

DOI:10.19651/j.cnki.emt.2210785

超轻量人脸关键点检测算法

朱望纯^{1,2} 张博^{1,2}

(1. 桂林电子科技大学电子工程与自动化学院 桂林 541004; 2. 智能综合自动化广西高校重点实验室 桂林 541004)

摘要: 随着深度学习网络研究的深入和网络模型精度的提高,网络层数及深度在逐渐增加,导致计算量增大。同时,基于深度学习模型人脸关键点检测在嵌入式设备上部署的需求,轻量化、高效和准确的网络模型成为研究关键。因此,本文设计了一个基于 Ghost Model 块和 Ghost Bottleneck 架构的超轻量型人脸关键点检测算法,在确保网络精度的同时,尽可能减小网络模型大小,降低计算量。在网络宽度因子为 1X 的情况下,与现有表现最好的轻量化网络模型 PFLD 1X 相比,归一化平均误差降低了 7%,参数量减小了 36%;在宽度因子为 0.25X 的情况下,本论文提出的网络模型大小仅 420 KB,归一化平均误差降低了 6.6%,参数量减小了 25%。

关键词: 超轻量化;深度学习;人脸关键点

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4

Ultra-lightweight facial landmark detector

Zhu Wangchun^{1,2} Zhang Bo^{1,2}(1. School of Computer Science and Engineering, Guilin University of Electronic Technology, Guilin 541004, China;
2. Key Laboratory of Intelligence Integrated, Automation in Guanxi Universities, Guilin 541004, China)

Abstract: The number of layers and the depth of the network are gradually increasing as the research on deep learning networks deepens and the accuracy of the network model improves, leading to an increase in computation. The lightweight, efficient and accurate network model becomes the key to research based on the need of deep learning model facial landmark detection for deployment on embedded devices. Therefore, an ultra-lightweight facial landmark detection network based on Ghost Model and Ghost Bottleneck is designed in this thesis to ensure the network accuracy while minimizing the network model size and reducing the computational effort. With a network width factor of 1X, the normalized mean error is reduced by 7% and the number of parameters is reduced by 36% compared to the best performing lightweight network model PFLD 1X; with a width factor of 0.25X, the proposed network model is only 420 KB in size, and the normalized mean error is reduced by 6.6% and the number of parameters is reduced by the average normalized error is reduced by 6.6% and the number of parameters is reduced by 25%.

Keywords: ultra-lightweight; deep learning; facial landmark

0 引言

人脸关键点检测又称脸部关键点定位,其中关键点代表了人脸眼睛、眉毛、鼻子、嘴巴及人脸轮廓这些关键特征,是面部图像处理的关键元素之一^[1]。该检测应用广泛,涉及到情感识别、虚拟人脸动画、驾驶员状态检测、人脸识别等领域。随着深度学习的发展,神经网络模型的精度得到了提升,随之而来的是网络模型深度越来越深,计算量越来越大,对设备性能的要求也越来越高。想要把深度学习网络模型部署在嵌入式设备上还具有很大的挑战。

目前,人脸及解剖结构的关键点检测一直是计算机视

觉领域的研究关键^[2-4]。在深度学习方法兴起之前,人脸关键点检测算法根据它们对人脸外观和形状进行建模主要分为主动形状模型(active shape models, ASM)^[5]、主动外观模型(active appearance models, AAM)^[6]、约束局部模型(constrained local models, CLM)^[7]。但这些方法大都存在对环境干扰(照明、正面、模糊以及遮挡)的鲁棒性较差、模型复杂度高、计算量大等局限。随着深度学习研究的发展,基于深度学习的人脸关键点检测得到较大提升。其中,多任务学习网络(tasks-constrained deep convolutional network, TCDCN)^[8]将多任务学习(multi-task learning, MTL)与 CNN 结合进行人脸关键点检测。由于 TCDCN

收稿日期:2022-07-22

多任务训练的性质,对性别、表情、姿势以及是否佩戴眼镜有不同的特点和收敛速度,在实际应用中难以实现。

随着深度学习网络模型在嵌入式设备上部署的需要,对网络模型的轻量化要求越来越高,在保证模型精度的前提下,压缩模型大小,降低参数量提高检测速度成为研究重点。Wu等^[9]提出了一种边界感知的人脸对齐算法,即LAB,先利用人脸热图捕捉人脸边界线,然后从边界线获得人脸关键点位置。2018年Sandler等^[10]提出了轻量级骨干网络MobileNet V2,利用深度可分离卷积减少网络参数量^[11],但网络本身特征提取能力不足,参数少,识别精度低。Guo等^[12]应用轻量化MobileNet V2作为主干网络进行人脸关键点定位,然后加入辅助网络进行人脸欧拉角计算,来提高检测精度,这也被称为实用人脸关键点检测器(practical facial landmark detector, PFLD)。尽管现有的深度学习算法在人脸关键点检测研究上取得了很大进步,但仍然存在着巨大的改进空间,特别是在保证网络模型精度和效率的前提下使网络模型轻量化,减少计算量,提高网络运行速度。

针对现有轻量化网络模型无法兼顾检测精度和检测速度的问题,本文设计了一个超轻量化人脸关键点检测算法,该网络模型在不同的光照、表情、部分遮挡等复杂情况下,可快速、高效识别人脸关键点位置,其在网络宽度为0.25X下模型大小仅420 kB,为计算能力和内存资源有限的嵌入式设备部署深度学习网络提供研究基础。文章以新型轻量化网络模块Ghost Module和Ghost Bottlenecks作为主干网络来保证算法速度和模型紧凑性,并在主干网络后加入多尺度全连接层来提高模型检测精度。在网络宽度因子为1X的情况下,与现有表现最好的轻量化网络模型PFLD 1X相比,本文提出的网络模型在模型大小上减小了4倍,归一化平均误差(normalized mean error, NME)降低了7%,参数量(Params)减小了36%;在宽度因子为0.25X的情况下,网络模型大小仅420 kB,比PFLD 0.25X(2.1 MB)减小了5倍,归一化平均误差降低了6.6%,参数量减小了25%。

1 研究方法和网络结构

主网络是提高处理速度和减小模型大小的关键,使其快速和紧凑至关重要。随着在嵌入式设备上部署神经网络模型的需要,近几年已经研究并提出了ShuffleNet^[13]、MobileNet^[14]、GhostNet^[15]等加速网络。由于GhostNet先通过普通卷积得到通道较少的特征图,然后再通过分层卷积得到Ghost特征图,最后将两个特征图进行拼接得到最终的特征图。在不改变输出特征图大小的情况下,GhostNet将传统卷积操作分成两部分运行,在保证测量精度的同时,减少计算量、提升运行速度;所以本文所有的卷积操作使用GhostNet中的Ghost Model来替换传统卷积操作。

1.1 卷积网络

深度卷积神经网络^[16-17]通常由大量卷积组成,这将导致计算成本增加。最近,MobileNet和ShuffleNet等网络架构通过引入深度卷积或混合卷积操作,以使用较小的卷积滤波器(浮点数操作)来构建高效CNNs,但剩余的 1×1 卷积层仍占用大量内存和计算量(floating point operations, FLOPs)。如图1所示,在传统卷积操作^[18]中,假如给定输入数据 $\mathbf{X} \in \mathbb{R}^{c \times h \times \omega}$,其中 c 是输入通道数, h 和 ω 是输入数据的高度和宽度,任意卷积层的操作于产生 n 个特征映射可表示为:

$$\mathbf{Y} = \mathbf{X} * \mathbf{f} + b \quad (1)$$

其中, $*$ 是卷积操作, b 是偏置项, $\mathbf{Y} \in \mathbb{R}^{h' \times \omega' \times n}$ 是具有 n 个通道的输出特征图, $\mathbf{f} \in \mathbb{R}^{c \times k \times k \times n}$ 是卷积滤波器。此外, h' 和 ω' 是输出数据的高度和宽度, $k \times k$ 是卷积核大小。卷积过程的计算量FLOPs如下所示:

$$\text{FLOPs} = n \cdot h' \cdot \omega' \cdot c \cdot k \cdot k \quad (2)$$

由于滤波器数 n 和通道数 c 通常非常大,因此计算量非常庞大。

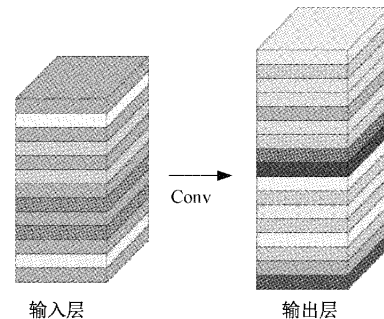


图1 传统卷积

Ghost模块的原理为其中 m 个特征图使用基本卷积完成,然后使用线性变换得到 s 个幻影特征图,最后把 m 个特征图和 s 个幻影特征图拼接成最终特征图,如图2所示。

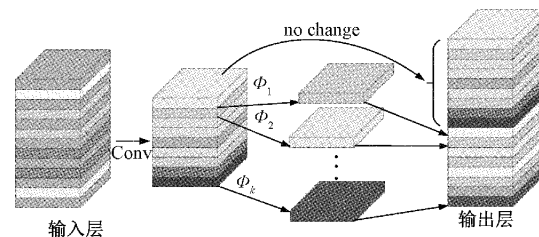


图2 Ghost Model 卷积操作

根据式(1)已知输入和输出特征图维度,要优化的参数是 f 和 b 。具体如下,基本卷积生成的特征图 $\mathbf{Y}' \in \mathbb{R}^{h' \times \omega' \times m}$

$$\mathbf{Y}' = \mathbf{X} * \mathbf{f}' \quad (3)$$

其中, $\mathbf{f}' \in \mathbb{R}^{c \times k \times k \times m}$, $m \leq n$

为了保证输出特征图空间大小 h' 和 ω' 一致,滤波器大小、步幅、填充等超参数与普通卷积(式(1))中参数相同。

幻影特征图由线性变换得到,方程式如下:

$$y_{ij} = \Phi_{ij}(y'_i), \forall i = 1, \dots, m \quad j = 1, \dots, s \quad (4)$$

其中, y'_i 是 Y' 的第 i 个特征图; Φ_{ij} 是除了最后一个线性变换的第 j 个线性变换,生成第 j 个幻影特征图 y'_{ij} , 即 y'_i 可以有一个或多个幻影特征图。最后一个 Φ_{is} 是保留基本特征图的恒等映射。通过式(4)可以得到 $n = m \cdot s$ 个特征图 $Y = [y_{11}, y_{12}, \dots, y_{ms}]$ 作为 Ghost 模块的输出。

幻影瓶颈(ghost bottlenecks, G-bneck)由两个堆叠的 Ghost 模块组成。第 1 个 Ghost 模块的作用是增加输入特征图通道数,第 2 个 Ghost 模块的作用是减少输出特征图通道数使其匹配输入捷径路径,然后将两个 Ghost 模块的输入和输出连接起来。如图 3(a)步长为 1 的情况,在第 1 个 Ghost 模块后的每一层都使用批量归一化和 ReLU 进行非线性激活,第 2 个 Ghost 模块后不使用 ReLU,只进行批量归一化。如图 3(b)对于步幅为 2 的情况,使用步幅为 2 的深度可分离卷积连接两个 Ghost 模块,通过减少采样层实现快捷路径。

本文的主干网络使用 Ghost Bottlenecks 架构代替传统卷积操作,利用多分支结构将单一尺度特征图扩展为多

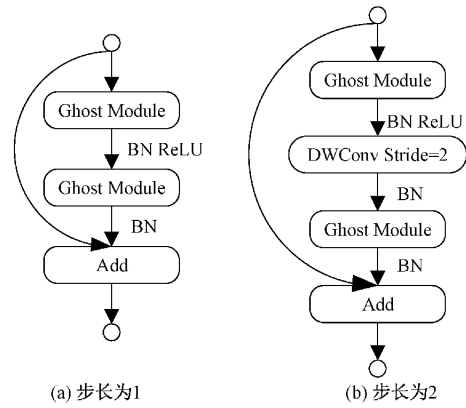


图 3 Ghost Bottlenecks 结构

尺度特征图,减少特征冗余来提升网络处理速度减小网络模型大小,图 4 展示了本网络的整体架构。由于眼睛、嘴巴、鼻子之间的空间关系和对称性,这样的全局结构更有利于精确定位关键点位置,因此通过执行带步幅的卷积操作来拓展多尺度特征图,扩大感受野。同时,保证速度不变的前提下在主干网络后使用多个多尺度全连接层提升检测精度。

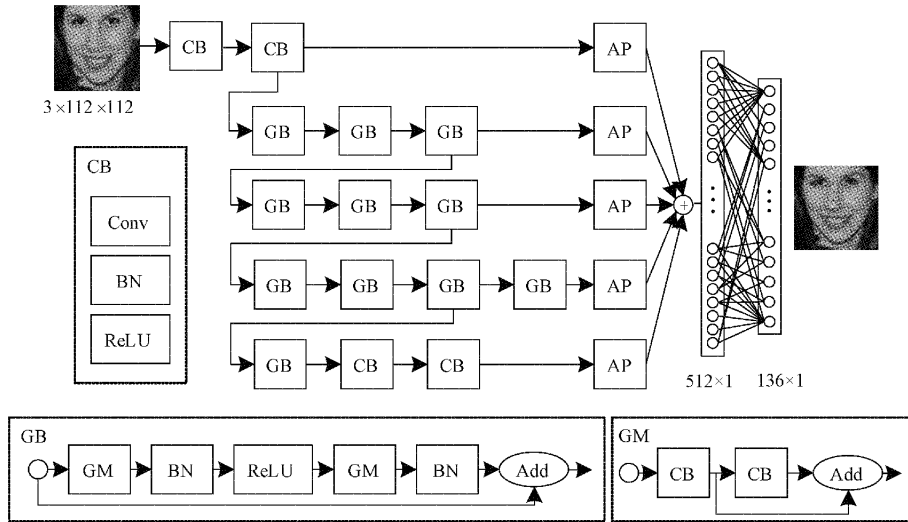


图 4 整体网络结构

图 4 所示的网络架构因计算量大幅度减少,从而加快检测速度。另外,文章中的网络还可以根据实际使用情况,调整 Ghost Model 的宽度因子来压缩模型大小,使网络模型更小、更快,但压缩模型会降低网络的检测精度。网络超参数宽度因子的选取是计算速度和模型准确率之间的折衷,文章旨在确保检测精度和速度的同时,最大化减小模型大小和参数量。网络结构详细配置如表 1 所示,其中 t 为扩张倍数, c 为输出通道数, n 为重复次数, s 为步长。

1.2 损失函数

损失函数的选取在很大程度上决定了训练质量,特别是当训练数据不够大的时候。在现有基于深度神经网络

的人脸特征点检测系统中,常用的损失函数有 $L1^{[19]}$, $L2^{[19]}$ 、smooth $L1^{[20]}$ 和 Wing Loss^[21] 损失函数,损失函数曲线如图 5 所示。通过图 5 可以直观看出 $L1$ 和 $L2$ 这两个函数的梯度大小分别为 1 和 $|x|$,应对最佳步长大小为 $|x|$ 和 1,且 $L2$ 损失函数对异常值较敏感。关键点位置差异小时, $L1$ 、 $L2$ 和 smooth $L1$ 产生的值也非常小,当网络预测几乎正确时,会阻碍网络训练。Wing Loss 损失函数通过对 $L1$ 和 $L2$ 损失函数进行优化,使用带偏移的对数函数应对小误差,使用 $L1$ 应对较大误差。Wing Loss 损失函数对中小误差值响应较为敏感,对于大误差值响应相对较慢。具体如式(5)所示。

Wing Loss 损失函数为:

表1 网络结构配置

输入	操作	t	c	n	s
$112 \times 112 \times 3$	Conv 3×3	—	64	1	2
$56 \times 56 \times 64$	Depthwise Conv 3×3	—	64	1	1
$56 \times 56 \times 64$	GhostBottleneck	2	80	3	2
$28 \times 28 \times 80$	GhostBottleneck	3	96	3	2
$14 \times 14 \times 96$	GhostBottleneck	4	144	4	2
$7 \times 7 \times 144$	GhostBottleneck	2	16	1	1
$7 \times 7 \times 16$	Conv 3×3	—	32	1	1
$7 \times 7 \times 16$	Conv 7×7	—	128	1	1
(S1) $56 \times 56 \times 64$	AvgPool	—	64	1	—
(S2) $28 \times 28 \times 80$	AvgPool	—	80	1	—
(S3) $14 \times 14 \times 96$	AvgPool	—	96	1	—
(S4) $7 \times 7 \times 144$	AvgPool	—	144	1	—
(S5) $1 \times 1 \times 128$	—	—	128	—	—
S1, S2, S3, S4, S5	Full Connection	—	136	1	—

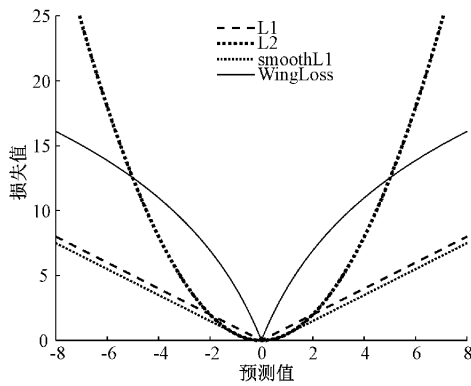


图5 损失函数

$$\text{Wing}(x) = \begin{cases} \omega \ln(1 + \frac{|x|}{\epsilon}), & |x| < 10 \\ |x| - C, & \text{其他} \end{cases} \quad (5)$$

其中,非负 ω 设置非线性部分的范围为 $(-\omega, \omega)$, ϵ 限制非线性区域的曲率和 $C = \omega - \omega \ln(1 + \frac{\omega}{\epsilon})$ 。

在本网络的实际应用中,选取 $\omega = 10$ 和 $\epsilon = 2$ 来具体应用,把参数代入方程中得:

$$\text{Wing}(x) = \begin{cases} 10 \ln(1 + \frac{|x|}{2}), & |x| < 10 \\ |x| - (10 - 10 \ln(6)), & \text{其他} \end{cases} \quad (6)$$

2 实验验证

2.1 数据集

WIDER Facial Landmarks In The Wild (WFLW)^[9] 包含 10 000 张人脸(7 500 张用于训练,2 500 张用于测试),和 98 个完全手动标注的关键点信息。除了人脸关键点标

注外,WFLW 数据集还包含不同属性的注释,如遮挡、姿势、化妆、照明、模糊和表情的标注信息。WFLW 数据集与之前的数据集相比,面孔在表情、姿势和遮挡方面更全面,补充了很多恶劣环境下拍摄的图片;只需在一个数据集上评估模型的鲁棒性,而不用在多个数据集上进行切换对比。因此,为了验证网络模型性能,使用具有挑战性的 WFLW 数据集进行测试。

本文使用的数据集,在原有 WFLW 数据集上进行调整,首先使用原数据集的 10 000 张图片(7 500 张用于训练,2 500 张用于测试),把原来 98 个关键点标注重新标注为 68 个关键点标注。其次去掉原数据集上姿态、遮挡、化妆、照明、模糊和表情等额外标注信息,只保留人脸关键特征如眼睛、鼻子、嘴巴和脸部轮廓的标注信息。改善后的数据集在减少网络模型计算量的同时,能够捕获足够人脸信息并进行关键点标注。

2.2 图像预处理

为了得到更多的训练样本,将 WFLW 数据集所有图片按照给定人脸边框信息进行处理,把所有图片裁剪成只保留人脸部分且尺寸为 112×112 pixels;对训练数据集每张图片进行 $-15^\circ \sim 15^\circ$ 随机旋转和随机翻转操作,使训练数据集扩大为原来的 10 倍。处理后的图像样本如图 6 所示。

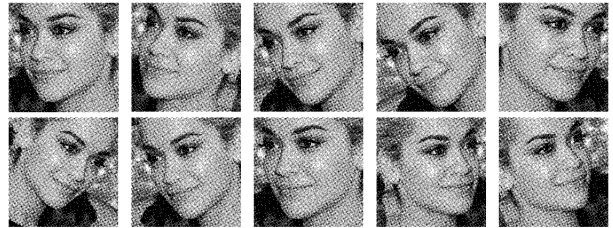


图6 图像预处理样本

2.3 评价指标

采用归一化平均误差(normalized mean error, NME)和失败率(failure rate, FR)来衡量模型准确性(式(7)),对所有带标注的关键点进行平均误差归一化,其中归一化系数使用外眼角距离^[22]。

$$\text{NME} = \frac{1}{K} \sum_{k=1}^K \text{NME}_k \quad (7)$$

$$\text{NME}_k = \frac{1}{N_L} \sum_{i=1}^{N_L} \frac{\|Y_i - \hat{Y}_i\|}{d} \quad (8)$$

其中, Y 是真实关键点位置矩阵, \hat{Y} 是预测关键点位置矩阵, d 是归一化系数, N_L 是数据集中每个图片面部关键点数量, K 是测试集中 batch_size 的数量。

$$\text{FR} = \frac{1}{K} \sum_{k=1}^K [\text{NME}_k \geq 10\%] \quad (9)$$

FR 表示归一化平均误差大于 10% 阈值的图像个数,其数值越低越好。

2.4 实验条件和参数细节

本文的网络模型在 AMD5800H CPU 和 NVIDIA RTX3060 GPU 的 PC 机上进行训练,训练数据集为重新标定的 68 个关键点的 WFLW 数据集,输入网络的图片大小为经过随机旋转和裁剪后 112×112 pixels 大小人脸图片。网络通过 PyTorch 框架构建,使用 Adam 技术进行优化,衰减率为 10^{-6} ,训练 echo 为 200,初始学习速率 10^{-4} ,训练过程中动态调整学习速率。

3 实验结果分析

3.1 模型大小与计算量

如表 2 所示,与现有轻量化网络表现最好的基于深度学习方法的 PFLD 网络模型相比,本论文网络模型在宽度因子为 0.25X 时,模型大小仅有 420 kB,比宽度因子同为 0.25X 的 PFLD 网络模型节省了 5 倍;宽度因子为 1X 时,网络模型大小为 3.3 M,比宽度因子同为 1X 的 PFLD 网络模型节省了将近 4 倍(3.79 倍)。

表 2 模型尺寸比较

模型名称	SDM	LAB	PFLD 0.25X	PFLD 1X	OURS 0.25X	OUR 1X
模型大小/MB	10.1	50.7	2.1	12.5	0.41	3.3

如表 3 所示,本论文网络模型在宽度因子为 0.25X 时与相同宽度因子 PFLD 网络模型相比参数量降低 25%,计算量减少 53.7%;宽度因子为 1X 时相比 PFLD 网络模型参数量降低 36%,计算量减少 63%。本网络模型在参数量和计算量上都要优于 PFLD 网络模型。

表 3 计算量和参数量

模型名称	PFLD 0.25X	PFLD 1X	OURS 0.25X	OUR 1X
计算量/MB	22.97	283.07	10.63	104.47
参数量/kB	94.35	1 220.00	70.73	780.08

3.2 检测精度

采用累积误差分布曲线(cumulative errors distribution curve,CED)分别在宽度因子为 0.25X 和 1X 时,将本论文的网络与 PFLD 网络模型进行对比(图 7 和 8)。两个不同的模型均使用 WFLW 数据集进行测试,在 WFLW 数据集上,通过归一化平均误差(NME)来测量精度,并对所有标注的关键点进行归一化,误差取平均值。从曲线中可以看出,在大于 0.02 之后,0.25X 模型比 PFLD 0.25X 在相同误差下面部标志比例最大高 19%;1X 模型比 PFLD 1X 在相同误差下面部标志比例最大高 15.8%。所以,本论文的网络模型均要优于 PFLD 网络模型。

本文网络模型与目前轻量化网络模型中效果最好的

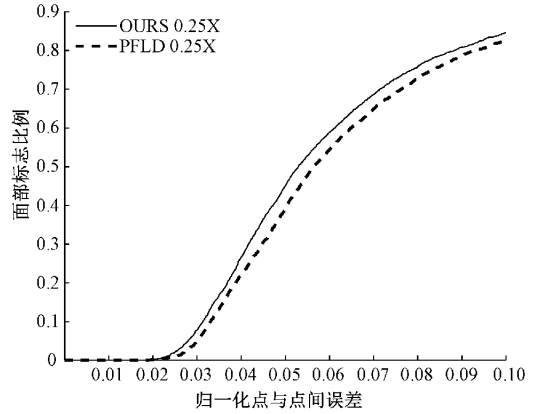


图 7 宽度因子 0.25 累积误差分布曲线

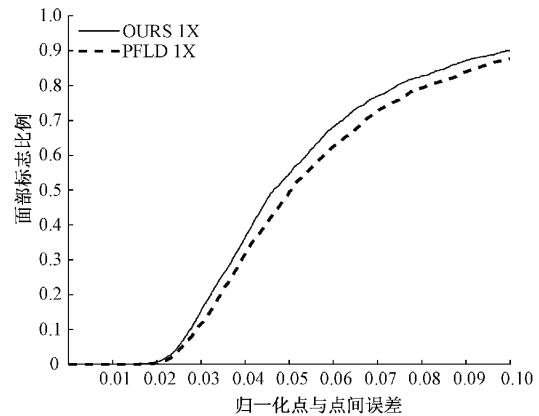


图 8 宽度因子 1 累积误差分布曲线

PFLD 进行对比。两个模型都在 WFLW 数据集上进行测试。从表中可以看出,我们的模型和 PFLD 在同一个宽度因子的情况下,归一化平均误差都要优于 PFLD 模型。其中,网络宽度因子为 0.25X 的归一化平均误差相比 PFLD 降低了 6.6%,网络宽度因子为 1X 的归一化平均误差相比 PFLD 降低了 7%。

从表 4 得出,本文网络在宽度为 0.25X 和宽度为 1X 下前向推理时间都要优于 PFLD 网络。其中,0.25X 的推理时间为 5.8 ms(172 fps),比 PFLD 0.25X 推理时间提高了 18.3%;1X 的推理时间比 PFLD 1X 提高了 29.2%。

表 4 精度和时间对比

模型名称	归一化平均误差	失败率	推理时间/ms
PFLD 0.25X	0.073 2	0.174 8	7.1
PFLD 1X	0.063 3	0.128 0	26.4
OURS 0.25X	0.068 4	0.154 0	5.8
OURS 1X	0.058 8	0.099 2	18.7

图 9 显示了使用本文网络模型宽度因子为 0.25 在 WFLW 数据集下人脸的一些测试结果,这些人脸处于不同的姿势、灯光、遮挡、模糊以及表情各异。本论文的网络模

型可以很好的定位到人脸关键点,对于遮挡的人脸可以识别出来并自动补全。



图9 极端条件下测试结果

为进一步验证模型鲁棒性,对一张图片中存在多张人脸进行验证,使用MTCNN^[23]对图像或输入视频进行人脸检测,然后把检测到的人脸定义为 112×112 pixels大小,输入到本文的模型中进行人脸特征点标定。图10为使用摄像头实时采集视频图像在不同姿态下和戴眼镜的截图,模型检测效果较好。图11和12为网上下下载的多张人脸的图片,模型的检测效果较好。

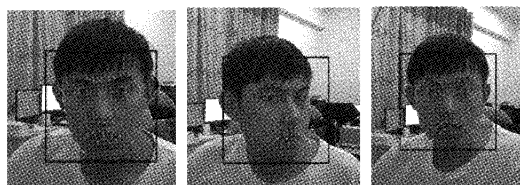


图10 视频效果截图



图11 多张人脸识别效果



图12 多张人脸识别效果图

4 结 论

为了方便在嵌入式设备上部署深度学习模型的人脸关键点检测,需要特别关注模型压缩即模型的轻量化,来减少计算量。本文提出了一个超轻量化人脸关键点检测算法,主网络架构使用轻量化的Ghost Model块构建,在保证检测精度的同时,极大的降低了模型的大小,很大程度上减少了卷积层的计算量,还可以调整宽度因子使模型的大小进一步降低。通过和PFLD网络模型的对比实验,本文提出的网络模型在模型大小和检测精度上都要优于PFLD,能够满足在嵌入式设备上实时检测的要求。本文为了达到模型超轻量化的目的,没有对人脸的旋转角度进行单独处理分析,在之后的工作中可以考虑添加角度信息进一步提高检测精度。

参考文献

- [1] CARYN F H, RAHADIANTI L. Driver drowsiness detection based on drivers' physical behaviours: A systematic literature review [J]. Computer Engineering and Applications Journal, 2021, 10(3): 161-175.
- [2] KHABARLAK K, KORIASHKINA L. Fast facial landmark detection and applications: A survey[J]. Journal of Computer Science and Technology, 2022, 22(1): 12-41.
- [3] KOPACZKA M, SCHOCK J, MERHOF D. Super-realtime facial landmark detection and shape fitting by deep regression of shape model parameters[J]. Arxiv Preprint, 2019, ArXiv: 1902. 03459.
- [4] 石高辉, 陈晓荣, 刘亚茹, 等. 基于卷积神经网络的人脸关键点检测算法设计[J]. 电子测量技术, 2019, 42(24): 125-130.
- [5] COOTES T F, TAYLOR C J, COOPER D H. Active shape models-their training and application [J]. Computer Vision and Image Understanding, 1995, 61(1): 38-59.
- [6] COOTES T F, EDWARDS G J, TAYLOR C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6): 681-685.
- [7] CRISTINACCE D, COOTES T. Automatic feature localisation with constrained local models[J]. Pattern Recognition, 2008, 41(10): 3054-3067.
- [8] ZHANG Z, LUO P, LOY C C. Facial landmark detection by deep multi-task learning[C]. European Conference on Computer Vision. Springer, Cham, 2014: 94-108.
- [9] WU W, QIAN C, YANG S, et al. Look at boundary:

- A boundary-aware face alignment algorithm [C]. Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2129-2138.
- [10] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [11] 李仁鹰, 钱慧芳, 郭佳豪, 等. 基于 M-YOLOv4 模型的轻量化目标检测算法[J]. 国外电子测量技术, 2022, 41(4): 15-21.
- [12] GUO X, LI S, ZHANG J, et al. PFLD: A practical facial landmark detector[J]. Arxiv Preprint, 2019, ArXiv: 1902.10859.
- [13] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:6848-6856.
- [14] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. Arxiv Preprint, 2017, Arxiv:170404861.
- [15] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020:1580-1589.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [17] 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
- [18] 张海超, 张闯. 融合注意力的轻量级行为识别网络研究[J]. 电子测量与仪器学报, 2022, 36(5): 173-179.
- [19] NATEKIN A, KNOLL A. Gradient boosting machines, a tutorial[J]. Frontiers in Neurorobotics, 2013, 7: 21.
- [20] GIRSHICK R. Fast R-CNN[C]. Proceedings of the Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [21] FENG Z-H, KITTLER J, AWAIS M, et al. Wing loss for robust facial landmark localisation with convolutional neural networks[C]. Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2235-2245.
- [22] TRIGEORGIS G, SNAPE P, NICOLAOU M A, et al. Mnemonic descent method: A recurrent process applied for end-to-end face alignment[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4177-4187.
- [23] ZHANG K, ZHANG Z, LI Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499-503.

作者简介

朱望纯(通信作者), 研究员, 主要研究方向为控制工程、相对导航、自动测试总线 VXI/PXI, 虚拟仪器及可互换式虚拟仪器(IVI)等。

E-mail: zhwch@guet.edu.cn

张博, 硕士研究生, 主要研究方向为图像处理与模式识别。