

# 基于卷积神经网络的多人脸识别

朱婷婷 周丽

四川大学锦城学院 计算机与软件学院 四川 成都 611731

DOI: 10.18686/jsjxt.v1i3.1267

**【摘要】**随着科技的不断发展,在这个大数据时代,人脸识别的应用越来越广泛。基于深度学习卷积神经网络的技术也越来越被认可和运用。卷积神经网络训练人脸识别不需要进行复杂的特征提取,只需要利用 OpenCV 库检测人脸,运用一个可行的网络模型进行自动训练进而得到一个较好的识别效果。

**【关键词】**人脸识别 OpenCv 卷积神经网络

## 1 引言

人脸识别,被大量运用到各个领域,例如人脸考勤、人脸社交、人脸支付,疑犯追踪、交通打卡等……人脸识别在宏观上分为两类,一类是人脸检测,就是识别出图像中人脸的位置和大小,一类是人脸识别,即与已有的收集到的图片库中的图进行对比然后判断是否是同一个人,最后完成身份验证。

人脸识别算法多种多样,对于传统方法则依赖于人工设计的特征以及机器学习的一些算法,比如提取图像中的边缘、纹理、线条、边界等特征,依据这些特征再进行下一步的处理,这样的处理效率较低。因此,人工提取特征以及机器学习技术的方法已经

被使用非常广泛的卷积神经网络训练数据所取代。

深度学习神经网络算法的特征与优点是可以对一个较为庞大的数据集进行训练,进而学习到其表面特征。除开人脸识别,CNN 还被广泛运用到面部表情识别、物体识别、年龄分析等方面。

## 2 数据收集

人脸识别的采集是通过摄像头等电子设备进行的。本文将采用的是班级的 33 个人,每个人使用 opencv+haar 特征通过摄像头提取 10 张彩色图片,每张照片都有不同的表情,不同的动作,不同位置,不同光线等。放在同一个目录下统一命名 name\_00.TIF~name\_09.TIF。



图 1 人脸图片数据收集

## 3 数据预处理

### 3.1 基于 opencv 的人脸检测

Opencv 是一个跨平台的计算机视觉库,可以实现大量的有关图像的处理<sup>[1]</sup>,在人脸检测方面也有很高的正确率。

利用 opencv 的 Face Detector 检测人脸,从视频中获取人脸,对于人脸检测部分,可以使用 Haar 特

征分类器,在这些分类器中有用于眼镜,头部,嘴巴,鼻子等地方。在本文中采用的是 haarcascade\_frontalface\_default.xml 分类器。

### 3.2 修改图片尺寸与数据增强

#### 3.2.1 修改尺寸

使用哈尔特征统一提取照片中人脸部分,并修改保存为 60 \* 60 大小的照片。

### 3.2.2 数据增强

对于一些图片其光照太弱导致照片里面的人脸和背景融为一体,它们的像素点就会很接近,还有人脸本身的变形等都会对计算机识别造成障碍,因此对原始图片进行增强在一定程度上提升任务的准确率。使用 opencv 库中的 flip() 函数对原始样本进行水平翻转、使用 contrast\_brightness\_demo() 函数调整亮度、并使数据扩充到 1320 张图片。

### 3.3 打标签并划分训练集、验证集、测试集

#### 3.3.1 打标签

使用 opencv 库中的 imread() 函数得到图片的矩阵信息,使用 pandas 库中的 get\_dummies() 函数为数据进行 one-hot 编码设置标签,最后将标签与图片信息相互对应整合转为 DataFrame 格式方便后续处理。

#### 3.3.2 划分训练集

将已经处理好的数据取 70% 作为训练集、20% 作为测试集、10% 作为验证集。为了使得每次训练的样本不同以及防止过拟合,使用 shuffle 函数对训练集进行打乱。

## 4 浅谈 CNN

### 4.1 CNN 简介

卷积神经网络(Convolutional Neural Network, CNN)<sup>[2]</sup>,是一类包含了卷积计算并且具有深度结构的神经网络,是深度学习的算法之一。

### 4.2 主要结构

#### 4.2.1 输入层

在图像处理的卷积神经网络中,它的输入层一般为一张图片的像素矩阵,设计成三个维度,例如  $28 \times 28 \times 3$ ,28 为图片大小,3 为图片的深度或者通道,彩色图片一般为 R、G、B 三个通道。

#### 4.2.2 卷积层

一般用来做特征处理。Filter,即卷积核或卷积层过滤器,是和输入图片深度一样的三维数据。设置步长,利用 filter 在输入数据的宽度和高度上滑动,然后计算每个通道的 filter 与输入数据任意一处的内积。在卷积核的输出后还会跟一个激活函数,是为了将非线性引入网络中。

#### 4.3.3 池化层

压缩矩阵大小,减少了计算资源耗费,同时可以有效控制过拟合提高一定泛化性。

#### 4.4.4 全连接层

将所有的局部部分通过全连接层将其连接起来,做最后的分类得出结果。

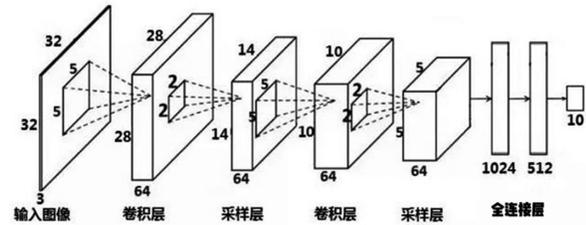


图 2 卷积神经网络结构图

## 5 搭建卷积网络

batch 值:初始训练批次以及每个批次训练样本数。

输入层:  $x$  为图片信息,  $y$  为分类信息。

卷积层:利用  $3 \times 3$  大小的卷积核去不断卷积图片,得到特征图,从第二层开始 filter 大小变为了前一层卷积后的图片深度,利用 relu 激活函数将非线性引入网络中。最后一层卷积步长由 1 变为 2。

池化层:矩阵信息大小缩小一半。

全连接层:图片在进行填充边界的情况下计算其大小与深度,在全连接层设置多个神经元,将卷积得到的特征图传入,最后一层输出层为 33 个神经元。

Softmax 层:使用 softmax 分类器将输出结果映射到 0-1 范围内,判定得分最高的为正确分类。

基本网络框架搭建好得到结果后,需要计算其损失值,然后将误差通过反向传播,采用梯度下降的方式不断更新权值,不断迭代训练网络,每次迭代都需要打乱训练集加强模型的泛化能力,将得到的分类与验证集对比输出验证正确率,最后再通过测试集输出测试正确率。

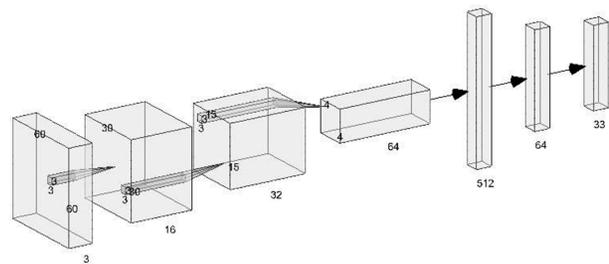


图 3 卷积过程结构图

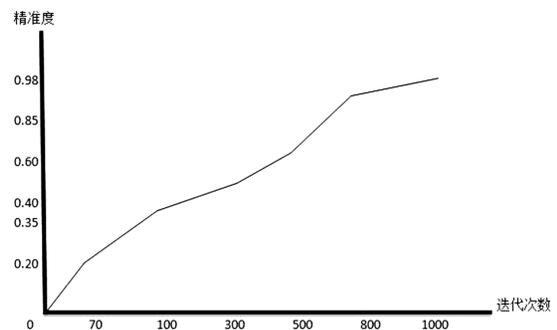


图 4 精准度变化曲线图

```
epoch: 930 valid_acc:0.98
epoch: 936 valid_acc:0.98
epoch: 937 valid_acc:0.98
epoch: 938 valid_acc:0.98
epoch: 939 valid_acc:0.98
epoch: 940 valid_acc:0.98

epoch: 941 valid_acc:0.98
epoch: 942 valid_acc:0.98
epoch: 943 valid_acc:0.98
epoch: 944 valid_acc:0.98
```

图 5 网络模型最后精度图

## 6 优化网络

### 6.1 优化训练批次

采用的是小批量训练数据, batch 初始设置大小, 因为训练集数量不算太大, 因此在设置 batch 时可适当减小, 但也不能过于太小以免训练速度过慢, 20, 30, 50, 100 等等大小都可以通过训练去选择一个最优, 本文训练出最好的批次为 100。

### 6.2 优化卷积核大小、个数以及卷积层数

在卷积层训练时, 可自行设置卷积核的长宽, 本文最后采用  $3 \times 3$  大小为较优, 在第一层卷积中卷积核的深度应保持和图片通道一致为 3, 使用 filter 个数自行调整。该网络第一层使用 16 个 filter, 第二层为 32 个 filter, 第三层为 64 个 filter。还可以通过增加或减少卷积层数去调整网络, 本文采用三层。

### 6.3 优化全连接层神经元个数以及层数

在全连接层中设置神经元个数以及全连接层数, 可适当增加提高精度, 过大同样会导致训练速度较慢, 该网络第一层设置了 512 个神经元, 第二层设置了 64 个神经元

### 6.4 优化学习率

在反向传播梯度下降中设置其学习率即学习步长, 一般选择 0.1、0.01、0.001 等, 过小会导致训练速度慢, 过大会导致跨过最优值, 通过调参找到最优, 该网络调试出最优的学习率为 0.001

### 6.5 优化迭代次数

在训练网络中, 需要多次迭代计算精度, 迭代次数需要自己设置, 太小会导致精度小, 选择一个较大的次数优化网络模型, 本文采用了 2000 次迭代。

## 【参考文献】

- [1]郭永平. 基于机器视觉缺陷检测的实践教学平台研究[J]. 宝鸡文理学院学报(自然科学版), 2017, 37(04): 59-61+67.
- [2]周明非, 汪西莉, 王磊, 陈粉. 高分辨卫星图像卷积神经网络分类模型[J]. 中国图象图形学报, 2017, 22(07): 996-1007.

## 7 实时人脸识别

在本项目中, 要做到实时的人脸识别, 还需要利用 OpenCv 库的方法实时获取需要识别的人脸数据。

已经通过卷积神经网络训练出了一个在验证集上准确率高达 98% 的模型, 现在需要用这个训练好的模型去实时预测人脸。所以首先需要用 OpenCv 调用摄像头, 捕获实时的视频流, 再使用 OpenCv 的人脸检测模型在视频窗口检测出人脸并截取人脸图片。将截取到的人脸图片做一些处理, 比如将尺寸修改为网络需要的大小。然后将这张图片传入网络, 使用模型的参数进行识别, 得到一个关于每个人的类别的概率, 最后返回最大的概率对应的人的名字。

再通过给图片标注的函数将人的名字显示在视频窗口中。

经过多次测试, 实时识别的准确率还是非常高的, 这也证明了所建立的模型对图片特征的学习能力还是很强的, 可以做到精确识别每一个人。

## 8 结束语

随着技术发展越来越快, 在这个大数据的时代, 人脸识别凭借着高效识别精确度给社会带来的极大的便捷。基于神经网络的多人脸识别有较良好的提取特征能力与较强学习能力, 通过使用 OpenCv 库去检测人脸, 再使用抓捕到的人脸数据去训练网络, 从而实现人脸识别。但由于图片数量太少, 训练样本得到的精度不够高, 后期精度会随着增加每个类别样本数量而增加。在训练网络过程中有时还会遇到过拟合的问题, 出现这个问题的原因在于数据量太少、模型太复杂, 可通过获取更多数据、减少网络层数、神经元个数等限制拟合能力来泛化性能。