

引文格式:

邓可望, 赵娟, 肖振中, 等. 基于多光谱和面部多区域联合的人脸活体检测算法 [J]. 集成技术, 2024, 13(1): 72-81.

Deng KW, Zhao J, Xiao ZZ, et al. Face anti-spoofing algorithm based on combination of multiple facial regions using multi-spectral images [J]. Journal of Integration Technology, 2024, 13(1): 72-81.

基于多光谱和面部多区域联合的人脸活体检测算法

邓可望^{1,2} 赵娟^{1*} 肖振中^{2*} 师少光² 朱亮³

¹(中国科学院深圳先进技术研究院 深圳 518055)

²(奥比中光科技集团股份有限公司 深圳 518062)

³(深圳奥芯微视科技有限公司 深圳 518062)

摘要 在常见的人脸活体检测应用场景中, 绝大多数相关技术聚焦于 RGB 图像或 IR 图像, 但是这些图像缺乏足够的生物特征, 容易受到层出不穷的假体人脸攻击。该文提出了一种基于面部多区域联合的 Transformer 模型, 并将多光谱成像技术引入人脸活体检测任务, 旨在获取人脸的独特生物特征, 增加与假体的可区分性, 进而提高活体检测准确率。多光谱图像拓宽了光谱范围, 可获取物体更为丰富的反射特性, 通过逐像素进行光谱归一化操作, 可降低光照强度变化带来的影响, 增强人脸反射特征区域的一致性。该文提出的算法选取多个人脸核心区域(如眼睛、鼻子、嘴巴、脸颊等)作为深度学习模型输入, 构建了基于 Transformer 的神经网络模型, 同时获取人脸局部区域特征和区域间关联特征, 整合成完备的人脸生物特征。在自建的多光谱人脸数据集上, 该文提出的方法获得了 95.72% 的活体检测准确率及 5.10% 的活体检测错分率, 优于常用的人脸活体检测模型。

关键词 活体检测; 多光谱; Transformer 模型; 深度学习

中图分类号 TP 391.41 文献标志码 A doi: 10.12146/j.issn.2095-3135.20230606001

Face Anti-Spoofing Algorithm Based on Combination of Multiple Facial Regions Using Multi-Spectral Images

DENG Kewang^{1,2} ZHAO Juan^{1*} XIAO Zhenzhong^{2*} SHI Shaoguang² ZHU Liang³

¹(Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

²(Orbbec Inc., Shenzhen 518062, China)

³(Shenzhen Oxin Technology Co., Ltd., Shenzhen 518062, China)

*Corresponding Authors: juan.zhao@siat.ac.cn; xiaozhenzhong@orbbec.com

Abstract In the research of face anti-spoofing (FAS), most related techniques are dependent to the RGB images

收稿日期: 2023-06-06 修回日期: 2023-07-06

作者简介: 邓可望, 博士, 研究方向为深度学习; 赵娟(通讯作者), 高级工程师, 研究方向为信息光学, E-mail: juan.zhao@siat.ac.cn; 肖振中(通讯作者), 高级工程师, 研究方向为机器视觉与数字图像处理, E-mail: xiaozhenzhong@orbbec.com; 师少光, 中级工程师, 研究方向为三维视觉; 朱亮, 博士, 研究方向为光学工程。

or IR images, which lack sufficient biometric features and are vulnerable to ever-advancing presentation attacks. In this paper, a Transformer model based on combination of multiple facial regions is proposed to introduce multi-spectral technology into the task of facial live detection, aiming to obtain unique biological features of the real faces and increase the distinguishability from the fake faces. In the proposed model, multi-spectral images are utilized to broaden the spectral dimension for more reflection information, which can identify various materials. Besides, a spectral normalization method is preprocessed pixel by pixel to reduce the impacts of the environmental illumination variations and enhance the consistency of facial reflection features regionally. Then multiple core facial regions, like eyes, nose, mouth and cheeks, are selected as input of the deep learning model. Furthermore, a Transformer-based model is constructed to obtain both local regional features and inter association features of different facial regions, which are integrated into complete facial biometric features to achieve facial live detection. On the author's self-built multi-spectral facial datasets, the results show that the proposed method achieved an accuracy of 95.72% for and a misclassification of 5.10% for live detection, which is superior to commonly used FAS models.

Keywords anti-spoofing; multi-spectral; Transformer model; deep learning

1 引言

在过去的几十年中, 人脸识别技术取得了里程碑式的发展, 尤其在智能安防、互联网支付等领域^[1-3]。在这些应用场景中, 摄像头模组需要具备体积小、帧率高、分辨率高的特性, 常用的模组包括 RGB 模组和 IR 模组。然而, 这两类摄像头模组仅能提供人脸的颜色和纹理特征, 缺乏相应的生物特征, 容易受到假体人脸攻击。多光谱滤波器阵列 (multi-spectral filter array, MSFA) 的推广使得多光谱摄像头在消费电子领域中的应用成为可能。与传统图像相比, 多光谱图像能够呈现物体的光谱反射特性, 而这些特性往往与物体自身的材质息息相关。因此, 可将多光谱图像获取的人脸反射特征作为人脸活体检测的重要判别依据。2001 年, Angelopoulo 等^[4]发现, 由于皮肤血管中的血红蛋白对不同波段光的吸收特性不同, 人脸反射光谱在 420~630 nm 处呈现明显的“W”形状。除此之外, Zhang 等^[5]研究发现, 人脸反射光谱在 850 nm 处具有更高

的反射率, 能够有效地将人脸与其余材质进行区分。

目前, 人脸活体检测方法主要分为交互式活体检测方法和静默式活体检测方法^[6]。由于交互式活体检测方法需要用户配合, 故学术研究主要集中在静默式活体检测方法。静默式活体检测方法主要分为基于纹理特征的活体检测方法、基于动态特征的活体检测方法、基于多模态的活体检测方法^[7]。基于纹理特征的活体检测方法关注人脸面部丰富的细节特征^[8-12], 采用局部二值模式、方向梯度直方图、频域转换等方法提取纹理特征, 但是容易受到逼真硅胶头模、高分辨率照片等的攻击。基于动态特征的活体检测方法将多帧图像作为时间序列, 利用神经网络获取人体动态生理特征 (如远程光容积描记术等^[13]), 但需要较长时间才能获得较好的活体检测结果, 且对高清视频防范效果较差。基于多模态的活体检测方法采用多个摄像头模组, 通常包括 RGB 摄像头、深度摄像头和 IR 摄像头, 通过建立神经网络模型处理多模态数据, 完成活体检测任务^[14-15]。但是,

其本质仍然是获取人脸的纹理细节和空间结构特征，并未考虑人脸生物特性，且成本较高，只有较大的网络模型才能获得比较好的收敛效果。

因此，针对上述问题，本文利用 MSFA 获取人脸多光谱图像，采用光谱归一化方法提高人脸光谱区域的一致性，形成稳定的人脸反射特性。此外，本文提出一种基于面部多区域联合的 Transformer 模型 (facial patch Transformer-based model, FPTransformer)，在利用注意力机制获取各区域关联性的基础上，同时提取人脸空间特征和光谱特征，提炼深层次人脸生物特征，提升人脸活体检测成功率。为验证方法的有效性，在自建的多光谱数据集上，将本文提出方法与人脸活体检测领域广泛应用的方法进行比较，并对实验结果进行总结分析。

2 基于面部多区域联合的活体检测算法

2.1 人脸反射特性研究及多光谱图像获取

皮肤是组成人脸面部的主要成分，它独特的生理特性决定了其对不同波段光的反射特性，可凭此作为判断人脸活体的依据。由于大部分物体的反射特性集中于可见-近红外波段^[16]，因此，本文选用手持式智能型高光谱相机 SPECIM-

IQ 对人脸皮肤与几种常见假体材质 (包括硅胶、树脂、乳胶和打印纸) 的反射率曲线进行研究。SPECIM-IQ 覆盖的波长范围为 400~1 000 nm，光谱分辨率为 7 nm，配备专业的白板，能够方便地获取物体反射率。本文将真人和不同材质假体脸颊的反射率曲线进行比较，同时将真人人脸不同区域组织的反射率进行比较，结果如图 1 所示。

由图 1(a) 可知，真人人脸皮肤反射率曲线在 500~600 nm 处呈现明显的“W”形状，且在 960 nm 处有明显的吸收峰，具有独特的可区分性。另外，由图 1(b) 可知，人脸不同区域组织由于血红蛋白及水分等含量不同，反射率也存在一定差异。在实际的人脸识别应用场景中，昂贵专业的高光谱仪很难得到广泛应用，而通过 MSFA 获取部分可区分性大的波段图像，从而实现活体检测将成为一种简单且高效的手段。考虑到市面上 MSFA 的波段排列以及滤光片的滤光曲线，本文选用的 MSFA 包含 9 个波段，排列方式以及对应的滤光曲线如图 2 所示。该滤光片组合覆盖了具有典型人脸活性特性的 500~600 nm 区域和 960 nm 区域。

MSFA 所获单帧马赛克图像的尺寸为 1 600×1 200 像素，为了便于在消费电子端进行部署，本文直接将马赛克图像重新采样为多光谱图像。

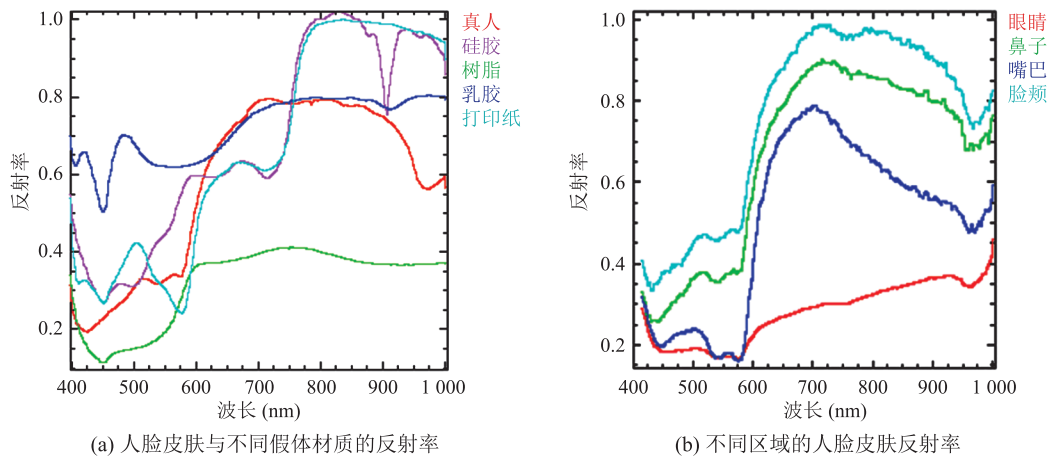


图 1 不同材质及不同区域的人脸皮肤反射率

Fig. 1 Facial skin reflectance of different materials and regions

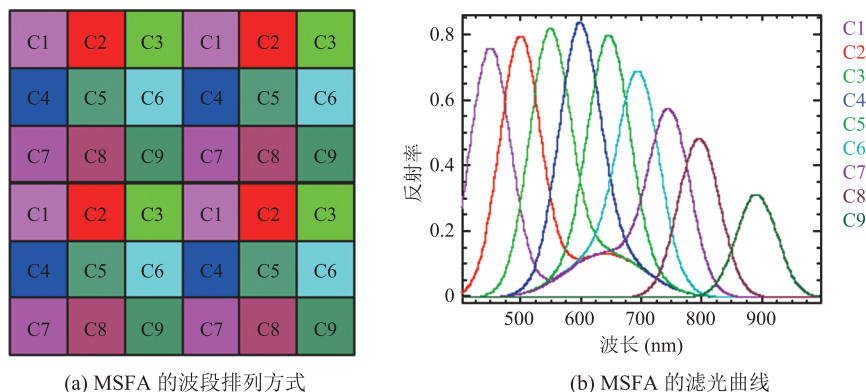


图 2 MSFA 的波段排列方式和滤光曲线图

Fig. 2 Band arrangement and filtering curves of MSFA

即令涂有相同滤光片的像素构成对应单波段图像, 经此操作后, 所获多光谱图像数据的大小为 $533 \times 400 \times 9$ 像素。

2.2 光谱归一化

在实际的智能安防场景中, 利用摄像头模组仅能获取物体表面的反射辐照度信息, 若缺乏相应的先验知识, 则无法计算反射率信息。因此, 直接利用 MSFA 捕获的多光谱图像容易受到外界光照条件的影响, 难以获取物体稳定的反射特性, 为后续的人脸活体检测任务带来挑战。本文引入多光谱成像技术的初衷是挖掘物体反射特性, 将其作为材质区分的重要判断依据。因此, 本文着重关注物体反射曲线, 假定组成成分相同的物体具有相近的反射曲线, 且一般不受外界光照条件影响。本文经过对 MSFA 成像原理进行研究, 提出了一种光谱归一化方法, 以获取物体的相对反射曲线, 提高多光谱图像区域一致性。令 MSFA 所获多光谱图像记作 $A = \{D_1, D_2, \dots, D_9\}$, 对于任一像素 $P = \{d_1, d_2, \dots, d_9\}$, 其中, d_i 为第 i 个波段的光强值, 光谱归一化计算公式如公式 (1) ~ (2) 所示:

$$E = \sqrt{\sum_{i=1}^9 d_i^2} \quad (1)$$

$$dn_i = \frac{d_i}{E} \quad (2)$$

其中, E 为像素总能量; dn_i 为第 i 通道的光强值所占总能量的比值。

由公式 (1) ~ (2) 得归一化后像素表达式为 $P_{\text{norm}} = \{dn_1, dn_2, \dots, dn_9\}$ 。当外界光照条件发生变化时, MSFA 所获各波段的灰度图像会随之变化, 而经过归一化后的多光谱图像能够降低这种影响, 并且能够使材质相同的物体具有相近的反射曲线, 极大地提升了后续活体检测的成功率。图 3 中显示了硅胶头模的原始伪彩色图和归一化后的伪彩色图(由多光谱图像中第 6、5、4 波段作为 R、G、B 波段组合而成), 同时还对比了不同区域的同种材质所对应的反射曲线。

由图 3 可知, 光谱归一化操作提升了同种材质的反射曲线的一致性, 凸显了物体的反射特性, 降低了光照条件的影响, 增加了不同材质的可区分性。除此之外, 光谱归一化后的反射曲线取值为 0 到 1, 降低了神经网络模型的收敛难度。

2.3 基于面部多区域联合的 Transformer 模型

神经网络模型能够提取图像的深层次特征, 在大规模的数据上进行训练后, 能够获得很好的鲁棒性, 被广泛应用于图像分类和识别任务。目前, 基于卷积神经网络(convolutional neural networks, CNN)模型的人脸活体检测技术获得了业界认可^[17-18], 但是, 该类模型仅能提取人脸

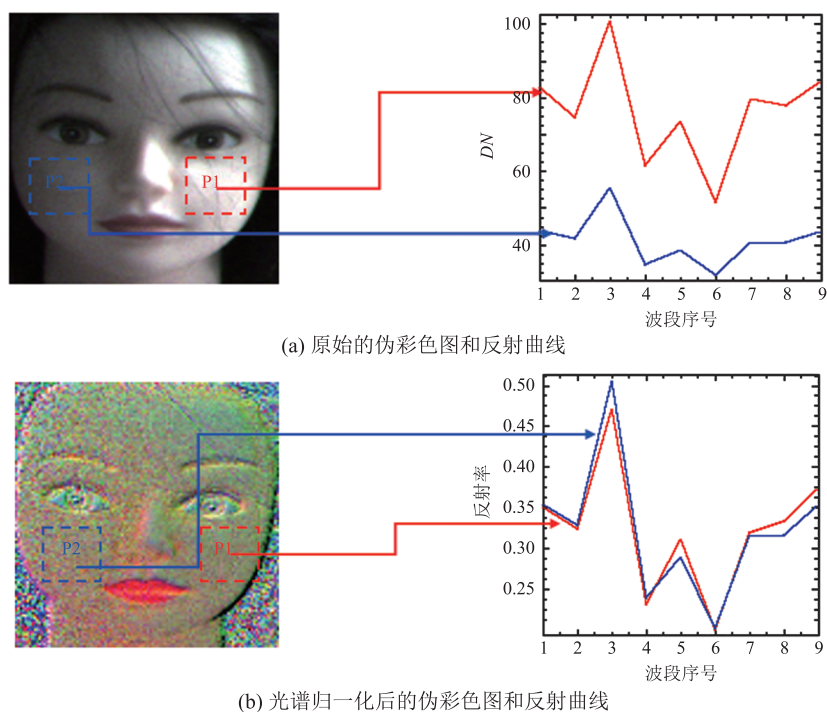


图3 光谱归一化前后的伪彩色图和反射曲线

Fig. 3 Pseudo-color images and reflection curves before and after spectral normalization

局部区域的特征，难以关注区域之间的关联性，仍会受到部分高逼真假体的攻击。Dosovitskiy等^[19]提出的VIT(vision Transformer)模型将基于Transformer的神经网络模型引入图像分类领域，将图像划分成多个图像块，利用注意力机制捕获图像块之间的关联，提取图像深层次特征，为视觉理解任务开辟了新的思路。由于人脸不同区域的组成成分及纹理结构不尽相同，反映在图像上表现为光谱特征和空间特征不同，故本文提出

FPTtransformer，以获取完备的人脸特征，并作为独特的生物特征，以提高人脸活体检测的准确率。考虑到活体检测易受周围环境干扰，且鉴于人脸典型的五官结构，本文选取6个人脸核心区域作为神经网络模型的输入，包括左眼、右眼、鼻子、左脸颊、右脸颊和嘴唇，按照一定顺序输入Transformer模型，最终输出活体检测结果，模型结构如图4所示。

由图4可知，本文提出模型的特征提取能力

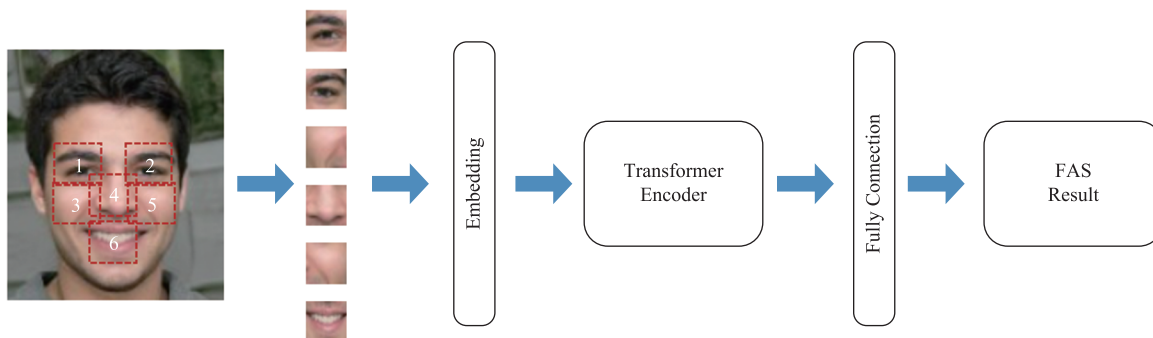


图4 FPTransformer 结构图

Fig. 4 Schematic diagram of FPTransformer

主要依赖于 Transformer Encoder 模块, 该模块利用独特的注意力机制, 关联各个人脸区域块, 综合提取人脸的光谱特征和空间特征, 提炼深层次的人脸生物特征。Transformer Encoder 模块的核心注意力机制如图 5 所示。

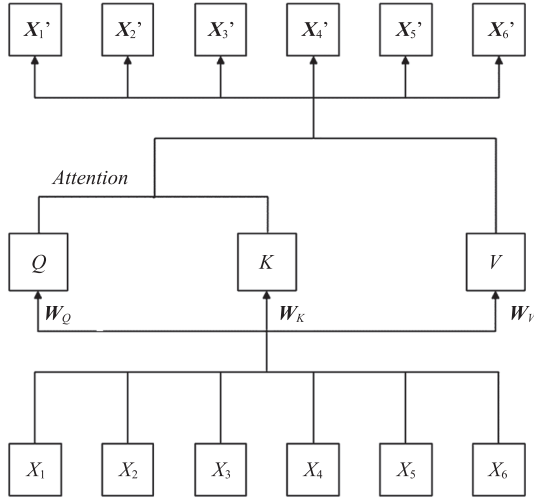


图 5 Transformer Encoder 模块的核心注意力机制图

Fig. 5 Core attention mechanism diagram of Transformer Encoder module

图 5 中的注意力机制计算如公式 (3) ~ (7) 所示。

$$Q = X \times W_Q \quad (3)$$

$$K = X \times W_K \quad (4)$$

$$V = X \times W_V \quad (5)$$

$$Attention(Q, K) = \text{Softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) \quad (6)$$

$$X' = Attention(Q, K) \times V \quad (7)$$

其中, $X = (X_1, X_2, \dots, X_6)$ 为输入数据; W_Q 、 W_K 、 W_V 分别为计算 Q 、 K 、 V 的权重矩阵; d_k 为缩放系数, X' 为光谱和空间联合特征。

由公式 (3) ~ (7) 可知, Transformer Encoder

模块根据输入图像块之间的关联计算相应的注意力矩阵, 并对图像进行特征提取。本文提出的 FPTransformer 通过多个 Transformer Encoder 的堆叠, 逐步获取深层次特征, 进而获取整个人脸的生物特征, 最终经过 Fully Connection 进行活体检测分类, 得到相应的活体检测结果。

综上所述, 本文对人脸光谱特性进行研究 (活体检测流程图如图 6 所示), 结合智能安防领域摄像头的特点, 利用 MSFA 获取多光谱图像, 经人脸检测模型提取人脸面部图像, 采用光谱归一化方法对其进行数据预处理, 分割多个人脸核心区域, 提出 FPTransformer, 以提取人脸面部生物特征, 完成高准确率的人脸活体检测。

与现有的人脸活体检测方案相比, 多光谱数据包含了丰富的光谱信息和空间信息, 提升了人脸的可区分性。本文提出的 FPTransformer 利用人脸核心区域之间的关联性提取人脸的本质生物特征, 摒弃了部分受成像环境影响较大的区域, 提升了模型的鲁棒性及活体检测准确率。

3 实验及分析

实验过程中将本团队提出的 FPTransformer 与几种常用的人脸活体检测方法在相同的数据集上进行了比较。并对实验结果进行讨论分析, 充分说明本文提出方法的有效性。同时, 本文将对光谱归一化方法的有效性进行进一步论证。

3.1 实验数据及实验环境

本文利用 MSFA 采集真人和不同材质假体的人脸数据, 通过人脸检测模型提取相应的人脸区域, 为了方便神经网络模型处理, 缩放人脸区域

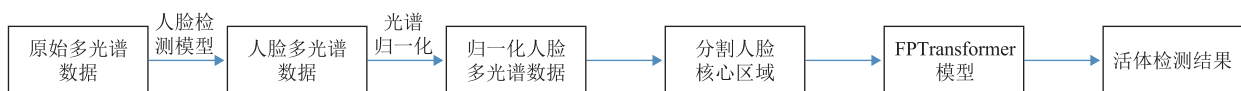


图 6 FPTransformer 模型流程图

Fig. 6 Flow chart of FPTransformer model

至相同尺寸 $112 \times 112 \times 9$ 像素。经统计,多光谱人脸数据集包含真人人脸数据 75 371 张,假体人脸数据 67 324 张。按照 3:1 的比例划分成训练集和验证集,训练集包含人脸图像 111 992 张,验证集包含人脸图像 30 703 张,真人和假体数据均涉及不同人脸姿态、不同采集距离及不同光照条件。其中,假体材质包含硅胶、乳胶、树脂、打印纸。

实验环境配置如下:操作系统为 Ubuntu 系统,深度学习框架为 Pytorch1.7.0 和 Cuda11.4,采用 4 张 NVIDIA GeForce RTX 3080 显卡训练神经网络模型。在模型训练时,最大迭代次数为 200 个 epochs,优化器为 Adam,损失函数为交叉熵损失函数。

3.2 实验结果分析

本小节将提出的方法与基于 CNN 的模型(包括采用差分卷积的 CDCN 模型^[20]、采用轻便网络结构的 Mobilenet 模型^[21]、采用密集连接的 Densenet 模型^[22])和传统的 VIT 网络模型进行比较,以攻击错误分类率(attack presentation classification error rate, APCER)、真实样本错误分类率(normal presentation classification error rate, NPCER)、平均分类错误率(average classification error rate, ACER)和准确率

(accuracy, ACC)为评价标准^[23],计算公式如公式(8)~(11)所示。

$$APCER = \frac{FN}{FN+TP} \quad (8)$$

$$NPCER = \frac{FP}{FP+TN} \quad (9)$$

$$ACER = \frac{APCER+NPCER}{2} \quad (10)$$

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

其中, TP 为模型预测为正的正样本; TN 为模型预测为负的负样本; FP 为模型预测为正的负样本; FN 为模型预测为负的正样本。

为了说明多光谱图像在人脸活体检测方面的优越性,本文将其中的第 6、5、4 波段组合成相应的 RGB 图像,与原始多光谱图像就活体检测效果进行比较。不同模型在验证集上的实验结果如表 1 所示,同一模型包含了多光谱图像数据集及 RGB 图像数据集的活体检测效果。

由表 1 可知,对于同一模型而言,多光谱图像的活体检测效果普遍优于对应的 RGB 图像,说明多光谱图像能够反映物体更多的反射特性,增强真人与假体的可区分性。对于不同的模型而言,基于 CNN 的模型在真人活体检测方面具有

表 1 不同神经网络模型的实验结果表

Table 1 Experimental results of different neural network models

Model	APCER (%)	NPCER (%)	ACER (%)	ACC (%)
Densenet(RGB)	3.35	14.88	9.11	92.66
Densenet	0.65	13.93	7.29	94.49
Mobilenet(RGB)	3.15	18.62	10.89	91.18
Mobilenet	1.72	15.44	8.58	93.25
CDCN(RGB)	2.06	14.62	8.34	93.34
CDCN	1.04	12.53	6.78	94.75
VIT(RGB)	4.62	11.90	8.26	92.71
VIT	2.67	11.32	6.99	94.17
FPTransformer(RGB)	3.75	13.95	8.85	92.52
FPTransformer	2.02	8.18	5.10	95.72

较好的表现, APCER 值均低于 2%, 而对假体的区分效果较差, 说明该类神经网络模型缺乏获取人脸本质特征的能力, 仅依据浅层特征对人脸进行活体检测, 在面对复杂的外部环境及多变的人脸攻击方式时, 模型泛化能力较差。而反观基于 Transformer 的神经网络模型, 虽然在真人活体检测方面的准确率存在一定程度下降, 但是对于假体的活体检测大幅提高, 且拥有更低的 ACER 值。值得注意的是, 本文提出的基于面部多区域联合的 Transformer 模型在 APCER 值仅下降 1% 左右的基础上, NPCER 值提升超过 4%, 且拥有最低的 ACER 值 5.10%, 以及最高的 ACC 值 95.72%, 说明本文提出的方法通过捕获人脸不同区域之间的关联性, 提取了人脸的本质生物特征, 获得了更好的人脸活体检测效果。为了进一步说明本文提出方法的泛化能力, 对不同材质的人脸假体数据分别进行统计, 如表 2 所示, 计算相应的准确率, 结果如图 7 所示, 模型输入均为

表 2 验证集假体分布表

Table 2 Distribution of the validation dataset of prosthesis

材料	数量
硅胶	2 182
乳胶	1 553
树脂	1 900
打印纸	5 000

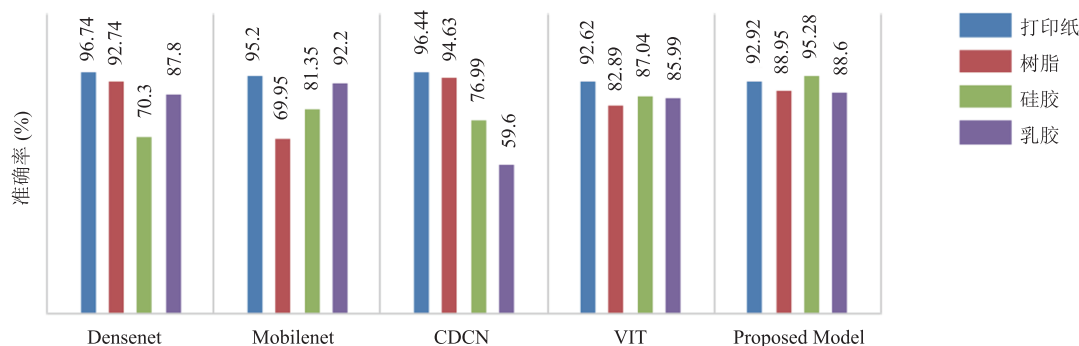


图 7 不同模型在假体验证集上的实验结果

Fig. 7 Experimental results of different models on the validation dataset of prosthesis

MSFA 采集的多光谱数据。

由图 7 可知, 基于 CNN 的模型对不同材质假体的活体检测准确率差异较大, 对 2D 打印假体普遍具有较高的准确率, 对树脂面具同样具有较好的可区分性, 而对于细节纹理丰富的 3D 硅胶、乳胶头模而言, 则表现较差, 说明该类神经网络模型对材质的可区分性较差, 仅根据浅层纹理特征进行人脸活体检测; 基于 Transformer 的神经网络模型对不同材质的假体均具有较好的检测效果, 且波动范围较基于 CNN 的模型更小。尤其是本文提出的模型, 对 3D 硅胶头模的活体检测准确率达到 95.28%, 充分说明该模型能够提取人脸的生物特征, 对不同材质的人脸进行准确划分, 实现高精度的人脸活体检测。

另外, 本文还将提出的光谱归一化方法与其余常见的数据归一化方法(包括最大值归一化方法和通道归一化方法)进行比较, 选用的神经网络模型为本文提出的基于面部多区域联合的 Transformer 模型, 实验结果如表 3 所示。

由表 3 可知, 与其他归一化方法相比, 光谱归一化方法的各项指标均有一定程度提升, APCER 值、ACER 值、ACC 值均提升不低于 2%, 说明光谱归一化方法在处理多光谱数据时能够提升不同材质之间的可区分性, 以及同种材质之间的一致性, 具有独特的优越性。

表 3 不同归一化方法结果表

Table 3 Results of different normalization methods

方法	APCER (%)	NPCER (%)	ACER (%)	ACC (%)
最大值归一化	4.26	15.87	10.07	91.49
通道归一化	5.37	8.83	7.10	93.36
光谱归一化	2.02	8.18	5.10	95.72

4 结 论

本文提出了一种基于面部多区域联合的 Transformer 模型，不仅获取了人脸不同核心区域的光谱特征和空间特征，还利用它们之间的关联性，综合提取了人脸的深层次生物特征，获得了高准确率的活体检测结果。与智能安防领域中现有的人脸活体检测方案相比，本文引入了多光谱人脸数据，拓展了光谱维度，增加了不同材质人脸的可区分性，并探索了基于 Transformer 的网络模型在人脸活体检测任务中的应用性。实验结果表明，在相同的人脸数据集上，本文提出的模型获得了高达 95.72% 的活体检测准确率，以及低至 5.10% 的错分率，优于其他常用的人脸活体检测模型，充分说明了该方法的有效性。在之后的研究中，不仅需要对面脸区域划分进行更多的研究，以求选择最优的区域组合，还需要对神经网络模型进行优化，获得更为丰富的人脸特征，提高真人活体检测准确率，以在消费电子领域获得更为广泛的应用。另外，由于多光谱滤波阵列的滤光片选择有限，不能针对性选择波段，因此无法获得可区分性最好的人脸特征波段。

参 考 文 献

- [1] Patel K, Han H, Jain AK. Secure face unlock: spoof detection on smartphones [J]. IEEE Transactions on Information Forensics and Security, 2016, 11(10): 2268-2283.
- [2] Ming ZH, Visani M, Luqman MM, et al. A survey on anti-spoofing methods for facial recognition with RGB cameras of generic consumer devices [J]. Journal of Imaging, 2020, 6(12): 139.
- [3] Ramachandra R, Busch C. Presentation attack detection methods for face recognition systems: a comprehensive survey [J]. ACM Computing Surveys, 2017, 50(1): 8.1-8.37.
- [4] Angelopoulou E, Molana R, Daniilidis K. Multispectral skin color modeling [C] // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001: 635.
- [5] Zhang ZW, Yi D, Lei Z, et al. Face liveness detection by learning multispectral reflectance distributions [C] // Proceedings of the 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), 2011: 436-441.
- [6] Boulkenafet Z, Komulainen J, Hadid A. Face antispoofing using speeded-up robust features and fisher vector encoding [J]. IEEE Signal Processing Letters, 2017, 24(2): 141-145.
- [7] Ji QB, Xu SG, Chen XD, et al. A cross domain multi-modal dataset for robust face anti-spoofing [C] // Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), 2021: 4309-4316.
- [8] Tirunagari S, Poh N, Windridge D, et al. Detection of face spoofing using visual dynamics [J]. IEEE Transactions on Information Forensics and Security, 2015, 10(4): 762-777.
- [9] Boulkenafet Z, Komulainen J, Hadid A. Face spoofing detection using colour texture analysis [J]. IEEE Transactions on Information Forensics and Security, 2016, 11(8): 1818-1830.
- [10] 顾梦霞. 基于最大 Gabor 相似度的多光谱人脸识别系统 [J]. 激光杂志, 2021, 42(1): 192-196.

Gu MX. Multi-spectral face recognition system based on maximum Gabor similarity [J]. Laser Journal, 2021, 42(1): 192-196.

- [11] 徐姚文, 毋立芳, 刘永洛, 等. 基于解耦空间异常检测的人脸活体检测算法 [J]. 信号处理, 2022, 38(12): 2469-2485.
Xu YW, Wu LF, Liu YL, et al. A face anti-spoofing algorithm based on anomaly detection in disentangling space [J]. Journal of Signal Processing, 2022, 38(12): 2469-2485.
- [12] 田野, 项世军. 基于 LBP 和多层 DCT 的人脸活体检测算法 [J]. 计算机研究与发展, 2018, 55(3): 643-650.
Tian Y, Xiang SJ. LBP and multilayer DCT based anti-spoofing countermeasure in face liveness detection [J]. Journal of Computer Research and Development, 2018, 55(3): 643-650.
- [13] Hernandez-Ortega J, Fierrez J, Morales A, et al. Time analysis of pulse-based face anti-spoofing in visible and NIR [C] // Proceedings of the Computer Vision and Pattern Recognition, 2018: 544-552.
- [14] Kuang HF, Ji RR, Liu H, et al. Multi-modal multi-layer fusion network with average binary center loss for face anti-spoofing [C] // Proceedings of the 27th ACM International Conference on Multimedia, 2019: 48-56.
- [15] 魏东, 岳许要, 章烈剡, 等. 基于注意力融合的多模态人脸活体检测算法 [J]. 现代信息技术, 2022, 6(20): 65-70.
Wei D, Yue XY, Zhang LB, et al. Attention-based multimodal fusion for face anti-spoofing [J]. Modern Information Technology, 2022, 6(20): 65-70.
- [16] Nowara EM, Sabharwal A, Veeraraghavan A. PPGSecure: biometric presentation attack detection using photoplethysmograms [C] // Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition, 2017: 56-62.
- [17] Souza GB, Papa JP, Marana AN. On the learning of deep local features for robust face spoofing detection [C] // Proceedings of the 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018: 258-265.
- [18] Shen T, Huang YY, Tong ZJ. FaceBagNet: bag-of-local-features model for multi-modal face anti-spoofing [C] // Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019: 1611-1616.
- [19] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale [C] // Proceedings of the International Conference on Learning Representations, 2021: 1-21.
- [20] Yu ZT, Zhao CX, Wang ZZ, et al. Searching central difference convolutional networks for face anti-spoofing [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 5295-5305.
- [21] Sandler M, Howard A, Zhu ML, et al. MobileNetV2: inverted residuals and linear bottlenecks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [22] Huang G, Liu Z, Maaten L, et al. Densely connected convolutional networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700-4708.
- [23] George A, Mostaani Z, Geissenbuhler D, et al. Biometric face presentation attack detection with multi-channel convolutional neural network [J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 42-55.