

广东工业大学硕士学位论文

(工程硕士)

基于深度学习的非接触式掌静脉识别技术研究

刘彪

二〇一九年五月

分类号:

学校代号: 11845

UDC:

密级:

学 号: 2111601311

广东工业大学硕士学位论文

(工程硕士)

基于深度学习的非接触式掌静脉识别技术研究

刘彪

校内导师姓名、职称: 谢康 教授

校外导师姓名、职称: 余孟春 工程师

学科(专业)或领域名称: 仪器仪表工程

学 生 所 属 学 院: 机电工程学院

论 文 答 辩 日 期: 2019 年 5 月

A Dissertation Submitted to Guangdong University of Technology
for the Degree of Master

(Master of Engineering)

Research on Contactless Palm Vein Recognition
Technology Based on Deep Learning

Candidate: Liu Biao

Supervisor: Prof. Xie Kang

May 2019

School of Electromechanical Engineering

Guangdong University of Technology

Guangzhou, Guangdong, P.R. China, 510006

摘要

随着当今经济和信息技术快速发展，如何有效地提升公民身份信息安全显得尤其重要。掌静脉识别技术是一种安全性高、防伪性高的生物特征识别技术，在未来将会有广阔的应用空间。然而，目前的掌静脉识别算法大都基于人工设计的特征进行识别，这些特征的鲁棒性不高，对图像质量以及手掌姿态等较为敏感；而且，目前掌静脉识别的研究普遍基于接触式设备采集的图像进行，对非接触式掌静脉识别的研究较少。非接触方式采集的掌静脉图像往往具有较大的形变，如手掌弯曲、手掌倾斜或者手掌过度伸展等，而且光照也不够均匀，因此具有更高的识别难度。但与接触式相比，非接触式的采集方式具有更好的用户体验，其应用前景更加广阔。

针对上述问题，本文采用了深度学习方法，对非接触式掌静脉识别过程中的手掌关键点定位问题和掌静脉图像的特征提取问题进行研究，取得了较好的提升效果。本文的主要研究内容如下：

1) 针对传统算法难以对非接触式手掌图像的关键点进行准确定位的问题，本文提出采用深度学习的方法进行手掌关键点定位，设计了两级卷积神经网络级联来定位手掌关键点；并且为了提高手掌关键点定位的精确度，对 MSE 损失函数进行改进，提出了 Modified MSE Loss 进行训练，有效地提升了手掌关键点的定位精度；

2) 针对非接触式掌静脉数据库样本较少，训练神经网络容易产生过拟合的问题，提出了类内扩增数据和类间扩增数据的方法，有效地扩充了原始数据库，为后续掌静脉特征提取网络的训练提供了数据支撑；

3) 对经典分类神经网络结构的设计原理进行分析后，自行设计了两个深层特征提取网络 Resnet18-Modified 和 VGG16-Modified，并对 Resnet18 网络的 Residual Block 模块结构进行改进，实验结果证明了改进后的 Residual Block 模块可以有效地提升识别精度；另外，探索了不同特征向量维度对识别精度的影响，实验证明，在大型深层网络上，512 维的特征向量精度最高，而在轻量级网络上，256 维的紧凑型特征向量精度更高；

4) 针对传统深层神经网络参数量大，计算耗时长的问题，本文设计了一个轻量级网络 TinyPVNet，主要采用深度可分离卷积结构，有效地降低了网络参数量和模型推理时间，并在 CASIA 数据集上取得了 0.51% 的等误率；另外，本文还对轻量级网络 TinyPVNet 进行了模型量化，有效地压缩了模型体积，从而能够将轻量级网络

部署在嵌入式等计算资源受限的平台上。

关键词：掌静脉识别；深度学习；关键点定位；特征提取

ABSTRACT

With the rapid development of the current economy and information technology, how to effectively improve the security of citizenship information is particularly important. The palm vein recognition technology is a biometric recognition technology with high safety and high anti-counterfeiting, and will have a broad application space in the future. However, most of the current palm vein recognition algorithms are based on artificially designed feature for recognition. These features are not robust and sensitive to image quality and palm posture. Moreover, the current research on palm vein recognition is generally based on images which acquired from contact devices. And there were few studies on contactless palm vein recognition. The palm vein image which collected from contactless devices often has large deformation, such as palm bending, palm tilt or palm over stretching, and the illumination is not uniform enough, so it has higher recognition difficulty. However, compared with the contact type, the contactless collection method has a better user experience, and its application prospect is broader.

In view of the above problems, this paper adopts the deep learning method to research the problem of contactless palm image landmark localization and feature extraction, and has achieved good improvement effect. The main research contents of this paper are as follows:

1) For the problem that the traditional algorithm is difficult to accurately locate the landmark of the non-contact palm image, this paper proposes a deep learning method to locate the landmark of the palm, and designs a two-level cascade convolutional neural network to locate the landmark of the palm; In order to improve the accuracy of the landmark localization, we improved the MSE loss function and proposed Modified MSE Loss function for training, which effectively improves the landmark localization accuracy of the palm;

2) Owing to the contactless palm vein database has fewer samples, training neural network is prone to over-fitting. We proposed intra-class data augmentation and inter-class data augmentation method, which effectively expands the original contactless palm database and provide support for training the palm vein feature extraction network;

3) After analyzing the design principle of the classical classification neural network structure, we designed two deep feature extraction networks which named Resnet18-Modified and VGG16-Modified, then improved the Resnet Residual Block module structure. The experimental results prove that the improved Residual Block module can effectively improve the recognition accuracy. In addition, we also explored the influence of different feature vector dimension on accuracy. Experiments show that the 512-dimension feature vectors achieve the highest accuracy on large deep network, while 256-dimension compact feature vectors have higher precision on lightweight network ;

4) In view of the problem that the traditional deep neural network has large parameters and high calculation consume, this paper designs a lightweight network named TinyPVNet, which mainly uses deep separable convolution structure, which effectively reduce network parameters and model inference time. we achieved 0.51% equal error rate in the CASIA dataset. In addition, we effectively compresses the TinyPVNet model volume by quantifying weight, and enabling the deployment of lightweight networks on platforms with limited computing resources such as embedded platform.

Keywords: Palm vein recognition; Deep learning; Landmark localization; Feature extraction

目录

摘要.....	I
ABSTRACT.....	III
目录.....	V
CONTENTS.....	VIII
第一章 绪论.....	1
1.1 研究背景及意义.....	1
1.2 静脉识别的国内外研究现状及难点.....	3
1.2.1 手指和手背静脉识别的研究现状.....	3
1.2.2 手掌静脉识别研究现状.....	4
1.2.3 非接触式掌静脉识别的研究难点.....	6
1.3 本文创新点.....	6
1.4 论文的结构安排及主要研究工作.....	7
第二章 掌静脉图像的获取及预处理.....	9
2.1 手掌静脉图像采集.....	9
2.2 掌静脉数据库介绍.....	10
2.3 掌静脉图像 ROI 选取.....	11
2.3.1 手掌关键点选择.....	11
2.3.2 图像旋转校正.....	12
2.3.3 掌静脉 ROI 选取及尺度归一化.....	13
2.4 掌静脉图像增强.....	14
2.4.1 直方图均衡.....	14
2.4.2 CLAHE(对比度限制自适应直方图均衡).....	16
2.5 掌静脉图像去噪.....	18
2.5.1 均值滤波.....	19
2.5.2 中值滤波.....	20
2.6 本章小结.....	21
第三章 手掌关键点定位网络.....	22
3.1 手掌关键点定位网络设计.....	22

3.1.1 粗略定位网络.....	22
3.1.2 精确定位网络.....	23
3.2 关键点回归损失函数设计.....	24
3.3 实验数据准备及结果分析.....	25
3.3.1 手掌图像数据集扩增.....	25
3.3.2 实验结果分析.....	26
3.4 本章小结.....	29
第四章 基于卷积神经网络的掌静脉特征提取模型.....	31
4.1 深层掌静脉特征提取网络.....	31
4.1.1 特征提取网络的输入输出维度设定.....	31
4.1.2 深层卷积神经网络结构介绍.....	32
4.1.3 深层掌静脉特征提取网络设计.....	34
4.2 轻量级掌静脉特征提取网络.....	37
4.2.1 轻量级神经网络结构介绍.....	37
4.2.2 轻量级掌静脉特征提取网络设计.....	39
4.3 损失函数设计.....	40
4.3.1 Softmax Loss.....	41
4.3.2 Center Loss.....	42
4.3.3 L-softmax Loss 和 A-softmax Loss.....	43
4.3.4 Additive Angular Margin Loss.....	45
4.4 特征向量的相似性度量方法.....	46
4.5 模型压缩.....	47
4.6 本章小结.....	48
第五章 实验结果及分析.....	49
5.1 实验数据准备.....	49
5.1.1 掌静脉数据扩增.....	49
5.1.2 实验数据准备.....	50
5.2 实验性能指标.....	52
5.2.1 类内比对和类间比对.....	52
5.2.2 误识率、拒识率和等误率.....	52

5.3 实验设置及模型训练参数设置.....	53
5.4 实验结果分析.....	53
5.4.1 不同特征向量维度的性能对比.....	53
5.4.2 不同网络结构的性能对比.....	55
5.4.3 改进 Residual Block 对算法精度的提升.....	56
5.4.4 不同损失函数的性能对比.....	57
5.4.5 与其他算法性能对比.....	58
5.4.6 模型压缩及网络时间测试.....	59
5.5 本章小结.....	59
总结与展望.....	61
参考文献.....	63
学位论文独创性声明.....	67
致谢.....	68

CONTENTS

ABSTRACT(Chinese)	I
ABSTRACT(English)	III
CONTENTS(Chinese)	V
CONTENTS(English)	VIII
Chapter 1 Introduction	1
1.1 Research background and significance.....	1
1.2 Research status and difficulties of vein recognition at domestic and foreign.....	3
1.2.1 Research status of finger and hand vein recognition.....	3
1.2.2 Research status of palm vein recognition.....	4
1.2.3 Research difficulties on contactless palm vein recognition.....	6
1.3 The article innovation.....	6
1.4 The structure and the main research work of the article.....	7
Chapter 2 Acquisition and preprocessing of palm vein images	9
2.1 Palm vein image acquisition.....	9
2.2 Palm vein database introduction.....	10
2.3 Palm vein image ROI selection.....	11
2.3.1 Palm landmark selection.....	11
2.3.2 Image rotation and correction.....	12
2.3.3 Palm vein ROI selection and image scale normalization.....	13
2.4 Palm vein image enhancement.....	14
2.4.1 Histogram equalization.....	14
2.4.2 CLAHE(Contrast-limited adaptive histogram equalization).....	16
2.5 Palm vein image denoising.....	18
2.5.1 Average filtering.....	19
2.5.2 Median filtering.....	20
2.6 chapter summary.....	21
Chapter 3 Palm landmark localization network	22
3.1 Palm landmark localization network design.....	22
3.1.1 Rough landmark localization network.....	22

3.1.2 Precise landmark localization network.....	23
3.2 landmark localization regression loss function design.....	24
3.3 Experimental data preparation and results analysis.....	25
3.3.1 Palm image data augmentation.....	25
3.3.2 Analysis of results.....	26
3.4 chapter summary.....	29
Chapter 4 Palm vein feature extraction model based on convolutional neural network.....	31
4.1 Deep palm vein feature extraction network.....	31
4.1.1 Feature extraction network input and output dimension settings.....	31
4.1.2 Introduction to deep convolutional neural network structure.....	32
4.1.3 Deep palm vein feature extraction network design.....	34
4.2 Lightweight palm vein feature extraction network.....	37
4.2.1 Introduction to lightweight neural network structure.....	37
4.2.2 Lightweight palm vein feature extraction network design.....	39
4.3 Loss function design.....	40
4.3.1 Softmax Loss.....	41
4.3.2 Center Loss.....	42
4.3.3 L-softmax Loss 和 A-softmax Loss.....	43
4.3.4 Additive Angular Margin Loss.....	45
4.4 Feature vector similarity measure.....	46
4.5 Model compression.....	47
4.6 chapter summary.....	48
Chapter 5 Experimental results and analysis.....	49
5.1 Experimental data preparation.....	49
5.1.1 Palm vein data augmentation.....	49
5.1.2 Experimental data preparation.....	50
5.2 Experimental performance index.....	52
5.2.1 Intra-class comparison and inter-class comparison.....	52
5.2.2 FAR、FRR and EER.....	52

5.3 Experimental setup and model training parameter settings.....	53
5.4 Analysis of results.....	53
5.4.1 Performance comparison of different feature vector dimensions.....	53
5.4.2 Performance comparison of different network structures.....	55
5.4.3 Improve the accuracy of the algorithm by improving the residual block	56
5.4.4 Performance comparison of different loss functions.....	57
5.4.5 Performance comparison with other algorithms.....	58
5.4.6 Model compression and network time testing.....	59
5.5 chapter summary.....	59
Conclusions and prospects.....	61
References.....	63
Dissertation originality statement.....	67
Acknowledgements.....	68

第一章 绪论

1.1 研究背景及意义

随着当今经济和信息技术快速发展，如何有效地提升公民身份信息的安全成为一项重要的研究课题。以前，人们常用一些证件(如身份证)、钥匙、密码锁等进行身份识别，这些身份识别技术普遍存在容易被盗用的缺点。生物特征识别技术是利用人体生物特征进行身份认证的一种技术，主要包括人类的各种生理和行为特征，如人脸、指纹、掌静脉、虹膜、声纹等。人体生物特征由于其稳定性和独特性，因此成为最理想的身份识别特征。

目前，人脸识别和指纹识别已经广泛地应用在人们的日常生活中，如公司及住宅门禁系统、安防监控系统、人流量较大的地铁、机场、火车站等领域。这两种生物识别技术在给人们带来便利的同时，也带来较大的安全隐患：指纹位于人体皮肤的表层，容易被伪造和泄露；人脸容易受年龄、表情、姿态、周围光照等影响，进而导致识别困难、识别率不高等问题，而且双胞胎也是困扰人脸识别的问题之一。在理想的环境中，语音识别能达到比较高的识别精度，但如果在地铁、火车站等环境嘈杂的区域，语音识别的效果就会大幅下降；虹膜在胎儿发育阶段形成后，在整个生命历程中将是保持不变的，因此虹膜识别曾被认为是最安全的身份识别技术，但其需要人眼正视采集装置，容易让用户产生抵触心理，用户体验性较差。近年来，指静脉识别也取得了很大的进步，但相比于掌静脉而言，指静脉含有的有效特征信息较少，因此其安全性还有待考验。

手掌静脉识别是对人体的手掌皮肤使用近红外光源进行照射，利用血液中血红蛋白对近红外光具有较强的吸收特性，使得手掌静脉处在近红外光照射后会出现较暗的纹路^[1]，从而利用该静脉特征进行身份识别。人体的手掌静脉具有唯一性、稳定性，隐藏在手掌表皮下，结构复杂故难以复制，且静脉特征在可见光下无法获取，具有较强的隐蔽性和防伪性，同时还具有活体验证功能（只有活体手掌才具有静脉特征），这些特点使得手掌静脉成为一种高安全性的生物特征，因此具有极大的研究价值，其实际应用将带来巨大的产业效益。相比于上述所提到的生物特征识别技术，掌静脉识别技术有以下一些优点：

- 1) 具有很强的普遍性和唯一性；

- 2) 非接触性方式认证, 不影响身体健康, 不会使人产生抵触情绪;
- 3) 与 DNA 识别和虹膜识别技术相比, 整个采集过程具有更好的用户体验;
- 4) 相比人脸、指纹等, 更难伪造, 特征更为稳定;
- 5) 设备成本低廉, 具有广泛的应用前景;
- 6) 容易与掌纹融合进行身份识别, 应用前景更为广阔。

由于人体静脉具有稳定性和安全性等独特性优势, 近年来被研究的越来越深入。掌静脉的采集与识别方式分为接触式与非接触式两种。目前, 国内的掌静脉研究数据库大多是基于接触式方式采集的图像, 其采集的掌静脉图像光照均匀, 而且手掌不会产生很大的偏移和旋转等, 因此能够达到比较高的识别精度。图 1-1 显示的是采用接触式方式采集手掌静脉图像。可以看到, 由于接触式装置一般都安装了辅助定位装置, 因此采集出来的图像光照均匀、偏移量小等, 进而容易取得较高的识别精度。

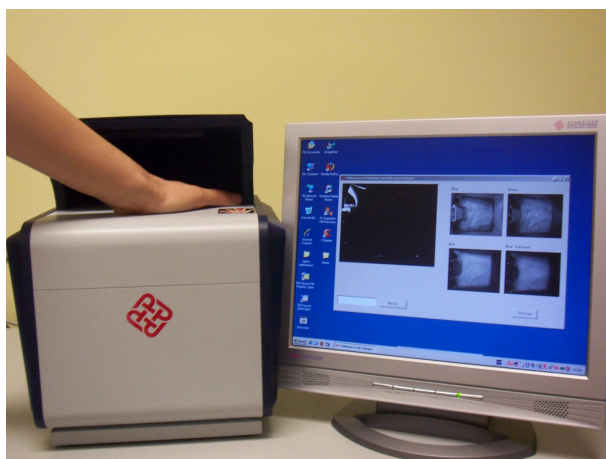


图 1-1 接触式手掌静脉采集装置^[2]

Fig.1-1 Contact palm vein collection device^[2]

与接触式采集的图像相比, 采用非接触式方式采集的手掌图像往往具有较大的形变, 如手掌弯曲、手掌倾斜或者手掌过度伸展等, 而且每次采集时手掌产生的偏移都不一样, 光照也不够均匀。因此, 使用非接触式方式采集的掌静脉图像, 采用目前的识别算法难以得到较高的识别精度。但使用非接触式方式采集和识别时, 用户的手掌不需要与采集设备接触, 整个过程干净卫生, 不会让用户产生抵触心理, 因此非接触式掌静脉识别具有更广阔的应用领域, 如地铁、公交、智能交通、智慧安防、图书馆等。综上, 本文所研究的内容具有实际的应用价值及现实意义。

1.2 静脉识别的国内外研究现状及难点

目前,静脉图像识别的主要研究分布在手指静脉、手掌静脉以及手背静脉领域。其中,手指和手背由于皮下皮肤比手掌皮肤薄,静脉更容易获取而被广泛研究,而手掌静脉由于其不易获取则研究的相对较少。

1.2.1 手指和手背静脉识别的研究现状

手背静脉由于比较容易获取,因此也较早地进入到科研人员的研究范围。韩国汉城大学电子工程系的研究人员首先提出了一种手背静脉识别算法^[3],探讨了手背静脉的图像采集、图像预处理以及图像匹配的流程,其采用了点对点匹配的方法,静脉识别准确率达到 94.8%;美国的 Badawi A.M.^[4]首次提出了用 3×3 的模板对手背静脉图像进行中值滤波,由于静脉匹配采用的是对图像的每个像素进行比较的方法,因此非常耗时,但精度提升比较明显,在 1500 张手背静脉图像上取得了 FRR 为 2.88%, FAR 为 0.16% 的优异成绩;新加坡南洋理工大学的 Wang L 和 Leedham G^[5]采集了 108 张手背的红热图像,通过对图像进行二值化、细化、中值滤波去噪等操作,然后进行匹配识别,得到了 FRR 和 FAR 都是零的结果。国内方面,清华大学的林喜荣、庄波等人利用自行设计的近红外血管图像采集仪提取了血管的原始图像,通过提取图像的点特征来进行匹配识别,文献^[6]给出了 65 个样本下的匹配结果,在拒识率为 4.6% 的条件下,误识率为 0%;吉林大学的韩笑等^[7]将神经网络和小波变换用于图像的二值化和增强,并提出了多分辨脊波特征匹配法和基于 FRAT 的静脉纹理方向特征匹配法,有效地提升了手背静脉图像的识别精度;哈尔滨工程大学王科俊、庄大燕、丁宇航等人对手背静脉识别也做了大量研究^{[8][9]}。在产品方面,韩国的 NEXTERN 公司研发出了 BK100、BK200、BK300、BK500 等手背静脉识别产品,已经成功应用在门禁系统、医疗等方面。

手指静脉方面,韩国的 Pham T D.^[10]等人提出了基于局部二元编码来进行指静脉特征提取,其通过计算静脉编码后的汉明距离进行特征的相似性匹配,取得了较好的识别效果。国内对手指静脉的研究则相对较晚,但也取得了一些不错的成果,发展速度较快。重庆工学院的余成波^[11]等人主要对指静脉的特征提取进行了相关探索,提出了一种 KFDA 方法进行静脉特征提取,解决了图像灰度不均问题带来的影响;哈尔滨工程大学王科俊教授带领的团队对指静脉识别技术进行了比较深入的研

究，其先后提出了基于 2DPCA 方法^[12]、基于相对位置和角度的静脉识别方法^[13]以及基于指纹和指静脉多特征融合的方法^[14]来进行静脉图像识别，取得了一系列的成果；华南理工大学的李德健^[15]提出了一种匹配主特征和软生物特征的混合匹配方法，有效地解决了手指旋转时手指静脉匹配不佳的问题；同团队的胡慧^[16]、唐溯^[17]等提出了基于深度学习的指静脉识别算法，通过使用卷积神经网络（CNN）直接对指静脉图像进行特征提取，取得了非常好的识别效果。在产品方面，最早开始进行研究的是日本日立公司的中央研究所，其指静脉识别相关的产品也已进入市场销售，2003 年，日立公司首次将指静脉识别应用到门禁系统、软件设计中心等。

1.2.2 手掌静脉识别研究现状

目前，国内外已经有一批高校在对手掌静脉识别等相关技术进行研究，并取得了一些不错的成果，如香港理工大学、华南理工大学、电子科技大学等，其对整个掌静脉识别技术的快速发展起到了巨大的推动作用。下面将对一些代表性成果进行简要介绍。

吉林大学的李铁刚博士^[18]在 Bandelet 理论和小波分析的基础上，研究了图像的多尺度几何分析问题，并提出了基于 Bandelet 变换的掌背静脉识别算法；此外，还深入研究了复值二维 Gabor 变换、相位编码和矢量线特征表示等静脉识别算法，并将这些算法应用在手掌静脉识别上，取得了较好的识别效果。但是相比于其他静脉识别算法，其提出的基于 Bandelet 变换的识别算法具有很高的计算复杂度，并且只用了少量的数据进行实验，因此算法的鲁棒性仍有待验证。

华中科技大学的李强^[19]博士对掌静脉图像的获取方法作了深入的研究，其搭建了被动式热红外成像系统和主动式近红外成像系统，其中热红外成像技术用于获取手背静脉图，最终采用 CLAHE 算法和多分辨曲波变换算法进行图像预处理。

北京邮电大学的马欣^[20]博士在手掌静脉身份识别关键技术研究中对图像获取与增强、自适应局部纹理提取、多模态生物特征信息融合、相关性分析以及稀疏编码等多个方面提出了优化算法和策略，使识别系统的可靠性得到了提高。

华南理工大学的颜学葵^[21]博士在掌静脉的 ROI 提取、掌静脉图像增强方面提出了一些有效的算法，并提出了基于 ORB 特征提取的识别算法和基于 AGAST 特征提取和 BRIEF 描述的识别算法，获得了较好的识别效果；另外，采用多采样和特征级融合的算法，在手掌姿态变化较大的情形下也能获得较高的识别率。

电子科技大学的徐红兵、凡时财教授等人带领的团队也对掌静脉识别算法进行了大量的深入研究^{[1][22][23]}。另外，国内其他高校的研究人员也对掌静脉识别技术领域的发展贡献了不可磨灭的力量。

在掌静脉识别产品方面，日本富士通公司于 2005 年 2 月成功研发出了第一套掌静脉识别产品^[24]，在日本图书馆进行了实际应用，在人数低于 1000 时，识别结果全部正确。在 2008 年，该公司实现了在 FRR 为 0% 的条件下，FAR 小于 0.008% 的卓越成绩。

纵观掌静脉识别技术的发展，主要有以下几种方法：

1. 基于结构特征的方法：主要根据静脉图像骨架的几何结构特征来识别，如点、线特征，利用这些点进行匹配、识别，这类方法的算法原理相对简单，计算复杂度不高，但识别精度不是很高；

2. 基于局部不变性特征方法：如 SIFT、ASIFT、SURF 等特征方法，这类方法对图像的平移、旋转以及尺度变换的鲁棒性较高，能取得不错的识别效果。

3. 基于纹理特征分析的方法：其主要是采用频谱方法来获取静脉图像的频率、方向、幅度以及相位等纹理特征，然后再对这些特征进行编码和识别。这是目前应用较多的静脉特征提取方法。

4. 基于子空间的方法：此方法是通过将高维的掌静脉图像进行投影或变换成低维的矩阵，然后在低维空间上进行掌静脉的表示和匹配。由于对图像空间进行降维操作，因此计算速度相对其他方法有显著优势，但是这种方法的鲁棒性一般，对图像的光线变化等比较敏感；

5. 基于统计特征的方法：通过提取静脉图像的全局统计特征进行识别，如 LBP、LDP 方法，这类静脉提取方法易于表征，但是会损失局部特征信息，因此识别性能不高。

上述几类特征提取方法性能各有优劣，针对这些特征提取方法的不足，现如今越来越多的科研人员也开始采用多特征融合的方法来提升掌静脉识别的整体精度，如 X Yan 和 F Deng 在文献^[25]中提出，通过将掌静脉图像的 SIFT 特征匹配分数和 ORB 特征的匹配分数进行评分级融合，在中科院的 CASIA 多光谱掌静脉数据库上取得了 0.36% 的最低等误率；后来 X Yan 和 W Kang 等人在文献^[26]中又提出了多采样和多特征融合的识别方法，通过融合多个掌静脉图像的特征，在 CASIA 掌静脉数据库上取得了 0.16% 的等误率；杨园园在基于多特征的掌静脉识别算法^[23]中提出了

一种融合双手掌 NBP 特征的识别算法,并在此基础上又提出两种多登记样本图时的双手掌融合识别算法等,因此通过掌静脉的多特征进行决策融合也成为目前提高掌静脉识别精度的一个研究热点。

1.2.3 非接触式掌静脉识别的研究难点

通过 1.2.2 节的论述可以看出,目前,国内的掌静脉技术研究大部分都是基于接触式设备采集的掌静脉图像来进行。接触式方式采集的图像给图像的预处理提供了便捷,同一个人的手掌图像相对稳定,不会产生较大的偏移和形变,因此能够取得比较理想的识别精度。但接触采集的方式不太好应用到实际产品中,其在用户体验方面还需改进。而采用非接触式方式进行采集和识别,不会让用户产生抵触心理,能让用户具有良好的使用体验,因而能很好地应用到实际产品中。但相对于接触式掌静脉识别,非接触式还存在一些难点需要攻克。

1. 手掌关键点定位问题

在使用非接触方式采集掌静脉图像的过程中,由于移除了定位装置,每个人在不同时刻采集的掌静脉图像都有很大的不同,而且会产生很大的偏移和旋转等。另外,在采集时手掌还容易发生较大的形变,如手掌弯曲、手掌倾斜或者手掌过度伸展等。这些因素就给图像预处理带来了难题,因此在非接触式情况下如何精准地定位手掌关键点是进行掌静脉识别前的一大难点。

2. 手掌形变和光照导致识别精度不高

由于手掌形变带来的影响,同一个人的手掌在挤压和过度伸展的情况下其 ROI 图像将有很大的差别,因此提取出来的特征也具有较大的类内差异。而且,在非接触式的情况下,手掌与相机镜头之间的距离也会不同,因而也带来了光照不均匀等问题。这些因素都会对掌静脉特征提取和匹配产生很大的影响。

由于上述这些影响因素,传统的掌静脉特征提取算法对使用非接触式采集的掌静脉图像识别效果不是很理想,因此,如何解决非接触方式对掌静脉识别技术带来的不利因素,提高非接触式掌静脉图像的识别精度,对于掌静脉识别技术的发展具有极大的意义。

1.3 本文创新点

针对非接触式掌静脉识别技术的难点,本文引入卷积神经网络进行手掌关键点

的定位和掌静脉的特征提取，并做了如下研究和改进工作：

1. 针对传统算法对非接触式手掌图像难以定位关键点的问题，采用了深度学习方法进行关键点定位，设计了两级卷积神经网络级联来定位手掌关键点；另外，由于损失函数(MSE Loss)对手掌关键点定位不准，本文在 MSE Loss 上进行改进，提出了 Modified MSE Loss，有效地提升了手掌关键点的定位精度；

2. 针对非接触式掌静脉数据库数据集较少，训练神经网络容易产生过拟合的问题，提出了类内扩增数据和类间扩增数据的方法，有效地扩充了原始数据库，为后续掌静脉特征提取网络的训练提供了数据支撑。

3. 对深层特征提取网络 Resnet18 的 Residual Block 模块结构进行改进，实验证明，改进后的 Residual Block 模块能有效地提升整体模型的精度；

4. 针对传统深层神经网络参数量大，计算耗时长的问题，本文设计了一个轻量级网络 TinyPVNet，网络主要采用深度可分离卷积结构，实现了在不降低识别精度的情况下，有效地降低了网络规模和参数量；并对模型进行量化压缩，大幅压缩了模型体积，从而让轻量级网络能够应用在嵌入式等计算资源受限的平台上。

1.4 论文的结构安排及主要研究工作

本文的主要研究内容是非接触式掌静脉识别，针对非接触式掌静脉识别中难以定位手掌关键点问题，本文提出了采用卷积神经网络进行回归定位的方法，并改进了 MSE 回归损失函数；另外，对于基于传统人工特征进行掌静脉识别方法的鲁棒性不足问题，本文提出采用卷积神经网络，从掌静脉图像中自动学习并提取具有高鲁棒性的特征，克服旋转、偏移以及光照等影响，达到较高的识别精度。同时，针对卷积神经网络模型体积过大，计算复杂度较高的问题，本文同时还研究了利用轻量化模型降低模型体积并加速网络运行的方法。本文的结构安排如下：

第一章，绪论。本章首先介绍了非接触式掌静脉识别的研究背景和意义，接着对静脉识别的国内外研究现状和非接触式掌静脉识别技术的研究难点进行了介绍，从而引出了本文的主要研究内容，最后对本文的创新点和研究工作进行了总结。

第二章，图像的获取与图像预处理。本章首先介绍了非接触式掌静脉图像的采集方法和本文所使用的掌静脉数据库，接着介绍了如何对掌静脉图像的 ROI 区域进行选取，最后介绍了掌静脉 ROI 图像的预处理方法：图像增强和图像去噪。

第三章，手掌关键点定位。本章详细介绍了手掌关键点定位网络的结构设计，

包括粗略定位网络和精确定位网络，然后介绍了改进的关键点回归损失函数 (Modified MSE Loss) 的原理，最后介绍了手掌图像数据集的扩增方法以及手掌关键点定位的实验结果。

第四章，掌静脉特征提取网络设计。本章首先对一些经典的分类神经网络架构作了简要介绍，然后自行设计了两个深层掌静脉识别网络 Resnet18-Modified、VGG16-Modified，并对 Resnet18 的 Residual Block 模块结构进行了改进；接着介绍了经典轻量级神经网络的结构和设计思路，并自行设计了一个轻量级掌静脉识别网络 TinyPVNet；然后介绍了一些经典的分类损失函数原理和特征向量的相似性度量方法，最后对模型量化压缩方法的原理进行了介绍。

第五章，实验结果及分析。本章首先介绍了本文提出的掌静脉数据扩增方法，包括类内扩增和类间扩增方法，然后简要介绍了掌静脉识别测试中用到的实验性能指标；接着介绍了实验数据的划分方式和模型训练时的参数设置；然后对多组实验结果进行了分析，主要包括不同特征向量维度、不同网络结构、Residual Block 模块改进以及不同损失函数所取得的实验结果进行分析，并与其他算法的性能进行对比；最后对模型量化压缩的结果和模型推理时间进行了测试。

本文掌静脉识别算法的整体框架图如图 1-2 所示。

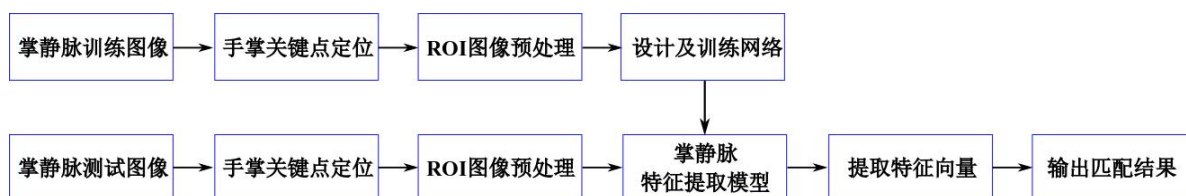


图 1-2 掌静脉识别算法整体框架图

Fig.1-2 Framework of palm vein recognition algorithm

第二章 掌静脉图像的获取及预处理

掌静脉图像的获取及预处理是进行掌静脉识别的前提条件，图像的预处理对后续掌静脉识别的精度也起着至关重要的作用。本章首先介绍了手掌静脉图像的非接触式采集方式，然后介绍了本文使用的数据库：CASIA 多光谱掌静脉数据库。由于数据库的手掌图片存在大量与静脉识别的无关区域，因此介绍了本文采用的关键点选择和掌静脉 ROI 选取方法，然后着重介绍了 ROI 图像的增强和去噪方法。

2.1 手掌静脉图像采集

掌静脉与人脸、指纹等位于皮肤表面的生物特征不同，它位于手掌皮肤的内部，所以用一般的可见光摄像头是无法得到掌静脉的图像信息。通过查阅相关生物医学资料得知，波长为 800-1000nm 的光线能够穿透人体的皮肤组织，但血管中的血红蛋白等物质会吸收大部分的光线，而其他的肌肉组织吸收的较少。根据这个特性可以推断出，当用波长为 800-1000nm 的光线照射手掌的同时，采用近红外摄像头对穿透手掌后的光线进行成像，由于血红蛋白吸收了大部分光线，会导致血管所在的部位亮度相对较低，而非血管组织部位的亮度则相对较高，这样我们便可以得到掌静脉的清晰纹路图像。

目前存在的成像方式较多，主要有光反射式、光透射式及光侧射式三种。由于人体的手掌组织较厚，光线不太容易穿透，因此光透射式及光侧射式采集的图像质量不佳，故我们采用光反射式采集图像，采集装置如图 2-1 所示。

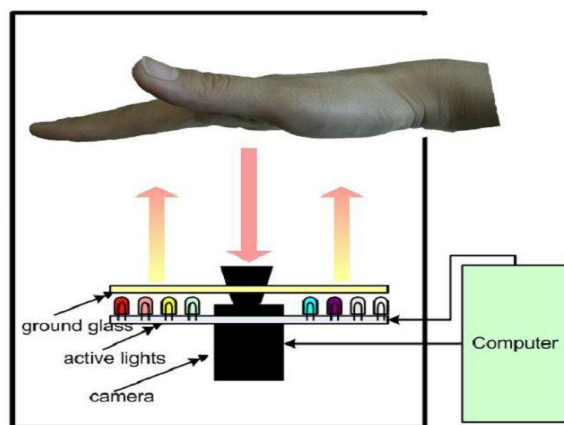


图 2-1 掌静脉采集装置图^[27]

Fig.2-1 Palm vein collection device^[27]

2.2 掌静脉数据库介绍

为了验证本文算法的有效性以及相比其他算法的优势，本文没有自采数据进行实验，而是采用了公开数据库：中国科学院生物识别研究所提供的 CASIA 多光谱掌静脉数据库，这也是目前非接触式掌静脉识别研究中最常用的公开数据库。

中国科学院自动化研究所(CASIA)提供的多光谱掌静脉图像公开数据库(CASIA-Multi-Spectral-PalmprintV1^[27])总共采集了 100 个志愿者的左右手手掌静脉图，其使用的是 2.1 节中介绍的光反射式采集装置，所有的掌静脉图像都是 8 bit 灰度图。整个数据库的采集过程共使用了 460nm、630nm、700nm、850nm、940nm 以及白光共 6 组不同的波长，图 2-2 中显示的是同一个人的左手在六组不同波长下的掌静脉图像。为了保证掌静脉图像质量的稳定性，整个采集过程分为两个时间段，时间间隔在一个月以上。每个时间段的采集过程中，在同一波长下对于每个人的左

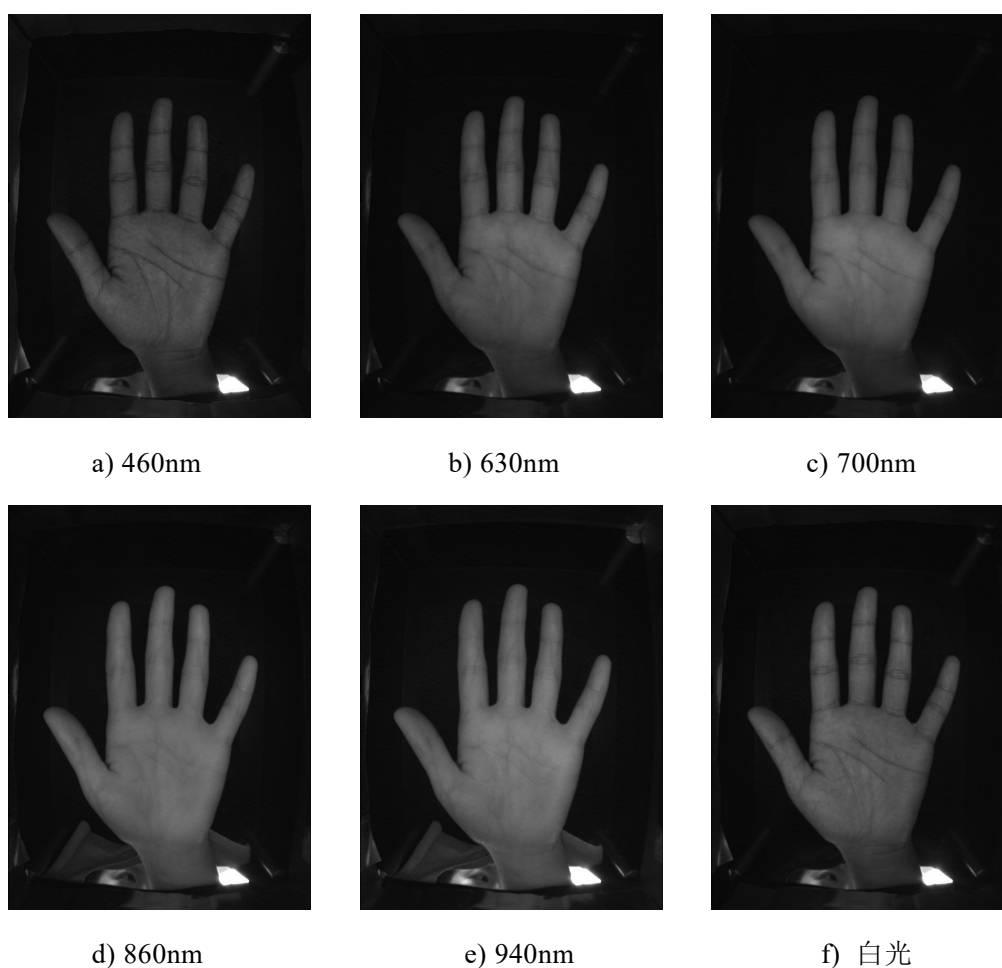


图 2-2 不同波长下的掌静脉图像

Fig.2-2 Images of palm vein at different wavelengths

右手分别采集 3 张图像，同时，为了增加类内样本的多样性及实用性，采集时每个人的手掌都进行了不同的旋转。在本文中，为了有效地利用数据集，我们规定同一个人的左右手图像属于不同类，因此该数据库总共采集了 200(100x2)类手掌，共 7200(100×2×6×6)张图像。

图 2-2 中，我们对 6 组不同波长下的掌静脉图像进行对比，发现在 940nm 波长下的掌静脉图像最清晰，因此本文实验采用的掌静脉图像集是该数据库中 940nm 波长红外光下采集到的掌静脉图像，总共有 200(100x2)类，共 1200(100x2x6) 张掌静脉图像。

2.3 掌静脉图像 ROI 选取

由图 2-2 可知，我们获取的是全手掌图像，由于同一个人在不同时刻获取的手掌图像都不相同，其手掌会有一些的位移，因此需要选取手掌的一部分固定区域作为 ROI(Region of Interest)，以便后续进行掌静脉识别。并且，掌静脉 ROI 图像区域的选择应以尽可能稳定地包含丰富的掌静脉信息为原则，从而为提升整个掌静脉匹配算法的性能提供前提条件。

2.3.1 手掌关键点选择

由于采集设备是非接触式的，用户的手掌可以有非常多的姿态变化，虽然采集设备对手掌的活动范围有一定的限制，但基于用户体验等方面的考虑，并没有设定太严格的活动限制，因此当人的手掌进入掌静脉采集装置时，手掌可能处于图像的不同位置，并可能出现手掌弯曲、手掌倾斜或者手掌过度伸展等多种姿态变化，这使得我们不能直接截取手掌图像中的固定区域作为 ROI。目前的方法都是先找到手掌的一些关键点，如指尖、指缝等，然后再根据这些关键点进行掌静脉图像 ROI 的选取。经过对整个手掌图像进行分析后发现，由于手掌图像亮度高，背景图像亮度低，两者之间的像素值差异明显，因此为了提高定位的准确性，我们应该把手掌和背景的连接区域作为关键点的候选区域，所以我们最终选择以手掌的每个手指间隙作为关键点参考。如图 2-3 所示，这里我们选取食指与中指间隙 A、中指与无名指间隙 B 以及无名指与小指间隙 C 作为手掌关键点。

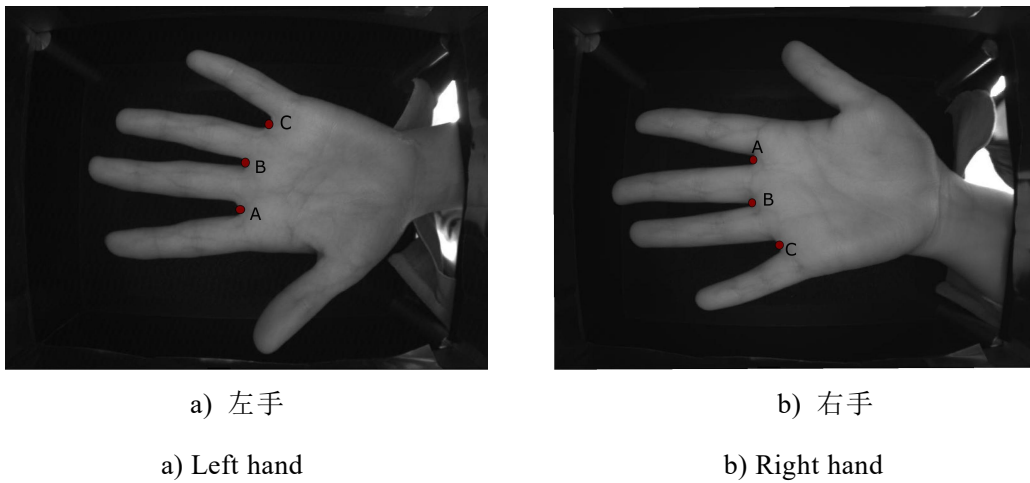


图 2-3 手掌关键点选择

Fig.2-3 Palm key point selection

2.3.2 图像旋转校正

在得到手掌关键点之后，我们需要对原图进行旋转校正，进而从原图中剪切出 ROI 图像。如图 2-4(a)所示，我们选取食指与中指间隙 A 以及无名指与小指间隙 C 两点，设 A 点坐标为 (x_1, y_1) ，C 点坐标为 (x_2, y_2) ，连接 AC，求出直线 AC 与垂直线之间的夹角 θ ，其中夹角 θ 的计算公式如下：

$$\theta = -\arctan\left(\frac{x_2 - x_1}{y_2 - y_1}\right) * 180 / \pi \quad (2-1)$$

然后将原图进行旋转 θ 角度，从而得到旋转校正后的图像。图 2-4(a)为原图像，图 2-4(b)为旋转校正后的手掌图像。

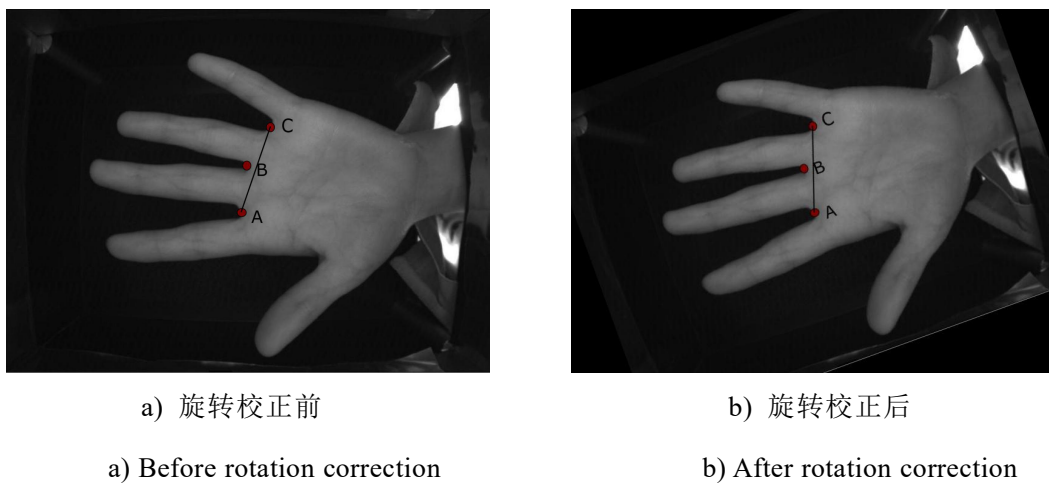


图 2-4 图像旋转校正结果

Fig.2-4 Image rotation correction result

2.3.3 掌静脉 ROI 选取及尺度归一化

经过图像旋转校正后，我们就可以根据这些关键点进行规则的区域（ROI）选取，如内切圆、矩形等。目前，掌静脉图像常用的 ROI 选取方法主要包括基于大鱼际区域^[28]、基于手掌内切圆^[29]和基于掌心矩形的方法^[30]等。其中，基于掌心矩形的 ROI 提取方法与其他两者相比，其计算比较简单，ROI 区域提取的位置比较精确，偏移量比较小，而且也有利于对后续卷积神经网络的输入图像进行处理，因此本文使用基于掌心矩形的 ROI 提取方法。如图 2-5 所示，我们首先得到食指与中指的间隙点 A 和无名指与小指的间隙点 C，设 AC 线段的长度为 m ，我们作出 AC 线段的中垂线，然后沿中垂线选取距离 AC 线段 $(L+m/2)$ 长度的点为中心，作出长度为 m 的正方形，即如图 2-5 所示的红色正方形区域，作为选取的掌静脉 ROI 区域。这里我们设置 L 为 20 个像素。

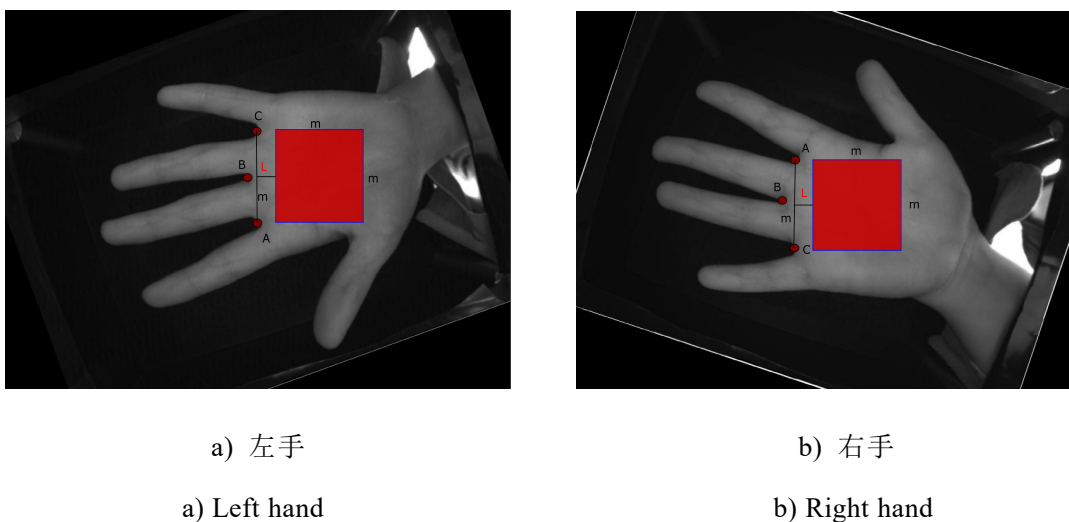


图 2-5 基于掌心矩形的 ROI 提取方法

Fig.2-5 ROI extraction method based on palm rectangle

从图 2-5 可以看出，即使是同一个人的手掌，在张开和闭合的情况下，其 AC 两点的长度 m 都不一样，因此我们要对提取到的 ROI 区域进行尺度归一化，将不同尺寸大小的 ROI 区域归一化成同样大小的 ROI 区域，这里我们采用双线性插值法进行尺度归一化。如图 2-6 所示，通过对所有的 ROI 区域的尺寸大小进行统计，我们得出其尺寸中位数为 149，因此我们将 ROI 区域的尺度大小归一化为 150×150 。

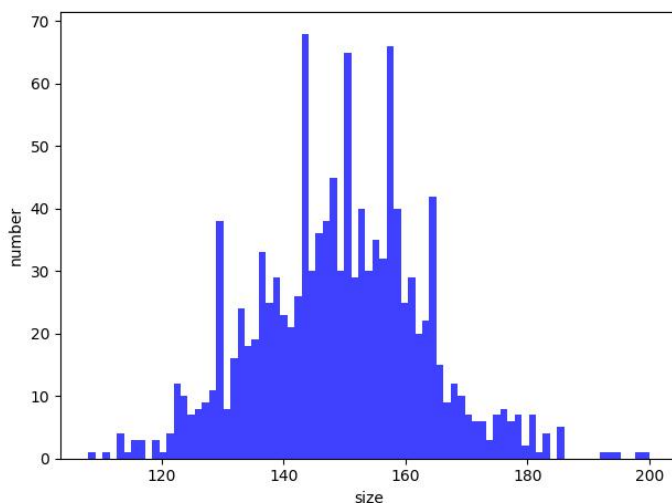


图 2-6 手掌 ROI 的边长统计结果

Fig.2-6 Side length statistics of palm ROI

2.4 掌静脉图像增强

在得到的近红外掌静脉 ROI 图像中，由于掌静脉与周围组织的区分度不明显，无法直接提取掌静脉信息，因此需要对图像进行增强处理，通过增强图像的对比度，以突出静脉结构的信息，使得静脉纹路清晰可见。常见的图像增强方法有很多，如对比度拉伸、直方图均衡、CLAHE(对比度受限自适应直方图均衡)等，我们将这些方法应用到掌静脉图像上，并对比它们的效果。

2.4.1 直方图均衡

直方图均衡化(Histogram Equalization)是图像处理领域中利用图像直方图对图像对比度进行调整的方法。其基本思想是将原始图像的灰度直方图映射到全部灰度范围内，使得原始图像的小区间灰度值均匀分布在整個灰度区间，从而完成灰度的拉伸。这样使得图像灰度的动态范围增加，从而提高了图像的对比度。

对于一幅图像，假设变量 r 代表原始图像中的灰度值，其范围为 $[0, L-1]$ ，则直方图均衡算法是寻找一个变换 $s = T(r)$ ， s 为变换后图像像素点的灰度值，则 $T(r)$ 满足下面的条件：

- 1) 在 $[0, L-1]$ 的区间内， $T(r)$ 是单值单调递增函数；
- 2) 当 $0 \leq r \leq L-1$ 时， $0 \leq T(r) \leq L-1$ 。

由于变换后不能改变原始图像的灰度值分布，单值是为了能够进行反变换，因

此需要满足条件(1)，这样才不会改变灰度值的排列顺序；而条件(2)则是要保证映射前后图像的灰度区间一致。

由于图像的灰度值是离散值的，在直方图均衡化时，我们首先计算每一个灰度值的像素点数量，从而求出每一个灰度值出现的概率，如式(2-2)所示，其中 MN 为图像中所有像素的点数， n_k 表示图像中灰度值是 r_k 的所有像素点数；图像进行直方图均衡化的映射变换函数如式(2-3)， s_k 代表变换之后的灰度级。

$$P_r(r_k) = \frac{n_k}{MN}, \quad k = 0, 1, \dots, L-1 \quad (2-2)$$

$$s_k = T(r_k) = (L-1) \sum_{j=0}^k P_r(r_j) = \frac{(L-1)}{MN} \sum_{j=0}^k n_j, \quad k = 0, 1, \dots, L-1 \quad (2-3)$$

直方图均衡的具体操作过程为：

- 1) 提取原始输入图像的直方图分布，结果为 $P_k(r_k)$ ；
- 2) 对原始输入图像的每一个灰度级 r_k 都进行式(2-3)的变换，得到变换之后的灰度级 s_k ；
- 3) 用式(2-3)变换之后的灰度值代替原始输入图像的灰度值，从而得到处理后的图像。

在近红外掌静脉图像中，由于静脉组织和周围组织的区分度低，因此图像中 80% 以上的像素值都集中在一个较小的范围内，如图 2-7(c)所示，原掌静脉图中的像素灰度值基本都集中在 80-120 范围内，很难区分出手掌静脉；而经过直方图均衡化以后，如图 2-7(d)所示，图像的灰度值基本上均匀地分布在 0-255 区间内，而且分布趋势也与原图一致，通过图 2-7(a)和图 2-7(b)的对比也可以明显的看出，在经过直方图均衡化后，掌静脉组织和周围图像的对比度增强了许多，效果不错。但从图 2-7(d)中也可以看出，经过直方图均衡后，图像的灰度值分布在接近 0 和 255 的部分较为集中，而中间分布则相对稀疏，可见拉伸并不均匀，会导致图像中的某些区域整体偏暗或偏亮，如图 2-7(b)所示，图像的右上部分偏亮，这样容易丢失原始掌静脉的信息。

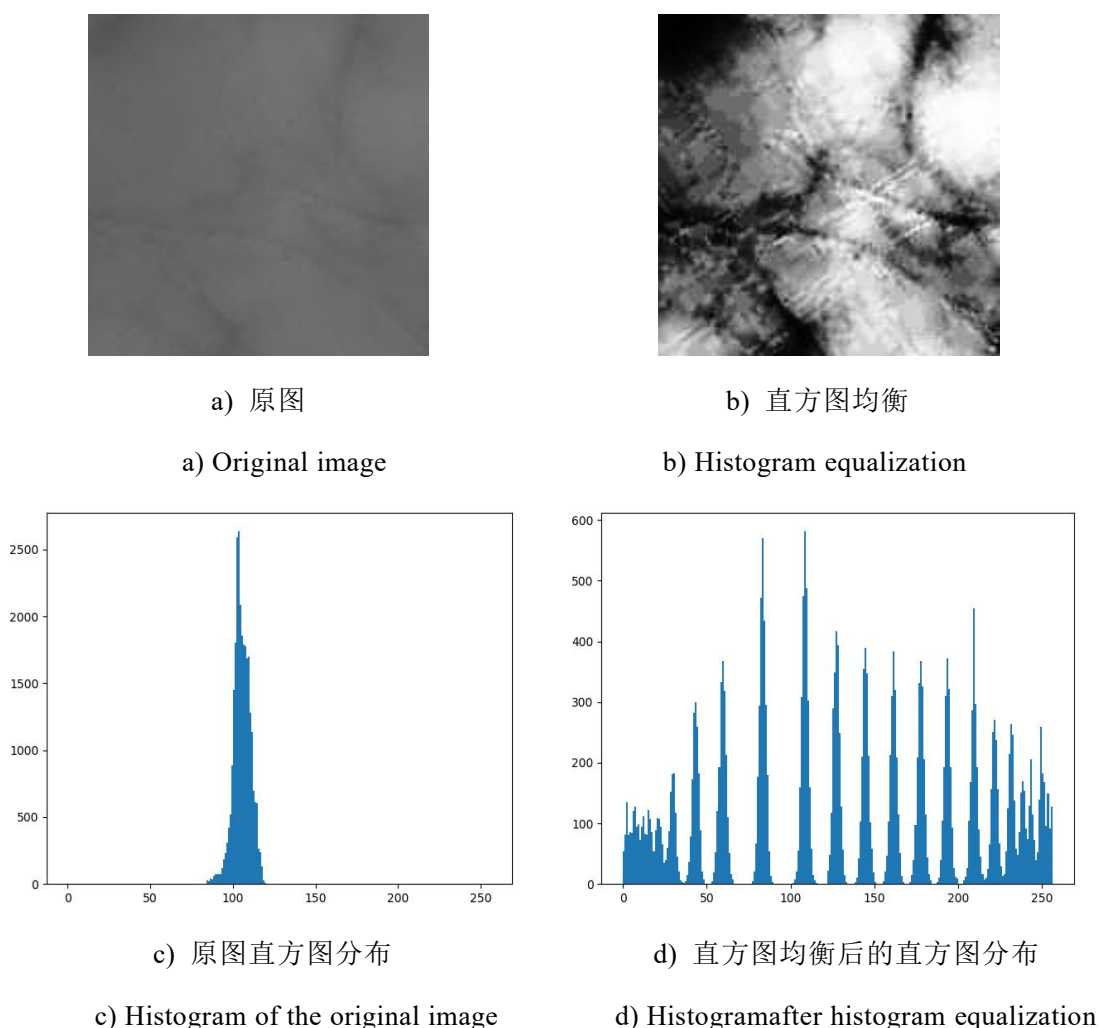
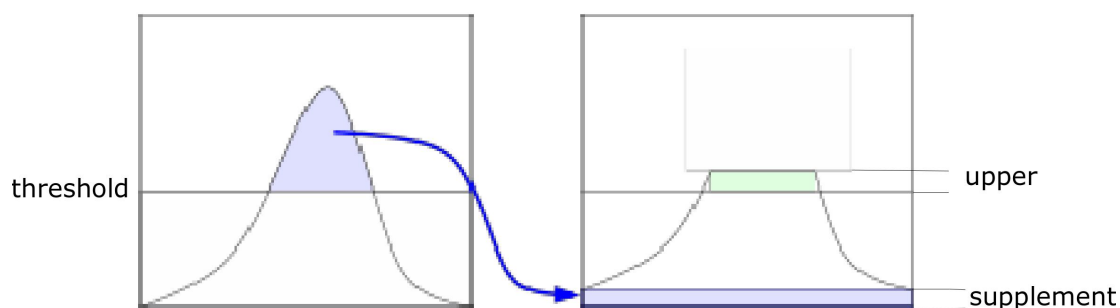


图 2-7 直方图均衡算法结果对比

Fig.2-7 Histogram equalization algorithm results comparison

2.4.2 CLAHE(对比度限制自适应直方图均衡)

对比度限制自适应直方图均衡 (CLAHE) 方法能有效解决对图像直接进行直方图均衡出现的过度增强现象。其基本思想是先把输入图像分成多个不重叠的小块图像, 分别统计每一块图像的直方图, 然后对每一块图像的直方图进行阈值剪切处理。如图 2-8 所示, 超出阈值的部分不会直接抛弃, 而是把超出的部分平均分布到各个灰度级上。从图 2-8 可以看出, 当把超出阈值的部分填充到各个灰度级之后, 整个直方图会出现一定的上移(图中 **upper** 部分), 从而使得这部分又重新超出阈值, 因此需要采用重复裁剪的方法, 逐渐缩小 **upper** 部分, 最终使 **upper** 部分趋近于零, 这样就完成了整个阈值裁剪过程, 然后再对每一块图像进行局部直方图均衡化处理。

图 2-8 直方图裁剪过程^[1]Fig.2-8 Histogram cutting process^[1]

经过局部直方图均衡化处理后，我们会发现原始图像被分成多个相互不重叠的区域，因此处理后的各子块区域会出现灰度的不连续，造成明显的分块现象。所以，我们还需要对各个子块的像素点按其和周围各块灰度值的映射关系进行插值化处理，从而消除这种“分块”现象。

CLAHE 算法的具体操作流程如下：

- 1) 首先将原始图像分成 $M \times N$ 个子块，而且 M 、 N 要分别能被原始图像的宽高整除，保证各子块之间互不重叠；
- 2) 对 $M \times N$ 个分块子图像分别进行直方图统计，得到子块对应的直方图分布；
- 3) 对每个子块图像的直方图进行重复的阈值裁剪后，再进行直方图均衡化处理，并获得映射关系矩阵；
- 4) 使用像素点周围子块的映射矩阵对原始图像中的每个像素点进行插值处理，最终得到处理后的图像。

我们对掌静脉原图进行受限自适应直方图均衡化处理后，得到的结果如图 2-9(b) 所示。与图 2-7(b) 中直接使用直方图均衡化处理的结果相比，可以看到 CLAHE 算法处理后的图像没有出现曝光过度的现象，掌静脉纹路更加清晰；通过图 2-7(d) 和图 2-9(d) 的对比也可以看出，使用 CLAHE 算法处理后的图像直方图分布更均匀平滑，效果更好。因此本文采用对比度限制自适应直方图均衡（CLAHE）方法对掌静脉的 ROI 区域进行图像增强。

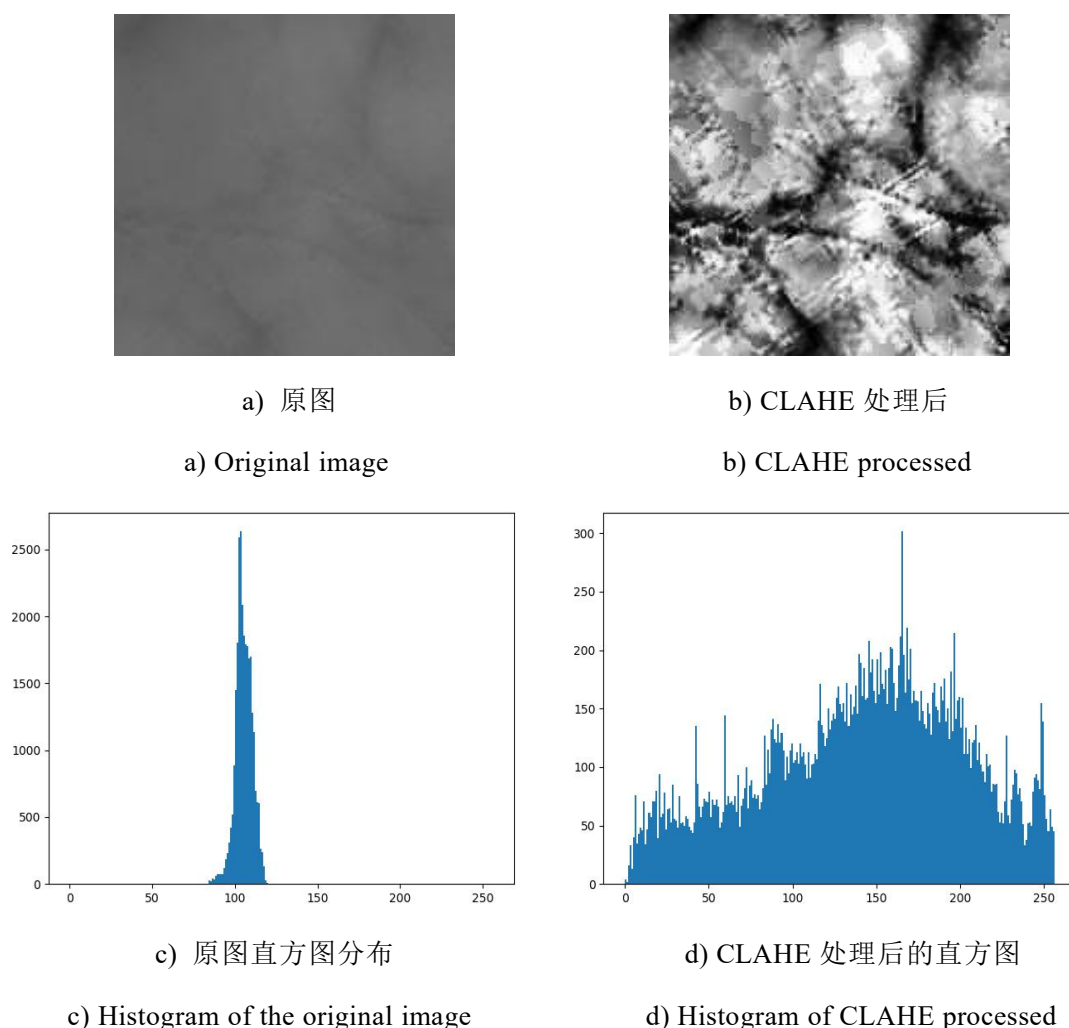


图 2-9 CLAHE 处理后的结果对比

Fig.2-9 Comparison of results of CLAHE processed

2.5 掌静脉图像去噪

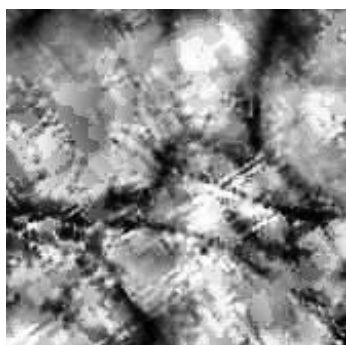
掌静脉图像在经过对比度限制自适应直方图均衡的方法得到对比度增强后，图像的噪声也得到了增强。图像噪声对后续的掌静脉特征提取及匹配会产生影响，因此需要运用图像去噪方法来降低图像噪声的干扰。常见的图像去噪方法分为空间滤波和频域滤波两种，空间滤波是采用图像模板对图像进行卷积处理，从而达到抑制噪声的目的，如中值滤波、均值滤波等；频域滤波则是先将图像变换到频率域，采用低通滤波器滤除高频噪声，然后再反变换得到平滑图像。下面将简要介绍一些典型的去噪方法，并对结果进行分析比较。

2.5.1 均值滤波

均值滤波属于线性滤波技术，其方法是采用整体像素邻域点的灰度平均值来代替中心点的值。其具体方法是：首先采用 3×3 或 5×5 的滤波模板在原始图像上依次移动，每次移动时，将模板范围内的像素点灰度值先进行求和再求平均，然后用平均值代替原中间像素点的灰度值。如式(2-4)所示， N 代表模板的大小， $f(x_i, y_i)$ 为原图中模板对应像素点的灰度值， $g(x, y)$ 为均值滤波后的像素值。利用均值滤波可以有效降低图像灰度值突变，从而去除图像的噪声。

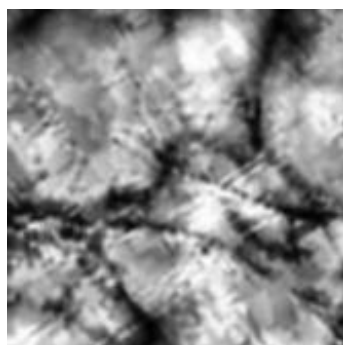
$$g(x, y) = \frac{1}{N * N} \sum_{i=0}^{N*N} f(x_i, y_i) \quad (2-4)$$

图 2-10(b)是使用 3×3 模板进行均值滤波处理后的掌静脉图像，图 2-10(d)是对



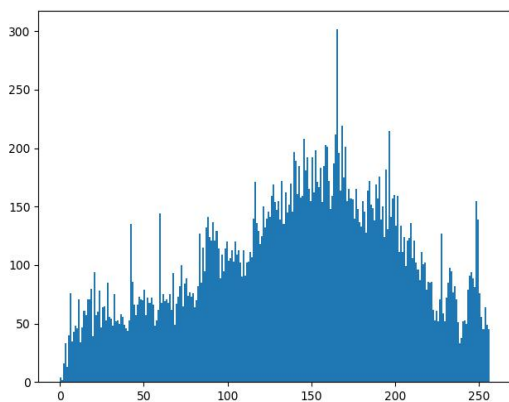
a) 原图

a) Original image



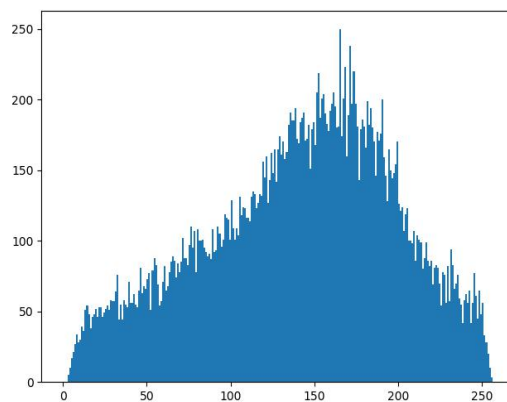
b) 均值滤波后的图像

b) Mean filtering image



c) 原图直方图分布

c) Histogram of the original image



d) 均值滤波后的直方图

d) Mean-filtered histogram

图 2-10 均值滤波结果分析

Fig.2-10 Analysis of Mean Filter Results

应图像进行均值滤波处理后的直方图分布。对比图 2-10(a)和图 2-10(b)后，我们可以发现原图中的噪声得到了很好的消除，但经过均值滤波后的图像也变得很模糊。通过图 2-10(d)也可以看出，图像非静脉区域的灰度值较原图有所增大，因此不易与静脉区域进行区分。

2.5.2 中值滤波

中值滤波是一种非线性空间滤波技术，其基本方法与均值滤波相似，不过与均值滤波采用平均值不同的是，其采用的是像素邻域点的统计排序中位值来代替中心点的值，从而消除孤立的噪声点。

图 2-11(a)和图 2-11(b)是原图经过增强后，再用 3×3 模板进行中值滤波的前后对比图。可以看到，经过中值滤波后，有效地消除了图像噪声，静脉结构有细微的

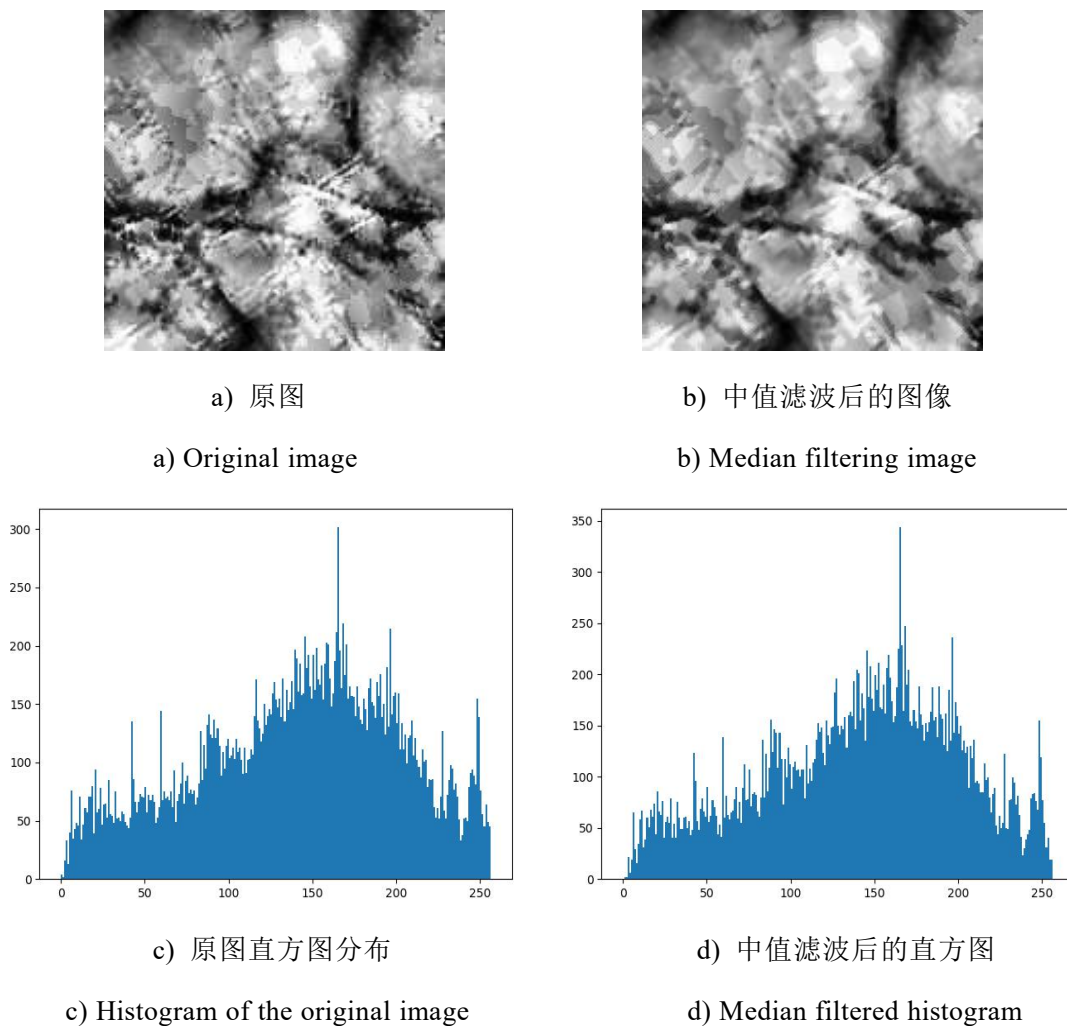


图 2-11 中值滤波结果分析

Fig.2-11 Analysis of median filtering results

模糊。通过图 2-11(d)也可以看出，与均值滤波相比，中值滤波处理后的图像静脉区域与非静脉区域更容易区分，有利于后续处理；而且中值滤波后的直方图分布与原图相比也没有太大的改变，因此本文选取中值滤波方法对掌静脉 ROI 区域进行去噪。

2.6 本章小结

本章主要对掌静脉图像的获取和预处理流程作了详细介绍。首先介绍了手掌静脉图像的获取及采集原理，然后介绍了本文所使用的 CASIA 掌静脉数据库；接着对如何选取手掌关键点以及 ROI 区域进行了介绍。针对掌静脉与周围组织的区分度不明显以及图像存在较强噪声的问题，对掌静脉图像进行了图像增强和图像去噪实验，并对不同算法之间的效果进行了对比，最后决定采用 CLAHE(对比度受限自适应直方图均衡)算法进行图像增强和中值滤波进行图像去噪。

第三章 手掌关键点定位网络

在传统的手掌关键点定位方法中，最常用的是先采用 OTSU 方法^[31]对高斯模糊后的图像进行二值化，获得手掌轮廓图，然后采用辐射距离函数 RDF(Radial Distance Function)算法计算手腕中心点到各个指间间隙的距离，根据各个参考点与手腕中心的距离及各个参考点之间的相对位置，最后得出手掌关键点的具体位置。这种方法在接触式采集的图像上表现良好，但当手掌出现倾斜、并拢或较大旋转角度时，传统方法就无法准确地定位出手掌关键点，尤其是对于左右手的食指与中指间隙 A 和无名指与小指间隙 C 会出现判断错误。近年来，卷积神经网络在目标检测和关键点定位领域也应用的越来越广泛，其表现出来的检测效果好、定位精度高等特性使得越来越多的图像处理任务也开始采用深度学习方法。因此本文决定引入卷积神经网络进行手掌关键点定位，提出了两级神经网络级联检测的方法，并通过改进 MSE Loss，有效地提升了手掌关键点的定位精度。

3.1 手掌关键点定位网络设计

由于关键点周围的图像占原图的比例很小，因此很难做到一次性的精确定位，所以我们采用了两级卷积神经网络级联检测的方法。首先设计一个粗略定位网络，粗略地定位手掌关键点区域；然后再将粗略定位的关键点周围图像输入到第二级精确定位网络，得到手掌关键点的精确位置。

3.1.1 粗略定位网络

由于输入图像的原图大小为 768×576，如果将整张原图按原始比例输入到神经网络中，整个网络前向推理的速度将会很慢。而我们在粗略定位网络中只需要得到关键点的粗略位置，因此为了缩短关键点定位的时间，我们将原图等比例 resize 成 64×48 进行输入。并且由于下一级精确定位网络的输入图像是关键点周围的局部图像，无法获取图像的全局信息，因此我们给粗略定位网络增加了一个左右手分类任务，其网络结构如图 3-1 所示。

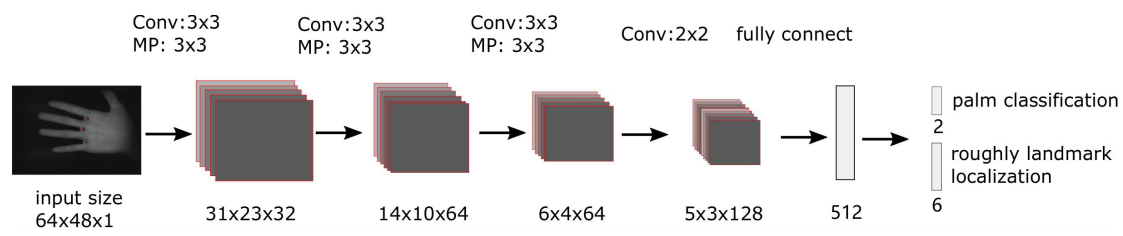


图 3-1 粗略定位网络结构

Fig.3-1 Rough localization network structure

首先，将原图等比例 `resize` 成 64×48 进行输入，依次经过 3×3 卷积层、 3×3 最大池化层以及 512 维全连接层后，进入两个不同的任务：左右手的分类和关键点的粗略定位。

3.1.2 精确定位网络

由于粗略定位网络的输入图像是整个原始图像，手掌关键点占的像素比例很小，因此不利于精确定位；但经过粗略定位网络后，我们可以得到 A、B、C 三点的粗略位置，所以这一级的精确定位网络，我们采用的是关键点周围的图像，如图 3-2 所示，以 B 点为矩形中心，AC 两点之间距离的两倍为边长，剪切出包含 A、B、C 三点的矩形区域，即图 3-2 中的绿色矩形框部分。这样手掌关键点部分所占图像的比例较大，并且能去除周围因素的干扰，因而更有利于关键点的精确回归。

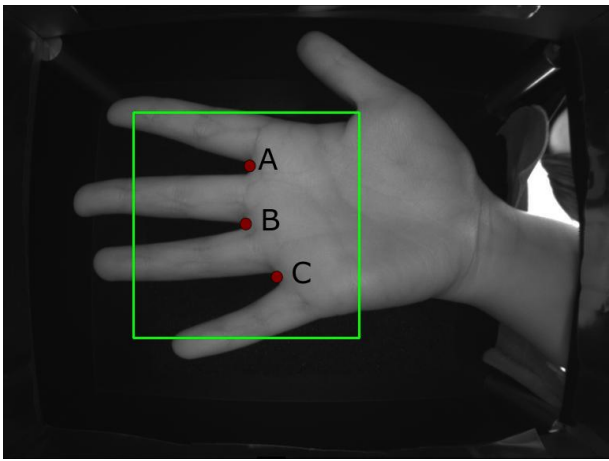


图 3-2 精确定位网络的输入图像

Fig.3-2 Input image of the precise localization network

精确定位网络的网络结构图 3-3 所示，首先，网络的输入为图 3-2 所示的绿色矩形框区域图像，为了缩短手掌关键点定位的时间，我们将该矩形区域统一 `resize` 成 64×64 大小。接着经过 3×3 卷积层、 3×3 最大池化层以及 512 维全连接层后，得

到 A、B、C 三个关键点的回归量，再经过简单的计算即可得到关键点的具体坐标。

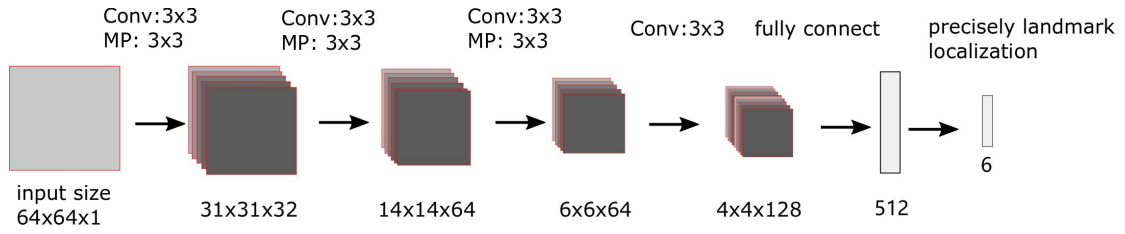


图 3-3 精确定位网络结构

Fig.3-3 Precise localization network structure

3.2 关键点回归损失函数设计

在关键点回归问题中，经常使用的损失函数是均方差(MSE)损失，其公式为式(3-1)，其中 y 表示真实值， \hat{y} 表示网络的预测值。从公式(3-1)可以看出，当预测值与真实值的差值越小，整体损失也就越小，点回归就越精确。但当网络的预测值与实际值相差较大时，其产生的 Loss 值也会更加大，如图 3-4 所示，这样离群值产生的 loss 将会主导整个 loss，整个梯度的更新也就受其很大的影响，因此整个网络将会收敛缓慢，甚至无法得到局部最优解，因而不利于关键点的精确回归。

$$L_{MSE} = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (3-1)$$

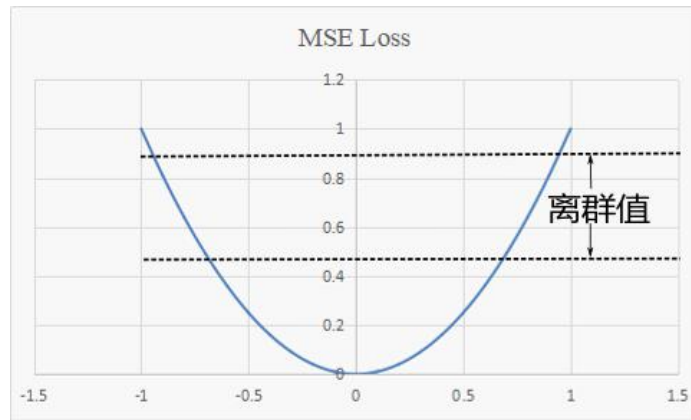


图 3-4 MSE Loss

Fig.3-4 MSE Loss

为解决均方差损失带来的局限性，本文在均方差损失函数的基础上进行改进，提出了式(3-2)的损失函数，命名为 Modified MSE Loss。其中 λ 和 w 是两个超参数， λ 控制函数线性段的斜率， w 控制函数线性与非线性的区域占比。整个函数的曲线如图 3-5 所示，可以看出，改进后的损失函数加大了正常区间的权重，这样相当于

$$L_{\text{modified LSE}} = \frac{1}{m} \sum_{i=1}^m \begin{cases} \lambda \cdot |y_i - \hat{y}_i|, & \text{if } |y_i - \hat{y}_i| < w \\ (y_i - \hat{y}_i)^2 + \lambda w - w^2, & \text{otherwise} \end{cases} \quad (3-2)$$

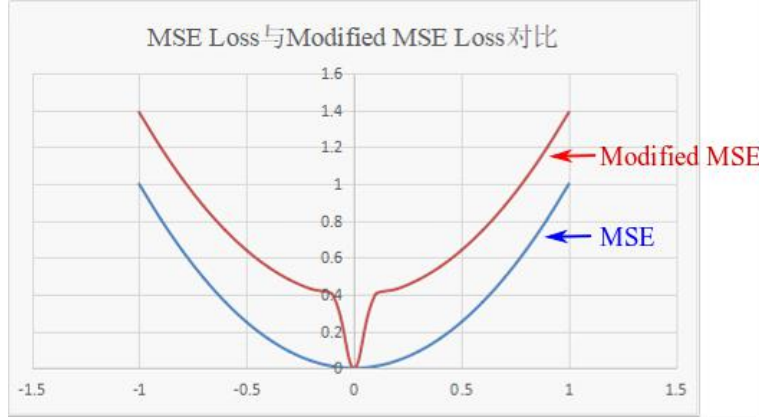


图 3-5 Modified MSE Loss($\lambda=4, w=0.1$)

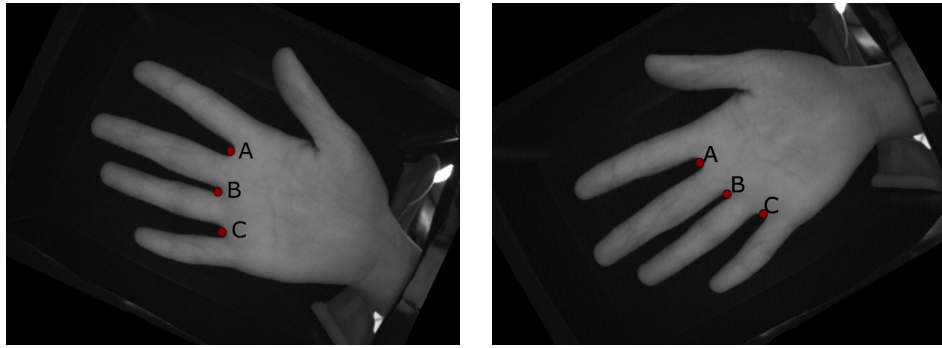
Fig.3-5 Modified MSE Loss($\lambda=4, w=0.1$)

减小了对离群值的惩罚程度，使得梯度更新受离群值的影响变小，而且存在两个超参数 λ 和 w ，便于调节两者之间的权重占比。另外，与原始 MSE Loss 相比，当预测值和真实值接近时，我们设计的损失函数增大了反向传播时的梯度值，从而使得关键点的回归更加精确。

3.3 实验数据准备及结果分析

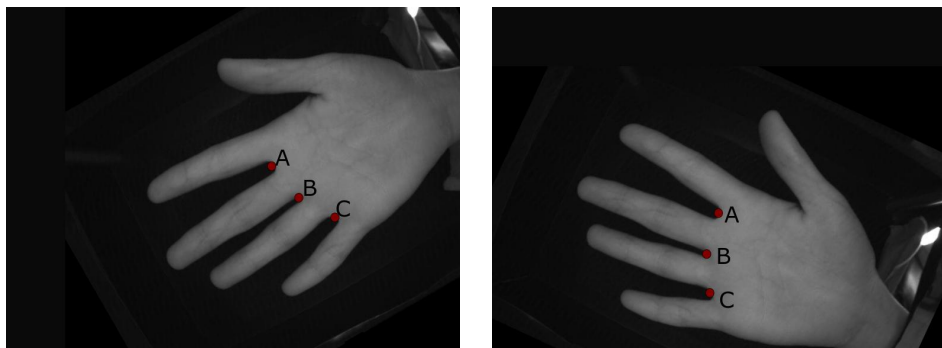
3.3.1 手掌图像数据集扩增

由于我们选取的波长为 940nm 的 CASIA 多光谱掌静脉数据库只有 1200 张手掌图像，如果只用这些数据训练神经网络，会导致严重的过拟合，即在测试集上的泛化能力很差。因此我们决定对数据库的手掌图像进行数据扩增。我们首先对原始数据集进行划分，将其中的 60% 作为训练集，10% 作为验证集，剩下 30% 作为测试集。这样我们的原始训练集图片数有 720(1200×60%) 张，我们对训练集的每张图像在 $(-30^\circ, 30^\circ)$ 范围内进行随机角度旋转 12 次，得到 8640(720×12) 张图像，然后在此基础上进行上下左右四个方向的随机像素平移 4 次，能得到 138240(8640×4×4) 张训练图像。图 3-6 显示的是对手掌图像进行数据扩增的示例，可以看到，经过随机平移和旋转后，手掌样本的多样性得到保证，能有效地防止卷积神经网络训练时出现过拟合问题，提高了卷积神经网络的泛化能力。



a) 手掌图像旋转

a) Palm image rotation



b) 手掌图像旋转+平移

b) Palm image rotation + translation

图 3-6 手掌图像数据扩增示例

Fig.3-6 Example of palm image data augmentation

3.3.2 实验结果分析

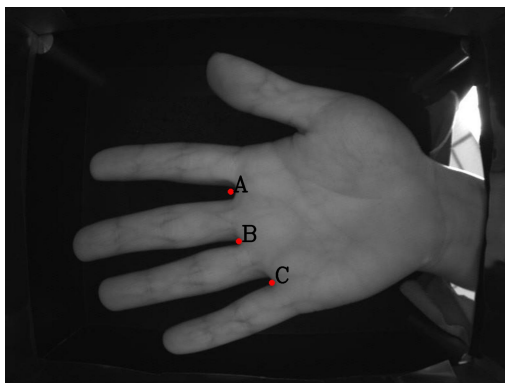
(1) 粗略定位网络定位精度分析

我们首先对粗略定位网络的定位结果进行分析，对于 Modified MSE Loss 中的两个超参数 λ 和 w ，我们实验了多组不同的参数值。为了便于观察和对比，我们在训练时将关键点坐标进行了归一化处理，即计算关键点坐标与图片长宽的比值，因此 w 的取值范围为 $0 < w < 1$ 。如表 3-1 所示，我们对 λ 和 w 各取 5 组不同的值进行实验，表 3-1 中计算的是粗略定位网络在不同参数下定位关键点的归一化绝对值平均误差（NMAE），单位为 10^{-2} 。从表 3-1 中可以得出，当 $\lambda=4$ 和 $w=0.1$ 时，粗略定位网络能取得最低的定位平均误差 0.573×10^{-2} 。

表 3-1 粗略定位网络定位的归一化绝对值平均误差（单位： 10^{-2} ）Table 3-1 Normalized mean absolute error of rough localization network(10^{-2})

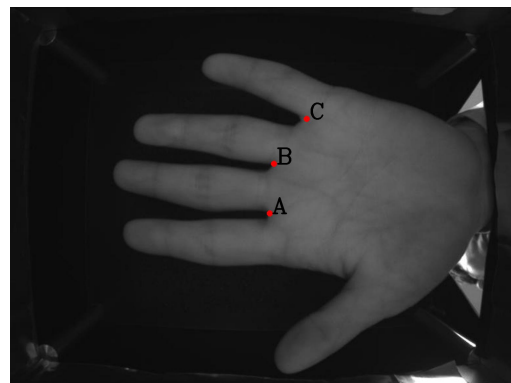
$\lambda \backslash w$	0.01	0.05	0.10	0.15	0.20
2	1.215	0.693	0.631	0.686	1.124
4	1.204	0.589	0.573	0.634	0.902
6	1.185	0.594	0.580	0.639	0.937
8	1.180	0.595	0.594	0.756	0.985
10	1.157	0.734	0.602	0.736	0.967

图 3-7 是粗略定位网络的关键点定位结果，可以看出，粗略定位网络的定位关键点结果比较理想，但有些点还存在一些偏差，如图 3-7(a)中的 A 点和图 3.7(b)中的 C 点定位偏差较大，所以仍需经过下一级网络进行精确定位。



a) 右手定位结果

a) Right hand localization result



b) 左手定位结果

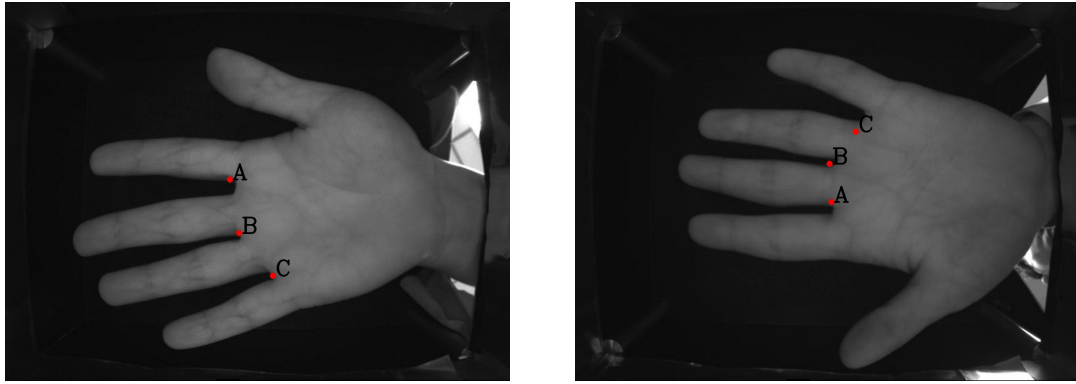
b) Left hand localization result

图 3-7 使用 Modified MSE Loss 的关键点定位结果

Fig.3-7 Landmark localization Results of Using Modified MSE Loss

(2) MSE Loss 和 Modified MSE Loss 的定位精度分析

本文对 MSE Loss 和 Modified MSE Loss 的定位精度进行了对比。当使用 MSE Loss 进行训练后，其定位关键点的归一化绝对值平均误差为 1.753×10^{-2} ，而使用 Modified MSE Loss 的归一化绝对值平均误差为 0.573×10^{-2} ，两者之间的差值为 0.0118。使用 MSE Loss 训练后的关键点定位结果如图 3-8 所示，可以看到，与图 3-7 使用 Modified MSE Loss 的训练结果相比，其关键点定位很不准确，无法满足定位要求，因此本文提出的 Modified MSE Loss 定位关键点效果更好，鲁棒性更强。



a) 右手定位结果

b) 左手定位结果

a) Right hand localization result

b) Left hand localization result

图 3-8 使用 MSE Loss 的关键点定位结果

Fig.3-8 Landmark localization Results of Using MSE Loss

(3) 精确定位网络定位精度分析

从图 3-7 的定位结果可以看出，粗略定位网络定位的部分关键点误差仍较大，因此需使用精确定位网络进行精确定位。在精确定位网络中，我们同样使用不同的参数来进行关键点定位实验，如表 3-2 所示，当 $\lambda=6$ 和 $w=0.1$ 时，精确定位网络能取得的最低定位平均误差为 0.391×10^{-2} 。与粗略定位网络的最低定位平均误差 0.573×10^{-2} 相比，定位精度有了很大的提升。

表 3-2 精确定位网络定位的归一化绝对值平均误差（单位： 10^{-2} ）

Table 3-2 Normalized mean absolute error of precise localization network(10^{-2})

$\lambda \backslash w$	0.01	0.05	0.10	0.15	0.20
2	0.967	0.573	0.501	0.531	1.202
4	0.934	0.468	0.424	0.546	0.735
6	0.868	0.474	0.391	0.489	0.744
8	0.876	0.505	0.442	0.535	0.867
10	0.891	0.684	0.481	0.506	0.993

精确定位网络的关键点定位结果如图 3-9 所示，可以看到，在经过精确定位网络后，图 3-9(a)中的 A 点和图 3-9(b)中的 C 点都比图 3-7 中定位的更准确。

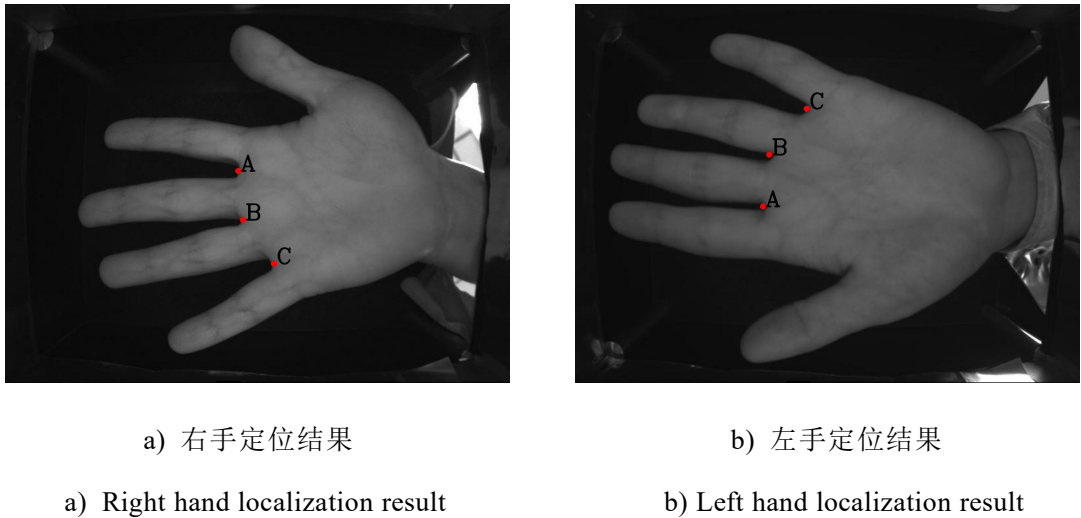


图 3-9 精确定位网络关键点定位结果(Modified MSE Loss)

Fig.3-9 Landmark localization Results of precise localization network(Modified MSE Loss)

(4) 两级网络的模型体积和网络耗时

我们对两级卷积神经网络的模型体积和网络耗时进行了统计，其结果如表 3-3 所示，其中网络耗时是在 CPU(Intel i5-3210M,2.5G 主频)上测试得到的。可以看到，两级网络的总耗时为 9.1ms，完全可以满足现实应用场景。

表 3-3 两级网络的模型体积和网络耗时

Table 3-3 Model volume and inference time of two cascade network

	模型体积(M)	网络耗时(ms)
粗略定位网络	4.11	3.6
精确定位网络	4.51	5.5

3.4 本章小结

本章主要解决了采用非接触方式采集的掌静脉图像手掌关键点难以准确定位的问题，首先介绍了设计的手掌关键点定位网络：粗略定位网络和精确定位网络。然后介绍了常用的回归损失函数 MSE Loss 及其在手掌关键点定位上的缺点及不足，接着介绍了我们改进后的 Modified MSE Loss 原理；针对手掌数据不足的问题，提出了数据集扩增方法，主要是对手掌图像进行平移和旋转，模拟手掌在真实场景下的各种姿态，从而防止训练网络时出现过拟合；最后对两级网络分别使用 MSE Loss

和 Modified MSE Loss 训练后的关键点定位结果进行了分析，通过实验证明了本文改进后的 Modified MSE Loss 对提升手掌关键点定位精度的有效性。

第四章 基于卷积神经网络的掌静脉特征提取模型

在掌静脉特征提取的过程中，传统的算法主要是利用人工设计的特征进行匹配，人工设计特征的过程繁琐而且效率低下，并且当手掌发生平移、旋转、挤压等复杂姿态变化时，这些人工设计的特征就容易存在鲁棒性不足的缺点。为了解决这个缺点，我们采用深层卷积神经网络从原始的掌静脉 ROI 图像中自动学习并提取具有区分力的特征。在现实的产品应用中，掌静脉识别算法通常需要部署在计算能力较弱的嵌入式平台上，因此需要对模型进行压缩。模型的压缩可以有多种方法进行实现，包括直接设计更加紧凑的网络模型、对网络进行剪枝以及参数量化等压缩方法，其中对网络进行剪枝处理需要在已训练好的模型上进行更改，流程繁琐，效率不高，因此本章探索了能够实现端到端的模型压缩方法，包括设计轻量级网络结构、模型参数量化等。

4.1 深层掌静脉特征提取网络

近年来，由于计算性能的提升和海量数据的出现，深度学习在计算机视觉方面的优势得以体现，自 2012 年 AlexNet^[32]提出后，卷积神经网络在图像分类、图像分割、目标检测等领域获得广泛应用。为了获得更好的性能，研究机构提出了性能更优越的卷积神经网络结构，并且网络层数也随之不断增加，从 7 层 AlexNet^[32]到 16 层 VGGNet^[33]，从 22 层的 GoogLeNet^[34]再到 152 层的 ResNet^[35]，甚至更有上千层的 ResNet^[35]和 DenseNet^[36]等。由此可见，要想在大规模数据集上取得较高的识别精度，我们需要采用深层神经网络进行有效的特征提取。而在本文的掌静脉识别任务中，由于掌静脉图像的纹理较复杂，因此需要采用深层神经网络才能提取出泛化能力强的特征。

4.1.1 特征提取网络的输入输出维度设定

在设计神经网络结构之前，我们首先需要确定神经网络的输入图片大小。输入图片的尺寸影响着整个识别算法的精度，如果对图片压缩的太严重，会导致图像的特征信息丢失。因此，根据 2.3.3 节中对手掌 ROI 区域大小统计的中位值，我们将网络输入图像的尺寸设置为 150×150；另一方面，本文中使用神经网络进行掌静脉识别的研究是属于开集识别，即测试集类别不会出现在训练集和验证集中，因此

测试算法性能时不能直接使用最后的分类层进行分类，而需要取分类层的前一层作为特征向量层，根据得到的特征向量再进行特征向量匹配，所以特征向量层的维度也影响着最终的匹配精度。唐溯^[17]在指静脉识别中将神经网络的输出维度设置为 256 维，取得了不错的效果。因此，本文将特征向量维度分别设置为 128 维、256 维和 512 维，并进行了对比实验，相关实验结果在第五章中进行了分析。

4.1.2 深层卷积神经网络结构介绍

目前，在各类图像识别大赛中，深层卷积神经网络取得的成绩都非常不错，如 VGGNet、Inception^{[37][38][39]}、Resnet 等都是几十层甚至上百层的大网络。由于这些深层神经网络的规模庞大且架构设计合理，因此能够提取到图像更深层的特征。这些大型网络取得良好成绩的前提是数据量足够大，而本文中掌静脉的数据量远远不够支持训练很深的网络模型。因此，本文在 5.1 节提出了掌静脉数据的扩增方法，通过对数据进行类内扩增和类间扩增，将原始数据集进行有效的扩充，从而为掌静脉特征提取模型的训练提供了数据支撑。

首先，我们对将要用到的深层神经网络做一个简要介绍，以便理解深层神经网络的工作原理和为什么深层神经网络能取得高精度的原因。

1. VGGNet

2014 年，牛津大学的视觉几何组(Visual Geometry Group)和 Google DeepMind 公司的研究员共同提出了 VGGNet^[33]，并且取得了 ILSVRC-2014 比赛的定位任务第一名和分类任务第二名的优异成绩。VGGNet 有 6 种结构，网络结构如表 4-1 所示，从左到右网络逐渐加深，其证明了大量使用很小的卷积核(3×3)后，能够使用更多的激活函数，从而产生更多的非线性变换，使 CNN 对特征的学习能力更强，而且参数量更小；另外，文中还对神经网络深度与精度之间的关系进行了实验探索，证明了增加网络深度可以有效地提升模型精度。

在表 4-1 的 6 种网络结构中，效果突出的还是较深的 VGG-16 和 VGG-19，其都是由传统的卷积层和全连接层构成。其特点是分成 5 组卷积，每一组卷积都有 2-3 个卷积层，且每组内的卷积核数量一致。同时，在每组卷积的尾部都会连接最大池化层(Max Pooling)，以减小特征图的大小。经过权衡，本文选择 VGG-16 作为基础网络。

表 4-1 VGGNet 网络结构^[33]

Table4-1 VGGNet network structure^[33]

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

2. Resnet

为了提升网络性能，网络的层数不断加深。但随着网络加深的同时，就容易出现梯度消失和梯度爆炸的问题，使得神经网络也变得越来越难训练。Kaiming He 等提出的 Resnet^[35](Residual Neural Network)在一定程度上解决了深度网络的优化问题，其通过使用 ResNet Unit 成功训练出了 152 层的神经网络，并在 ILSVRC-2015 比赛中取得冠军，在 top5 上的错误率为 3.57%，效果非常突出，其核心在于所提出的残差单元。如图 4-1 所示，ResNet 结构增加了将输入直接连接到输出的结构，这样网络输出层的梯度可以通过这个连接直接回传到输入层，在一定程度上解决了深度网络中存在的梯度消失问题。从图中结构也可以看出，ResNet 做了一个线性变换，

使得优化目标由学习潜在的理想映射 $H(x)$ 转变为学习 $F(x)=H(x)-x$, 由于梯度能够直接回传到输入层, 也从侧面证明了优化后者更为容易。

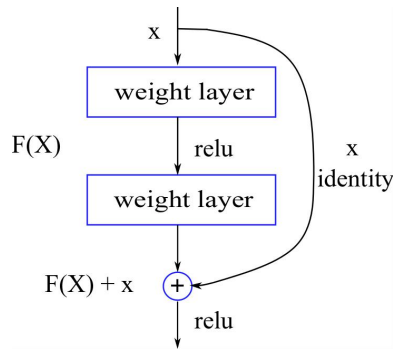


图 4-1 残差网络结构

Fig.4-1 Residual network structure

如表 4-2 所示, 文中共提出了 18、34、50、101 以及 152 层五种网络结构, 由于本文掌静脉图像的数据较少, 因此选择较浅的 Resnet18 网络作为基础网络。

表 4-2 Resnet 网络结构^[35]

Table4-2 Resnet network structure^[35]

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

4.1.3 深层掌静脉特征提取网络设计

上一节介绍了一些经典的深层神经网络结构, 接下来我们将以 Resnet18 和 VGG16 网络为基准, 自行设计深层卷积神经网络结构进行掌静脉特征提取。首先, 针对 Resnet18 网络, 我们对原始的 Residual block 结构进行了改进, 如图 4-2 所示, 原始的 Residual Block 是先进行步长为 2 再进行步长为 1 的卷积, 这里我们将两个卷积的顺序进行调换, 先进行步长为 1 的 3×3 卷积后再进行步长为 2 的 3×3 卷积,

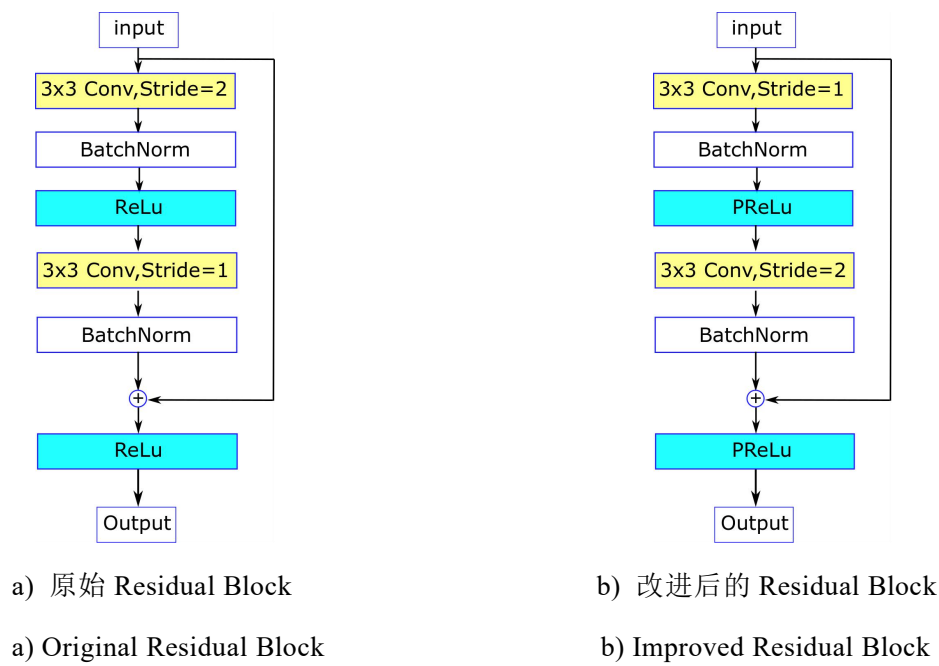


图 4-2 Residual Block 对比

Fig.4-2 Comparison of Residual Block

这样设计的原因是原始 Residual Block 中先进行步长为 2 的卷积会使特征平面缩小一倍，这样输入到下一层卷积层的特征将会有所损失，而调换顺序后是先提取特征再进行降维，这样提取的特征信息更加丰富；另外，原始 Residual Block 采用的是 ReLu 激活函数，我们这里采取 PReLU 激活函数，防止训练中产生梯度消失的现象。

表 4-3 Resnet18-Modified 网络结构

Table4-3 Resnet18-Modified network structure

层类型	卷积核大小/数量/步长	Output
Input	—	150×150×3
Conv1	3×3/64/1	150×150×64
Max Pooling	3×3/-/2	75×75×64
Residual block 1	$\begin{bmatrix} 3\times3 & 64 & 1 \\ 3\times3 & 64 & 2 \end{bmatrix} + \begin{bmatrix} 3\times3 & 64 & 1 \\ 3\times3 & 64 & 1 \end{bmatrix}$	38×38×64
Residual block 2	$\begin{bmatrix} 3\times3 & 128 & 1 \\ 3\times3 & 128 & 2 \end{bmatrix} + \begin{bmatrix} 3\times3 & 128 & 1 \\ 3\times3 & 128 & 1 \end{bmatrix}$	19×19×128
Residual block 3	$\begin{bmatrix} 3\times3 & 256 & 1 \\ 3\times3 & 256 & 2 \end{bmatrix} + \begin{bmatrix} 3\times3 & 256 & 1 \\ 3\times3 & 256 & 1 \end{bmatrix}$	10×10×256
Residual block 4	$\begin{bmatrix} 3\times3 & 512 & 1 \\ 3\times3 & 512 & 2 \end{bmatrix} + \begin{bmatrix} 3\times3 & 512 & 1 \\ 3\times3 & 512 & 1 \end{bmatrix}$	5×5×512
Fully Connected	—/—/—	Embedding
Fully Connected	—/—/—	训练类别数

为了验证改进 Residual Block 的有效性，我们进行了实验对比，具体实验结果见 5.4.3 节。我们将在 Resnet18 基础上改进的网络结构命名为 Resnet18-Modified。其具体的网络结构设计如表 4-3 所示。

对于 VGG16 网络，由于其都是由一些传统卷积层和池化层构成，因此我们对其网络结构并没有进行大的改变。但是由于其参数量很大，而且掌静脉 ROI 图像的输入大小是 150×150 ，网络的原始输入为 224×224 ，因此我们删除了最后一组卷积，其具体的网络结构设计如表 4-4 所示，网络命名为 VGG16-Modified。

表 4-4 VGG16-Modified 网络结构

Table4-4 VGG16-Modified network structure

层类型	卷积核大小/数量/步长	Output
Input	—	$150 \times 150 \times 3$
Conv1	$3 \times 3 / 64 / 1$	$150 \times 150 \times 64$
Conv2	$3 \times 3 / 64 / 1$	$150 \times 150 \times 64$
Max Pooling	$2 \times 2 / \text{—} / 2$	$75 \times 75 \times 64$
Conv3	$3 \times 3 / 128 / 1$	$75 \times 75 \times 128$
Conv4	$3 \times 3 / 128 / 1$	$75 \times 75 \times 128$
Max Pooling	$2 \times 2 / \text{—} / 2$	$37 \times 37 \times 128$
Conv5	$3 \times 3 / 256 / 1$	$37 \times 37 \times 256$
Conv6	$3 \times 3 / 256 / 1$	$37 \times 37 \times 256$
Conv7	$3 \times 3 / 256 / 1$	$37 \times 37 \times 256$
Max Pooling	$2 \times 2 / \text{—} / 2$	$18 \times 18 \times 256$
Conv8	$3 \times 3 / 512 / 1$	$18 \times 18 \times 512$
Conv9	$3 \times 3 / 512 / 1$	$18 \times 18 \times 512$
Conv10	$3 \times 3 / 512 / 1$	$18 \times 18 \times 512$
Max Pooling	$2 \times 2 / \text{—} / 2$	$9 \times 9 \times 512$
Fully Connected	—/—/—	4096
Fully Connected	—/—/—	Embedding
Fully Connected	—/—/—	训练类别数

另外，我们在每个卷积层之后都做了 Batch Normalization^[37]，以加快网络的训练速度和防止模型过拟合。

4.2 轻量级掌静脉特征提取网络

从 4.1.2 节可以看到，深层神经网络网络结构和参数规模都很庞大，而目前掌静脉识别算法大多运行在搭载中低端处理器的嵌入式设备上，无法实时处理深层神经网络这么庞大的计算量，因此有必要设计轻量级网络，实现与深度神经网络相同甚至更高的精度，从而满足神经网络在嵌入式设备上的低时延运行要求。目前，很多轻量级网络都已经能够达到甚至超越深层神经网络的精度，本节我们先对一些轻量级基础网络结构进行简要介绍，通过这些优异的网络结构设计方法来启发我们设计轻量级的掌静脉特征提取网络。

4.2.1 轻量级神经网络结构介绍

1. MobileNet-v1

2016 年 4 月，Google 在 CVPR 会议上发表的 mobilenet-v1^[41]网络提出了用深度分离卷积(depthwise separable convolution^[42])来代替传统卷积，在不降低精度的同时有效地减少了参数量，深度分离卷积由逐层卷积和逐点卷积两部分操作构成。

首先，逐层卷积(depthwise convolution)的操作方式是一个卷积核只与一个输入通道进行卷积，这种卷积方式可以减少大量的参数，但是这种做法仅仅整合了每个通道内部的信息，忽略了输入特征通道之间的相关性；所以在逐层卷积之后需要加上逐点卷积(pointwise convolution)，通过对通道之间的信息进行融合来提升精度。逐点卷积采用的是 1×1 卷积核，因此其参数数量只与输入通道和输出通道的数量有关，从而有效地减少了网络的参数量。

我们通过公式计算来对比深度分离卷积和传统卷积的计算量。如图 4-3(a)所示，假设输入特征图大小为 $D_F \times D_F$ ，其中 M 为输入通道数， N 为输出通道数， D_K 为卷积核的大小，则传统卷积的计算量为 $D_K \times D_K \times D_F \times D_F \times M \times N$ ，深度分离卷积的计算量为 $D_K \times D_K \times D_F \times D_F \times M + D_F \times D_F \times M \times N$ ，两者的比值为式(4-1)所示，

$$\frac{D_K \times D_K \times D_F \times D_F \times M + D_F \times D_F \times M \times N}{D_K \times D_K \times D_F \times D_F \times M \times N} = \frac{1}{N} + \frac{1}{D_K^2} \quad (4-1)$$

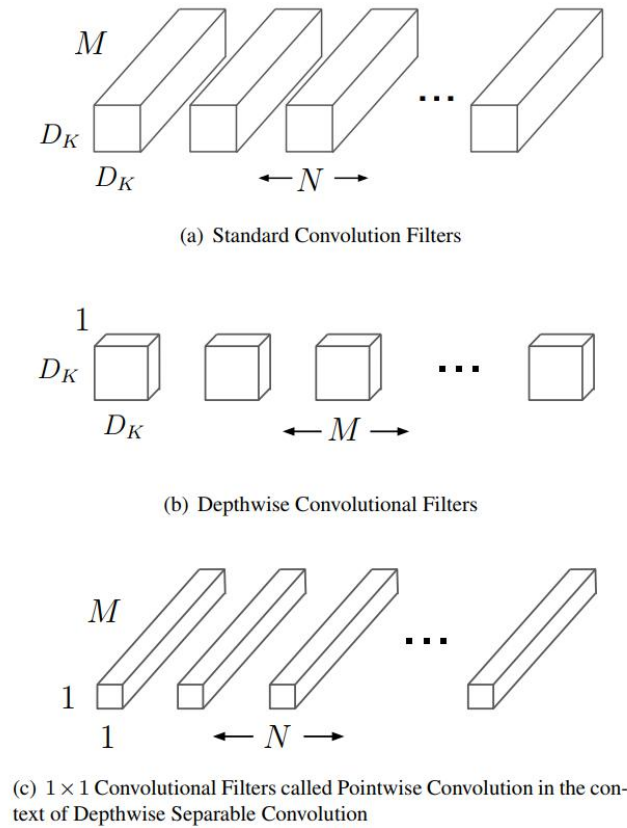


图 4-3 mobilenet-v1 的改进之处^[41]

Fig.4-3 Improvements of mobilenet-v1 network^[41]

因此，与传统的 3×3 卷积相比，深度分离卷积的计算资源消耗可以降到传统卷积的 $1/9$ 左右，同时也减少了参数量，有效地提升了网络的推理速度。

2. Mobilenet-v2

2018 年 1 月，Google 提出了 Mobilenet-v1 的改进版：Mobilenet-v2^[43]，在计算量和参数量均小于 Mobilenet-v1 的情况下，实现了精度的提升。

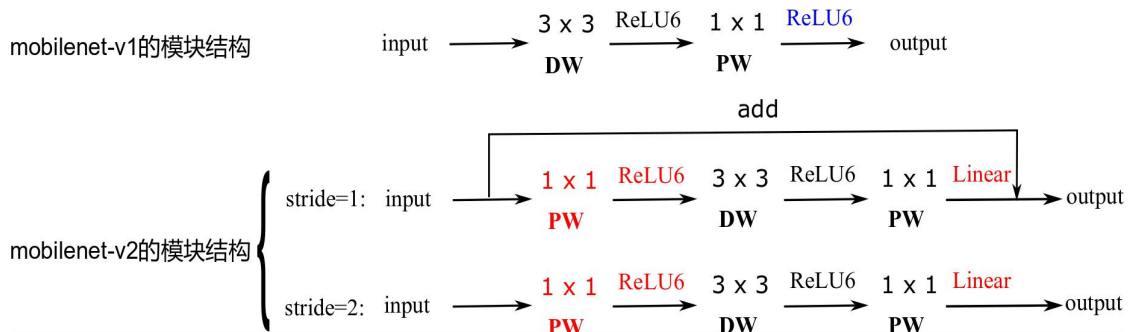


图 4-4 Mobilenet-v2 的改进之处

Fig.4-4 Improvements to mobilenet-v2 network

如图 4-4 所示,与 Mobilenet-v1 的模块相比,Mobilenet-v2 主要有两点改进之处:

1)、由于深度分离卷积是先逐层卷积再进行逐点卷积,通道之间的信息融合没有传统卷积好,因此为了提取到更丰富及有效的特征,在深度分离卷积之前加了一层 1×1 的卷积层进行特征提取,而且这层卷积的输出通道数远大于输入通道数;

2)、为了防止 ReLU 激活函数破坏提取到的特征,在逐点卷积之后不采用 ReLU,而是采用线性激活函数,使提取到的特征能较完整的保存下来。

4.2.2 轻量级掌静脉特征提取网络设计

上文对目前流行的轻量级网络结构的重点模块做了一个简要的回顾,这些网络的设计方法都对接下来的神经网络设计起到了很好的借鉴作用。因此,本节我们将结合上述轻量级基础网络的优点来设计轻量级掌静脉特征提取网络。

首先,通过神经网络的可视化^[40],我们可以了解到,神经网络的底层网络主要是检测一些线条等简单的边缘结构信息,这些基础信息对后续的高维特征信息提取非常重要,因此我们在输入层仍然采用传统卷积对输入层进行有效的特征提取;根据 MobileNet-v1 中分析的,深度分离卷积可以大大减少模型的计算量,同时精度不会产生很大的损失,因此在传统卷积之后的卷积层我们都采用深度分离卷积;另外,由于 Mobilenet-v2 采用了先升维再降维的方法,有效地提升了图像识别的精度,因此决定采用这种有效的 Bottleneck 模块, Bottleneck 模块结构如图 4-5 所示。

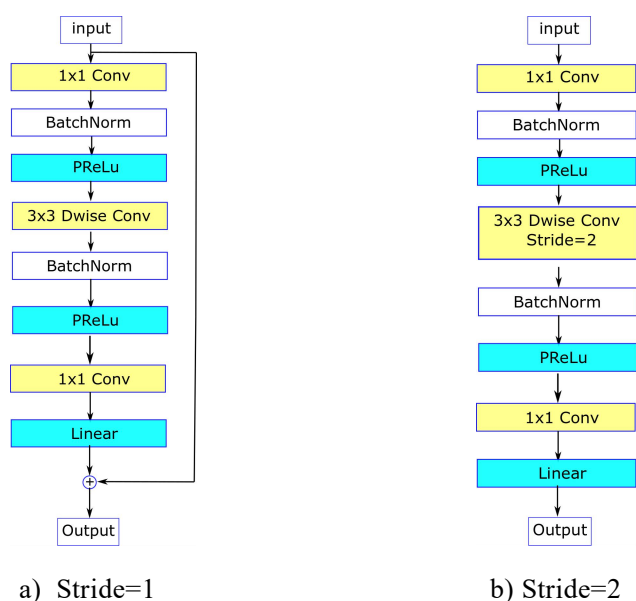


图 4-5 TinyPVNet bottleneck 结构

Fig.4-5 TinyPVNet bottleneck structure

整个轻量级网络 TinyPVNet 的框架设计如表 4-5 所示，表中的 n 代表对应的层重复 n 次，Expansion 代表的是 Bottleneck 模块中第一个 1×1 卷积层升维的倍数。除了 Bottleneck 模块中最后一层采用线性(Linear)激活函数外，网络的其他卷积层之后均采用 PReLU 非线性激活函数。为了加快训练和防止过拟合，每个卷积层之后都进行了 Batch Normalization 操作。

表 4-5 轻量级网络结构 TinyPVNet

Table4-5 Lightweight network structure of TinyPVNet

层类型	n	Expansion	Stride	卷积核大小/数量	Output
Input	—	—	—	—	150×150×3
Conv1	1	—	2	3×3/64	75×75×64
Depthwise Conv1	1	—	1	3×3/64	75×75×64
Bottleneck	5	2	2	—	38×38×64
Bottleneck	1	4	2	—	19×19×128
Bottleneck	6	2	1	—	19×19×128
Bottleneck	1	4	2	—	10×10×128
Bottleneck	2	2	1	—	10×10×128
Conv1×1	1	—	1	1×1/512	10×10×512
Depthwise Conv2	1	—	1	7×7/512	4×4×512
Fully Connected	1	—	—	—	Embedding
Fully Connected	1	—	—	—	训练类别数

4.3 损失函数设计

损失函数 (Loss function) 反映了模型的预测值 $f(x)$ 与真实值 Y 的不一致程度，它是一个非负实值函数，通常用 $L(Y, f(x))$ 来表示。损失函数越小，模型的预测值与真实值越接近，说明模型的鲁棒性就越好。训练时通过每次得到的损失函数值进行反向传播，以此来更新模型的参数，最终使模型的输出尽可能地接近真实值。本节将简要介绍一些经典的分类损失函数，并对其设计原理进行分析。

4.3.1 Softmax Loss

在卷积神经网络用于分类任务时，Softmax 函数经常做分类器使用，它是从适用于二分类问题的 Logistic 分类器扩展而来。Softmax 函数是将最后多个神经元的输出映射到 $(0, 1)$ 区间内，且所有输出值和为 1。公式为：

$$S(i) = \frac{e^{f_{y_i}}}{\sum_{j=1}^K e^{f_{y_j}}} \quad (4-2)$$

如图 4-4 所示，我们假设 X 是网络最后一层全连接层的输入，也就是提取的特征。其中特征 X 是 $N \times 1$ 的向量， W 是全连接层的参数，是个 $T \times N$ 的矩阵，这个 N 和 X 的 N 对应， T 表示最终的类别数。最后全连接层就是执行 $W \times X$ 来得到一个 $T \times 1$ 的向量（也就是图中的 $\text{logits}[T \times 1]$ ），这个向量的值域是 $(-\infty, +\infty)$ 。然后经过 softmax 函数将向量映射到 $(0, 1)$ 区间，也就是图中的 $\text{prob}[T \times 1]$ ，这个向量的每个值表示这个样本属于每个类的概率。

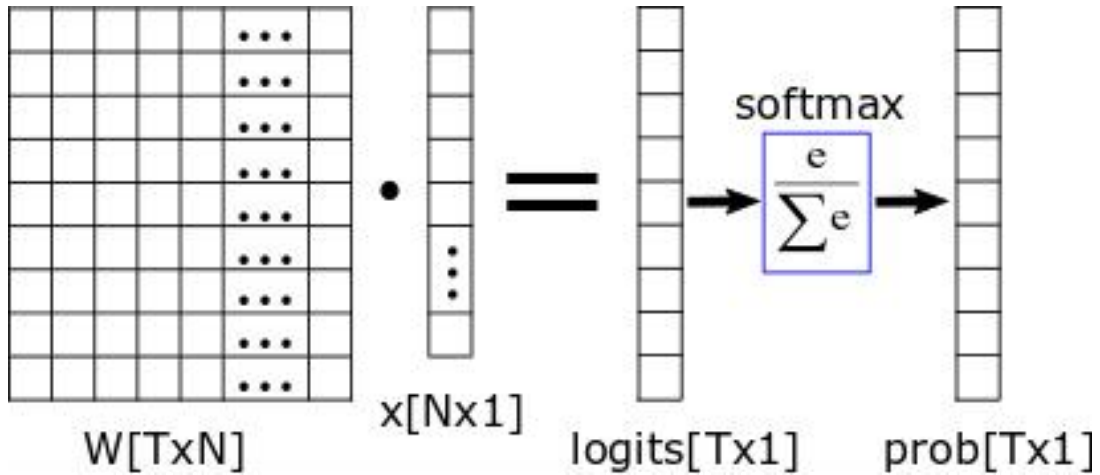


图 4-4 Softmax 分类

Fig.4-4 Softmax classification

在得到每个类别的 softmax 输出概率值之后，我们采用交叉熵损失函数计算损失值，其表达式如式(4-3)所示，其中 m 代表每批训练的样本数， $j \in [1, n]$ ，其中 n 代表类别数， f_{y_i} 代表全连接层最后一层的输出，即 $f_{y_i} = W_{y_i}^T x_i + b_{y_i}$ ；所以交叉熵损失函数的最终表达式为式(4-4)所示。

$$L_{\text{softmax}} = \frac{1}{m} \sum_{i=1}^m L_i = \frac{1}{m} \sum_{i=1}^m -\log \frac{e^{f_{y_i}}}{\sum_{j=1}^n e^{f_j}} \quad (4-3)$$

$$L_{\text{softmax}} = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}} \quad (4-4)$$

交叉熵损失函数在分类问题上应用的非常广泛，很多模型只使用交叉熵损失就可以达到不错的精度，但其缺点也很明显：只关注类间可分，类内方差依然较大，因此很容易产生类间距离比类内距离小的情况。

4.3.2 Center Loss

为了改进交叉熵损失类内方差大的缺点，研究学者提出了 **Center Loss**^[44]，通过聚拢同一类别的特征，减小类内方差，从而相当于增加了不同类之间的间距，能够有效地提升分类精度。**Center Loss** 的函数表达式如式(4-5)所示，其中 x_i 代表训练样本的特征向量， c_{y_i} 为其对应类别的特征中心，这样就能通过约束样本特征向量与特征中心的距离来减小类内方差。

$$L_{\text{Center}} = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (4-5)$$

在整个神经网络训练过程中，如果每次都通过计算所有样本特征来更新类别的特征中心，其计算量以及占用内存都太大，因此文献^[44]提出了每次只统计一个 **mini-batch** 类别中心的方法，但由于 **mini-batch** 中不能包含训练集的所有类别，所以作者采用类似滑动平均统计的方法来更新类别中心，使得类别中心的更新更加稳定。如图 4-5(a)所示，交叉熵损失只要求类间可分，同类之间的样本波动性大，处于决策边界附近的样本会出现类间距离小于类内距离的情况；而加入 **Center Loss** 后，同类样本之间的方差明显减小，不同类样本之间的距离也相应地增加，样本离决策边

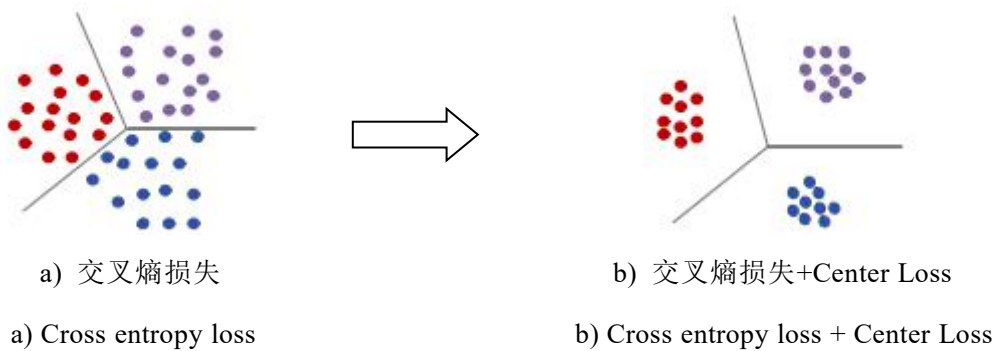


图 4-5 交叉熵损失与 Center Loss 的区别

Fig.4-5 Difference between cross entropy loss and Center Loss

界的距离也更大，结果更稳定。

本文中，我们将综合使用交叉熵损失和 Center Loss 来训练神经网络，新的损失函数如式(4-6)所示，其中超参数 λ 用于调节 Center Loss 占总损失的比重， λ 越大，则同类之间的特征聚集越紧凑，但优化难度也随之增加。因此本文将 λ 设置为经验值 0.003。

$$L_{total} = L_S + \lambda L_{Center} = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}} + \frac{\lambda}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2 \quad (4-6)$$

4.3.3 L-softmax Loss 和 A-softmax Loss

为了使同类之间的特征聚集更紧凑，L-softmax Loss^[45]提出了一种更严格的条件。具体地，我们首先回顾交叉熵损失的函数表达式(4-4)，由于向量的内积等于对应向量模与向量之间夹角余弦值的乘积，所以我们可以将式(4-4)重写成式(4-7)。

$$L = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{\|W_{y_i}\| \cdot \|x_i\| \cdot \cos \theta_{y_i} + b_{y_i}}}{\sum_{j=1}^n e^{\|W_j\| \cdot \|x_i\| \cdot \cos \theta_j + b_j}} \quad (4-7)$$

这里我们以二分类为例，原始 softmax loss 是要能满足式(4-8)，从而得到 x 的正确分类结果（ x 所属类别为 1）。

$$W_1^T x > W_2^T x \quad (4-8)$$

即

$$\|W_1\| \|x\| \cos(\theta_1) > \|W_2\| \|x\| \cos(\theta_2) \quad (4-9)$$

然后 L-softmax Loss 引入了一个角度裕量 m （ m 为正整数），从而能严格约束不等式(4-9)，即

$$\|W_1\| \|x\| \cos(\theta_1) > \|W_1\| \|x\| \cos(m\theta_1) > \|W_2\| \|x\| \cos(\theta_2), 0 \leq \theta_1 < \frac{\pi}{m} \quad (4-10)$$

这样，特征向量与权重向量之间的夹角要是原夹角的 $1/m$ 才能被正确分类，分类的条件变得更加严格。因此式(4-7)可以写成式(4-11)，如图 4-6(b)所示，L-softmax Loss 使得同类之间的特征更加紧凑，不同类之间有更宽的决策边界。式(4-12)中，当 m 越大时，分类条件越严格，同时也越难收敛。这里我们设置 m 为 4。

$$L = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{\|W_{y_i}\| \cdot \|x_i\| \cdot \phi(\theta_{y_i})}}{e^{\|W_{y_i}\| \cdot \|x_i\| \cdot \phi(\theta_{y_i})} + \sum_{j \neq y_i} e^{\|W_j\| \cdot \|x_i\| \cdot \cos \theta_j}} \quad (4-11)$$

其中

$$\varphi(\theta) = (-1)^k \cos(m\theta) - 2k, \quad \theta \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m}\right]; k = 0, 1, 2, \dots, m-1 \quad (4-12)$$

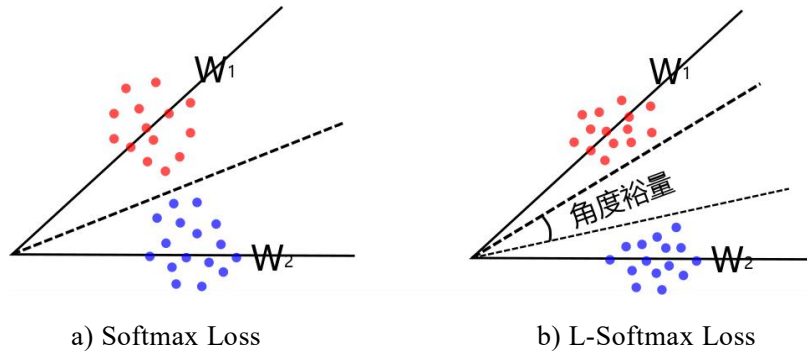


图 4-6 Softmax Loss 和 L-Softmax Loss 的特征分布

Fig.4-6 Feature distribution of Softmax Loss and L-Softmax Loss

与 L-Softmax Loss 相比，A-softmax Loss^[46]进一步提出了更严格的分类条件。如图 4-7(a)所示，L-Softmax Loss 在训练时可能受到夹角角度和权重向量模长的共同影响，导致分类条件不严格。为了解决这个缺点，A-Softmax Loss 提出将权重进行归一化以及偏置置为零，即 $\|W\|=1$ 和 $b=0$ ，改进后的公式为式(4-13)。如图 4-7(b)所示，对权重向量进行归一化后，训练分类时则只取决于权重向量 W 和特征向量 x 之间的角度，分类条件更加严格，精确度也随之提升。

$$L = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{\|x_i\| \cdot \phi(\theta_{y_i})}}{e^{\|x_i\| \cdot \phi(\theta_{y_i})} + \sum_{j \neq y_i} e^{\|x_i\| \cdot \cos \theta_j}} \quad (4-13)$$

其中

$$\varphi(\theta) = (-1)^k \cos(m\theta) - 2k, \quad \theta \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m}\right]; k = 0, 1, 2, \dots, m-1 \quad (4-14)$$

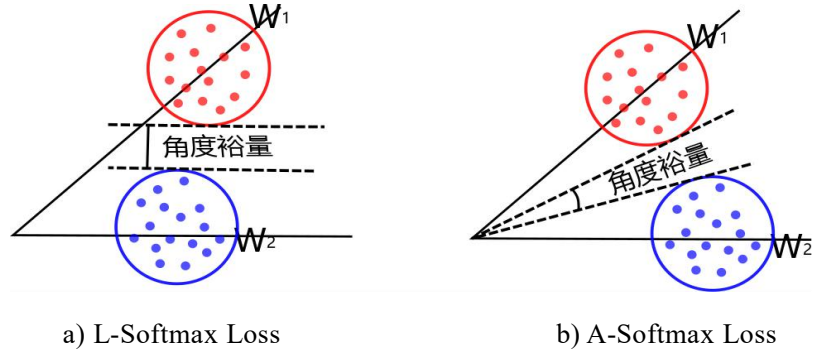


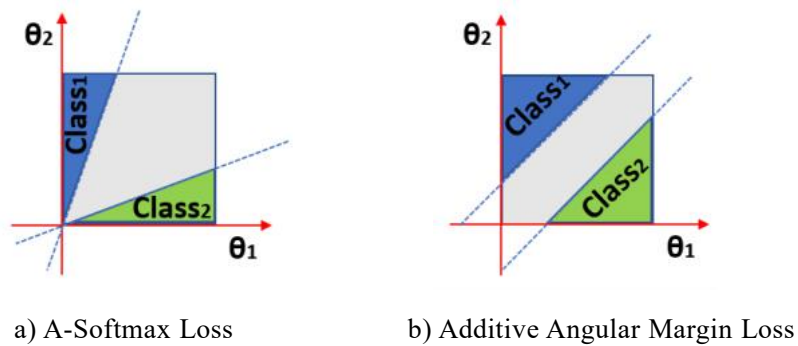
图 4-7 L-Softmax Loss 和 A-Softmax Loss 的特征分布

Fig.4-7 Feature distribution of L-Softmax Loss and A-Softmax Loss

4.3.4 Additive Angular Margin Loss

与 A-Softmax Loss 相似，Additive Angular Margin Loss^[47]提出通过增加角度距离进行更严格的分类，并对特征向量 x 也进行归一化，这样训练时仅注重从数据中得到的角度信息，而不考虑特征向量的值。我们知道，不同清晰度的图片其特征范数也不相同，高清晰度图片提取出来的特征范数大，低清晰度图片提取出来的特征范数小，Additive Angular Margin Loss 对特征向量 x 也进行归一化以后，相当于增加了低清晰度图片的权重，使得低清晰度图片获得网络更多的注意力，这样网络进行反向传播时，低清晰度的图片会产生更大的梯度，因而有利于网络对难分样本的学习。Additive Angular Margin Loss 的表达式如式(4-15)所示，式中 s 是一个超参数，通过映射到更大的超球面来助于分类。并且通过图 4-8(b)可以看出，Additive Angular

$$L = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{s \cdot (\cos(\theta_{y_i} + m))}}{e^{s \cdot (\cos(\theta_{y_i} + m))} + \sum_{j \neq y_i} e^{s \cdot \cos \theta_j}} \quad (4-15)$$

图 4-8 A-Softmax Loss 和 Additive Angular Margin Loss 的决策边界^[47]Fig.4-8 Decision boundary of A-Softmax Loss and Additive Angular Margin Loss^[47]

Margin Loss 的决策边界比 A-Softmax Loss 更合理, 其不受角度 θ 的影响, 始终存在 m 大小的决策边界, 这样使得训练时更容易收敛。这里我们将超参数 s 和 m 分别设置为 64 和 0.5。

另外, 还有一些其他的常用 Loss 函数, 如对比损失函数(Contrastive Loss)、Triplet Loss 等度量学习的损失函数。但这些 Loss 在训练时都需要构造大量的样本对, 过程繁琐且复杂, 很容易造成训练时不稳定以及无法收敛, 在实际应用中其效果并不是很好。因此, 本文采用 Softmax Loss、Softmax Loss+Center Loss、A-Softmax Loss 以及 Additive Angular Margin Loss 四种 Loss 函数进行实验, 5.4.4 节详细分析了其对掌静脉数据的分类效果。

4.4 特征向量的相似性度量方法

在一般的图像分类等任务中, 训练集和测试集所包含的图像类别数是相同的, 这种情况称作闭集识别(Close-set Identification)。此时, 测试集直接使用网络的最后一层(分类层)即可得出具体的类别。而本文的掌静脉识别属于开集识别(Open-set Identification)任务, 即训练集与测试集的样本类别不一致, 此时无法直接使用最后一层得出样本所属的类别。由于网络的最后一层其实是对网络倒数第二层输出的样本特征进行分类, 因此在这种开集识别任务中, 我们提取网络倒数第二层的输出(Embedding)作为特征向量, 然后使用距离度量函数来计算样本之间的距离, 通过距离来度量特征向量之间的相似性, 并以此判断样本是否属于同一类。

对于特征向量的相似性度量, 常用的距离度量函数有欧式距离、余弦距离以及汉明距离等。其中汉明距离经常用于对向量编码后的相似性度量, 主要应用在基于 LBP 以及 NBP 特征匹配的掌静脉识别算法中^[48]; 而神经网络提取的特征向量是以浮点数表示, 因此不适合使用汉明距离。这里我们主要采用余弦距离和欧式距离进行特征向量的相似性度量。具体地, 对于 Softmax Loss 和 A-Softmax Loss, 其分类概率是基于权重向量和特征向量的内积进行计算的, 因此采用余弦距离进行度量; 而 Center Loss 是通过欧式距离来计算类内方差, 因此我们采用欧式距离进行计算; 对于 Additive Angular Margin Loss 来说, 由于对特征向量进行了归一化, 我们可以根据式(4-16)和式(4-17)得出, 采用欧式距离和余弦距离是等价的, 因此我们也采用欧式距离进行度量。

$$\text{European_distance} = \sqrt{(x_1 - x_2)^2} = \sqrt{\|x_1\|^2 + \|x_2\|^2 - 2x_1 \cdot x_2} = \sqrt{2 - 2x_1 \cdot x_2} \quad (4-16)$$

$$\text{Cosine_distance} = \frac{x_1 \cdot x_2}{\|x_1\| \cdot \|x_2\|} = x_1 \cdot x_2 \quad (4-17)$$

4.5 模型压缩

卷积神经网络在实现高精度的同时，也带来了占用空间大的缺点。本文中如 VGG16-Modified 的模型大小为 432 MB，其中模型的体积几乎全部来自于神经元连接的权重，因此我们要对模型的权重进行有效的压缩，模型才能较好地应用到嵌入式等移动设备上。一般地，神经网络模型的参数都是略微不同的 32 位浮点数，常见的如 zip 这样的压缩方法并不能对模型体积进行有效的压缩。如果能在不影响模型准确率的情况下，对模型内部的参数采用其他简单数值类型进行存储，就能大大地减少模型的存储空间。因此我们这里采用量化压缩的方法对模型进行压缩。

另外，值得注意的是，神经网络的每层神经元权重都在一定范围内趋向于正态分布。如图 4-9 所示，我们统计了 TinyPVNet 模型第一层卷积层的权重分布图，可以得出其权重的分布范围为(-1.0,0.75)，并且趋于高斯正态分布。

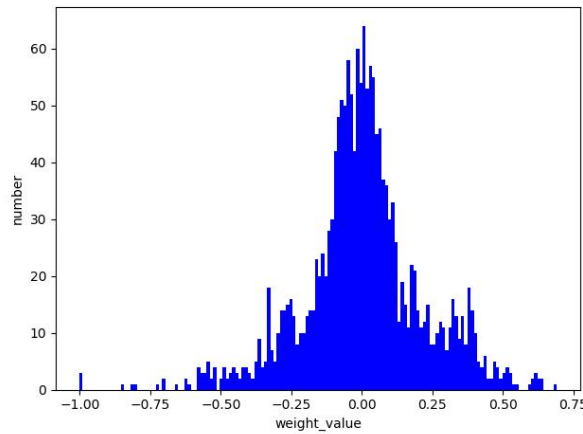


图 4-9 TinyPVNet 第一层卷积层的权重分布图

Fig.4-9 Weight distribution of TinyPVNet first convolution layer

因此我们采用的量化压缩方案如下：首先存储每层参数的最小值和最大值，然后把最小值和最大值形成的区间线性分成 256 个离散值，这样每个浮点数都可以用 8 位（二进制）整数来表示，即将每个浮点数近似为离得最近的那个离散值。具体的计算方法如式(4-18)所示，其中 W_{input} 为原始权重， W_{out} 为量化后的权重。

$$W_{\text{out}} = \frac{W_{\text{input}} - W_{\text{min}}}{W_{\text{max}} - W_{\text{min}}} \times 255 \quad (4-18)$$

经过量化操作后，就能将 32 位浮点数参数转换成 8 位整数来存储，理论上可以将模型体积缩小 75%，在网络推理时再将 8 位整数转换回 32 位浮点数进行推理计算，得到模型的输出值。由于神经网络对噪声具有很好的鲁棒性，所以经过量化后模型的精度往往不会下降的很明显。具体的实验结果见 5.4.6 节。

4.6 本章小结

本章首先介绍了一些经典深层卷积神经网络和轻量级神经网络结构的设计思路，并在此基础上自行设计了两个深层卷积神经网络 Resnet18-Modified、VGG16-Modified 和一个轻量级网络 TinyPVNet 用于掌静脉图像的特征提取，对于 Resnet18 网络，我们对其原始的 Residual Block 模块进行了改进；然后介绍了分类问题中常用损失函数，如 Softmax Loss、Center Loss、A-Softmax Loss 以及 Additive Angular Margin Loss 等，并分析了它们的设计原理；接着介绍了本文中所使用的特征向量相似性度量方法；最后对本文中所使用的模型量化压缩方法原理进行了阐述。

第五章 实验结果及分析

5.1 实验数据准备

5.1.1 掌静脉数据扩增

由于本文使用的 CASIA 数据集的每一类只有 6 张静脉图像,整个数据集也只有 1200(200*6)张图像,这些数据完全不能满足 CNN 训练的要求。因此,为了让 CNN 学习出具有区分力的特征,本文需要对掌静脉图像进行数据扩增,其中类内数据的扩增要求扩增后的数据保持类内差异性不大,而类间数据的扩增则要求扩增后的数据能够保持类间差异性较大,以达到可区分的目的。下面将简要介绍本文提出的数据扩增方法。其中,类内扩增主要包括平移、剪切、旋转、伽马变换等方法,增强后的效果图如图 5-1 所示,采用这些类内扩增方法主要是考虑到掌静脉图像在 ROI 截取的过程中会出现一定的误差与偏移,因此可以利用这些增强方法生成新样本去模拟可能出现的误差,从而提高模型对这些误差和偏移的鲁棒性;

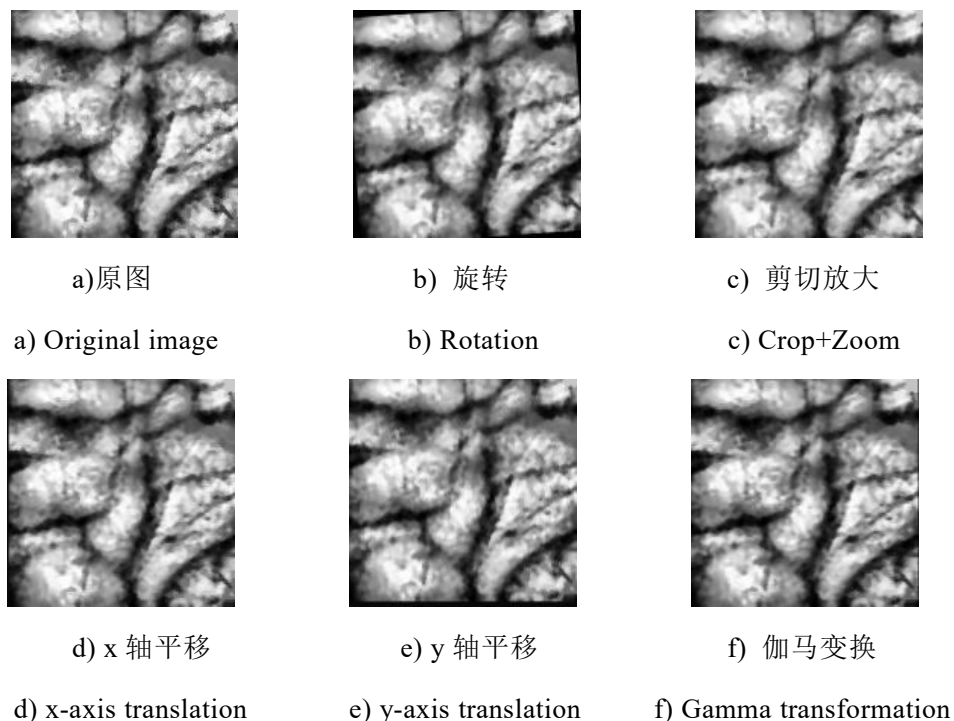


图 5-1 类内扩增效果图

Fig.5-1 Intra-class data augmentation result

在进行类内扩增后，每一类的数据量都得到了有效的扩充，但整个数据集仍然只有 200 类。类别数太少会导致神经网络学习出来的特征没有很强的区分力。因此为了提高 CNN 的泛化能力，本文还提出了类间扩增的方法，包括对图像进行水平镜像和垂直镜像，图像旋转等，有效地扩充了掌静脉图像类别数，类间增强的效果图如图 5-2 所示。

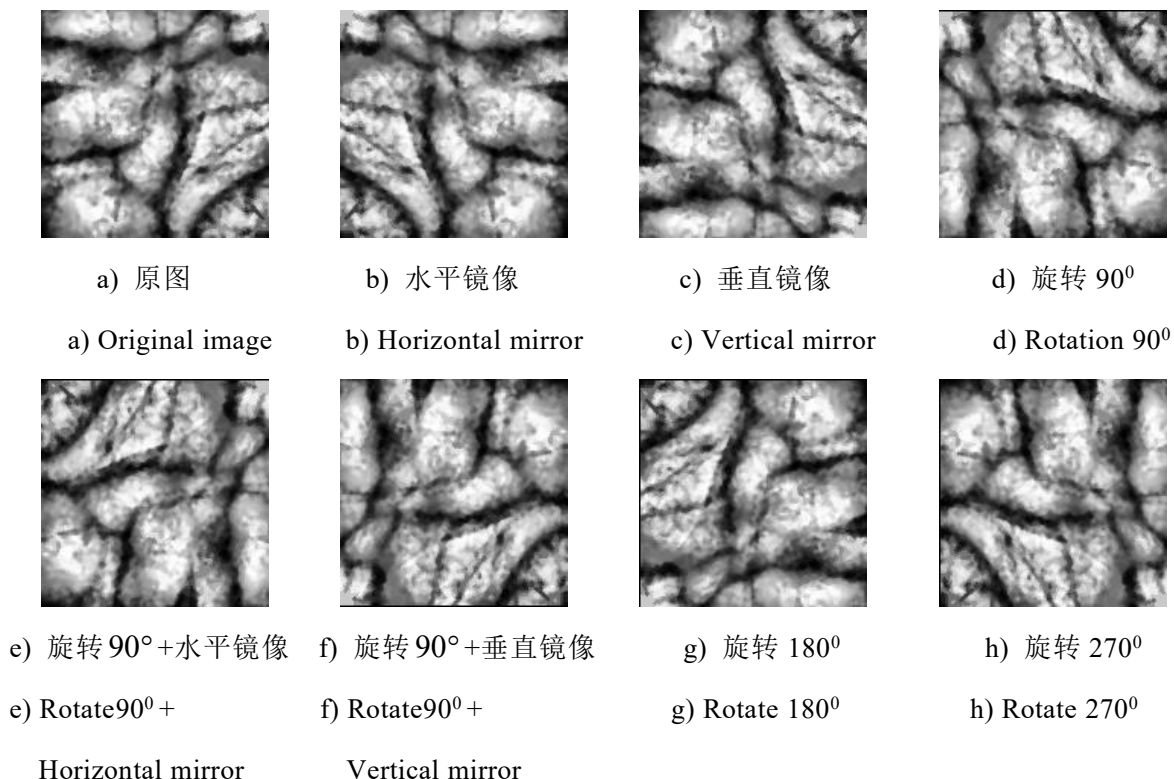


图 5-2 类间扩增效果图

Fig.5-2 Inter-class data augmentation result

将掌静脉数据进行扩增后，能有效地防止卷积神经网络训练时出现过拟合问题，提高模型的泛化能力。

5.1.2 实验数据准备

本文首先使用训练好的关键点定位网络对手掌进行关键点定位，然后按照 2.3 节的方法对掌静脉图片进行 ROI 提取，然后采用 2.4 节介绍的对比度限制自适应直方图均衡(CLAHE)方法对 ROI 图像进行图像增强，采用中值滤波进行图像去噪，最终得到处理后的 ROI 掌静脉图像。接着，本文对原始 CASIA 数据库进行训练集和测试集的划分，如表 5-1 所示，其中训练集占总类别数的 70%，测试集占总类别数的 30%。

表 5-1 CASIA 数据库的训练集和测试集划分

Table5-1 Training set and test set of CASIA database

训练集类别数	140
训练集图片数	840 (140×6)
测试集类别数	60
测试集图片数	360 (60×6)

划分好数据集后，再按照 5.1.1 节中的方法对训练集图片进行数据扩增。首先进行类内数据扩增，将原图分别进行 2 次旋转，2 次随机剪切以及 2 次 gamma 变换，这样能生成 7 倍数据量(包括原图)；在此基础上对生成的所有数据再分别进行 1 次 x 轴平移和 1 次 y 轴平移，最后每一类的数据量都扩充到原来的 21 倍；然后再在类内扩增的基础上进行类间扩增，由于图像转置的效果与将图像逆时针旋转 90° 再进行垂直镜像的效果相同，因此对每一类图片，分别进行逆时针旋转 0、90、180、270 度，在此基础上分别对旋转 0 度和 90 度的图片分别进行水平镜像和垂直镜像，这样每一类就能扩增到 8 类(4+2*2)。

扩增完数据后，再对扩增完的数据进行划分，结果如表 5-2 所示。数据扩充完后，训练集总共有 1120 类，每一类 120 张掌静脉图像。其中验证集是从数据扩充后的每一类中随机抽取 10% 的图片。由于数据扩增产生的是伪静脉数据，目的是用来训练神经网络，防止模型过拟合。因此为了提升算法的说服力，本文中的测试集不参与数据扩增。

表 5-2 数据扩增后的训练集和验证集划分

Table 5-2 Training set and verification set after data augmentation

训练集类别数	1120 (140×8)
训练集总数	127680 (1120×114)
验证集类别数	1120 (140×8)
验证集总数	13440 (1120×12)
测试集类别数	60
测试集总数	360 (60×6)

经过数据扩增后，图像分布的多样性得到了保证，这样便有利于 CNN 学习到更丰富的特征，从而提升 CNN 的泛化能力。

5.2 实验性能指标

5.2.1 类内比对和类间比对

类内比对是指同一只手掌的两张不同图片进行比对，类间比对是指不同类的两张手掌图片进行比对。掌静脉识别的目标就是要使同类之间的最大距离要比不同类之间的最小距离要小，因此存在寻找最优阈值问题。本文中，对测试集分别进行类内比对和类间比对，并保证两者比对次数的平衡性。假设有 n 类手掌，每一类有 m 张图片，则所有的类内比对组合数为 $n \times C_m^2$ ，而所有的类间比对组合数为 $m \times m \times n \times (n-1)/2$ 。可以看到，当 $n > m$ 时，类间比对的组合数将远远大于类内比对组合数，此时若将全部的类间组合和类内组合都进行比对测试，得到的整体准确率将远大于真实准确率，例如当类间比对与类内比对的组合比为 9:1 时，如果将所有的比对组合都预测为不同类，也能得到 90% 的准确率。因此在测试时需要保证两者比对次数相等，才能得到真实的准确率。

本文以类内比对组合数 $n \times C_m^2$ 为基数，在所有类间比对组合中随机选取 $n \times C_m^2$ 对，组成测试集。具体地，本文使用的测试集中每一类有 6 张掌静脉图像，因此每一类可以进行 15 次类内比对，整个测试集可进行 900(60×15)次类内比对；而总类间比对组合有 31320(6×6×60×59/2)对，因此本文从中随机抽取 900 对，组成类间比对。因此本文实验对测试集分别进行了 900 次类内比对和 900 次类间比对。

5.2.2 误识率、拒识率和等误率

误识率(False Accept Rate, FAR)指的是在比对过程中将不同类的图片判断成相同类的概率，而拒识率(False Rejection Rate, FRR)则指的是在比对过程中将相同类的图片判断成不同类的概率，因此误识率和拒识率都是越低，代表掌静脉识别算法的性能越好。但两者是属于对立面的，当不断降低误识率时，拒识率会不断升高；而不断降低拒识率的同时，误识率则会不断升高；因此本文采用等误率(Equal Error Rate, EER)来衡量算法的性能。等误率越小，说明掌静脉识别算法的识别精度越高。

对于测试集，本文采用十折交叉验证的方法来计算模型的最终等误率。首先将测试集的数据平均分成十份，每次取其中九份作为训练集，剩余一份作为测试集，这样总共能进行十次交叉验证，每次验证都能在测试集上得到一个等误率，最后求取十次等误率的平均值，即为最终的等误率。

5.3 实验设置及模型训练参数设置

本文中使用的深度学习框架为 Tensorflow，GPU 型号为 GeForce GTX1080Ti，显存为 12G。

在参数设置方面，对于深层网络和轻量级网络的训练，本文都使用相同的参数。参数设置如下：

- 1) Batch Size 大小设置为 64；
- 2) 权重初始化采用 Xavier 方法，参数服从正态分布，权重正则化系数为 0.0005；
- 3) 优化策略选择 Momentum，其冲量系数设置为 0.9；
- 4) 初始学习率为 0.1，每经过 6 个 epoch 后学习率衰减 5 倍，整个训练迭代过程为 30 个 epoch；
- 5) 对于 VGG16-Modified 网络，由于其全连接层参数量太大，为了防止过拟合，本文在全连接层中加入了 dropout，比例为 0.8。

5.4 实验结果分析

5.4.1 不同特征向量维度的性能对比

本文训练和测试的网络是第四章中提出的三种网络结构(深层网络 VGG16-Modified 和 Resnet18-Modified 以及轻量级网络 TinyPVNet)。本文首先对网络采用不同的特征向量维度进行测试，其中特征向量的维度分别设置为 128 维、256 维和 512 维，测试的等误率曲线如图 5-3 所示。可以看到，当特征向量维度设置为 128 维时，其在三个网络上的精度都不如 256 维和 512 维。其原因是掌静脉图像中蕴含的特征信息丰富，128 维的低维特征向量难以完整地表征丰富的掌静脉特征信息。因此后面进行分析将只对比 256 维和 512 维的特征向量。

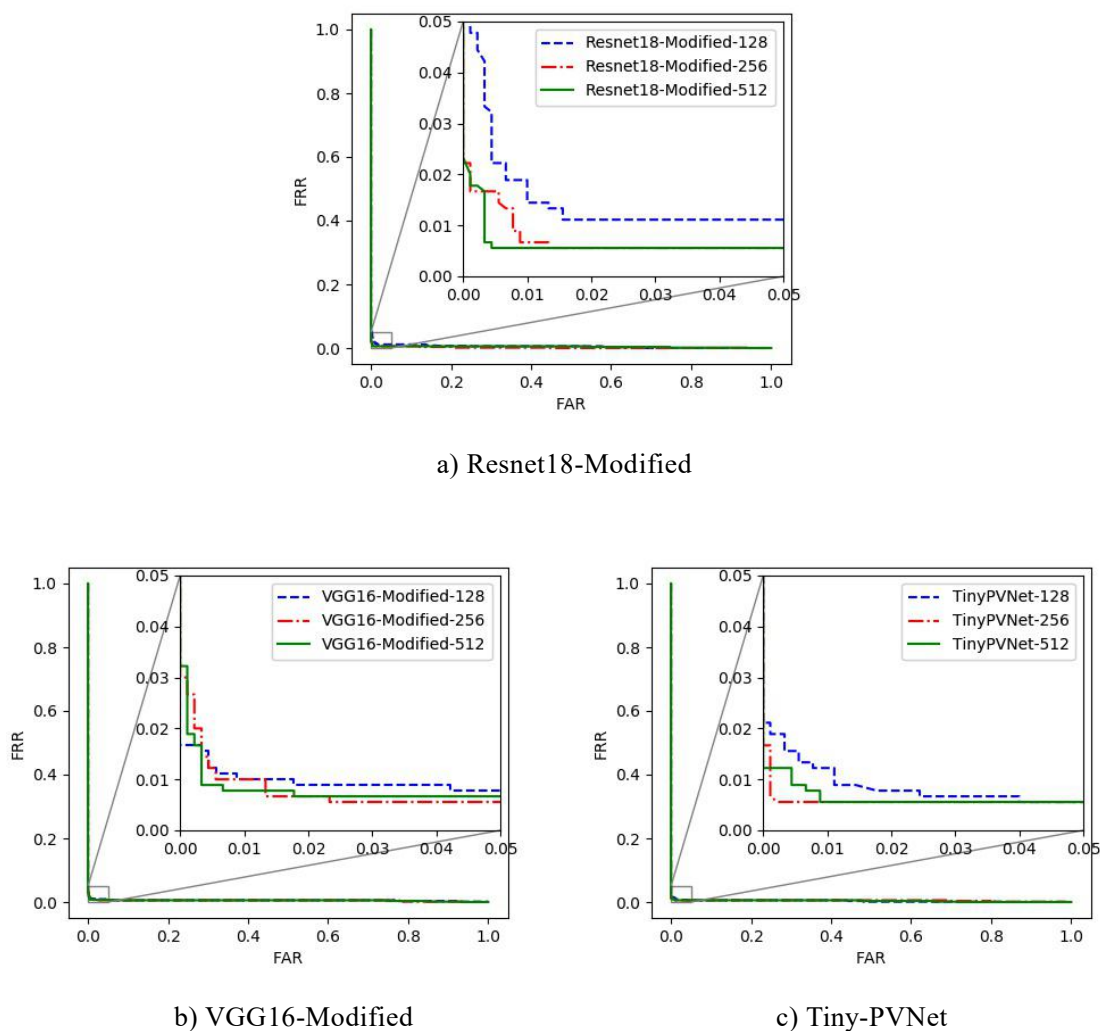


图 5-3 不同维度特征向量的性能对比

Fig.5-3 Performance comparison of feature vectors with different dimensions

具体的等误率数值如表 5-3 所示，从表中可以看出，当采用深层网络 Resnet18-Modified 和 VGG16-Modified 时，512 维的识别精度都优于 256 维；而采用轻量级网络 TinyPVNet 时，采用 256 维的特征向量却比 512 维特征向量精度要高。其原因是由于轻量级网络 TinyPVNet 的网络结构设计得更加精炼，因此网络提取的特征向量更加紧凑。而深层网络由于网络规模较大，因此提取得特征较为冗余，所以在深层网络 Resnet18-Modified 和 VGG16-Modified 上，512 维比 256 维的识别精度要高。

表 5-3 不同网络结构的等误率

Table5-3 Equal error rate of different network structures

网络结构	特征向量维度	等误率(%)
Resnet18-Modified	128	1.45
Resnet18-Modified	256	0.83
Resnet18-Modified	512	0.56
VGG16-Modified	128	1.00
VGG16-Modified	256	0.85
VGG16-Modified	512	0.74
TinyPVNet	128	1.08
TinyPVNet	256	0.51
TinyPVNet	512	0.67

因此，本文将深层网络 Resnet18-Modified 和 VGG16-Modified 的特征维度设为 512 维，而轻量级网络 TinyPVNet 的特征维度设置为 256 维。

5.4.2 不同网络结构的性能对比

本文中提出了两个深层网络 Resnet18-Modified、VGG16-Modified 以及一个轻量级网络 TinyPVNet，下面将对比不同网络结构的性能。

从表 5-4 中可以得到，在特征向量为 512 维时，Resnet18-Modified 的等误率要低于 VGG16-Modified，说明 Resnet 的残差结构设计还是要优于传统卷积组成的网络。另外，轻量级网络 TinyPVNet 在 256 维时取得了最好的识别精度，其等误率为

表 5-4 不同网络结构的性能对比

Table5-4 Performance comparison of different network structures

网络	Resnet18-Modified(512)	VGG16-Modified(512)	TinyPVNet(256)
等误率(%)	0.56	0.74	0.51

0.51%，甚至超越了深层网络 Resnet18-Modified 的识别精度，证明了本文设计的轻量级网络 TinyPVNet 具有良好的特征提取能力。所以在计算资源紧张的情况下，本文设计的轻量级网络 TinyPVNet 完全可以替代一些基础的深层网络，甚至取得更好的识别精度。

5.4.3 改进 Residual Block 对算法精度的提升

在 4.1.3 节中，我们对 Resnet18 中的 Residual Block 模块进行了改进，下面我们将通过实验结果分析改进 Residual Block 后对精度提升的有效性。如图 5-4 所示，蓝色曲线代表原始 Residual Block 模块的等误率，红色曲线代表改进 Residual Block 模块后的等误率。可以看到，无论特征向量设置成 256 维还是 512 维，我们改进 Residual block 模块后的等误率都要低于原始 Residual Block 模块的等误率，证明了本文提出的改进 Residual Block 模块对掌静脉识别精度提升的有效性。具体的等误率数值如表 5-5 所示。

表 5-5 改进 Residual Block 与原始 Residual Block 的等误率对比 (%)

Table5-5 Comparison of the equal error rate between the improved Residual Block and the original

Residual Block(%)

	256 维	512 维
原始 Block(Resnet18)	0.94	1.00
改进 Block 后(Resnet18-Modified)	0.83	0.56

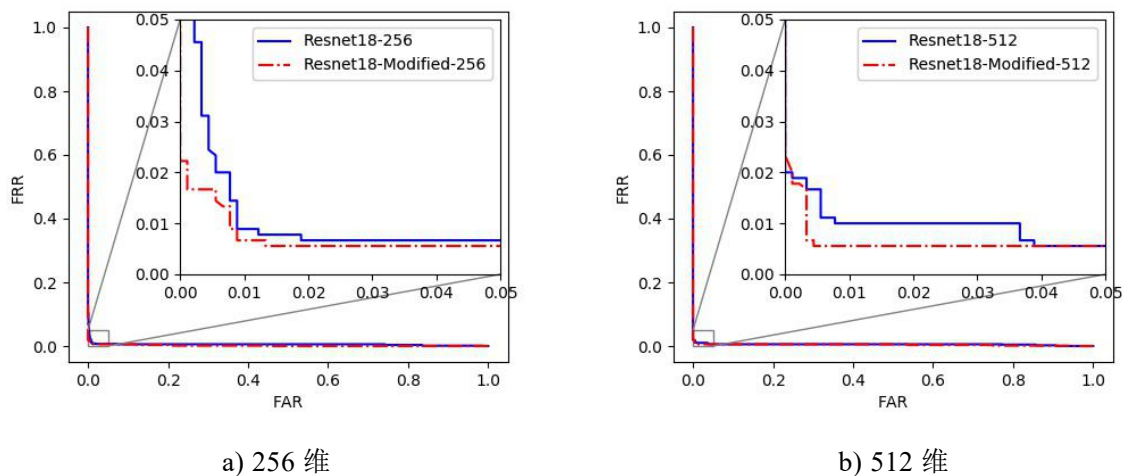
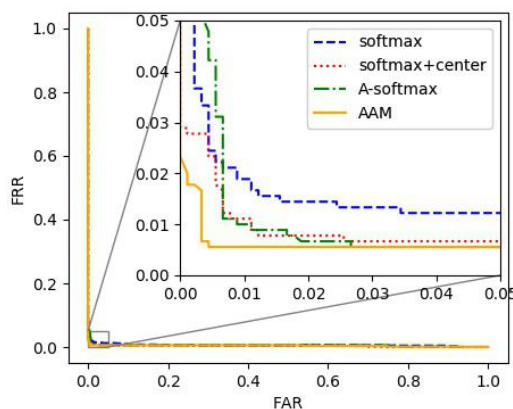


图 5-4 改进 Residual Block 与原始 Residual Block 的等误率对比

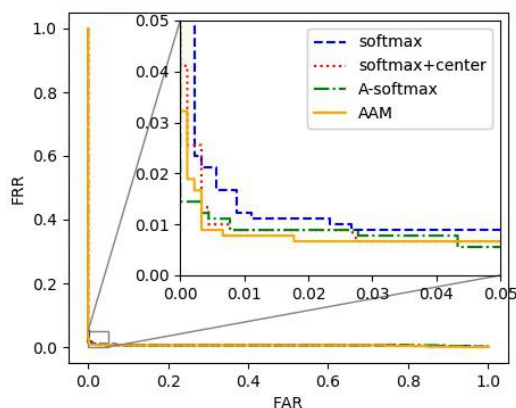
Fig.5-4 Comparison of the equal error rate between the improved Residual Block and the original Residual Block

5.4.4 不同损失函数的性能对比

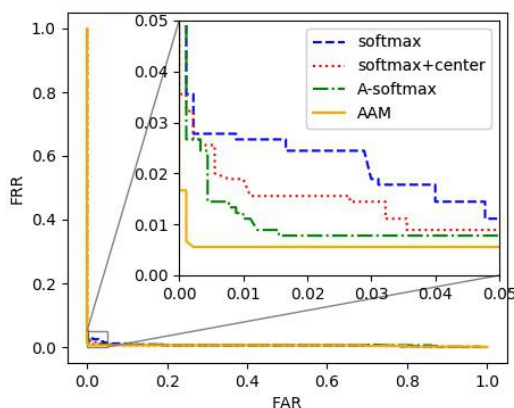
本文 4.3 节已经对 Softmax Loss、Center Loss、A-Softmax Loss 以及 Additive Angular Margin Loss 等损失函数的原理作了详细介绍，下面将对比这些损失函数在不同网络结构上的性能。不同损失函数的等误率曲线如图 5-5 所示。



a) Resnet18-Modified



b) VGG16-Modified



c) TinyPVNet

图 5-5 不同损失函数在不同网络上的等误率对比

Fig.5-5 Comparison of equal error rates of different loss functions on different networks

可以看到，在三个不同的网络上，Softmax Loss 的精度与其他三种 loss 函数相比都有较大的差距，说明传统的 Softmax Loss 仅仅是在训练时将样本进行正确分类，缺乏对类内和类间距离的约束。而非接触式掌静脉图像属于类内特征差异较大的图像，因此 Softmax Loss 在测试集上的识别效果不是很好。另外，分类条件更加严格的 Additive Angular Margin Loss 在三个网络上的精度都优于 A-Softmax Loss 和

Center Loss，证明了特征向量归一化对提升识别精度的有效性。不同损失函数的等误率具体数值如表 5-6 所示。

表 5-6 不同损失函数在不同网络上的等误率（%）

Table5-6 Equal error rate of different loss functions on different networks(%)

	Resnet18-Modified(512)	VGG16-Modified(512)	TinyPVNet(256)
Softmax Loss	1.46	1.18	2.42
Softmax + Center Loss	1.11	0.95	1.53
A-Softmax Loss	1.07	0.96	1.19
Additive Angular Margin Loss	0.56	0.74	0.51

5.4.5 与其他算法性能对比

我们对本文的算法性能与其他常用算法的性能进行对比，对比结果如表 5-7 所示。可以看到，本文使用神经网络进行特征提取后取得的等误率低于大部分传统算法，这也证明了本文采用卷积神经网络进行关键点定位和特征提取方案的可行性和有效性。另外，神经网络一般要求大数据集才能取得较高的精度，而本文在 CASIA 单数据集上训练已经能够取得 0.51% 的测试等误率，证明了本文提出的数据扩增方法可以有效地降低模型过拟合的风险。

表 5-7 与其他算法性能对比(CASIA 掌静脉)

Table 5-7 Performance comparison with other algorithms (CASIA)

文献	方法	等误率(EER)
Zhou Y, Ajay Kumar. ^[51] (2011)	Hessian Phase	2.24%
Zhou Y, Kumar A. ^[52]	Local Radon transform	1.03%
Yuan W, Wu W. ^[53] (2013)	BLOCK+PLS	0.66%
Kang W, Wu Q. ^[49] (2014)	Improved LBP method on mutual foreground	2.53%
Yan X, Kang W, Deng F ^[26] (2015)	Multi-sampling and feature-level fusion	0.16%
Wang P ^[50] (2017)	Principal curvature and template matching	1.965%
本文方法	CNN+Additive Angular Margin Loss	0.51%

5.4.6 模型压缩及网络时间测试

首先，我们对各个网络的推理时间进行统计，我们对每个网络进行 100 次时间测试，然后求其平均值。其统计结果如表 5-8 所示，其中推理设备采用的是 CPU(Intel i5-3210M,2.5G 主频)。可以看到，由于深层网络 VGG16-Modified 中有大量的传统卷积层，因此整个网络的耗时非常长。另外，轻量级网络 TinyPVNet 的推理时间要低于深层网络 Resnet18-Modified 和 VGG16-Modified，其推理耗时为 58ms，即使在嵌入式等中低端处理器上，轻量级网络也完全可以满足实时性的要求。

表 5-8 网络推理时间

Table5-8 Network inference time

网络	Resnet18-Modified(512)	VGG16-Modified(512)	TinyPVNet(256)
推理时间(ms)	75	343	58

接着，我们采用 4.5 节中的模型压缩方法对轻量级网络 TinyPVNet 进行量化压缩，表 5-9 显示的是量化前和量化后的精度、推理时间和模型体积对比，可以发现，经过量化操作后，模型 TinyPVNet 的体积由原来的 12.1M 降低到 3.4M，压缩了 71.9%，而等误率增加了 0.16%，在可接受的误差范围内。因此，这也证明了模型量化压缩对模型部署到嵌入式等移动设备上提供了很强的实用性。

表 5-9 TinyPVNet 量化前后对比

Table5-9 Comparison of original TinyPVNet and TinyPVNet quantization

	模型体积(M)	推理时间(ms)	等误率(%)
量化前	12.1	58	0.51
量化后	3.4	55	0.67

5.5 本章小结

本章通过对掌静脉特征提取网络进行多组实验结果分析，证明了自行设计的卷积神经网络提取特征的有效性。首先，针对掌静脉训练数据量不足的问题，本章提出了掌静脉数据扩增方法，包括类内扩增和类间扩增方法，有效地扩充了数据集，为掌静脉特征提取网络的训练提供了数据支撑。接着介绍了模型训练时的参数设置，最后对多组实验结果进行了分析。首先对不同特征向量维度、不同网络结构所取得的实验结果进行了分析，对于深层网络 Resnet18-Modified、VGG16-Modified，512

维特征向量的效果最好，而对于轻量级网络 TinyPVNet，256 维的紧凑型特征向量效果更好。另外，本文设计的轻量级网络 TinyPVNet 取得的精度甚至超越了深层网络，证明了轻量级网络 TinyPVNet 提取掌静脉特征的有效性；接着分析了本文对 Resnet18 网络中 Residual Block 模块改进的有效性，通过改进 Residual Block 模块，提升了 Resnet18-Modified 网络的识别精度；然后对比了不同损失函数训练的结果，与 Softmax Loss、A-Softmax Loss 和 Center Loss 相比，Additive Angular Margin Loss 通过对特征向量归一化可以有效地提升识别精度；最后，我们还对模型压缩结果进行了测试，通过模型压缩，我们能将轻量级网络 TinyPVNet 的模型体积压缩到 5M 以下，这也为轻量级网络应用到嵌入式等移动设备上提供了可行方案。

总结与展望

传统的接触式掌静脉识别方法在手掌出现较大偏移和旋转时性能较差，无法很好地应用在非接触式掌静脉识别中。因此，本文重点研究了非接触式掌静脉识别过程中的手掌关键点定位问题和掌静脉特征提取问题。本文的主要工作总结如下：

1) 针对传统算法对非接触式手掌图像难以准确定位关键点的问题，采用深度学习方法进行关键点定位，设计了两级卷积神经网络级联来定位手掌关键点；并且为了提高手掌关键点定位的精确度，在 MSE Loss 上进行改进，提出了 Modified MSE Loss 进行训练，有效地提升了手掌关键点的定位精度，证明了提出的 Modified MSE Loss 的有效性；

2) 针对非接触式掌静脉数据库数据集较少，训练神经网络容易产生过拟合的问题，提出了类内扩增数据和类间扩增数据的方法，有效地扩充了原始数据库，为后续掌静脉特征提取网络的训练提供了数据支撑；

3) 在分析了一些经典分类神经网络结构的优点后，自行设计了两个深层特征提取网络 Resnet18-Modified 和 VGG16-Modified，并对 Resnet18 的 Residual Block 模块结构进行改进，实验结果证明了模块改进后对提升识别精度的有效性；另外，探索了不同特征向量维度对识别精度的影响，实验证明，在大型深层网络上，512 维的特征向量精度最高，而在轻量级网络上，256 维的紧凑型特征向量精度更高；

4) 针对传统深层神经网络参数量大，计算耗时长的问题，本文设计了一个轻量级网络 TinyPVNet，主要采用深度可分离卷积结构，有效地降低了网络规模和参数量，在 CASIA 数据集上取得了 0.51% 的等误率；另外，本文还对轻量级网络 TinyPVNet 进行了模型量化，有效地压缩了模型体积，从而能够将轻量级网络部署在嵌入式等计算资源受限的平台上。

通过本文的研究工作，解决了非接触式掌静脉识别过程中的手掌关键点定位问题和掌静脉特征提取问题。但整个研究工作还有进一步的提升空间：

1) 提升掌静脉图像的质量。掌静脉图像的质量是进行识别的基础，尤其是对于非接触式采集方式而言，其容易受光照等外界因素影响。因此，高质量的掌静脉图像是提高识别准确率的关键，可以从光源、传感器等方面进行改进；

2) 扩充掌静脉数据库规模。本文采用的 CASIA 数据库，每个波长段只有 1200 张图像，虽然经过我们提出的数据扩增方法进行扩增后，增加了样本的多样性和复

杂性。但使用扩增方法扩增出的数据毕竟是伪数据，与真实场景下采集的数据仍有所差别。因此，扩充掌静脉数据库的样本量，才能进一步提升算法的性能；

3) 提升图像预处理效果。图像的预处理质量对掌静脉识别的精度也有很大的影响，本文采取了传统的图像预处理方法，后续可以考虑尝试深度学习的方法对掌静脉图像进行预处理。

参考文献

- [1] 张钊. 掌静脉识别算法研究[D]. 电子科技大学, 2017.
- [2] Zhang D, Guo Z, Lu G, et al. An Online System of Multispectral Palmprint Verification[J]. Instrumentation and Measurement, IEEE Transactions on, 2010, 59(2):480-490.
- [3] Im S.K.,Choi H.S.,Kim S.W. A direction-based vascular pattern extraction algorithm for hand vascular pattern verification[J].ETRI Journal, 2003, 25(2): 101-108.
- [4] Badawi A M. Hand Vein Biometric Verification Prototype:A Testing Performance and Patterns Similarity[J]. International Journal of Biomedical Sciences,2008:34-38.
- [5] Wang L., Leedham G. A thermal hand vein pattern verification system[M]. Berlin:Springer, 2012:58-65.
- [6] 林喜荣,庄波. 人体手背血管图像的特征提取及匹配[N]. 清华大学学报, 2003, 43(2):13-16.
- [7] 韩笑. 人体手背静脉识别算法研究[D]. 吉林大学, 2007.
- [8] Kejun Wang, Yan Zhang, Zhi Yuan, Dayan Zhuang. Hand Vein Recognition Based on Multi Supplemental Features of Multi-Classifer Fusion Decision[J]. 2006 IEEE International Conference on Mechatronics and Automation, 2006:1790-1795.
- [9] Yuhang Ding, Dayan Zhuang, Kejun Wang. A Study of Hand Vein Recognition Method[J]. Proceedings of IEEE. International Conference on Mechatronics and Automation. Niagara Falls, Canada, July 2006:2106-2110.
- [10] Pham T D, Park Y H,Nguyen D T, et al. Nonintrusive Finger-Vein Recognition System Using NIR Image Sensor and Accuracy Analyses According to Various Factors[J]. Sensors, 2015,15(7):1866-1894.
- [11] 余成波, 张进, 张一萌, 等. 基于核 Fisher 鉴别分析的指静脉识别. 重庆邮电大学学报:自然科学版[J]. 2012,24:90-95.
- [12] 王科俊, 管凤旭, 刘靖宇, 等. 归一双向加权静脉识别方法[J]. 模式识别与人工智能, 2011,24(3):417-424.
- [13] 王科俊, 刘靖宇. 基于相对位置和角度的静脉识别方法[J]. 华中科技大学学报:自然科学版, 2011,(5):96-99.

- [14] 王科俊, 马慧. 指纹与指静脉双模态识别融合方法[J]. 模式识别与人工智能, 2013,25(4):669-675.
- [15] 李德健. 具备抗旋转能力的指静脉识别算法研究[D]. 华南理工大学, 2017.
- [16] 胡慧. 基于卷积神经网络的手指静脉认证算法研究[D]. 华南理工大学, 2018.
- [17] 唐溯. 基于深度学习的指静脉识别算法研究[D]. 华南理工大学, 2018.
- [18] 李铁刚. 静脉识别算法研究[D]. 吉林大学, 2007.
- [19] 李强. 掌静脉身份识别技术的理论与实验研究[D]. 华中科技大学, 2010.
- [20] 马欣. 手掌静脉身份识别关键技术研究[D]. 北京邮电大学, 2018.
- [21] 颜学葵. 掌静脉识别算法研究[D]. 华南理工大学, 2015.
- [22] 罗敏. 手指静脉认证系统的设计与实现[D]. 电子科技大学, 2016.
- [23] 杨园园. 基于多特征的掌静脉识别算法研究[D]. 电子科技大学, 2018.
- [24] Shigeru Sasaki, et al. Business Expansion of Palm Vein Pattern Authentication Technology[J]. FUJITSU Sci. Tech.J. 2008, 41(3):341-347.
- [25] Yan X, Deng F, Kang W. Palm Vein Recognition Based on Multi-algorithm and Score-Level Fusion[C]. Seventh International Symposium on Computational Intelligence and Design. IEEE, 2015:441-444.
- [26] Yan X, Kang W, Deng F, et al. Palm vein recognition based on multi-sampling and feature-level fusion[J]. Neurocomputing, 2015, 151(151):798-807.
- [27] CASIA Palmprint Database version [J]. <http://biometrics.idealtest.org/>.
- [28] 吴微, 苑玮琦, 林森, 等. 手掌静脉识别中感兴趣区域的选择与定位研究[J]. 光电子·激光, 2013(1):152-160.
- [29] 李威. 非接触成像方式下手掌特征提取方法研究[D]. 沈阳工业大学, 2013.
- [30] 佟海滨. 手掌静脉识别系统[D]. 沈阳工业大学, 2012.
- [31] Sivanandam S. A Novel Biometric system for Person Recognition Using Palm vein Images [J]. International Journal on Computer Science & Engineering, 2013, 5(8).
- [32] Krizhevsky A., Sutskever I., Geoffrey E.H. ImageNet Classification with Deep Convolutional Neural Networks[A]. Advances in Neural Information Processing Systems[C], 2012:1-9.
- [33] Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. ImageNet Challenge, 2014:1-10.

- [34] Christian Szegedy, Wei Liu, Yangqing Jia, et al. Going deeper with convolutions[J]. arXiv preprint, 2014:1409.4842.
- [35] He K., Zhang X., Ren S., et al. Deep Residual Learning for Image Recognition[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. 2016:770-778.
- [36] Gao Huang,Zhuang Liu,et al. Densely Connected Convolutional Networks[J]. arXiv preprint, 2016:1608.06993.
- [37] Ioffe S., Szegedy C. Batch Normalization:Accelerating Deep Network Training by Reducing Internal Covariate Shift[A]. International Conference on Machine Learning[C]. 2015:448-456.
- [38] Szegedy C., Vanhoucke V., Ioffe S., et al. Rethinking the Inception Architecture for Computer Vision[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. 2016:2818-2826.
- [39] Szegedy C., Ioffe S., Vanhoucke V., et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[A]. AAAI Conference on Artificial Intelligence[C], 2017:4278-4284.
- [40] Zeiler M.D., Fergus R. Visualizing and Understanding Convolutional Networks[A]. European Conference on Computer Vision[C] 2014:818-833.
- [41] Howard A.G., Zhu M., Chen B., et al. MobileNets:Efficient Convolutional Neural Networks for Mobile Vision Applications[J]. arXiv preprint, 2017: 1704.04861.
- [42] Chollet F. Xception:Deep Learning with Depthwise Separable Convolutions[J]. arXiv preprint, 2016:1610.02357.
- [43] Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L.C. MobileNetV2: Inverted Residuals and Linear Bottlenecks[J]. arXiv preprint, 2018:1801.04381.
- [44] Wen Y., Zhang K., Li Z., et al. A discriminative feature learning approach for deep face recognition[A]. European Conference on Computer Vision[C]. Springer, 2016:499-515.
- [45] Liu W, Wen Y, Yu Z, et al. Large-Margin Softmax Loss for Convolutional Neural Networks[C]. Proceedings of The 33rd International Conference on Machine Learning. 2016: 507-516.
- [46] Weiyang Liu, Yandong Wen, et al. SphereFace: Deep Hypersphere Embedding for

Face Recognition[A].Computer Vision and Pattern Recognition[C].2017:6738-6746.

[47] Jiankang Deng, Jia Guo, et al. ArcFace: Additive Angular Margin Loss for Deep Face Recognition[J]. arXiv preprint, 2018:1801.07698.

[48] 张文辉. 掌静脉识别系统的研究与实现[D]. 电子科技大学, 2017.

[49] Kang W, Wu Q. Contactless Palm Vein Recognition Using a Mutual Foreground-Based Local Binary Pattern[J].IEEE Transactions on Information Forensics and Security, 2014,9(11):1974-1985.

[50] 王鹏壮. 非接触式掌静脉识别研究[D]. 北京交通大学, 2017.

[51] Zhou Y, Ajay Kumar. Human Identification using Palm-Vein Images[J]. IEEE Transactions on Information Forensics and Security, 2011.

[52] Zhou Y, Kumar A. Contactless palm vein identification using multiple representations[C]. International Conference on Biometrics: Theory Applications & Systems. 2010:1-6.

[53] Yuan W, Wu W. Non-contact palm Vein biometric recognition based on block and partial least squares[J]. Chinese Journal of Scientific Instrument, 2013,34(7):1470-1478.

学位论文独创性声明

本人郑重声明：所呈交的学位论文是我个人在导师的指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明，并表示了谢意。本人依法享有和承担由此论文所产生的权利和责任。

论文作者签名：刘彪 日期：2019年5月31日

学位论文版权使用授权声明

本学位论文作者完全了解学校有关保存、使用学位论文的规定：“研究生在广东工业大学学习和工作期间参与广东工业大学研究项目或承担广东工业大学安排的任务所完成的发明创造及其他技术成果，除另有协议外，归广东工业大学享有或特有”。同意授权广东工业大学保留并向国家有关部门或机构送交该论文的印刷本和电子版本，允许该论文被查阅和借阅。同意授权广东工业大学可以将本学位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、扫描或数字化等其他复制手段保存和汇编本学位论文。保密论文在解密后遵守此规定。

论文作者签名：刘彪 日期：2019年5月31日

指导教师签名：刘彪 日期：2019年5月31日

致谢

光阴似箭，三年的研究生生活即将结束。一路走来，我收获了很多，在此论文完成之际，真心感谢我的老师、家人和同学。

首先我要对我的导师谢康教授表达衷心的感谢，感谢您这三年来对我的悉心关怀和教导。谢康老师渊博的专业知识，严谨的工作态度，刻苦钻研的精神都对我产生了深远的影响，让我受益匪浅。谢康老师不仅在科研上为我指点迷津，在生活上也对我悉心关照，教给我很多人生道理。

其次，我还要感谢实验室的姜海明教授，夏宏燕副教授以及彭迪、申方成老师，感谢你们在我撰写论文期间的悉心指导。感谢实验室同学刘文、谭义和师弟师妹们（陈廷阔、胡畅、肖红周、高飞、孟令龙、胡丹、张木儒、袁伟超、丘焕然、唐常钦、李锦豪、文博）的支持与陪伴，感谢你们对我课题实验的帮助。

再次，我还要感谢三年的研究生室友刘汉健、邹智胜和李宽元，正因为有了你们的包容和理解，才造就了寝室和谐的氛围。在撰写论文阶段，我们也互相激励，共同进步。

另外，我还要感谢我的家人和女友，感谢你们多年来的默默理解与支持，让我能顺利完成三年的学习生涯。

最后，衷心感谢百忙之中参与论文评审和答辩的各位专家，向您致以我最诚挚的谢意。