

兰秋松, 杨成福, 李俊玮, 等. 人脸增强: DeepFake 检测任务中一种简单有效的数据处理方法[J]. 智能计算机与应用, 2026, 16(4): 20-27. DOI: 10.20169/j.issn.2095-2163.24062001

# 人脸增强: DeepFake 检测任务中一种简单有效的数据处理方法

兰秋松<sup>1</sup>, 杨成福<sup>1,2</sup>, 李俊玮<sup>1</sup>, 阿琴花<sup>1</sup>, 赵建龙<sup>1</sup>

(1 云南师范大学 信息学院, 昆明 650500; 2 云南省教育厅计算机视觉与智能控制技术工程研究中心, 昆明 650500)

**摘要:** DeepFake 技术尽管在游戏、社交媒体等领域有着促进意义,但也因安全和隐私等问题引发了人们的担忧。对 DeepFake 内容的精确检测与识别,已成为计算机视觉领域的一个重要研究方向,涌现出众多的深度辨伪网络模型。然而,在 DeepFake 检测任务中对输入图像进行预处理以提升模型性能方面却鲜有研究。鉴于此,本文提出了一种简单有效的人脸增强方法,即通过对输入图像的人脸部分进行提取,并将其余区域填充为黑色以形成填充图像,再将原始图像与填充图像进行加权叠加以作为检测网络的输入图像。此外,本文还设计了一个分类头,用于将检测网络提取出的特征转换为网络输出的分类。通过在 Celeb-DF-V2 上对 DeepFake 伪造内容的检测实验证明,与单纯的 EfficientNet、CapsuleNet 以及 MesoNet 这几个知名网络相比,融合了本文所提出的人脸图像增强方法后,检测网络模型的 AUC (Area Under the Curve) 值、准确率、精确率以及召回率等性能指标都得到了有效提升。本文所提的人脸增强方法对推动 DeepFake 的辨伪检测工作将产生积极意义。

**关键词:** DeepFake 检测; 数据预处理; 人脸增强; Celeb-DF-V2

中图分类号: TP183

文献标志码: A

文章编号: 2095-2163(2026)04-0020-08

## Face enhancement: A simple and effective data processing method for DeepFake detection

LAN Qiusong<sup>1</sup>, YANG Chengfu<sup>1,2</sup>, LI Junwei<sup>1</sup>, A Qinhua<sup>1</sup>, ZHAO Jianlong<sup>1</sup>

(1 School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China ;

2 Computer Vision and Intelligent Control Technology Engineering Research Center, Yunnan Provincial Department of Education, Kunming 650500, China)

**Abstract:** Although DeepFake technology has promoted significant advancements in areas such as gaming and social media, concerns regarding security and privacy have also arisen. Accurate detection and identification of DeepFake have become important research directions in the field of computer vision, leading to the emergence of numerous deep forgery detection network models. However, there has been little research on preprocessing input images to improve model's performance in DeepFake detection. In light of this, this paper proposes a simple and effective facial enhancement method, wherein the facial part of the input image is extracted, and the remaining areas are filled with black to create a padded image. The original image is then weighted and overlaid with the padded image to serve as the input image for the detection network. Additionally, a classification head is designed to transform the features extracted by the detection network into the network's output classification. Experimental results on Celeb-DF-V2 demonstrate that, compared to well-known networks such as EfficientNet, CapsuleNet, and MesoNet, the detection network model's performance metrics, including AUC (Area Under the Curve), accuracy, precision, and recall, are effectively improved when incorporating the facial image enhancement method proposed in this paper. The facial enhancement method proposed in this paper will have a positive impact on advancing DeepFake forgery detection work.

**Key words:** DeepFake detection; data preprocessing; Face Enhancement; Celeb-DF-V2

## 0 引言

DeepFake 技术是一种图像生成技术,是通过使

用深度学习模型(如生成对抗网络 GAN (Generative Adversarial Network)<sup>[1]</sup>等来合成或修改图像、音频及视频内容,以创建看起来很真实的虚假内容。近

**基金项目:** 云南省基础研究面上项目(202301AT070065); 云南省教育厅项目(2023J0208)。

**作者简介:** 兰秋松(1999—),男,硕士,主要研究方向:深度学习,伪造图像识别;李俊玮(1998—),男,硕士,主要研究方向:深度学习;阿琴花(1999—),女,硕士,主要研究方向:深度学习,天线;赵建龙(2001—),男,硕士研究生,主要研究方向:图像辨伪。

**通信作者:** 杨成福(1986—),男,博士,讲师,硕士生导师,主要研究方向:信息超材料,人工智能,计算机视觉。Email: yangchengfu@ynnu.edu.cn。

收稿日期: 2024-06-20

年来,DeepFake 技术发展迅速,被广泛应用于游戏、动画电影以及社交媒体等领域。然而,由于该技术存在新闻及色情视频恶意造假、网络诈骗等事关安全与隐私的问题而备受争议。因此,如何对 DeepFake 内容进行检测、识别与辨伪具有重要研究价值。

当前,DeepFake 内容的生成技术已具备较高的水平<sup>[2-7]</sup>,利用传统图像辨伪方法<sup>[8-12]</sup>已经很难对其进行高效率、高准确率的检测。为此,许多研究人员开始致力于使用深度学习技术来对 DeepFake 内容进行真伪检测。由于卷积神经网络(Convolutional Neural Networks, CNN)具备强大的特征提取能力,对图像进行检测及辨伪的研究工作也较多集中于多层 CNN 技术的使用<sup>[13-22]</sup>。近年来,随着 Transformer (尤其是 Vision Transformer (ViT)) 的提出,又有不少研究者将目标放在使用 ViT 模型来解决 DeepFake 伪造检测的问题<sup>[23-25]</sup>。

然而在现有大多数研究中,研究者均是将原图或人脸图像 Resize 到特定尺寸之后直接输入网络,未对输入数据进行过针对性预处理。事实上,在对 DeepFake 图像或视频检测与辨伪的过程中,通过对关键特征区域进行针对性增强处理,可提升该区域对最终检测结果的贡献度,从而提高整体识别的准确性。现有 DeepFake 检测网络主要采用两种典型的输入处理策略,一种是直接将视频或数据集中提供的原始图像作为输入,另一种是以人脸区域作为输入进行训练与检测。这两种策略各有优劣:前者能够保留图像的完整背景信息,但因缺乏明确的关

注焦点,难以有效提升模型性能;后者虽可将计算资源集中在人脸区域,但会因丢失背景信息而削弱模型的泛化性能。

针对上述问题,本文融合两种策略的优势,提出一种简洁高效的“人脸增强”方法,该方法在保留原始图像信息的前提下,对人脸区域的关键特征进行增强处理,生成融合图像并作为检测网络的输入,供后续网络模型进行特征提取与辨伪检测。所提方法的有效性在 Celeb-DF-V2 数据集上进行了充分验证。

本文的主要贡献包括:(1)提出了一种简单且高效的图像融合方法,用于实现核心检测区域的图像增强;(2)设计了一个分类头,用于将检测网络提取出的特征转换为网络输出的分类;(3)构建了多种性能优异的 DeepFake 检测网络,所提网络在 AUC (Area Under the Curve)值、准确性、精确度及召回率等指标上均表现出色。

## 1 模型与方法

### 1.1 人脸图像增强方法

前文提到,未经预处理的原始图像或视频帧,虽然保留了完整的信息,但未对人脸区域和背景区域进行区分,若直接将其作为模型训练的输入会导致模型训练耗时长、性能表现不佳等问题。而裁剪掉图像的背景部分,仅保留人脸区域,虽能使模型聚焦于人脸区域,但会因丢失背景信息而影响模型的检测准确性与泛化能力。基于此,本文提出了一种“人脸增强”方法,其具体实现流程如图 1 所示。

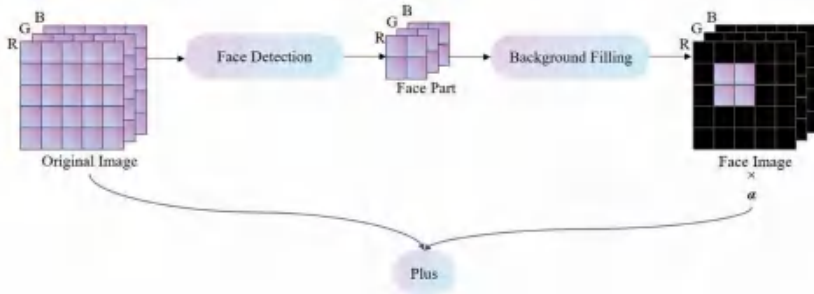


图 1 人脸增强方法

Fig. 1 Face enhancement method

具体地,首先从原始图像中提取出人脸部分,并且将人脸之外的区域填充 0 像素以保证人脸图像和原始图像的尺寸一致。其次,将得到的人脸图像乘上一个可学习的系数  $\alpha$ ,并将其与原始图像进行叠加融合。该人脸图像增强方法可由如下公式表征:

$$E = O + \alpha \cdot F \quad (1)$$

$$F = F_i(D(O)) \quad (2)$$

其中,  $\alpha$  是一个可学习的参数;  $E$  代表增强后的图像;  $O$  代表原始图像;  $F$  代表经过处理之后的人脸区域图像;  $F_i$  代表填充全 0 像素的操作;  $D$  代表人脸区域检测的操作。

经过前述操作,所得到的新图像会在人脸部分予

以增强,而背景区域则保持不变,从而实现增强图像中的人脸区域的效果。另外,由于可学习参数  $\alpha$  的存

在,这种人脸增强方法也可被视作一层简单的神经网络。人脸增强方法的实际效果如图 2 所示。

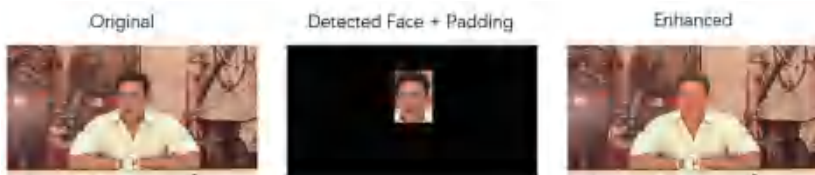


图 2 实际效果

Fig. 2 Actual effect

图 2 左侧为原图,中间为提取后的人脸部分以及将人脸之外的区域填充为黑色,右侧为将前面二者融合之后的效果。

## 1.2 Deepfake 检测网络

本文采用了 3 个在 Deepfake 检测领域比较知名的神经网络方法,将其与本文所提出的人脸增强方法进行有机融合,共同完成伪造检测任务,同时将融合后的检测结果与这些方法单独使用时的结果进行比对分析,以此验证本文所提方法的有效性。

### 1.2.1 MesoNet+Face Enhancement

MesoNet 是 Darius Afchar 于 2018 年提出的一种基于中层语义的 DeepFake 伪造检测网络架构<sup>[26]</sup>。本文采用 MesoInception-4 作为特征提取架构,其核心思想是使用两个 Inception 模块来替换 Meso-4 模

型的前两个卷积层。Inception 模块由 4 组并行的卷积操作以及最后的拼接操作组成,这 4 组并行的卷积操作分别为: $a$  个  $1 \times 1$  卷积; $b$  个  $1 \times 1$  卷积经过 ReLU 函数激活之后再经过  $b$  个  $3 \times 3$  卷积; $c$  个  $1 \times 1$  卷积经过 ReLU 函数激活之后再经过  $c$  个  $3 \times 3$  空洞卷积; $d$  个  $1 \times 1$  卷积经过 ReLU 函数激活之后再经过  $d$  个  $3 \times 3$  空洞卷积。其中  $a, b, c, d$ , 为可以自行设计的卷积数量参数<sup>[25]</sup>。本文将 MesoInception-4 与所提的 Face Enhancement 方法相结合来实现检测任务。

为了验证人脸增强方法的有效性,本文沿用 MesoInception-4 的网络结构进行检测任务。MesoNet 与 Face Enhancement 相结合的整体流程如图 3 所示。

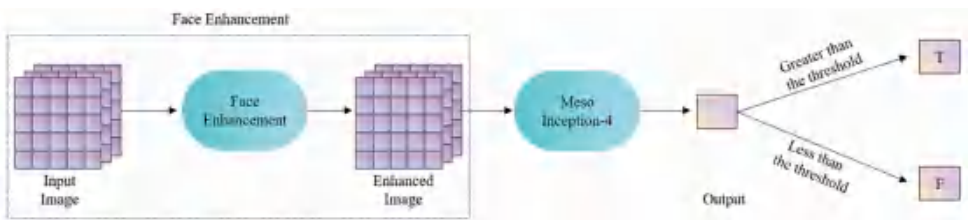


图 3 MesoInception-4+Face Enhancement 整体流程

Fig. 3 The overall process of MesoInception-4+Face Enhancement

### 1.2.2 DFDT+Face Enhancement

DFDT 是 Khormali 等<sup>[27]</sup>于 2022 年提出的一种基于视觉 Transformer 的深度学习网络架构,旨在克服传统卷积神经网络(CNN)在提取全局特征方面的局限性。

DFDT 框架主要由补丁提取与嵌入模块(Patch Extraction & Embedding)、多流 Transformer 块(Multi-stream Transformer Block)、基于注意力的 Patch 选择模块(Attention-based Patch Selection)和多尺度分类器(Multi-scale Classifier)4 个部分组成。其中,补丁提取与嵌入模块将图像转换为  $D$  维的较小 Patch 嵌入序列。多流 Transformer 块具有包含两个分支,即每个 Patch 更小的低级别 Patch 分支和每个 Patch 更大的高级别 Patch 分支,该模块可从细小的 Patch

当中学习更多的局部特征,从粗大的 Patch 当中学习全局特征。基于注意力的 Patch 选择模块则通过注意力权重迫使模型在训练和推理阶段更加关注携带判别性信息的 Patch,并忽略信息量较少或所携带的信息对于最终的分类贡献较少的 Patch。多尺度分类器对不同尺度的信息进行整合并进行最终分类,其通过对不同尺度下分类器的预测进行平均操作得出最终的决策<sup>[27]</sup>。

本文中,仍然沿用这种网络架构,并在将图像拆分为若干个 Patch 之前对图像的人脸部分进行增强,其具体的网络架构如图 4 所示。

### 1.2.3 EfficientNet+Face Enhancement

EfficientNet 最初由 Tan 等<sup>[28]</sup>在 2019 年提出。简单来说,Tan 等<sup>[28]</sup>使用  $d, r, w$  这 3 个参数来分别

表征缩放深度、缩放分辨率和缩放宽度,并提出一个混合缩放方法,即使用一个混合因子 $\phi$ 对 $d, r, w$ 这三个参数进行统一的缩放,如下式:

$$d = \alpha^\phi, w = \beta^\phi, r = \gamma^\phi \quad (3)$$

随后,在 EfficientNet-B0 网络架构的基础上,进一步使用 NAS 技术对  $\alpha, \beta, \gamma$  这三个技术进行搜索。其搜索的大致流程为:首先固定  $\phi = 1$ , 得出基于

EfficientNet-B0 的最佳  $\alpha, \beta, \gamma$  值。然后将  $\alpha, \beta, \gamma$  的值固定,再在 EfficientNet-B0 的基础上使用不同的  $\phi$  分别得到 EfficientNet-B1 至 EfficientNet-B7 网络<sup>[28]</sup>。

在本文中,选择预训练好的 EfficientNet-B7 作为 Base Model,并将其与所提出的 Face Enhancement 方法相结合,并通过分类器(Classifier)完成最终的任务,具体的网络架构如图 5 所示。

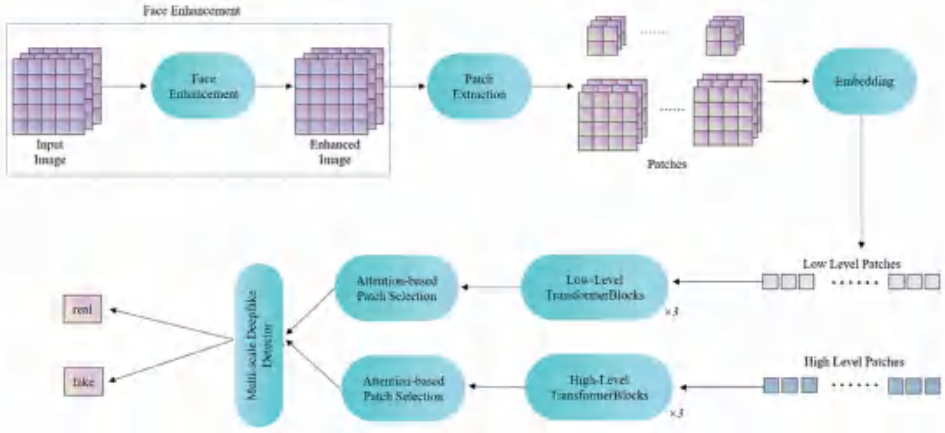


图 4 DFDT+Face Enhancement 整体流程

Fig. 4 The overall process of DFDT+Face Enhancement

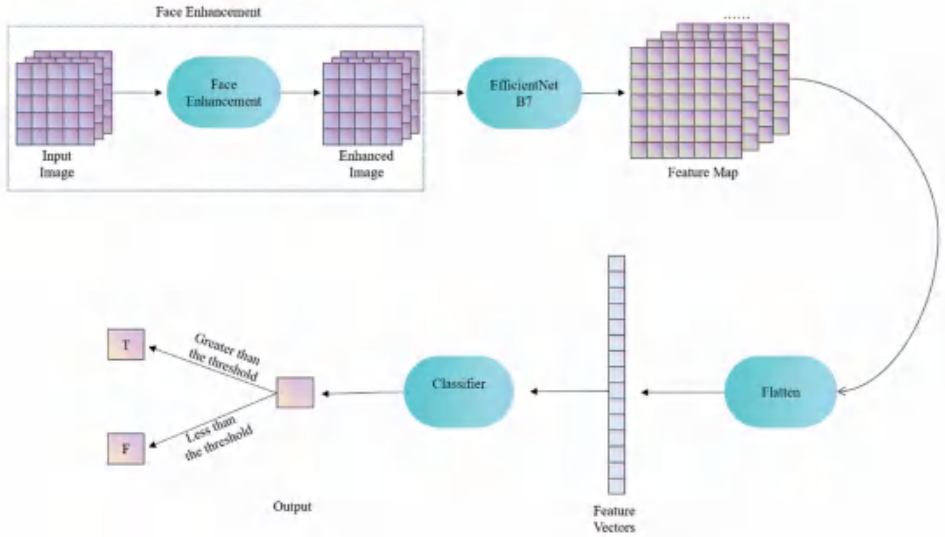


图 5 EfficientNet-B7+Face Enhancement 整体流程

Fig. 5 The overall process of EfficientNet-B7+Face Enhancement

## 2 实验结果及分析

### 2.1 实验细节

本文首先使用 DLIB 库<sup>[29]</sup>对图像的人脸部分进行提取,并将人脸之外的区域填充成黑色像素以保持人脸图像和原始图像大小的一致。接着,对于 Face Enhancement 的人脸图像和原始图像进行一样的 Resize、Normalize 等预处理。对于所有的实验,本文

都选用 SGD 优化器并将其初始学习率设置为 0.1,再使用 ReduceLROnPlateau 方法,根据模型在验证集上的性能来进行学习率衰减。在每个 epoch 中,本文对训练集进行划分,将其中的 80%用于训练,20%用于验证以检验是否需要早停或更改学习率。

### 2.2 实验数据集

本文使用 Celeb-DF-V2 进行所有网络的训练和测试,该网络包含 590 个真实视频以及 5 639 个

伪造视频,这些视频质量与在线传播的视频非常接近<sup>[30]</sup>。为了进行图像辨伪检测,本文对 Celeb-DF-V2 所划分的训练集和测试集视频按 1 帧/s 的帧速率进行抽帧处理,得到 71 678 张伪造图像和 71 693 张真实图像,其中有 4 340 张伪造图像和 4 340 张真实图像用于测试,其余均作为训练集使用。

### 2.3 实验结果评估

本文提出人脸增强方法的目的是,为了在保留图像背景信息的同时,让网络更加关注图像中的人脸部分,因此本文选取提升最大的 MesoNet 网络来对模型的最后一层卷积层进行可视化以直观地分析人脸增强对于模型关注区域的影响。此外,为了探究本文所提的人脸增强方法有助于提升图像辨伪能

力,本文选用了 MesoNet、DFDT 以及 EfficientNet-B7 网络作为基准模型,开展了对照实验。对照组中,本文将原始图像直接加载至这 3 个网络,以计算出其对应的 ROC 曲线、准确率、精度、召回率以及混淆举证等;实验组中,本文将经过 Face Enhancement 的图像分别作为前述 3 个网络的输入,并同样计算了其相应的性能指标。

首先,本文选取通过人脸增强方法获得最大提升的 MesoNet 网络来对其进行可视化分析,选用网络的最后一层卷积层来对其进行可视化,这是因为其包含了网络学习到的最高级别的抽象特征,通过对其进行可视化可以清楚地看到模型在做出决策时关注的具体区域。可视化结果如图 6 所示。



图 6 原始图像与人脸增强图像的 MesoNet Grad-CAM 可视化结果

Fig. 6 Comparison of Grad-CAM visualization results of MesoNet on original images and Face-Enhancement images

图 6 第一行左边的图像为一张示例图。MesoNet 未能正确识别,但经过人脸增强之后的 MesoNet 准确识别了该示例图,右边是对网络的最后一层卷积层进行可视化得到的模型关注区域。第二行右边是对经过人脸增强之后的网络的最后一层卷积层进行可视化得到的模型关注区域。其中可以明显地看出未经人脸增强之前的 MesoNet 模型关注的区域比较杂乱,且背景部分的占比较多,而经过人脸增强之后的 MesoNet 模型,仍然保留了对背景区域的关注,但是模型明显地将更多注意聚集于人脸之中。这说明人脸增强方法能够迫使模型更关注于人脸区域而不丢失背景部分携带的信息。

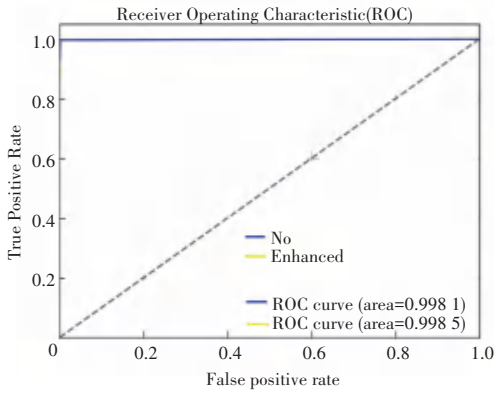
此外,本文选用 EfficientNet 和 MesoNet 作为网络模型,分别绘制了采用人脸增强方法和未使用人脸增强时各自的 ROC 曲线 (Receiver Operating Characteristic),如图 7 所示。其中,图 7 (a) 为 EfficientNet 的 ROC 曲线,图 7 (b) 为 MesoNet 的

ROC 曲线。从图中可以看出,对于 EfficientNet 网络,在使用和未使用人脸增强技术时的 ROC 曲线几乎一致,呈一条直线。这是因为原始的 EfficientNet 加上本文所设计的 Classifier 之后,本身的 AUC 已经达到了 99.81%。尽管叠加上人脸增强及 Classifier 之后,AUC 进一步得到提升,达到 99.85%,但二者非常接近,且趋近于 1,这说明网络已具有很高的性能。然而,对于 MesoNet 网络,在使用本文所提的人脸增强技术以后,其 AUC 值得到了明显提升,由原来的 94.14% 提升至 96.85%。这说明,使用人脸增强方法有助于提升网络的 AUC 性能,且对原有性能较低的网络提升效果更明显。

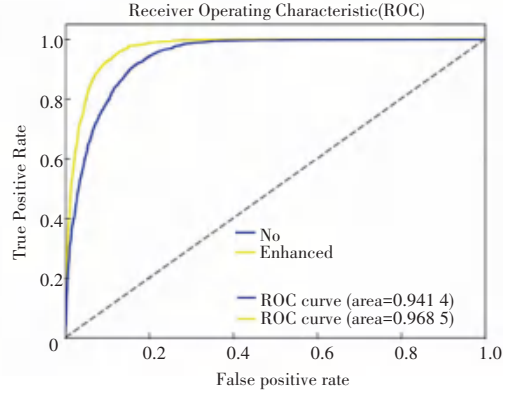
随后,本文将 EfficientNet、DFDT 以及 MesoNet 作为基准模型,分别对其叠加本文所提出的人脸增强技术,并在阈值为 0.5 的条件下对叠加前后网络准确率 (ACC)、精确率以及召回率进行比对分析 (见表 1)。从表 1 可以看出,在将原有的伪造内容

检测网络与本文提出的人脸增强方法相结合之后,模型的准确率获得较好的提升,其中 EfficientNet 的准确率提升至 99.60%,较原来提升了 0.12%,DFDT 的准确率提升至 99.52%,较原来提升了 0.21%,MesoNet 的准确率达到 91.38%,较原来提升了 4.19%。此外,从精确率上来看,使用人脸增强方法后, EfficientNet、CapsuleNet 以及 MesoNet 的精确率分别提高到了 99.56%、97.88%和 88.74%,较原有网

络分别提升了 0.2%、2.32%和 5.3%。从召回率来看,使用了人脸增强方法后, CapsuleNet 网络略有下降(较原来下降 0.61%)外, EfficientNet 和 MesoNet 都有所提升,尤其是 MesoNet,提升至 94.79%(较原来提升了 2%)。由此可见,使用本文所提的人脸增强方法,可以有效提升辨伪检测网络的准确率、精确率以及召回率,优化网络性能。



(a) EfficientNet 的 ROC 曲线



(b) MesoNet 的 ROC 曲线

图 7 EfficientNet 组和 MesoNet 组的 ROC 曲线

Fig. 7 ROC curves of EfficientNet and the MesoNet group

表 1 各方法在阈值为 0.5 时主要性能数据对比

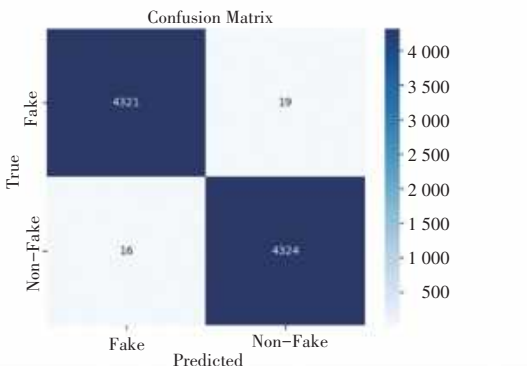
Table 1 Comparison of main performance metrics for each method at a threshold of 0.5

Method	Score		
	ACC	Precision	Recall
EfficientNet+Face Enhancement	0.996 0	0.995 6	0.996 3
Original EfficientNet	0.994 9	0.993 6	0.996 3
DFDT+Face Enhancement	0.995 2	0.993 6	0.996 8
Original DFDT	0.993 1	0.989 9	0.996 3
MesoNet+Face Enhancement	0.913 8	0.887 4	0.947 9
Original MesoNet	0.871 9	0.834 4	0.927 9

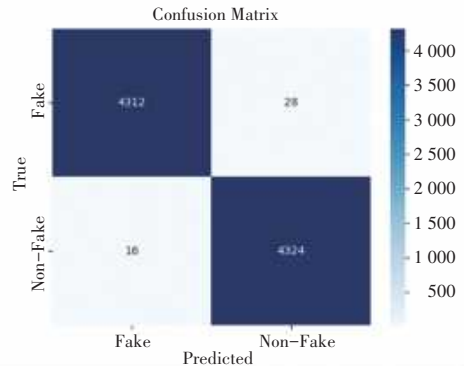
络的正类鉴别能力和负类鉴别能力进行的影响,本文仍然以 EfficientNet、CapsuleNet 以及 MesoNet 作为基准模型,分别考察 3 个网络在使用了人脸增强方法前后的混淆矩阵。

EfficientNet 网络在叠加人脸增强算法后的混淆矩阵如图 8 所示,其中,图 8(a) 为 EfficientNet+ Face Enhancement 的混淆矩阵,图 8(b) 为 Original EfficientNet 的混淆矩阵。从图中可以看出,在与本文所提的人脸增强算法进行结合之后, EfficientNet 对于真实样本的分类能力保持不变,但是对伪造样本的分类能力有所提升。

为了更直观地揭示本文所提人脸增强算法对网



(a) EfficientNet+Face Enhancement



(b) Original EfficientNet

图 8 EfficientNet+Face Enhancement and Original EfficientNet 混淆矩阵

Fig. 8 Confusion matrix for EfficientNet+Face Enhancement and Original EfficientNet

DFDT 在使用人脸增强算法前后的混淆矩阵如图 9 所示,其中,图 9(a)为 DFDT+ Face Enhancement 的混淆矩阵,图 9(b)是 Original DFDT 的混淆矩阵。从图 9 中可以看出,在与本文所提的人脸增强方法相结合之后,DFDT 对于真实样本以及伪造样本的分类能力均有所提升,且对伪造样本的分类能力有了较大的提升。同时,从图 9(b)中可以看出: Original DFDT 对于负类样本的鉴别能力明显低于对正类样本的鉴别能力,但与人脸增强方法结合后,尽管该网络对二者的鉴别能力仍然存在差距,但差距得到大幅缩减,鉴别能力更趋向于平衡。

MesoNet 网络在使用人脸增强方法前后的混淆

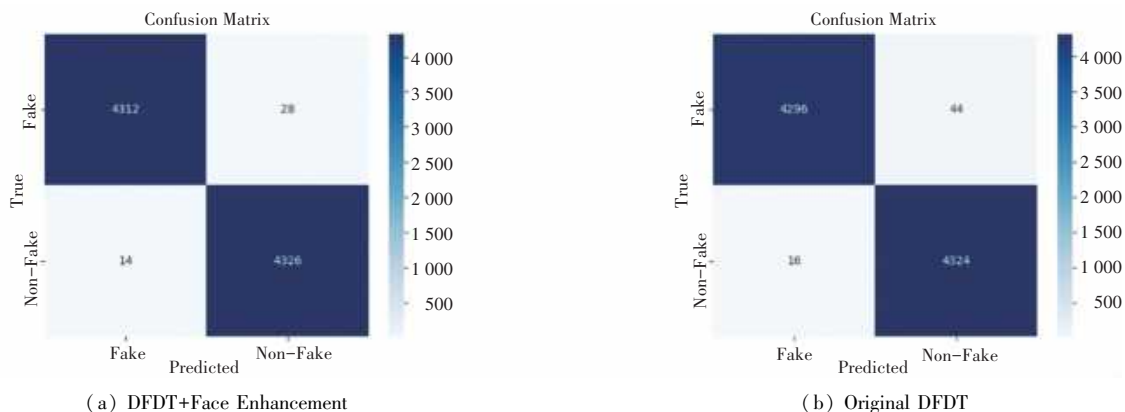


图 9 DFDT+Face Enhancement and Original CapsuleNet 混淆矩阵

Fig. 9 Confusion matrix for DFDT+Face Enhancement and Original CapsuleNet

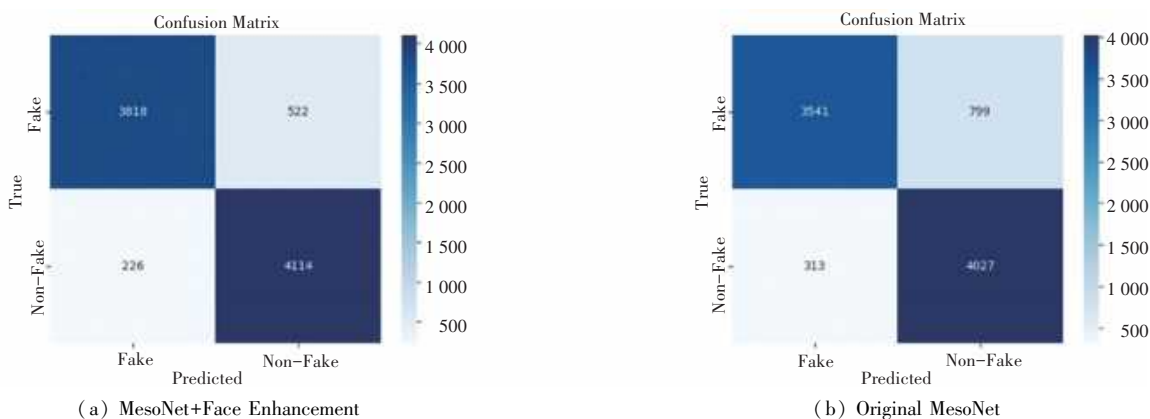


图 10 MesoNet+Face Enhancement and Original MesoNet 混淆矩阵

Fig. 10 Confusion matrix for MesoNet+Face Enhancement and Original MesoNet

由此可见,本文所提的人脸增强方法可有效提升 EfficientNet、CapsuleNet 以及 MesoNet 等的分类鉴别能力和网络性能,为推进 DeepFake 的辨伪检测工作提供技术支撑。

### 3 结束语

本文提出了一种简单有效的人脸增强方法,用

矩阵如图 10 所示。其中,图 10(a)是 MesoNet+ Face Enhancement 的混淆矩阵,图 10(b)是 Original MesoNet 的混淆矩阵。从图中可以看出,在与人脸增强方法融合后,MesoNet 对于真实样本的分类能力以及对伪造样本的分类能力都有一定程度的提升,尤其是对于伪造样本的分类能力提升更为显著。同样地,从图 10(b)中可以直观看出: Original MesoNet 对于负类样本的鉴别能力明显低于对正类样本的鉴别能力,但与本文所提人脸增强方法相结合之后,该网络对二者的鉴别能力尽管存在差距,但这种差异明显变得更加平衡。

于对输入到鉴别网络的图像进行预处理。通过将该方法应用于 EfficientNet、CapsuleNet 以及 MesoNet 等 Deepfake 伪造内容检测网络,结果显示该方法对模型的准确率、精度以及召回率等性能指标都取得了较好的提升,并对平衡模型的真实样本检测能力和伪造样本检测能力产生积极影响。本文所提方法对 DeepFake 内容的高效、高精度辨伪检测提供一定的

参考价值。

## 参考文献

- [1] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the Advances in Neural Information Processing Systems. NeurIPS, 2014; 2672-2680.
- [2] MIRZA M, OSINDERO S. Conditional generative adversarial nets [J]. arXiv preprint arXiv, 1411.1784, 2014.
- [3] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv, 1511.06434, 2015.
- [4] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks [C]//Proceedings of the International Conference on Machine Learning. PMLR, 2017;214-223.
- [5] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ;IEEE, 2017;2223-2232.
- [6] BROCK A, DONAHUE J, SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis [J]. arXiv preprint arXiv, 1809.11096, 2018.
- [7] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ;IEEE, 2019;4401-4410.
- [8] KORSHUNOV P, MARCEL S. Deepfakes: A new threat to face recognition? Assessment and detection[J]. arXiv preprint arXiv, 1812.08685, 2018.
- [9] MATERN F, RIESS C, STAMMINGER M. Exploiting visual artifacts to expose deepfakes and face manipulations [C]//Proceedings of the 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW). Piscataway, NJ;IEEE, 2019; 83-92.
- [10] YANG X, LI Y, LYU S. Exposing deep fakes using inconsistent head poses [C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ;IEEE, 2019;8261-8265.
- [11] JUNG T, KIM S, KIM K. Deepvision: Deepfakes detection using human eye blinking pattern [J]. IEEE Access, 2020, 8: 83144-83154.
- [12] AGARWAL S, FARID H, GU Y, et al. Protecting world leaders against deep fakes [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway, NJ;IEEE, 2019;38.
- [13] PATEL Y, TANWAR S, BHATTACHARYA P, et al. An improved dense cnn architecture for deepfake image detection[J]. IEEE Access, 2023, 11: 22081-22095.
- [14] ROY R, JOSHI I, DAS A, et al. 3D CNN architectures and attention mechanisms for deepfake detection [M]. Handbook of Digital Face Manipulation and Detection; From DeepFakes to Morphing Attacks. Cham:Springer, 2022;213-234.
- [15] IKRAM S T, CHAMBIAL S, SOOD D. A performance enhancement of deepfake video detection through the use of a hybrid CNN deep learning model [J]. International Journal of Electrical and Computer Engineering Systems, 2023, 14(2): 169-178.
- [16] RANA M S, NOBI M N, MURALI B, et al. Deepfake detection: A systematic literature review [J]. IEEE Access, 2022, 10: 25494-25513.
- [17] AHMED S R, SONUÇ E, AHMED M R, et al. Analysis survey on deepfake detection and recognition with convolutional neural networks[C]//Proceedings of the 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). Piscataway, NJ;IEEE, 2022; 1-7.
- [18] YU C M, CHEN K C, CHANG C T, et al. SegNet: A network for detecting deepfake facial videos [J]. Multimedia Systems, 2022, 28(3): 793-814.
- [19] ZHANG T. Deepfake generation and detection, A survey [J]. Multimedia Tools and Applications, 2022, 81(5): 6259-6276.
- [20] SAIKIA P, DHOLARIA D, YADAV P, et al. A hybrid CNN-LSTM model for video deepfake detection by leveraging optical flow features [C]//Proceedings of 2022 International Joint Conference on Neural Networks (IJCNN). Piscataway, NJ: IEEE, 2022; 1-7.
- [21] AHMED S R A, SONUÇ E. Retracted article: Deepfake detection using rationale-augmented convolutional neural network [J]. Applied Nanoscience, 2023, 13(2): 1485-1493.
- [22] NGUYEN H H, YAMAGISHI J, ECHIZEN I. Use of a capsule network to detect fake images and videos [J]. arXiv preprint arXiv, 1910.12467, 2019.
- [23] HEO Y J, YEO W H, KIM B G. Deepfake detection algorithm based on improved vision transformer [J]. Applied Intelligence, 2023, 53(7): 7512-7527.
- [24] WANG T, CHENG H, CHOW K P, et al. Deep convolutional pooling transformer for deepfake detection [J]. ACM Transactions on Multimedia Computing, Communications and Applications, 2023, 19(6): 1-20.
- [25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv, 1409.1556, 2014.
- [26] AFCHAR D, NOZICK V, YAMAGISHI J, et al. Mesonet: A compact facial video forgery detection network [C]//Proceedings of 2018 IEEE International Workshop on Information Forensics and Security (WIFS). Piscataway, NJ;IEEE, 2018; 1-7.
- [27] KHORMALI A, YUAN J S. DFDT: An end-to-end deepfake detection framework using vision transformer [J]. Applied Sciences, 2022, 12(6): 2953.
- [28] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks [C]//Proceedings of the International Conference on Machine Learning. IMLS, 2019: 6105-6114.
- [29] KING D E. Dlib-ml: A machine learning toolkit [J]. The Journal of Machine Learning Research, 2009, 10: 1755-1758.
- [30] LI Y, YANG X, SUN P, et al. Celeb-df: A large-scale challenging dataset for deepfake forensics [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ;IEEE, 2020; 3207-3216