

**2023 届毕业设计(论文)**

**课 题 名 称: 基于深度学习的笑脸检测**

**课 题 名 称（英文）：Smiling Face Detection**

**based Deep Learning**

**学 生 姓 名： 王姿尹 学 号： 2019082417**

**专 业 名 称： 计算机科学与技术**

**指 导 教 师： 张雄涛 职 称： 副教授**

**所 在 学 院： 信息工程学院**

**完 成 日 期： 2023 年 3 月 26 日**

**教务处制表**

**基于深度学习的笑脸检测**

**摘要：**近年来，深度学习技术在各个领域都得到了广泛的应用和发展，其中包括人脸识别、机器翻译等。深度学习技术的一个重要目标是实现对图像或视频中目标的识别和分类。而笑脸作为一种特殊的图像或视频流，经常出现在各种应用场景中，例如人脸识别、人机交互等。因此，如何实现对笑脸的识别和检测一直是一个有意义的课题。传统的笑脸检测通常需要人工收集和处理大量图像或视频数据，存在时间和人力成本高、准确性低等问题。深度学习技术可以自动提取图像中的特征，并且具有较高的准确性和鲁棒性。因此，基于深度学习的笑脸识别检测应运而生。本文提出了一种基于深度学习的笑脸识别检测系统，该系统采用GENKI-4K表情数据集和卷积神经网络（CNN）的方法，实现对人脸图像中笑脸的识别和检测。

首先，论文分析了当前笑脸识别的背景和现状，然后详细阐述了该系统采用的卷积神经网络的理论知识，接下来，论述了系统框架、软件、笑脸数据集、模块实现。实验结果表明，该系统能够有效识别人脸图像中的笑脸，并能检测出人脸图像中是否存在笑脸。

**关键词：**深度学习，卷积神经网络，笑脸检测

**Smiling Face Detection based Deep Learning**

**Abstract**:In recent years, deep learning techniques have been widely used and developed in various fields, including face recognition and machine translation. An important goal of deep learning techniques is to achieve recognition and classification of targets in images or videos. And smiling faces, as a special image or video stream, often appear in various application scenarios, such as face recognition, human-computer interaction, etc. Therefore, it has been an interesting topic to realize the recognition and detection of smiling faces. Traditional smiling face detection usually requires manual collection and processing of large amounts of image or video data, which has problems such as high time and labor costs and low accuracy. Deep learning techniques can automatically extract features from images and have high accuracy and robustness. Therefore, smiling face detection based on deep learning comes into being. In this paper, we propose a deep learning-based smiling face detection system, which uses GENKI-4K expression dataset and convolutional neural network (CNN) approach to achieve recognition and detection of smiling faces in face images.

First, the paper analyzes the current background and status of smiling face recognition, then elaborates the theoretical knowledge of the convolutional neural network used in the system, next, discusses the system framework, software, the smiling face dataset, and module implementation. The experimental results show that the system can effectively recognize smiling faces in face images and detect the presence or absence of smiling faces in face images.

**Keywords:**deep learning, smiley face recognition, convolutional neural network

**目 录**

[第一章 绪论 1](#_Toc2144)

[1.1 研究背景和目的 1](#_Toc9366)

[1.2 深度学习和人脸识别研究的现状 2](#_Toc382)

[1.2.1 深度学习的研究 2](#_Toc6876)

[1.2.2 基于深度学习的人脸识别的研究 2](#_Toc12405)

[1.3本文章节内容安排 3](#_Toc13499)

[第二章 卷积神经网络概述 4](#_Toc23078)

[2.1输入层 4](#_Toc3809)

[2.2卷积层 4](#_Toc3749)

[2.3池化层 5](#_Toc20141)

[2.4全连接层 6](#_Toc17679)

[2.5输出层 6](#_Toc8493)

[2.6其他层 7](#_Toc26354)

[2.6.1 Flatten层 7](#_Toc32627)

[2.6.2 随机失活层 7](#_Toc20466)

[2.7激活函数 7](#_Toc18974)

[第三章 笑脸检测系统的设计与应用 9](#_Toc11153)

[3.1 实验平台的配置 9](#_Toc7810)

[3.2 系统框架 9](#_Toc10822)

[3.3 系统软件介绍 9](#_Toc6625)

[3.4笑脸数据集 10](#_Toc4170)

[3.4.1笑脸数据库建立 10](#_Toc9100)

[3.4.2笑脸数据预处理 11](#_Toc29632)

[3.5基于CNN的笑脸检测 13](#_Toc23781)

[3.6 系统模块设计 13](#_Toc27938)

[3.6.1 笑脸识别模块 13](#_Toc3738)

[3.6.2 系统界面模块 14](#_Toc21531)

[3.7实验结果 14](#_Toc25744)

[3.7.1混淆矩阵 14](#_Toc13368)

[3.7.2分类报告 15](#_Toc5722)

[3.7.3笑脸检测效果图 16](#_Toc16405)

[第四章 总结与展望 18](#_Toc18480)

[致谢 19](#_Toc32493)

[参考文献 20](#_Toc4282)

# 第一章 绪论

## **1.1 研究背景和目的**

中国信息通研究院和深度学习技术及应用国家工程研究中心在2022世界数字经济峰会“AI引领未来”主题下，共同发布《深度学习平台发展报告》。可见深度学习是目前人工智能较为前沿的研究热点，也是当今学科发展趋势。

深度学习，以生物科学为基础，模仿人类大脑的工作机理，使机器执行可以像人类那样对所学习的数据进行分析，拥有很强的特征学习和特征表达能力。很多技术突破都离不开深度学习，尤其在计算机视觉、自然语言处理、生物信息等领域都有着广泛运用，一些具体应用如下：

（1）人脸识别门禁。人脸识别门禁系统是一种应用于智能门禁系统的人工智能技术。它可以通过收集和分析大量人脸图像数据，识别门禁系统中人员的身份，并在设备上进行自动开关门等操作。同时，人脸识别门禁系统还可以根据需要进行人工验证，以保证数据的准确性和可靠性。

（2）智能聊天机器人。智能聊天机器人是一种能够自动学习和理解用户问题并给出答案的机器人。在深度学习领域，智能聊天机器人可以用于用户提问、自然语言处理、知识库管理、对话式交互等方面。智能聊天机器人可以通过对话式交互技术来与用户进行交互。这可以提高用户体验，减少交互时间和成本。另外还有很多其他类型的智能聊天机器人，例如智能问答机器人、智能写作机器人等。

（3）疾病风险评估。疾病风险评估利用机器学习和计算智能技术，对疾病的风险进行评估和预测。深度学习技术可以用于训练特定的神经网络来模拟大脑中神经元的活动，从而预测某种疾病的发病风险。这种方法可以用于预测某种特定疾病或病症的发病率或患病率，例如，通过对X光照片进行学习，可以预测肺癌的发病率。

美国一位心理学教授Albert Mehrabian[1]研究显示，人与人之间有效的沟通，是根据自己的语言、声音和表情决定的，其中，表情等肢体动作占55%，另外的语言占7%，声音占38%。由此可知，表情在人类交流过程中起着传递信息的重要作用，是传递人类情绪状态的最直接方式之一，是语言交流的重要补充。人类面部表情含有大量的情感信息，是人们相互理解情感的重要途径。

人脑具有识别不同面部表情的先天能力，但开发机器自动识别人类面部表情的技术是一项重大挑战。人机交互中，如果计算机能够理解人类的情绪，自主地适应环境并对人类的情绪作出适当的反应，那么人类和计算机之间的关系就会发生巨大的变化。因此面部表情识别作为一种新兴的技术，在深度学习中具有广泛的应用前景。而笑脸是自身喜悦的外在表现，通过对笑脸表情进行研究，可以更好地理解人们的情感状态，并在日常生活中与人们更好地沟通。也可以帮助识别出人们的心理状态，例如当人们感到开心时，会表现得更加明显。通过对笑脸进行识别，可以帮助研究人员更好地理解人们的情感认知能力和思维水平。

**1.2 深度学习和人脸识别研究的现状**

### **1.2.1 深度学习的研究**

反向传播算法，也称BP算法，是由Werbos[2]等多位研究者在提出的。而目前最有效的BP算法则是Geoffrey Hinton[3]于1986年等人在此基础上明确提出。BP算法的基本思想是通过将输入数据映射到输出数据，并计算误差，然后调整网络参数，使误差最小化。在传统神经网络正向传播的基础上，增加了误差的反向传播过程，完美的解决了非线性分类问题。然而，计算机技术有限，出现的 "梯度损失 "问题，严重限制了BP算法的发展。2006年，Geoffrey Hinton，Yoshua Bengio，Yann LeCun [4]采用无监督的学习方式，然后一层一层进行训练，再利用有监督的反向传播算法对其进行优化，从而克服了“梯度消失”问题。这一学习方法的提出，立即在学术圈引起了轩然大波，被越来越多人所关注，并在更多领域开始进行广泛研究。2012年，Hinton和他的学生Alex Krizhevsky在ILSVRC竞赛中，采用深度学习模型AlexNet[5]。AlexNet使用ReLU激活函数，防止了“梯度消失”问题，大大提高了运算效率，将top-5的错误率降低到15.3%，一举夺冠。2016年3月，谷歌DeepMind研发的AlphaGo[6]在围棋人机对弈中战胜了韩国顶尖棋手李世乭。后一年，又以3:0的战绩战胜中国顶尖棋手柯洁。

### **1.2.2 基于深度学习的人脸识别的研究**

《Integrating Deep Learning Algorithms to Overcome Challenges in Big Data Analytics》[7]一书中提到大量基于深度学习的数据分析。可见当下深度学习已经成为一个热门话题，而人脸检测和识别技术的研究则是当前深度学习中计算机视觉领域的一个热门课题。2019年，中国科学院大学刘小沣[8]基于增强学习对图片集进行了人脸识别。其不足之处在于样本是静态图片，而人的表情却时刻变化的。于是2020年，西南科技大学刘成攀[9]便针对动态人脸进行研究。此观点的提出，有利于我们对动态人脸视频进行实时识别，但是这未考虑到人脸被遮挡方面的实际情况。2022年，上海电子信息职业技术学院贾璐[10]便基于关键点定位提取面部属性特征和文本特征，研究出可以很好识别佩戴口罩的人脸图像，在所有实验中的图像结构相似度均达到了0.9以上。同年贵州民族大学赵雨青[11]更是对局部遮挡人脸图像进行了修复，引入注意力机制，提高了修复质量。中国电子科技集团第三十八研究所顾立春[12]针对不同角度，利用仿射变换进行人脸姿态的引导转换，实现人脸识别。随着深度学习的不断深入，国内外已经渐渐有了一些研究成果，比如通过人脸识别来进行人员出入记录或是在线考勤等。但是，我们可以很容易地看到，随着整个社会的智能化水平的提升，人与人之间的互动也越来越频繁。如果一台电脑，可以像人那样，去了解和表达自己的情感，更好地为人服务，那将会是一场从根本上的变革。

**1.2.3 基于深度学习和人脸识别的表情识别的研究**

人脸表情识别技术是将某一静止或运动的视频序列中的某一特定的表情状态进行分离，通过计算机来判断识别对象的心理，从根本上改变人与电脑之间的关系，使其能够更好地进行人机互动。因此，人脸表情识别在深度学习领域有着巨大的潜力。从上个世纪七十年代起，很多学者都在进行脸部表情分析，但大多是从心理层面进行的。随着科技发展，人脸表情识别技术逐渐提高，同时深度学习的应用，也让人们对其的关注度越来越高。2016年卢官明等人[13]利用卷积神经网络，对样本图像的表情进行分类识别。2019年南京信息工程大学王珂[14]，能够提取出视频中人脸表情特征，但未完全达到实时的效果。上海师范大学信息与机电工程学院李传江等人[15]于2020年针对这一问题，在研究中采用点阵红外夜视摄像机进行数据采集，实现实时视频与照片的拍摄。同年，上海师范大学张秋秋[16]构建了大量儿童笑脸数据库,并指出利用数据增强技术，将数据库图片又一次进行扩增。2022年，Oday A等人[17]更是将笑脸进行了更为精确地分类，比如微笑、大笑等，利用eBagging协助训练，可达到98%的识别精度。

## **1.3本文章节内容安排**

本文共有四个章节，每个章节的内容安排如下:

第一章，首先阐述了当前热门深度学习的研究背景与表情识别的研究意义，然后介绍了深度学习和人脸识别的国内外研究现状，最后说明本文的内容安排。

第二章，讲解了卷积神经网络的组成结构，为之后的研究内容提供理论支持。

第三章，笑脸检测系统的设计与应用。在此基础上，给出了整个系统的总体结构，软件，同时选择合适的笑脸数据集，并对数据集进行预处理，接着给出了各子模块的详细说明，最后对系统进行了测试与应用。

第四章，总结与展望。对全文的研究成果和方法进行了总结。

# 第二章 卷积神经网络概述

卷积神经网络（Convolutional Neural Networks, CNN）是一种机器学习模型，它通过对图像进行特征提取来实现图像分类、目标检测等任务，利用反向传播来训练模型。同时，卷积神经网络能够简化参数，并且可以做到逐层提炼图像数据的局部数据特征，是一种非常有潜力的深度学习网络。与传统的分类算法不同，卷积层与全连接层不单单只把图像数据映射至低维空间，而是利用卷积层与全连接层来提取图像特征，然后通过多个卷积层和全连接层进行组合，实现更准确的分类识别。

卷积神经网络的基本架构包括输入层、卷积层、池化层、全连接层、输出层五个部分。输入层获取图像后，卷积层对图像进行卷积运算，得到图像的特征量，池化层再一次特征提取，然后把特征量映射到一个低维空间上，全连接层再来实现对特征的处理和分类，最后，输出层会将提取到的特征值通过一个标准的神经网络进行分类和预测。卷积神经网络结构示例如图2-1所示。本小节将对这些重要部分进行介绍。

|  |
| --- |
| wps  图2-1 CNN结构图 |

## **2.1输入层**

输入层（Input Layer）就是把数据输入卷积神经网络完成特征提取，最终得到想要的结果。在卷积神经网络的输入层中，数据可以是一维的，比如语音，也可以是多维的，比如图像。输入层的输入量保持了图像自身的结构。

## **2.2卷积层**

卷积层是卷积神经网络的最基本层之一，也是卷积神经网络中最重要的一层。它将输入数据映射到一个低维空间中，然后使用卷积层提取特征。卷积层通常采用在数据区域的四周进行旋转、缩放和翻转等操作来实现。

卷积层的具体架构如下：

1. 卷积核由多个小尺寸的像素点组成。卷积核通常采用圆形或方形的形状，以便可以更好地提取图像中的特征。

2. 参数设置：卷积层在模型中发挥着重要作用，它可以通过调整卷积核的尺寸、深度、卷积核的方向等参数来改变模型的输出。在设置卷积层时，需要注意卷积层与激活函数的组合，以确保模型的性能和效率。

①卷积核

卷积核是网络模型中经常出现的一种结构，它是一种可以将输入的图像特征从多个像素点中提取出来的函数。卷积核将图像数据经过卷积处理，然后将它们从图像中提取出来，从而实现特征提取的功能。

卷积神经网络中的卷积核通常由多个小尺寸的像素点组成，它们通过全连接层进行连接，然后将特征值以一定的规则映射到图像中的特定区域。例如，一个卷积核可以被映射到图像中的一个点或者多个点，然后在这个点上进行池化或全连接。卷积核在卷积神经网络中的作用非常重要，在图像识别任务中，卷积核可以提取图像中的一些简单特征，如局部均值、局部方差等。通过使用卷积核，网络模型可以大大简化，从而提高模型性能和效率。

②卷积运算

卷积运算是指在卷积神经网络中，输入数据和参数之间的变换关系式通过卷积运算得到的结果。也是一种互相关运算，不涉及旋转180度。卷积运算有多种常用的方法和公式，其中最常见的有转置卷积、扩张卷积等。

总的来说，卷积层是卷积神经网络中最重要的一种转换层操作之一，它在模型中发挥着重要作用。它可以简化计算和降低计算成本，同时也可以提高模型的识别准确率和效率。

## **2.3池化层**

池化层（Pooling Layer)，也叫作子采样或或下采样层，即经过卷积处理后，对输出的图像再一次进行特征提取，保留关键信息，减少不必要的计算[18]。通常会分别作用于每个输入的特征并减小其空间大小。它可以将低维的特征向量投影到一个低维空间中，从而简化计算和降低计算成本，提升卷积神经网络的性能和效率。

在卷积神经网络中，池化层通常用于将低维特征向量转化为高层语义信息，使其更适合分类任务。池化层常用的操作如：

1. 范围选取：在卷积神经网络中，池化层可以将高维特征向量投影到一个更低的空间，即池化空间。池化空间通常是一个三维或四维的空间，也就是说，它可以将某个特征向量的取值范围缩小到一个三维或四维的空间。

2. 最大值限制：在卷积神经网络中，池化层还可以设置最大值限制，以确保池化特征向量不会因为过大而失真。如果池化空间的尺寸超过了最大值，那么卷积神经网络输出的特征向量中就会出现一些白噪声，导致特征向量中的信息损失。

3. 缩放池化：卷积神经网络的池化层还可以设置缩放池化操作，以改变卷积神经网络的特征图尺寸。在缩放池化操作中，池化层将特征图缩小到一个二维或三维的尺寸。这种方式可以使卷积神经网络的特征图更容易计算和拟合原始数据。

4. 转换层：在卷积神经网络中，池化层还可以将高维的特征向量转换为低维的特征向量。

常见的池化转换层如下：

1. 全连接层到空间池化：卷积网络的全连接层通过在输入空间中进行卷积操作来提取特征。但是，如果特征图过于复杂或者维度太高，那么该层就会损失很多信息。这个时候就需要进行空间池化操作，将特征图转换为一个较低维度的空间，从而减少信息损失。

2. 通道池化：通道池化是卷积神经网络中常见的一种转换层操作。它将低维特征向量映射到一个二维或三维的通道中，以增加特征向量的视觉吸引力。此外，通道池化还可以减少模型计算复杂度，因为它可以将多个通道转换为一个通道进行计算。

3. 上采样：卷积网络的上采样层也可以使用通道池化操作。这种操作通过在输入数据中加入通道数相同但方向相反的通道来减少模型的计算量。

4. 反池化层：在卷积神经网络中，反池化层可以用来调整卷积神经网络的特征图大小和方向。它可以使特征向量在某个特定位置聚集，以减少模型中梯度扩散、提高计算速度。

## **2.4全连接层**

全连接层（dense layer）是卷积网络的核心层，它连接输入和输出节点（例如图像或语音），以提取原始数据的特征。该层的主要作用是将原始数据通过卷积操作从特征空间中采样，然后使用全连接层进行特征提取。

常见的全连接层网络如下：

1. ReLU激活函数：ReLU函数通常在全连接层作为激活函数，原因是它比其他激活函数更容易解释，计算复杂度低，并且能够更好地处理非线性问题。

2. 全连接层：卷积网络的全连接层通常由3到7个神经元组成。它们接收从特征空间采样的特征向量，并将其转换为参数。这个层的作用是将原始数据的特征表示与参数进行组合，得到一个新的特征向量。

3. 多分类器层是全连接层网络中最常见的应用之一。该层通过在输入数据中加入多个标签，来识别不同类别的样本，从而提高分类准确率。

4. 残差连接是卷积网络的全连接层中常用的一种连接方式。它通过在输入数据中加入残差单元来解决网络退化问题。残差单元是由低层输出和中间层输入通过残差连接形成的单元，以减少特征图中的梯度问题[19]。

5. 上采样：卷积神经网络的全连接层还可以通过上采样操作，即将特征向量沿某一特定方向进行采样来进一步减少特征图中的梯度问题。

6. 剪枝是卷积神经网络中的一个常用操作，它通过将过多参数进行压缩而减少了神经网络的规模[20]。通过剪枝，卷积网络可以得到更清晰、特征更显著，性能也更好。

总的来说，在卷积神经网络中，全连接层是其最关键的一环，对特征提取过程中的非线性组合起着至关重要的作用。它主要用于特征图尺寸较小、维度较低、不易过多地拟合原始数据等特点。

## **2.5输出层**

由于卷积神经网络的输出层在全连接层之后，所以它的结构和工作方式与输出层类似。在对影像分类的处理过程中，输出层采用一种逻辑函数或一种标准化的指数函数来实现类别标记的输出。在目标识别方面，可以将目标的中心坐标、尺寸、类别等信息进行输出。在语义分割方面，通过输出层将每一像素的分类器的结果直接输出。

## **2.6其他层**

### **2.6.1 Flatten层**

将输入的多维数据变成一维，常用在从卷积层到全连接层的过渡。

### **2.6.2 随机失活层**

随机失活（dropout）[是一种用于神经网络的深度结构优化的方法](javascript:;" \o "" \t "_blank)。在学习过程中，相互连接的神经元会以一定的概率随机地减少或删除。“删除”指的是将该部分神经元的激活函数的输出设置为0，不计算部分神经元。降低节点间的依赖，使网络中过度拟合效应最小化，从而达到神经网络的正则化。在某种意义上，层内的神经元学习的权重值不在是基于其相邻神经元的合作，而是每个训练中的所有不同神经元连接组合的平均值。

## **2.7激活函数**

激活函数（Activition Function）是指一种可以用来模拟神经元的工作情况，从而影响神经网络输出的函数。激活函数可分为线性和非线性。线性激活函数（Linear Activities Function）是指一种可以将输入数据转化为相应的输出数据的函数。这种激活函数在神经网络中通常用来模拟神经元的工作情况，从而影响网络的输出。非线性激活函数（Non-linear Activities Function）是指一种不能将输入数据转换为相应的输出数据，而是通过非线性变换将输入数据变换成一个固定大小的向量，这种激活函数在神经网络中通常用来模拟神经元的工作情况，从而影响网络的输出。

在实际应用中，激活函数通常由训练数据集、激活强度、参数等因素构成。需要注意的是，不同的激活函数在不同领域和任务中会发挥不同的作用，因此需要根据具体情况选择合适的激活函数。

常见激活函数[21]有：

① Sigmoid函数：



|  |
| --- |
| 9  图2-2 |

② Tanh函数：

|  |
| --- |
| 7  图2-3 |

③ ReLU函数：

|  |
| --- |
| 8  图2-4 |

# 第三章 笑脸检测系统的设计与应用

## **3.1 实验平台的配置**

本文采用PYCHARM软件进行程序编写及测试，具体硬件环境如表3-1所示：

表3-1 硬件环境

|  |  |
| --- | --- |
| 操作系统 | 64 位操作系统 |
| CPU处理器 | Intel(R) Core(TM) i5-8265U |
| 内存 | 7.9GB |

## **3.2 系统框架**

|  |
| --- |
| wps  图3-1 系统框架 |

## **3.3 系统软件介绍**

本次模型使用keras库和Tensorflow框架进行卷积网络模型的搭建、训练和优化。

① Tensorflow

Tensorflow[22]是谷歌公司研发的深度学习框架，具有可移植性、可延伸性、高效性、灵活性等特点。它受到了来自国内外科研机构、高校、企业等多方的支持和帮助，允许用户使用Python和TensorFlow编写深度学习算法，并在Python程序中对其进行训练和推理。Tensorflow主要用于构建深度学习模型训练框架，它可以帮助用户使用Python程序对图像、语音、自然语言处理等领域的数据进行训练。

Tensorflow支持多种类型的模型，比如卷积神经网络、循环神经网络、自注意力机制等，还支持多种数据集，包括图片、音频、视频等多种类型的数据，以及数据预处理、特征提取、池化等功能，并且支持多种类型的数据格式和并行计算。TensorFlow提供了多种训练神经网络的工具，包括梯度下降、随机收敛、反向传播等。这些工具可以帮助用户快速构建和训练深度学习模型，并在实际应用中取得更好的性能。

② Keras

Keras是一个开源的深度学习框架，提供了一系列工具和 API，包括模型训练、优化、部署和调试等模块，这些工具和 API可以帮助开发者轻松地构建和部署自己的深度学习模型，并提供了大量的实验环境和示例来帮助开发者更好地了解Keras的工作原理。例如，Keras提供了一个名为 Refine Transformer的模块，该模块可以对输入的序列进行编码，然后通过注意力机制来自动提取特征，最后使用全连接层进行输出。同时可以利用分布式计算和存储来加速模型的训练和推理过程。此外，Keras还提供了一个名为flow release和flow stacking的工具。flow release可以将训练好的模型与实时运行时的数据进行结合，并在训练过程中进行模型优化，从而实现模型的无缝迁移，提高模型性能。flow stacking通常用于模型训练，模型结构更加灵活，可以根据具体需求进行调整。

## **3.4笑脸数据集**

面部表情的分类是人机交互及相关领域的一个关键方面。笑容是一种非常重要的面部表情，它能传达出激动、满足、吸引、友好等多种情绪。随着数字相机、数字影像、社交机器人等商用多媒体产品的普及，笑容识别已逐渐深入到人们的日常生活中。伴随着笑脸识别技术不断进步，国内外学者为了方便研究，收集并创建许多表情数据库。许多比较成熟的数据库如：JAFFE[23]、CK+[24]、GENKI-4K[25]等等都被广大研究学者进行广泛使用。

### 3.4.1笑脸数据库建立

本文采用的是GENKI-4K表情数据集。该数据集每张照片大小、光亮、姿势均不同，专门由美国实验室研究人员收集用于研究笑脸识别。实验过程中，将对4000张图片进行分类，其中包含2162张笑脸，1838张非笑脸。训练1332张图片，测试1332张图片，剩余的用于验证模型的准确率。其中笑脸部分样本图片和非笑脸部分样本图片如下：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | file (603) | file (604) | file (605) | file (606) | file (607) | | file (121) | file (379) | file (288) | file (493) | file (86) |   图3-2 笑脸部分样本 |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | file (3571) | file (2553) | file (2569) | file (2559) | file (2549) | | file (2485) | file (2493) | file (2513) | file (2525) | file (2505) |   图3-3 非笑脸部分样本 |

### 3.4.2笑脸数据预处理

图像预处理是为了去除图像中的无关信息，尽可能地将数据进行简化，方便日后进行其他操作。卷积神经网络中的图像预处理方法包括以下几个方面：

1. 图像大小归一化：在卷积神经网络中，如果在图像上人脸的大小、角度和位置，会因为拍摄方向、位置、距离、焦距等不同的拍摄条件而发生变化，那么，图像的尺寸一般会被重新归一化，以确保图像大小符合模型要求。常见的归一化方法包括将图像大小重新归一化为0或1,以防止模型训练过程中的过拟合。

2. 图像尺寸变换：在卷积神经网络中，图像的尺寸可以通过变换来改变。常见的变换包括平移、放缩、旋转等。

①平移变换:在图片所处的位置上，按照特定的方法，使图片按照X轴，Y轴，或者两个方向都可以同时进行，可以随意改变图片的位置。

②放缩变换：放大或者缩小图像的运算。放大图像可在保持原有大小的情况下加以裁切，而缩小图像一般需常量填充。

③旋转变换：将平面上的图像以规定角度转动或翻转。

3. 图像数据增强：在卷积神经网络中，图像数据增强是一种常用的方法，它可以，增加训练集数量，提高模型泛化性，防止模型过度拟合。常见的数据增强方法包括：随机抽取、正例加强、随机缺失等。

4. 模型训练：卷积神经网络中的模型训练是必不可少的一部分，它可以提高模型的准确率和减少模型复杂度。常用的模型训练方法包括：批量标准化、梯度下降法、随机梯度下降等。

进行数据增强时必须符合实际情况。比如不能将图片旋转180度，这不利于模型的训练。缩放也不能过大或过小，过大会造成人脸的缺失，过小对人脸增强的影响很小。总的来说，卷积神经网络中的图像预处理方法包括图像大小归一化、图像数据增强和模型训练等几个方面，它们在卷积神经网络中发挥着重要作用。通过对这些方法进行合理的组合和应用，可以提高模型的性能和效率。

为保证深度学习过程中图像的大小一致性，对输入图像、训练集、验证集进行了预处理，如图2-3、2-4所示。

|  |
| --- |
| IMG_256  图3-4 图像预处理 |
| IMG_256  图3-5 训练集、验证集预处理 |

## **3.5基于CNN的笑脸检测**

笑脸表情识别技术的关键在于其分类问题，而如何有效地使用所提取的特征，实现其快速、有效的分类，是笑脸识别的关键。因为人类面部表情的产生是以复杂的神经机理为基础的，所人体脸部肌肉的控制能力、表情的表现程度等多方面的因素都会对脸部表情产生微妙而复杂的变化，这就造成了表情变化的复杂性非常高，所以这就对笑脸识别分类算法的设计提出了很高的要求。

本文将采用的是CNN网络模型。CNN进行特征提取有别于传统的人工选取特征，它可以自动习得重要特征。网络各层参数如表4-2所示。

表3-2 网络各层参数

|  |  |  |
| --- | --- | --- |
| Layer (type) | Output Shape | Param # |
| conv2d (Conv2D) | (None, 148, 148, 32) | 896 |
| max\_pooling2d (MaxPooling2D) | (None, 74, 74, 32) | 0 |
| conv2d\_1 (Conv2D) | (None, 72, 72, 64) | 18496 |
| max\_pooling2d\_1 (MaxPooling2D) | (None, 36, 36, 64) | 0 |
| conv2d\_2 (Conv2D) | (None, 34, 34, 128) | 73856 |
| max\_pooling2d\_2 (MaxPooling2D) | (None, 17, 17, 128) | 0 |
| dropout (Dropout) | (None, 17, 17, 128) | 0 |
| conv2d\_3 (Conv2D) | (None, 15, 15, 256) | 295168 |
| max\_pooling2d\_3 (MaxPooling2D) | (None, 7, 7, 256) | 0 |
| dropout\_1 (Dropout) | (None, 7, 7, 256) | 0 |
| flatten (Flatten) | (None, 12544) | 0 |
| dense (Dense) | (None, 512) | 6423040 |
| dense\_1 (Dense) | (None, 2) | 1026 |

## **3.6 系统模块设计**

### 3.6.1 笑脸识别模块

本文通过对深度神经网络的学习，根据要求进行设计。当摄像头获取实时视频时，需对采集到的人脸图像进行笑脸识别。在卷积神经网络中对第二章经过处理过的笑脸数据集进行加载，经过训练后对模型进行保存model.save('smile\_detection\_model.h5')，然后再用该模型对测试集进行测试，以检查该模型的准确率。

当系统与摄像头同时开启后,摄像头则开始不断地扫描识别实时视频中的人脸是否为笑脸。当smile\_prob < no\_smile\_prob即笑脸概率低于非笑脸概率，则根据检测到的人脸区域进行标注为No smile: {no\_smile\_prob:.2f}%。

### 3.6.2 系统界面模块

本文利用PyQT5搭建了一个简单笑脸检测系统平台。界面底部设计了两个按钮，一个“开启摄像头”，一个“关闭检测”，如图4-2。

|  |
| --- |
| 1679725387(1)  图3-6 笑脸检测系统初始界面 |

点击“开启摄像头”按钮即可开启摄像头进行笑脸检测，同时按钮文本切换为“关闭摄像头”。点击“关闭检测”按钮，即可结束当前检测，同时按钮文本切换为“开启检测”。

|  |
| --- |
| C:/Users/529/AppData/Local/Temp/wps.bJaBjHwps  图3-7 笑脸检测流程图 |

## **3.7实验结果**

### **3.7.1损失率折线图**

在深度学习中，训练集和验证集的损失率是一个重要的指标，因为它们反映了模型在训练过程中损失的数据量。

训练集的损失率说明模型在训练过程中损失的数据量。当模型被训练时，它会学习到有用的特征，并将它们用于预测未知数据。然而，在训练过程中，可能会因为各种原因而损失数据，例如模型参数不正确、数据异常、计算资源不足等。验证集的损失率说明模型在验证集上的准确性和可靠性。验证集的损失率是验证模型预测结果与真实结果的百分比。通过计算测试集中的损失，我们可以评估模型的准确性和可靠性。需要注意的是，训练集和验证集的损失率并不是绝对的，因为它们是在不同数据上学习得到的。在某些情况下，训练集的损失率可能高于验证集的损失率，但这并不会影响模型的性能。

据图3-8可知，经过60次训练，训练集和验证集损失率呈现明显的下降趋势，损失率均在0.5以下，意味着模型已经学到了有用的特征，并且在训练过程中表现良好，然而，随着模型的不断学习和改进，在训练至65次时，验证集的损失率在0.4-0.6之间上下波动，可能是因为模型训练得太快或太慢。太快的训练速度可能会导致模型过拟合，使得验证集的损失率出现波动。

|  |
| --- |
| loss  图3-8 损失率折线图 |

### **3.7.2混淆矩阵**

在Python中，混淆矩阵是一种数学工具，它可以帮助我们计算复杂度。混淆矩阵是由一些混淆运算符组合而成的，例如floating、false和aliying可以用来计算复杂度，而不会影响结果。

|  |
| --- |
| 866beb8cc1fecb83656d5b62deab46e  图3-9 混淆矩阵 |

根据图3-9可知，摄像头实际识别到的笑脸，模型预测出来的结果为0.86，非笑脸的为0.88。因此该模型还是有一定的准确率。

### **3.7.3分类报告**

表3-3 分类报告

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| no smile | 0.82 | 0.87 | 0.84 | 612 |
| smile | 0.88 | 0.84 | 0.86 | 720 |
| accuracy |  |  | 0.85 | 1332 |
| macro avg | 0.85 | 0.85 | 0.85 | 1332 |
| weighted avg | 0.85 | 0.85 | 0.85 | 1332 |

表3-3是一个针对二元分类任务的分类报告。其中包括了准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F1 分数（F1-score）等指标。下面是对每个指标的解释：

precision: 精确率，指的是在所有预测为正例的样本中，真正的正例占的比例。也就是说，它描述的是分类器预测为正例的样本中，有多少是真正的正例。计算公式为 TP / (TP + FP)，其中 TP 表示真正的正例，FP 表示误判为正例的负例。

recall: 召回率，指的是在所有实际为正例的样本中，分类器正确预测为正例的比例。也就是说，它描述的是分类器能够识别多少真正的正例。计算公式为 TP / (TP + FN)，其中 TP 表示真正的正例，FN 表示被误判为负例的正例。

f1-score: F1 分数，综合考虑了精确率和召回率两个指标。F1 分数是精确率和召回率的调和平均值，计算公式为 2 \* (precision \* recall) / (precision + recall)。F1 分数的取值范围是 [0, 1]，数值越高表示分类器的性能越好。

support: 每个类别的样本数量。

accuracy: 准确率，指的是分类器正确分类的样本数量与总样本数量的比例。计算公式为 (TP + TN) / (TP + FP + TN + FN)，其中 TP 表示真正的正例，TN 表示真正的负例，FP 表示误判为正例的负例，FN 表示被误判为负例的正例。

此外，还有一个宏平均（macro average）和加权平均（weighted average）指标。宏平均是对每个类别的指标取平均，不考虑样本数量的影响。而加权平均则是对每个类别的指标取加权平均，权重为每个类别的样本数量占总样本数量的比例。

根据表3-3给定的数据，可以看出，分类器在 "no smile" 和 "smile" 两个类别上的召回率分别为 0.87 和 0.84。这表示分类器对于实际上是 "no smile" 和 "smile" 的样本，分别能够识别出 87% 和 84% 的比例。

召回率是指分类器正确预测出所有正例样本（在本例中即 "no smile" 和 "smile" 样本）所占的比例，因此召回率的高低反映了分类器对于真实的正例样本是否能够做出正确的判断。在本例中，两者的召回率都达到了 0.84 或以上，这表明分类器对于这两个类别的样本能够较为准确地进行分类。

### **3.7.3笑脸检测效果图**

为了对真实人脸表情进行判别测试，我们进行实时笑脸识别。程序会调用电脑自带摄像头首先对拍摄到的画面进行笑脸检测，如果检测到笑脸就会框选出人脸，并实时地展示表情识别结果，以及当前识别结果的准确率，如图3-10、3-11所示。

|  |
| --- |
| 1679882204(1)  图3-10 检测笑脸界面 |
| 1679725496(1)  图3-11 检测非笑脸界面 |

# 第四章 总结与展望

本文提出了一种基于深度学习的笑脸识别检测系统，该系统采用卷积神经网络（CNN）的方法，实现对人脸图像中的笑脸识别和检测。

然而，基于深度学习的笑脸识别检测仍然存在一些展望和挑战。以下是未来值得关注的几个方面：

1. 数据集扩大：随着深度学习技术的不断发展和推广，扩大数据集是未来研究的一个重要方向。需要进一步扩大数据集，以验证算法在实际应用中的普遍性。

2. 提高算法精度：深度学习技术虽然可以自动提取图像中的特征，但在实际应用中仍有一些不确定因素影响。未来研究应该更加注重算法精度的提高，以提高系统的实用性和准确性。

3. 多模态融合：深度学习技术不仅可以对人脸图像进行识别检测，还可以对视频、语音等多模态数据进行识别检测。未来研究应该综合运用多种技术，如图像超分辨率、特征选择、表情分析等，以提高系统的整体性能。

4. 硬件设备改进：深度学习技术需要高性能、低功耗的硬件设备支持。未来研究应该不断改进硬件设备，以满足模型训练、检测精度等要求。

总的来说，基于深度学习的笑脸检测系统具有很好的前景和潜力。通过不断改进和优化系统，该系统将在笑脸检测领域取得更好的效果和成绩。

**致谢**

写到“致谢”，便意味着我即将大学毕业，很快就要踏入社会开始新的生活。回想起大学生活，百感交集，总以为四年大学时光很长，但经过疫情之后，就觉得时间飞逝。纵使有千般遗憾、万般不舍，也将勇敢的踏出这一步迈向一段崭新的旅程。

饮其流者怀其源，学其成时念吾师。感谢大学期间的任课老师对我学业上的帮助，以及班主任、辅导员对我生活上的关心。在校这段经历，让我收获了很多东西，也让我找到了自己的热爱。愿所有老师们工作顺利，教泽绵长。

感谢我的家人。即使远在千里，也默默支持着我，给我无限的包容与爱，愿意支持我并尊重我任何一个决定。愿我的家人们身体健康，万事如意。

感谢我的舍友们。虽然我们来自不同的地方，但是我们经过四年的相处，相互照顾，彼此爱护。我们一起上课，一起逛街，一起吃饭，那些开怀大笑的美好瞬间都和你们在一起。愿我们年年约，常相见；但无事，身强健。

感谢我素未蒙面互联网的朋友们。在我感到郁闷或者难过的时候，倾听我的述说，分享我的担忧，给我安慰和帮助。正是你们的出现，让我渡过每一个艰辛又难熬的夜晚。愿我们能从容奔赴理想，成为更好的自己。

至此，愿一切顺遂。

**参考文献**

[1]Mehrabian A,Ferris S R. Inference of attitudes from nonverbal communication in two channels[J]. Journal of consulting psychology,1967,31(3)：248-252.

[2]P. Werbos. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences[D]. Harvard University, Cambridge, MA, 1974.

[3]David E. Rumelhart, Geoffrey E. Hinton & Ronald J. Williams . Learning representations by back-propagating errors[J].Nature,1986,323:533–536

[4]G. E. Hinton,R. R. Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks[J]. Science,2006,313(5786)：504-507.

[5]Alex Krizhevsky, Ilya Sutskever,Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks[J].Communications of the ACM,2017,60(6)：84-90.

[6]Silver David, Huang Aja, Maddison Chris J, et al. Mastering the game of Go with deep neural networks and tree search[J].Nature,2016,529(7587)：484-489.

[7]R.Sujatha,S.L.Aarthy,R.Vettriselvan. Integrating Deep Learning Algorithms to Overcome Challenges in Big Data Analytics[M].CRC Press,2021.

[8]刘小沣.基于深度学习的人脸图像识别技术的研究[D].中国科学院大学(中国科学院长春光学精密机械与物理研究所),2019.

[9]刘成攀.基于深度学习的动态人脸识别[D].西南科技大学,2020.

[10]贾璐.复杂校园场景下佩戴口罩人脸动态识别方法[J].互联网周刊,2022(16):62-64.

[11]赵雨青. 局部遮挡人脸图像的识别与修复研究[D].贵州民族大学,2022.

[12]顾立春,雷鸣.多姿态的人脸识别算法研究[J].电脑知识与技术,2022,18(21):70-72.

[13]卢官明,何嘉利,闫静杰,李海波.一种用于人脸表情识别的卷积神经网络[J].南京邮电大学学报(自然科学版),2016,36(01):16-22.

[14]王珂. 基于深度学习的人脸表情识别算法研究与实现[D].南京信息工程大学,2019.

[15]李传江,张秋秋,张崇明,尹仕熠,韩浩志.笑脸识别鼓励系统在中小学心理健康教育中的应用[J].中国教育技术装备,2020(07):9-11.

[16]张秋秋. 基于深度卷积神经网络的儿童笑脸识别[D].上海师范大学,2020.

[17]Oday A. Hassen, Nur Azman Abu, Zaheera Zainal Abidin, et al. Realistic Smile Expression Recognition Approach Using Ensemble Classifier with Enhanced Bagging[J]. Computers, Materials & Continua,2022,70(2)：2453-2469.

[18]丁金洋.基于残差卷积神经网络的握笔手势识别研究与应用[D].东华大学,2022.

[19]王进,李颖,蒋晓翠等.基于层级残差连接LSTM的命名实体识别[J].江苏大学学报(自然科学版),2022,43(04):446-452.

[20]王兆丰年,欧中洪,宋美娜.基于剪枝的卷积神经网络压缩算法综述[J].中国基础科学,2022,24(03):54-62.

[21]张港.基于深度学习轻量化卷积神经网络的遥感图像场景分类研究[D].南京邮电大学,2022.

[22]何梓林.基于Tensorflow的车型识别系统研究[J].农业装备与车辆工程,2022,60(08):135-138.

[23]陈俊.笑脸表情分类识别的研究[D].华南理工大学,2011.

[24]唐武宾.基于深度学习的视频人脸表情识别研究[D].南京邮电大学,2022.

[25]甘俊英,白振峰,吴必诚等.面向多任务学习的改进十字绣网络在人脸美丽预测中的研究[J].信号处理,2021,37(05):825-834.