

## 基于汇聚级联卷积神经网络的旋转人脸检测方法

齐悦<sup>1</sup>, 董云云<sup>2</sup>, 王溢琴<sup>3\*</sup>

1. 太原开放大学 计算机网络中心, 山西 太原 030024;
2. 太原理工大学 软件学院, 山西 太原 030600;
3. 晋中学院 信息技术与工程系, 山西 晋中 030619)

**摘要:** 针对大规模姿态变化和大角度人脸平面旋转 (Rotation-in-Plane, RIP) 等复杂条件下, 多尺度旋转人脸检测精度低的问题, 提出了一种基于汇聚级联卷积神经网络 (Convolutional Neural Networks, CNN) 的旋转人脸检测方法。采用由粗到精的级联策略, 在主网络 SSD 的多个特征层上汇聚级联了多个浅层的卷积神经网络, 逐步完成人脸/非人脸检测、人脸边界框位置更新和人脸 RIP 角度估计。该方法在 Rotate Fddb 和 Rotate Sub-WIDER FACE 数据集上取得了较好的检测效果。在 Rotate Sub-WIDER FACE 数据集出现 100 次误报时的检测精度为 87.1%, 速度为 45 FPS, 证明该方法可在低时间损耗下完成精确的旋转人脸检测。

**关键词:** 旋转人脸检测; 汇聚级联; 卷积神经网络; 尺度变换; 平面旋转

**中图分类号:** TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20220176

## Rotating face detection based on convergent cascaded convolutional neural network

Qi Yue<sup>1</sup>, Dong Yunyun<sup>2</sup>, Wang Yiqin<sup>3\*</sup>

1. Computer Network Center, Taiyuan Open University, Taiyuan 030024, China;
2. College of Software, Taiyuan University of Technology, Taiyuan 030600, China;
3. Department of Information Technology and Engineering, Jinzhong University, Jinzhong 030619, China)

**Abstract:** To solve the problem of low accuracy of multi-scale rotating face detection under complex conditions such as large-scale pose change and large-angle face rotation-in-plane, a rotating face detection method based on parallel cascade convolution neural network is proposed. Using a coarse-to-fine cascading strategy, multiple shallow convolutional neural networks are cascaded in parallel on multiple feature layers of the main network SSD. Face/non-face detection, face boundary box position update and face RIP angle estimation are gradually completed. Experimental results on Rotate Fddb dataset and Rotate Sub-WIDER FACE dataset show that the proposed method achieves advanced face detection. The detection precision of the method is 87.1% and the speed is 45 FPS when 100 false positives occur in the rotating Sub-WIDER FACE dataset, which proves that the method can achieve accurate rotating face detection with low time loss.

**Key words:** rotating face detection; parallel cascade; CNN; scale transformation; rotation-in-plane

收稿日期:2022-03-13; 修订日期:2022-03-30

基金项目:山西省高等学校科技创新项目 (2021L490); 山西省教育科学“十四五”规划课题 (GH-21105)

作者简介:齐悦,男,副教授,硕士,主要从事计算机视觉、深度学习方面的研究。

通讯作者:王溢琴,女,副教授,硕士,主要从事图像处理、深度学习方面的研究。

## 0 引言

近年来,伴随着深度学习技术的广泛应用,基于卷积神经网络的人脸检测技术也得到了快速发展,其被广泛应用于人脸识别<sup>[1]</sup>、人脸对齐<sup>[2]</sup>和人脸超分辨率重建<sup>[3]</sup>等研究领域。虽然近年来人脸检测的精度不断提高,但多尺度旋转人脸检测的问题还没有得到较好解决。

多尺度旋转人脸检测<sup>[4-5]</sup>是指在输入图像中检测出具有尺度变化和任意角度平面旋转的人脸,在刷脸闸机通行、人脸注册登录、智能视频监控、互动娱乐美颜等场景得到了广泛应用。传统基于卷积神经网络(Convolutional Neural Networks, CNN)的人脸检测方法<sup>[6-8]</sup>受益于 CNN 自动高效的非线性特征提取能力,可以实现快速准确的人脸检测,但是在处理具有尺度变化和平面旋转的人脸图像时,受多尺度和多角度条件下人脸特征差异明显的影响,模型训练收敛速度慢,检测精度一般。因此,最简单的方法就是在不同平面旋转角下对训练样本进行扩充,形成包含各个平面旋转角度的训练数据,然后再训练人脸检测模型。但是,更大规模的训练样本会在模型训练过程中降低分类器的性能,导致人脸检测时间增加<sup>[9-10]</sup>。参考文献[11]通过训练多个覆盖各种角度的分类器完成旋转人脸的检测,每一个分类器仅负责小范围内的角度变化,降低了模型的收敛难度。参考文献[12]首先分别训练眼、鼻、嘴等四个检测器,然后组合起来构建目标区域以进行人脸检测,但是训练多个检测器会增加耗时并且对检测精度提高有限。参考文献[13]提出了一种渐进式校准网络(Progressive Calibration Networks, PCN)以进行旋转人脸检测,该方法采用三级级联的方式逐步估计人脸旋转角度,虽然检测速度

很快,但是检测精度不高,需要在速度和精度上进行权衡调整。参考文献[14]提出了一种基于通用目标检测框架的人脸检测方法,在准确率上有所提升,但计算消耗比较大。参考文献[15]提出了基于单射多尺度检测器(Single Shot MultiBox Detector, SSD)的人脸检测算法,该算法直接在多尺度卷积层进行人脸/非人脸鉴别和人脸框位置调整,可快速高效地完成多尺度人脸检测,进而采用 SSD 人脸检测模型解决各种人脸检测问题迅速成为研究热点。

文中汲取了 SSD 人脸检测模型精准高效的特点,同时针对多种复杂环境下的多尺度旋转人脸检测问题进行了以下改进:(1)采用由粗到精的级联策略,在主网络 SSD 的多个特征层上级联多个浅层的卷积神经网络,逐步完成人脸/非人脸鉴别、人脸 RIP 角度估计和人脸框位置调整;(2)在主干网络的卷积特征层,通过汇聚级联多个浅层人脸分类网络和人脸回归网络,加快对旋转人脸的检测实时性;(3)对浅层的卷积神经网络采用不同的分辨率设计,可以实现比单分辨率 CNN 更强的检测能力。

## 1 SSD 多尺度人脸检测模型

相比于 Faster-RCNN 等需要进行多个人脸检测阶段的算法,SSD 算法可在单阶段检测中同时进行目标类别预测和候选框位置回归,进一步提高了检测速度。基于 SSD 算法的多尺度人脸检测模型如图 1 所示,该模型采用 3×3 尺寸卷积核对每一层特征层进行遍历卷积,在加速检测的同时,实现了多尺度检测效果。当输入图像包含被检测人脸时,经过多个特征层的特征提取,可基于每一特征层生成不同尺度的离散人脸边界框和类别得分,最终经过非极大值抑制处理,输出检测人脸。

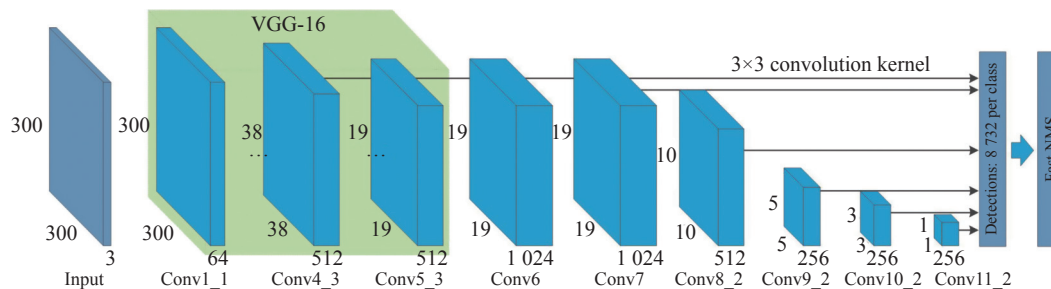


图 1 基于 SSD 算法的多尺度人脸检测模型

Fig.1 Multi-scale face detection model based on SSD algorithm

SSD 算法采用倒金字塔式的特征图检测来实现多尺度人脸检测,从浅层特征图 Conv4\_3、Conv7、Conv8\_2到深层特征图Conv9\_2、Conv10\_2、Conv11\_2,特征图尺寸不断减小。以  $38 \times 38 \times 512$  尺寸的特征图为例,采用  $3 \times 3 \times 512$  大小的卷积核对其进行卷积,得到人脸/非人脸得分和人脸边界框的位置回归。假设  $38 \times 38 \times 512$  尺寸的特征图上有四种不同比例的检测框,那么该特征图就会产生  $38 \times 38 \times 4$  个默认检测框,并且每个默认检测框会被预测出一个人脸/非人脸得分和四个位置偏移量  $(cx, cy, w, h)$ 。

SSD 算法在训练时需要将真实标注的位置信息和人脸/非人脸得分进行预处理,并将其对应到每个默认检测框上。首先计算默认框  $R_d$  和真实框  $R_g$  的交并比,接着取交并比大于 50% 的默认框放入候选正样本集,取交并比小于 50% 的默认框经策略筛减后放入候选负样本集。

$$IoU = \frac{R_d \cap R_g}{R_d \cup R_g} \quad (1)$$

## 2 基于汇聚级联卷积神经网络的多尺度旋转人脸检测模型

### 2.1 整体框架

所提出的基于汇聚级联卷积神经网络的多尺度旋转人脸检测模型如图 2 所示,它的主网络部分采用 SSD 在不同卷积层上提取多个尺度的特征图,并采用密集连接的策略生成候选人脸特征图集合,作为附加网络部分的输入。附加网络部分由两个浅层的人脸分类网络 24-classification-net 和 12-classification-net,以及两个浅层的人脸回归网络 46-regression-net 和 22-regression-net 组成。人脸分类网络和人脸回归网络内部是平行关系,人脸分类网络与人脸回归网络之间是级联关系。主网络和附加网络中采用汇聚级联的设置方法,可以进一步加快对旋转人脸的检测速度。

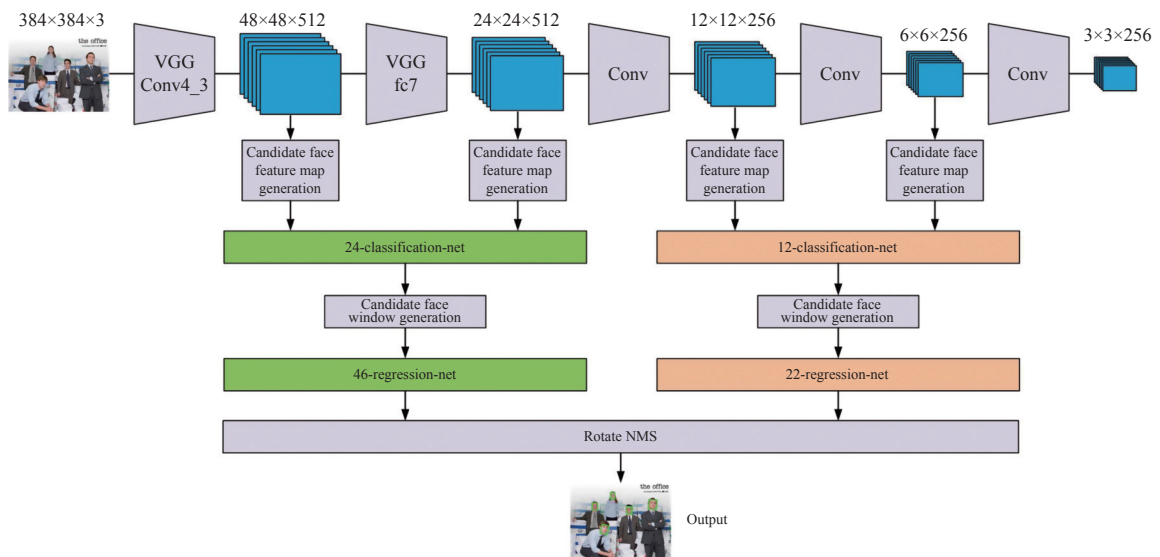


图 2 基于汇聚级联卷积神经网络的多尺度旋转人脸检测模型

Fig.2 Multi-scale rotating face detection model based on convergent cascaded convolutional neural network

对于浅层人脸分类网络 24-classification-net,它的输入是缩放为  $24 \times 24 \times 512$  大小的候选人脸特征图。在该网络中直接完成人脸/非人脸检测、人脸边界框位置更新和人脸 RIP 角度粗估计。通过 24-classification-net 后,消除了大部分置信度较低的候选人脸窗,大大降低了后续人脸回归网络的检测工作量。为了进一步加快速度,24-classification-net 仅粗略估计候

选人脸的 RIP 主方向,输出  $[-45^\circ, 45^\circ]$ 、 $(45^\circ, 135^\circ]$ 、 $(135^\circ, 225^\circ]$ 和 $(225^\circ, 315^\circ]$ 四个主方向区间,并将输出的人脸边界框旋转到 $[-45^\circ, 45^\circ]$ 区间内,该过程通过直接对人脸边界框旋转 $90^\circ$ 、 $180^\circ$ 或者 $270^\circ$ 度完成。在 24-classification-net 之后,利用更新的人脸边界框在原始图像上采样三通道的候选人脸窗口作为 46-regression-net 的输入,并将候选人脸窗口缩放为

46×46×3 大小。在 46-regression-net 中,进一步完成人脸/非人脸检测、人脸边界框位置更新和人脸 RIP 角度精确回归,回归是指在[-45°, 45°]区间内连续估计的人脸 RIP 角度。在 24-classification-net 和 46-regression-net 之后,采用旋转 NMS 合并高重叠率的人脸边界框。

另一条平行线路将 SSD 的两个更深层特征图与 12-classification-net 和 22-regression-net 级联。与 24-classification-net 和 46-regression-net 的组合类似, 12-classification-net 和 22-regression-net 仅在输入尺寸和网络深度上略有调整。

### 2.2 候选人脸特征图生成

对于一个  $m \times n \times k$  尺寸的特征图,在每一个单元格内生成四个固定大小的矩形框,一共可以产生  $m \times n \times 4$  个矩形框。这些矩形框内包围的特征图区域,

经过缩放后产生候选人脸特征图作为下一层人脸分类网络的输入。

以图 2 中  $12 \times 12 \times 256$  尺寸的特征图为例,每个单元格对应四个固定大小的矩形框,所有单元格共可以产生  $12 \times 12 \times 4 = 576$  个矩形框。将这些矩形框内包围的特征图区域缩放为  $12 \times 12 \times 256$  尺寸后,就产生了候选人脸特征图集合,该集合中共包含 576 个候选人脸特征图。

### 2.3 人脸分类网络

人脸分类网络包括 24-classification-net 和 12-classification-net,它们之间是平行关系。这两个网络的结构如图 3 和图 4 所示。图中,“Conv”、“MP”、“FC”和“ReLU”分别表示卷积层、最大池化层、全连接层和 ReLU 层,并且在卷积层要进行补边操作。

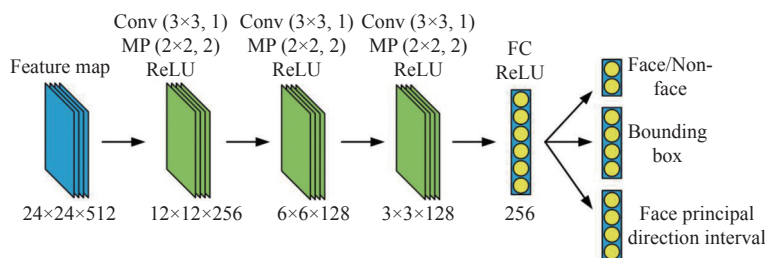


图 3 24-classification-net 结构

Fig.3 24-classification-net structure

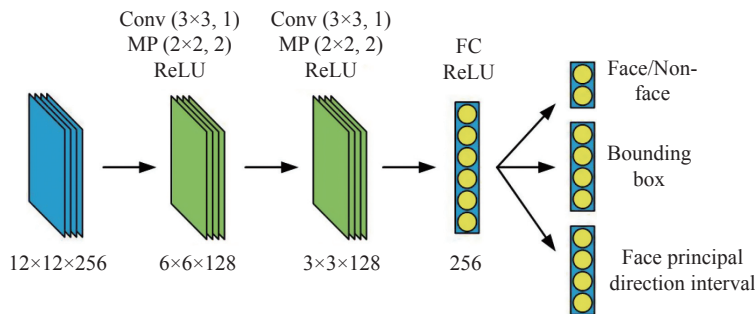


图 4 12-classification-net 结构

Fig.4 12-classification-net structure

对于输入的每一幅候选人脸特征图  $w$ ,人脸分类网络只完成三个目标:人脸/非人脸检测,人脸边界框位置更新和人脸 RIP 角度粗估计,归纳如下:

$$[c, t, d] = F(w) \tag{2}$$

式中:  $F$  表示人脸分类网络模型;  $c$  表示人脸置信度得分;  $t$  表示人脸边界框的回归向量;  $d$  表示人脸主方向

粗估计得分。

对于人脸/非人脸检测,设置 softmax 损失函数:

$$L_{Face} = y \log c + (1 - y) \log(1 - c) \tag{3}$$

式中:当输入的候选人脸特征图  $w$  是人脸时,  $y = 1$ , 否则  $y = 0$ 。

对于人脸边界框位置更新,设置第二个损失函数:

$$L_{BBox}(t, t^*) = S(t - t^*) \quad (4)$$

式中:  $t$ 和 $t^*$ 分别表示预测人脸边界框和真实人脸标注框;  $S$ 表示Smooth L1函数。

人脸主方向粗估计以离散四元分类的方式快速估计人脸主方向,其损失函数为:

$$L_{Pose} = - \sum_{i=1}^4 y_i \ln d_i \quad (5)$$

式中:  $i$ 表示人脸主方向标签, 1~4 分别表示正向脸、右向脸、反向脸和左向脸。当 $y_i$ 对应正确的人脸主方向时,  $y_i = 1$ , 其余都等于 0。

将以上三个目标函数组合, 得到人脸分类网络的总损失函数:

$$\min_F L = L_{Face} + k_1 L_{BBox} + k_2 L_{Pose} \quad (6)$$

式中:  $k_1$ 和 $k_2$ 为平衡 $L_{Face}$ 、 $L_{BBox}$ 和 $L_{Pose}$ 的权重。

通过优化公式 (6), 过滤掉大部分低置信度候选人脸窗。对于保留的候选人脸窗, 先对其进行位置更新, 接着根据粗估计的人脸主方向对其进行旋转。人脸分类网络粗估计的人脸 RIP 角度 $\theta_1$ 计算如下:

$$\theta_1 = \begin{cases} 0^\circ, & id = 1 \\ -90^\circ, & id = 2 \\ -180^\circ, & id = 3 \\ -270^\circ, & id = 4 \end{cases} \quad (7)$$

式中:  $d_i$ 表示人脸主方向粗估计得分。根据分类得到

的 $id$ 不同, 候选人脸窗相应旋转 $0^\circ$ 、 $-90^\circ$ 、 $-180^\circ$ 和 $-270^\circ$ 。经过人脸分类网络之后, 人脸 RIP 角度的可能存在范围从 $360^\circ$ 减小到 $-45^\circ \sim 45^\circ$ 之间。

在人脸分类网络的训练阶段, 由于大部分人脸数据集中包含的人脸都是正向脸, 需要在  $360^\circ$ 范围内对人脸数据集进行旋转扩充。对于扩充后的人脸数据集中的人脸, 当它的 RIP 角度在 $[-45^\circ, 45^\circ]$ 内时, 定义为正向脸; 在 $[45^\circ, 135^\circ]$ 内时, 定义为右向脸; 在 $[135^\circ, 225^\circ]$ 内时, 定义为反向脸; 在 $[225^\circ, 315^\circ]$ 内时, 定义为左向脸。此外, 还需要定义三种类型的人脸窗口: 正样本、负样本和中间样本。正样本与人脸标注框的 $IoU \geq 0.7$ , 负样本与人脸标注框的 $IoU \leq 0.3$ , 中间样本与人脸标注框的 $0.3 < IoU < 0.7$ 。正负样本用以进行人脸得分计算, 正样本和中间样本用以进行人脸边界框位置更新。

### 2.4 人脸回归网络

经过人脸分类网络 24-classification-net 和 12-classification-net 处理之后, 利用旋转更新的人脸边界框在原始图像上采样三通道的候选人脸窗, 此时所有候选人脸窗都处于 $[-45^\circ, 45^\circ]$ 区间内。对候选人脸窗进行缩放后(缩放尺寸为 $46 \times 46 \times 3$ 或 $22 \times 22 \times 3$ ), 送入平行设置的人脸回归网络 46-regression-net 和 22-regression-net, 其网络结构如图 5 和图 6 所示, 需要注意的是人脸回归网络的卷积层不执行边缘填充操作。

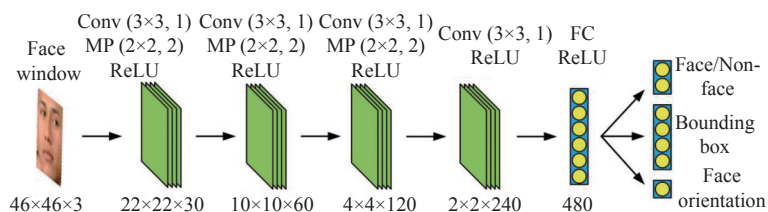


图 5 46-regression-net 结构

Fig.5 46-regression-net structure

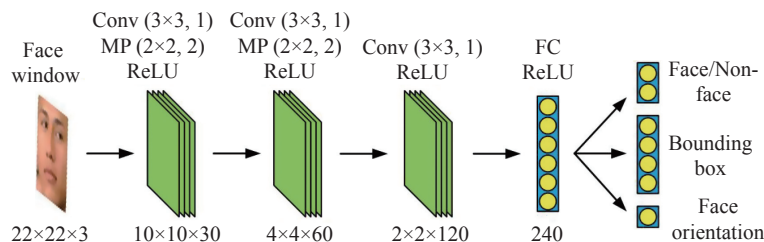


图 6 22-regression-net 结构

Fig.6 22-regression-net structure

人脸回归网络在人脸分类网络的基础上更准确地进行人脸得分计算和人脸边界框位置更新。由于人脸分类网络已将人脸 RIP 角度范围缩小到 $[-45^\circ, 45^\circ]$ 区间内,人脸回归网络连续回归候选人脸窗的精确 RIP 角度 $\theta_2$ ,  $\theta_2$ 的连续回归可采用参考文献 [16] 的级联回归模式。

结合人脸分类网络估计的 RIP 角度 $\theta_1$ 和人脸回归网络计算的 RIP 角度 $\theta_2$ , 得到候选人脸窗的总 RIP 角度 $\theta_{all}$ :

$$\theta_{all} = \theta_1 + \theta_2 \quad (8)$$

在人脸回归网络的训练阶段,需要在 $[-45^\circ, 45^\circ]$ 范围内均匀旋转三通道的候选人脸窗集合以生成训练样本。

### 2.5 旋转 NMS

将人脸回归网络检测的人脸边界框按照得分进行降序排序。对此刻遍历的人脸边界框与剩余的人脸边界框执行交集运算,得到相应的相交点集合,通过计算相交点集合包围的凸多边形面积,得到每两个人脸边界框的 *IoU*, 去掉大于设定阈值的人脸边界框,保留小于设定阈值的人脸边界框,以得到最终的多尺度旋转人脸检测结果。

## 3 实验及结果分析

### 3.1 网络设置

文中算法基于 TensorFlow 深度学习框架实现。为对比文中算法的有效性和优越性,设置 PCN<sup>[13]</sup>、Faster R-CNN<sup>[14]</sup>、SSD<sup>[15]</sup> 和 Cascade CNN<sup>[16]</sup> 等多个对比算法。训练汇聚级联卷积网络所需的数据在经过旋转的 Fddb 人脸数据集<sup>[17]</sup> 和 WIDER FACE 人脸数据集<sup>[18]</sup> 上采集,并且人脸真实标注都被调整为正方形。网络采用反向传播的随机梯度下降 (Stochastic Gradient Descent, SGD) 进行优化,最大迭代次数设置为 $10^5$ ,前 $7.5 \times 10^4$ 次迭代,学习率设置为 $10^{-3}$ ,后 $2.5 \times 10^4$ 次迭代,学习率设置为 $10^{-4}$ 。权重衰减系数设置为 0.0005,动量参数设置为 0.9,所有层均由零高斯分布初始化,标准差设置为 0.01,每批次包含正样本,负样本和中间样本的比例约为 3:3:2。

### 3.2 Rotate Fddb 数据集测试结果分析

Fddb 人脸数据集在具有遮挡、大角度偏转、低

分辨率和模糊对焦等复杂场景中标记了 5171 个人脸框。为了更好地评估各个算法对旋转人脸的检测性能,对 Fddb 人脸数据集中的图像分别进行了 $90^\circ$ 、 $180^\circ$ 和 $270^\circ$ 旋转,得到 Rotate Fddb 数据集。

整理文中算法和对比算法在 Rotate Fddb 数据集上的实验结果,形成 ROC 曲线如图 7 所示。通过分析各算法的 ROC 曲线走向可知,文中算法的检测性能优于 SSD 和 Cascade CNN 算法,在假正率较低时,文中算法的检测性能比 PCN 和 Faster R-CNN 更强,在假正率较高时,文中算法的检测性能与 PCN 和 Faster R-CNN 基本持平。这说明在主干网络的卷积特征层,通过汇聚级联多个浅层人脸分类网络和人脸回归网络,逐步完成人脸得分计算、人脸 RIP 角度估计和人脸边界框位置更新,可提升精确检测旋转人脸的能力。

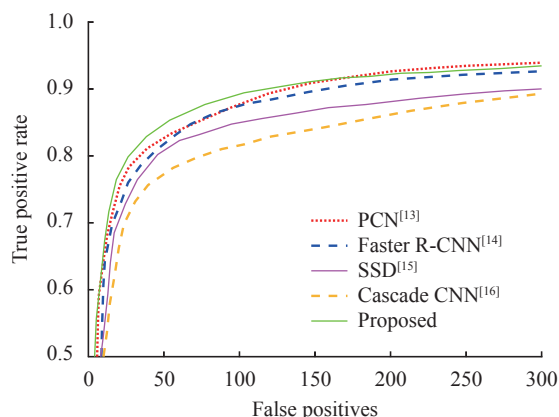


图 7 Rotate Fddb 数据集上的 ROC 曲线对比

Fig.7 Comparison of ROC curves on the Rotate Fddb dataset

### 3.3 Rotate Sub-WIDER FACE 数据集测试结果分析

WIDER FACE 人脸数据集在具有遮挡、模糊对焦、平面旋转和尺度姿态变化等多种复杂场景中标记了 393703 个人脸框,是目前人脸检测领域最具挑战性的基准数据集。从该数据集中选择部分包含平面旋转人脸的图片,建立了一个包含 391 张图片和 1027 个旋转人脸的 Rotate Sub-WIDER FACE 数据集。

整理文中算法和对比算法在 Rotate Sub-WIDER FACE 数据集上的实验结果,形成 ROC 曲线如图 8 所示。通过分析各算法的 ROC 曲线走向可知,与其他四种旋转人脸检测算法相比,文中算法达到了最佳的检测效果,进一步证明了所提出的汇聚级联卷积神经

网络的有效性。图 9 为文中算法在 Rotate Sub-WIDER FACE 数据集中的部分检测效果图。

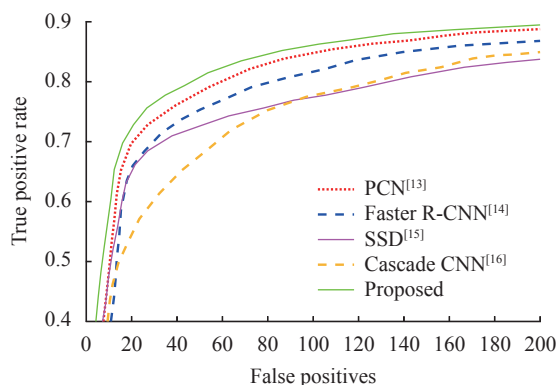


图 8 Rotate Sub-WIDER FACE 数据集上的 ROC 曲线对比

Fig.8 Comparison of ROC curves on the Rotate Sub-WIDER FACE dataset



图 9 文中算法在 Rotate Sub-WIDER FACE 数据集上的部分检测效果图

Fig.9 Partial detection effect diagram of proposed algorithm on the Rotate Sub-WIDER FACE dataset

### 3.4 速度和精度对比

文中算法旨在低时间损耗下完成精确的旋转人脸检测。为了定量比较文中算法与对比算法的精度和速度,选择 Rotate Sub-WIDER FACE 数据集进行实验,并统计文中算法与 PCN、Faster R-CNN、SSD 和 Cascade CNN 等算法在 100 次误报时的精度与速度,如表 1 所示。

由表 1 可以看出,文中算法的运行速度几乎与 PCN 和 Cascade CNN 相同,且远高于 Faster R-CNN 和 SSD 算法,这主要受益于文中网络采用的汇聚级联框架和人脸分类网络对人脸主方向的快速分类。此外,人脸回归网络对人脸 RIP 角度的连续精确回归

表 1 文中算法与 PCN、Faster R-CNN、SSD 和 Cascade CNN 等算法在 100 次误报时的速度和精度对比

Tab.1 Comparison of speed and precision of PCN, Faster R-CNN, SSD, Cascade CNN and proposed algorithm in 100 false positives

Method	Precision	Speed/FPS
PCN <sup>[13]</sup>	0.852	47
Faster R-CNN <sup>[14]</sup>	0.818	13
SSD <sup>[15]</sup>	0.771	24
Cascade CNN <sup>[16]</sup>	0.778	51
Proposed algorithm	0.871	45

也使文中算法的检测精度高于其他四种对比算法,进一步证明文中算法具有较高的精确性和实时性。

## 4 结束语

文中提出的基于汇聚级联卷积神经网络的多尺度旋转人脸检测方法采用由粗到精的级联策略,在主网络 SSD 的多个特征层上汇聚级联了多个浅层的卷积神经网络,逐步完成人脸/非人脸检测、人脸边界框位置更新和人脸 RIP 角度估计。同时,浅层的卷积神经网络还采用了多分辨率设计,可以实现比单分辨率 CNN 更强的检测能力。所提出的方法在保持低时间损耗的同时,进一步提高了多尺度旋转人脸的检测精度,同时在 Rotate Fddb 和 Rotate Sub-WIDER FACE 数据集上也取得了较好的检测效果。

### 参考文献:

- [1] Guo J, Zhu X, Zhao C, et al. Learning meta face recognition in unseen domains [C]//Computer Vision and Pattern Recognition, IEEE, 2020: 6163-6172.
- [2] Liu H Z, Yang S P, Yuan J Z, et al. Multi-scale face detection based on single neural network [J]. *Journal of Electronics and Information Technology*, 2018, 40(11): 63-70. (in Chinese)
- [3] Fang Y, Chi Z, Yang W, et al. Blind visual quality assessment for image super-resolution by convolutional neural network [J]. *Multimedia Tools & Applications*, 2018, 77(10): 29829-29846.
- [4] Xiao Y, Cao D, Gao L. Face detection based on occlusion area detection and recovery [J]. *Multimedia Tools and Applications*, 2020, 79(2): 16531-16546.
- [5] Ma D, Lai H C. In-plane rotation face detection based on improved SURFs [J]. *Computer Engineering and Design*, 2018,

- 39(9): 125-129. (in Chinese)
- [6] Hu P, Ramanan D. Finding tiny faces [C]//Computer Vision and Pattern Recognition, IEEE, 2017: 1522-1530.
- [7] Rezaei M, Ravanbakhsh E, Namjoo E, et al. Assessing the effect of image quality on SSD and faster R-CNN networks for face detection [C]//27th Iranian Conference on Electrical Engineering (ICEE), IEEE, 2019: 1589-1594.
- [8] Zhang H T, Li M L, Dong S H. Two-layer cascaded convolutional neural network for face detection [J]. *Journal of Image and Graphics*, 2019, 24(2): 49-60. (in Chinese)
- [9] Hao Z, Liu Y, Qin H, et al. Scale-aware face detection [C]//Computer Vision and Pattern Recognition, IEEE, 2017: 1913-1922.
- [10] Wu S, Kan M, He Z, et al. Funnel-structured cascade for multi-view face detection with alignment-awareness [J]. *Neurocomputing*, 2017, 221(C): 138-145.
- [11] Priya G N, Banu R S D W. A robust rotation invariant multiview face detection in erratic illumination condition [J]. *International Journal of Computer Applications*, 2012, 57(20): 46-51.
- [12] Kylberg G, Sintorn I M. On the influence of interpolation method on rotation invariance in texture recognition [J]. *Eurasip Journal on Image and Video Processing*, 2016(1): 17.
- [13] Shi X, Shan S, Kan M, et al. Real-time rotation-invariant face detection with progressive calibration networks [C]//Computer Vision and Pattern Recognition, IEEE, 2018: 2295-2303.
- [14] Jiang H, Learned-Miller E. Face detection with the faster R-CNN [C]//IEEE International Conference on Automatic Face & Gesture Recognition, IEEE, 2017: 650-657.
- [15] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]//European Conference on Computer Vision, 2016: 21-37.
- [16] Li H, Lin Z, Shen X, et al. A convolutional neural network cascade for face detection [C]//Computer Vision and Pattern Recognition, IEEE, 2015: 5325-5334.
- [17] Jain V, Learned-Miller E. FDDB: A benchmark for face detection in unconstrained settings [R]. US: UMass Amherst, 2010.
- [18] Yang S, Luo P, Chen C L, et al. WIDER FACE: A face detection benchmark [C]//IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016: 5525-5533.