

改进型 MobileNetV3 轻量级人脸活体检测算法

李侯彤¹ 宋伟² 南新元¹

¹(新疆大学电气工程学院 新疆 乌鲁木齐 830047)

²(新疆维吾尔自治区特种设备检验研究院 新疆 乌鲁木齐 830002)

摘要 为了解决人脸识别系统中的欺诈攻击问题,以及目前基于深度学习的活体检测方法大多以大型卷积网络作为主干网络,导致模型结构复杂、计算量大等问题,提出一种改进型 MobileNetV3 的轻量级人脸活体检测算法。对 MobileNetV3 中利用全局平均池化计算通道注意力权重和使用双非线性全连接层存在的不足进行讨论,提出新的注意力机制 EFCANet,并利用 EFCANet 网络对 MobileNetV3 轻量级卷积神经网络进行改进。实验结果表明,改进后轻量级活体检测算法,在检测精度、网络模型大小、损失值和等错误率等方面有着不错的表现。

关键词 活体检测 深度学习 注意力机制 轻量级网络

中图分类号 TP391.4

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.09.028

LIGHTEWIGHT FACE ANTI-SPOOFING ALGORITHM BASED ON IMPROVED MOBILENETV3

Li Yutong¹ Song Wei² Nan Xinyuan¹

¹(College of Electrical Engineering, Xinjiang University, Urumqi 830047, Xinjiang, China)

²(Xinjiang Special Equipment Inspection and Research Institute, Urumqi 830002, Xinjiang, China)

Abstract In order to solve the problem of fraud attacks in face recognition systems and the problem that most deep learning based in-vivo detection methods use large convolutional networks as the backbone network, which leads to complex model structure and large computation amount, an improved lightweight face detection algorithm based on MobileNetV3 is proposed. This paper discussed the shortcomings of using global average pooling to calculate channel attention weight in MobileNetV3 and using double nonlinear fully connected layer. A new attention mechanism EFCANet was proposed, and the EFCANet network was used to improve MobileNetV3 lightweight convolutional neural network. The experimental results show that the improved light-weight face anti-spoofing algorithm has a good performance in detection accuracy, network model size, loss value, and equal error rate.

Keywords Anti-spoofing Deep learning Attention mechanism Lightweight network

0 引言

人脸作为人的一种特殊生物特征,有着非常广泛的应用前景,特别是近年来深度学习的快速发展,使得人脸识别技术越来越成熟与完善。人脸识别具有非强制性、非接触式及隐蔽性等特点,在智能门禁、移动支付、医疗卫生、身份认证甚至公安刑侦上都有着广泛的

应用^[1],但随着互联网技术和移动设备的发展,人脸识别还存在的问题。

人们的个人信息充斥在各大社交软件,甚至还有流氓软件在用户不知情的情况下收集个人信息,这些人脸信息容易被不法分子所利用,人脸的活体检测问题越来越突出。人脸活体检测指的是判断当前检测的人脸是活体人脸还是假体人脸^[2],即是用户本身还是不法分子制作的仿制人脸,人脸活体检测是确保人脸

识别安全性的重要防线,具有很高的研究价值。

当前常见的假体人脸包括照片类假体人脸、视频类假体人脸和3D假体人脸^[3]。活体检测方法可分为交互式和非交互式人脸活体检测。交互式常见于银行等用户身份认证的场景,通过随机动作如眨眼、转头或声音检测等方式来检测用户身份的真实性,该方法准确度高且通用性强,但需要用户根据提示进行配合,导致用户体验差同时不适用于对验证速度有要求的场景。非交互式方法有基于颜色纹理、生命信息、图像质量、基于近红外相机等硬件设备及深度学习等方法,非交互式的活体检测方式具有便捷性、无须用户配合、检测速度快等特点,受到更多学者的关注。

随着深度学习在图像领域取得的巨大成功,越来越多的学者提出基于深度学习的活体检测方法。Zhang等^[4]以ResNet为主干网络,设计出融合RGB、深度图像和红外图像多模态人脸反欺诈网络,但深度相机的价格比较昂贵,造成检测成本的增加。Zhang等^[5]提出FeathersNets网络,给出活体检测的网络轻量化的解决方案,但红外和深度图像的实时获取依然需要较大的数据量。Guo等^[6]认为获取假体攻击数据的代价是比较高昂的,故提出一种三维虚拟欺诈数据的合成方法,利用合成的数据进行网络的训练。Song等^[7]在纹理和颜色特征的基础上,提出一种新颖的非侵入式人脸活体检测方法。陈放等^[8]从HSV和YCb-Cr颜色空间提取LBP特征直方图,并利用SVM进行活体假体分类。邓雄等^[9]提出利用卷积神经网络分别从RGB、HSV、LBP图中提取特征,对各通道提取的特征进行融合,但三支的网络结构使得模型的复杂度和计算量增大。

通过查阅近年来的文献,常见的思路是构建多流卷积神经网络,从不同颜色空间、模态进行特征的提取,这样虽然会带来不错的性能提升,但相应地会增加模型的复杂度和计算量,不利于模型在算力有限的嵌入式设备上应用。同时随着注意力机制的出现,使网络有着显著的性能提升,故在模型搭建中得到了广泛的应用。但通道注意力机制使用全局平均池化进行注意力权值的提取,这样的做法存在信息缺失的情况。针对上述问题,本文以MobileNetV3^[10]为主干网络构建人脸活体检测模型,进行网络的轻量化设计,在注意力机制的设计上,从频域角度对通道注意力权重进行提取,解决信息缺失的问题,并通过实验对改进后的模型进行验证,取得了不错的效果。

1 神经网络与注意力机制

深度学习是机器学习领域的一个重要分支,是人

们实现人工智能的重要手段。随着互联网的快速发展,特别是GPU的发展使得计算机的算力得到了飞跃,让深度学习得到了长足的发展。卷积神经网络是深度学习中常见的模型之一,其主要结构为输入层、卷积层、激励层、池化层和全连接层。输入层为整个网络输入初始数据,对输入数据进行去均值、归一化等处理。卷积层通过多维度多通道卷积核与输入数据进行卷积运算,达到提取特征的目的。激励层会根据任务的不同需求,选择不同的激活函数对卷积层的输出进行非线性映射。池化层的主要作用为降低特征维度、对数据进行压缩减少参数量、防止过拟合。全连接层相当于网络的输出层,在综合网络所提取的特征后利用softmax等函数对结果进行输出。随着卷积神经网络在计算机视觉上取得的巨大成就,越来越多的卷积神经网络模型被提出,如:VGG^[11]、ResNet^[12]、Inception^[13]、MobileNets^[14]等网络。

注意力机制是让网络有重点学习的一种机制,希望模型能够关注有用信息同时抑制无用信息,是提高网络模型性能的高效做法^[15]。计算机视觉领域常用的注意力机制可分为空间注意力机制、通道注意力机制和混合域注意力机制。空间注意力机制通过利用卷积将原始输入图像的空间信息转换至其他空间,保留关键信息,达到提高模型性能的目的,如STN^[16]空间转换网络。通道注意力机制按通道对特征图进行权值提取,用权值表征不同特征图的重要度,网络根据权值大小对不同通道的特征图进行有重点的学习,如SENet^[17]通过全局平均池化对通道进行压缩,再通过双全连接层结构进行激励获取注意力权值,最后再将注意力权值加权到原网络中;ECANet^[18]注意力机制为解决注意力模块给主干网络带来的计算负担,设计出超轻量级注意力模块;FCANet^[19]从频域的角度重新阐释了注意力机制的作用。混合域注意力机制模型是空间域和通道域的组合形式,如:CBAM^[20]利用最大池化和平均池化对通道进行特征提取,再利用串联空间注意力模块获取特征空间的重要度;DANet^[21]通过并联空间和通道注意力机制进行模块的设计。

2 轻量级网络 MobileNetV3

2.1 轻量级网络 MobileNet

MobileNetV3网络是在MobileNetV2^[22]基础上,调整模型结构中冗余的部分,引入SENet注意力机制,对网络进行改进。同时为了减轻计算量设计出hard_swish

激活函数。

$$h\text{-swish}[x] = x \frac{\text{ReLU} 6(x+3)}{6} \quad (1)$$

使用 hard_swish 激活函数替换 Sigmoid,降低网络的计算开销,利于模型在移动端的应用。

模型首先利用 1×1 的卷积对通道进行扩展,使得网络能够在更高的维度进行特征的提取,使用 depthwise 卷积和 pointwise 卷积的组合形式替换标准卷积,让模型的参数量和计算量得到控制,更便于模型能够在移动端上得到应用。网络模型的深度是影响模型效果的重要因素之一,随着深度学习的快速发展和计算机算力的不断完善与提高,卷积神经网络的模型结构越来越复杂、层数越来越深。但随着模型的深度加深,伴随着梯度弥散和梯度爆炸等问题,导致模型在训练的过程中无法收敛、误差增大,故 MobileNetV2 在设计时,引入 ResNet 网络所使用的残差结构。

在 MobileNetV2 具有残差结构的深度可分离卷积基础上,MobileNetV3 将 SENet 通道注意力机制合并到网络中,MobileNetV3 基本结构如图 1 所示。

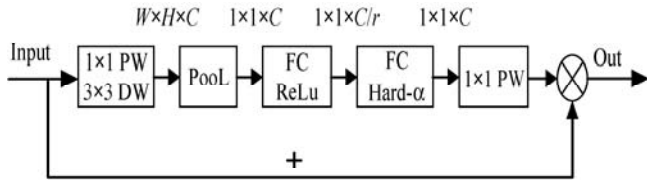


图 1 MobileNetV3 block

在深度可分离卷积后进行 SENet 操作,利用注意力机制来提高模型的性能。SENet 网络包括 Squeeze 和 Excite 两个部分,利用全局平均池化对 $C \times W \times H$ 的特征图进行注意力权重的提取,得到具有全局感受野且通道为 C 的 1×1 特征图,再经过两次全连接层得到最终的通道注意力权重,对每个通道进行打分,最终每个通道的注意力权重通过乘法加权到原卷积核中,实现模型有重点地学习,增强重要特征,抑制无用特征,从而使得网络在特征的提取上具有更强的指向性。同时,使用分段函数 h-swish 模拟 Sigmoid 激活函数,提高计算速度的同时,对模型的精度没有明显影响。

2.2 MobileNetV3 的改进

MobileNetV3 利用 SENet 注意力机制来达到提高模型性能的目的,但还存在不足之处。在 MobileNetV3 网络中,利用全局平均池化层对特征图的所有特征值相加求平均,利用得到的数值表示对应的特征图。但从频域的角度分析,利用 2D-DCT 反变换公式进行推导:

$$x_{i,j}^{2d} = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} f_{h,w}^{2d} \cos\left(\frac{\pi h}{H}(i+0.5)\right) \cos\left(\frac{\pi w}{W}(j+0.5)\right) \quad (2)$$

s. t. $i \in \{0, 1, \dots, H-1\}, j \in \{0, 1, \dots, W-1\}$

式中: $f_{h,w}^{2d}$ 为 2D-DCT 频谱; $x^{2d} \in \mathbf{R}^{H \times W}$ 为网络的输入; H 和 W 分别对应 x^{2d} 的高和宽。

$$\text{令 } B_{h,w}^{i,j} = \cos\left(\frac{\pi h}{H}(i+0.5)\right) \cos\left(\frac{\pi w}{W}(j+0.5)\right) \quad (3)$$

$B_{h,w}^{i,j}$ 为 2D-DCT 权重。将式(3)代入式(2)中得到式(4)。

$$x_{i,j}^{2d} = f_{0,0}^{2d} B_{0,0}^{i,j} + f_{0,1}^{2d} B_{0,1}^{i,j} + \dots + f_{H-1,W-1}^{2d} B_{H-1,W-1}^{i,j} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} x_{i,j}^{2d} B_{0,0}^{i,j} + f_{0,1}^{2d} B_{0,1}^{i,j} + \dots + f_{H-1,W-1}^{2d} B_{H-1,W-1}^{i,j} = \text{gap}(x^{2d}) HW B_{0,0}^{i,j} + f_{0,1}^{2d} B_{0,1}^{i,j} + \dots + f_{H-1,W-1}^{2d} B_{H-1,W-1}^{i,j} \quad (4)$$

由式(4)可知网络的特征图可以通过不同频域分量来进行表示。而全局平均池化 $\text{gap}(x^{2d})$ 只是 2D-DCT 的一种特殊形式,其与 2D-DCT 最低频分量成比例关系。通道注意力机制的数学表达如式(5)所示。

$$a_u = \text{sigmoid}(fc(\text{gap}(X))) \quad (5)$$

式中: X 为网络输入; $\text{gap}(X)$ 为全局平均池化层; fc 为全连接层或一维卷积; sigmoid 为激活函数。结合式(4)、式(5)可知网络在使用全局平均池化进行注意力权重提取时,存在信息缺失的情况。

同时在 MobileNetV3 网络中,通道注意力使用双非线性全连接结构,在使用第一个全连接层计算通道的注意力权重时,对通道进行压缩,之后再通过第二个全连接层将压缩的通道恢复到原通道的大小。这么做的目的,其一是增加通道之间的非线性跨通道作用,其二是降低维度,减少全连接带来的高计算量。但降维会对通道关注度造成影响,同时每个通道都进行交互作用,效率是低效且非必要的,并且两层全连接的使用会带来较大的计算量。

针对以上两点,结合 FCANet 和 ECANet 提出一种高效的频域通道注意力机制 (EFCANet),用于改进 MobileNetV3。EFCANet 注意力机制结构如图 2 所示。

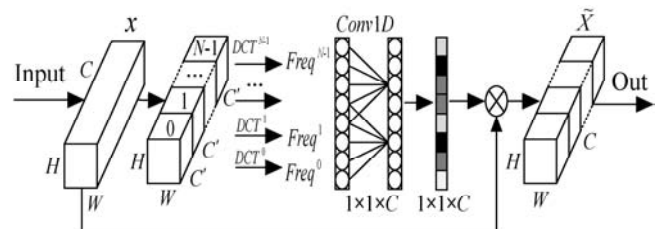


图 2 EFCANet 注意力机制结构

首先,将 $C \times W \times H$ 特征图按通道分成 N 组,每组通道 $C' = C/N$,利用式(6)计算每组的离散余弦变换权重。

$$F_{\text{req}}^i = 2DDCT^{U,V}(X^i) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X_{:,h,w}^i B_{h,w}^{u,v} \quad (6)$$

$$\text{s. t. } i \in \{0, 1, \dots, n-1\}$$

式中: U, V 表示频率分量的二维索引; X^i 为第 i 组输入的特征图。在得到所有分组的频率权值后, 将所有的频率分量进行拼接。经过 DCT 变换后生成 $1 \times 1 \times C$ 的新特征图, 通过由式 (4) 可知 DCT 能够获取更多的频率信息。其次, 使用长度为 K 的一维卷积代替 MobileNetV3 中双 FC 结构, K 值根据通道 C 进行自适应的调整, 在卷积神经网络中, 通道的维度通常是 2 的倍数, 所以 K 与通道数 C 的非线性关系如下:

$$C = \theta(k) = 2^{(r \times k - b)} \quad (7)$$

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{r} + \frac{b}{r} \right\rfloor_{\text{odd}} \quad (8)$$

式中: θ 和 ψ 表示二次函数; 通过调整超参数 r, b , 可以调节通道之间相互作用的范围, 在通道的维度较高时, 通过调整超参数 r, b , 增大 K 值, 提高交互的距离, 而在通道的维度较低时, 通过调整超参数 r, b , 降低 K 值; $\lfloor T \rfloor_{\text{odd}}$ 表示最接近 T 的奇数。根据通道的 K 个邻域进行局部非线性跨通道交互, 这样降低模型复杂度和计算量的同时, 还能提高非线性跨通道交互的效率。最后在获取最终的通道注意力机制权重后, 经过 h-swish 函数重新加权到 $C \times W \times H$ 特征图上。

2.3 算法流程与网络结构

基于 MobileNetV3 的人脸活体检测算法流程主要包括三部分: 输入数据的预处理、网络的特征提取、活体假体人脸分类。首先对输入图像进行人脸检测, 提取人脸, 对输入的人脸图像进行缩放, 以满足网络的输入要求; 其次通过主干网络对输入图像提取特征; 最后根据提取的特征进行分类。改进型 MobileNetV3 基本模块如图 3 所示。

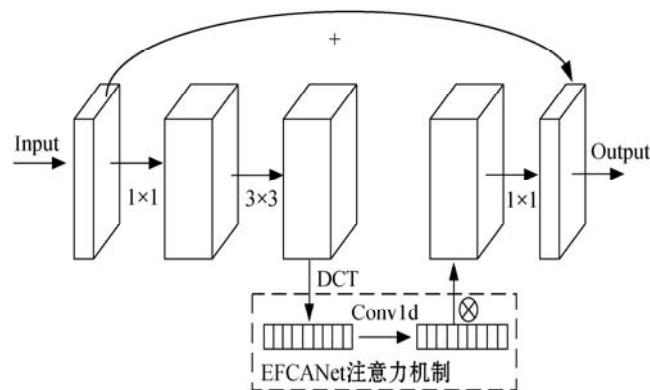


图3 MobileNetV3 + EFCANet 基本 block

表 1 为改进后的 MobileNetV3 网络结构, 其中: Input 表示输入; Op 表示运算方式; E 表示扩展层的通

道数; out 表示输出; Att 表示是否使用 EFCANet 注意力机制; NL 表示非线性激活函数, 其中 HS 表示 h-swish 激活函数, RE 表示 ReLU 激活函数; s 表示步长系数。

表 1 MobileNetV3 + EFCANet 网络结构

Input	Op	E	Out	Att	NL	s
$224^2 \times 3$	Conv2d, 3×3	—	16	—	HS	2
$112^2 \times 16$	block, 3×3	16	16	Y	RE	2
$28^2 \times 24$	block, 3×3	72	24	Y	RE	2
$28^2 \times 24$	block, 3×3	88	24	Y	RE	1
$28^2 \times 24$	block, 5×5	96	40	Y	RE	2
$14^2 \times 40$	block, 5×5	240	40	Y	HS	1
$14^2 \times 40$	block, 5×5	240	40	Y	HS	1
$14^2 \times 40$	block, 5×5	120	48	Y	HS	1
$14^2 \times 48$	block, 5×5	144	48	Y	HS	1
$14^2 \times 48$	block, 5×5	288	96	Y	HS	2
$7^2 \times 96$	block, 5×5	576	96	Y	HS	1
$7^2 \times 96$	block, 5×5	576	96	Y	HS	1
$7^2 \times 96$	Conv2d, 1×1	—	576	Y	HS	1
$7^2 \times 576$	Pool, 7×7	—	—	—	—	1
$1^2 \times 576$	Conv2d, 1×1	—	1280	—	HS	1
$1^2 \times 1\ 280$	Conv2d, 1×1	—	2	—	—	—

3 实验

3.1 数据库与训练设备

实验数据采用 NUAA^[23]、Replay-Attack^[24] 人脸活体检测数据集。NUAA 数据集共包含 15 个用户, 活体人脸数据 5 120 幅, 假体人脸数据 7 523 幅。采用可见光摄像头采集不同情景、不同光照条件下的人脸信息, 假体攻击方式为打印攻击方式, 包括照片的扭曲、左右、上下旋转。Replay-Attack 数据集的数据为 1 300 个视频, 每段视频帧宽高为 320×240 , 帧速率为 25 帧/s, 假体人脸攻击方式为手持或者固定的照片与视频。实验环境为 Windows 10 操作系统, 16 GB 内存, 英伟达 GeForce RTX2060 显卡, 在 PyCharm2020 上使用 Pytorch 深度学习框架进行仿真实验。

3.2 实验数据处理

NUAA 数据集的数据形式为 640×480 分辨率的照片, 故需要对数据集进行人脸检测, 提取照片中的人脸信息, 保留人脸信息, 去除无用的背景信息, 这样有

利于提高模型的精度,对检测到的人脸缩放至 $224 \times 224 \times 3$ 以满足网络的输入需求。在对 Replay-Attack 数据集的数据进行处理时,将视频文件按帧提取图片,去除无效帧,共提取 14 166 幅图片,训练集和验证集分别占比 70% 与 30%。训练时使用随机梯度下降法 (SGD) 进行模型训练,计算如下:

$$l_i = l_{r_{start}} \times \frac{1.0}{(1.0 + d_{ecay} \times i)} \quad (9)$$

$$v = -d_x \times l_r + v \times m_{omentum} \quad (10)$$

式中: $l_{r_{start}}$ 为初始学习率,设置为 0.01; d_{ecay} 为权值衰减系数; d_x 为输入维度; v 为参数更新速度; $m_{omentum}$ 为动量参数,设置为 0.5,使用等间隔调整学习率,每迭代 20 次调整一次学习率,每次缩小为十分之一。

使用交叉熵损失函数计算 Loss 值,计算如下:

$$Loss(x, c_{class}) = -\log\left(\frac{\exp(x[c_{class}])}{\sum_j \exp(x[j])}\right) = -x[c_{class}] + \log\left(\sum_j \exp(x[j])\right) \quad (11)$$

式中: c_{class} 为预测值。

3.3 结果分析

表 2、表 3 为各模型在 NUAA、Replay-Attack 数据集上进行的实验仿真结果,主要从参数量 (Params)、计算量 (FLOPs)、准确度、损失值和等错误率 (EER) 对网络模型进行评价。

表 2 模型在 NUAA 数据集上的仿真结果

指标	V3	V3 + FCANet	V3 + EFCANet
Params/MB	1.66	1.66	1.21
FLOPs/MB	60.03	60.03	59.14
Acc	0.967 1	0.974 9	0.983 0
Loss	0.107 1	0.087 4	0.083 2
EER/%	1.101 6	0.806 8	0.709 2

表 3 模型在 Replay-Attack 数据集上的仿真结果

指标	V3	V3 + FCANet	V3 + EFCANet
Acc	0.972 3	0.982 4	0.985 8
Loss	0.076 7	0.038 5	0.035 1
EER/%	3.141 7	2.262 2	2.112 1
模型大小/MB	6.45	6.53	3.25

根据表 2 中参数量和计算量等评价指标发现,使用 DCT 离散余弦变换计算通道注意力权重,相比于全局平均池化,并不会带来额外的计算量和参数量,同时从频域角度对特征图提取权值,包含的信息更加全面

和准确。使用一维卷积代替双全连接层结构,使得改进后的 MobileNetV3 网络在计算量和参数量上有明显的降低。在 NUAA 数据集上准确度为 98.3%,等错误率 EER 达到 0.71%。

由表 3 可以看出,改进后的模型在 Replay-Attack 数据集上依然有着较好的性能,准确度为 98.58%,等错误率 EER 为 2.11%。利用一维卷积替换 V3 网络中的双非线性全连接层能极大地降低模型大小,改进后的模型大小比 V3 原网络模型减少了 49.6%。图 4 和图 5、图 6 和图 7 分别为改进后和未改进的 MobileNetV3 网络模型,其在 NUAA 和 Replay-Attack 数据集上损失和精度随迭代次数增加的变化情况。结果表明改进后的网络精度更高,Loss 值更低。

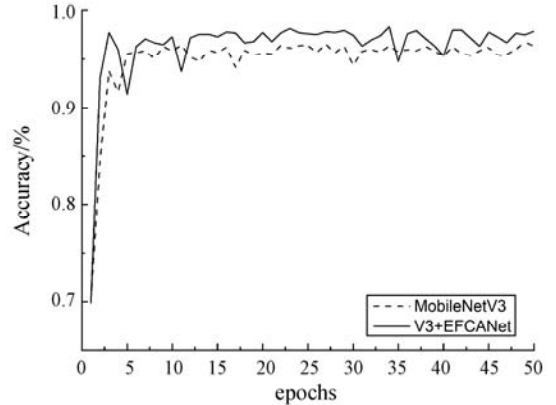


图 4 NUAA 数据集上训练精度随迭代次数增长的变化情况

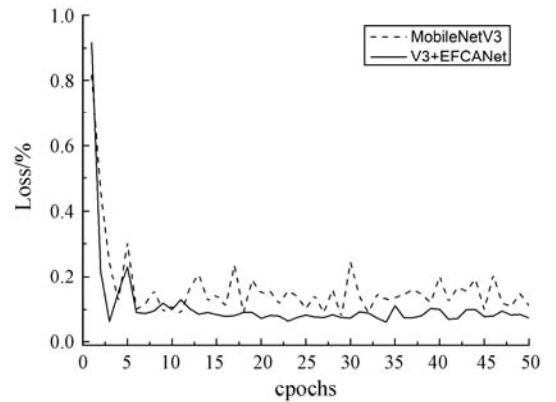


图 5 NUAA 数据集上损失值随迭代次数增长的变化情况

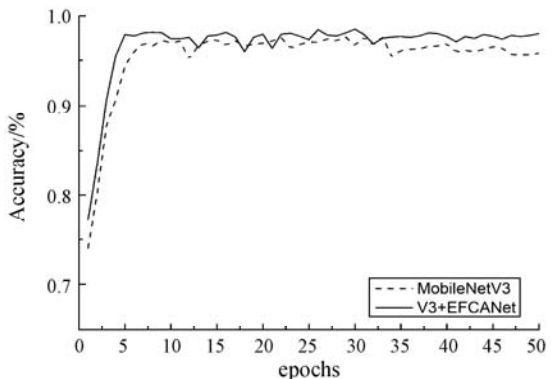


图 6 Replay-Attack 数据集上训练精度随迭代次数增长的变化情况

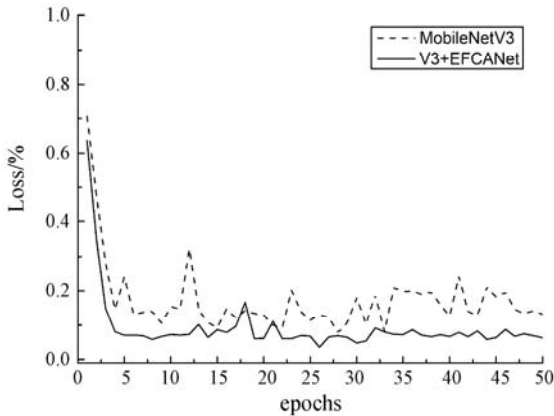


图7 Replay-Attack 数据集上损失值随迭代次数增长的变化情况

4 结 语

本文在 MobileNetV3 轻量级网络的基础上进行改进,提出一种改进型 MobileNetV3 的人脸活体检测算法。对原网络结构存在的不足进行分析,利用离散余弦变换计算通道注意力权重,解决全局平均池化在提取通道注意力权重时存在信息缺失的问题。同时对网络中的双非线性全连接层进行改进,达到提高通道交互效率以及降低模型参数量和计算量的目的。改进后的模型在 NUAA 和 Replay-Attack 活体检测数据集上进行实验验证,通过实验证明了改进后模型在性能上得到了提升,同时模型大小大幅降低。但活体检测算法不断升级改进的同时,人脸的假体攻击方式也是不断变化的,所以后续在模型面对新型假体攻击时能够自适应调整的问题上还需进一步研究。

参 考 文 献

- [1] 董艳花,张树美,赵俊莉. 有遮挡人脸识别方法综述[J]. 计算机工程与应用,2020,56(9):1-12.
- [2] 蒋方玲,刘鹏程,周祥东. 人脸活体检测综述[J]. 自动化学报,2021,47(8):1799-1821.
- [3] 邓雄,王洪春,赵立军,等. 人脸识别活体检测研究方法综述[J]. 计算机应用研究,2020,37(9):2579-2585.
- [4] Zhang S, Liu A, Wan J, et al. CASIA-SURF: A large-scale multi-modal benchmark for face anti-spoofing[EB]. arXiv:1908.10654,2019.
- [5] Zhang P, Zou F, Wu Z, et al. FeatherNets: Convolutional neural networks as light as feather for face anti-spoofing[EB]. arXiv:1904.09290,2019.
- [6] Guo J, Zhu X, Xiao J, et al. Improving face anti-spoofing by 3D virtual synthesis[C]//2019 International Conference on Biometrics,2019:1-8.
- [7] Song L, Ma H. Face liveness detection based on texture and color features[C]//2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis,2019:418-422.
- [8] 陈放,刘晓瑞,杨明业. 基于活体检测和身份认证的人脸识别安防系统[J]. 计算机应用,2020,40(12):3666-3672.
- [9] 邓雄,王洪春. 基于深度学习和特征融合的人脸活体检测算法[J]. 计算机应用,2020,40(4):1009-1015.
- [10] Howard A, Sandler M, Chu G, et al. Searching for MobileNetV3[EB]. arXiv:1905.02244,2019.
- [11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB]. arXiv:1409.1556,2014.
- [12] He K, Zhang X, Ren S. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition,2016:770-778.
- [13] Szegedy C, Wei L, Jia Y, et al. Going deeper with convolutions[EB]. arXiv:1409.4842,2014.
- [14] Howard A, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB]. arXiv:1704.04861,2017.
- [15] 任欢,王旭光. 注意力机制综述[J]. 计算机应用,2021,41(S1):1-6.
- [16] Jaderberg M, Simonyan K, Zisserman A, et al. Spatial transformer networks[EB]. arXiv:1506.02025,2015.
- [17] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,2018:7132-7141.
- [18] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition,2020:11531-11539.
- [19] Qin Z, Zhang P, Wu F, et al. FcaNet: Frequency channel attention networks[EB]. arXiv:2012.11879,2020.
- [20] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[EB]. arXiv:1807.06521,2018.
- [21] Fu J, Liu J, Tian H, et al. Dual attention network for scene segmentation[EB]. arXiv:1809.02983,2018.
- [22] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottleneck[EB]. arXiv:1801.04381,2018.
- [23] Tan X, Yi L, Liu J, et al. Face liveness detection from a single image with sparse low rank bilinear discriminative model[C]//11th European Conference on Computer Vision; Part VI,2010:504-517.
- [24] Chingovska I, Anjos A, Marcel S. On the effectiveness of local binary patterns in face anti-spoofing[C]//IEEE International Conference of the Biometrics Special Interest Group, 2012:1-7.