



基于增强卷积神经网络的尺度不变人脸检测方法

李保华 王海星

Scale-invariant face detection method based on enhanced convolutional neural network

Li Baohua, Wang Haixing

在线阅读 View online: <https://doi.org/10.3788/IRLA20210586>

您可能感兴趣的其他文章

Articles you may be interested in

基于3D卷积神经网络的PolSAR图像精细分类

Fine classification of polarimetric SAR images based on 3D convolutional neural network

红外与激光工程. 2018, 47(7): 703001 <https://doi.org/10.3788/IRLA201847.0703001>

基于卷积神经网络的反无人机系统图像识别方法

Image recognition method of anti UAV system based on convolutional neural network

红外与激光工程. 2020, 49(7): 20200154 <https://doi.org/10.3788/IRLA20200154>

多尺度卷积神经网络的噪声模糊图像盲复原

Blind deblurring of noisy and blurry images of multi-scale convolutional neural network

红外与激光工程. 2019, 48(4): 426001 <https://doi.org/10.3788/IRLA201948.0426001>

流场测速中基于深度卷积神经网络的光学畸变校正技术

Aberration correction for flow velocity measurements using deep convolutional neural networks

红外与激光工程. 2020, 49(10): 20200267 <https://doi.org/10.3788/IRLA20200267>

基于深度谱卷积神经网络的高效视觉目标跟踪算法

Efficient visual target tracking algorithm based on deep spectral convolutional neural networks

红外与激光工程. 2018, 47(6): 626005 <https://doi.org/10.3788/IRLA201847.0626005>

采用卷积神经网络的红外和可见光图像块匹配

Infrared-visible image patches matching via convolutional neural networks

红外与激光工程. 2021, 50(5): 20200364 <https://doi.org/10.3788/IRLA20200364>

基于增强卷积神经网络的尺度不变人脸检测方法

李保华¹, 王海星²

(1. 郑州升达经贸管理学院 信息工程学院, 河南 郑州 451191;
2. 河南理工大学 电气工程与自动化学院, 河南 焦作 454003)

摘要: 针对非约束场景下小尺寸人脸检测困难的问题, 提出了一种基于增强卷积神经网络的尺度不变人脸检测方法。首先, 在 SSD 基础检测网络的两个浅层特征图上, 通过协调聚合当前层特征图和前后两层特征图的特征信息, 对当前层特征图的鉴别性和稳健性进行增强。然后, 对两个增强特征图进行负样本筛选, 通过增加分类的难度来降低由小尺寸锚框引起的人脸检测假正率上升。最后, 为原始特征图和增强特征图设置了两种基于锚框尺寸的损失函数, 并通过加权求和的方式对其进行融合。在 Fddb 和 WIDER FACE 数据集上的测试结果表明, 文中所提方法比目前主流人脸检测方法具有更高的检测精度。

关键词: 多尺度人脸检测; 卷积神经网络; 特征图增强; 负样本筛选

中图分类号: TN911.73; TP391 文献标志码: A DOI: 10.3788/IRLA20210586

Scale-invariant face detection method based on enhanced convolutional neural network

Li Baohua¹, Wang Haixing²

(1. School of Information Engineering, Zhengzhou Shengda University, Zhengzhou 451191, China;
2. School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454003, China)

Abstract: Aiming at the difficulty of small-scale face detection in unconstrained scenes, the proposes a scale-invariant face detection method based on enhanced convolutional neural networks. Firstly, On the two shallow feature maps of the SSD basic detection network, the discrimination and robustness of the current layer feature map was enhanced by blending the feature information of the current layer feature map and adjacent layer feature map. Then, the negative sample screening was performed on the two enhanced feature maps, and the false positive rate of face detection caused by the small-scale anchor box was reduced by increasing the difficulty of classification. Finally, two loss function based on anchor boxes size were set for the original feature map and the enhanced feature map, and they were merged by weighted summation. The test results on the Fddb and WIDER FACE datasets show that the proposed method has higher detection accuracy than the current mainstream face detection methods.

Key words: multi-scale face detection; convolutional neural network; feature map enhancement; negative sample screening

收稿日期: 2021-12-20; 修订日期: 2022-01-20

作者简介: 李保华, 男, 讲师, 硕士, 主要从事计算机网络、区域信息化规划建设方面的研究。

王海星, 副教授, 博士, 主要从事电机及其驱动系统的理论、应用和控制方面的教学和科研工作。

0 引言

人脸检测作为面部信息处理的基础技术,被广泛的应用到人脸识别^[1]、人脸跟踪^[2]、人脸特征点定位^[3]和人脸属性分析^[4]等任务,其检测结果直接影响后续步骤的精确性和稳健性,具有重要的研究意义。人脸检测可分为约束场景下的人脸检测和非约束场景下的人脸检测^[5],后者包含跨光谱^[6]、多人密集环境^[7]、局部遮挡^[8]等多种复杂环境。由于约束场景下的人脸检测较为简单,而非约束场景下的人脸检测受人脸尺度多样性、姿态多样性、遮挡、模糊等条件影响比较困难,因此当前人脸检测的研究热点主要集中于非约束场景。

作为人脸检测的先驱,参考文献 [9] 提出了一种基于 Haar-like 特征和 AdaBoost 级联分类器的人脸检测方法,取得了相对较好的检测效果,但其仍局限于人工设计特征的弱描述能力。

在神经网络迅速发展的今天,研究人员已经开始使用更强特征描述能力的卷积神经网络 (Convolutional Neural Networks, CNN)^[10-11] 来取代传统方法进行人脸检测,并且在检测精度、速度和适应性等方面取得了较大的提高。比如 Cascade CNN^[12],就是 AdaBoost 人脸检测方法的卷积神经网络实现,和传统方法相比, Cascade CNN 具有更高的检测精度和速度。虽然基于 CNN 的人脸检测方法进一步提高了人脸检测的准确率,但对非约束场景下的小尺寸人脸检测仍具难度。

基于卷积神经网络的人脸检测方法主要分为两种:第一种是两阶段人脸检测方法,先在区域建议网络^[13](Region Proposal Network, RPN) 中经过端到端的训练,再将训练得到的区域候选图送入到 Faster RCNN 进行分类和回归;另一种是源于单一神经网络物体检测器 (Single Shot MultiBox Detector, SSD) 的人脸检测方法^[14],它不需要进行候选区域的建议,直接在多尺度卷积层上执行人脸/背景分类和人脸边界框位置回归。参考文献 [15] 提出了一种融合面部区域上下文信息的人脸检测方法,对小尺寸人脸检测效果较好,但运算时效性低。参考文献 [16] 通过等比例补偿锚匹配策略提升对小尺寸人脸的正样本分辨能力,通过 max-out 标签增强了人脸分类能力和检测准确度。参考文献 [17] 提出了一种基于背景协作描述的单一神

经网络人脸检测方法,通过扩大检测框的感受野提高了对小尺寸人脸的检测精度。

文中基于 SSD 人脸检测模型,针对非约束场景下的小尺寸人脸检测,从以下 3 个方面进行了改进:(1) 提出了一种特征增强网络。在 SSD 基础检测网络的两个浅层特征图上,通过协调聚合当前层特征图和前后两层特征图的特征信息,增强了当前层特征图的鉴别性和稳健性,提高了对小尺寸人脸的检测精度。(2) 对两个检测小尺寸人脸的增强特征图,进行负样本筛选,通过增加分类的难度来降低由小尺寸锚框引起的人脸检测假正率升高。(3) 为增强特征图和原始特征图分别建立两种基于锚框尺寸的损失函数,并进行加权融合。

1 基于 SSD 的人脸检测模型

与先生成 RPN 再进行目标检测的算法不同, SSD 算法可直接预测目标类别和检测框位置,具有较高的检测精度和很快的检测速度。该算法通过在不同尺度特征图上进行 3×3 尺寸的滑动卷积,可实现精度较高的尺度不变人脸检测结果。基于 SSD 基础检测网络的人脸检测模型如图 1 所示。

由图 1 可知, Conv1_1-Conv5_3 是 VGG16 的前 5 个卷积层, Conv6 和 Conv7 是由 VGG16 的 2 个全连接层 FC6 和 FC7 改造的卷积层, Conv8_2-Conv11_2 是新增的 4 个卷积层。该模型的输入是包含一个或多个面部的图像,输出是一系列人脸检测框及其对应的面部得分,当所有人脸检测框执行非最大值抑制后,得到最终的人脸检测结果。

该 SSD 人脸检测模型的损失函数包含了人脸/背景分类的 softmax 损失 L_{conf} 和人脸检测框回归的 smooth L1 损失 L_{loc} :

$$L_{face}(x, c, t, t^*) = \frac{1}{N} (L_{conf}(x, c) + \lambda L_{loc}(x, t_i, t_i^*)) \quad (1)$$

式中: N 为匹配到人脸检测框的数量; x 为检测框属于人脸的得分,是则取 1,不是则取 0; c 为检测目标的人脸置信度; t_i 为预测框; t_i^* 为真实框; λ 为平衡 softmax 损失和 smooth L1 损失的权重。

在训练时,为了提高运行效率和检测效果,对正负训练样本比例进行控制。首先计算先验框和真实框的交并比 (Intersection over Union, IoU), 取 $\text{IoU} > 0.5$

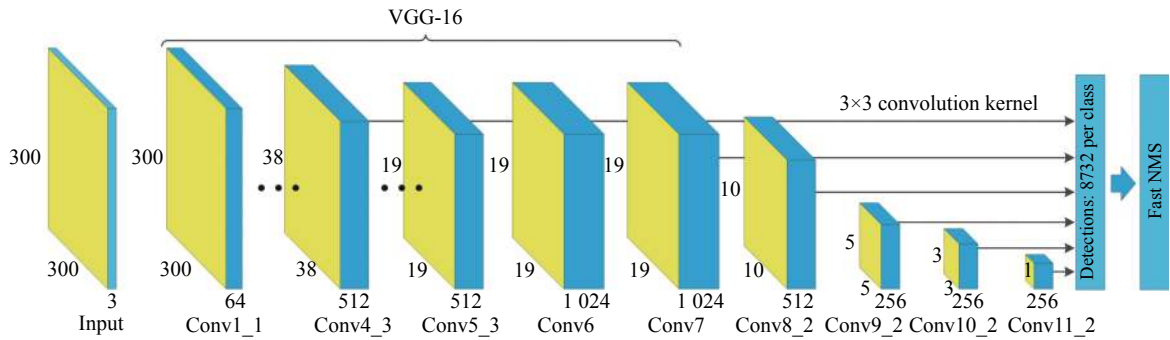


图 1 基于 SSD 基础检测网络的人脸检测模型

Fig.1 Face detection model based on SSD basic detection network

的先验框为正样本,其余先验框为负样本。接着,将检测框与负样本集进行一一匹配,删除匹配失败的检测框以平衡正负样本数量,提高训练的收敛速度。

2 基于增强卷积神经网络的尺度不变人脸检测方法

2.1 总体结构

文中提出的基于增强卷积神经网络的尺度不变人脸检测模型如图 2 所示,该模型主要分为基础卷积

层、特征增强网络、增强特征层和多任务损失层 4 个部分。

基础卷积层采用 SSD 模型从 Conv3_3~Conv7_2 的 6 个多尺度卷积层,每层对应生成的特征图为 $f_1(160 \times 160)$ 、 $f_2(80 \times 80)$ 、 $f_3(40 \times 40)$ 、 $f_4(20 \times 20)$ 、 $f_5(10 \times 10)$ 和 $f_6(5 \times 5)$ 。其中,前 2 个特征图 f_1 和 f_2 被送入特征增强网络进行特征增强。后 4 个特征图直接被用来进行人脸/背景分类和人脸边界框位置回归。

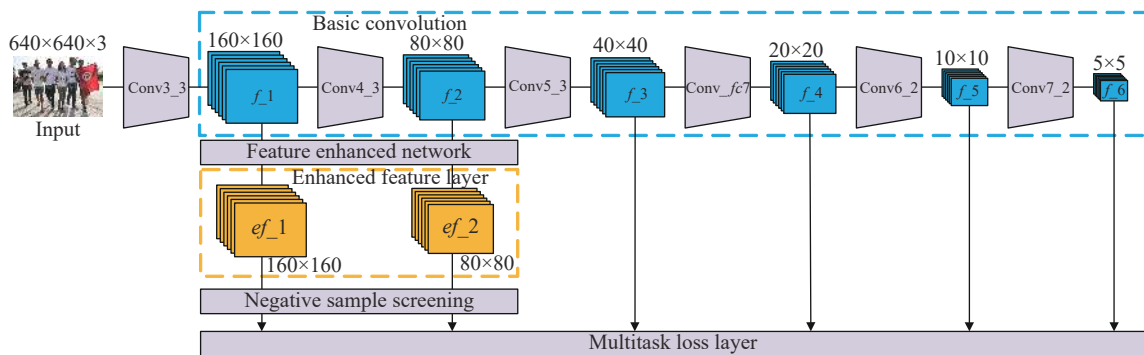


图 2 基于增强卷积神经网络的尺度不变人脸检测模型

Fig.2 Scale-invariant face detection model based on enhanced convolutional neural network

特征增强网络负责协调组合当前层特征图和前后两层特征图的上下文信息。对这 3 层特征图进行归一化卷积、上采样、下采样、点乘、连接和扩张卷积等操作后,得到当前层特征图的增强特征图,并且它们的尺寸与当前层特征图保持一致。由于浅层特征图更适合进行小尺寸目标检测,因此增强特征图 ef_1 和 ef_2 被用来进行小尺寸人脸检测。另外,在训练时还为 ef_1 和 ef_2 加入了负样本筛选,通过增加分类的难度来降低由小尺寸锚框引起的人脸检测假正率升高,进一步提高了人脸检测精度。

2.2 特征增强网络

参考文献 [13] 指出,浅层特征图更适合被用来检测小尺寸人脸,因此,文中提出了特征增强网络来增强浅层特征图 f_1 和 f_2 的鉴别性和稳健性,以进一步提高对小尺寸人脸的检测精度。在增强第 k 层特征图时,需要协调组合第 $k-1$ 层、第 k 层和第 $k+1$ 层特征图的上下文信息,其网络结构如图 3 所示。

由图 3 可以看出,第 $k-1$ 层、第 k 层和第 $k+1$ 层特征图的尺寸和维度均不相同,此时不能直接进行特征图的增强融合。因此,可先遍历采用 1×1 尺寸卷积

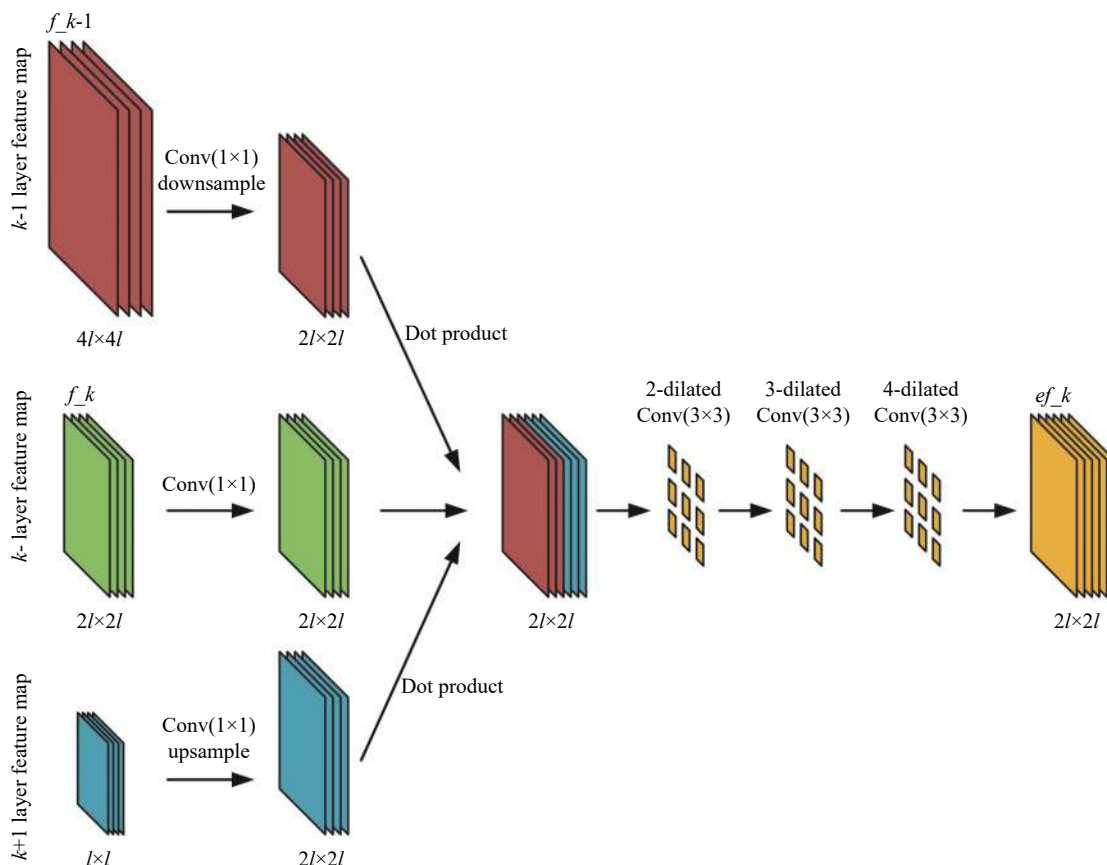


图 3 特征增强网络

Fig.3 Feature enhancement network

核对上述特征图进行尺寸和维度的归一化处理,使 3 层特征图的维度相同。然后分别对第 $k-1$ 层特征图进行下采样,对第 $k+1$ 层特征图进行上采样,使 3 层特征图的尺寸也都为 $2l \times 2l$ 。接着,对第 $k-1$ 层特征图和第 k 层特征图进行逐元素点乘,对第 k 层特征图和第 $k+1$ 层特征图进行逐元素点乘,并将点乘后的 2 个新特征图连接,得到 $2l \times 2l$ 大小的融合特征图。最后,对融合特征图进行扩张卷积(扩张率分别是 2、3、4),得到特征图 f_k 的增强特征图 ef_k 。以上步骤可归纳如下:

$$\begin{aligned} nf_{k-1} &= f_{\text{Dot}}(f_{\text{down}}(f_{\text{conv1}}(f_{k-1})), f_{\text{conv1}}(f_k)) \\ nf_{k+1} &= f_{\text{Dot}}(f_{\text{up}}(f_{\text{conv1}}(f_{k+1})), f_{\text{conv1}}(f_k)) \\ ef_k &= f_{2,3,4\text{-dilated_conv3}}(f_{\text{connect}}(nf_{k-1}, nf_{k+1})) \end{aligned} \quad (2)$$

式中: f_{conv1} 表示遍历采用 1×1 尺寸卷积核执行归一化卷积; f_{down} 表示对第 $k-1$ 层特征图进行下采样; f_{up} 表示对第 $k+1$ 层特征图进行上采样; f_{Dot} 表示对两个特征图进行逐元素点乘; f_{connect} 表示对两个新特征图 nf_{k-1} 和 nf_{k+1} 进行连接; $f_{2,3,4\text{-dilated_conv3}}$ 表示扩张率

分别是 2、3、4,卷积核大小为 3×3 的扩张卷积。

2.3 负样本筛选

基于锚框的人脸检测方法可以被视为一个二元分类问题,即确定锚框所包围的区域是人脸还是背景。由于人脸在整张图像上所占的面积较小,这会导致正负训练样本的严重不匹配,只有一小部分锚框对应的是人脸,而绝大多数锚框对应的是背景,这会增加人脸检测的假正率。这种严重的正负训练样本不匹配主要是由小尺寸人脸检测引起,为了检测小尺寸人脸,必须在图像上密集平铺数量庞大的小尺寸锚框,这会导致负训练样本的急剧增加。为了解决这个问题,文中对 ef_1 和 ef_2 这两个检测小尺寸人脸的增强特征图,进行了负样本筛选,通过增加分类的难度来降低由小尺寸锚框引起的人脸检测假正率升高,该过程的示意图如图 4 所示。

由图 4 可知,对于 ef_1 和 ef_2 这两个增强特征图,输出通道数设置为 $(N_c+1)+4$,其中 N_c+1 对应 N_c 个负样本概率和 1 个正样本概率,4 对应人脸边界框回

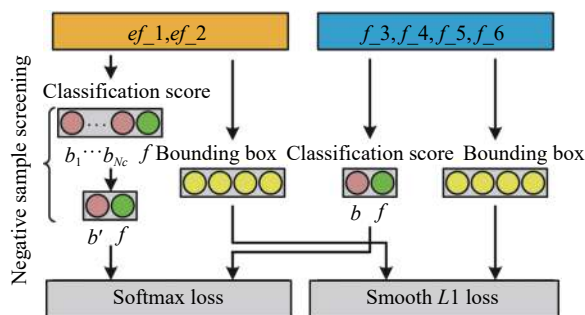


图 4 负样本筛选示意图

Fig.4 Schematic diagram of negative sample screening

归的位置坐标 $[x, y, \Delta x, \Delta y]$ 。对于每个匹配到真实人脸框的锚框,同时预测 N_c 个该锚框的背景锚框,然后选择其中得分最高的背景锚框作为负样本与正样本一起完成 softmax 二分类:

$$b' = \max(b_1, \dots, b_{N_c}) \quad (3)$$

式中: b' 表示得分最高背景锚框的概率; $b_1 \sim b_{N_c}$ 表示 N_c 个预测背景锚框的概率。通过这种局部优化方法,可以降低负训练样本比例,同时增强对小尺寸人脸的检测能力。对于从 $f_3 \sim f_6$ 这四个特征层,输出通道数设置为 2+4,其中 2 表示人脸/背景的二元分类概率,4 表示人脸边界框回归的位置坐标。

2.4 分层损失函数

在浅层特征图上更适合进行小尺寸人脸检测,因此,文中算法在训练时为 ef_1 和 ef_2 这两层增强特征图,设置了基于小尺寸锚框 \hat{r}_i 的损失函数,为 f_3 、 f_4 、 f_5 和 f_6 这四层特征图设置了基于正常尺寸锚框 r_i 的损失函数,并通过加权求和来融合上述两种损失函数。

ef_1 和 ef_2 这两层增强特征图对应的损失函数为:

$$L_{ef}(p_i, t_i, \hat{r}_i) = \frac{1}{N_{conf}} \left(\sum_i L_{conf}(p_i, p_i^*) \right) + \frac{\alpha}{N_{loc}} \left(\sum_i p_i^* L_{loc}(t_i, t_i^*, \hat{r}_i) \right) \quad (4)$$

式中: N_{conf} 表示正负样本锚框数量; L_{conf} 表示对人脸和背景的分类损失; p_i 表示预测为人脸的概率; p_i^* 表示正负样本标签; N_{loc} 表示正样本锚框的数量; L_{loc} 表示使用锚框 \hat{r}_i 时人脸预测框 t_i 和真实框 t_i^* 之间的参数化 smoothL1 损失,当 $p_i^* = 1$ 时,激活 L_{loc} ,并通过权重 α 来平衡 L_{conf} 和 L_{loc} 。

f_3 、 f_4 、 f_5 和 f_6 这四层特征图对应的损失函

数为:

$$L_f(p_i, t_i, r_i) = \frac{1}{N_{conf}} \left(\sum_i L_{conf}(p_i, p_i^*) \right) + \frac{\alpha}{N_{loc}} \left(\sum_i p_i^* L_{loc}(t_i, t_i^*, r_i) \right) \quad (5)$$

式中: 锚框 r_i 的尺寸是 \hat{r}_i 的两倍。将 ef_1 和 ef_2 的损失函数与 f_3 、 f_4 、 f_5 和 f_6 的损失函数加权融合,得到均衡化处理后的新损失函数 L_{Sum} :

$$L_{Sum} = L_{ef}(p_i, t_i, \hat{r}_i) + \beta L_f(p_i, t_i, r_i) \quad (6)$$

式中: β 表示均衡 $L_{ef}(p_i, t_i, \hat{r}_i)$ 和 $L_f(p_i, t_i, r_i)$ 的权重。

3 实验

为了验证所提出算法的精确性和稳健性,选择 Fddb^[18] 和 WIDER FACE^[19] 人脸检测基准数据集进行实验,选择 HR^[15]、SFD^[16] 和 PyramidBox^[17] 算法作为比较算法。实验中所有算法都在 PyCharm 上基于 python 代码实现,并且代码运行环境为配置 Nvidia GTX Titan X 的计算机。在训练时,对所有的卷积层参数都使用 Xavier 方法^[20] 随机初始化,模型优化方法采用随机梯度下降,批大小设置为 32,权重衰减设置为 0.0005,动量设置为 0.9,最大迭代次数设置为 12×10^4 ,前 8×10^4 次迭代,学习率设置为 10^{-3} ,后 4×10^4 次迭代,学习率设置为 10^{-4} 。文中算法使用 WIDER FACE 的训练集进行训练,并在两个数据集上进行测试。

3.1 Fddb 数据集测试结果分析

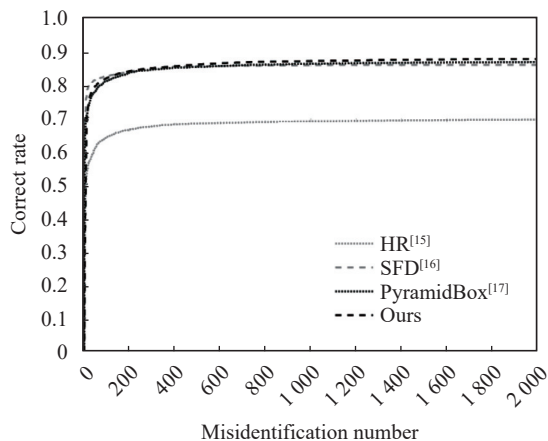
首先,在 Fddb 数据集上测试在 WIDER FACE 数据集上训练好的模型。Fddb 数据集共有 2845 个图像,其中标记了 5171 张具有模糊,遮挡,低分辨率和姿势变化的人脸。在测试之前,需要将 Fddb 数据集的人脸椭圆标注转化为矩形框。

文中算法与 HR, SFD 和 PyramidBox 算法在 Fddb 数据集上的连续 ROC 曲线和离散 ROC 曲线如图 5 所示。由图 5 可知,连续 ROC 曲线在离散 ROC 曲线之下,因为连续 ROC 曲线关注于 IoU 越大越好,而离散 ROC 曲线仅关注 IoU 是否大于 0.5。Fddb 数据集上的检测精度可通过不同误识别数下的正确率来描述,在低误识别数时,文中算法的正确率略低于 SFD 算法,在高误识别数时,文中算法的正

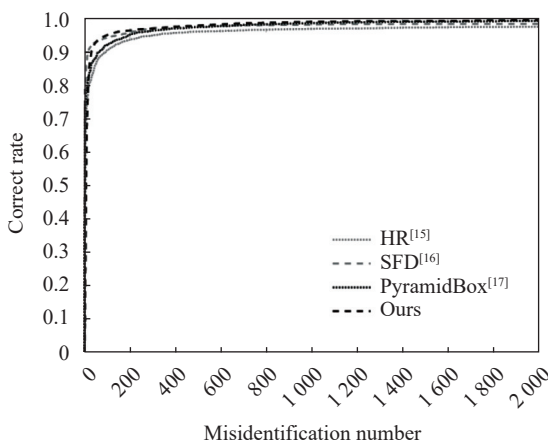
准确率均高于其他 3 种主流人脸检测算法。这说明文中算法通过引入特征图增强、负样本筛选和基于锚框尺寸的损失函数,进一步提高了人脸检测的精确性。

中选择 WIDER FACE 的测试子集进行实验,实验结果如图 6 所示。

观察图 6 可以发现,对于简单、适中和困难这三



(a) 连续 ROC 曲线
(a) Continuous ROC curve



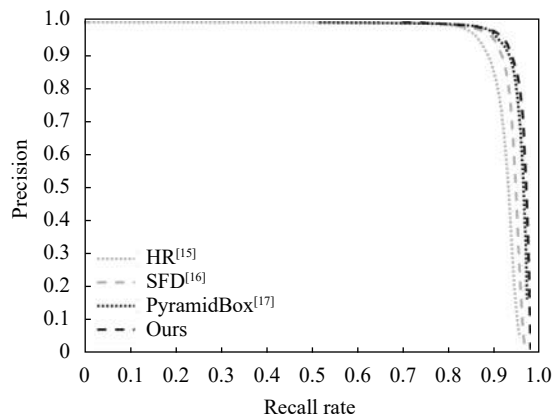
(b) 离散 ROC 曲线
(b) Discrete ROC curve

图 5 文中算法与 HR、SFD 和 PyramidBox 算法在 FDDB 数据集上的 ROC 曲线

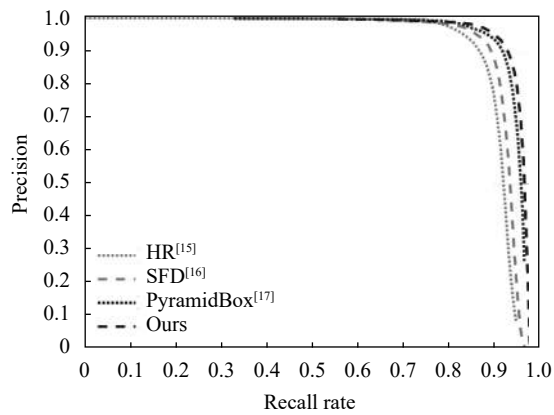
Fig.5 ROC curves of HR, SFD, PyramidBox and proposed algorithm on FDDB dataset

3.2 WIDER FACE 数据集测试结果分析

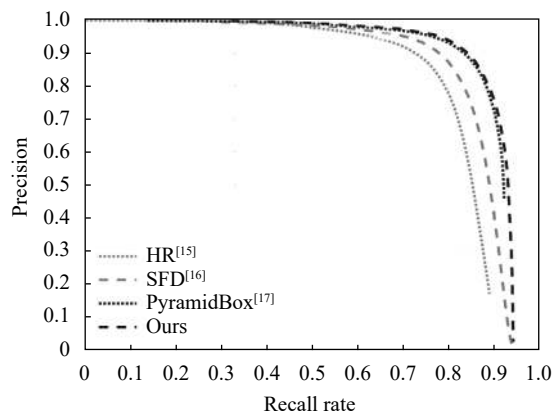
WIDER FACE 数据集共有 32203 个图像,其中标记了 393703 张具有模糊、多尺度、多姿态、局部遮挡和不同光照等多种要素的人脸。该数据集共包含 61 种情景,每种情景分别随机选择 40%、10%、50% 的数据作为训练,验证和测试子集。其中,验证子集和测试子集根据检测难度,被划分为 3 种难度级别,分别是简单、适中和困难。WIDER FACE 数据集上的检测精度为真正人脸占全部检测人脸的比例。文



(a) 简单
(a) Simple



(b) 适中
(b) Middle



(c) 困难
(c) Difficult

图 6 文中算法与 HR、SFD 和 PyramidBox 算法在 WIDER FACE 测试子集上的精度和召回率曲线

Fig.6 Precision-Recall curves of HR, SFD, PyramidBox and proposed algorithm on WIDER FACE test subset

个难度级别的测试,文中算法的精度和召回率曲线(Precision-Recall, PR)都在其他 3 种主流人脸检测算法之上。根据 PR 曲线的定义,如果一种检测算法的 PR 曲线高于另外一种检测算法,则前者具有更强的

检测性能,这说明文中算法具有比其他 3 种对比算法更强的检测性能。文中算法在 WIDER FACE 测试子集上的部分检测结果如图 7 所示。由图 7 可知,文中算法实现了较好的小尺寸人脸检测效果。



图 7 文中算法在 WIDER FACE 测试子集上的部分检测结果

Fig.7 Partial detection results of proposed algorithm on WIDER FACE test subset

4 结 论

文中提出了一种基于增强卷积神经网络的尺度不变人脸检测方法。该方法在 SSD 基础检测网络的两个浅层特征图上,通过协调聚合当前层特征图和前后两层特征图增强了当前层特征图的辨识度。接着对经过增强的两个特征图进行负样本筛选,降低了由小尺寸锚框引起的人脸检测假正率升高。最后,为原始特征图和增强特征图设置了两种基于锚框尺寸的损失函数以增强对小尺寸人脸的检测效果。在 FDDB 和 WIDER FACE 数据集上的测试结果表明,文中算法的检测精度高于当前主流的人脸检测方法。下一步将注重于增强网络模型对人脸可见区域的特征响应,以提升文中方法对遮挡人脸的检测效果。

参考文献:

- [1] Bruno R. Humans are visual experts at unfamiliar face recognition [J]. *Trends in Cognitive Sciences*, 2018, 22(6): 471-472.
- [2] Dadgostar F, Sarrafzadeh A. Gesture-based human-machine interfaces: a novel approach for robust hand and face tracking [J]. *Iran Journal of Computer Science*, 2018, 1(1): 47-64.
- [3] Lai H, Xiao S, Pan Y, et al. Deep recurrent Regression for facial landmark detection [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(5): 1144-1157.
- [4] Li H Y, Dong W M, Hu B G. Facial image attributes transformation via conditional recycle generative adversarial networks [J]. *Journal of Computer Science & Technology*, 2018, 33(3): 511-521.
- [5] Jiang H, Learned M E. Face detection with the faster R-CNN[C]//International Conference on Automatic Face Gesture Recognition, 2016: 650-657.
- [6] Yan Mengkai, Qian Jianjun, Yang Jian. Weakly aligned cross-spectral face detection [J]. *Acta Automatica Sinica*, 2022, 48(8): 1-13.
- [7] Zhou Yan, Xiang Yang, Zeng Fanzhi, et al. Tiny face recognition algorithm based on super-resolution reconstruction [J]. *Journal of Chinese Computer Systems*, 2020, 11: 2340-2346.
- [8] Cheng Shuhong, Ma Xiaofei, Zhang Shijun, et al. Smoking detection algorithm based on multitask classification [J]. *Acta Metrologica Sinica*, 2020, 41(5): 538-543.
- [9] Viola P, Jones M J. Robust real-time face detection [J]. *International Journal of Computer Vision*, 2004, 57(2): 137-154.

- [10] Najibi M, Samangouei P, Chellappa R, et al. SSH: Single stage headless face detector[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 4885-4894.
- [11] Yang S, Luo P, Loy C C, et al. Faceness-Net: Face detection through deep facial part responses [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(8): 1845-1859.
- [12] Li H, Lin Z, Shen X, et al. A convolutional neural network cascade for face detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE Computer Society, 2015: 5325-5334.
- [13] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(6): 1137-1149.
- [14] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//Proceedings of the European Conference on Computer Vision, 2016: 21-37.
- [15] Hu P, Ramanan D. Finding tiny faces[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1522-1530.
- [16] Zhang S, Zhu X, Lei Z, et al. S³FD: Single shot scale-invariant face detector[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 192-201.
- [17] Tang X, Du D K, He Z, et al. Pyramidbox: A context-assisted single shot face detector[C]//Proceedings of the European Conference on Computer Vision, 2018: 797-813.
- [18] Jain V, Learned-miller E. Fddb: A benchmark for face detection in unconstrained settings[R]. Amherst: University of Massachusetts, 2010.
- [19] Yang S, Luo P, Chen C L, et al. WIDER FACE: A face detection benchmark[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016: 5525-5533.
- [20] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[C]//Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010: 249-256.