

DOI:10.19651/j.cnki.emt.2107045

VGG网络与多特征融合的遮挡人脸检测*

何其霖 穆平安

(上海理工大学 光电信息与计算机工程学院 上海 200082)

摘要: 实际生活中的人脸图像大部分带有遮挡,常常导致待检测人脸的关键信息丢失。针对人脸识别过程中由于遮挡所导致的人脸特征难以提取的问题,本文设计了一种基于VGGNet和多特征点融合的遮挡人脸检测算法。该方法以VGG-16框架为特征提取的骨干网络,在传统VGG网络的全连接层输入前增加遮挡处理单元OCC-Net。在该层中首先采用多特征融合的方法,增强网络对人脸特征的提取;然后利用尺度不变特征转换(SIFT)算法,扩大网络中小尺度特征图,得到更为丰富的互补信息,改善了传统VGG网络由于多次卷积、池化操作后所导致的小尺度特征损失严重的问题;最后改进回归框参数以降低损失函数对遮挡区域的敏感度,通过边框回归得到遮挡区域的位置信息,提升了有遮挡情况下的人脸检测精度。实验结果表明,相较于PCANet、Faster RCNN及未添加OCC-Net的传统VGGNet等常用算法,本文算法在常用的FDDB以及RMFD等遮挡数据集上,均能更精确地定位被遮挡的人脸,证实了该算法的有效性和鲁棒性。

关键词: 深度学习;人脸检测;遮挡识别;特征提取

中图分类号: TP911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Occlusion face detection based on VGG network and multi-feature fusion

He Qilin Mu Pingan

(School of Optoelectronic Information and Computer Engineering, Shanghai University of Technology, Shanghai 200082, China)

Abstract: Most face images in real life are occluded, which often leads to the loss of key information of the face to be detected. Aiming at the problem of difficulty in extracting facial features due to occlusion in the process of face recognition, this paper designs a occluded face detection algorithm based on VGGNet and multi-feature point fusion. This method uses the VGG-16 framework as the backbone network for feature extraction, and adds the occlusion processing unit Occlusion-Net (OCC-Net) before the input of the fully connected layer of the traditional VGG network. In this layer, the method of multi-feature fusion is first adopted to enhance the network's extraction of facial features; then the scale-invariant feature transformation (SIFT) algorithm is used to expand the small-scale feature maps in the network to obtain richer complementary information and improve the traditional VGG network has a serious problem of small-scale feature loss caused by multiple convolution and pooling operations; finally, the regression box parameters are improved to reduce the sensitivity of the loss function to the occluded area, and the position information of the occluded area is obtained through border regression. Improved the accuracy of face detection in the presence of occlusion. The experimental results show that compared with commonly used algorithms such as PCANet, Faster RCNN, and traditional VGGNet without OCC-Net, the algorithm in this paper can more accurately locate the occluded face on the commonly used occlusion data sets such as FDDB and RMFD, which confirms the effectiveness and robustness of the algorithm.

Keywords: deep learning; face detection; occlusion recognition; feature extraction

0 引言

人脸检测作为常用的无接触生物特性鉴别手段,一直

是计算机视觉领域的重要研究方向,即从静止的图像或动态视频中^[1]定位所需要的人脸信息,是视频人像追踪,人像对比等图像视觉领域的技术基础。人脸检测的发展始于

收稿日期:2021-06-24

* 基金项目:2021年学位点引导布局与建设培育项目(XWDB2021105)资助

20世纪70年代中期,但受制于硬件设备的性能不佳,检测结果在精度与速度方面均表现欠佳。直到90年代,随着计算机硬件性能的提升以及相当多优秀算法的提出,如多视觉人脸检测器(multi-view face detection, MVFD)^[2],生成对抗网络(generative adversarial network, GAN)^[3]等。郭伟等^[4]基于PCANet,提出尺度不变特征变换方法(scale-invariant feature transform, SIFT),解决了特征在小尺度变换时易丢失的问题。童小彬等^[5]基于梯度方向直方图(histogram of oriented gradient, HOG)的轮廓匹配算法,解决了轮廓噪声点、非线性光照下匹配效果差的问题。此外,基于支持向量机(support vector machine, SVM)的图像识别方法也得到了广泛应用。田冬雨等^[6]基于SVM,提出了基于灰狼优化算法的图像去噪算法,提升了SVM在人脸识别领域的表现,处理精度与速度上取得了显著进展。

尽管人脸检测算法已经有了一定的发展,但针对复杂场景下的检测精度与速度仍有待提高。在现实生活中,人脸遮挡问题普遍存在,例如在新冠疫情(COVID-19)期间,人们都佩戴上了口罩,护目镜等物品,对人脸的部分关键位置造成了遮挡,导致了人脸特征提取的不完整,从而使得非深度学习的常规检测算法产生错检甚至是漏检。Wright等^[7]基于稀疏矩阵,利用局部空间改进的思路,将未遮挡区域处理为新的投影空间,将识别过程放在新的投影空间上进行。但此算法对于特征并不敏感,识别效果依赖于图像的维度,计算复杂度较大。胡正平等^[8]提出了针对Gabor特征集结合判别式字典学习的稀疏表示图像识别,提高了对图像光照遮挡、表情遮挡的识别能力,但对于遮挡噪声较多的情况,算法性能不太理想。刘洪缘等^[9]依据特征构造方法的不同将遮挡人脸检测分为基于手工设计特征的经典方法和基于深度学习的现代方法两大系列。

张莺露等^[10]总结了局部遮挡人脸识别研究综述,冯友兵等^[11]针对传统SVM和CNN(convolutional neural network)在有遮挡识别中存在的检测精度低,抗干扰能力较差等问题,本文参考王洋^[12]对YOLOv5框架的改进方法,提出一种基于VGGNet(visual geometry group network)的有遮挡人脸检测算法,在常用人脸检测算法VGG网络结构基础上,增加一层用于处理遮挡的OCC-Net,该处理层对VGG网络的输出特征进行全局融合,改善了传统VGG网络多次卷积、池化操作后,小尺度特征损失较为严重的情况,并对不同的特征通道重新进行不同的赋值,基于区域先验信息^[13],改变损失函数对遮挡区域的敏感度,提升该网络鲁棒性。本文算法在部分遮挡的情况下,仍旧有较好的表现。能更高效地识别佩戴口罩的人脸信息,提升无接触检测的准确度。

1 VGG网络框架

1.1 VGGNet模型介绍

VGGNet^[14]是机器学习常用的人脸检测框架。该框

架便于从图像中提取CNN特征,具有优秀的迁徙学习能力。本文拟在VGGNet框架的基础上,嵌入遮挡信息检测与处理网络。

VGGNet,也叫做VGG-16,即该网络包含了16个卷积层与全连接层。该算法没有太多的超参数,仅专注于卷积层的构建,具有相对较快的收敛速度。该网络可大致分为4个级联网络,采用卷积层加池化层加全连接层组合的思路,进行目标检测。VGG网络的整体架构如图1所示。

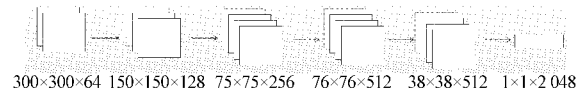


图1 VGG网络架构

4个级联网络分别为快速从输入图像到一次特征的First-Stage,保持原始图像平移不变性而做池化(pooling)的Second-Stage,对图像特征进行高精度过滤提取的Third-Stage,增大待测图像感受野,生成最终检测框的Fourth-Stage。

1.2 VGG-16检测流程

VGG-16图像检测流程如下:

1)原始图像输入 $300 \times 300 \times 3$ 的图片,通过VGG-16网络结构中的第1层(Conv1_1)和第2层(Conv1_2)。Conv1层含有两个 $3 \times 3 \times 3$ 的卷积层,主要用于生成候选框。同时用小卷积核取代之前的大卷积核,减少了近一半参数。原始图像被处理为了 $300 \times 300 \times 64$ 的矩阵,该层网络结构如图2所示。

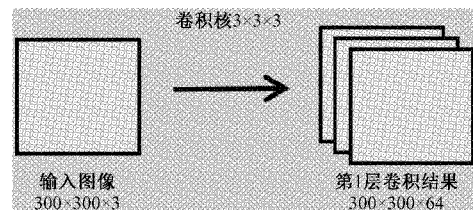


图2 Conv1层网络结构

2)对于Conv1层中置信度高于阈值的候选框,输入Conv2层进行平均池化操作,利用平移不变性保留图片的背景信息,同时将图像大小压缩。Conv2层网络结构如图3所示。

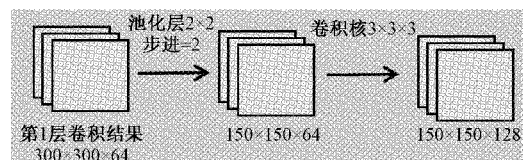


图3 Conv2层网络结构

3)Conv3层是由3个 $3 \times 3 \times 3$ 的卷积层构成,对图像特征进行提取,对人脸位置进行更精确的回归,利用非最大值抑制进行去重。Conv3网络结构如图4所示。

4)Conv4由1个池化层和3个全连接层组成,Conv4

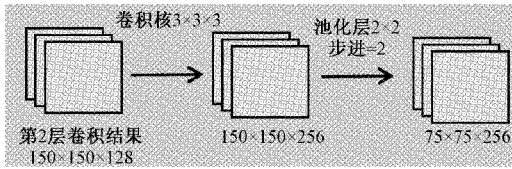


图 4 Conv3 层网络结构

将会输出最终的人脸框位置以及人脸检测信息。Conv4 网络结构如图 5 所示。

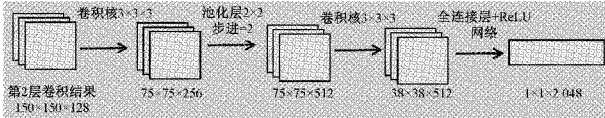


图 5 Conv4 层网络结构

VGG16 网络由 13 个卷积层与 3 个全连接层构成,随着网络的加深,特征的提取也更加的精细,同时网络中的每一个块(block)中均有多个卷积(Conv)操作,减少了参数的同时还获得了同样的感受野。传统 VGG 网络人脸检测结果如图 6 所示。



(a) 无遮挡情况



(b) 有遮挡情况

图 6 传统 VGG 网络人脸检测

图 6 实验结果表明,对于未有遮挡的目标检测,传统 VGG 网络已经有较好的识别度;而当目标面部有部分遮挡时,VGG 的识别准确度会受到遮挡物的影响,从而导致部分特征的提取不够完整,影响识别精度。

2 研究内容

2.1 网络模型设计

针对实名认证,生物识别等检测业务的需要,在现有的人脸检测框架基础上,提出了基于 VGG16 网络的有遮挡人脸识别算法。旨在提升被检测对象在佩戴口罩,墨镜等遮挡物时的识别精度,避免了脱去遮挡物的繁琐,提升了体验。

该算法在原有 VGG16Net 的输出结果基础上,重新设计的针对有遮挡物的识别算法。基础网络部分与本文检测算法共享,在原版 VGG 模型的全连接层(fully connection)前增加一个 OCC-Net 对遮挡物信息进行识别。OCC-Net 针对 5 个不同的特征区域(左眼,右眼,鼻子,左嘴角,右嘴角),5 组特征图像分别经过 3 层卷积后,通过 concat 进行特征融合,输出 5 个不同的特征分支。根据特征通道值与网络所设阈值所比较的结果,使用 SENet 的网络结构对各特征通道的权重进行重新赋值。对能直接检测到的 5 点特征赋予高权重,而对无法直接检测到的点,利用 SIFT 算法来侦测图像中的局部性特征,在空间尺度寻找极值点,将最终得到的五点特征进行全局融合,降低损失函数对遮挡区域的敏感度。OCC-Net 网络结构如图 7 所示。

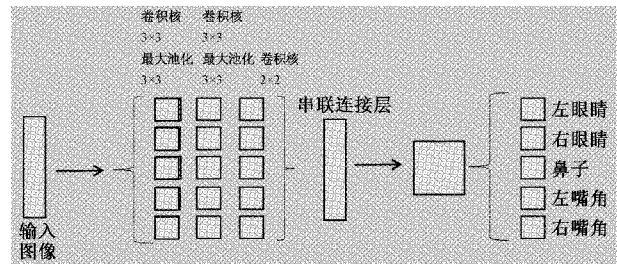


图 7 OCC-Net 网络结构

2.2 确定损失函数

1)人脸分类:属于二分类问题,因此采用交叉熵损失函数:

$$L_i^{\text{det}} = -(y_i^{\text{det}} \log(p_i)) + (1 - y_i^{\text{det}})(1 - \log(p_i)) \quad (1)$$

式中: y_i^{det} 表示第 i 个样本的标签; p_i 表示第 i 个样本是人脸的概率。

2)生成候选框:候选框是预测与它最接近的人脸框,属于回归问题,采用均方差损失函数:

$$L_i^{\text{box}} = \arg \min_{(w, b)} \sum_{i=1}^m (f(x_i) - y_i)^2 \quad (2)$$

式中: $f(x_i)$ 表示网络预测第 i 个样本人脸框的位置; y_i 表示第 i 个人脸框的实际位置。

3)人脸关键点位置检测:属于特征点的回归问题,采用支持向量回归(support vector regression, SVR),关键点位置检测可形式化为:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m l_e(f(x_i) - y_i) \quad (3)$$

式中: w 为模型参数; l_e 为 e-不敏感损失(e-insensitive loss)函数; $f(x_i)$ 表示网络预测第 i 个样本特征点的位置; y_i 表示第 i 个样本特征点的实际位置。

4)遮挡信息:预测关键特征点是否被遮挡,属于多元分类问题,采用交叉熵损失函数:

$$L_i^{\text{occ}} = -(y_i^{\text{occ}} \log(p_i)) + (1 - y_i^{\text{occ}})(1 - \log(p_i)) \quad (4)$$

式中: y_i^{occ} 表示第 i 个样本的真实遮挡信息; p_i 表示第 i 个样本关键点被遮挡的概率。

3 实验结果与分析

3.1 实验环境

本文在训练好的模型上进行实验,基于 paddle 深度学习框架搭建,实验环境为 Inter I7 9700K CPU, GTX1070 Ti 显卡,内存为 256 G RDIMM DDR4,集成开发环境为 Pycharm。

3.2 训练集

由于没有公开的整合了口罩,墨镜,帽子于一体的遮挡测试数据集,因此在实验时以常用人脸数据集 Fddb 及 RMFD 为基础,自行制作的数据集共 600 张,将其分为帽子遮挡,口罩遮挡,墨镜遮挡 3 类数据各 200 张,如图 8 所示。



图 8 自制数据集

每一数据集包括 3 部分:训练集 train、验证集 val、测试集 test,分别占样本数的 40%、20%、40%。对所测试的数据集进行手工标注、模型训练与预测。使用训练好的 VGGNet 模型进行特征提取、融合后送入 OCC-Net 遮挡处理层,确定五官特征点是否被遮挡,进而通过遮挡分类器,使用基于 k 近邻学习的标签预测进行识别。并与常用 PCANet 算法、Faster-RCNN 和传统 VGGNet 进行对比。实验平均耗时如表 1 所示。

表 1 不同算法处理有遮挡的数据集平均耗时

编号	算法	平均每一张图片耗时/ms
1	PCANet	39.5~72.3
2	Faster-RCNN	23.8~92.8(时间随图片尺寸增加)
3	传统 VGGNet	32.5~78.4
4	本文算法	36.1~86.3

表 1 结果表明,在添加了 OCC-Net 遮挡信息处理层后,本文算法在处理遮挡图像的时间性能上相较于常用 PCANet、Faster-RCNN 以及未添加 OCC-Net 的传统 VGG 网络,并未有明显增加,算法的有效性较好。

为进一步验证本文算法的可靠性,采用 crossvalid 函数^[15]对 3 类数据集共 600 张进行混合后交叉分包,分为 3 组混合数据集,每一集合包含不同遮挡部位(口罩、墨镜、帽子)的共 200 张灰度图像。重新训练集后,再次进行实验比较,验证算法识别准确率。口罩遮挡实验如图 9 所示。



图 9 口罩遮挡识别

墨镜遮挡实验如图 10 所示。

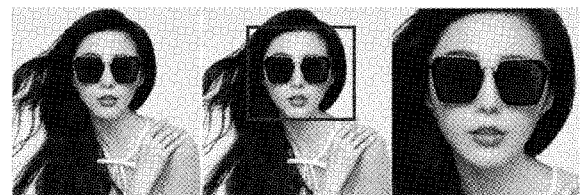


图 10 墨镜遮挡识别

帽子遮挡实验如图 11 所示。

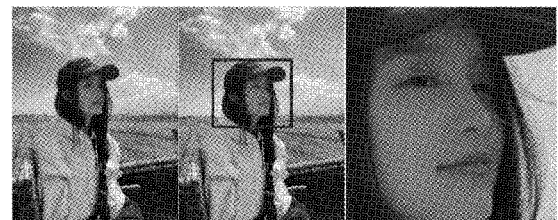


图 11 帽子遮挡识别

平均识别准确率如表 2 所示。

表 2 不同遮挡下条件下实验结果 %

编号	算法	口罩遮挡	墨镜遮挡	帽子遮挡
		平均识别率	平均识别率	平均识别率
1	PCANet	78.0	75.0	80.0
2	Faster-RCNN	83.0	80.0	87.0
3	传统 VGGNet	93.5	93.0	98.0
4	本文算法	98.0	96.0	98.5

实验结果表明,在有遮挡情况下,本文算法通过对遮挡信息进行处理,较其他算法性能有明显提升。同时,从实验数据可以得出,不同的遮挡部位对实验的结果有不同的影响,佩戴墨镜的平均识别准确率普遍偏低,这或许是因为人眼^[16]是作为生物识别的一个相对关键的特征,墨镜对人眼特征的提取起到了较大的干扰;而佩戴帽子则影响相对较小,原因可能为帽子的遮挡对五官特征的影响最小,因此识别准确度相对较高。为验证本文算法的动态范围,人为改

变待测人像被遮挡比例,绘制遮挡比例-识别精度曲线,横坐标为遮挡比例,纵坐标为相应的遮挡比例下算法的平均识别准确率。不同遮挡比例下各算法实验结果如图 12 所示。

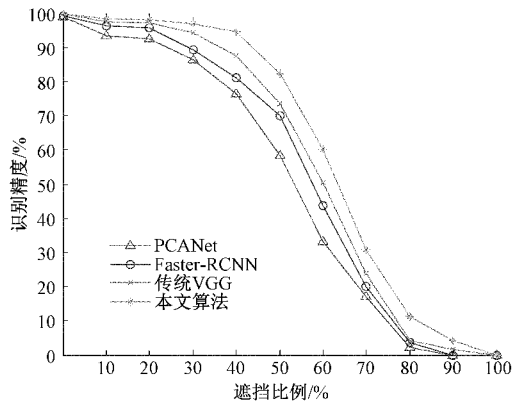


图 12 不同遮挡比例下各算法实验结果

最后的实验结果表明,本文在遮挡程度不太高的情况下,识别准确率较其他算法都有更好的性能;但当遮挡比例超过 70% 时,识别效率骤降,原因与特征点的过多丢失有较大关系,日后可结合特征丢失后的修复^[17]进行进一步的研究。

总的来说,本文算法在处理有遮挡的人脸检测问题上,相较于常用 PCANet、Faster-RCNN 及未添加 OCC-Net 的传统 VGG 网络而言,在时间复杂度增加不大的情况下,有效的提升了算法精度。本文所提“深度学习+遮挡信息处理”的框架,可推广至其他有遮挡的目标检测问题上,如标签被污染,车牌被遮挡等。具有一定的实践意义。

4 结 论

本文提出了一种基于深度学习的有遮挡人脸检测算法,该算法在现有 VGG-16 网络结构的基础上,增加了基于多特征融合的遮挡信息处理层 OCC-Net,增强了网络对遮挡干扰的处理能力,提升检测准确度。本文算法保留了 VGGNet 超参数少,计算量小以及网络收敛速度快等优势。同时 OCC-Net 作为加性算法,并未显著增加算法的时间复杂度。

实验结果表明,本文方法通过关键点定位以及优化损失函数对遮挡区域的敏感度,在检测有遮挡的目标时,检测精度较传统方法有所提升。基于“现有网络框架”与“遮挡信息处理层”相结合的检测模型可以推广至其他特征缺失与噪点混叠的图像检测工作中。

参考文献

- [1] 兰胜坤. 基于 Adaboost 算法的人脸检测实现[J]. 电脑与信息技术, 2021, 29(2): 16-19.
- [2] 郭耸, 顾国昌, 蔡则苏, 等. 基于特征融合与决策树级联

结构的多姿态人脸检测[J]. 沈阳工业大学学报, 2012, 34(2): 203-208.

- [3] 胡龙辉, 王朝立, 孙占全, 等. 基于 WGAN 的图像识别方法[J]. 控制工程, 2020, 27(12): 2168-2175.
- [4] 郭伟, 白文硕, 曲海成. PCANet 下的遮挡定位人脸识别算法[J]. 计算机科学与探索, 2019, 13(12): 2149-2160.
- [5] 童小彬, 孟婷, 孙以泽, 等. 基于 HOG 特征描述的轮廓匹配算法[J]. 东华大学学报(自然科学版), 2020, 46(5): 787-792.
- [6] 田东雨, 何玉珠, 宋平. 基于灰狼优化算法的 SVM 的图像噪声识别[J]. 电子测量技术, 2019, 42(4): 90-94.
- [7] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [8] 胡正平, 徐波, 白洋. Gabor 特征集结合判别式字典学习的稀疏表示图像识别[J]. 中国图象图形学报, 2013, 18(2): 189-194.
- [9] 刘淇缘, 卢树华, 兰凌强. 遮挡人脸检测方法研究进展[J]. 计算机工程与应用, 2020, 56(13): 33-46.
- [10] 张莺露, 史璐媛, 董馨格, 等. 局部遮挡人脸识别技术综述[J]. 科技经济导刊, 2021, 29(8): 54-55.
- [11] 冯友兵, 陆轶秋, 仲伟波. 基于 CNN 和 SVM 的人脸识别系统的设计与实现[J]. 计算机与数字工程, 2021, 49(2): 378-382, 420.
- [12] 王洋. 改进的 Yolov5 口罩和安全帽佩戴人工智能检测识别算法[J]. 建筑与预算, 2020(11): 67-69.
- [13] 陈先锋, 郭正华, 伍俊龙, 等. 基于区域先验信息的去遮挡立体匹配算法[J]. 激光与光电子学进展, 2019, 56(19): 95-101.
- [14] 陈津微, 张元良, 尹泽睿. 基于改进的 VGG19 网络的面部表情识别[J]. 电脑知识与技术, 2020, 16(29): 187-188.
- [15] 于万波, 王香香, 王大庆. 基于离散余弦变换基函数迭代的人脸图像识别[J]. 图学学报, 2020, 41(1): 88-92.
- [16] 张胜虎, 马惠敏. 遮挡对于目标检测的影响分析[J]. 图学学报, 2020, 41(6): 891-896.
- [17] 方小艳. 基于 SIFT 特征点检测与维纳滤波的图像复原算法[J]. 电子测量技术, 2017, 40(6): 105-108.

作者简介

何其霖, 硕士研究生, 主要研究方向图像处理、深度学习。

E-mail: ybqilin5@163.com

穆平安, 博士, 教授, 硕士生导师, 研究方向为图像处理、测试信息获取与处理、在线检测技术与装置。

E-mail: mupingan@163.com