

基于改进的 ResNet 手指静脉识别

汪凯旋^{1*}, 陈光化^{1,2}, 褚洪佳¹

¹上海大学微电子研究与开发中心, 上海 200444;

²上海大学机电工程与自动化学院, 上海 200444

摘要 针对指静脉提取困难和识别精度不够高等问题, 提出了一种基于 ResNet 改进的手指静脉识别方法。首先, 使用深度超参数化卷积(DO-Conv)代替网络中的传统卷积, 在减少模型参数的同时提高了网络识别率。然后, 将空间注意力模型(SAM)和挤压激励块(SE-Block)融合, 应用于改进的残差网络(ResNet)提取图像在通道和空间域上的细节特征。最后, 使用标签平滑的交叉熵(LSCE)损失函数来训练模型, 实现自动校准网络防止分类出现误差。实验结果表明, 改进后的模型不易受到图片质量的影响, 在公开数据库 FV-USM 和 SDUMLA 上的识别精度分别达到 99.4919% 和 99.4485%, 较之前的网络在精度上有明显提高。

关键词 图像处理; 指静脉识别; 注意力机制; 残差网络; 标签平滑; 损失函数

中图分类号 TP389.1

文献标志码 A

doi: 10.3788/LOP202158.2010002

Finger Vein Recognition Based on Improved ResNet

Wang Kaixuan^{1*}, Chen Guanghua^{1,2}, Chu Hongjia¹

¹Microelectronics R&D Center, Shanghai University, Shanghai 200444, China;

²School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China

Abstract An improved finger vein recognition method based on ResNet is proposed to solve the problem of finger vein extraction difficulties and insufficient recognition accuracy. First of all, depthwise over-parameterized convolution (DO-Conv) is used to replace the traditional convolution in the network, while reducing the model parameters and improving the network recognition rate. Then, the spatial attention module (SAM) and squeeze-and-excitation block (SE-block) are fused and applied to an improved ResNet to extract the detailed features of the image in the channel and spatial domain. Finally, label smoothed cross entropy (LSCE) loss function is used to train the model in order to automatically calibrate the network to prevent errors in classification. The experimental results show that the improved network model is not easily affected by the image quality. And the recognition accuracy of FV-USM and SDUMLA can reach 99.4919% and 99.4485%, respectively, which is significantly higher than that of the previous network. Compared with other models, the proposed method has a significant improvement on the recognition accuracy.

Key words image processing; finger vein recognition; attention mechanism; residual network; label smoothing; loss function

OCIS codes 100.3008; 100.4996; 150.1135

1 引言

随着社会的发展, 身份识别变得尤为重要, 传统

的身份识别技术已经难以满足高安全性要求, 发展生物特征识别技术是大势所趋。现有的生物特征识别技术有人脸、虹膜、指纹、掌纹和指静脉识别等^[1]。

收稿日期: 2020-11-30; 修回日期: 2020-12-15; 录用日期: 2021-01-02

基金项目: 国家自然科学基金(61671285)

通信作者: *18361258215@163.com

指静脉具有高防伪性,隐藏在皮肤下面难以窃取,可以为用户提供更好的安全性和隐私性。但是,在实际应用时,指静脉识别^[2-3]会受到光照、遮挡、角度等因素的影响。因此,有效提取辨识信息是指静脉识别的难点之一。

传统的识别方法存在着一些问题,如在低质量图像上识别效果不佳,容易受噪声、光照等因素的影响。使用一些降维的方法,也可能会漏掉重要的信息^[4-7]。与传统方法相比,卷积神经网络对噪声具有鲁棒性,因此卷积神经网络的方法受预处理的影响小^[8]。卷积神经网络使用大量数据来训练模型,训练后的模型更具有泛化性。近几年,随着深度学习的发展,基于卷积神经网络(CNN)的指静脉识别方法开始崭露头角^[9-12]。Hong 等^[13]将 VGG-16 网络框架用在指静脉识别上,在训练大量数据后,尽管取得了很好的结果,但是模型的计算复杂度较大,训练时间过长。Qin 等^[14]提出使用全卷积网络提取静脉特征,在特征提取上有很强的能力,但是对图像的细节不敏感,容易出现分类错误的情况。Yang 等^[15]首次将生成对抗网络用于静脉的提取和验证,在精度和等误率上有显著的提高,但是网络存在训练困难且不稳定的问题。Song 等^[16]使用了一种深度 DenseNet 网络,将差分图像作为网络输入,使训练和验证不易受到噪声的影响,但是训练时间较长。Radiz 等^[17]提出一种有融合卷积的四层 CNN 进行指静脉识别,虽然获得了良好的识别效果,但是他们的方法和实验没有使用公共数据库进行评估。袁丽莎等^[18]研究的结合随机森林和神经网络进行掌静脉识别,减少了存储空间,降低了分类误差,有较好的表现,但是图像需要进行预处理,耗时较长。陶志勇等^[19]提出了改进的 AlexNet 进行手指静脉识别,在训练耗时和识别准确率上有较好表现,但是识别结果容易受到图像质量的影响。

为了获得更好的指静脉识别效果,本文在 ResNet 网络基础上进行改进,首先将 ResNet 网络减少至 8 层,使用深度超参数化卷积(DO-Conv)替代传统的卷积,减少网络计算,提升网络性能。然后,将空间注意力模型(SAM)和挤压激励块(SE-Block)双注意力机制进行融合,在通道和空间上进行关注,有效地提取了图像在信道和空间上的细节信息。最后对交叉熵损失函数使用标签平滑,使分类概率结果向正确分类靠近,进一步提高图像分类的准确性。

2 ResNet 网络模型

ResNet 是由微软研究院发布的一种深层卷积神经网络架构。在网络模型训练时,网络越深,越容易出现梯度消失和梯度爆炸现象,从而影响模型的性能。为了解决这种问题,使用残差模块(Residual block)构建直连通道,直接将输入信息绕道传到输出,提高了网络性能。残差网络也因为这个优点被广泛地应用于图像识别的领域。残差块结构如图 1 所示,输入 X_L 直接经过两个卷积层就可以得到输出 X_{L+1} ,图中的 $F(X_L)$ 为求解网络的残差映射函数,使用残差块得到的结果为

$$X_{L+1} = X_L + F(X_L) \quad (1)$$

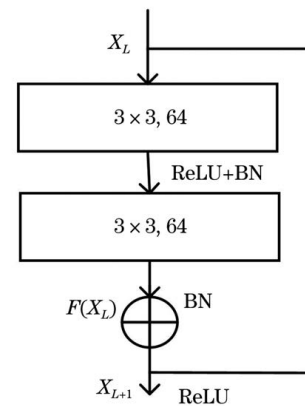


图 1 残差模块

Fig. 1 Residual block

在进行手指静脉识别时,一般用纹理特征作为特征识别。但是有些不同的个体在手指静脉纹理特征上有较高的相似度,所以需要更加细微的细节特征进行区分。残差网络在网络性能和网络参数保持不变的基础上,可以学到新的特征,这种特点就使残差网络更适用于手指静脉图像数据库,因此本文在 ResNet 网络的基础上进行改进。基于 ResNet-18,在运行时间和识别精度方面进行优化。

3 基于 ResNet 网络的改进

3.1 深度超参数化卷积

卷积神经网络用卷积核来提取手指静脉的特征,对传统卷积层增加一个额外的深度卷积构成深度超参数化卷积,这样的组合构成过度参数化,增加了网络可学习参数,提高了提取手指静脉纹理特征的质量,从而使网络模型更适合于指静脉识别。

传统卷积层如图 2 所示,图中 P 是一个二维的张量, $P \in \mathbf{R}^{M \times C_{in}}$,其中 M 表示特征图的空间维数,

C_{in} 是输入特征图的通道数。卷积核 \mathbf{K} 是一个三维的张量, $\mathbf{K} \in \mathbf{R}^{C_{out} \times M \times C_{in}}$, 其中 C_{out} 是输出特征图的通道数, 经过卷积运算后的输出是一个 C_{out} 维的特征:

$$\mathbf{O} = \sum_i^{M \times C_{in}} \mathbf{K}_{C_{out}i} \mathbf{P}_i \quad (2)$$

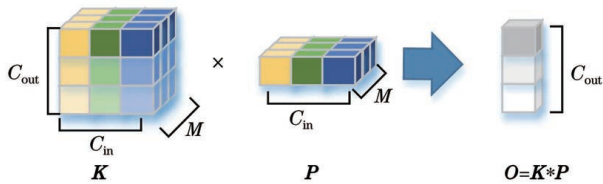


图 2 传统卷积

Fig. 2 Conventional convolution

不同于传统卷积, 深度卷积的一个卷积核负责一个通道, 一个通道也只能被一个卷积核卷积。在传统卷积核中, 每个 C_{out} 通道的卷积核与整个特征图进行点积。而在深度卷积中, 特征图 \mathbf{P} 的每个输入通道都和卷积核的 D 通道进行点积。因此, 每个输入特征图的通道(一个 M 维的特征)被卷积成一个 D 维的特征, D 就被称作是深度乘数。深度卷积如图 3 所示, \mathbf{K} 是一个三维的张量, 每个输入通道都经过卷积换成 D 维特征, 输出为

$$\mathbf{O} = \sum_i^M \mathbf{K}_{iDC_{in}} \mathbf{P}_{iC_{in}} \quad (3)$$

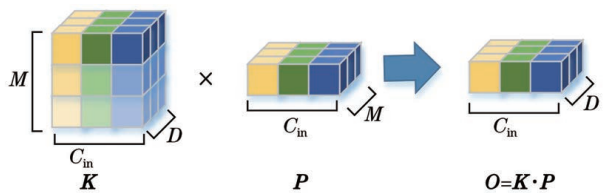


图 3 深度卷积

Fig. 3 Depthwise convolution

深度超参数化卷积是由一个深度卷积核 \mathbf{J} 和一个传统卷积 \mathbf{K} 组合而成, 先用深度卷积核和传统卷积核卷积形成新的卷积核, 再用新的卷积核卷积对特征图进行卷积得到最终特征。

$$\mathbf{O} = (\mathbf{J}^T \cdot \mathbf{K}) * \mathbf{P} \quad (4)$$

由图 4 可知, \mathbf{J} 和 \mathbf{K} 经过运算得到 $\mathbf{K}' = \mathbf{J}^T \cdot \mathbf{K}$, \mathbf{K}' 是一种增强的卷积核, 它对指静脉图像的每个通道都进行学习, 使提取的指静脉纹理特征质量更好, 更具有区分性。又因为 \mathbf{K}' 与传统卷积核的尺寸完全相同, 所以计算量与使用传统卷积核是一样的。只有当 $D \geq M$ 时, \mathbf{K}' 才能与传统卷积中的 \mathbf{K} 进行相同的线性变换。深度超参数化卷积使网络呈现出一种过参数化, 不仅增加了可学习的参数, 加速了网

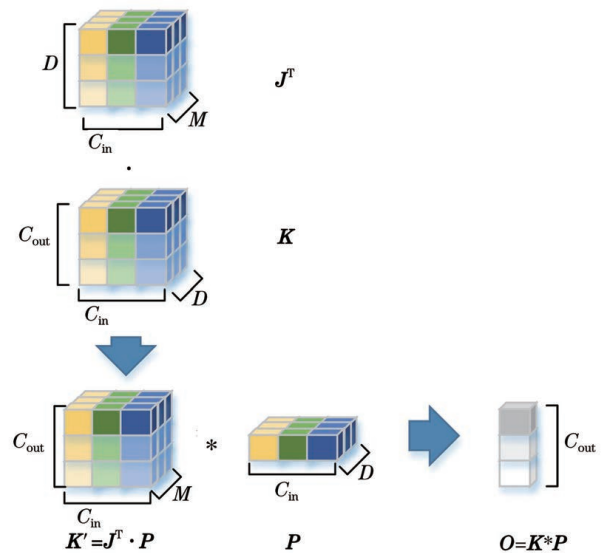


图 4 深度超参数化卷积

Fig. 4 Depthwise over-parameterized convolution

络的训练, 而且使网络在保持原有计算量的基础上, 提高了提取纹理特征的质量。

3.2 注意力机制

手指静脉特征具有很大的相似性, 仅使用纹理特征进行识别会导致分类错误, 所以需要更细微的特征来区分不同类别, 注意力机制可以加强网络对细小特征提取的能力。本文使用 SAM 和 SE-Block 双注意力机制, 将其加入到 ResNet 网络中, 可以进一步提高网络在通道和空间域中提取指静脉更深细节特征的能力。与原始的卷积模块的注意力机制相比, SAM 和 SE-Block 双注意力机制使用一个全局池化层代替最大池化层和平均池化层对特征进行压缩, 可以避免模块的参数损失过大, 从而完成精准的预测。图 5 是改进后的注意力机制结构图, 首先利

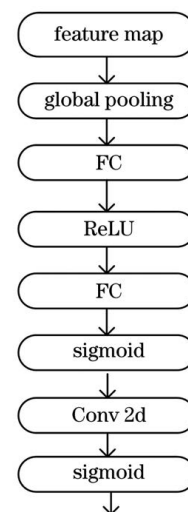


图 5 双注意力机制结构

Fig. 5 Structure of dual attention mechanism

用一个全局平均池化将特征图全局空间信息压缩到一个通道中,

$$Z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), \quad (5)$$

式中: H 和 W 表示输入特征图的维度, 经过压缩后得到 Z 的第 c 个元素值。经过压缩模块得到了聚集的信息, 然后是激励操作, 由两层全连接层, 线性整流函数 (ReLU) 层和 Sigmoid 函数组成, 利用参数 w 为特征通道产生权重, 表示为

$$s = \sigma [w_2 \delta(w_1 z)], \quad (6)$$

式中: δ 表示 ReLU 函数; σ 表示 Sigmoid 函数; w_1 和 w_2 分别为第一层和第二层的权重参数。SE-Block 的主要作用是利用提取的特征对网络进行重新校准, 使有效的特征权重增大, 从而提升网络性能, 但是却忽略了空间上的细节特征, 因此在 SE-Block 后面接上 SAM 模块, 空间注意力机制侧重于位置信息, 是对通道注意力的补充, 可以增强网络对空间细节的提取。通过一个 3×3 的标准卷积层 $f^{3 \times 3}$ 得到空间注意力图:

$$M_s(s) = \sigma [f^{3 \times 3}(s)]. \quad (7)$$

将改进后的注意力机制放在残差模块中得到新的残差模块, 并用 SELU 代替 ReLU 将细小的变化进行放大, 结合注意力机制的残差模块结构如图 6 所示。

3.3 标签平滑的交叉熵损失函数

手指静脉的特征比较简单, 不同的手指静脉在提取到的特征上, 可能会有比较高的相似度, 因此即使提取了高质量的指静脉特征, 也可能出现一种过度自信的现象, 最后导致分类错误。为了解决过度自信的问题, 本文对交叉熵函数进行标签平滑。

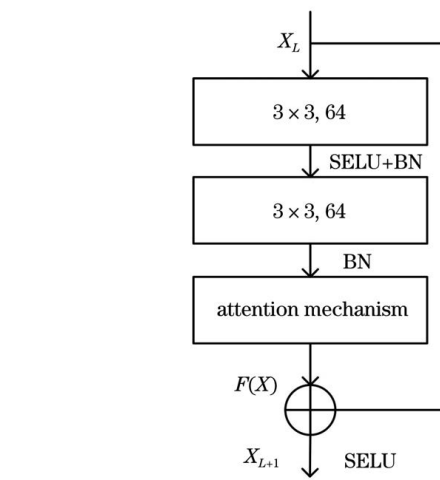


图 6 改进型残差模块

Fig. 6 Improved residual block

标签平滑交叉熵 (LSCE) 损失函数公式为

$$L_{ce} = - \sum y_{true} \ln y_{pred}. \quad (8)$$

对交叉熵函数使用标签平滑, 实质上就是添加一个平滑系数 ϵ , 使用标签平滑后的交叉熵函数为

$$L_{lsce} = - \sum [(1 - \epsilon) * y_{true} + \epsilon / N] \ln y_{pred} = (1 - \epsilon) \sum y_{true} \ln y_{pred} - \epsilon / N \sum \ln y_{pred}, \quad (9)$$

式中: y_{true} 表示真实的结果; y_{pred} 表示预测的结果; N 是指静脉的类别数量。从 (9) 式可以看出, 标签平滑使最大预测与其他指静脉类别平均值之间的差距变得更加平滑, 用在网络中可以防止过拟合, 增强网络的预测和泛化能力。

本文针对手指静脉类间差异小和细微特征多的特点, 进行了改进。改进后的残差网络结构如图 7 所示, 与传统的网络模型相比, 具有明显的优势。首先将网络层数简化至 8 层, 减少了模型参数, 减少了

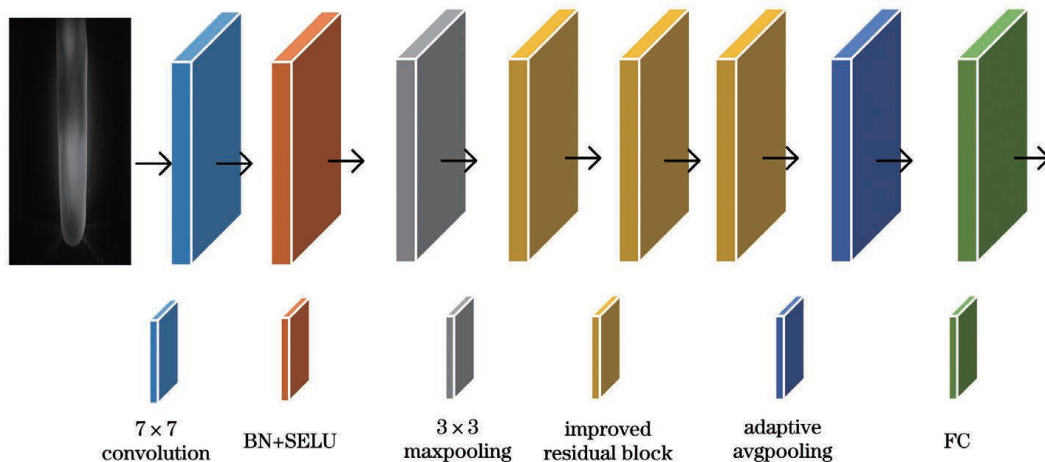


图 7 改进后的残差网络

Fig. 7 Improved ResNet

运行时间,使用深度超参数化卷积(DO-Conv)可以提高网络提取手指静脉纹理特征的质量。然后用注意力机制增强网络对细微特征的提取能力,通过提取到的细微特征来区分不同类,提升了识别精度。最后使用标签平滑的交叉熵函数,解决了过度自信的问题,有效地区分出相似度较大的手指静脉图像,进一步提升了识别准确率。

4 实验和分析

4.1 手指静脉图像数据库

为了对本文改进的网络进行验证,本文在两个公开的数据库上进行实验。1) 马来西亚理工大学手指静脉数据集(FV-USM)^[20],从 123 名志愿者手上提取,每个志愿者提供四根手指:左食指,左中指,右食指和右中指,共 492 个手指类别,分两个阶段进行提取,每一个阶段都捕获六次,图片大小为 640 pixel×480 pixel,共 5904 个图片。2) 山东大学手指静脉数据库(SDUMLA)^[21],从 106 名志愿者中采集,每个志愿者提供两只手的食指、中指和无名

指,每个手指捕获六次,共有 636 个手指类别,3816 张手指图片,图片大小为 320 pixel×240 pixel。两个公开数据库分别从多个志愿者的不同手指上进行采集,并且考虑到了不同光照、多角度等因素,数据库满足本次实验要求,按照 1:9 的比例划分为测试数据库和训练数据库。

4.2 实验环境及参数设置

本文使用 Pycharm 作为实验平台,所有的方法都是在 Pytorch 框架下进行完成的,中央处理器(CPU)为 AMD Ryzen 5 2600CPU,内存 8 GB,图形处理器(GPU)为 NVIDIA GeForce GTX1660 (6 GB)。在训练中训练轮数 epoch 设置为 500,学习率设置为 0.0001,batchsize 设置为 128,利用反向传播算法和 Adam 优化器对模型进行训练。

4.3 实验结果

使用 ResNet-18 作为基础网络,在此基础上进行改进,将改进的网络在数据集上进行测试得到表 1 的结果。

表 1 不同模型的结构及性能指标对比

Table 1 Comparison of structure and performance indicators of different models

Method	Accuracy / %		Time / ms		Parameter
	FV-USM	SDUMLA	FV-USM	SDUMLA	
VGG-16	95.8333	95.7721	49	49	44.0×10^6
DenseNet	96.9512	98.1618	34	34	8.5×10^6
AlexNet	92.2764	96.3235	16	16	26.7×10^6
ResNet-18	96.0366	97.7941	16	16	19.7×10^6
Improved ResNet+DO-Conv	98.4756	98.7132	15	15	14.8×10^6
Improved ResNet+DO-Conv+LSCE	99.0854	98.8971	15	15	14.8×10^6
Improved ResNet+DO-Conv+LSCE+(SE+SAM)	99.4919	99.4485	15	15	14.8×10^6

从表 1 中可以看出,在几种经典的网络模型中,DenseNet 的识别精度最高,其次是 ResNet。AlexNet 和 ResNet 耗时最短。DenseNet 的模型参数最少,其次是 ResNet。综合三个方面考虑,ResNet 的性能最好。将每一个模块加入到 ResNet 中,精度均有提升,证明了改进后的模块和模型的有效性。两个数据库的准确度曲线如图 8 所示,横轴为训练迭代轮数,纵轴为精确度。

5 结论

针对指静脉的识别困难,本文在 ResNet-18 网络的基础上进行改进,使用 DO-Conv 代替传统卷

积,在没有增加计算复杂性的基础上,提高了提取特征的质量,从而提升了网络性能。其次,引入注意力机制增强网络对手指静脉图像通道和空间域上细节特征的提取。最后,使用标签平滑的交叉熵函数作为损失函数,解决了过度自信问题,提升了分类的准确性。实验结果显示,使用改进后的网络在 FV-USM 和 SDUMLA 两个数据库上识别准确度分别为 99.4919% 和 99.4485%,与基础网络相比有 3.4553% 和 1.6544% 的提升,证明了改进模型的有效性。在未来的工作中,我们将对攻击检测和软硬件优化进行深入的研究。

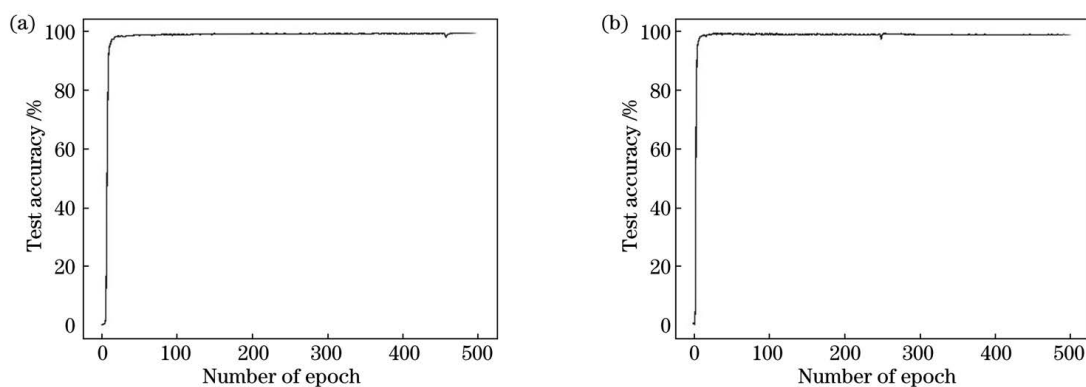


图 8 测试精度图。(a) FV-USM; (b) SDUMLA

Fig. 8 Test accuracy. (a) FV-USM; (b) SDUMLA

参 考 文 献

- [1] Jain A K, Ross A A, Nandakumar K. Introduction to biometrics[M]. Boston: Springer, 2011: 32-52.
- [2] Yang J F, Shi Y H, Jia G M. Finger-vein image matching based on adaptive curve transformation[J]. Pattern Recognition, 2017, 66: 34-43.
- [3] Shi Y H, Yang J F. Image restoration and enhancement for finger-vein recognition[C] // 2012 IEEE 11th International Conference on Signal Processing, October 21-25, 2012, Beijing, China. New York: IEEE Press, 2012: 1605-1608.
- [4] Wu J D, Liu C T. Finger-vein pattern identification using SVM and neural network technique[J]. Expert Systems With Applications, 2011, 38(11): 14284-14289.
- [5] Raghavendra R, Surbiryala J, Busch C. An efficient finger vein indexing scheme based on unsupervised clustering[C] // IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015), March 23-25, 2015, Hong Kong, China. New York: IEEE Press, 2015: 1-8.
- [6] Zhao X H, Yin L F, Zhu Y N, et al. Improved image classification algorithm based on principal component analysis network[J]. Laser & Optoelectronics Progress, 2019, 56(2): 021004.
赵小虎, 尹良飞, 朱亚楠, 等. 基于主成分分析网络的改进图像分类算法[J]. 激光与光电子学进展, 2019, 56(2): 021004.
- [7] Li X C, Ma H Y, Lin S. Palmprint recognition based on subspace and texture feature fusion[J]. Laser & Optoelectronics Progress, 2019, 56(7): 071007.
李新春, 马红艳, 林森. 基于子空间与纹理特征融合的掌纹识别[J]. 激光与光电子学进展, 2019, 56(7): 071007.
- [8] Wu J D, Ye S H. Driver identification using finger-vein patterns with Radon transform and neural network[J]. Expert Systems With Applications, 2009, 36(3): 5793-5799.
- [9] Chen L K, Wang J, Yang S Y, et al. A finger vein image-based personal identification system with self-adaptive illuminance control[J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(2): 294-304.
- [10] Das R, Piciucco E, Maiorana E, et al. Convolutional neural network for finger-vein-based biometric identification[J]. IEEE Transactions on Information Forensics and Security, 2019, 14(2): 360-373.
- [11] Xie C H, Kumar A. Finger vein identification using convolutional neural network and supervised discrete hashing[J]. Pattern Recognition Letters, 2019, 119: 148-156.
- [12] Wang J, Pan Z Y, Wang G Q, et al. Spatial pyramid pooling of selective convolutional features for vein recognition [J]. IEEE Access, 2018, 6: 28563-28572.
- [13] Hong H G, Lee M B, Park K R. Convolutional neural network-based finger-vein recognition using NIR image sensors[J]. Sensors, 2017, 17(6): 1-21.
- [14] Qin H F, El-Yacoubi M A. Deep representation-based feature extraction and recovering for finger-vein verification[J]. IEEE Transactions on Information Forensics and Security, 2017, 12(8): 1816-1829.
- [15] Yang W M, Hui C Q, Chen Z Q, et al. FV-GAN: finger vein representation using generative adversarial networks [J]. IEEE Transactions on Information Forensics and Security, 2019, 14(9): 2512-2524.
- [16] Song J M, Kim W, Park K R. Finger-vein recognition based on deep DenseNet using composite image[J]. IEEE Access, 2019, 7: 66845-66863.
- [17] Radzi S A, Khalil-Hani M, Bakhteri R. Finger-vein biometric identification using convolutional neural network [J]. Turkish Journal of Electrical Engineering & Computer Sciences, 2016, 24(3):

- 1863-1878.
- [18] Yuan L S, Lou M Y, Liu Y Q, et al. Palm vein classification based on deep neural network and random forest[J]. *Laser & Optoelectronics Progress*, 2019, 56(10): 101010.
袁丽莎, 娄梦莹, 刘娅琴, 等. 结合深度神经网络和随机森林的手掌静脉分类[J]. *激光与光电子学进展*, 2019, 56(10): 101010.
- [19] Tao Z Y, Hu Y L, Lin S. Finger vein recognition based on improved AlexNet[J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 081005.
陶志勇, 胡亚磊, 林森. 基于改进 AlexNet 的手指静脉识别[J]. *激光与光电子学进展*, 2020, 57(8): 081005.
- [20] Asaari M S M, Suandi S A, Rosdi B A. Fusion of band limited phase only correlation and width centroid contour distance for finger based biometrics [J]. *Expert Systems With Applications*, 2014, 41(7): 3367-3382.
- [21] Yin Y L, Liu L L, Sun X W. SDUMLA-HMT: a multimodal biometric database[M]//Sun Z N, Lai J H, Chen X L, et al. *Biometric recognition. Lecture notes in computer science*. Heidelberg: Springer, 2011, 7098: 260-268.