# 基于卷积神经网络对人脸性别识别算法性能研究

# 

# 摘要

随着时代的发展和变革，现在人脸检测成为了目前机器视觉领域中最重要的课题之一。基于神经网络的人脸检测与生物特征提取，不仅仅在学术界是一个热点，同时也是社会安全、生活的重要关注点。在社会的治安保障、购物支付等方方面面，准确快速的人脸检测是其最主要的技术支撑，也成为了当今各大科技公司最关心的领域。在人口密集的娱乐休闲场所，性别识别也是应用广泛的一个技术，可以有效地提供给人们更加优质、高效率的统计和服务。

近年来在人工智能领域，卷积神经网络由于其局部连接和全职共享的特点，大大的改进了传统全连接参数过多、网络层数限制的缺点，成为了当下最稳定快速的神经网络，在社会各个领域都发挥了重要的作用。

本文通过对卷积神经网络原理的研究，介绍了图像处理、人脸建模的相关技术概念，提出了一种人脸检测和性别识别协同技术，实现了对人脸识别并检测其性别的整套方案。

# ABSTRACT

In the field of artificial intelligence in recent years, the convolutional neural network has greatly improved the shortcomings of traditional full connection parameters and network layer limitation due to its local connection and full-time sharing characteristics, and has become the most stable and fast neural network at present. All areas of society have played an important role.

With the development and change of the times, face detection has become one of the most important topics in the field of machine vision. Face detection and biometric extraction based on convolutional neural networks are not only a hot topic in academia, but also It is also an important concern for social security and life. In terms of social security and shopping payment, accurate and rapid face detection is its main technical support, and it has become the most concerned area of today's major technology companies. In population gathering, entertainment and leisure venues, gender identification is also a widely used technology that can effectively bring people more high-quality and efficient statistics and services.

Based on the research of the principle of convolutional neural network, this paper introduces the related technical concepts of image processing and face modeling, and proposes a face detection and gender recognition collaboration technology, which realizes the whole face recognition and detection of gender. Program.

目录

[基于卷积神经网络对人脸性别识别算法性能研究 1](#_Toc9362767)

[摘要 2](#_Toc9362768)

[ABSTRACT 3](#_Toc9362769)

[第一章 绪论 6](#_Toc9362770)

[1.1 研究背景 6](#_Toc9362771)

[1.2 卷积神经网络研究现状 7](#_Toc9362772)

[1.3 国内外研究现状 7](#_Toc9362773)

[1.4 文章组织结构 8](#_Toc9362774)

[第二章 神经网络相关理论与技术综述 9](#_Toc9362775)

[2.1 人工神经网络技术 9](#_Toc9362776)

[2.1.1 神经元 9](#_Toc9362777)

[2.1.2 常见的人工神经网络模型 10](#_Toc9362778)

[2.1.3 神经网络的学习方式 11](#_Toc9362779)

[2.2 训练网格时所用到的基础技术原理 12](#_Toc9362780)

[2.2.1 学习准则 12](#_Toc9362781)

[2.2.2 优化算法 14](#_Toc9362782)

[2.3 设计语言与学习框架的对比与选择 15](#_Toc9362783)

[总结 16](#_Toc9362784)

[第三章 卷积神经网络原理 17](#_Toc9362785)

[3.1 前馈神经网络原理 17](#_Toc9362786)

[3.1.1 前馈神经网络的传播原理 17](#_Toc9362787)

[3.1.2 通用近似定理 18](#_Toc9362788)

[3.1.3 参数学习 18](#_Toc9362789)

[3.1.4 反向传播算法 19](#_Toc9362790)

[3.2 卷积神经网络结构和模型 20](#_Toc9362791)

[3.2.1 卷积神经网络的改进之处 20](#_Toc9362792)

[3.2.2 卷积网络中各个层的概念与原理 21](#_Toc9362793)

[3.3 卷积神经网络的参数学习 23](#_Toc9362794)

[3.4 卷积神经网络处理图片的优势 24](#_Toc9362795)

[总结 24](#_Toc9362796)

[第四章 卷积神经网络对人脸的性别识别 25](#_Toc9362797)

[4.1 人脸检测相关技术及仿真 25](#_Toc9362798)

[4.1.1 基于Haar特征的Adaboost人脸检测方法 25](#_Toc9362799)

[4.1.2 基于opencv的人脸检测及其仿真结果 27](#_Toc9362800)

[4.2 用于性别识别的卷积神经网络结构设计 28](#_Toc9362801)

[4.3 卷积神经网络训练过程 30](#_Toc9362802)

[4.4 结果仿真和分析 31](#_Toc9362803)

[总结 33](#_Toc9362804)

[第五章 总结与展望 34](#_Toc9362805)

[附录 35](#_Toc9362806)

# 第一章 绪论

## 1.1 研究背景

随着科学技术的发展，特别是在近十年里面，人工智能领域飞速发展，一系列智能算法应运而出，切实的解决了我们生活中、社会上许多问题。所谓的人工智能，即是一种赋予机器生物(人类)行为与思考的技术，如蚁群算法、粒子群算法、模拟遗传算法等等，但是现在应用范围最广阔、技术最成熟的，便是神经网络算法。

人工智能领域中有着许多的实现方式，均可以通过设计优秀的算法来达到一定的目的，而其中人机交互是其中最直观，也是其现在应用程度最高的一类实现方式。所谓的人机交互，便是通过人们对计算机输入文字、图像、声音来训练以获得优质的模型，解决人们生活中现实存在的问题，由此也诞生了一系统的学科————自然语言处理、数字图像处理、数字语音处理等。而其中图像类别的应用最为丰富，现在从事这类科研与研发的研究人员和公司也是最多的。因为图像中包含了丰富的信息，单是人脸识别这一领域，便有年龄、性别、民族等极其丰富的信息。

人脸识别技术包含多个任务子分支，首先要对一个图片进行人脸检测。如果这张图片中含有人脸，那么再进一步进行人脸图片的预处理，将图片去噪，重点突出人脸部分。之后进行关键特征提取，从而进行人脸身份识别，其中生物特征识别可包括诸多的特征，根据这些特征的不同技术也不尽相同，如性别检测、年龄检测等。

性别检测在其中是较为简单的一个二分类问题，但是在社会日常生活中，这个问题任然扮演了十分重要的角色，如下列情境下：

1）在社会治安中，公安部门在相关地方搜寻嫌疑犯的情况下，使用性别识别技术可以将人群先进行性别过滤，有助于侦查工作取得较好的效率； 2）在娱乐购物场景下，我们可以使用人脸性别检测技术，对顾客的性别进行检测，实现性别偏向的智能推荐，提高商场环境下的用户满意度；

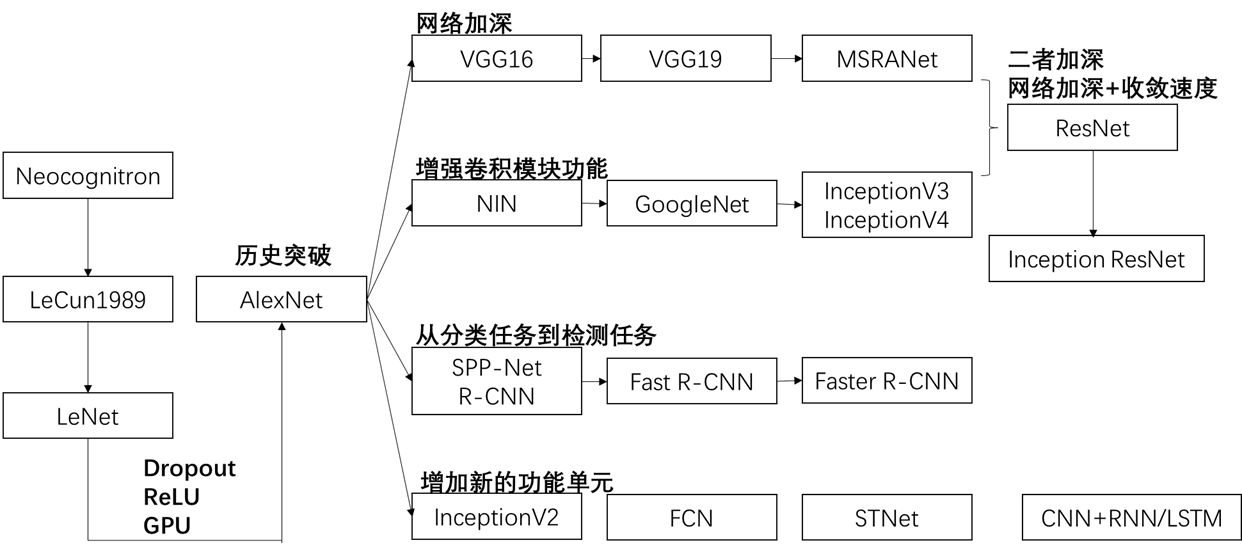
故性别识别在人类日常生活中有着很重要的实用价值。

## 1.2 卷积神经网络研究现状

上个世纪六十年代，Hubel和Wiesel的实验揭示了视觉皮质神经元如何编码图像，帮助我们理解了视觉通路中的神经元如何从视网膜上投射的光图案中提取复杂的信息以构建图像。

最初的卷积神经⽹络要追溯到上个世纪七十年代，建⽴起现代卷积⽹络学科的开创性论⽂为“Gradient-based learning applied to document recognition”[1]，作者为 Yann LeCun，Léon Bottou，Yoshua Bengio，和Patrick Haffner，其中他们使用梯度下降算法训练了我们今天所熟知的卷积神经网络，并建立了LeNet，且将其应用在了手写体字母识别上。

但是在此之后，受限于计算机的算力和广泛的数据集支持，卷积神经网络一直停留在小范围的图片处理阶段，使用价值并不高。但ILSVRC2012中Alex和Hinton提出的的AlexNet[2]开启了卷积神经网络在计算机视觉上大规模应用的先河，在ImageNet上的错误率也在人们对网络结构的一步步改进中降低至个位数百分点。

下面给出卷积神经网络的进化图示： 

之后在ILSVRC各年的竞赛中，各大研究机构提出了改进后的版本，如Inception V1、GoogLeNet等，继续降低了识别的错误率，从此卷积神经网络具有了很高的使用价值，催生了一系统机器视觉公司，为我们的生活生产带来了诸多方便。

## 1.3 国内外研究现状

基于卷积神经网络的性别识别在人脸识别领域在今年成为热点之一，国内外的学者对此进行了相关的研究，取得了不错的效果，产生了一些具有实际应用价值的成果。

国外的Gil Levi和Tal Hassner[3]设计了一种基于卷积神经网络的方案，并在Adience基准下取得了最佳的成绩；Ahmed等人直接将图像的RGB像素输入了卷积神经网络，在FRGC人脸数据路中取得了95%的识别准确率[4]；Aasma Aslam等人将图像在YCbCr空间中的Y分量经过小波变换，达到了较高的识别准确率[5]。

国内的黄勇提出了一种结合Gabor特征和基于Fisher准则的卷积神经网络的性别识别方法[6]，裴子龙提出了一种使用PSO粒子群优化CNN网络的方法[7]，陆丽提出了采用AdaBoost算法提 取脸部整体特征，主动外观模型提取局部特征（脸部几何特征），融合局部与整体特征后用支持向量机进行分类的方法[8]。

## 1.4 文章组织结构

针对性别识别这一课题，本文从基础理论知识、卷积神经网络原理、人脸检测、图片与处理、性别特征分类这几个方面进行研究，搭建出用于性别识别的神经网络，并进行数据集的测试。

其中文章安排如下：

第一章：对文章的研究背景、目的和意义进行叙述，说明了卷积神经网络的现状，以及对性别识别领域中国内外的研究现状；

第二章：对基础的理论和工具做了简单的介绍，对人工神经网络基础概念做了叙述，之后对实现语言的选取以及数据集的采用进行了说明；

第三章：对卷积神经网络的原理做了详细的介绍，从其结构到模型，以及求解方法均进行了叙述；

第四章：对人脸检测技术进行了说明，详细说明了其原理和实现方法，之后对图片做了预处理，方便投入神经网络进行学习；

第五章：设计了用于性别识别的卷积神经网络，详细介绍了训练过程中各个参数的变化，并进行了数据集测试，得出了准确率的数据；

第六章：总结了全文，指出了自己准确率的不足之处，以及改进的方法。

# 第二章 神经网络相关理论与技术综述

本章节首先对人工神经网络的技术与原理做阐述，之后对人工神经网络中较简单的BP神经网络首先进行分析，并分析了现阶段人工智能与机器学习领域常用的语言选择，且对人脸数据库的选择进行说明。

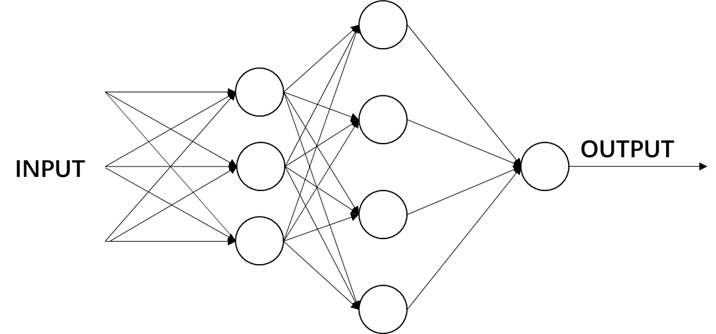
## 2.1 人工神经网络技术

人工智能在近几十年高度发展，形成了机器学习、知识推理、智能算法等庞大的学科体系,其中发展的最好、成果最为丰富的为机器学习一科。机器学习拥有传统计算机行业不具备的优势，即无需明确的编程指定便可以进行学习，且经过网络结构的设计与调参的整理工作，一部分神经网络模型显示出了十分优秀的学习水平。

根据第一章1.2节卷积神经网络的研究现状，人工神经网络为20世纪80年代出现的一种新型的智能技术，其基于生物学中神经网络构造原理，通过模拟人的大脑来构建一个智能化的信息处理系统。神经网络由大量的神经元相互连接而成，其拥有高度的非线性特点，在处理非线性逻辑操作方面显示出了强大的能力，且其作为一中仿生生智能型，拥有并行计算、高容错率的特点，可以进行智能化学习。近年内，人工神经网络在机器视觉、语音识别、自然语言处理等诸多领域发挥了重要作用，下面文章对其基本原理进行说明。

### 2.1.1 神经元

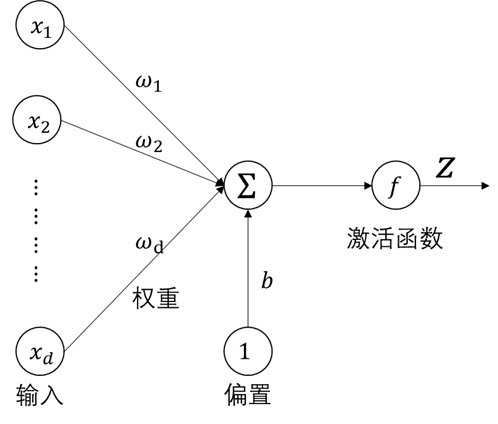
在生物学中，人脑中的最基本的处理单元为神经元。而对于人工神经网络，其模仿人脑神经元建立了多个彼此连接的结点，即人工神经网络中的神经元。如图2-1所示：



我们假设一个神经元拥有个输入，用向量表示这组输出，并用来表示输出，我们有：

其中为维的权重向量，为神经网络的偏置，非线性函数为激活函数，常用的为sigmoid函数：

综上所述，典型的神经元结构如下[9]：



### 2.1.2 常见的人工神经网络模型

单个人工神经元的功能较为简单，要模拟人脑的功能，必须使用多个神经元进行协作以完成复杂的功能。这样通过一定的连接方式或信息传递方式进行协作的神经元可以看作是一个网络，就是神经网络。

根据所处理问题的不同，目前研究者们发明了各种不同的神经网络结构，其中最常用的为以下三种：

* 前馈型神经网络

在前馈网络中，各层神经元按照接受信息的先后分成不同的组，其中每一组都可以看成一个神经层。除第一层与最后一层外，每一层网络都接受前一层的信息，并传递给下一层。整个网络的传播方向均为正向，无反向传播信息，根据图论的知识可以将其视为一个有向无环路图。

前馈型神经网络中常见的网络类型有：全连接前馈型网络、BP神经网络、径向基函数RBF网络与卷积神经网络等。

* 反馈型神经网络

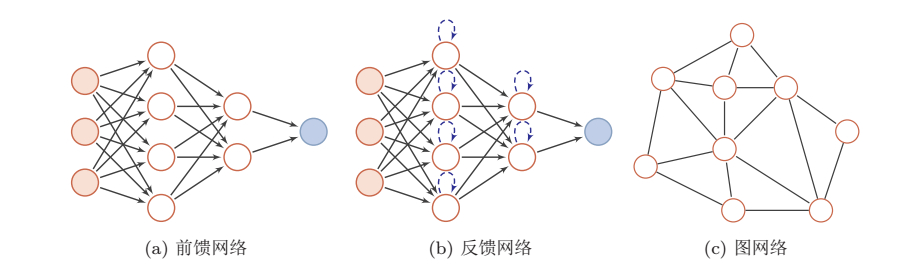
与前馈神经网络不同，反馈型神经网络不仅仅可以接受其他神经元传来的信号，也可以接受自己传来的信号。根据图论的知识可以将其视为一个无向图或者有向循环图。

反馈型神经网络常见的网络类型有：循环卷积网络、Hopfield网络、玻尔兹曼机、神经图灵机等。

* 图网络

由于前馈与反馈型的神经网络的输入都可以用向量来表示，而有一部分数据，如知识图谱、社交分子网络等机构只能以图的形式进行处理，故人们设计了基于图结构的神经网络。其节点既可以接收来自相邻结点的数据，也可以接收自己的数据。

图类型神经网络常见的网络类型有：图卷积网络、消息传递网络等。

三种网络的构造示意如下图所示： 

### 2.1.3 神经网络的学习方式

机器学习中根据不同准测的划分，可以将学习算法划分成不同的分类。比如根据激活函数的不同，我们将其分为线性模型与非线性模型；以学习准则的不同，可以分为统计方法与非统计方法；以训练样本提供的信息以及反馈方式的不同，我们可以将其分为有监督学习、无监督学习记忆强化学习。其中第三种方案也是我们最常使用的方案，下面文章对其分类原理与依据进行介绍。

* 有监督学习(Supervised Learning)

若在机器学习中我们的目标为建立样本特征与便签之间的关系：，且训练中每一个样本均有标签，我们称之为监督学习。且根据其标签类型的不同，又可以将监督学习分为回归与分类两类。

监督学习过程中，我们先将样本数据投入神经网络，得到输出后与期望值进行比较以得到误差信号，再根据误差函数对输入向量的值进行调整，在多次训练后将误差信号收敛到一个可接受的范围，便最终确定了网络的输入向量与网络结构。我们可以将其表示为：

常见的监督学习方法有：纠错学习规则等。

* 无监督学习(Unsupervised Learning)

无监督的学习是指在没有目标训练标签的训练样本中自适应的学习，以获取有价值的信息，这种算法由于没有人为的干预，可以更好的体现出自适应学习的特点。

常见的无监督学习有Hebb学习规则、竞争型学习规则等。

* 强化学习(Reinforcement Learning)

强化学习是一类根据交互来学习的机器学习算法，其根据环境的状态给出自己的判断，收到即使或延时的奖励。网络在与环境的交互中不断学习，以得到最大化的回报与奖励。

在第四章具体的卷积神经网络构造中，论文采取有监督的学习机制进行网络训练。

## 2.2 训练网格时所用到的基础技术原理

在上节文章介绍了神经网络的相关元素，在本节文章重点讨论在训练网络过程中，所使用的技术原理，为第三章卷积神经网络的原理介绍做铺垫。

### 2.2.1 学习准则

所谓的学习准则即为制定一种规则，定量的衡量我们训练模型的优劣。

网络学习过程中，假设输入空间为输出空间为，训练集满足独立同分布，若存在一模型，使得：

则称这是一个好的模型，其中为真实映射函数。

其中模型的好坏我们可以用期望风险予以衡量:

其中为真实概率分布，为损失函数。

常见的损失函数有:

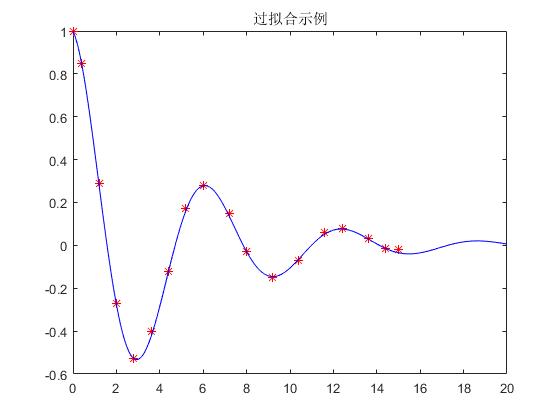
* 0-1损失函数
* 平方损失函数
* 交叉熵损失函数

由于我们实际上并不能确定真实概率分布与映射函数，所以我们无法计算期望风险，而使用经验风险进行代替，即训练集上的平均损失。

我们可以确定一组参数使得上式即经验风险最小，即经验风险最小化原则，

由大数定理可知当时，经验风险接近期望风险：

但是在实际的训练过程中，我们做不到以无限的数据进行训练，而是在数据集中拿出一部分进行训练，一部分进行验证。故经验风险最小化原则经常出现过拟合的现象，即在训练过程中准确率极高但是在非训练数据验证时正确率较低。



通常我们选择在经验最小化的基础上，引入正则项以限制模型的能力，达到避免过拟合的目的：

### 2.2.2 优化算法

神经网络模型为非凸模型，我们通常使用**梯度下降算法**来求解其局部最优值。

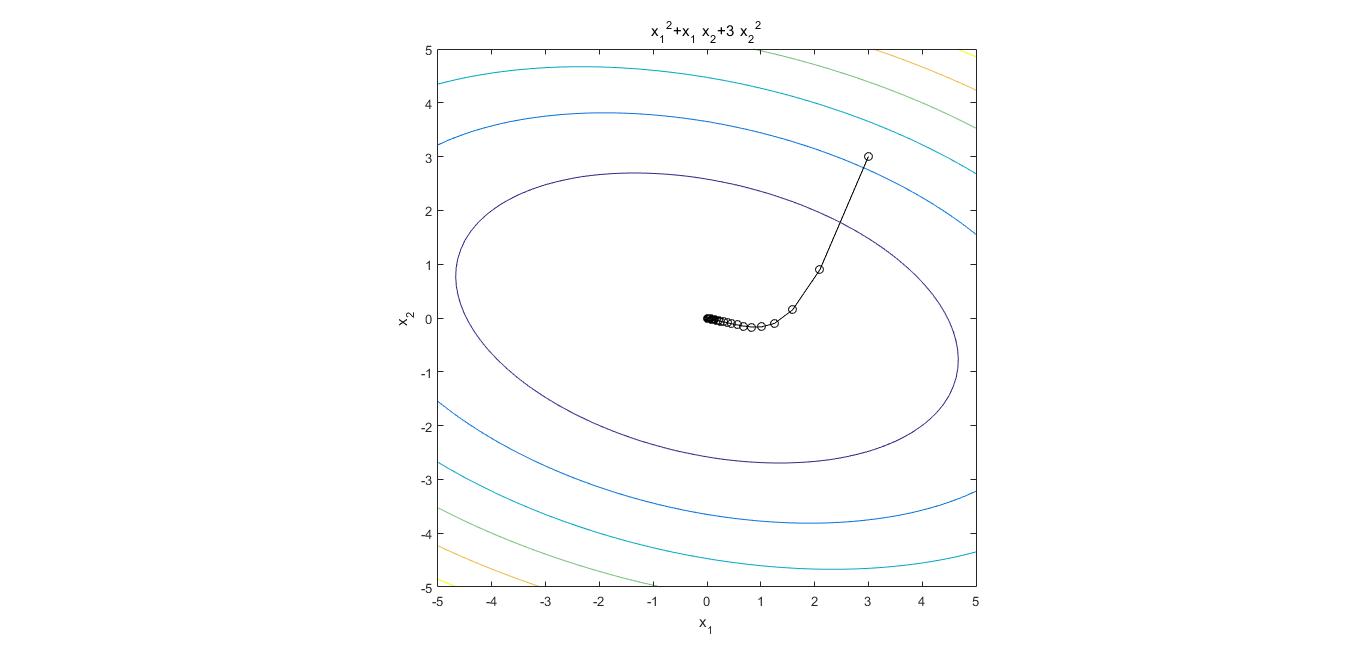
假设函数于处连续可微，则下降最快的方向是函数于处梯度方向的反方向。由泰勒公式对函数进行一阶展开可得：

取，则当且数值较小时，满足,函数呈下降趋势。

通过迭代公式我们可以生成序列，使得：

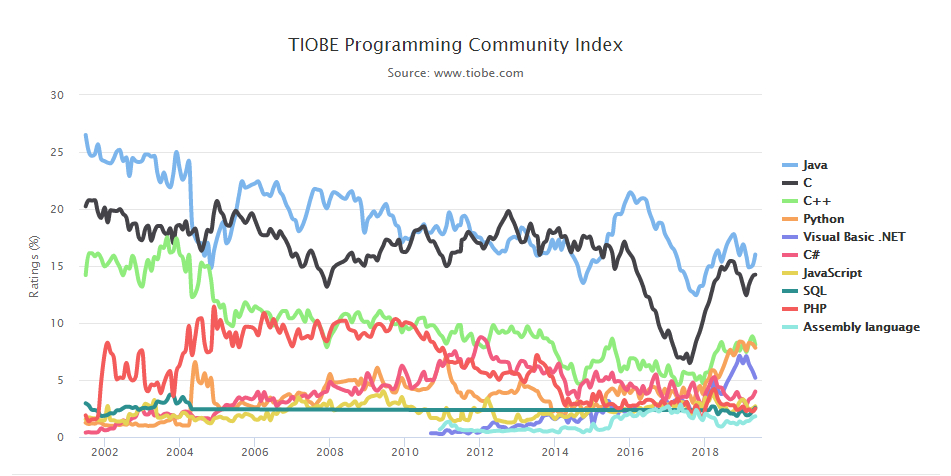
以逐渐逼近函数的最小值。

梯度下降过程如下图所示，以函数为例：



## 2.3 设计语言与学习框架的对比与选择

正如传统的工具一般，一种编程语言只是实现框架的一种工具，无论使用如中工具均可以达到自己的目标。但是，工具使用起来是否顺手，即是否可以快速、高效率的将自己的想法以清晰、有条理的代码呈现，仍是十分重要的。下面我们进行常见语言的对比以进行选择。



正如TOBIE index[10]语言指数反映的一样，在榜单前列的为JAVA、C/C++、Python等。但JAVA更偏重于框架开发，工业级别应用的制作，而C/C++更贴近底层的开发，虽然前者有其较好的机器学习开源JAR类，后者更是所有编程人员的基本功，其在人工智能领域的方便、受欢迎程度远不及Python、Matlab。但Matlab又为一闭源的平台，更偏重于工具类的调取，其强大的矩阵运算也在一定程度上被Numpy开源库所替代。

Python中携带大量的开源库文件，可以较为简单的调用以实现自己构想的功能，使编程人员将时间精力投入在AI/ML相关的算法、框架的设计上，故文章中所设相关的的脚本选择采用Python语言进行编写。

在网络架构的编写上，人们更倾向于使用框架进行实现。一方面开源框架提供了诸多的函数调用，不必要再进行重写，由于在参数的学习过程中一般通过反向传播算法进行学习，采用手工计算梯度的算法经常陷入低效、鲁棒性低的难题之中；另一方面，使用框架可以更为简单的使用CUDA进行加速，提高网络的训练速度。常用的机器学习框架有：Tensorflow以及其的包装框架keras、caffe、PyTorch等，本文选用Tensorflow框架进行开发。

## 总结

本章首先对人工神经网络中的基础概念进行了阐述，之后分别介绍了神经网络中学习方式的划分方法，以及常见的人工神经网络结构模型，最后对文章采用的编程设计语言和框架多了简单的介绍。

# 

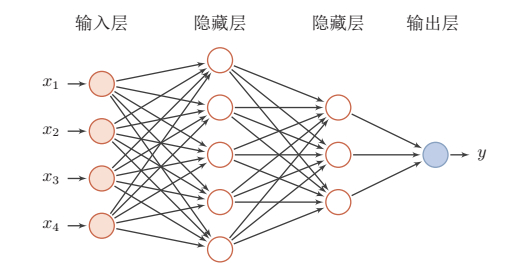
# 第三章 卷积神经网络原理

在本章，论文中重点从数学角度介绍卷积神经网络的原理，为论文的性别识别网络模型的设计工作做好理论铺垫。

## 3.1 前馈神经网络原理

### 3.1.1 前馈神经网络的传播原理

由第二章2.1.3小节可知，卷积神经网络为前馈神经网络的一种，下面论文给出前馈神经网络的通用模型，并对其中原理进行研究。下图为前馈神经网络的结构图示[9]：



论文用下面的符号表示一个前馈神经网络：

* ：神经网络的层数；
* ：第层神经元的个数；
* ：第层神经元的激活函数；
* ：第层到第层的权重矩阵；
* ：第层的偏置；
* ：第层神经元的输入；
* ：第层神经元的输出；

故前馈神经网络的信息传播可以用如下公式表示：

将作为第一层的输入，作为整个神经网络的输出，则其各个神经元层之间的状态变化图示如下：

### 3.1.2 通用近似定理

通用近似定理从数学角度说明了：常见的连续非线性函数均可以使用前馈神经网络来近似。此问题来源于数学中的KST定理，Cybenko[11]于1989年证明了通用计算定理，给与神经网络严格的理论依据。简单而言，定理说明明只要有一个隐藏层，前馈神经网络就能一致逼近任意的连续函数。

从数学角度而言，若给定和连续函数，则存在一权重矩阵,使得网络神经的输出满足：

通用近似定理指出了前馈神经网络有能力去近似任一函数，我们在训练神经网络时，并不知道真实的映射函数，一般采用正则化与经验风险最小化的原则进行学习。

### 3.1.3 参数学习

假设在学习过程中我们采用交叉熵损失函数，样本的损失函数为：

上式为标签的one-hot向量表示。

若选择监督学习，其中训练集为，其中的结构化风险函数为：

其中一般使用Frobenius范数：。

当我们有了足够的样本和以及学习准则后，网络参数便可以通过梯度下降算法来进行学习更新，其中第层的参数更新方式为：

### 3.1.4 反向传播算法

在上一节中，我们说明了使用梯度下降算法来更新参数。但是如果根据链式法则逐一计算的话，不仅工作量大而且易出错，十分低效，在此我们通常使用反向传播算法来高效的计算梯度以更新参数。

现以第层的参数权重与阈值为例，因3.1.3小节中的涉及向量矩阵微分，故我们先行对其偏导进行计算，根据链式法则有：

上市中的第二项为误差项，根据书籍[9]中的推导，偏导数、以及误差项的结果如下所示：

误差项反映了最终损失对于第层神经元的敏感程度，亦间接反映了不同层神经元的贡献程度。由上述三式可以看出，我们可以从第层的误差项得出第层的误差项，即反向传播原理。

在计算出了上述三个偏导数之后，我们可以将第层的参数值表达如下：

至此，我们可以将看基于误差反向传播算法的前馈网络的训练过程总结如下：

1，前馈计算每一层的的输入与输出;

2，反向传播每一层的误差项;

3，利用误差项计算每个神经层的偏导数更新参数;

## 3.2 卷积神经网络结构和模型

### 3.2.1 卷积神经网络的改进之处

在全连接前馈网络中，假设第层有个神经元，第层有个神经元，则权重矩阵有个参数。在特别大的时候，网络训练起来会十分麻烦，效率非常低下。

我们为解决这一问题，使用卷积来替代全连接，假设第层的输入为为第层的输出与滤波器的卷积，则有：

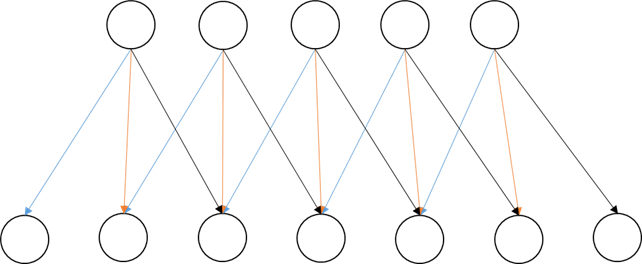
其中代表卷积运算，为可学习的权重，为可学习的偏置。根据卷积的数学性质，卷积层有两个十分重要的性质：

1)**局部连接**

模仿人类的视觉系统，在卷积神经网络中，卷积层中第层的每一个神经元只与第层的部分神经元相连接，连接数值取决于卷积核的大小，若将其设为，则连接数可以表示为。

2)**权重共享**

卷积神经网络中作为参数的卷积核对于第层的所有神经元均相同，如下图所示，相同颜色的线的权重一致。



### 3.2.2 卷积网络中各个层的概念与原理

在卷积神经网络中，通常含有三种类型的隐藏层，分别为卷积层、池化层以及全连接层。由于全连接网络与常见的前馈网络一致，故下面重点对这卷积层、池化层这两种类型的隐藏层进行介绍：

(1)卷积层

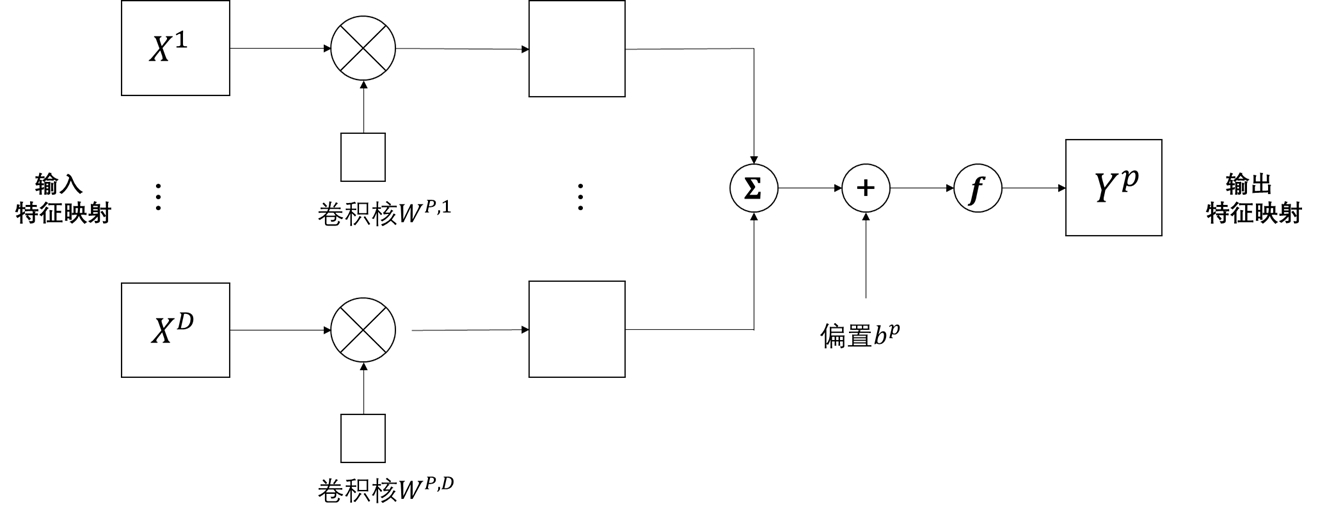
卷积层的作用为提取一个局域区域的特征，在第一层神经网络中提取一些边缘、线条，在更多层的卷积层中可以提取更加高级的特征。其中一个重要的概念为feature map：feature map为在一幅图像或者其他的feature map经过卷积提取的特征。

由于图像为二维结构，所以我们构建三层的神经层予以表示，其中大小可以表示为高度宽度深度，即为个大小的feature map。

在输入层中，若图像为彩色图像，以RGB颜色空间表示，则有一个特征映射，深度。

我们可以将卷积神经网络卷积层作用用一般性的结论予以表示，根据文献[9]所示，假设输入feature map sets为，输出feature map sets为，卷积核为，卷积层净输入为，输出为，则有：

其中，且其中每个slice矩阵为一输入feature map，；，且其中每个slice矩阵为一输出feature map，，。卷积层从输入feature map 到输出feature map 的计算过程图示如下：



(2)池化层

池化层又称汇聚层、子采样层，其作用为降低特征维度，减少参数数量，避免过拟合。

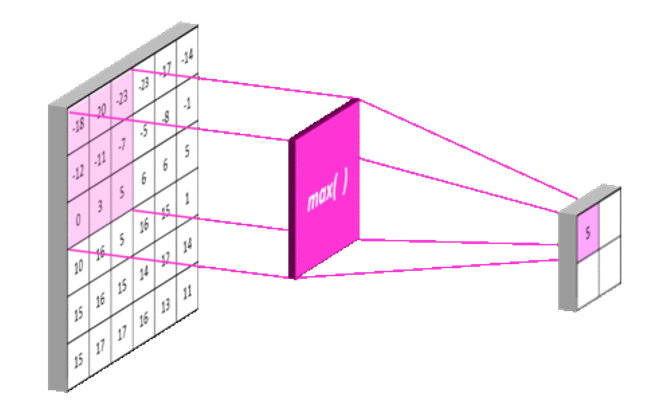
与卷积层类似，我们给出池化层的一般性结论，假设池化层的输入featurem map为，将划分为多个区域，，。我们在池化层中常用的汇聚函数有三种：

1.Maximum Pooling，即取一个区域内所有神经元的最大值。

2.Mean Pooling，即取一个区域内神经元的平均值。

3.L2 Pooling，即取一个区域内神经元的平方和的平方根。

我们以Maximun Pooling的图示[11]如下，其余的汇聚函数的映射方式也是大致如此：



## 3.3 卷积神经网络的参数学习

与常见的前馈神经网络一致，卷积神经网络的训练过程也可以使用基于梯度下降的误差反向传播算法予以完成。

如3.2小节所叙述，卷积神经网络一般含有卷积层、池化层、全连接层这三种类型的隐藏层，但是其卷积核、偏置两个重要参数只在卷积层中涉及，故我们只需要计算卷积层之中的梯度。

假设第层的输入Feature Map sets为，第层的Feature Map sets净输入为。则第层第个Feature Map净输入为：

其中损失函数关于第层卷积核、偏置的偏导数如下：

与其他前馈神经网络相同，卷积层每一层的卷积核、偏置均依赖于所在层的误差项。而第层卷积层中第个Feature Map的误差项的计算与传播更新方式如下所示：

## 3.4 卷积神经网络处理图片的优势

在处理图片这种信息的时候，传统的全连接网络会存在两个问题：

1)参数过多:如果图像大小为，即长度宽度为100，3个颜色通道(RGB)。则在第一个隐藏层即有30000个互相独立的连接，且每个连接都有自己的参数。这样致使网络训练量过大，也非常容易过拟合；

2)局部不变形特征:自然界中物体常常具有局部不变性，在旋转、平移、尺度变换中不损失语义信息，但是全连接网络常常提取不到这些性质。

而卷积神经网络与传统的全连接神经网络对比，有以下诸多优势：

(1)在神经网络中输入的图像是二维结构，而卷积神经网络中各层神经元也是二维平面，两者可以在拓扑结构上直接相连；

(2)权值共享减少了诸多的训练参数，并且简化了网络的结构，极大程度上降低了运算的时间与空间复杂度；

(3)卷积神经网络中卷积层与池化层两层结合进行特征提取的方式，对输入样本有较高的畸变容忍能力。

## 总结

本章主要讲解了卷积神经网络的原理，包括通用的前馈网络模型与卷积神经网络结构本身。

第一节通过前馈网络中传播原理、通用近似定理、参数学习方式以及反向传播算法的介绍，将前馈型神经网络的数学模型用严格的数学模型进行了推导，给出了其通用的模型原理；第二节介绍了卷积神经网络相对于传统前馈神经网络的改进之处，并着重介绍了其隐藏层中的新概念；第三节结合卷积神经网络的特点与我们在前文谈论的基于梯度下降的反向传播算法，探讨了卷积神经网络的参数学习过程；第四节介绍了卷积神经网络在进行图片处理时候的独特优势，为第四章的性别识别模型打下理论基础。

# 第四章 卷积神经网络对人脸的性别识别

在使用卷积神经网络实现性别识别过程中，大致分为两大步骤，一为人脸检测，二为性别识别。故本章在第一节简要介绍人脸监测的相关技术与原理，并且给出仿真结果；在本章的二、三、四节进行具体的性别识别程序实现。

## 4.1 人脸检测相关技术及仿真

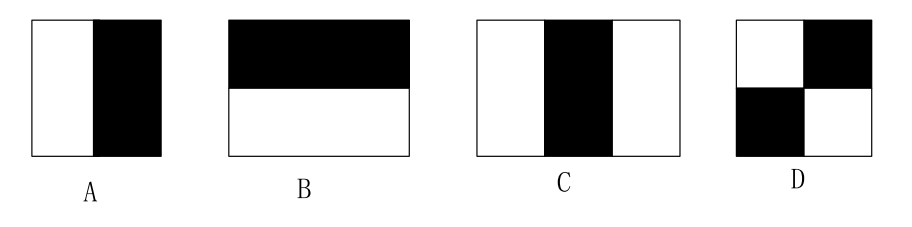
人脸检测技术是一种检测图片中是否含有(多个)人脸的算法，由于人脸在自然环境下存在着诸多的复杂情景，故采取一个具有鲁棒性的方法是十分必要的。人脸检测有基于先验知识、统计模型、人脸特征等多种方法，其中统计模型是现在的主流方向，本文采用基于统计模型的Adaboost方法来进行人脸检测。

### 4.1.1 基于Haar特征的Adaboost人脸检测方法

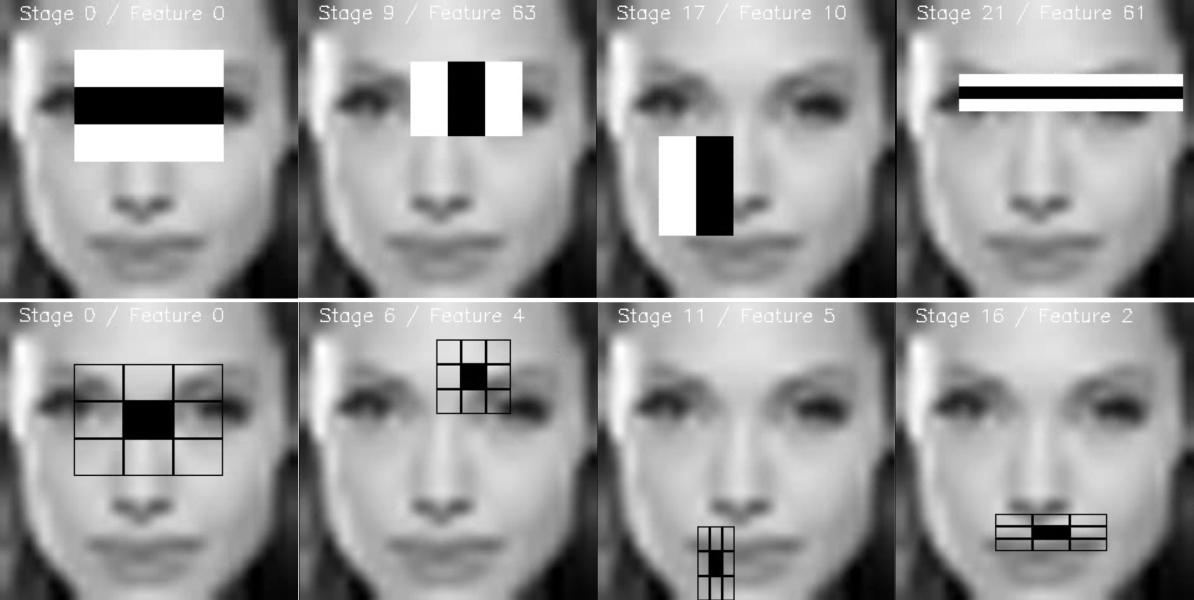
本方法来源于2001年Viola与Jones于文献[12]中提出的方法，整个人脸检测方法包括大致三个步骤，分别为Haar-like特征提取、基于Adaboost的分类器迭代训练。下面对其进行简单的介绍：

(1)Haar-like特征极其计算

Haar-ike定义为图像之中相邻像素和的差，在本文中可知它的特征值反映的是人脸图像的灰度变化情况，其特征大致一般分为四种类型：边缘特征、线性特征、对角线特征以及中心特征。[13]如下图所示：



在人脸之中，Haar特征可以用下图[15]清晰地体现出来：



在图像预处理形成灰度图像之后，我们可以将Haar-like特征值定义为图像之中一块区域内黑色像素之和与白色像素之和的差值，并且可以用积分图来提高其计算速度。

假设为点左上角所有像素之和，为于积分图中的值，为于这点的像素值，为点纵坐标像素值之和。故我们有Haar积分图算法如下：

(2)Adaboost算法训练分类器

Adaboost(Adaptive Boosting)即为自适应增强算法，其原理为迭代运算，由多个弱级联器组成强级联器进行检测。其理论基础为：若每一个弱级联器的分类水平都优于随机猜测的水平，将弱级联器的数量增加至无穷大，便可以得到误差率接近于零的强级联器。Adaboost算法伪代码如下所示：

1.给定，其中

2.初始化权重

3.

* 使用初始化权重进行训练
* 得到弱分类器及其错误率
* 其中
* 计算
* 更新权重

4.输出强级联器:

Adaboost算法即为：将Haar特征向量输入强级联器中，结果为正，即为人脸。

### 4.1.2 基于opencv的人脸检测及其仿真结果

opencv为一开源机器视觉库，包含着一系列的C函数与C++类，实现了一系列常用的机器视觉算法；并且提供了Matlab、Python、Ruby等其他编程语言的接口。优秀的算法实现与开源精神，使得opencv成为了在机器视觉中使用最为广泛应用之一。

在此我们参考opencv中Adaboost算法源码，实现人脸检测。下以弱分类器源码[16]为例，进行解读：

1. **typedef** **struct** CvCARTClassifier
2. {
3. CV\_CLASSIFIER\_FIELDS()
4. **int**  count;                     //弱分类器数量
5. **int**\*compidx;                    //对应特征
6. CvTHaarFeature\* feature;        //特征值
7. CvFastHaarFeature\* fastfeature;
8. **float**\*threshold;                //阈值
9. **int**\*left;                       // 左子节点
10. **int**\*right;                      // 右子节点
11. **float**\*val;                      //输出
12. }CvCARTClassifier;

此结构体定义的CvCARTClassifier即构造了一个二叉决策树，体现在机器学习中即为一个预测模型，指向目标对象的属性。按照Adaboost算法，多个弱分类器按照二叉决策树模型连接，构成强分类器。

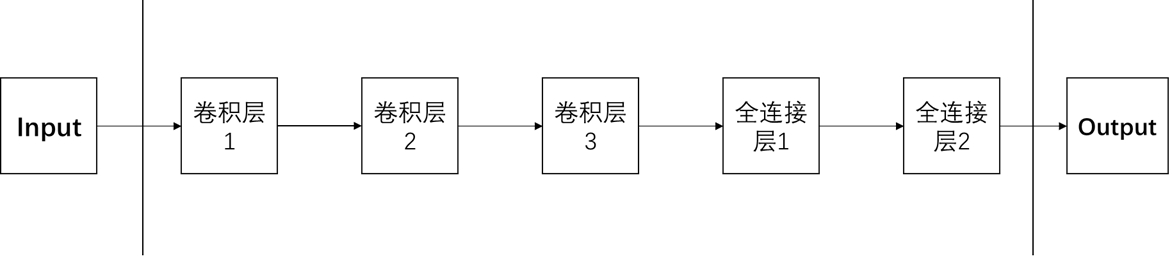
在此我们利用opencv中训练完成的强分类器[19]进行人脸检测，人脸检测结果如下：



我们可发现在这张著名的照片中，与会者29人有27人正确被检测出，正确率为，达到了较高的水平。在本章的后续章节，将在本节讨论的基础上进行后续的展开工作。

## 4.2 用于性别识别的卷积神经网络结构设计

对于论文的性别预测而言，为一个二分类问题。本文中采取Gil Levi、Tal Hassner于文献[3]中提出的网络结构，其为AlexNet卷积神经网络的修改版本，且仅含三个卷积层和两个具有较少神经元的全连接层。具体的结构如下所示：



这种具有较少结构的设计可以有效的避免过拟合现象，并且在训练速度较为快速。下面结合模型的python示意代码具体说明模型各层参数与功能：

(1)**卷积层1**(Convolutional Layer 1)

1. conv1 = convolution2d(images, 96, [7,7],[4,4],padding='VALID',...)
2. pool1 = max\_pool2d(conv1, 3, 2,...)
3. norm1 = tf.nn.local\_response\_normalization(pool1, ...)

第一层卷积层中采用96个卷积核，且每个卷积核的参数大小为:；补0策略为只执行有效卷积，对边界数据不处理；激活函数默认为ReLU，即:

其池化层执行Maximum Pooling，且池化的大小为，strides为2。且其后添加了一个局部响应归一化层，目的为提高网络的泛化能力，其公式[17]为：

(2)**卷积层2**(Convolutional Layer 2)

1. conv2 = convolution2d(norm1, 256, [5, 5], [1, 1], padding='SAME', ...)
2. pool2 = max\_pool2d(conv2, 3, 2, ...)
3. norm2 = tf.nn.local\_response\_normalization(pool2, ...)

第二层卷积层中选择使用256个卷积核，且每个卷积核的参数大小为；补0策略为保留边界处的卷积结果。

池化层与局部响应归一化层参数与第一层卷积层的参数一致。

(3)**卷积层3**(Convolutional Layer 3)

1. conv3 = convolution2d(norm2, 384, [3, 3], [1, 1],padding='SAME',...)
2. pool3 = max\_pool2d(conv3, 3, 2,...)
3. flat = tf.reshape(pool3, [-1, 384\*6\*6], name='reshape')

第三层卷积层中选择使用384个卷积核，且每个卷积核的参数大小为；补0策略为保留边界处的卷积结果。

池化层与第一层卷积层的参数一致；而第三层的tf.reshape函数为一变换函数，作用为将张量tensor变换为参数shape的形式。

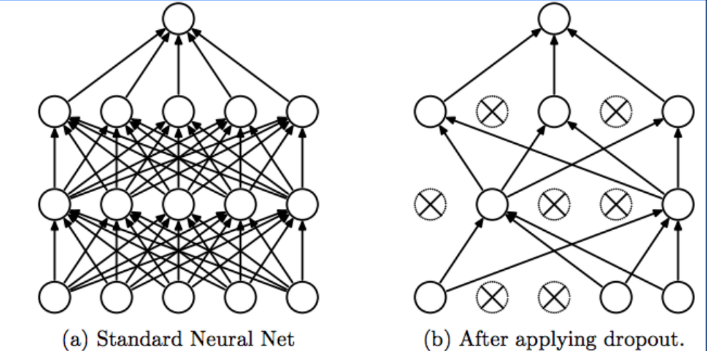
(4)**全连接层1**(Full connected Layer 1)

1. full1 = fully\_connected(flat, 512,...)
2. drop1 = tf.nn.dropout(full1, pkeep, name='drop1')

第一层全连接层神经元个数为512。

dropout层为训练样本模型时，训练样本过少而防止过拟合加入的一层神经元。其作用为在训练更新参数的过程中，按照一定概率随机断开输入神经元，其原理可以下面公式予以表示：

Dropout层网络结构图示如下：



(5)**全连接层2**(Full connected Layer 2)

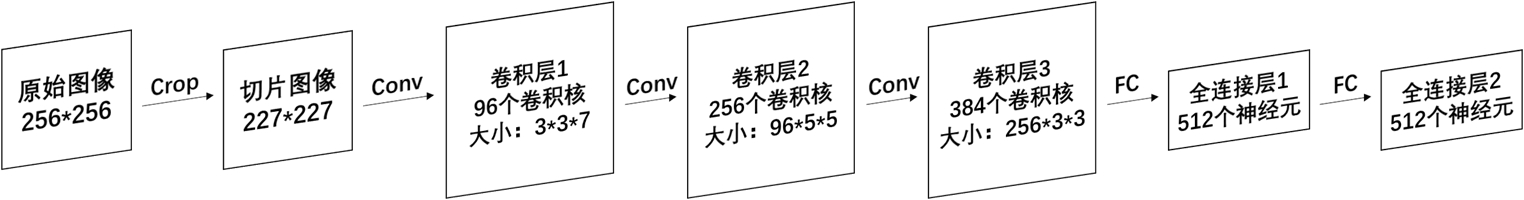
1. full2 = fully\_connected(drop1, 512,...)
2. drop2 = tf.nn.dropout(full2, pkeep, name='drop2')

第二层全连接层，其神经元个数也是512个。

(6)**输出层**(Output)

对于本文，性别识别为二分类问题，故其输出神经元个数为2。

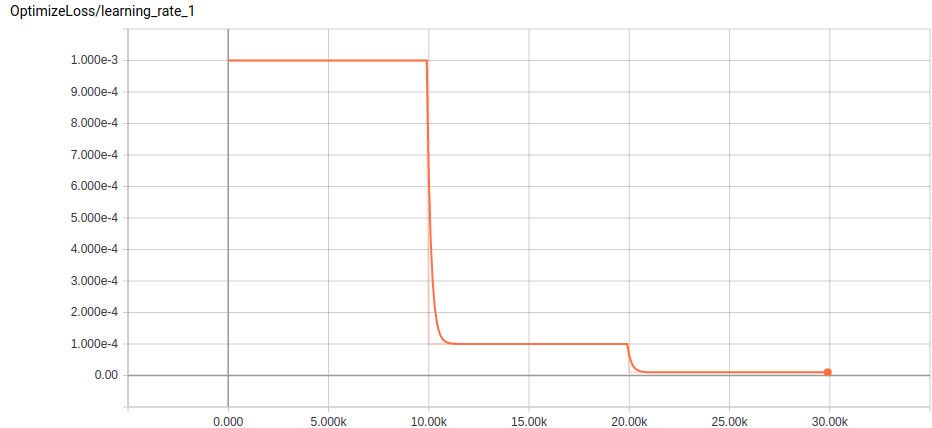
综上所述，其网络结构可以由下图表示：



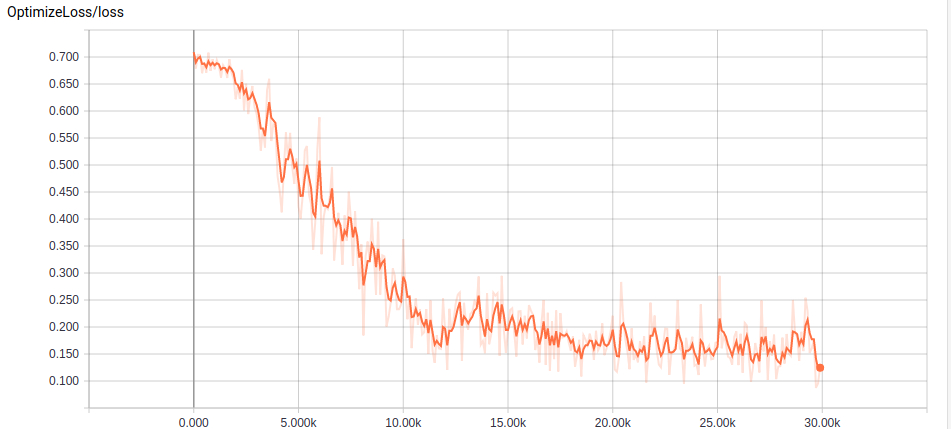
## 4.3 卷积神经网络训练过程

对4.2节提出的神经网络进行训练，论文使用的人脸图像数据为Adience data[18]，论文训练次数设置为30000次，具体训练参数变化如下图所示：

* 学习率



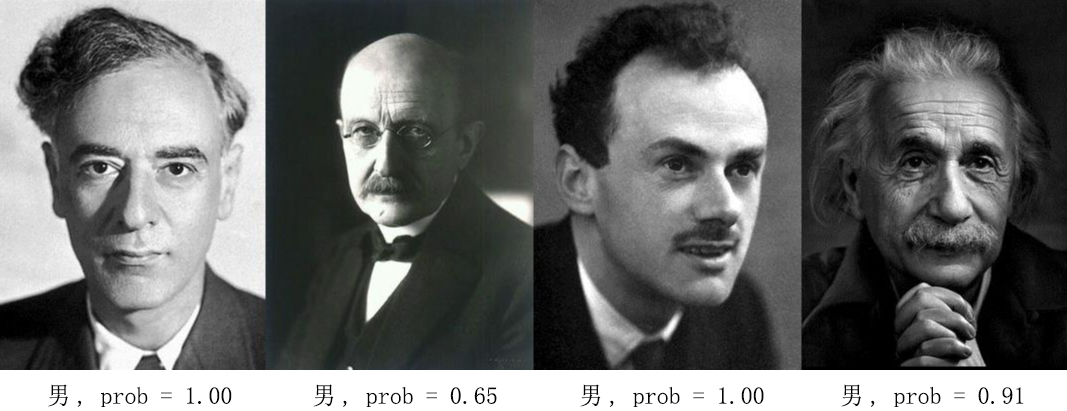
* 误差



可以观察到当学习次数到达20000次时，学习率已经趋于0，而最终的误差也在0.15左右震荡，网络训练在30000次时已经趋于稳定。

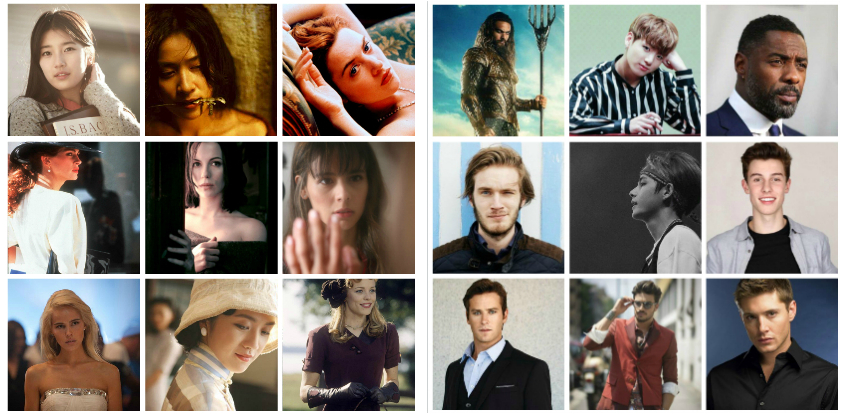
## 4.4 结果仿真和分析

在此我们使用由4.3节训练完成的模型对人脸图像进行性别测试，以验证模型的优劣。如下图实例：

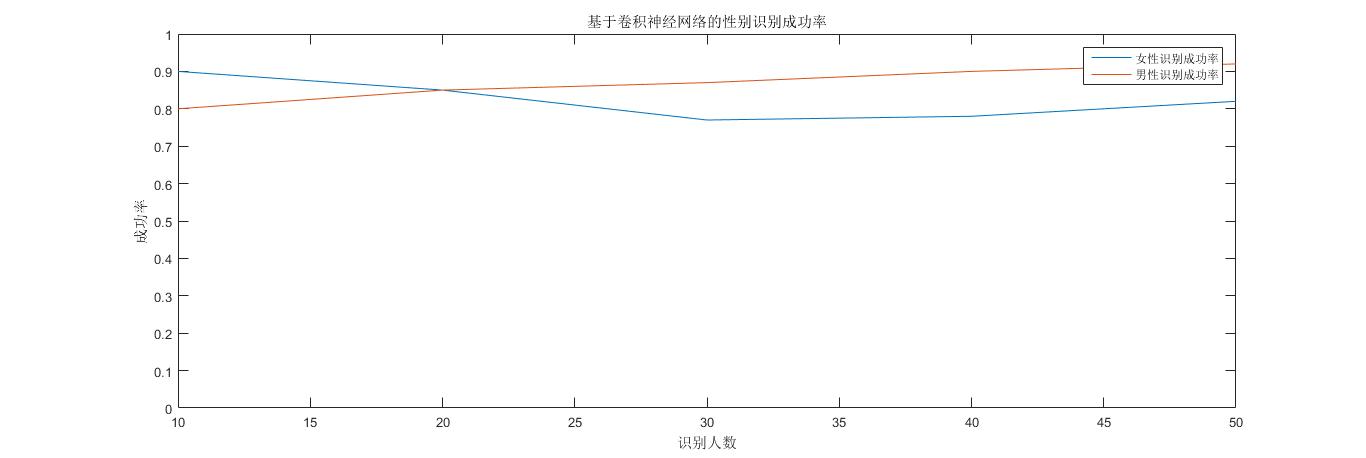


我们可以看出，在以上的若干测试图片中，三张图片变现良好，一张图片虽然给出正确性别，但是其概率并不高。在证明此模型为有效模型的基础上，我们制作一个含有大量图片的数据集去检测其识别率。

在此我们选用在网络抓取的100张图片进行验证，其中男女各为50张。其中一部分如下所示：



经过测试，我们得出正确率曲线如下：



我们可以发现，男性的识别成功率在左右，而女性的识别成功率在左右，均取得了不错的结果。这个结果说明了本文中基于AlexNet改进的卷积神经网络模型较为成功，可以在合理的精度内有效地完成解决性别识别的任务。

## 总结

在本章内，文章首先研究了基于Haar特征与Adaboost算法的人脸检测方法，之后通过opencv的相关开源代码对之进行了仿真。在本章的第二小节重点对卷积神经网络的网络架构——配合实际应用的python代码——进行了说明，三、四小节对网络的训练过程以及训练之后模型的结果进行了仿真与说明。测试图片为在网络中抓取的明星图片，在大多数为正面照、少部分为侧面照的情况下，均取得了在识别率之上(男性图片达到)的结果，说明了此卷积神经网络在解决性别识别的问题上是构建成功的。

# 

# 第五章 总结与展望

在人工智能浪潮席卷世界的今天，机器学习作为其中最为重要的一个部分，在目前显示出了最为强大的能力，切实地解决了各行各业中无数之前难以解决的问题。而机器视觉在机器学习中占有重要的一个位置，近年受到了全球科技工作者的注意。人脸监测与性别识别作为机器视觉中的一项重要内容，对社会治安、用户服务等方方面面都有着重要的研究价值。

文章使用了基于AlexNet卷积神经网络的修改版本，较少的神经层使得网络不易陷入过拟合，而在最后的仿真结果测试中也证明了这一点，性别的识别率达到了较高的水平。结果说明，在正面照的情况下，该网络模型可以取得较高的正确识别率，但是在侧面照、环境噪音较多的情况下，识别率较低，但是也显著高于。

文章的不足以及可以改进的地方在于：

1)对环境噪音较多环境的图片识别率的提升。对于这种情况，本文的网络层数便过于稀少，如何合理的安排与设计网络层数、结构，是核心的问题；

2)对于训练速度的改进。如论文[7]中提到的PSO-CNN算法，其可以加速梯度下降算法以达到节省时间的效果，但是其作者训练次数在50左右，能否在上万的量级中也达到加速的效果需要进一步的检验；

# 附录

[1]:Gradient-based learning applied to document recognition

[2]:ImageNet Classification with Deep Convolutional Neural Networks

[3]:Age and Gender Classification using Convolutional Neural Networks

[4]:Training Hierarchical Feed-Forward Visual Recognition Models Using Transfer Learning from Pseudo-Tasks

[5]:Wavelet-based convolutional neural networks for gender classification

[6]:基于人脸图像的性别识别研究

[7]:基于卷积神经网络的人脸性别识别研究

[8]:基于人脸图像的性别识别与年龄估计研究

[9]:神经网络与机器学习

[10]:[TIOBE Index for May 2019](https://www.tiobe.com/tiobe-index/)

[11]:Machine Learning Notebook:A resource for machine learning with Python

[12]:[Mathematical physics applications in present-day image processing](https://mathoverflow.net/questions/268418/mathematical-physics-applications-in-present-day-image-processing)

[13]:Rapid Object Detection using a Boosted Cascade of Simple Features.Viola&Jones

[14]:基于人脸图像的性别识别研究

[15]:[Cascade Classifier Training](https://docs.opencv.org/master/dc/d88/tutorial_traincascade.html)

[16]:[\_cvhaartraining.h](https://github.com/npinto/opencv/blob/master/apps/haartraining/_cvhaartraining.h)

[17]:ImageNet Classification with Deep Convolutional Neural Networks

[18]:[Unfiltered faces for gender and age classification](https://talhassner.github.io/home/projects/Adience/Adience-data.html)

[19]:[haarcascade\_frontalface\_default.xml](https://github.com/opencv/opencv/blob/master/data/haarcascades/haarcascade_frontalface_default.xml)