

# 基于增强并行级联卷积神经网络的人脸检测方法

朱富丽<sup>1</sup> 杨磊<sup>2\*</sup> 姬波<sup>3</sup>

<sup>1</sup>(河南牧业经济学院 河南 郑州 450044)

<sup>2</sup>(河南信息统计职业学院 河南 郑州 450008)

<sup>3</sup>(郑州大学信息工程学院 河南 郑州 450001)

**摘要** 针对复杂场景下,小尺度、模糊和遮挡人脸检测精度低的问题,提出一种基于增强并行级联卷积神经网络的人脸检测方法。在主网络 SSD 的多层特征图上,通过融合前后层特征图增强原始特征图的辨识度。将多个增强特征图组合成附加增强网络,与主网络并行设置,加快对小尺度,模糊和遮挡人脸的检测速度。在训练阶段为主网络和附加增强网络设置两种基于锚框尺寸的损失函数,并通过加权求和的方式对这两种损失函数进行融合。在 Fddb 和 WIDER FACE 数据集上的实验结果表明,该方法实现了比当前主流人脸检测方法更高的检测精度。

**关键词** 人脸检测 卷积神经网络 增强特征图 附加增强网络 多尺度

**中图分类号** TP391 **文献标志码** A **DOI**:10.3969/j.issn.1000-386x.2020.11.018

## FACE DETECTION METHOD BASED ON ENHANCED PARALLEL CASCADED CONVOLUTIONAL NEURAL NETWORK

Zhu Fuli<sup>1</sup> Yang Lei<sup>2\*</sup> Ji Bo<sup>3</sup>

<sup>1</sup>(Henan University of Animal Husbandry and Economics, Zhengzhou 450044, Henan, China)

<sup>2</sup>(Henan Information and Statistics Vocational College, Zhengzhou 450008, Henan, China)

<sup>3</sup>(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, Henan, China)

**Abstract** To solve the problem of low precision of small scale, blur and occlusion face in complex scenes, this paper proposes a face detection method based on enhanced parallel cascaded convolutional neural network. On the multi-layer feature map of the main network SSD, the recognition of the original feature map was enhanced by the fusion of the front and back layer feature maps; the multiple enhanced feature maps were combined into an additional enhanced network, which was set in parallel with the main network, speeding up the detection of small scale, fuzzy and occluded faces; two loss functions based on the anchor size were set for the main network and the additional enhanced network in the training phase, and the two loss functions were combined by weighted summation. The experimental results on the Fddb and WIDER FACE datasets show that the proposed method achieves higher detection accuracy than the current mainstream face detection methods.

**Keywords** Face detection Convolutional neural network Enhanced feature map Additional enhanced network Multiscale

## 0 引言

人脸检测是人脸信息处理领域的一项关键性技

术,也是人脸特征点定位<sup>[1]</sup>、人脸比对<sup>[2]</sup>、人脸识别<sup>[3]</sup>、人脸超分辨率重建<sup>[4]</sup>等任务的前置操作,其检测效果直接影响后续操作的准确性,具有重要的研究价值。文献[5]提出了基于 Haar 特征和 AdaBoost 的

人脸检测方法,取得了较好的人脸检测效果。文献[6]提出了基于树的集成学习方法来进行人脸检测,提高了检测精度,但仍受限于人工提取特征的表达能力和分类器的分类能力。目前,该方法已经被基于卷积神经网络(Convolutional Neural Networks, CNN)<sup>[7-8]</sup>的深度学习方法取代,并且取得了较大的提高。尽管基于CNN的人脸检测方法得到了广泛的关注和研究,人脸检测准确率也不断提高,但是在复杂场景下,对小尺度、模糊和遮挡人脸的检测问题仍然存在困难。

目前基于CNN的人脸检测方法主要分为两类。一类是基于Faster R-CNN<sup>[9]</sup>中的区域建议网络(Region Proposal Network, RPN),它采用两阶段检测方案,在RPN中经过端到端的训练,生成高质量的区域建议图,然后送入Faster R-CNN执行进一步检测。另一类是基于单一网络(Single Shot MultiBox Detector, SSD)的人脸检测方法<sup>[10]</sup>,它去除了RPN,直接在不同尺度的卷积层上进行人脸/非人脸判定和人脸边界框位置回归。文献[11]提出了一种利用人脸周边信息进行辅助的人脸检测方法,对小尺度人脸检测效果较好,但计算消耗较大。文献[12]通过规模补偿锚匹配策略提高了小尺度人脸的召回率,通过max-out背景标签降低了小尺度人脸的误报率。文献[13]提出了一种基于环境辅助的单一网络人脸检测方法,通过引入环境敏感预测模块提升了预测网络的人脸分类能力和最终输出的准确度。近年来,单一网络人脸检测框架因其高效的检测效率和直观的网络结构而备受关注。

本文在SSD人脸检测模型的基础上,针对复杂场景下的小尺度、模糊和遮挡人脸检测问题,进行三方面的改进:(1)提出一个特征增强网络。在主网络SSD的多层特征图上,通过融合前后层特征图增强原始特征图的辨识度,并引入监督信息去学习上下文融合特征。(2)在网络结构的设计上,集合多个增强特征图以组成附加增强网络,主网络与附加增强网络采用并行设置方式,加快对小尺度、模糊和遮挡人脸的检测速度。(3)在训练阶段,为主网络和附加增强网络设置两种基于锚框尺寸的损失函数,并通过加权求和的方式对这两种损失函数进行融合。预测阶段仅对附加增强网络的特征图进行检测,在保持人脸检测精度的同时,降低时间损耗。

## 1 SSD人脸检测模型

与Faster R-CNN等生成RPN的算法不同,SSD算法直接预测目标类别和候选框位置,具有很快的检测

速度和精度。由于图像中的人脸具有多尺度特性,传统的人脸检测算法在检测人脸时,会先建立图像金字塔,然后在不同尺度的图像上利用滑动窗口进行人脸检测。而SSD算法另辟蹊径,通过在不同尺度特征图上进行 $3 \times 3$ 大小的滑动卷积,实现了效果更佳的人脸检测效果。

基于SSD算法的人脸检测模型如图1所示,其中:Conv1\_1到Conv5\_3是VGG16的前5个卷积层;Conv6和Conv7是由2个全连接层改造的卷积层;Conv8\_2到Conv11\_2是新增的4个卷积层。该模型的输入是一幅 $300 \times 300$ 的包含单人或多人脸的图像,输出一组离散人脸边界框和人脸/非人脸得分,对所有输出的人脸边界框进行非极大值抑制(Non-Maximum Suppression, NMS)后,得到最终的人脸检测结果。SSD算法在检测人脸时,利用了不同尺度特征图的信息,被检测的特征图为conv4\_3、conv7、conv8\_2、conv9\_2、conv10\_2和conv11\_2,由图1可知,这些特征图的尺寸是依次减小的,这样就实现了多尺度人脸检测效果。

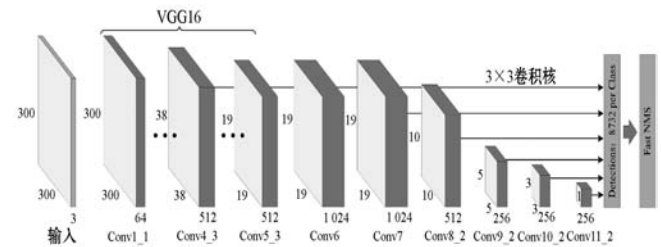


图1 基于SSD算法的人脸检测模型

该模型的损失函数包含了用于人脸/非人脸分类的log损失和用于人脸边界框回归的smooth  $L_1$ 损失,并对正负样本比例进行控制,提高了训练速度和检测结果的稳定性。

$$L_{SSD}(x, c, t, t^*) = \frac{1}{N} (L_{cont}(x, c) + \alpha L_{loc}(x, t_i, t_i^*)) \quad (1)$$

式中: $N$ 表示匹配的人脸边界框个数; $x$ 表示匹配的人脸边界框是否属于人脸,取值为 $\{0, 1\}$ ;  $c$ 表示所框选目标属于人脸的置信度; $t_i$ 表示预测人脸框; $t_i^*$ 表示真实标注框; $\alpha$ 表示平衡权重系数。

在训练时,通过计算默认检测框和真实标注框的交并比(Intersection over Union, IoU)来确定正负训练样本。取 $IoU > 0.5$ 的默认检测框为候选正样本集,取 $IoU \leq 0.5$ 的默认检测框为候选负样本集。此时,候选负样本集的数量远大于候选正样本集,这会造成正负样本数量不平衡,降低训练时的收敛速度。因此,对于候选负样本集,用实际选取要投入训练的检测框与候选负样本集匹配,剔除掉匹配失败的检测框,这是因为它们的损失过小,容易被识别为背景,对整体网络训练的提升不大。

## 2 方法设计

### 2.1 总体框架

基于增强并行级联卷积神经网络的人脸检测模型如图 2 所示,它由主网络、特征增强网络和附加增强网络三部分组成。

主网络采用 SSD 算法在 6 个连续的卷积层上提取

多尺度特征图,在不同的特征图上分别进行人脸/非人脸分类和人脸边界框回归。六个连续的卷积层为 Conv3\_3、Conv4\_3、Conv5\_3、Conv6、Conv7 和 Conv8\_2,对应提取的特征图为  $f_1$ 、 $f_2$ 、 $f_3$ 、 $f_4$ 、 $f_5$  和  $f_6$ 。由图 2 可知,输入图像的尺寸为  $384 \times 384$ ,因此从  $f_1$  到  $f_6$  的特征图尺寸分别为: $96 \times 96$ 、 $48 \times 48$ 、 $24 \times 24$ 、 $12 \times 12$ 、 $6 \times 6$  和  $3 \times 3$ 。

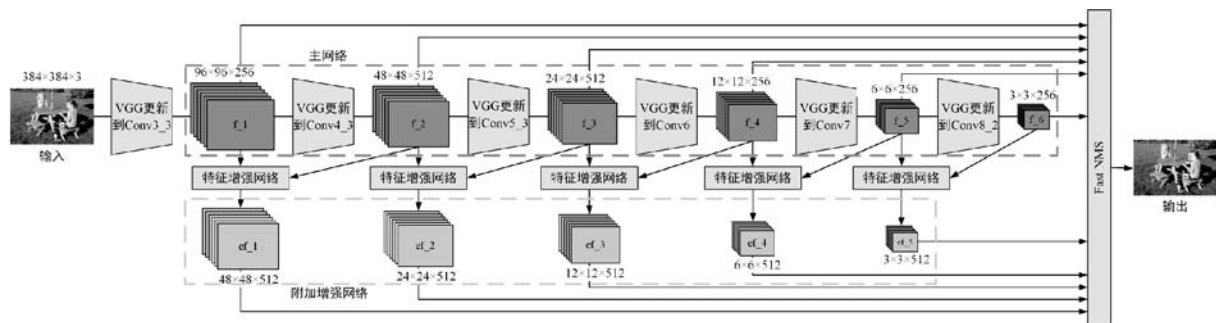


图 2 基于增强并行级联卷积神经网络的人脸检测模型

特征增强网络对当前层的特征图和后一层的特征图进行归一化、卷积、池化、融合和激活等操作后,得到当前层的增强特征图。附加增强网络的检测层是经过特征增强网络处理后得到的,它由  $ef_1$ 、 $ef_2$ 、 $ef_3$ 、 $ef_4$  和  $ef_5$  这 5 个增强特征图组成。当前层增强特征图  $ef_k$  ( $k = 1, 2, \dots, 5$ ) 的长和宽都只有当前层特征图  $f_k$  ( $k = 1, 2, \dots, 5$ ) 的一半,这是由于特征增强网络对当前层特征图进行了池化操作。另外,主网络和附加增强网络设置了并行的运行方式和不同的损失函数,进一步提高了人脸检测的速度和精度。

### 2.2 特征增强网络

设置特征增强网络的目的是增强原始特征图的辨识度,以进一步提高对小尺度、模糊和遮挡人脸的检测精度。为了增强原始特征图,特征增强网络融合了第  $k$  层特征图和第  $k + 1$  层特征图的特征信息,如图 3 所示。

维度不同,为了后续特征信息的融合,首先利用  $1 \times 1$  大小的卷积核对第  $k$  层特征图和第  $k + 1$  层特征图进行归一化处理,归一化处理后,第  $k$  层特征图和第  $k + 1$  层特征图的维度都是 512。此时,第  $k$  层特征图的长和宽都是第  $k + 1$  层特征图的 2 倍,因此还要对第  $k$  层特征图进行  $2 \times 2$  的最大池化。池化处理后两层特征图的尺寸完全一样,对其进行连接以得到  $a \times a \times 1024$  大小的融合特征图。最终利用 512 个  $3 \times 3$  大小的卷积核进行边缘补 0 的卷积,得到原始特征图  $f_k$  的增强特征图  $ef_k$ 。上述操作步骤可被归纳为:

$$ef_k = f_{Conv_3} [f_{connect} (f_{pool} (f_{Conv_1} (f_k)) + f_{Conv_1} (f_{k+1}))] \quad (2)$$

式中:  $f_{Conv_1}$  表示归一化卷积;  $f_{pool}$  表示对第  $k$  层特征图进行  $2 \times 2$  最大池化;  $f_{connect}$  表示对第  $k$  层特征图和第  $k + 1$  层特征图的连接操作;  $f_{Conv_3}$  表示  $3 \times 3$  的卷积。

### 2.3 基于锚框尺寸的损失函数

与传统目标检测算法的损失函数不同,本文算法在训练时为主网络和附加增强网络设置两种基于锚框尺寸的损失函数,并通过加权求和的方式对这两种损失函数进行融合。文献[9]指出,在较浅层特征图上更适合进行较小尺度的人脸检测,因此在设置主网络损失函数时引入较小尺度的锚框  $\hat{a}_i$  :

$$L_{M-Net}(p_i, t_i, \hat{a}_i) = \frac{1}{N_{cls}} (\sum_i L_{cls}(p_i, p_i^*)) + \frac{\lambda}{N_{reg}} (\sum_i p_i^* L_{reg}(t_i, t_i^*, \hat{a}_i)) \quad (3)$$

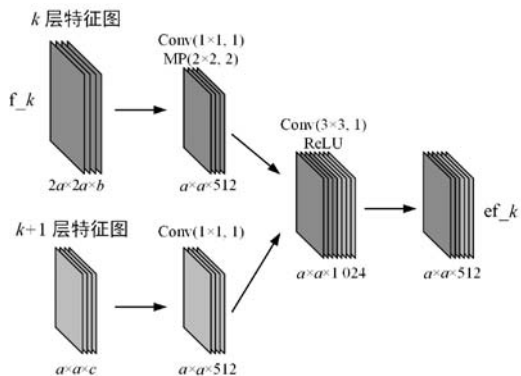


图 3 特征增强网络结构

由图 3 可知,第  $k$  层特征图与第  $k + 1$  层特征图的

式中:  $N_{\text{cls}}$  表示正样本锚框和负样本锚框的数量;  $N_{\text{reg}}$  表示正样本锚框的数量;  $L_{\text{cls}}$  表示人脸/非人脸分类的 Softmax 损失;  $L_{\text{reg}}$  表示使用锚框  $\hat{a}_i$  时预测人脸框  $t_i$  和真实标注框  $t_i^*$  之间的参数化 Smooth  $L_1$  损失;  $p_i$  表示预测为人脸的概率;  $p_i^*$  表示正样本和负样本的标签, 当  $p_i^* = 1$  时,  $L_{\text{reg}}$  被激活;  $\lambda$  作为权重来平衡  $L_{\text{cls}}$  和  $L_{\text{reg}}$ 。

在设置附加增强网络的损失函数时, 引入较大尺寸的锚框  $\hat{a}_i$  :

$$L_{\text{AE-Net}}(p_i, t_i, \hat{a}_i) = \frac{1}{N_{\text{cls}}} \left( \sum_i L_{\text{cls}}(p_i, p_i^*) \right) + \frac{\lambda}{N_{\text{reg}}} \left( \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*, \hat{a}_i) \right) \quad (4)$$

式中: 锚框  $\hat{a}_i$  的尺寸是锚框  $\hat{a}_i$  的 2 倍。将主网络的损失函数与附加增强网络的损失函数进行加权求和, 得到训练阶段的融合损失函数  $L_{\text{Fusion}}$  :

$$L_{\text{Fusion}} = L_{\text{M-Net}}(p_i, t_i, \hat{a}_i) + \beta L_{\text{AE-Net}}(p_i, t_i, \hat{a}_i) \quad (5)$$

式中:  $\beta$  是平衡  $L_{\text{M-Net}}(p_i, t_i, \hat{a}_i)$  和  $L_{\text{AE-Net}}(p_i, t_i, \hat{a}_i)$  之间的权重系数。在预测阶段, 本文仅对附加增强网络的特征图进行检测, 这样并不会引入额外的时间损耗。

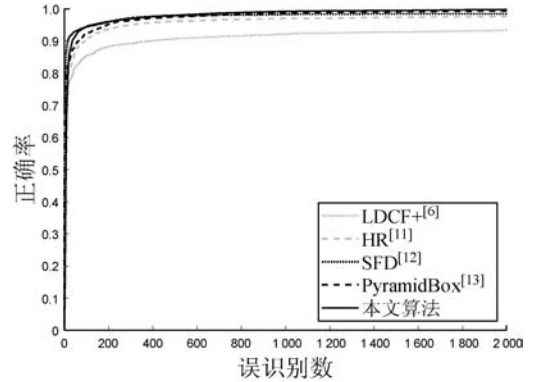
### 3 实验

为了验证本文算法的精确性和稳健性, 选择 FDDB<sup>[14]</sup> (Face Detection Data Set and Benchmark) 和 WIDER FACE<sup>[15]</sup> 人脸检测基准数据集进行实验, 实验中选择 LDCF + <sup>[6]</sup>、HR<sup>[11]</sup>、SFD<sup>[12]</sup> 和 PyramidBox<sup>[13]</sup> 算法作为对比算法。本文网络仅使用 WIDER FACE 的训练数据进行训练, 在两个数据集上分别进行测试。在训练时, 使用反向传播的随机梯度下降来优化, 最大迭代次数设置为  $10^5$ , 前  $7 \times 10^4$  次迭代, 学习率为  $10^{-3}$ , 后  $3 \times 10^4$  次迭代, 学习率为  $10^{-4}$ 。批量大小设置为 16, 权重衰减设置为 0.000 5, 动量设置为 0.9, 所有卷积层参数都使用 MSRA 方法进行初始化。本文算法及其对比算法均采用 Python 代码实现, 并在配置 Nvidia GTX Titan X 的电脑上运行。

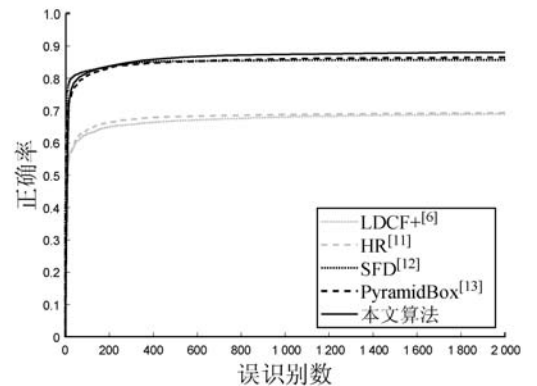
#### 3.1 FDDB 数据集测试结果分析

将在 WIDER FACE 数据集上训练好的模型再在 FDDB 数据集上进行测试。FDDB 人脸数据集共包含 2 845 幅图片, 其中标记了 5 171 个具有遮挡、模糊、低分辨率和姿态变化的人脸图像。由于 FDDB 数据集用的是椭圆标注, 而本文训练的模型输出的是矩形框, 因此在测试之前需要将椭圆标注转化矩形框。本文算法

与 LDCF +、HR、SFD、PyramidBox 算法在 FDDB 数据集上的 ROC 曲线如图 4 所示。



(a) 连续 ROC 曲线



(b) 离散 ROC 曲线

图 4 ROC 曲线

可以看出, 在误识别数较低时, 本文算法的检测性能略低于 SFD 算法, 在误识别数较高时, 本文算法的检测性能优于其他 4 种主流人脸检测算法。这说明本文算法在主网络 SSD 的多层特征图上, 通过融合前后层特征图增强原始特征图的辨识度, 并引入有监督信息去学习上下文融合特征, 进而在多尺度增强特征图上完成人脸/非人脸分类和人脸边界框位置更新, 可以有效提高人脸检测的精度。

#### 3.2 WIDER FACE 数据集测试结果分析

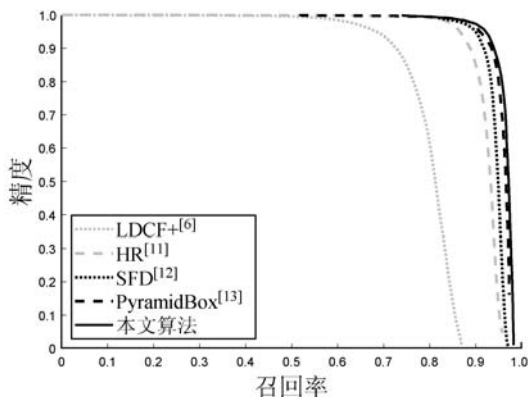
WIDER FACE 人脸数据集在 32 203 幅图像中, 标记了 393 703 个具有遮挡、模糊、尺度变化、光照变化和姿态变化等因素的人脸图像。它包含了 61 种不同的事件类别, 对于每种事件类别, 随机选择 40%、10%、50% 的数据作为训练、验证、测试子集。此外, 每个子集中的数据根据人脸检测的难易程度被划分为简单、适中和困难 3 个难度级别。

本文算法与 LDCF +、HR、SFD 和 PyramidBox 算法进行了对比, 在 WIDER FACE 简单、适中和困难 3 个子测试数据集上的 mAP 及速度对比结果如表 1 所示。在 WIDER FACE 简单、适中和困难 3 个子测试数据集上的精度和召回率 (Precision-Recall, PR) 曲线如图 5

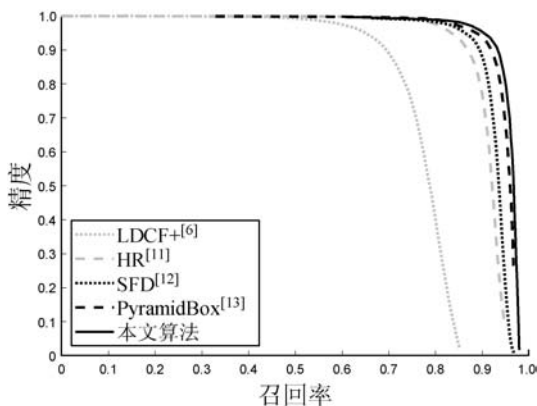
所示。

表 1 mAP 对比结果

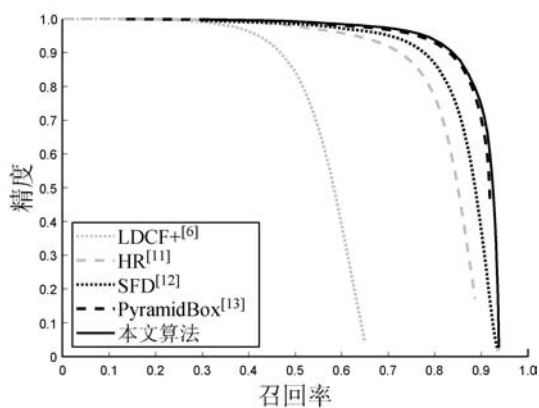
方法	mAP			速度/(帧·s <sup>-1</sup> )
	简单	适中	困难	
LDCF + [6]	0.797	0.772	0.564	21
HR [11]	0.923	0.910	0.819	6
SFD [12]	0.928	0.913	0.840	36
PyramidBox [13]	0.956	0.946	0.887	27
本文算法	0.962	0.951	0.903	34



(a) 简单



(b) 适中



(c) 困难

图 5 精度和召回率曲线

由表 1 可以看出,本文算法在保持低时间损耗的

同时,提高了人脸检测精度。从图 5 的对比结果可以发现,本文算法的 PR 曲线均高于其他 4 种主流人脸检测算法。如果一种算法的 PR 曲线能够包围住另外一种算法的 PR 曲线,则前者的检测性能优于后者,这说明本文算法的人脸检测性能优于所对比的 4 种主流人脸检测算法。本文算法在 WIDER FACE 数据集上的部分检测结果如图 6 所示。



图 6 本文算法在 WIDER FACE 数据集上的部分检测结果

## 4 结 语

本文提出一种基于增强并行级联卷积神经网络的人脸检测方法。该方法在主网络 SSD 的多层特征图上,通过融合前后层特征图增强原始特征图的辨识度,并引入有监督信息去学习上下文融合特征,进而在多尺度增强特征图上完成人脸/非人脸分类和人脸边界框位置更新。此外,在训练阶段为主网络和附加增强网络设置两种基于锚框尺寸的损失函数,增强对小尺度人脸的检测效果。在 Fddb 数据集和 WIDER FACE 数据集上的实验结果表明,本文算法具有比当前主流人脸检测方法更高的检测精度。

## 参 考 文 献

[ 1 ] 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.

[ 2 ] Sun Y, Ren L, Wei Z, et al. A weakly supervised method for makeup-invariant face verification [J]. Pattern Recognition, 2017, 66 : 153 - 159.

[ 3 ] 叶剑锋,王化明. AdaBoost 检测结合 SOM 的自动人脸识别方法 [J]. 哈尔滨工程大学学报, 2018, 39 ( 1 ) : 129 - 134.

[ 4 ] Jiang J, Ma J, Chen C, et al. Noise robust face image super-resolution through smooth sparse representation [J]. IEEE Transactions on Cybernetics, 2017, 47 ( 11 ) : 3991 - 4002.

- [ 5 ] Zhao K, Liu B, Li W H, et al. Anomaly detection and localization: A novel two-phase framework based on trajectory-level characteristics [ C ] // 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2018:1-6.
- [ 6 ] Peng P X, Tian Y H, Wang Y W, et al. Robust multiple cameras pedestrian detection with multi-view Bayesian network [ J ]. Pattern Recognition, 2015, 48 ( 5 ): 1760-1772.
- [ 7 ] Tian L, Du Q. Parallel multi-view low-rank and sparse subspace clustering for unsupervised hyperspectral image classification [ C ] // 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, 2018:618-621.
- [ 8 ] Qian B, Gu X G, Shu Z Q, et al. Constrained multi-view NMF with graph embedding for face clustering [ C ] // 2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). IEEE, 2018:103-106.
- [ 9 ] Marsden M, McGuinness K, Little S, et al. ResnetCrowd: A residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification [ C ] // 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2017:1-7.
- [ 10 ] Shuaibu A N, Malik A S, Faye I. Adaptive feature learning CNN for behavior recognition in crowd scene [ C ] // 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE, 2017:357-361.
- [ 11 ] Keçeli A S, Kaya A. Violent activity detection with transfer learning method [ J ]. Electronics Letters, 2017, 53 ( 15 ): 1047-1048.
- [ 12 ] Zhang W J, Tian L H, Li C, et al. A SSD-based crowded pedestrian detection method [ C ] // 2018 International Conference on Control, Automation and Information Sciences (ICCAIS). IEEE, 2018:222-226.
- [ 13 ] Tzelepi M, Tefas A. Graph embedded convolutional neural networks in human crowd detection for drone flight safety [ J / OL ]. IEEE Transactions on Emerging Topics in Computational Intelligence. (2019-03-04) [ 2019-07-18 ]. DOI: 10.1109/TETCI.2019.2897815.
- [ 14 ] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size [ EB ]. arXiv:1602.07360, 2016.
- [ 15 ] Motamedi M, Gysel P, Akella V, et al. Design space exploration of FPGA-based deep convolutional neural networks [ C ] // 2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC). IEEE, 2016:575-580.
- [ 16 ] 蒋华,季丰,王慧娇,等.改进 Kmeans 算法的海洋数据异常检测 [ J ]. 计算机工程与设计, 2018, 39 ( 10 ): 3132-3136.
- [ 17 ] 左进,陈泽茂.基于改进 K 均值聚类的异常检测算法 [ J ]. 计算机科学, 2016, 43 ( 8 ): 258-261.
- [ 18 ] Shao J, Loy C C, Wang X G. Scene-independent group profiling in crowd [ C ] // 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2014:2219-2226.
- [ 19 ] Zhou B L, Tang X O, Zhang H P, et al. Measuring crowd collectiveness [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36 ( 8 ): 1586-1599.
- [ 20 ] Solera F, Calderara S, Cucchiara R. Socially constrained structural learning for groups detection in crowd [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38 ( 5 ): 995-1008.
- [ 21 ] Zhou B L, Tang X O, Wang X G. Coherent filtering: Detecting coherent motions from crowd clutters [ C ] // 2012 European Conference on Computer Vision. Springer, 2012:857-871.
- ~~~~~
- (上接第 105 页)**
- [ 5 ] Viola P, Jones M J. Robust real-time face detection [ J ]. International Journal of Computer Vision, 2004, 57 ( 2 ): 137-154.
- [ 6 ] Ohn-bar E, Trivedi M M. To boost or not to boost? on the limits of boosted trees for object detection [ C ] // 2016 23rd International Conference on Pattern Recognition, 2016.
- [ 7 ] 刘宏哲,杨少鹏,袁家政,等.基于单一神经网络的多尺度人脸检测 [ J ]. 电子与信息学报, 2018, 40 ( 11 ): 63-70.
- [ 8 ] Jiang H, Learned-miller E. Face detection with the faster R-CNN [ C ] // 2017 12th International IEEE Conference on Automatic Face Gesture Recognition, 2016.
- [ 9 ] 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.
- [ 10 ] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [ C ] // European Conference on Computer Vision, 2016.
- [ 11 ] Hu P, Ramanan D. Finding tiny faces [ C ] // 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [ 12 ] Zhang S, Zhu X, Lei Z, et al. S3FD: single shot scale-invariant face detector [ C ] // 2017 IEEE International Conference on Computer Vision, 2017.
- [ 13 ] Tang X, Du D K, He Z, et al. Pyramidbox: a context-assisted single shot face detector [ C ] // European Conference on Computer Vision, 2018.
- [ 14 ] Jain V, Learned-miller E. Fddb: a benchmark for face detection in unconstrained settings [ M ]. UMass Amherst Technical Report, 2010.
- [ 15 ] Yang S, Luo P, Chen C L, et al. WIDER FACE: a face detection benchmark [ C ] // 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.