

DOI:10.13364/j.issn.1672-6510.20210281

基于特征融合的双模态生物识别方法

周卫斌, 王 阳, 吉书林

(天津科技大学电子信息与自动化学院, 天津 300222)

摘要: 针对单一生物特征识别存在准确率低、安全性差的问题,提出了一种基于特征融合的指静脉和人脸双模态生物识别方法.该方法通过双通道卷积神经网络提取生物特征,并采用自注意力机制更新特征权重.首先不同的生物特征通过特征融合模块形成新的混合特征;然后借助残差结构将新的混合特征与指静脉特征、人脸特征再次融合,最大化地保留生物特征信息.在公开数据集上对该方法进行训练和验证,并与 AlexNet、VGG-19 和 MobileNetV2 这三种单模态生物识别方法进行对比,该方法识别准确率均超过 98.80%,表明该方法是一种高效的双模态生物识别方法.

关键词: 生物识别; 卷积神经网络; 双模态; 特征融合

中图分类号: TP389.1 **文献标志码:** A **文章编号:** 1672-6510(2022)04-0044-05

Bimodal Biological Recognition Method Based on Feature Fusion

ZHOU Weibin, WANG Yang, JI Shulin

(College of Electronic Information and Automation, Tianjin University of Science & Technology,
Tianjin 300222, China)

Abstract: In order to solve the problems of low accuracy and poor security in single biometric feature recognition, a biometrics method of finger vein and face based on feature fusion is proposed in this article. This method extracts biometrics through two-channel convolution neural network and updates feature weights by self-attention mechanism. In this method, first, new mixed features were formed by feature fusion module, and then the new mixed features were fused with finger vein and face features by residual structure to maximize the biometric information. The method was trained and validated on the public datasets and compared with the three single-mode biometrics AlexNet, VGG-19 and MobileNetV2. The accuracy of the method is over 98.80%, which shows that the method is an efficient biometrics method.

Key words: biometrics; convolutional neural network; bimodal; feature fusion

近年来,信息技术飞速发展、智能设备不断普及,指纹识别^[1]、人脸识别^[2]、虹膜识别^[3]、语音识别^[4]等身份识别技术早已应用于日常生活中的各个方面.然而,单模态生物特征识别技术在实际应用中不仅受外部环境的影响,还受自身局限性的影响,极大地限制了其应用场景,降低了身份识别的准确率.例如,指纹破损或者沾水会降低指纹识别的准确率;佩戴口罩会降低人脸识别的准确率;佩戴眼镜会降低虹膜识别的准确率.因此,双模态融合识别对于弥补单一生物特征识别的缺陷、提高身份识别的准确率、增

强信息的安全性具有重要意义.相比于三模态、四模态等多模态的生物识别方法^[5],指静脉和人脸的双模态识别不仅减少了计算量、降低了算法复杂度,还使体内生物特征(指静脉)与体外生物特征(人脸)的优势得以互补,打破了单一生物特征识别的应用局限,提高了身份信息的安全性和身份识别的准确率.

双模态生物特征识别技术是融合两种单一生物特征作为新的特征进行识别的方法^[6].指静脉和人脸属于两种不同的生物特征,根据融合发生的位置不同,可分为数据层融合、特征层融合、匹配层融合和

收稿日期: 2021-12-01; 修回日期: 2022-04-10

作者简介: 周卫斌(1981—),男,湖北孝感人,副教授, zhouweibin@tust.edu.cn

决策层融合,如图1所示。

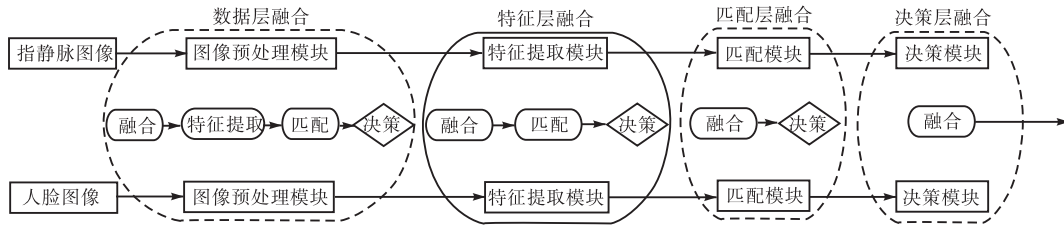


图1 双模态生物特征融合框图

Fig. 1 Block diagram of bimodal biological feature fusion

数据层融合属于最低层次的融合^[7];匹配层融合又称分数层融合,不同样本经匹配层得到与模板库的匹配分数,经过归一化处理后进行匹配分数融合^[8];决策层融合类似投票表决机制,对不同样本的决策结果(接受/拒绝)进行融合^[9];特征层融合发生在特征提取模块,不同样本经过特征提取获得各自的特征向量,然后将特征向量通过特征串联、级联等方式拼接成一个新的特征向量,将新的特征向量作为生物识别系统的输入进行匹配和决策^[10].特征层具有丰富的有效特征信息,是众多学者研究的热点.例如, Yang等^[11]利用统一的 Gabor 滤波器框架提取指纹和指静脉特征,提出一种新的监督局部保持典型相关分析方法(SLPCCAM),用于生成特征层融合的指纹指静脉特征向量(FPVFs);只使用一个数据库对该方法的性能进行评估,在实际应用中具有很大的局限性. Veluchamy等^[12]提出将指关节和手指静脉进行特征层融合,并采用支持向量机(SVM)分类器进行多模态生物识别,但是该方法在特征融合后产出的新特征向量的维数过高. Pan等^[13]提出一种基于核 Fisher 判别分析(KFDA)方法,并结合平均规则、加权求和规则和乘法规则实现人脸和人耳的特征层融合,但是该方法同样也出现了“维度灾难”现象。

近年来,卷积神经网络(CNN)的发展在生物识别系统中产生了非常大的影响,并取得了优异的成果^[14].传统算法提取图像特征需要大量手工标注,而卷积神经网络对图像逐像素卷积的方式,能够快速、高效地提取到图像多尺度特征。

1 本文方法的实现

在众多的生物特征中,人脸图像是最自然、最明显的个人识别特征;指静脉是体内生物特征,不易被损坏、伪造和复制,并且每个人、每个手指之间指静脉特征都具有差异,具有极高的安全特性^[15].鉴于卷积神经网络在图像识别和图像特征提取任务中的出

色表现,本文融合指静脉和人脸特征,提出一种高效的双模态生物识别方法.首先获取指静脉图像和人脸图像的感兴趣区域(ROI),并针对指静脉图像数据规模较小的问题,采用数据增强方法用于扩充数据.然后将指静脉和人脸 ROI 图像输入到双通道卷积神经网络中提取图像特征.在全连接层之前,根据各特征的置信度分配权重,并进行特征融合,形成新的特征用于身份识别.该方法的实现过程如图2所示。

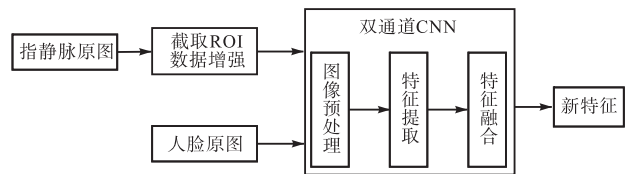


图2 双通道卷积神经网络结构图

Fig. 2 Structure diagram of two-channel convolutional neural network

1.1 特征提取网络

在图像分类识别领域,特征提取网络的好坏决定着识别结果的准确率. AlexNet 网络是一种简单的卷积神经网络,共有 8 层,其中前 5 层用于特征提取^[16].相对较浅的网络层数决定了 AlexNet 网络的准确率较低. VGG 网络通过堆叠多个尺寸为 3×3 的卷积核来代替大尺度卷积核,并且与大尺度卷积核拥有相同的感受野. VGG-19 网络有 19 层,前 16 层用于特征提取^[17].相比于前两种传统的卷积神经网络, MobileNet 网络内存需求小、运算量小,可以部署到移动设备以及嵌入式设备上. MobileNetV2 网络引入深度可分离卷积,由逐通道卷积和逐点卷积组成,减少了模型参数和运算成本^[18].此外, MobileNetV2 网络使用反向残差结构,在增加网络层数的同时,也使特征表达能力得到增强。

1.2 特征融合方法

本文提出的双模态特征融合方法在特征层采用双通道卷积神经网络模型实现特征融合,其实现过程如图3所示。

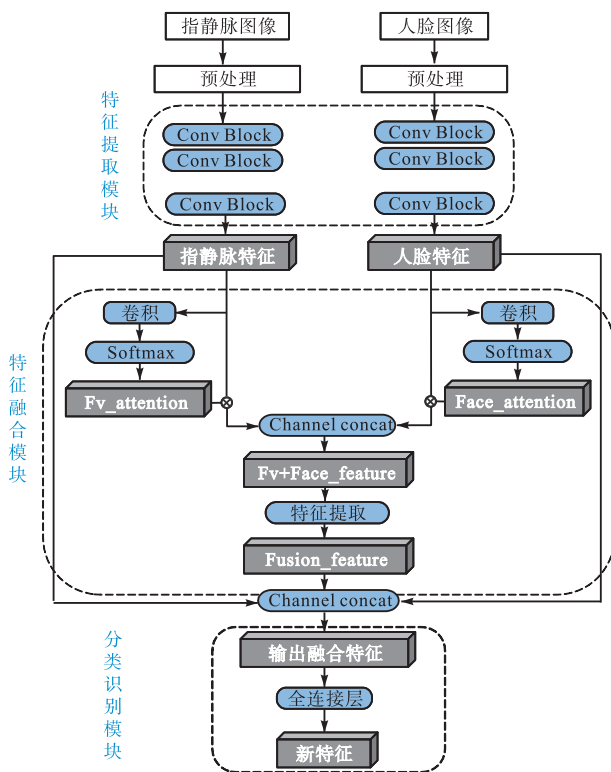


图3 双模态特征融合框架
Fig. 3 Framework of bimodal feature fusion

整体框架分 3 个模块：特征提取模块、特征融合模块和分类识别模块。在特征提取模块之前，对指静脉图像和人脸图像进行预处理，预处理过程包括对指静脉图像感兴趣区域进行截取、数据增强和数据扩容。每个生物特征都被输入到卷积神经网络模型中，经过多层卷积层和池化层，提取图像的特征信息，得到指静脉特征 (Fv_feature) 和人脸特征 (Face_feature)。特征融合模块是该框架的核心模块。将经过特征提取模块后的指静脉特征和人脸特征进行卷积操作降维，再经过 Softmax 层，分别得到自注意力权重，并与特征提取所得到的特征相乘；将两特征融合到一起，得到指静脉与人脸的融合特征 (Fv + Face_feature)；经过特征提取后得到更深层的融合特征 (Fusion_feature)。为防止在特征融合时出现部分特征信息丢失的情况，将特征提取后的指静脉特征、人脸特征、融合特征三者再次融合到一起。两次融合保证了特征信息最大化。最后经过全连接层进行分类识别。

本文分别选择 AlexNet 网络、VGG-19 网络和 MobileNetV2 网络的特征层，用于提取指静脉和人脸图像特征，作为特征融合模块的输入。为避免在融合模块出现过拟合现象和减少融合后的运算量，本文采

用自适应均值池化的方法和 Dropout 技术。

2 数据处理及实验

实验软件环境：Python 3.8、CUDA 11.0、cuDNN 8.0、Pytorch 1.7.1、Ubuntu 18.04；硬件平台：NVIDIA GeForce RTX 2080 Ti GPU。

2.1 公开数据集

为了验证本文所提出的双模态特征融合方法的有效性，同时表明它与单模态生物特征识别的优势，实验过程选用指静脉公开数据集 SDUMLA-FV^[19]和 Finger Vein USM (FV-USM)^[20]、人脸公开数据集 CASIA-WebFace^[21]，对本文提出的双模态特征融合方法的有效性进行测试，并与单模态生物特征识别方法进行对比。

SDUMLA-FV 数据集是由山东大学机器学习与数据挖掘实验室创建，包含 106 人的左、右手的食指、中指和无名指中每根手指的 6 张指静脉图像，共 636 类手指图像，总计 3 816 张。FV-USM 数据集是马来西亚理工大学创建，包含 123 人的左、右手的食指和中指中每根手指的 6 张图像，共 492 类手指图像，总计 2 952 张。FV-USM 数据集提供了已经截取好的 ROI 图像，为后续指静脉图像的预处理提供了方便。CASIA-WebFace 数据集是应用于人脸识别领域最广泛的公开数据集之一，该数据集收集了网络上的人脸图像，共 10 575 类，494 414 张图像。在实验中，根据指静脉图像的类别数，随机从人脸数据集中选择相同的类别数。

2.2 图像预处理

由于 SDUMLA-FV 数据集没有提供指静脉的 ROI 图像，所以需要对该数据集中图像进行 ROI 的截取，以去除过多的背景无用信息。采用 Prewitt 边缘检测算子对指静脉原图垂直方向上的上下边缘进行边缘检测，对于存在伪边缘的现象，通过设置连通域阈值去除伪边缘。使用最小二乘法线性回归拟合手指的中轴线，根据拟合直线与水平线的夹角将图像旋转矫正；拟合手指上下边缘的内切线；根据图像水平方向上的亮度变化趋势，选择指关节处（即亮度峰值处）；最后截取手指静脉 ROI 图像。

为了获得清晰的指静脉纹路，还需要对截取的 ROI 图像进行对比度受限自适应直方图均衡化 (contrast limited adaptive histogram equalization, CLAHE)，并在 CLAHE 图像增强之后添加 Gabor 滤

波器,用于去除图像增强后的噪声. SDUMLA-FV 数据集图像的预处理过程如图 4 所示,ROI 原图经过 CLAHE 图像增强、Gabor 滤波去噪后,相比于原图可以得到清晰的静脉纹路. SDUMLA-FV 和 FV-USM 数据集只提供了每根手指 6 张静脉图像,为防止卷积神经网络模型训练过程中出现过拟合现象,对每一类

指静脉进行数据扩增,其中包括对图像随机平移、旋转、裁剪、亮度调节、对比度调节,将原本每类 6 张图像扩充至每类 36 张图像. FV-USM 数据集提供了 ROI 图像,因此只需要对该数据集进行图像增强和扩充.

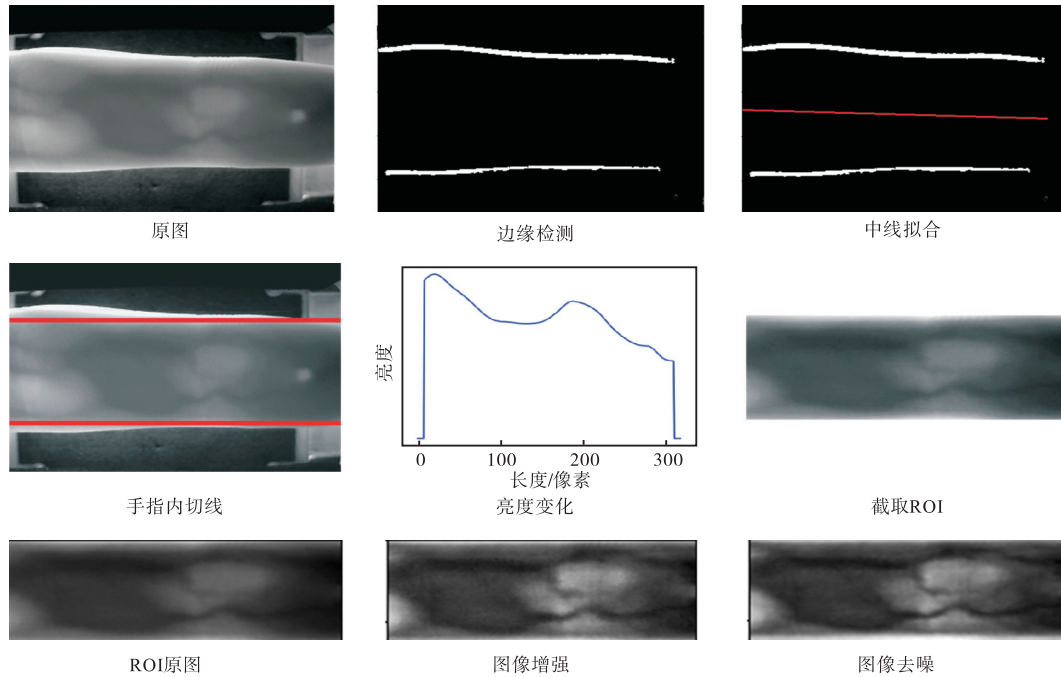


图 4 SDUMLA-FV数据集图像的预处理

Fig. 4 Image preprocessing of SDUMLA-FV datasets

2.3 实验结果与分析

实验中将每一类别中的数据(36 张)按照 7 : 3 的比例划分训练集和测试集. 为了提高网络的泛化能力,在融合模块中,每张指静脉图像与每张人脸图像一一对应. 单模态实验和双模态特征融合实验数据分布结构见表 1.

实验中,所有的模型输入尺寸均为 224 像素 × 224 像素,激活函数为 ReLU,优化器选用 Adam,学习率设置为 0.000 1. 分别对不同数据集、不同特征提

取网络进行对比实验,识别准确率的实验结果见表 2.

表 1 实验数据分布

Tab. 1 Experimental data distribution

数据集	图像数量/张	
	训练集	测试集
SDUMLA-FV	16 536	6 360
FV-USM	12 792	4 920
SDUMLA-FV + CASIA-WebFace	429 936	63 600
FV-USM + CASIA-WebFace	332 592	49 200

表 2 不同数据集、不同特征提取网络的识别准确率

Tab. 2 Recognition accuracy of different datasets and different feature extraction networks

模型	参数量	识别准确率/%				
		SDUMLA-FV	FV-USM	CASIA-WebFace	SDUMLA-FV + CASIA-WebFace	FV-USM + CASIA-WebFace
AlexNet	16 630 440	79.92	52.95	53.95	—	—
AlexNet 融合	9 858 994	—	—	—	99.80	99.95
VGG-19	143 667 240	94.36	83.80	55.75	—	—
VGG-19 融合	45 229 938	—	—	—	99.95	99.94
MobileNetV2	3 504 872	54.81	39.59	21.32	—	—
MobileNetV2 融合	30 426 738	—	—	—	98.84	99.40

由表 2 可知:在单模态识别实验中,只有 VGG-19 模型对 SDUMLA-FV 数据集的识别准确率较高,为 94.36%,其他单模态识别实验的识别准确率均处于较低水平.在双模态特征融合实验中,VGG-19 融合模型对 SDUMLA-FV + CASIA-WebFace 数据集的识别准确率为 99.95%,对 FV-USM + CASIA-WebFace 数据集的识别准确率为 99.94%.其他双模态特征融合实验的识别准确率均高于 98.80%,相比单模态下的识别准确率均得到了很大提升.对比 Alay 等^[22]使用 VGG-16 基础网络进行的虹膜、人脸和指静脉三模态特征融合方法,本文所提出的方法在只使用两种生物特征融合的情况下,识别准确率相比

于三种生物特征融合均有所提升.为了进一步验证所提出的双模态特征融合方法的有效性,将融合模态与 MobileNetV2 轻量级网络结合并考察其识别准确率.由表 2 可知:轻量级网络模型(MobileNetV2)相比于 AlexNet、VGG-19 网络模型,参数量大幅减少,并且单模态识别准确率也大幅降低,但双模态特征融合识别实验可以达到与 AlexNet、VGG-19 网络相当的识别准确率.

每个模型识别一张图像(融合模型为识别指静脉和人脸两张图像)的时间性能对比见表 3.从表 3 中可以看出,融合模型的耗时与单模态识别的耗时差异不明显.

表 3 时间性能
Tab. 3 Time performance

模型	识别时间/s				
	SDUMLA-FV	FV-USM	CASIA-WebFace	SDUMLA-FV + CASIA-WebFace	FV-USM + CASIA-WebFace
AlexNet	1.44	1.40	1.42	—	—
AlexNet 融合	—	—	—	1.42	1.41
VGG-19	1.40	1.45	1.48	—	—
VGG-19 融合	—	—	—	1.39	1.40
MobileNetV2	1.36	1.36	1.38	—	—
MobileNetV2 融合	—	—	—	1.36	1.36

3 结 语

本研究提出了一种基于卷积神经网络模型的指静脉和人脸特征融合识别方法,解决了实际应用场景中单一生物特征识别准确率低、安全性差的问题.对比实验结果表明本文提出的方法可以有效提高生物特征识别的准确率.此外,本文所提出的 AlexNet 融合模型和 VGG-19 融合模型,参数量均少于 AlexNet、VGG-19 网络,但是轻量级网络 MobileNetV2 上参数量并没有减少.对于这一问题,下一步还需要继续对网络结构进行优化.同时,本研究还需要进一步在实际复杂场景的数据集上进行对比实验.

参考文献:

[1] 于晓明. 指纹识别应用研究[J]. 山东工业技术, 2019(9): 168.
 [2] 王浩. 当今人脸识别技术的现状与未来研究方向[J]. 计算机与数字工程, 2021, 49(9): 1905-1911.
 [3] 史春蕾, 杨丹丹, 韩飞. 虹膜识别综述[J]. 科技展望, 2016, 26(5): 307.
 [4] 郭宗昱, 刘博, 吴可欣, 等. 端到端语音识别研究综述[J]. 科技风, 2020(34): 7-9.

[5] 聂昊, 鲁玺龙, 郭文志, 等. 多模态生物特征识别技术的研究进展[J]. 生命科学仪器, 2020, 18(5): 20-28.
 [6] 苗迪. 基于特征层与分数层的多模态生物特征融合[D]. 合肥: 中国科学技术大学, 2017.
 [7] 李芹. 基于特征提取的图像拼接方法[J]. 现代计算机, 2020(9): 69-72.
 [8] 阮彦浪. 多模态生物特征识别[D]. 鞍山: 辽宁科技大学, 2018.
 [9] GUPTA K, WALIA G S, SHARMA K. Quality based adaptive score fusion approach for multimodal biometric system[J]. Applied intelligence, 2020, 50: 1086-1099.
 [10] 刘奕. 多模生物特征融合关键技术研究[D]. 济南: 山东大学, 2017.
 [11] YANG J F, ZHANG X. Feature-level fusion of fingerprint and finger-vein for personal identification[J]. Pattern recognition letters, 2012, 33(5): 623-628.
 [12] VELUCHAMY S, KARLMARX L R. System for multimodal biometric recognition based on finger knuckle and finger vein using feature-level fusion and k-support vector machine classifier[J]. IET Biometrics, 2017, 6(3): 232-242.
 [13] PAN X Q, XU X N, LU Y, et al. Feature fusion in multimodal recognition based on ear and profile face[C]// (下转第 54 页)

- [4] HANNAN M A, LIPU M S H, HUSSAIN A, et al. A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: challenges and recommendations[J]. *Renewable and sustainable energy reviews*, 2017, 78: 834–854.
- [5] JIANG Z Y, QU Z G, ZHOU L, et al. A microscopic investigation of ion and electron transport in lithium-ion battery porous electrodes using the lattice Boltzmann method[J]. *Applied energy*, 2017, 194: 530–539.
- [6] 葛云龙, 陈自强. 基于 STF&LM 算法的串联锂离子电池组不一致性辨识与状态估计[J]. *中国电机工程学报*, 2018, 38(14): 4271–4280.
- [7] 范文杰, 徐广昊, 于泊宁, 等. 基于电化学阻抗谱的锂离子电池内部温度在线估计方法研究[J]. *中国电机工程学报*, 2021, 41(9): 3283–3293.
- [8] LI W H, FAN Y, RINGBECK F, et al. Electrochemical model-based state estimation for lithium-ion batteries with adaptive unscented Kalman filter[J]. *Journal of power sources*, 2020, 476: 228534.
- [9] TRAN M K, MEVAWALA A, PANCHAL S, et al. Effect of integrating the hysteresis component to the equivalent circuit model of lithium-ion battery for dynamic and non-dynamic applications[J]. *Journal of energy storage*, 2020, 32: 101785.
- [10] ZHANG C P, JIANG Y, JIANG J C, et al. Study on battery pack consistency evolutions and equilibrium diagnosis for serial-connected lithium-ion batteries[J]. *Applied energy*, 2017, 207: 510–519.
- [11] KIM M, KIM K, KIM J, et al. State of charge estimation for lithium ion battery based on reinforcement learning[J]. *IFAC Papers on line*, 2018, 51(28): 404–408.
- [12] SBARUFATTI C, CORBETTA M, GIGLIO M, et al. Adaptive prognosis of lithium-ion batteries based on the combination of particle filters and radial basis function neural networks[J]. *Journal of power sources*, 2017, 344: 128–140.
- [13] SHEN J N, SHEN J J, HE Y J, et al. Accurate state of charge estimation with model mismatch for Li-ion batteries: a joint moving horizon estimation approach[J]. *IEEE Transactions on power electronics*, 2018, 34(5): 4329–4342.
- [14] 廖晓闽, 严少虎, 石嘉, 等. 基于深度强化学习的蜂窝网资源分配算法[J]. *通信学报*, 2019, 40(2): 11–18.

责任编辑: 周建军

(上接第 48 页)

- Proceedings of the Seventh International Symposium on Instrumentation and Control Technology. Beijing: International Society for Optics and Photonics, 2008.
- [14] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. *计算机学报*, 2017, 40(6): 1229–1251.
- [15] 牛妞王洁, 程君霞, 程瑛璠, 等. 指静脉识别技术的应用前景研究[J]. *电脑知识与技术*, 2021, 17(12): 190–191.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 26th Annual Conference on Neural Information Processing Systems 2012. Nevada: NIPS, 2012.
- [17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of the 3rd International Conference on Learning Representations. San Diego: Computational and Biological Learning Society, 2015.
- [18] SANDLER M, HOWARD A, ZHU M L, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE, 2018: 4510–4520.
- [19] YIN Y L, LIU L L, SUN X W. SDUMLA-HMT: a multimodal biometric database[C]//Proceedings of the Chinese Conference on Biometric Recognition. Heidelberg: Springer, 2011: 260–268.
- [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] ZHOU J C, JIA X, SHEN L L, et al. Improved softmax loss for deep learning-based face and expression recognition[J]. *Cognitive computation and systems*, 2019, 1(4): 97–102.
- [22] ALAY N, AL-BAITY H H. Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits[J]. *Sensors*, 2020, 20(19): 5523.

责任编辑: 周建军