西 南 交 通 大 学

本科毕业设计（论文）

基于红外摄像头的活体检测技术研究

年 级： 2016 级

学 号： 2016112422

姓 名： 魏伟

专 业：计算机科学与技术

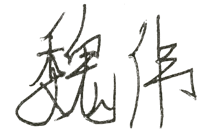
指导教师： 胡晓鹏

二零二零年五月

西南交通大学

本科毕业设计（论文）学术诚信声明

本人郑重声明：所呈交的毕业设计（论文），是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。



作者签名：

日期： 2020 年 5 月 30 日

西南交通大学

本科毕业设计（论文）版权使用授权书

本毕业设计（论文）作者同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权西南交通大学可以将本毕业设计（论文）的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本毕业设计（论文）。

**保密**□，在 年解密后适用本授权书。

本论文属于

**不保密**☑。

（请在以上方框内打“🗸”）

A close up of a hanger

Description automatically generatedA drawing of a tree

Description automatically generated

作者签名： 指导教师签名：

日期： 2020 年 5 月 30 日 日期： 2020年 5 月 30 日

**指导教师评语**

院 系 信息科学与技术学院 专 业 计算机科学与技术

年 级 2016级 姓 名 魏伟

题 目 基于红外摄像头的活体检测技术研究

指导教师

评 语 论文在介绍人脸识别场景中基于红外摄像头的活体检测相关背景的基

础上，讨论了活体检测理论及数据集。按照应用专题的研究思路，首先讨论

基于SVM的红外活体检测模型，包括人脸提取、LBP二值模式提取、GLCM

灰度共生矩阵提取以及模型训练参数设置。然后讨论了基于CNN的红外活

体检测模型，包括基于dlib开源人脸识别库的人脸提取、去背景噪声和模

模型训练参数设置。接着进行了基于SVM模型和CNN模型的红外活体测

试工作。论文作者学习认真，完成一些人脸识别场景中基于红外摄像头的活

体检测研究工作，为后续的学位深造奠定了研究基础。论文结构清晰，达

A close up of a hanger

Description automatically generated 到了本科毕业设计论文水平，同意提交答辩。

指导教师 (签章)

**评阅人评语**

院 系 信息科学与技术学院 专 业 计算机科学与技术

年 级 2016级 姓 名 魏伟

题 目 基于红外摄像头的活体检测技术研究

评 阅 人

评 语 该论文作者调研了活体检测技术的知识技术要点，设计实现了基于红

外摄像头的人脸图像活体检测。重点研究了如何使用支持向量机（SVM）

和卷积神经网络（CNN）这两种分类算法来抵御纸质照片攻击。论文利用红

外图像能够抑制环境光影响，更加清晰地反映人脸轮廓和皮肤纹理信息的

特点，结合纹理检测方法，优化了活体检测技术。论文训练了检测模型并

进行测试，分析了两种算法的有效性。

论文工作较完整，学生工作态度认真。写作规范，完成了任务书要求，

 达到了本科毕业设计论文水平，同意提交答辩。

评阅人 (签章)

**答辩成绩**

院 系 信息科学与技术学院 专 业 计算机科学与技术

年 级 2016级 姓 名 魏伟

题 目 基于红外摄像头的活体检测技术研究

成 绩 良好（84）

答辩委员会主任 (签章)

2020 年 5 月 30 日

**毕业设计（论文）任务书**

班 级 计算机2016-1班 学生姓名 魏伟 学 号 2016112422

发题日期：2019年11月22日 完成日期：2020年5月30日

题 目 基于红外摄像头的活体检测技术研究

1、本论文的目的、意义

最近一些年人脸识别技术取得了突破性进展并且应用越来越广泛。例如，使用人脸识别技术进行身份甄别、人员考勤、人员计数、重点人口监控等。与其它有物理接触的生物特征识别技术相比，人脸识别技术有不易磨损、可靠性高、后期运行维护成本相对低等优势。但是在实际的人脸识别应用中，存在着非监督环境下被识别人员使用纸质照片或者其它数码设备（手机、Pad等）中的照片欺骗摄像头的问题，因此在人脸识别前必须确定被识别人员是活体（即被识别对象是一个真正的人）。为了识别活体，在一些应用场合要求被识别人员按应用系统的指令做出肢体动作（摇头、眨眼等），以确定被识别人员是活体，而在另外一些应用场合则专门安装了红外摄像头，通过红外功能来检测活体，然后再通过可见光摄像头拍照进行人脸识别。目前，在研究领域已经有一些开源的红外活体检测项目，本题目希望对开源的红外活体检测项目进行对比、测试、封装，使之能满足一般人脸识别的应用场合。

2、学生应完成的任务

本论文要求学生对人脸识别应用中的活体检测原理进行研究，并对开源的活体检测项目进行对比、测试、封装，具体要完成的工作包括：

1. 对人脸识别应用中基于红外的活体检测原理进行研究；
2. 对开源的基于红外的活体检测项目进行编译、测试、对比；
3. 根据测试、对比结果选择合适的开源项目进行剪裁、封装，使之可以实现活体检测功能。

3、本设计（论文）与本专业的毕业要求达成度如何？（如在知识结构、能力结构、素质结构等方面有哪些有效的训练。）

本论文支撑本专业以下毕业要求的达成：**（1）**能够通过查阅和分析文献，为计算机系统及工程的问题求解寻找方案，并认识到所求解的问题具有多种可能的解决途径（指标点2.3）；**（2）**能够针对特定需求确定目标，设计计算机系统框架、组成模块，合理组织/存储数据，基于适当的模型进行系统设计与实现，并体现一定的创新意识（指标点3.3）；**（3）**能够在解决方案中从技术、非技术（如经济、社会、健康、安全、法律、文化以及环境等）角度，对设计方案的可行性进行评价和分析（指标点3.4）；**（4）**能够采用科学方法对计算机系统及工程问题进行研究，通过实验对比、文献综合、归纳整理得到合理有效结论，并对其进行规范表述（指标点4.3）；**（5）**能够利用开发环境和工具，对计算机系统及工程问题进行模拟仿真和数据分析（指标点5.3）；**（6）**能识别、分析、评价特定需求的计算机系统在设计和实现中对社会、健康、安全、法律以及文化的影响，并明确自己应承担的责任（指标点6.2）；**（7）**能够评价计算机系统设计、开发、运行和维护对环境保护和社会持续发展的影响（指标点7.2）；**（8）**能够通过口头、文稿、图表等方式、陈述和表达自己的观点，能够就计算机系统及工程问题与同行和相关人员进行交流（指标点10.1）；**（9）**能够根据对工作内容和过程的记录与整理，撰写技术报告和设计文稿、陈述发言或回应质询（指标点10.2）；**（10）**了解计算机系统工程管理原理与经济决策方法，理解计算机系统项目的组织模式和实施过程，掌握项目管理原理和内容（指标点11.1）；**（11）**正确认识自主学习的必要性和重要性，认识到本专业是一个发展迅速的学科，具有自主学习和终身学习的意识（指标点12.1）；**（12）**具备自主学习新技术和新方法的能力，能够通过学习不断提高、适应信息技术和职业的发展（指标点12.2）。

4、论文各部分内容及时间分配：（共18周）

第一部分 学习人脸识别及活体检测技术，熟悉开发工具 (2 周)

第二部分 对开源人脸识别项目进行编译、测试 (4周)

第三部分 对开源人脸识别项目进行封装、测试 (4周)

第四部分 与具体需要人脸识别的应用项目系统进行联调测试 (3周)

第五部分 论文撰写、修改、打印与装订 (4 周)

评阅及答辩 (1周)

备 注

A close up of a hanger

Description automatically generated

指导教师： 2019年11月22日

摘 要

近年来人脸识别技术发展迅速，相关产品在生产生活中应用广泛，逐渐成为个人身份验证的重要手段之一。但是随之而来也出现了一些使用纸质或电子照片欺骗摄像头，冒用人脸进行身份造假的案例。如何对被测者进行活体检测以抵御这些欺骗手段成为人脸识别应用中的关键问题。本文在红外摄像头的基础上，对基于两种分类方法的活体检测技术进行了研究。

本文利用红外摄像头的特点，结合两种分类方法，实现红外人脸图像的活体检测。由于红外图像能够完全抵御电子照片的攻击，所以本文重点研究如何使用支持向量机 （SVM） 和卷积神经网络（CNN）这两种分类算法来抵御纸质照片攻击。本文研究的基于红外图像的活体检测重点利用了红外图像能够抑制环境光影响，更加清晰地反映人脸轮廓和皮肤纹理信息这一特点，同时结合传统活体检测中已经成熟的纹理检测方法，通过提取红外图像中更为明显的纹理特征来优化活体检测技术。在使用 SVM 分类方法完成活体检测的研究中，本文针对红外图像的特点，使用局部二值模式和灰度共生矩阵将人脸图像中的纹理特征提取为特征值，然后对这些特征值使用 SVM 进行二元分类，完成活体检测功能。而使用 CNN 分类方法完成活体检测功能时，本文重点在对深度学习的理解和参数的调试，在实际搭建模型时利用 keras和 Tensorflow提供的接口，结合红外图像对其进行参数设置，以更好完成活体检测。本文在充分理解 SVM 和 CNN 理论的基础上，训练并测试了基于两种算法的模型，并对测试结果作出了细致分析。

**关键词：**活体检测；支持向量机；卷积神经网络

Abstract

In recent years, face recognition technology has developed rapidly, and related products have been widely used in production and life. It has gradually become one of the important means of personal identity verification. But then there have been some cases of using paper or electronic photos to deceive the camera and use the face to fake identity. How to perform live detection on the subject to resist these frauds has become a key issue in face recognition applications. In this thesis, based on the infrared camera, the living detection technology based on two classification methods is studied.

This thesis uses the characteristics of infrared camera and combines two classification methods to realize the live detection of infrared face image. Because infrared images can completely resist the attack of electronic photos, this thesis focuses on how to use support vector machine (SVM) and convolutional neural network (CNN) two classification algorithms to resist paper photo attacks. In this thesis, the infrared image-based living body detection focuses on the use of infrared images to suppress the influence of ambient light and more clearly reflect the contours and skin texture information of the face. Infrared images have more obvious texture features to optimize live detection technology. In the study of using SVM classification method to complete live detection, this thesis focuses on the characteristics of infrared images, using local binary mode and gray level co-occurrence matrix to extract the texture features in face images as feature values, and then use SVM to perform these feature values. Binary classification, complete living body detection function. When using the CNN classification method to complete the live detection function, this thesis focuses on the understanding of deep learning and parameter debugging. When the model is actually built, the interface provided by keras and Tensorflow is used to set the parameters in combination with infrared images to better complete Biopsy. Based on a full understanding of SVM and CNN theory, this thesis trained and tested the models based on the two algorithms, and made a detailed analysis of the test results.

**Keywords:** Liveness detection; support vector machine; convolutional neural network

目 录

[第1章 绪 论 1](#_Toc41514222)

[1.1背景与意义 1](#_Toc41514223)

[1.2国内外发展（应用）现状 1](#_Toc41514224)

[1.3论文所做工作及思路 3](#_Toc41514225)

[1.4论文章节安排 4](#_Toc41514226)

[第2章 活体检测相关理论及数据集 5](#_Toc41514227)

[2.1 SVM分类方法概述 5](#_Toc41514228)

[2.1.1支持向量分类机 5](#_Toc41514229)

[2.1.2核方法 9](#_Toc41514230)

[2.2 CNN分类方法概述 11](#_Toc41514231)

[2.2.1卷积层 12](#_Toc41514232)

[2.2.2池化层 13](#_Toc41514233)

[2.2.3 激活函数 14](#_Toc41514234)

[2.2.4全连接层 16](#_Toc41514235)

[2.3数据集 16](#_Toc41514236)

[2.3.1 中科院近红外人脸数据集 16](#_Toc41514237)

[2.3.2 CASIA-Surf多模态活体检测数据集 16](#_Toc41514238)

[2.3.3自制红外数据集 19](#_Toc41514239)

[2.4本章小结 19](#_Toc41514240)

[第3章 基于SVM的红外活体检测模型训练 21](#_Toc41514241)

[3.1图像预处理 21](#_Toc41514242)

[3.1.1人脸提取 21](#_Toc41514243)

[3.1.2 LBP 二值模式提取 22](#_Toc41514244)

[3.1.3 GLCM灰度共生矩阵提取 22](#_Toc41514245)

[3.2模型训练 24](#_Toc41514246)

[3.2.1训练过程 24](#_Toc41514247)

[3.2.2训练结果 25](#_Toc41514248)

[3.3本章小结 25](#_Toc41514249)

[第4章 基于CNN的红外活体检测模型训练 26](#_Toc41514250)

[4.1图像预处理 26](#_Toc41514251)

[4.2模型训练 26](#_Toc41514252)

[4.2.1参数设置 26](#_Toc41514253)

[4.2.2训练结果 28](#_Toc41514255)

[4.3本章小结 29](#_Toc41514256)

[第5章 基于两种模型的实际测试工作 30](#_Toc41514257)

[5.1 SVM 测试 30](#_Toc41514258)

[5.2 CNN测试 32](#_Toc41514259)

[5.3本章小结 32](#_Toc41514260)

[结 论 35](#_Toc41514261)

[致 谢 36](#_Toc41514262)

[参考文献 37](#_Toc41514263)

第1章 绪 论

1.1背景与意义

最近一些年人脸识别技术取得了突破性进展并且应用越来越广泛，已逐渐成为个人身份验证的重要工具。例如，使用人脸识别技术进行身份甄别、人员考勤、人员计数、重点人口监控等。与其它有物理接触的生物特征识别技术相比，人脸识别技术有不易磨损、可靠性高、后期运行维护成本相对低等优势。但是在实际的人脸识别应用中，存在着非监督环境下被识别人员使用纸质照片或者其它数码设备（手机、Pad等）中的照片欺骗摄像头的问题，包括人脸照片攻击 、人脸视频回放攻击以及三维人脸模型攻击[1,2]。照片攻击是指非法用户翻拍合法用户的图片或从视频里截取图像，通过打印或电子屏幕呈现给认证系统的攻击。回放攻击就是非法用户重放合法用户的视频以欺骗识别系统。三维人脸模型攻击即通过三维建模制造一合法用户的3D面具来欺骗识别系统。

因此，在人脸识别前必须确定被识别目标是活体（即被识别目标是一个真正的人）。在这种背景下，活体检测技术应运而生[3]。通过在人脸识别之前加入一道活体检测关卡，只有在识别目标被检测为活体的情况下，才会进行下一步的人脸识别。通过对目标进行双重检测，大大加强了身份验证系统的安全性和可靠性。活体检测技术为人脸识别应用消除了潜在的安全隐患，并能够加速人脸识别技术的进一步普及和应用。

为了识别活体，在一些应用场合要求被识别目标按应用系统的指令做出肢体动作（摇头、眨眼等），以确定被识别目标是否为活体，如银行自助柜台。这种方式检测精度高，但是需要检测目标配合，时间要求比较长，不适合需要短时间进行大量人脸识别的应用场合。而在另外一些应用场合则专门安装了红外摄像头，通过红外功能来检测活体，然后再通过可见光摄像头拍照进行人脸识别。这种活体检测方法非常适合非监督情况下的检测，且检测时间一般很短，适用于即检即过。目前，在研究领域已经有一些开源的红外活体检测项目，本题目便针对开源的红外活体检测项目进行对比、测试、封装，使之能满足一般人脸识别的应用场合。

1.2国内外发展（应用）现状

活体检测作为伴随人脸检测兴起的一项技术，近年来吸引了众多研究者的兴趣，也提出了许多算法。文献[4]总结了三种活体检测方法：基于光谱或纹理信息的活体检测、基于深度信息的活体检测和基于多帧动态视频的活体检测。文献[5]则对近十年来生物活体检测技术做了细致的总结，在该文献中Galbally J等将活体检测技术分为三大类：基于静态细节纹理的技术、基于动态细节纹理的技术和基于传感器层次的技术。其中基于静态细节纹理的技术就是文献[4]所说的基于光谱或纹理信息的活体检测，而基于动态细节纹理的技术则包含了文献[4]所说的基于深度信息的活体检测和基于多帧动态视频的活体检测。同时文献[5]还详细列出了各项技术的代表文献、具体方法、攻击方式、数据集以及识别错误率。

基于静态纹理检测的活体检测算法，它通过分析图像局部或全局纹理信息，找出真假人脸的差异。其使用的图像分析方法有很多，包括使用傅立叶频谱、高斯多重差分来提取图像中的频率信息[6]；使用偏最小二乘法、梯度方向直方图来分析图像的一些小的细节信息[7]，例如边缘（edge)、角(corner)、颜色(color)、像素(pixels)、梯度(gradients)等；使用局部二值模式来进行细节分析以抵挡攻击[8]；使用灰度共生矩阵[9]研究灰度图像的空间相关性进而获取相关纹理等。这类算法通常实时性较好，但对环境光照敏感，对高清照片无法鉴别，算法鲁棒性不高。

基于动态纹理检测的活体检测算法有两类：基于深度信息的活体检测和基于多帧动态视频的活体检测。基于深度信息的活体检测方法有：基于人脸关键点的稀疏重建方法来恢复人脸的三维结构[10]、基于双目获得人脸三维点云信息。而基于多帧动态视频的活体检测则包括检测人的眨眼动作、检测面部和头部姿势（点头、微笑、眼光角度等）以及使用光流法[11]等。这些算法因为同时利用了时间和空间上图片（连续帧）的细节，使得它能有效抵御平面模型的攻击，尤其是应对照片和三维模型攻击成效显著。但这些算法无法针对单张图片进行处理，录制视频存储容量也上升了一个数量级，且无法应对合成视频攻击，使用时需要用户花一定时间配合，用户体验并不好。同时某些算法（三维重建算法）相对耗时、系统实时性不高。

基于传感器层次的技术则是在硬件上对活体检测技术进行革新，其中最典型的就是本课设研究的基于红外图像的活体检测。近红外图像(NIR-Near infrared)是通过红外摄像头主动发出红外光探照目标，接收端加上只允许近红外光通过的窄带滤光片得到的与环境光相对无关的图片。人脸NIR图不仅保留了除色彩以外的形状、轮廓信息，并且还有效抑制了环境可见光的影响，并能体现一定的人脸深度变化、皮肤纹理信息，在人脸活体检测方面具有很大的应用前景。另外，红外摄像头主要是靠接收目标反射的红外光，表面光滑平整的物体如塑封照片、手机屏幕、平板电脑等极易反光，在红外摄像头中不能清晰地成像，这就很容易鉴别此类攻击。除此之外，近红外人脸活体检测无需识别目标配合，能够在短时间中完成检测，且检测成功率较高。

因为现实应用的需求，目前国内外对基于红外摄像头的活体检测逐渐增加。文献[12]探讨了一种结合RGB彩色图像和NIR近红外图像来对目标进行活体检测的方法。在该方法中，主要利用近红外图像中人瞳孔的明亮特征（Bright Pupil Feature）以及人脸鼻子区域的梯度方向直方图来得到分类（真实/虚假）依据，最后使用SVM分类器分类，得到活体检测判断结果。但是近红外图像中人瞳孔的明亮特征对光线、摄像条件等比较严格，且要求目标未佩戴眼镜，通常来说这些要求在现实生活之中难以满足。与之相似的，文献[13]也使用红外图像和可见光图像的结合来进行活体检测，不过在该文献给出的方法中，研究者使用热红外TIR(Thermal IR)而不是近红外。其主要方法是将人脸划分为8块区域：左睫毛、左脸颊、左眼、右睫毛、右脸颊、右脸、鼻子和嘴巴。然后对可见光和红外光条件下这八块区域的特征进行交叉分析建模，再进行活体检测。文献[13]的方法相比于传统活体检测方法要新奇巧妙，但是对于佩戴眼镜的目标检测效果不好，而且热红外摄像头价格昂贵， 不适合实际应用，并且文献中着重于理论的研究，缺乏实证分析，实验结果也不是很令人信服。

1.3论文所做工作及思路

本文基于人脸识别应用中的活体检测原理，结合红外图像自身的特点，选取了支持向量机（Supporting Vector Machine, SVM）和卷积神经网络（Convolutional Neural Networks, CNN）两种分类方法来对红外活体检测技术进行了研究。主要研究工作如下：

1.基于 SVM 的红外活体检测技术研究

SVM 是一种广义的二元线性分类器，其基本思想是在特征空间中寻找一个能够将数据集划分为两个不同集合的“超平面”。考虑到活体检测在本质上就是一个简单的二元分类问题，故本文首先对使用SVM来实现活体检测进行了研究。

人脸识别中的活体检测是以图像为基础的，而图像在计算机中作为一个矩阵数据，无法直接输入到SVM分类器中进行分类。需要提取图像的某些关键特征，然后以这些关键特征代表该图像进行分类。因此，对图像关键特征的选择显得尤为重要。本文经过查阅相关文献和了解计算机图形学的相关知识，选择先对人脸图像进行局部二值模式(Local Binary Patterns, LBP)处理，提取出图像的大体灰度特征，然后再使用灰度共生矩阵(Gray-level co-occurrencematrix, GLCM）特征提取来提取图像的四个关键纹理特征值。这四个关键纹理特征值分别是：均值（Mean）、逆差分矩（Inverse Difference Moment）、信息熵（Entropy）和角二阶矩（Angular Second Moment）。这个四个关键纹理反映了图像纹理变化和灰度分布的相关信息，在活体图像和非活体图像之间具有显著差别，适合用作分类指标。

选择好分类特征值之后，算法实现相对简单。首先编写LBP和GLCM的算法，对图像进行预处理，获得关键特征值。然后使用python的机器学习第三方库sklearn来训练数据，获得训练模型。最后将模型封装好，在主函数调用模型进行活体检测。

2. 基于 CNN 的红外活体检测技术研究

CNN 是一类前馈神经网络的统称，在深度学习和图像处理中具有广泛应用。因为活体检测也是对图像的识别和分类过程，所以本文的第二个工作就是使用CNN的方法进行活体检测的研究。

CNN 的输入相较于 SVM 简单许多，只需将图像由正常的图像格式（如jpg格式）转为矩阵格式就可以。但运行时环境的搭建和参数的设置则较为复杂。本文搭建的CNN运行环境是基于CPU的Tensorflow底层库，使用开源人工神经网络库Keras来作为Tensorflow的高阶应用接口。通过该接口对数据集进行训练，获得训练模型。将模型封装之后，通过简单的调用就可以完成活体检测功能。

1.4论文章节安排

论文共分5章。

第1章对论文内容做了整体概述，介绍了研究课题的相关背景以及研究意义，同时描述了对于本课题当前国内外相关研究的状况。本章还对论文的整体架构做了介绍。

第2章主要对本文涉及到的理论基础做了梳理。包括SVM分类方法的基本概念、公式推导以及推广后的重要概念：核方法；CNN 分类方法的四个重要概念：卷积层、池化层、激活函数和全连接层。同时第 2 章还介绍了本文使用的 3 个数据集，包括两个公开红外人脸数据集和一个自制红外攻击人脸数据集。

第3章介绍基于SVM的红外活体检测模型的训练过程，包括图像的预处理和模型的训练两部分。其中图像预处理用到了dlib、LBP、GLCM 等相关图像处理技术，在第 3 章中也对这些技术做了介绍。最后第3章给出了基于SVM的红外活体检测模型的训练结果。

第4章介绍基于CNN的红外活体检测模型的训练过程，同样包括图像的预处理和模型的训练两部分。其中图像预处理十分简单，而模型训练则重点介绍了CNN中一些重要的超参数图和设置。第3章也给出了基于CNN的红外活体检测模型的训练结果。

第5章介绍针对第4章和第3章训练的模型进行的测试工作。第五章给出了对基于SVM 的训练模型和基于CNN的训练模型的测试结果，同时分析了测试结果与训练结果差异产生的原因，指出了后续工作的方向。同时，第5章也说明了训练的模型在实践中的应用能力。

第2章 活体检测相关理论及数据集

2.1 SVM分类方法概述

2.1.1支持向量分类机

作为机器学习中监督学习的典型方法，SVM分类方法最早由Corinna Cortes和Vladimir Vapnik在 1995 年的论文中提出[14]，在文中称之为支持向量网络（Support-Vector Network, SVN）。其基本思想是：将输入空间通过非线形变换投射到一个高维的特征空间，并在该特征空间中寻找能够对样本进行最佳划分的超平面（Hyperplane）这种超平面在二维空间中是一条线，在三维空间中是一个平面…，而离超平面最近的特征向量就被称为支持向量（Support Vector）。支持向量到超平面的距离称为间隔。SVM的核心就是使超平面和支持向量之间的间隔尽可能的大，如此一来超平面才可以将两类样本准确的分开，使分类器误差尽可能的小，尽可能的健壮，如图2-1所示。

A picture containing clock

Description automatically generated

图 2-1 二维 SVM 分类示意图，其中灰色点代表支持向量

令二元分类的输入样本集T为 {(x1,y1),(x2,y2),...,(xn,yn)} ∈ T ，并且xi ∈ Rd，i ∈ [1, n]，即输入数据为d维向量，xi ∈ Rd 为第i个特征向量，yi ∈ {−1,+1}，即分类标签为二元，-1 表示负样本，+1表示正样本。Corinna Cortes和Vladimir Vapnik指出，如果样本向量h 线性可分，则必有以下两个不等式恒成立：

其中 w 是一个d维向量，对式子（2-1）和式子（2-2）进行归一化处理后即可得：

故就是所求的超平面，其超平面方程为：

式（2-4）即是特征空间中存在的能使得划分间隔最大的唯一超平面。假设P(x1,x2,...,xn) 为特征空间中的一个点，则该点到超平面的欧氏距离为：

所以在特征空间中，最优超平面划分得出的最大间隔即是两个分类标签分别为 1 和-1 的支持向量到最优超平面的距离之和，结合式（2-4）和（2-5）可以得出：

对式（2-6）再度归一化后可以得到分类间隔.：

要求出最优的超平面方程，即要求式（2-7）最大值，而要使（2-7）最大，就必须要求||W||最小。为了后面求极值方便，我们将求||W||的最小值放缩为求A close up of a logo

Description automatically generated 的最小值。于是我们得到：当前的已知条件为式（2-4），目标函数为A close up of a logo

Description automatically generated。这是一个含有不等式约束的凸二次规划问题，通常我们可以用拉格朗日乘子法来求解。首先，我们将有约束的原始目标函数转换为无约束的新构造的拉格朗日目标函数：

其中为拉格朗日乘子，且≥ 0。根据拉格朗日乘子法求有约束条件函数极值方法，对L求关于w,b的偏导数得：

将上述（2-9），（2-10）两式带入拉格朗日目标函数（2-8）得到：

即：

此时问题即为求对的极大值。将式（2-11）乘以-1，把求最大值转化为求w，b最小值：

且该问题的KKT(Karush–Kuhn–Tucker)条件[15]为：

将优化问题变为（2-13）的形式之后，便可以使用序列最小优化算法[16]（Sequential Minimal Optimization, SMO）对其求最优。通过序列最小优化算法得到最优解。与式（2-9）和式（2-10）结合，可知在中至少有一个，对此j有：

根据该j便可以求出，，进而得到超平面的表达式：

从超平面的表达式（2-18）可以看出，在输入的训练样本中，只有那些α∗ 中大于y的分量所对应的训练样本才会决定最终的超平面。这些训练样本就是支持向量，即图 2-1中灰色样本点，而其他的大部分训练样本在训练完成之后都不会被保留，最终的模型（超平面）仅与支持向量有关。

式（2-18）给出了在训练数据线性可分的情况下支持向量分类机的工作原理，但是在很多情况下并不能保证输入的训练数据线性可分，如图2-2所示，此时的处理方法有两种。第一种方法同样是Corinna Cortes和Vladimir Vapnik在他们的论文[14] 中提出的，即所谓的“软间隔”方法。第二种方法则是下一节将会详细介绍的核方法。

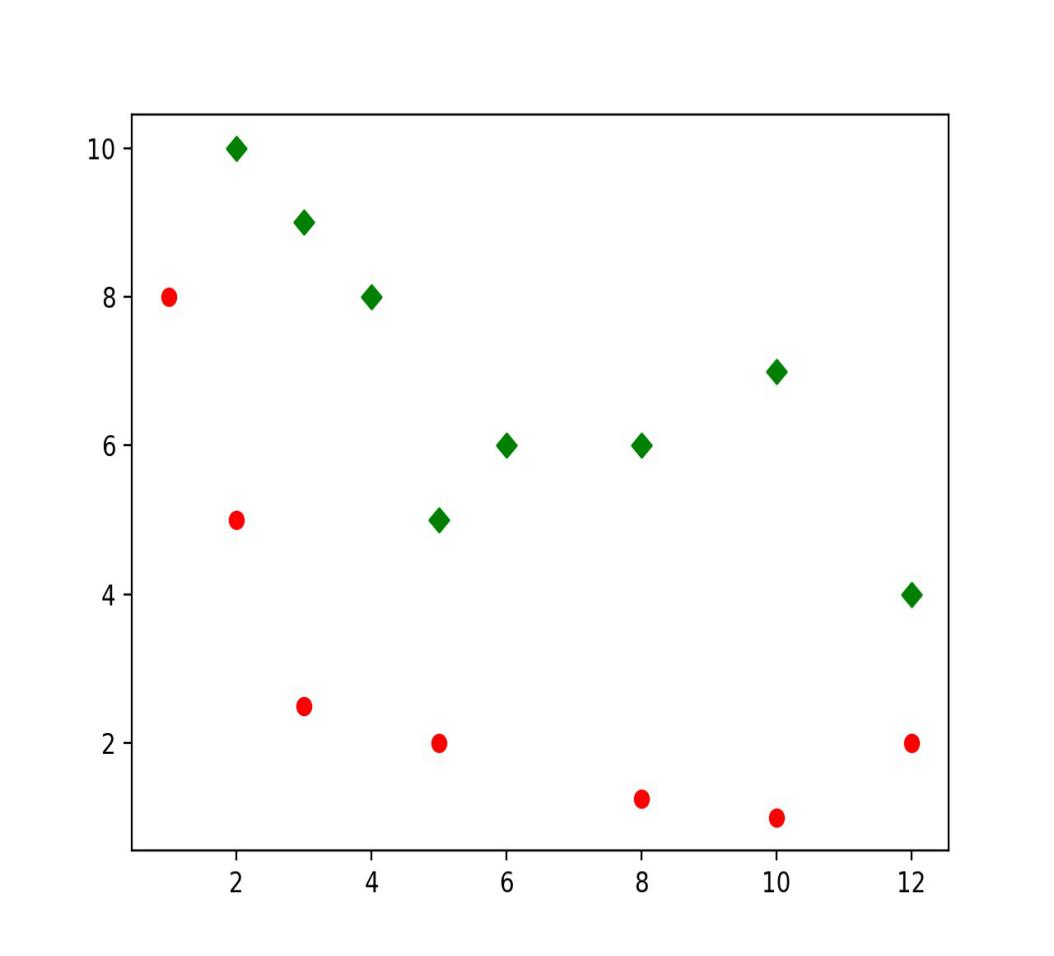


图 2-2 无法进行线性分类的样本

“软间隔”的思想是允许某些点不满足式（2-3）的约束，引入一个松弛变量ξ 来允许一些数据可以处于分隔面错误的一侧。此时，新的约束条件变为：

其中ξi 称为“松弛变量”，它的含义为允许第i个训练数据偏离分隔面的距离，ξi = max(0,1−yi(w ∗ xi + b)) 如果ξ 趋于无穷大，那么任意的超平面都将会符合要求。所以此时的目标是让ξ 的总量尽可能的小。新的目标函数变为：

其中的C是一个权重值，也被称为“惩罚因子”。它是用来控制“最大化间隔”和“保证大部分的点的函数间隔都小于1”这两个目标。经过计算化简之后得到目标函数：

该目标函数与线形可分时的目标函数一样，但由于添加了松弛变量，此时可以解决更为复杂的问题，通过修改参数C，可以得到不同的结果。C的具体大小则需要根据实际问题进行调节。

2.1.2核方法

上一节提到，在处理不能线性分类的数据时，可以使用“软间隔”方法，即引入松弛变量。但是很多情况下，这种方法效果也不是很好，特别对于那些分布明显不是线性分布的数据时，如图2-3。该图所表示的输入数据对应的分割超平面明显应该是一个圆，如果采用线性划分，就算引入松弛变量，效果也很差。Boser[17]等人因此提出了一种全新的方法来解决这一问题。他们的主要思想是通过一个映射函数将低维度下线性不可分的数据映射到高维空间中，使得他们在高维空间中线性可分，从而在高维中进行支持向量机分类。以图2-3为例，在图2-3 所示的二维空间中，很明显它是线性不可分的，但是如果将它映射到三维空间中，如图2-4所示，此时的数据就可以被一个平面线性划分。核方法处理SVM的基本思路为：

1.对于一个输入空间为E维的数据集，通过一个非线性映射，将该数据集映射到D维的高维空间，在这个高维的特征空间中实现线性划分。

2.在高维特征空间中使用映射后的数据进行线性划分，得到最优超平面，即得到分类决策函数。

3.对于待分类的新数据，在调用SVM进行分类时，先把数据映射到高维特征空间，再利用训练得到的分类决策函数进行分类。

在引入映射函数之后，式（2-21）所表示的目标函数将会变为：

A picture containing glass, star

Description automatically generated

图 2-3 非线性分布数据

A close up of a map

Description automatically generated

图 2-4 投射到高维空间后的图像

至此，如果知道的具体定义，便可以直接使用线性可分的方法来求出决策函数，完成分类。

但是在大多数情况下，由于维度灾难的原因，都不会显式定义的具体表达，也不会直接去计算的结果。观察式（2-22）我们需要计算的是, 并不需要单独将和计算出来。由此引入了核技巧[18] (Kernel Trick）。

核技巧是通过核函数来实现的。设xi ∈ χ，χ 是输入空间，即χ 是Rd的子集，又设H为特征空间，如果存在一个从 χ 到 H 的映射,

使得对于所有的 x，z ∈ χ，函数 K(x，z) 满足：

那么就称K为核函数，φ(x) 为映射函数，为内积。核函数的意义是：对于输入核函数的两个向量，核函数将返回一个值，该值与两个向量分别作φ映射然后点积的结果相同。核技巧通过直接计算 的结果，避开分别计算 φ(x) 和 φ(z)，以加速计算过程。使用核技巧，式（2-22）就可以表示如下：

核函数的选择是SVM最大的变数，如何选择合适的核函数一直以来没有定论。就像现实生活中的诸多选择一样，核函数的选择需要结合训练样本的实际情况进行不同的选择。在实验中尝试积累的比较好的核函数有：

1.线性核函数：

2.多项式核函数：

3.S型核函数：

4.高斯核函数：

其中的，β，σ 都是参数，需要根据实际情况在训练过程中选择。在上述的核函数中，以高斯核函数的泛化能力最好，同时高斯函数也是在实践中应用最广泛的核函数，本文的实验所采用的核函数就是高斯核函数。

2.2 CNN分类方法概述

CNN是一种前馈型的神经网络，在大型图像处理方面有出色的表现，CNN在应对图像的平移、缩放和旋转时有很好的不变性[19]。1962年Hubel以及Wiesel[20]通过生物学研究表明，从视网膜传递到脑中的视觉信息是通过多层次的感受野（Receptive Field）激发完成的。在视觉皮层中有一系列细胞对视觉输入空间的局部区域很敏感，Hubel和Wiesel将这些细胞称为感受野，并将这些细胞分为简单细胞和复杂细胞。这一视觉神经机制的发现对后续神经网络的设计启发很大。1980年，Fukushima根据Hubel和Wiesel的视觉层级感知模型提出了结构类似的神经认知机(Neocognitron)。它是一个自组织的多层神经网络模型，每一层的响应都由上一层的局部感受野激发得到，在识别模式时不受位置、较小形状变化以及尺度大小的影响。神经认知机已经有了卷积神经网络的雏形。到1998年，LeCun等人在Fukushima的基础上使用BP算法设计了LeNet-5[21]。LeNet-5是典型的CNN结构，后续的CNN模型基本上都在LeNet-5的基础上进行改进，尤其是在2006年Hinton[22]提出深度学习之后，对CNN的研究呈现井喷之势，相关成果也在现实应用中取得了不俗的效果。

CNN基本结构包括输入层(Input Layer)、卷积层（Convolutional Layer）、池化层（Pooling Layer）、全连接层（Fully Connected Layers）及输出层（Output Layer）构成。一般会使用多个卷积层和池化层，采用卷积层和池化层交替设置的方法，即一个卷积层连接一个池化层，池化层后再连接一个卷积层，依此类推。这样连续的卷积-池化过程又被统称为隐藏层（Hidden Layer）[23]。

A picture containing clock, table

Description automatically generated

图 2-5 CNN 流程

2.2.1卷积层

卷积是分析数学的概念，设f(x)、g(x)是实数域上的两个可积函数，则：

称为f(x)与g(x)的卷积。如果将式（2-26）离散化，则可得到：

进一步推广到离散二维形式：

可以看出，卷积表征函数f(x)与g(x)经过翻转和平移的重叠部分的面积。在CNN中使用的卷积层是离散二维形式的卷积，但是又与分析数学中的离散卷积有一些区别，其卷积操作更类似于一种线性运算：CNN中的卷积是将卷积核与输入数据进行移位相乘，并进行加权求和。

卷积核又被称为权重矩阵，式（2-29）展示了CNN卷积的运算过程。等式左边的3\*3矩阵表示一个卷积核，5\*5矩阵表示输入数据，“\*”运算符代表进行卷积运算。经过运算我们得到一个 3\*3 的特征图（Feature Map）。该特征图经过激励函数处理后将会被作为池化层的输入。

A close up of a logo

Description automatically generated（2-29）

2.2.2池化层

池化层又称汇合层，是卷积神经网络中一般紧跟在卷积层之后的操作。同卷积不同，池化并不含有需要进行学习的参数，只需指定池化类型（最大值池化或者均值池化）、池化核的大小以及移动池化核的步长。

同卷积层类似，在池化层中，用来与输入数据进行运算的矩阵称为池化核。不过该池化核非常简单，一般只用平均（最大）池化核。平均（最大）池化核在进行每次操作时，将池化核覆盖区域中所有值的平均值 （最大值） 作为汇合。图2-6展示了大小为 2\*2，步长为1的最大值池化的效果。

从图2-6可以看出，同卷积层一样，池化层也有对数据降维的作用。经过池化操作，可以降低输入数据的规模，对特征数据进一步抽象。其实池化操作实际上就是一种“降采样”（Down-Sampling）操作，这也是在某些文章中池化层又称为“降采样层”的原因。池化层的引入是模拟生物的视觉系统对视觉输入对象进行降维和抽象的过程。通常认为池化层有如下作用：

A close up of a screen

Description automatically generated

图 2-6 最大值池化操作

1.特征降维。因为池化操作的降采样作用，使得池化结果中的每个元素都对应原输入数据中的一个子区域，就相当于在空间范围内对数据做了维度约减，从而使得模型可以抽象更加广泛的特征，同时也减小了下一层（通常是卷积层）的输入大小，进而减小了计算量和未知参数个数。

2.特征不变性。池化操作能够使得模型更关注特征是否存在而不是特征的具体位置，从而让CNN能够包含一定的容忍自由度，能容忍图像中特征微小的位移，让模型的功能更加泛化。

3.池化层还有一定的防止过拟合[24]（Overfitting）的作用，更方便优化。

2.2.3 激活函数

和 SVM 一样，CNN 也需要考虑非线性划分的情况，上面介绍的卷积层和池化层得到的结果都是线形变化，如果输入数据的分类不是线性可分，那么训练的模型就会出现较大偏差。基于此，提出了激活函数的概念。

激活函数（Activation Function）又称为非线性映射，引入激活函数是为了增加整个网络的表达能力（即增强神经网络对非线性的适应能力）。如果没有激活函数，若干个线性操作层的叠加只能起到线性映射的作用，难以形成复杂函数。最经典的激活函数是 Sigmoid型函数，它是对生物神经元的模拟[25]：接受一组输入信号并产生输出。在神经科学中，生物神经元一般会有一个阈值，神经元所收到的输入信号没有达到该阈值时，神经元就处于抑制状态。当神经元所获的输入信号累计效果超过了该阈值，神经元就被激活处于兴奋状态。Sigmoid型函数就是模拟了这一过程。

Sigmoid型函数又被称为Logistic函数：

其函数图像如图 2-7所示。从图像可以看出，不论输入为多少，Sigmoid型函数的输出范围只能在[0,1]之间，而0对应了生物神经元的“抑制状态”，1对应着“兴奋状态”。由此实现了对数据的非线性映射。

A screenshot of a cell phone

Description automatically generated

图 2-7 Sigmoid函数图像

另一个经常被用到的激活函数是Nair和Hinton在2010年提出的修正线性单元（Rectified Linear Unit, ReLU）[26]。ReLU函数是一个分段函数，又被称为线性整流函数：

其函数图像如图2-8所示。ReLU函数最大的优点在于它十分简单的计算过程[27]，相比Sigmoid型函数，它没有复杂的指数函数和除法计算，使得整个神经网络计算成本下降。

A screenshot of a cell phone

Description automatically generated

图 2-8 ReLU函数图像

此外，ReLU函数还促进了更加有效率的梯度下降和反向传播，避免了梯度爆炸和梯度消失的问题。

2.2.4全连接层

经过多个卷积层和池化层之后，CNN 完成了特征提取的任务，此时需要一个对提取到的信息进行整合并映射到样本的标记空间的功能块。全连接层就是这样的一个功能块。在 CNN 中，通常会在卷积核池化完成之后连接1个或多个全连接层。之所以叫全连接层，是因为全连接层中的每一个神经元都与前一层所有神经元进行全连接。全连接层整合了卷积层和池化层中具有类别区分性的信息[28]。全连接层可由卷积操作进行实现：当前一层也是全连接层时，全连接层使用卷积核为1\*1的卷积；而当前一层是卷积层时，全连接层使用卷积核为h\*w的全局卷积，其中h、w分别是卷积层输出结果的高和宽。最后一层全连接层输出后，将会被传递到一个输出层，该输出层通过分类函数对全连接层的结果进行分类。通常，输出层会使用softmax逻辑回归进行分类。

另外，当使用一个大的前馈神经网络去训练一个较小的数据集时，训练模型容易出现过拟合，造成模型表现不佳。为了避免这一情况，常在全连接层中采用正则化方法——丢失数据（Dropout）技术。它使隐藏层（即多个卷积层和池化层）神经元的输出值以一定的概率变为0，所以又被称为随机失活技术。通过该技术让隐藏层的部分节点失效，对于每次输入到网络中的样本，由于 Dropout 技术，它所对应的网络结构都不相同，降低了神经元之间相互适应的复杂性，使得神经元学习能更加鲁棒[29]。

2.3数据集

在本文中使用了三个红外人脸数据集，本节对这三个数据集进行介绍。

2.3.1 中科院近红外人脸数据集

该数据集由中国科学院自动化研究所 （CASIA） 于 2013 年制作，包含 725 个不同年龄段的人的活体红外图像，每个人拍摄了5-50张不同光照，不同角度的红外图像，共计17580张图像，每张图像分辨率为640\*480[30]。数据集示例如图2-9。

2.3.2 CASIA-Surf多模态活体检测数据集

该数据集由中国科学院自动化研究所（CASIA）与澳门科技大学以及巴塞罗那大学共同制作，包含1000个人的活体图像以及每人的攻击图像。制作者设计了6种不同的攻击方式[31]：

攻击方式一：剪去被拍照者的人脸照片中的眼睛部分后，由该被拍照者举着这张照片对着摄像头拍摄。

攻击方式二：剪去被拍照者的人脸照片中的眼睛部分后，由该被拍照者将照片向两侧弯曲后紧贴面部（使得真实的眼睛部分更加突出）对着摄像头拍摄。

A picture containing lined, row, different, many

Description automatically generated

图 2-9 NIR-VIS红外人脸图像

攻击方式三：剪去被拍照者的人脸照片中的眼睛以及鼻子部分后，由该被拍照者举着这张照片对着摄像头拍摄。

攻击方式四：剪去被拍照者的人脸照片中的眼睛以及鼻子部分后，由该被拍照者将照片向两侧弯曲后紧贴面部（使得真实的眼睛以及鼻子部分更加突出）对着摄像头拍摄。

攻击方式五：剪去被拍照者的人脸照片中的眼睛，鼻子以及嘴巴部分后，由该被拍照者举着这张照片对着摄像头拍摄。

攻击方式六：剪去被拍照者的人脸照片中的眼睛，鼻子以及嘴巴部分后，由该被拍照者将（使得真实的眼睛，鼻子以及嘴巴部分更加突出）对着摄像头拍摄。

该数据集同时提供上述活体图像以及攻击图像的不同版本：RGB图像，近红外图像以及深度图像。

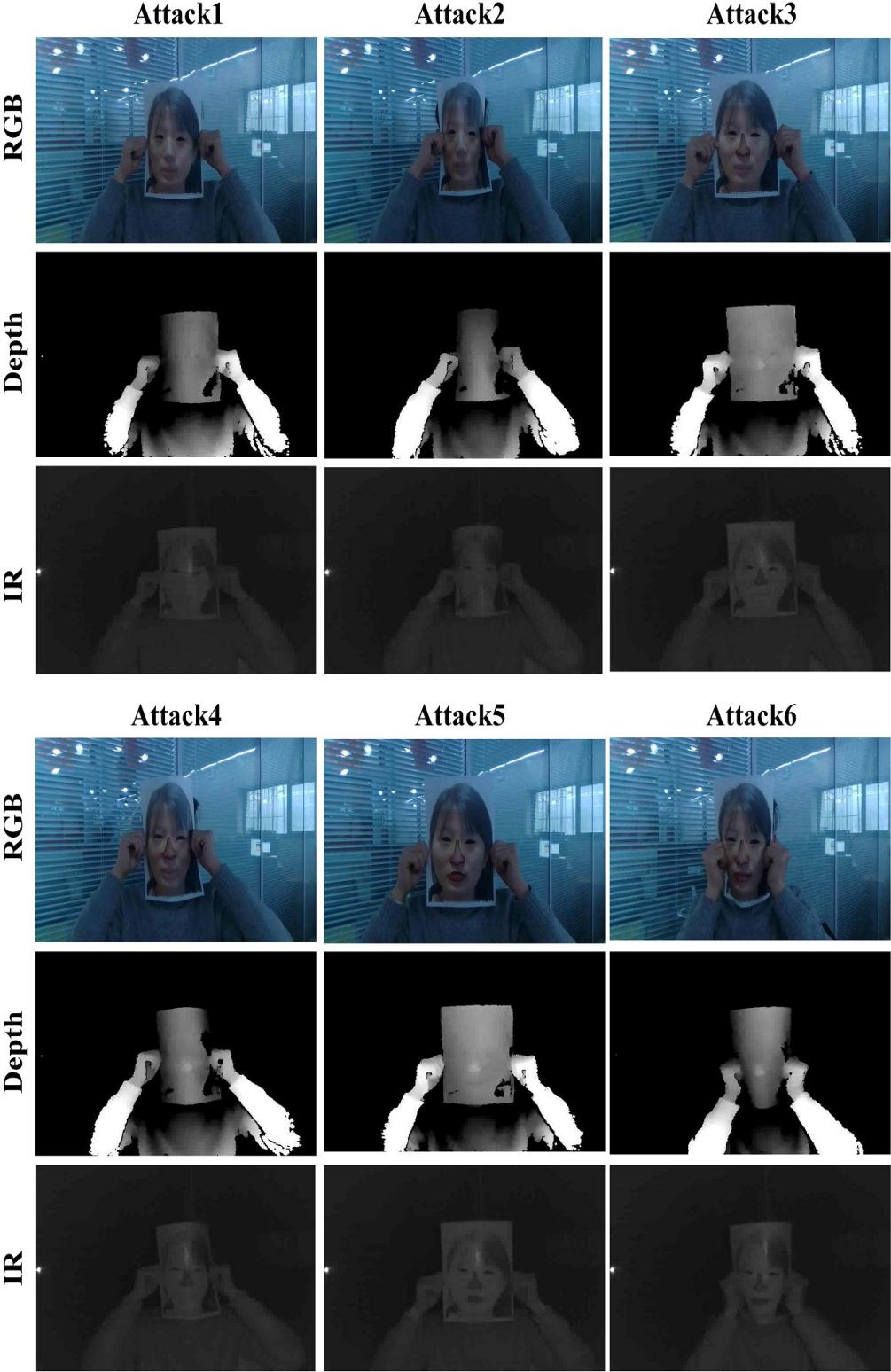


图 2-10 三种版本下的六种不同攻击方式

CASIA-Surf数据集还主动将整个数据分为了训练集（Training），验证集（Validation）和测试集（Testing），合计492522张人脸图像。由于本文只讨论红外图像，所以对CASIA-Surf数据集进行了简单的提取，将训练集，验证集和测试集中的红外图像分别提取出来，共计53531张图像。数据集示例如图2-11。

A picture containing view, old, display, train

Description automatically generated

图 2-11 CASIA-Surf 数据集中的红外图像

2.3.3自制红外数据集

由于NIR-VIS近红外图像中只有活体图像而没有攻击样本图像，所以为了使用NIR-VIS中的图像进行模型训练，本文使用红外摄像头自制了一个小规模的攻击样本数据集。

红外摄像头选用的是添加了850nm窄带滤光片的USB低照度红外摄像头。该摄像头自带红外灯，利用红外灯人为产生红外辐射，产生人眼看不见而摄像头能捕捉到的物体反射的红外光，从而获得被拍摄物体的图像。本文在拍摄攻击图像时，采取了一半在白天拍摄，一半在夜晚拍摄的策略，以使得图像样本的来源更加多元，从而增强模型的泛化能力。

由于条件限制，无法得到一定数量的真人纸张照片，本文使用杂志上的人物图像来作为攻击人脸样本，经过大量拍摄后得到240张可以识别出人脸的图像，分辨率与NIR-VIS保持一致，为640\*480，如图2-12所示。将这240张负样本与NIR-VIS中随机选取的240张图像结合，就得到了本文的第三个红外人脸数据集。

2.4本章小结

本章主要对两种活体检测分类方法的理论知识进行了详细介绍。首先是 SVM 的理论基础、公式推导以及优化方法-核方法；其次是CNN分类方法的理论基础和CNN的四个重要部分：卷积层、池化层、激活函数和全连接层。最后，本章还介绍了在本文中使用到的三个红外人脸数据集，包括两个公开数据集和一个自制攻击数据集。

A picture containing covered

Description automatically generated

图 2-12 自制攻击红外图像

第3章 基于SVM的红外活体检测模型训练

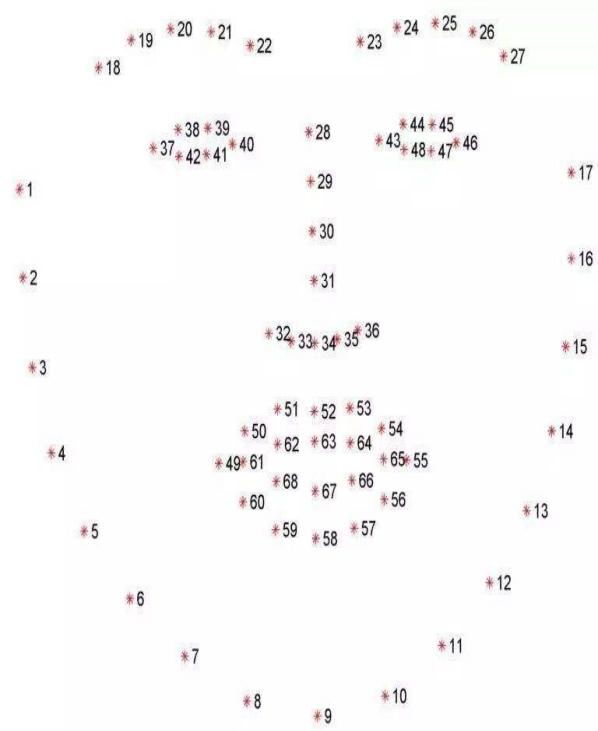
3.1图像预处理

在完成基于SVM的活体检测时，首先对输入图像进行了预处理操作，主要包括：对输入图像提取人脸；对人脸图像进行LBP二值模式提取[32]；在LBP图的基础上再提取GLCM 灰度共生矩阵的特征值[33]。本节对上述功能进行详细介绍。

3.1.1人脸提取

由于活体检测是为了给人脸识别做预处理工作[34]，所以在最开始应当对输入图像进行人脸提取的工作。如果没有提取到人脸，那么就不需要进行后续的分类工作，直接输出图像非活体。

人脸提取的成熟方法有很多，本文使用 dlib 开源库[35]中的人脸检测部分来完成人脸提取。dlib库基于方向梯度直方图[36]（Histogramof Oriented Gridients，HOG）来检测人脸，它会在一张图像上寻找68个人脸特征点，框定人脸，如图3-1(a)所示。得到人脸之后，将人脸之外的像素置为0，使其为黑色，以便背景像素不影响分类结果，如图3-1（b）。这样就完成了人脸提取的步骤。



1. dlib 68个特征点 （b） 人脸提取结果

图 3-1 dlib人脸检测示意

3.1.2 LBP 二值模式提取

LBP 局部二值模式图像是通过 LBP 算子对图像进行处理之后得到的特征图，它能够迅速地计算出图像的灰度特征。LBP 算子最早是由Ojala为了分析图像纹理特征在1996年提出[37]。作为一种非参数描述子，LBP对灰度变化不敏感且计算速度快，目前已经被广泛用于人脸检测、人脸识别、特征提取等技术当中。Ojala提出的 LBP 算子对于图像中每个像素点通过其紧邻的3\*3邻域像素值进行描述，计算公式为：

其中表示需要表示的像素点，表示其邻域8个像素。当Threshold取0时，LBP算法流程如图3-2。

A screenshot of a cell phone

Description automatically generated

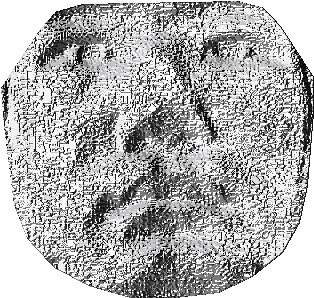
图 3-2 LBP 算法流程

对于中心像素点，将其邻域8个像素点分别与该中心像素点进行比较，如果邻域像素点灰度值大于中心像素点灰度值，就将该邻域像素点标记为1，否则标识为0。当这一过程结束后，按照一定的顺序将二值化的结果组成一个8位的二进制数，以该二进制数对应的整数来代表中心像素点的LBP值。当对整个图像完成这一操作之后，就实现了对图像局部二值模式的提取[38]。图3-3展示了对人脸图像提取LBP之后的效果。

3.1.3 GLCM灰度共生矩阵提取

灰度共生矩阵（Gray-Level Co-occurrence Matrix）是由Haralick于1973年定义的用于纹理分析的图像特征参数[9]，通过研究图像灰度的空间相关性来描述图像纹理特征。Haralick 对图像上相隔一定距离一定方向的两个像素的灰度值进行统计分析，来反映像素灰度值的空间分布关系，进而通过一系列特征参数来量化这种分布关系以显示图像的纹理信息。

假定两个灰度点i，j的空间距离为d，位置方向为θ。则表示当起始点灰度值为i时，终点灰度值为j的条件概率。对于一灰度层级为N的矩形图像区域，Haralick基于GLCM 定义了14种不同的纹理特征参数。本文选取了其中四种参数—均值（Mean）、逆差分矩（Inverse Difference Moment，IDM）、信息熵（Entropy）和角二阶矩（Angular Second Moment，ASM）—来确定一个图像的纹理信息。



（a） 原始人脸灰度图像 （b） LBP 处理结果

图 3-3 提取 LBP 效果展示

其中，均值的计算方式为：

逆差分矩的计算方式为：

信息熵的计算方式为：

角二阶矩的计算方式为：

其中，均值表征了图像的纹理变化信息，反映了纹理的规则程度，一般来说，纹理越规则有序，均值越大，而纹理越杂乱，均值就越小；逆差分矩又称同质度，反映了纹理的局部变化，其值越大，说明图像局部越均匀，变化较小；信息熵是对图像信息量的度量，反映了纹理的复杂程度，信息熵越大，图像越复杂；角二阶矩则反映图像灰度分布的均匀程度，角二阶矩越大，说明图像纹理越粗糙，反之则越细腻[39]。

3.2模型训练

在完成了对输入图像的预处理后，每张图像现在都由均值，逆差分矩，信息熵和角二阶矩四个纹理特征来唯一描述，对应到SVM中就是四维特征空间的一个特征向量。使用SVM对这些特征向量进行训练，找到最优“超平面”，即可实现分类模型的建立。

3.2.1训练过程

对SVM模型训练基于Python开源数据挖掘库sklearn，该库自带SVM分类算法，并提供了多个核方法以供使用，本文选择高斯核来训练，高斯核的函数在2.1.2节已经给出。同时使用sklearn的joblib模块完成对模型的保存，以便后面封装测试时使用。

由于公开人脸图像集NIR-VIS只有正样本（即真人图像），而没有负样本（即攻击图像），而 CASIA-Surf 多模态人脸数据集既有正样本也有负样本。所以在训练中将数据分为了两组，

第一组为 CASIA-Surf 的图像，第二组为 NIR-VIS的正样本加上 CASIA-Surf 的负样本。对这些图像进行LBP模式提取，得到LBP模式图之后再进行GLCM特征值提取，得到包含四个GLCM纹理特征值的csv文件。表3-1、3-2分别展示了对两个输入数据集提取到的特征值的部分结果。

表 3-1 CASIA-Surf 提取到的特征值表（局部）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean | IDM | Entropy | ASM | Live\_label |
| 10.36384651 | 0.424621397 | 4.346355915 | 0.030335502 | 1 |
| 10.23901922 | 0.360432046 | 4.613733707 | 0.051240167 | 1 |
| 12.23352011 | 0.656916236 | 2.979900367 | 0.27214661 | 0 |
| 11.37727273 | 0.622021635 | 3.224293211 | 0.225273603 | 0 |

表 3-2 NIR-VIS+CASIA-Surf 提取到的特征值表（局部）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean | IDM | Entropy | ASM | Live\_label |
| 9.765309364 | 0.451315239 | 4.590994164 | 0.050140437 | 1 |
| 9.392626357 | 0.463315557 | 4.476001913 | 0.060130942 | 1 |
| 12.23352011 | 0.656916236 | 2.979900367 | 0.27214661 | 0 |
| 11.37727273 | 0.622021635 | 3.224293211 | 0.225273603 | 0 |

3.2.2训练结果

将提取到的GLCM特征数据作为SVM分类的输入数据，对数据进行训练，得到训练结果如表3-3。可以看出两个数据集的分类训练准确率都在0.96以上，说明GLCM提取到的四个纹理特征值较好的反映了活体红外图像和非活体红外图像的区别。当然 NIR-VIS+CASIA-Surf 这一组训练数据的结果接近100%正确，准确率偏高，可能显示了在进行分类时，两个数据集图像由于拍摄设备不同，拍摄环境不同等非分类要素的差异对分类结果的影响。

表 3-3 SVM 训练结果

|  |  |  |
| --- | --- | --- |
| 数据集 | 分类准确率 | |
| 训练集 | 测试集 |
| CASIA-Surf | 0.96736 | 0.96423 |
| NIR-VIS+CASIA-Surf | 0.99593 | 0.99496 |

3.3本章小结

本章主要介绍了本文对公开红外人脸数据集进行SVM活体检测模型训练的过程。重点讲述了人脸提取、LBP局部二值模式提取和GLCM灰度共生矩阵特征值提取这三步图像预处理的方法和结果。同时也给出了训练的最终结果，训练结果总体显示使用SVM对红外人脸图像进行活体检测准确率较高，具有实践意义。

第4章 基于CNN的红外活体检测模型训练

4.1图像预处理

基于CNN分类方法的活体检测第一步仍然是对图像进行人脸提取以及去背景操作。同样是基于dlib开源人脸识别库来完成这一工作，具体可见3.1.1节。

在完成人脸提取之后，为了适应不同距离下采集的不同尺寸的人脸照片并加快神经网络的收敛性，还对人脸图像进行了归一化处理。具体来说，就是将所有的输入图像都归一化为 256\*256大小的图像。经过这一步之后，所有的图像矩阵大小都是256\*256，方便神经网络进行卷积时计算。

在完成了对人脸的提取以及图像大小的归一化之后，还需要将图像转为矩阵格式。在 Python中有多种方式可以实现将图像读取为矩阵形式，包括matplotlib开源绘图库；opencv的python接口；开源图像处理库Scipy；近乎事实上的官方图像处理库 PIL(Python Image Library)等。本文采取的是PIL中的Image类接口来读取图像，并通过其自带的convert函数将3通道图像转为单通道的8位像素图像。完成灰度转换之后，使用开源数值计算库numpy的array函数来获取灰度图的矩阵形式。获取到每个图像的矩阵形式之后，再将所有图像矩阵存储到一个新的3维矩阵中，得到一个N\*256\*256的矩阵，其中N是图像个数。

CNN图像预处理的最后一步是将N\*256\*256的图像矩阵数据存为二进制文件，由于这相当于保存多个数组到同一个文件中，所以使用 numpy的savez函数，将数据存为.npz 格式，方便后续处理。

4.2模型训练

本文基于 Tensorflow底层库来搭建 CNN 运行环境，使用开源人工神经网络库 keras来作为 Tensorflow的高阶应用接口。通过该接口对数据集进行训练，调整重要参数，获得训练模型。

4.2.1参数设置

在CNN分类器训练中，参数设置是最为重要的部分，在本文训练的CNN分类器中，有batch\_size、epoch、filters、pool\_size、kernel\_size五个重要参数，下面对其进行逐一说明。

batch\_size意为“批大小”，指的是每批数据量的大小，即一次一起训练batch\_size个样本，CNN 通过计算它们的平均损失函数值，来更新参数。最初在神经网络训练中，直接一次对全体训练集中的样本进行训练，计算损失函数值，更新权重，这样 batch\_size 即为样本数。但是随着训练数据增大，这种方法一次开销大，占用大量内存，运行速度过慢，而且 batch\_size 过大时，训练出来的模型泛化性不好[40]，容易收敛到一些不好的局部最优点。后来产生了一次只训练一个样本的方法—即batch\_size为1。这种方法会根据每一个样本的情况更新一次权值，系统开销小，计算速度快，但由于单个样本的随机性，使得全局优化能力较差，有时很难收敛到一个最优点[41]。现在的普遍做法是寻找一个折中的 batch\_size，使得其既具有很好的泛化性，又具有较好的全局优化能力。于是寻找一个合适的batch\_size相当重要。

epoch意为“世代”，1个epoch表示训练过了1遍训练集中的所有样本。通常一个epoch 中会进行A close up of a logo

Description automatically generated 次训练，其中N表示训练集中样本的数目。通常，我们会对样本数据集训练多次以使模型获得更好的泛化能力，所以一般训练 CNN 时会设置 epoch 为一个大于 1 的整数。但是同时，盲目增大epoch也不一定会能使神经网络模型更好。epoch会与batch\_size 一起影响CNN的分类结果，对epoch和batch\_size的选择需要结合具体数据和模型大小。本文在自制攻击图像和 NIR-VIS的联合数据集上测试不同 batch\_size和 epoch对分类准确率的影响（由于硬件设备是普通笔记本，训练时考虑到 CPU 的能力，epoch控制在20以内），结果如表4-1所示。

表 4-1 不同 epoch 和 batch\_size 的分类准确率

batch\_size

|  |  |  |  |
| --- | --- | --- | --- |
| 分类准确率  epoch | 5 | 10 | 15 |
| 8 | 0.9375 | 0.875 | 0.9875 |
| 16 | 0.9624 | 0.9625 | 0.975 |
| 32  batch\_size | 0.8625  batch\_size | 0.975  batch\_size | 0.9875  batch\_size |
| 64 | 0.9375 | 0.925 | 0.5 |

表 4-2 不同 epoch 和 batch\_size 的损失函数值

batch\_size

|  |  |  |  |
| --- | --- | --- | --- |
| 损失函数值  epoch | 5 | 10 | 15 |
| 8 | 0.289 | 0.102 | 0.119 |
| 16 | 0.285 | 0.135 | 0.315 |
| 32 | 0.697 | 0.101 | 0.109 |

续表

|  |  |  |  |
| --- | --- | --- | --- |
| 64 | 0.411 | 0.415 | 0.693 |

可以看出在batch\_size为8和epoch 为15以及 batch\_size 为32和epoch为 15 时模型分类准确率达到最大值 0.9875。另外也记录了模型损失函数值随 batch\_size 和 epoch 变化而变化的数据，如表 4-2所示。在分类准确率相同的情况下，batch\_size 为 32 和 epoch为15的损失函数值要比batch\_size为8和epoch为15小，说明其分类效果更好，泛化能力也更强，故本文对自制攻击图像数据集和240张NIR-VIS图像组成的联合数据集选取batch\_size为32，epoch为15。而采用同样的方法对 CASIA-Surf 数据集进行测试后，选取batch\_size为32，epoch为20。

filters定义的是卷积滤波器的数量，直观来说就是图像特征的个数。pool\_size定义的是池化层中池化窗口的大小，kernel\_size定义的是卷积滤波器的窗口大小。作为三个重要的超参数，对于这三个参数的确定目前也多是凭直觉和不断测试，没有严格的理论支撑来指导如何确定这三个参数。在本文中，根据相关文献和测试情况，结合自身的硬件水平，对两个数据集都选取filters为32，pool\_size为2\*2，kernel\_size为3\*3。

4.2.2训练结果

在设定好了4.2.1节提到的相关超参数之后，便可对数据集进行训练了。同基于SVM的模型一样，选取了两个数据集来训练，得到两个分类模型。第一个数据集是CASIA-Surf数据集，另一个则是自制攻击图像数据集和240张NIR-VIS图像的联合数据集。同时，为了给后面实际测试使用，还应当将训练好的模型保存。由于本文使用的是底层Tensorflow+keras接口的 CNN 运行环境，所以直接使用keras中的 model.to\_json()函数存储模型为序列化的.json格式。另外CNN训练得到权重值和重要参数也需要保存[42]，同样使用keras中的model.save\_weights函数将权重值存储为.h5格式。后面需要使用该模型时，只需加载这两个文件，编译模型，就可以使用了。

经过训练，最终得到训练结果如表4-3所示。可以看出，对两个数据集的分类准确率都在表 4-3 两个数据集的训练结果0.95以上，说明训练的模型效果较为优异。同时，由自制攻击图像数据集和240 张 NIR-VIS图像组成的联合数据集仍存在分类准确率偏高的问题。

表 4-3 两个数据集的训练结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 数据集 | batch\_size | epoch | 分类准确率 | 损失函数值 |

续表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CASIA-Surf | 32 | 20 | 0.957 | 0.135 |
| 自制攻击数据集+NIR-VIS | 32 | 15 | 0.937 | 0.103 |

4.3本章小结

本章介绍了使用CNN来训练红外活体检测模型的过程。同使用SVM来训练红外活体检测模型相似，也分为图像预处理和模型训练两部分。其中图像预处理较为简单，仅仅只要进行人脸提取和图像矩阵化两步。将提取到的人脸存为.npz格式后即可进行模型训练。在模型训练部分，CNN 最为重要的是参数设置，本章详细介绍了本文确定 CNN 分类模型中五个重要参数的过程。最后，本章给出了使用CNN来训练红外活体检测模型的结果。

第5章 基于两种模型的实际测试工作

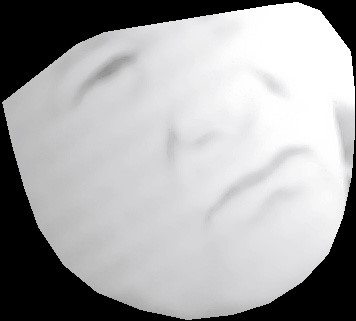
在完成两种方法的模型训练之后，本章将介绍使用训练好的模型进行实际测试的有关内容。测试环境均为Inteli5-8257U CPU @ 1.40GHz\*4，Intel Iris Plus Graphics 645 GPU，MacOS 10.15.4 (19E287)操作系统。

5.1 SVM 测试

如第3章所言，本文基于SVM训练了两种不同数据集下的模型。实践测试时对这两个模型分别进行了测试。

首先是针对基于 CASIA-Surf 训练的模型的测试。该模型在测试时分类准确率在 0.96左右。测试时使用真人和照片来检测该模型，最终获得的准确率只有0.72左右。对于使用NIR-VIS+CASIA-Surf训练的模型，测试的准确率也只有 0.75 左右，与模型训练准确率区别较大。

通过对测试结果的分析，发现基于SVM的红外活体检测模型对于照片攻击具有较强的识别能力，对于图5-1中的这种照片攻击都能有效抵御。



（a） 照片攻击测试示例 1 （b） 照片攻击测试示例 2

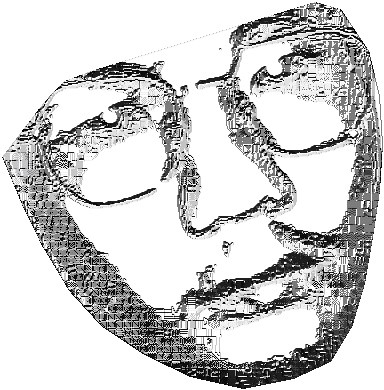
图 5-1 照片攻击示例

但是该模型的问题在于容易将真实人脸误识别为非活体。图5-2展示了三张典型的测试时误识别为非活体的真实人脸图像。经过分析，主要还是因为LBP模式的特征导致真实人脸被误识别。如果仅观察这三张原始图像，比较难理解为什么SVM分类模型会将它们判别为非活体，但如果结合LBP图就能够发现问题所在。本文在3.1.2节提到，LBP能够显示图像的灰度特征，从而反映出图像的纹理特征。对图 5-2的图形进行 LBP 局部二值模式图提取后，可以得到图 5-3。与之相对比，图 5-4给出了 NIR-VIS中人脸的 LBP 图以及测试时一个被模型判定为活体的人脸LBP图。



（a） 真实人脸图像 1 （b） 真实人脸图像 2 （c） 真实人脸图像 3

图 5-2 真实人脸示例



（a） LBP 图 1 （b） LBP 图 2 （c） LBP 图 3

图 5-3 测试人脸对应LBP图

对比图5-4和图5-3，可以很明显看出LBP图对分类结果的影响。

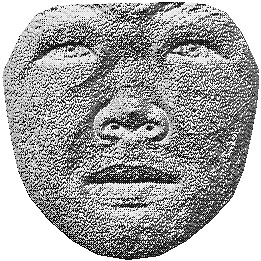
对于被误分为非活体的三张图像，它们的LBP图中都有大量空白，说明这些图像在很多区域中像素值没有显著变化，经过训练的模型认为这种现象是照片攻击的特征，所以将它们判定为非活体。而图图5-4中的两张LBP图则空白较少，甚至没有空白，说明这两张图纹理丰富，特征明显，故被模型判定为活体。

进一步研究发现，这种因为图像纹理不丰富而判定被检测者非活体的情况并不是模型的问题，因为基于SVM的活体检测模型本身就是通过学习活体与非活体的纹理特征来进行分类的。造成本次测试准确偏低的原因是本文使用的红外摄像头效果不佳，对红外灯的控制不够灵敏，导致有时被测者被大量红外光照射，形成如图5-2（a）中的过度曝光，从而产生测试图像纹理特征不明显，LBP图大量空白的问题[43]，进而导致测试准确率下降。此外，NIR-VIS数据集中包含了不同姿态的人脸图像，而自制的攻击数据集由于条件限制，无法获得多种姿态人脸图像，使得数据集整体不够完整，最后训练出来的模型对于多姿态人脸活体识别能力较差。

5.2 CNN测试

本文同样使用真人和照片攻击来测试基于CNN训练的两个红外活体检测模型。如第4章所述，两个基于CNN训练的红外活体检测模型分类准确率分别为0.957和0.987。经过测试，两个模型的实际分类准确率都在 0.9左右，其中基于自制攻击数据集和 NIR-VIS数据集的训练模型分类准确率为0.91，而基于CASIA-Surf的训练模型分类准确率为0.88。

。



（a） NIR-VIS人脸 LBP 图 （b） 被判定为活体的人脸 LBP 图

图 5-4 典型LBP图示例

经过分析，与摄像头相关的因素仍然是影响测试结果的重要原因。图5-5展示了四种典型的测试结果。其中5-5（b）、（c）被正确识别，而5-5(a）、（d）被错误识别。与上一节相似，可以看到两张被错误识别的图像都存在过度曝光的问题。虽然基于 CNN 的模型并没有使用 LBP 特征，但是观察如图5-6所示的NIR-VIS数据集和CASIA-Surf中的图像可以看出，用来训练模型两个数据集图像普遍不存在过度曝光的问题，这使得训练好的模型对于过度曝光的图像较为敏感，比较容易将这类图像识别为非活体。但是CNN在存在过度曝光的情况下，分类准确率依然达到了0.9，相比于SVM，效果要优异许多，说明了训练的模型泛化性能较佳，在特征提取时综合了很多因素，能够较好应对不同的实际环境。

5.3本章小结

本章介绍了本文的最后一个部分的工作：对训练好的红外活体检测模型进行测试。分别对基于SVM方法的模型和基于CNN方法的模型进行了测试，得出了测试结果。

测试结果显示，两个模型的实际应用分类准确率都要低于模型训练时的准确率。其中基于 SVM 方法的两个模型在测试中的准确率与训练时有较大差异，训练时准确率高达0.96 以上，但是实际测试只有0.7左右。经过分析对比，本文认为测试采用的红外摄像机对此具有主要责任，同时，数据集的拼接以及自制数据集不够完善也在一定程度上影响了训练结果偏高，测试结果偏低的现象。



（a）被错误识别为非活体的活体图像 （b）被正确识别为活体的活体图像



（c）被错误识别为活体的非活体图像 （d）被正确识别为非活体的非活体图像

图 5-5 四种典型测试结果



（a）NIR-VIS示例图像 （b） CASIA-Surf 示例图像

图 5-6 数据集示例图像

相比于基于SVM的分类模型，基于CNN的分类模型在测试中准确率与模型训练的分类模型在测试中准确率与模型训练时准确率较为接近，在0.9左右。这说明了CNN模型的泛化性能较好，同时通过对测试结果进行分析，也可以得出测试时使用的红外摄像头确实在红外灯的控制上存在一些问题。

另外，针对测试时拍摄的红外图像有时存在高光现象导致模型对真实人脸的误判，本文认为主要原因是在测试时使用的是主动式红外摄像头，通过主动发射红外光线并接受接收物体反射的红外光线来实现成像。而数据集中使用的红外摄像头效果更好，图像不存在高光现象，是因为数据集中拍摄时对环境要求较高，在实际应用中难以达到这种环境。后续可以在已有数据集的基础上，拍摄更为贴合实际应用的数据集图像，从而进一步增强模型的泛化能力。

结 论

本文主要对基于红外图像的活体检测进行了相关研究。活体检测是一个二元分类问题，本文采用了两种分类方法来进行活体检测，一个是 SVM 分类方法，另一个是 CNN 分类方法。

在对两种分类方法的理论进行细致研究之后，分别训练了基于两种方法的分类模型，并对模型进行测试，分析测试结果。

根据实际测试的结果，基于SVM分类方法的模型虽然训练时准确率较高，但是实际应用中准确率只有0.7左右，显示出该方法泛化性能不佳，不能较好地适应测试环境。基于 CNN 分类方法的模型虽然准确率也与训练时准确率有差异，但差异较小，说明训练得到的CNN模型泛化能力较强，提取到了更深刻的图像特征，能够应对测试环境的变化。

本文的工作还需要进一步完善的地方有：

1.由于时间和硬件设备的原因，没能制作一个独立的包含正负样本的大型红外人脸数据集，而在进行模型训练时使用的公开红外人脸数据集都不是很理想。后续应该在这方面做进一步的工作，制作出一个含正负样本的红外人脸数据集。

2. 在进行 CNN 训练时，由于本机GPU不支持相关神经网络框架，所以只能使用基于CPU的TensorFlow框架，所以运行能力受到限制。同时也不能对其他成熟的深度学习框架（如caffe）进行研究，也无法对 CNN 的底层做更深入的了解和优化。今后可以在这方面做更多的研究，加深对神经网络训练的深入理解。

致 谢

本文的完成得益于论文指导老师胡晓鹏老师的精心辅导，没有胡老师的日常督促和耐心解答，本文一定无法顺利完成。同时，胡老师严谨的治学态度和渊博的专业知识也深深地影响了我的论文写作和研究工作，让我能够以认真的姿态，严格的要求来完成这篇论文。在此，我想对胡老师表示真挚的感谢，感谢胡老师半年来的指导。

此外，我还想感谢我的父母，是他们给了我进入高等教育学习的机会，也是他们教会了我严格认真，努力向上的精神的。在论文写作的过程中，他们也一直陪伴在我身边，支持我，鼓励我。没有我的父母，就没有今日的我。

参考文献

1. 孙霖, 潘纲. 人脸识别中视频回放假冒攻击的实时检测方法[J]. 电路与系统学报, 2010, 15(02):39-46.
2. 谷小婧, 付传卿, 顾幸生. 基于面部生命特征的3D假面欺骗攻击检测方法[J]. 系统仿真学报, 2016, 28(02):361-368.
3. 孙霖. 人脸识别中的活体检测技术研究[D]. 博士论文. 浙江大学, 2010.
4. 张自友. 一种基于LBP和GLCM的单帧人脸近红外图活体检测方法[J]. 乐山师范学院学报, 2018, 33(12):5-12+73.
5. Galbally J, Marcel S, Fierrez J. Biometric Antispoofing Methods: A Survey in Face Recognition[J]. IEEE Access, 2014, 2:1530-1552.
6. Zhang Z, Yan J, Liu S, et al. A face antispoofing database with diverse attacks[C]. In Biometrics (ICB), 2012 5th IAPR International Conference on. IEEE, 2012.
7. Schwartz W R, Rocha A, Pedrini H. Face spoofing detection through partial least squares and low-level descriptors[C]. In International Joint Conference on Biometrics. IEEE, 2011.
8. Chingovska I, Anjos A, Marcel S. On the Effectiveness of Local Binary Patterns in Face Anti-spoofing[C]. In IEEE International Conference of the Biometrics Special Interest Group (BIOSIG). IEEE, 2012.
9. Haralick R M. Texture features for image classification[J]. IEEE Trans.syst.man Cybern, 1973, 3.
10. 杨健伟. 面向人脸识别的人脸活体检测方法研究[D]. 博士论文. 北京邮电大学, 2014.
11. Kollreider K, Fronthaler H, Bigun J. Non-intrusive Liveness Detection By Face Images[J]. Image & Vision Computing, 2009, 27(3): p. 233-244.
12. Lingxue S, Changsong L. Face Liveness Detection Based on Joint Analysis of RGB and Near-Infrared Image of Faces[J]. Electronic Imaging, 2018, 2018(10):3731-3736.
13. Sun L, Huang W B, Wu M. TIR/VIS Correlation for Liveness Detection in Face Recognition[C]. Computer Analysis of Images and Patterns - 14th International Conference, CAIP 2011, Seville, Spain, August 29-31, 2011, Proceedings, Part II. Springer-Verlag, 2011.
14. Cortes C, Vapnik V N. Support Vector Networks[J]. Machine Learning, 1995, 20(3):273-297.
15. Boyd, Vandenberghe, Faybusovich. Convex Optimization[J]. IEEE Transactions on Automatic Control, 2006, 51(11):1859-1859.
16. 赵洪波, 赵丽红. 支持向量机学习算法——序列最小优化(SMO)[J]. 绍兴文理学院学报(自然科学), 2003(10):25-28.
17. BOSER,Bernhard E. A Training Algorithm for Optimal Margin Classifiers[J]. proceedings of annual acm workshop on computational learning theory, 2008, 5:144--152.
18. Schölkopf, B, Smola A, Müller, K. Nonlinear Component Analysis as a Kernel Eigenvalue Problem[J]. Neural Computation, 1998, 10(5):1299-1319.
19. 叶浪. 基于卷积神经网络的人脸识别研究. 博士论文[D]. 东南大学, 2015.
20. David H. Hubel and Torsten N. Wiesel. Brain mechanisms of vision[J]. Scientific American,1979, 241(3):150–162.
21. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied todocument recognition. [J] Proceedings of the IEEE, 1998, 86(11):2278–2324.
22. Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm fordeep belief nets[J]. Neural Computation, 2006, 18(7):1527–1554.
23. Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedfor-ward neural networks[J]. Journal of Machine Learning Research, 2010, 9:249–256.
24. 李俭川, 秦国军, 温熙森, 胡茑庆. 神经网络学习算法的过拟合问题及解决方法. 振动[J]. 测试与诊断, 2002(4):16–20+76.
25. Jun Han and Claudio Moraga. The influence of the sigmoid function parameters onthe speed of backpropagation learning[C]. In International Workshop on from Natural toArtificial Neural Computation, 1995.
26. Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmannmachines vinod nair[C]. In Proceedings of the 27th International Conference on MachineLearning (ICML-10), June 21-24, 2010, Haifa, Israel, 2010.
27. M.D. Zeiler, M. Ranzato, R. Monga, M. Mao, and G. E. Hinton. On rectified linear unitsfor speech processing[C]. In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEEInternational Conference on, 2013.
28. Soltau, Hagen, Dahl, George, Kingsbury, Brian, Sainath, Tara, N., and Saon and. Deepconvolutional neural networks for large-scale speech tasks[J]. Neural Networks the OfficialJournal of the International Neural Network Society, 2015.
29. Alex Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep con-volutional neural networks[J]. Advances in neural information processing systems, 2012,25(2).
30. Stan Z. Li, Yi Dong, Lei Zhen, and Shengcai Liao. The casia nir-vis 2.0 face database[C]. InIEEE Conference on Computer Vision Pattern Recognition Workshops, 2013.
31. Ajian Li, Zichang Tan, Xuan Li, Jun Wan, Sergio Escalera, Guodong Guo, and Stan ZLi. Casia-surf cefa: A benchmark for multi-modal cross-ethnicity face anti-spoofing[J]. 2020.
32. 程雪峰. 基于LBP特征的人脸识别算法研究[D]. 博士论文. 厦门大学, 2014.
33. 薛翠红, 于洋, 张朝, 杨鹏, 李扬. 融合lbp与glcm的人群密度分类算法[J]. 电视技术, 2015, 039(24):7–10.
34. 汤德俊. 人脸识别中图像特征提取与匹配技术研究[D]. 博士论文. 大连海事大学, 2013.
35. Davis E King. Dlib-ml: A machine learning toolkit. Journal of Machine Learning Re-search[M], 2009, 10(Jul):1755–1758.
36. Xiaoyu Wang, Tony X. Han, and Shuicheng Yan. An hog-lbp human detector with partialocclusion handling[C]. In IEEE International Conference on Computer Vision, 2009.
37. Timo Ojala, Matti PietikInen, and David Harwood. A comparative study of texturemeasures with classification based on featured distributions[J]. Pattern Recognition, 1996,29(1):51–59.
38. 赵玉丹. 基于LBP的图像纹理特征的提取及应用[D]. 博士论文. 西安邮电大学, 2015.
39. 薄华, 马缚龙, 焦李成. 图像纹理的灰度共生矩阵计算问题的分析[J]. 电子学报, 2006, 34(1):155–158.
40. Priya Goyal, Piotr Dollár, Ross Girshick, Pieter Noordhuis, Lukasz Wesolowski, AapoKyrola, Andrew Tulloch, Yangqing Jia, and Kaiming He. Accurate, large minibatch sgd:Training imagenet in 1 hour[J]. 2017.
41. Nitish Shirish Keskar, Dheevatsa Mudigere, Jorge Nocedal, Mikhail Smelyanskiy, and Ping Tak Peter Tang. On large-batch training for deep learning: Generalization gap andsharp minima[J]. 2016.
42. 杨巨成, 刘娜, 房珊珊,等. 基于深度学习的人脸识别方法研究综述[J]. 天津科技大学学报, 2016(6).
43. 宁星. 基于LBP的人脸识别研究[D]. 博士论文. 东华理工大学, 2012.