

学校编码: 10384

分类号 _____ 密级 _____

学号: 24320141152402

UDC _____

廈門大學

碩 士 學 位 論 文

基于卷积神经网络的人脸属性识别研究

Research of Facial Attributes Recognition Based on
Convolutional Neural Networks

许焰瑜

指导教师: 夏侯建兵 副教授

专业名称: 计算机科学与技术

论文提交日期: 2017 年 月

论文答辩日期: 2017 年 月

学位授予日期: 2017 年 月

指导教师: _____

答辩委员会主席: _____

2017 年 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

人脸属性识别,是指对利用计算机来辨别和分析图像中关于人脸的各种属性信息,其包括生物属性识别及非生物属性识别。前者是指对个体的年龄、性别、种族等一般生物信息进行测定,对发色、眉形等特定类别进行分类,以及对微笑、愤怒等情绪/脸部表情进行识别。而后者是指识别眼影、口红等妆容以及眼镜、帽子等配饰。近年来,人脸属性识别因其在人脸验证、人脸识别和人脸检索的广泛应用,受到了越来越多的关注。传统的人脸属性识别方法,一般遵循手工提取人脸特征表示,继而训练分类器的步骤。众所周知,分类器的泛化能力取决于人脸特征表示的优劣。但考虑到手工方式只能够提取具有低层次视觉概念的人脸表征,因而无法应对由于光照条件、姿势以及遮挡等引起的复杂的面部变化,这使得在无约束条件下进行人脸属性的识别仍旧是一个极具挑战性的研究课题。而依托于大数据基础以及云计算技术的发展,借助卷积神经网络的“自学能力”来自动地学习具有高层次语义概念的人脸表征成为了可能。

综上所述,本文确立主要研究内容为基于卷积神经网络的人脸属性识别研究。首先,本文阐述了基于卷积神经网络的人脸属性识别研究的背景与意义,继而对面脸属性识别研究的现状进行了综述。其次,本文介绍了传统的人脸属性识别方法,并对其中的常见方法进行了具体介绍。再次,本文概述了人工神经网络和卷积神经网络的基础理论,并着重地介绍了其核心理念和基本结构。然后,本文提出了一个在无约束条件下用于多标签人脸属性识别的深度卷积神经网络模型,其对于复杂背景和面部变化具有鲁棒性,并获得了分别高于两种传统的人脸属性识别方法 9.638%和 7.605%的预测准确率。最后,针对与年龄和性别相关的训练数据集规模狭小的现状,本文提出了两种网络模型的调整策略。两种策略的平均预测准确率均超过了先进的年龄和性别分类方法。

关键词: 卷积神经网络; 人脸属性识别; 网络调整策略

Abstract

Facial Attributes Recognition means using computers to identify and analyze a variety of facial attribute information in images, including biological recognition and non-biological recognition. The former includes the determination of general biological factors such as age, gender, race and specific biological factors like hair color and eyebrow shape, besides facial expression for instance smile, anger and etc. Non-biological determination contains the identification of accessories like glasses and hat. Facial Attributes Recognition has raised growing attention since its wide use on Face Verification, Face Identification and Face Retrieval in recent years. The traditional methods for Facial Attributes Recognition generally follow the procedure that extracting facial characteristic artificially firstly then training the classifier. It is well known that the generalization of classifier is subject to the representation of facial characteristic. Facial Attributes Recognition under unrestraint condition is still a challenged research subject because the traditional methods mentioned earlier failed to respond to facial changes influenced by lightings, poses and occlusions resulting from that manual method is only able to extract low-level facial features limited in visual concept. However, the automatic learning of high-level facial features in semantic concept with self-learning ability of Convolutional Neural Networks becomes to be possible based on the development of Big Data and Cloud Computing technology.

In conclusion, this thesis mainly researches on the Facial Attributes Recognition based on Convolutional Neural Networks (hereinafter referred to as CNNs). Firstly, this thesis explains the background and significance of CNNs combining the summary of existing research on Facial Attributes Recognition. Secondly, this thesis introduces the traditional method of Facial Attributes Recognition and basic theory of Artificial Neural Networks and CNNs. Thirdly, this thesis proposed a deep CNNs with robustness towards complex background and facial changes for multi-label Facial

Attributes Recognition under unrestraint condition, which received two prediction accuracy rate higher than that by traditional Facial Attributes Recognition method on 9.638% and 7.605% separately. Lastly, this thesis proposed two strategies for adjusting the network architecture, which aim at status quo that the training dataset related to age and gender is narrow. The average prediction accuracy generated by two strategies mentioned before all exceeded that by most advanced method of age and gender classification.

Key Words: Convolutional Neural Networks; Facial Attributes Recognition; Strategies for Adjusting the Network

目录

第一章 绪论	1
1.1 研究背景与意义	1
1.2 研究现状	1
1.3 主要研究内容与组织结构	3
第二章 相关理论介绍	4
2.1 传统的人脸属性识别方法	4
2.1.1 LBP 特征	4
2.1.2 HOG 特征	9
2.2 人工神经网络	12
2.2.1 人工神经网络的层结构	12
2.2.2 激活函数	13
2.3 卷积神经网络	16
2.3.1 卷积神经网络的核心概念	16
2.3.2 卷积神经网络的基本结构	19
2.4 本章小结	21
第三章 基于深度卷积神经网络的多标签人脸属性识别研究	22
3.1 问题分析	22
3.1.1 遇到的问题	22
3.1.2 解决方法	23
3.2 深度卷积神经网络模型	24
3.2.1 网络模型概述	26
3.2.2 网络模型细节	26
3.3 实验与分析	29
3.3.1 实验环境	29
3.3.2 数据的准备	29

3.3.3 网络的训练	31
3.3.4 比较实验与结果分析	33
3.3.5 预训练策略	37
3.4 特征图谱的可视化.....	41
3.5 本章小结.....	44
第四章 基于卷积神经网络的年龄和性别分类研究.....	45
4.1 问题分析.....	45
4.1.1 遇到的问题	45
4.1.2 解决方法	46
4.2 基于卷积神经网络的年龄和性别分类.....	46
4.2.1 基于浅层卷积神经网络的年龄和性别分类	46
4.2.2 调整 AlexNet 模型用于年龄和性别分类	47
4.2.3 调整 VGG-Face 模型用于年龄和性别分类.....	49
4.3 实验与分析.....	50
4.3.1 实验环境	50
4.3.2 数据的准备	50
4.3.3 网络的训练	51
4.3.4 实验结果与分析	51
4.4 本章小结.....	54
第五章 总结与展望	56
5.1 总结.....	56
5.2 展望.....	56
参考文献.....	58
攻读硕士学位期间已发表的论文	62
致谢.....	63

Contents

Chapter 1 Introduction	1
1.1 Research Background and Significance.....	1
1.2 Research Status	1
1.3 Main Research Content and Organizational Structure.....	3
Chapter 2 Relevant Theories	4
2.1 Traditional Facial Attribute Recognition Method	4
2.1.1 LBP Feature	4
2.1.2 HOG Feature.....	9
2.2 Artificial Neural Network.....	12
2.2.1 Layer Structure of Artificial Neural Network	12
2.2.2 Activation Function	13
2.3 Convolutional Neural Networks	16
2.3.1 Core Concept of Convolutional Neural Networks	16
2.3.2 Basic Structure of Convolutional Neural Networks	19
2.4 Summary	21
Chapter 3 Research on Multi-label Facial Attribute Recognition	
Based on Deep Convolutional Neural Networks.....	22
3.1 Problem Analysis.....	22
3.1.1 Problems Encountered	22
3.1.2 Solution.....	23
3.2 Architecture of Deep Convolution Neural Networks.....	24
3.2.1 Overview of Network Architecture	26
3.2.2 Details of Network Architecture.....	26
3.3 Experiments and Analysis	29
3.3.1 Experimental Environments	29
3.3.2 Data Preparation	29

3.3.3 Training the Network.....	31
3.3.4 Comparative Experiments and Result Analysis	33
3.3.5 Pre-training Strategy	37
3.4 Visualization of Feature Maps	41
3.5 Summary	44
Chapter 4 Research on Age and Gender Classification Based on Convolutional Neural Networks	45
4.1 Problem Analysis.....	45
4.1.1 Problems Encountered	45
4.1.2 Solution.....	46
4.2 Age and Gender Classification Based on Convolutional Neural Networks	46
4.2.1 Age and Gender Classification Based on Shallow Convolutional Neural Networks.....	46
4.2.2 Adjust the AlexNet Model for Age and Gender Classification.....	47
4.2.3 Adjust the VGG-Face Model for Age and Gender Classification.....	49
4.3 Experiments and Analysis	50
4.3.1 Experimental Environments	50
4.3.2 Data Preparation	50
4.3.3 Training the Network.....	51
4.3.4 Experimental Results and Analysis	51
4.4 Summary	54
Chapter 5 Conclusion and Future Work.....	56
5.1 Conclusion.....	56
5.2 Future Work	56
References	58
Papers Published During Master-Degree Study.....	62
Acknowledgements	63

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

人脸是人类个体身份的重要组成部分,时刻反映着与其个体密切相关的许多重要的可感知信息。每一张面孔下是一个独特的人,其具有因不同偏好、经历、习惯、甚至个性特征形成的独特人脸属性。人脸属性识别,是指利用计算机来辨别和分析图像中关于人脸的各种属性信息。其不仅包括对诸如性别、年龄、种族等一般生物信息进行测定,也包括对发色、脸型等特定类别进行区分。此外,人脸属性识别还可以捕捉饰品(例如帽子、眼镜、耳环),妆容(例如眼影,口红)以及情绪/面部表情。近几年,人脸属性识别受到了极大的关注,取得了一系列进步,许多应用领域受益于人脸属性识别的最新进展。例如,人脸验证^{[1][2][3]},使用一系列属性分类器在每张人脸图像上独立计算,每个属性分类器可认作是人脸图像的某一视觉特征,从而将人脸验证转换成视觉特征的简单比较。此外,随着手机、相机、摄像机的普及以及诸如朋友圈、微博、Facebook、Instagram 等社交网络的出现,每天有数以亿计的图片上传网络,其中包含人脸的图像尤其吸引人们的注意。随着图像数量的指数级增长,人们迫切需要一项应对大规模数据的基于内容的人脸图像检索技术。利用包含人脸图像语义线索的人脸属性,可构造用于高效、大规模人脸检索的语义特征^{[4][5]}。

传统的人脸属性识别方法,一般为手工提取人脸特征表示,继而训练分类器。但是,这些人脸特征表示通常是具有有限的语义概念的低层次视觉特征。由于受到诸如照明、姿势(转动角度)、表情(面部形变)等不利因素的影响,人脸属性识别仍然是一个巨大的挑战。而卷积神经网络,在大数据的基础上,借助 GPU 技术的提高,其可以通过自我学习的方式自动地提取具有高级语义概念的人脸表征。

1.2 研究现状

近年来,基于机器学习的发展,与人脸属性识别相关的工作可以分为两类。

其一，利用常见的机器学习模型（浅层机器学习模型）来解决人脸属性识别问题。作为人脸属性识别的早期研究研究者之一，Kumar N 等人^[1]在每张人脸图像的局部区域手工地提取低级特征以计算视觉特征来训练属性分类器，并将人脸属性用于人脸验证，从而在 LFW 数据集上分别将当时的最佳误差率减少了 23.92% 和 26.34%。此外，Chen B C 等人^[5]提出通过稀疏表示来学习局部的人脸属性，继而借助估计人脸属性的方式来提高人脸检索的性能，该方法在 MAP 上实现了高达 43.5% 相对改进。Luo P 等人^[6]认为人脸属性与特定的人脸局部区域具有相互依赖性，并且某些人脸属性间具有高阶相关性。因此，Luo P 等人使用决策树方法分别对人脸子区域进行建模，其中每个树节点由定位该特定子区域的检测器和与该子区域中相对应的分类器组成，而属性间的相关性通过将所有决策树组成 SPN 的方式表示。考虑到人脸属性主要由人脸局部区域确定，Berg T 等人^[2]提出了以局部人脸区域替代全局人脸区域的方式来学习人脸特征表示。通常，许多着重于人脸属性识别的工作将人脸属性限制为二值分类任务。例如，人是“微笑着”和“没有微笑着”。然而，“微笑”表情几乎无法分为两类，因为人的“微笑”表情通常介于“微笑着”和“没有微笑着”之间。因此，Parikh D 等人^[7]定义了人脸相对属性的建模方式，首先学习每个人脸属性的排名函数，使用学习的排名函数来预测图像中每个属性的相对强度，然后在属性排序输出的联合空间上构建生成模型，为人脸图像提供了更丰富的语义描述。

其二，使用深度学习模型来识别多种类型的人脸属性。受益于深度学习的强大的学习能力，这一类方法通常训练和测试更多数量的人脸属性。由于增加了更多的人脸属性，所以其对人脸的描述更加的全面。文献^[8]较早地将深度学习的方法运用于人脸属性识别，提出深度属性网络模型（DAN）以获得紧凑、有识别力的特征，该模型在已标注属性标签的训练集上进行训练，不执行任何额外的属性分类，而是直接输出人脸属性。Levi G 等人^[9]表明，通过使用卷积神经网络学习抽象的人脸特征表示，可以在年龄和性别识别中获得性能的显著提高。此外，Liu Z 等人^[10]认为，与人脸属性有关的位置信息对于人脸属性的识别也很重要，因此文献^[10]级联了两个卷积神经网络，通过使用不同的预训练策略以及利用属性标签进行微调的方法，分别学习 40 个人脸属性的位置信息和特征值，实现了在无约束条件下的人脸图像的属性识别任务。

1.3 主要研究内容与组织结构

本文主要研究基于卷积神经网络的人脸属性识别。本文设计了一个深度卷积神经网络模型，用于自动地提取人脸图像的具有高级语义概念的人脸表征，从而实现在无约束条件下的端到端的多标签人脸属性识别技术。此外，面对与年龄和性别分类有关的数据集的小规模现状，本文提出了两种对网络模型的改进策略。其一，本文压缩了深度卷积神经网络模型从而提出了一个浅层卷积神经网络模型，实现了在训练数据集规模有限的情况下的年龄和性别分类技术。其二，本文证明了经典的卷积神经网络模型通过调整后可以用于年龄和性别分类。

本节余下部分将介绍本文的组织结构。

第一章，绪论。首先阐述了人脸属性识别研究的背景与意义，其次介绍了人脸属性识别研究的现状，最后明确了本文的主要研究内容。

第二章，相关理论介绍。首先，介绍了传统的人脸属性识别方法，并着重地介绍了两种传统的人脸属性识别方法中的常用特征。然后，介绍了人工神经网络和卷积神经网络的基础理论，并着重地介绍了其核心理念和基本结构。

第三章，基于深度卷积神经网络的多标签人脸属性识别研究。提出了一个用于在无约束条件下进行端对端的多标签人脸属性识别的深度卷积神经网络模型。

第四章，基于卷积神经网络的年龄和性别分类研究。提出了一个解决小规模的无约束人脸图像数据的年龄和性别分类问题的浅层卷积神经网络模型。此外，还证明了传统的深度卷积神经网络模型通过调整可以用于年龄和性别分类。

第五章，总结与展望。总结了本文的主要研究内容及其成果，并展望了下一阶段的主要研究内容。

第二章 相关理论介绍

2.1 传统的人脸属性识别方法

人脸属性识别技术，是指利用计算机来辨别和分析图像中关于人脸的各种属性信息，其试图对人脸图像进行包括诸如性别、年龄、种族等一般生物信息的分析；也对发色、脸型等特定类别进行区分；此外人脸属性识别技术还捕捉饰品（例如帽子、眼镜、耳环），妆容（例如眼影，口红）以及情绪/面部表情。

传统的人脸属性识别方法，首先检测输入图像的面部区域或面部特征点，然后从面部区域中使用手工方式来提取人脸特征表示，最后将提取的人脸表征输入属性分类器进行训练。

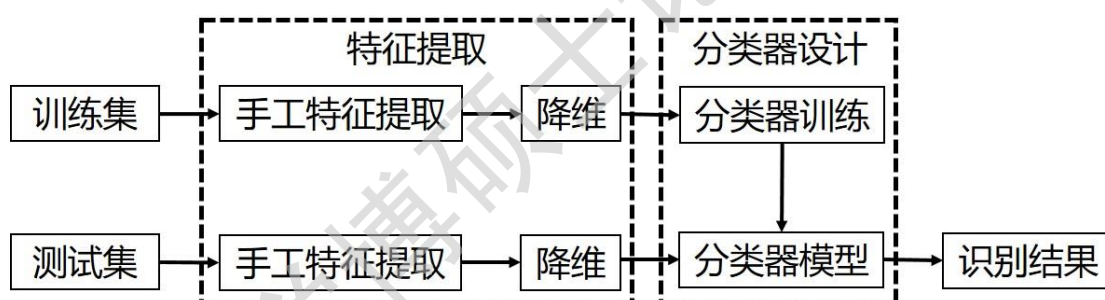


图 2-1 传统的人脸属性识别方法流程图

本节余下部分将介绍两种传统的人脸属性识别方法中用于手工地提取人脸表征的常见算子。

2.1.1 LBP 特征

(1) LBP 的基本模式

LBP (Local Binary Patterns, 局部二进制模式) 算子由文献[11]提出，被用于提取图像的局部纹理特征，即图像在像素水平上的分布关系。

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.

厦门大学博硕士学位论文摘要库