法将人脸 68 个特征点检测问题分为预测外部轮廓点与人脸内部特征点的检测问题,并对人脸内部特征点进一步的分类为各个面部组件(左右眉毛、左右眼、鼻子、嘴巴),这种从整体到局部的预测方法能够有效预测 68 点人脸特征点检测问题。本文将算法 DCNNSF-CFC 同其他七个代表性经典算法在 LFPW、Helen、AFW 无约束人脸数据库中进行了实验对比分析,结果表明:本文提出的算法 DCNNSF-CFC 比起大部分经典算法有更高的精确度。

总之,本论文主要针对不同特征点检测问题,应用卷积神经网络的方法,提出两个有效的解决方法,相比于传统人脸关键点检测算法,本文提出的算法在精度上有较大的提升。这些工作在一定程度上促进了卷积神经网络在人脸对齐工作上的深入研究。

关键词: 小滤波器; 特征点检测; 卷积神经网络

Abstract

In recent years, the understanding of the importance of facial landmark detection, how to optimize facial landmark detection technology has become a research hotspot in face recognition. A lot of research scholars have proposed some face facial landmark detection algorithm used to solve various types of face recognition problems, some of them have been successfully applied in the actual project. However, in the complex environment, subject to a variety of external factors, such as posture, lighting conditions, and occlusion factors, the traditional facial landmark detection algorithm will appear a significant decline in accuracy. Based on the comprehensive introduction of the theoretical knowledge and research status of facial landmark detection, this dissertation mainly studies how to use the convolution neural network to design a significant facial landmark detection algorithm. The major researching work of this dissertation can be summarized as follows:

(1) Aiming at the shortcomings of the traditional convolution neural network model in dealing with the problem of facial landmark detection and redesigning a convolution neural network model suitable for facial landmark detection. For example, the traditional convolution neural network model is used to deal with classification problems and easy to over-fit, we proposed Deep Convolution Neural Network with Small Filter (DCNNSF). This algorithm introduces the idea of small filter and the deep convolution neural network model with the depth of "network depth", and redesigns the training for facial landmark detection to improve the effectiveness and applicability of the algorithm. In this dissertation, the proposed algorithm DCNNSF is used to predict the 5-point face feature points on the ALFW and AFW face data sets and to compare them with other classical algorithms. The experimental results show that the deep convolution neural network DCNNSF based on small filter has good accuracy and robustness in predicting face facial landmark (5 points) detection.

(2) Aiming at the problem that the single convolution neural network model is not suitable for predicting multiple feature landmarks, we propose DCNNSF with Coarse-to-fine Cascade on the basis of the original DCNNSF (DCNNSF-CFC). The

DCNNSF-CFC algorithm divides the 68 feature points detection problems into the detection problem of the external contour points and the internal feature points of the face. The facial landmarks are further classified into various facial components (left and right eyebrows, left and right eyes, Nose, mouth), this from the whole to the local prediction method can effectively predict the 68 points of facial landmark detection problems. In this dissertation, the algorithm DCNNSF-CFC is experimentally compared with the other eight representative classical algorithms in the 300-W unconstrained face database containing LFPW, Helen, AFW and IBUG. The results show that: our proposed algorithm, DCNNSF-CFC, has higher accuracy and better robustness than most classical algorithms.

Keywords: Small Filter; Facial Landmark Detection; Convolution Neural Network

目 录

| | |
|---|-----------|
| 第一章 绪论 | 1 |
| 1.1 研究背景与意义 | 1 |
| 1.2 人脸特征点检测算法的发展与研究现状 | 3 |
| 1.2.1 基于生成模型的方法 | 4 |
| 1.2.2 基于辨别模型的方法 | 6 |
| 1.2.3 国内研究现状 | 8 |
| 1.3 本文主要研究内容 | 9 |
| 1.4 本文组织架构 | 9 |
| 第二章 人脸对齐特征点检测算法的理论基础 | 11 |
| 2.1 人脸特征点估计问题介绍与描述 | 11 |
| 2.2 传统特征点检测算法 | 12 |
| 2.2.1 基本主动外观模型算法 | 12 |
| 2.2.2 基本约束局部模型算法 | 14 |
| 2.3 卷积神经网络 | 16 |
| 2.3.1 卷积神经网络的发展历史 | 16 |
| 2.3.2 卷积神经网络结构 | 18 |
| 2.4 人脸特征点检测算法的性能评价指标 | 19 |
| 2.5 本章小结 | 20 |
| 第三章 用于 5 点人脸特征点检测的卷积神经网络检测算法 | 21 |
| 3.1 引言 | 21 |
| 3.2 网络描述 | 22 |
| 3.2.1 网络深度的提升 | 22 |
| 3.2.2 小型卷积滤波器 | 23 |
| 3.2.3 基于小滤波器的深卷积神经网络具体架构 | 24 |
| 3.3 实验 | 26 |
| 3.3.1 模型训练与评估指标 | 27 |

| | |
|--------------------------------------|-----------|
| 3.3.2 测试数据集 | 27 |
| 3.3.3 实验环境与参数设置 | 28 |
| 3.3.4 实验结果与分析 | 29 |
| 3.3.5 实验展示与分析 | 33 |
| 3.4 本章小结 | 33 |
| 第四章 用于 68 点人脸特征点检测的卷积神经网络检测算法 | 35 |
| 4.1 引言 | 35 |
| 4.2 级联卷积神经网络介绍 | 35 |
| 4.2.1 级联回归 | 35 |
| 4.2.2 级联卷积神经网络 | 36 |
| 4.3 从整体到局部检测的级联卷积神经网络 | 37 |
| 4.4 网络实施细节 | 40 |
| 4.5 DCNSF-CFC 具体架构 | 42 |
| 4.5.1 人脸特征点初始化检测层 | 42 |
| 4.5.2 边界框估计层 | 42 |
| 4.5.3 人脸轮廓边界框预测 | 43 |
| 4.5.4 人脸轮廓点检测 | 45 |
| 4.5.5 人脸内部特征点边界框预测 | 46 |
| 4.5.6 人脸内部特征点检测 | 47 |
| 4.5.7 内部组件边界框估计层 | 48 |
| 4.5.8 内部组件特征点预测 | 49 |
| 4.5.9 人脸特征点输出层 | 49 |
| 4.6 实验 | 50 |
| 4.6.1 数据集 | 50 |
| 4.6.2 实验环境与参数设置 | 51 |
| 4.6.3 实验结果与分析 | 52 |
| 4.6.4 实验展示与分析 | 54 |
| 4.7 本章小结 | 56 |
| 第五章 总结与展望 | 57 |

| | |
|-------------------|----|
| 5.1 总结 | 57 |
| 5.2 展望 | 58 |
| 参考文献 | 59 |
| 硕士期间发表学术论文情况..... | 67 |
| 致谢 | 68 |

厦门大学博硕士论文摘要库

Contents

| | |
|---|-----------|
| Chapter1 Introduction | 1 |
| 1.1 Research Background and Significance..... | 1 |
| 1.2 State of Facial Landmark Detection Algorithms..... | 3 |
| 1.2.1 Generative Methods | 4 |
| 1.2.2 Discriminative Methods..... | 6 |
| 1.2.3 Status of Domestic DCNN | 8 |
| 1.3 Major Work..... | 9 |
| 1.4 Outline of Dissertation..... | 9 |
| Chapter2 Basic Theories of Facial Landmark Detection Algorithms | 11 |
| 2.1 Introduce of Facial Landmark Detection | 11 |
| 2.2 Traditional Facial Landmark Detection algorithm..... | 12 |
| 2.2.1 Basic AAM Algorithm | 12 |
| 2.2.2 Basic CLM Alogrithm..... | 14 |
| 2.3 Convolutional Neural Networks | 16 |
| 2.3.1 The Hestory of Convolutional Neural Network..... | 16 |
| 2.3.2 Convolution Neural Network Structure | 18 |
| 2.4 Performance Measures of Facial Landmark Detection algorithm..... | 19 |
| 2.5 Summary..... | 20 |
| Chapter3 CNN for 5 Point Face Feature Point Detection | 21 |
| 3.1 Introduction..... | 21 |
| 3.2 Description..... | 22 |
| 3.2.1 Increased Network Depth | 22 |
| 3.2.2 Small Convolution Filter..... | 23 |
| 3.2.3 The Structure of DCNNSF | 24 |
| 3.3 Experiments..... | 26 |
| 3.3.1 Training and Performance Measures | 27 |
| 3.3.2 Test Datas | 27 |
| 3.3.3 Parameters Settings | 28 |
| 3.3.4 Experimental Results and Analysis..... | 29 |

| | |
|---|-----------|
| 3.3.5 Experimental Presentation and Analysis..... | 33 |
| 3.4 Summary..... | 33 |
| Chapter4 CCNN for 68 Point Face Feature Point Detection..... | 35 |
| 4.1 Introduction..... | 35 |
| 4.2 Introduction of Cascade Convolution Neural Network | 35 |
| 4.2.1 Cascade Regression | 35 |
| 4.2.2 Cascade Convolution Neural Network | 36 |
| 4.3 DCNNSF with Coarse-to-fine | 37 |
| 4.4 Network description..... | 40 |
| 4.5 The structure of DCNNSF-CFC | 42 |
| 4.5.1 Initial Prediction Layer | 42 |
| 4.5.2 Bounding Box Estimation Layer | 42 |
| 4.5.3 Contour-box Layer..... | 43 |
| 4.5.4 Contour-points Layer..... | 45 |
| 4.5.5 Contour-Box Layer | 46 |
| 4.5.6 Inner-Points Layer | 47 |
| 4.5.7 Component Bounding Box Estimation Layer..... | 48 |
| 4.5.8 Component-Points Layer | 49 |
| 4.5.9 Output Layer..... | 49 |
| 4.6 Experiments..... | 50 |
| 4.6.1 Datas | 50 |
| 4.6.2 Parameters Settings..... | 51 |
| 4.6.3 Experimental Results and Analysis..... | 52 |
| 4.6.4 Experimental Presentation and Analysis..... | 54 |
| 4.7 Summary..... | 56 |
| Chapter5 Conclusions and Futrue Work | 57 |
| 5.1 Conclusions..... | 57 |
| 5.2 Future Work..... | 58 |
| References..... | 59 |
| Publications | 67 |
| Acknowledgements | 68 |

第一章 绪论

1.1 研究背景与意义

近年来，随着计算机技术的不断发展，人工智能技术取得了突破性的发展，智能社会的到来指日可待。人类作为社会的主体，如何实现人类身份的数字化将是重中之重，而人脸识别技术更是实现这一目标的关键技术。虽然人类本身可以很容易的从图片中分辨出人脸以及人脸的一些特征（例如男女、笑与哭等），但对于计算机而言，这并不是一件容易的事情。简单来说，人脸识别的主要过程可以分为三个步骤^[1]：

（1）人脸检测

根据一定算法，确认图片中的人脸数量以及人脸的大致位置，作为人脸识别的第一步，其效果直接关系到后续步骤的效果

（2）人脸特征点检测

人脸特征点主要位于面部组件周围，例如眼睛，嘴，鼻子和人脸轮廓(图 1.1)，根据输入的坐标数据和人脸检测得出的数据，分析出人脸的具体位置，并定位出特征点的具体位置。

（3）人脸特征提取与分类

根据图像信息以及人脸特征点的定位信息，提取出人脸表情等相关特征，并将其与数据库中已有特征信息进行比对、分类。

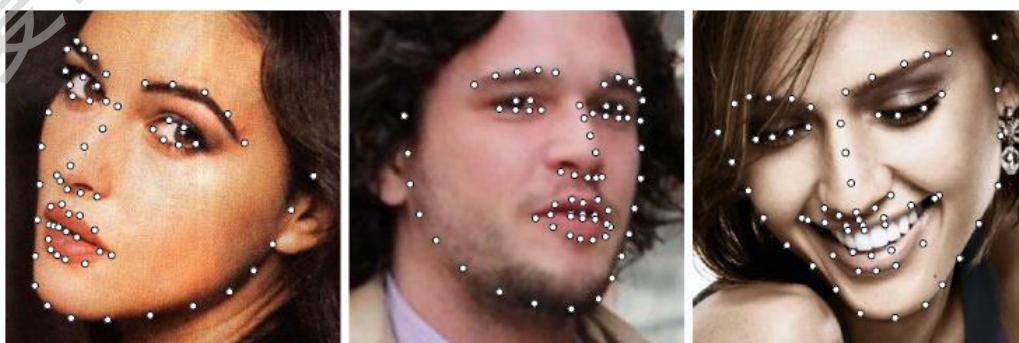


图 1-1: 来自 IBUG 数据库^[2]的 68 个手动注释点的一些示例性面部图像

在上述三个主要步骤中，虽然人脸检测通常被认为是所有人脸识别任务的起点^[2]，但人脸特征点检测可以被认为是对于从生物识别到精神状态理解的许多后续面部分析起到决定性作用的中间步骤。许多重要的任务，如面部识别，面部跟踪，面部表情识别，头部姿态估计，可以从精确的面部特征点定位中受益。下面给出一些关于人脸对齐起着重要作用的三个典型任务的细节：

(1) 人脸面部识别：面部对准被人脸识别算法广泛使用，以提高它们对姿势变化的鲁棒性。例如，在人脸对齐(Face Registration)的阶段，第一步通常是定位一些主要面部特征点并将其用作用于仿射变形的锚点、支撑点，而其他人脸识别算法，例如基于结构特征的匹配[3,4]则依靠精确的面部对准来建立要匹配的局部特征（例如，眼睛，鼻子，嘴等）之间的对应关系。

(2) 属性计算：人脸对齐也有利于面部属性计算，因为许多面部属性，如眼镜和鼻子形状与脸部的特定空间位置密切相关。在文献[5]中，六个面部点被局部化以计算定性属性和微笑，然后用于在无约束条件下的鲁棒面部验证。

(3) 表情识别：面部特征点（通常在 20-60 之间）的配置就能够可靠地表示出由表情引起的变形，并且随后的分析将揭示可能导致这种变形的特定表情的表达。许多工作如文献[6,7,8,9,10]都遵循这个想法，并从这些特征点中提取出各种特征进行表情识别。

在受限环境或不太具有挑战性的数据库中，人脸对齐的问题已被很好地解决，并且一些算法甚至实现接近人类的性能^[11,12]。然而，在无约束条件下，由于面部组件的内在动态特征（例如眼睛和嘴）或周围环境变化引起的高度面部外观变化，这项任务是非常具有挑战性且远未得到很好的解决。

一般来说，对人脸外观形状和人脸局部特征的变化有明显影响的因素有：

(1) 姿势：局部面部特征的外观在不同的拍照姿势（例如，正面，正面，上下）之间差别很大，并且一些五官（例如面部轮廓的一侧）甚至可以完全被侧脸的另一边挡在。

(2) 遮挡：对于在无约束条件下捕获的面部图像，遮挡经常发生，并且对人脸对齐带来很大挑战。例如，眼睛可能被头发，太阳镜或带有黑色框架的近视眼镜遮挡。

(3) 表情：一些局部面部特征，例如眼睛和嘴巴对各种表情的变化是敏感

的。例如，笑可能导致眼睛完全闭合，并且嘴巴的形状会极度变形。

(4) 光照：光照（光谱，光源分布和强度不同）的多样性可能会明显改变整个脸部的外观，并使某些五官的详细纹理丢失，使人脸呈现阴影部分。

(5) 其他情况：图片的模糊不清等情况。

这些问题不难从由 IBUG 数据库^[2]提供的图 1-2 看出。理想的人脸对齐系统应该对这些面部变化一方面是鲁棒的；而另一方面，尽可能有效地满足实际应用（例如，实时面部跟踪系统）的需要。



图 1-2：从左到右（每两列）无约束（IBUG [6]）人脸对齐的巨大挑战

从上所述中不难看出人脸对齐特征点检测问题的困难所在，这些问题也使其成为一个很有挑战性的问题。由于其很大的理论意义以及应用意义，如何设计一个鲁棒算法，近年来受到国内外人脸识别领域学者的广泛关注与研究。

1.2 人脸特征点检测算法的发展与研究现状

人脸对齐的问题在计算机视觉中有很长的历史，并且已经提出了大量的方法来处理它，并且具有不同程度的成功。从整体的角度来看，人脸对齐可以被称为在面部图像上搜索预定面部特征点（也称为面部形状）的问题，其通常从粗糙的初始形状开始，并且通过确定形状逐步估计直到收敛。在搜索过程中，通常使用两个不同的信息源：面部外观和形状信息。后者旨在明确地建模面部点位置之间的空间关系，以确保估计的面部点可以形成有效的面部形状。虽然一些方法不明确使用形状信息，但通常是将这两个信息源组合。

经过长期的研究，人脸特征点检测有较坚实的研究基础。从使用的方法和遵循模式识别中的基本建模原则，并大致将现有方法分为两类：基于生成模型的方

法和基于辨别模型的方法两类。

(1) 基于生成模型方法：这些方法根据人脸面部纹理形状和外观建立生成模型。它们通常将人脸对齐作为优化问题来处理，从而产生形状和外观参数，根据形状和外观参数生成给予测试面最适合的外观模型。注意，面部外观可以由整个（扭曲的）面部或由以面部点为中心的局部图像表示。

(2) 基于辨别模型方法：这些方法从面部外观直接推断目标位置。这通常通过针对每个面部点研究学习独立的局部检测器或回归器并且采用全局形状模型来对它们进行预测特征点，或者通过直接学习向量回归函数来推断整个面部形状来完成，其中形状约束被隐式编码。

1.2.1 基于生成模型的方法

通常，人脸被建模为可以在形状和外观方面变化的可变形对象，因此生成模型方法根据人脸面部纹理形状和外观建立生成模型。在生成模型中，主要可以根据生成模型建模方式的不同，可以分为基于部分的生成变形模型和主动外观模型两子类。

(1) 主动外观模型

1998年，初始主动外观模型 AAM (Active Appearance Models) 是由 Cootes 等人^[13-14]提出，是在 ASM^[15,16,17,18]的基础上，进一步对纹理进行统计建模，最后在 2001 年^[19]形成最终 AAM 模型，并将面部形状和面部纹理两个模型融合成为外观模型。主动外观模型根据拟合方式的不同又可以分类为基于回归的拟合 (Regression-based Fitting) 以及基于梯度下降的拟合 (Gradient Descent-based Fitting)。

在基于回归的拟合中，2007 年，Liu^[20]提出增强外观模型 (Boosted Appearance Model)，首次将迭代图像对准问题视为分类器能够区分正确对准（正样本）和不正确对准（负样本）的过程，该过程采用得分最大化的方法。同年，于是 Saragih 等人^[21]提出了 AAM 拟合的非线性判别方法 (A nonlinear Discriminative Approach to AAM Fitting)，针对由于 AAM 拟合的大多数方法所使用的线性更新模型的简单性，利用非线性更新模型，从而加强了算法性能。2011 年，Sauer 和 Cootes 等人^[22]提出了主动外观模型的准确的回归程序 (Accurate Regression Procedures

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 for delivery details.

厦门大学博硕士论文摘要库