

融合深度字典学习和特征重建的遮挡人脸检测研究

戴惠丽

(闽南科技学院计算机信息学院 福建 泉州 362332)

摘要 针对复杂场景中遮挡人脸检测精度低的问题,提出一种融合深度字典学习和特征重建的遮挡人脸检测方法。利用一个浅层 CNN(Convolutional Neural Networks)生成人脸候选区域并用预训练好的 VGG16 网络对其进行特征描述。采用稀疏编码建立一个由典型性的人脸和非人脸构成的深度检索字典。采用局部保留投影的方法,利用检索字典将人脸候选区域的特征描述符重建为一个基于相似性的特征向量。将重建后的特征向量送入到深层神经网络以同时进行人脸/非人脸分类和人脸边界框位置回归。在 MAFA 遮挡人脸数据集上的实验结果表明,该方法的检测精度比当前主流人脸检测方法提高了约 12.3 个百分点。

关键词 人脸检测 遮挡人脸 卷积神经网络 字典学习 特征重建

中图分类号 TP3 **文献标志码** A **DOI**:10.3969/j.issn.1000-386x.2024.11.032

AN OCCLUSION FACE DETECTION METHOD COMBINING DEPTH DICTIONARY LEARNING AND FEATURE RECONSTRUCTION

Dai Huili

(College of Computer Information, Minnan Science and Technology University, Quanzhou 362332, Fujian, China)

Abstract Aimed at the low accuracy of occluded face detection in complex scenes, an occlusion face detection method combining depth dictionary learning and feature reconstruction is proposed. A shallow CNN was used to generate face candidate regions, and the pre-trained VGG16 network was used to characterize them. A sparse coding method was used to establish a deep retrieval dictionary composed of typical faces and non-faces. Using the locality preserving projections method, the feature descriptor of the face candidate region was reconstructed into a similarity-based feature vector by using the retrieval dictionary. The reconstructed feature vector was sent to the deep neural network to perform face/non-face classification and face bounding box location regression simultaneously. The experimental results on the MAFA occlusion face dataset show that the detection accuracy of this method is about 12.3 percentage points higher than the current mainstream face detection method.

Keywords Face detection Occlusion face Convolutional neural network Dictionary learning Feature reconstruction

0 引言

随着机器学习和计算机视觉技术的快速发展,简单的人脸检测问题已经得到了较好的解决。比如,文献[1]提出的人脸检测器通过融合尺度、分辨率和上下文等多种信息来检测人脸,在 WIDER FACE^[2] 人脸检测数据集简单、中等和困难三个测试子集上分别实现了 92.3%、91.0% 和 81.9% 的平均精度。在先进人

脸检测方法的启发下,人脸检测被广泛地应用于人脸对齐^[3]、人脸比对^[4]、人机交互^[5] 和人脸重建^[6] 等研究领域。

作为人脸检测的经典方法,文献[7]提出了一种基于 Haar 特征和级联弱分类器的人脸检测方法,取得了良好的检测效果。文献[8]提出了一种使用分层的可变形组件模型(Deformable Part Model, DPM)来联合进行人脸检测和特征点定位的方法,虽然达到了较好的检测精度,但是时间损耗大。受限于人工设计特征

的表达能力和分类器的分类能力,上述方法已经被基于卷积神经网络(CNN)^[9-10]的深度学习方法取代。

与基于特征增强和基于可变形组件模型的方法不同,基于 CNN 的人脸检测方法直接从数据中学习人脸表示,并采用深度学习范式在滑动窗口中检测人脸。文献[11]提出了一种基于级联 CNN 的人脸检测方法,它是经典的 Viola-Jones 人脸检测算法的深度卷积神经网络实现。文献[12]提出了一种基于特征聚合框架^[13]的人脸检测方法,而这些特征由 CNN 生成。文献[14]提出了一种基于多任务级联卷积神经网络的人脸检测方法,该方法采用三级级联卷积神经网络进行由粗到细的人脸检测,并且还通过在线困难样本生成策略提升了人脸检测性能。文献[15]通过将多个深度回归网络与去遮挡自编码网络相耦合,恢复了人脸遮挡区域的真实外观,并将恢复的人脸遮挡区域与未被遮挡区域联合训练,提高了对遮挡人脸的定位精度。文献[16]提出了一种基于上下文辅助的单步人脸检测方法,通过引入上下文敏感的预测模块,提升了对小尺度、模糊和部分遮挡人脸的检测性能。文献[17]提出了一种基于身份保持的生成式对抗网络(Identity-Preserving Generative Adversarial Networks, IP-GAN),通过生成符合真实人脸分布的图像来填充人脸被遮挡部分,并加入身份属性以融合人和机器的感知,提高了修复和识别遮挡人脸的精度。文献[18]基于长短期记忆网络(Long Short-Term Memory, LSTM)提出了一种自动编码器模型,该模型由两个 LSTM 组件组成,分别对遮挡人脸进行稳健编码和循环遮挡去除,提升了户外场景下对遮挡人脸的检测性能。

虽然基于 CNN 的人脸检测方法已经被广泛地关注和研究,但是在复杂场景下,对遮挡人脸的检测问题仍然存在困难。如图 1 所示,遮挡人脸通常具有多方向性、多种遮挡程度和不同的遮罩类型,这使得最先进的人脸探测方法^[19-20]也难以对其精确定位。当前主流的人脸检测方法在检测遮挡人脸时性能下降明显,主要是因为遮挡区域存在较大程度的人脸特征缺失,并且遮挡的存在会不可避免地带来噪声干扰。



图 1 遮挡人脸示意图

本文针对复杂场景中的遮挡人脸检测问题,进行以下两个部分的改进:1) 提出一个通用特征提取模块,使用一个浅层 CNN 生成人脸候选区域并用预训练好的 VGG16 网络对人脸候选区域进行特征描述,该模块可直接移植到其他人脸检测算法中。2) 提出一种对遮挡人脸的特征重建方法,既恢复了由遮挡引起的人脸信息缺失,又抑制了遮挡噪声影响。首先,利用稀疏编码的方法建立了一个由典型性的人脸和非人脸构成的深度检索字典。接着,采用局部保留投影的方法,利用该检索字典将人脸候选区域的特征描述符重建为一个基于相似性的特征向量。

1 VGG16 人脸特征提取模型

VGG16 是一个经典的深度卷积神经网络,它通过反复堆叠 3×3 大小的卷积核和 2×2 大小的最大池化层,构建了深度为 16 层的卷积神经网络。到目前为止,很多网络都是在 VGG16 的基础上通过增加和删减其他层来进行改进,此外,该网络也经常被用来进行图像特征的提取。

基于 VGG16 的人脸检测模型如图 2 所示,它包含 13 个卷积层、5 个最大池化层、3 个全连接层和一个 Softmax 层。所有卷积层的卷积核大小都是 3×3 ,并且步长为 1,该模型通过堆叠多个 3×3 卷积核来获得与大尺寸卷积核相同的感受野并捕获更多的空间上下文信息。比如堆叠 2 个 3×3 卷积核可以实现与 5×5 卷积核相同的感受野,堆叠 3 个 3×3 卷积核可以实现与 7×7 卷积核相同的感受野。这种卷积核组织方式既有助于提高模型的分类能力,又可以减少网络参数的数量。5 个最大池化层分别设置在第 2、4、7、10、13 个卷积层后,经过池化,特征图的长和宽减少为原来的一半,通道数增加为原来的两倍,需要注意的是,第 4 个和第 5 个最大池化层保持通道数不变。在最大池化层之后,有 3 个全连接层,前两个全连接层有 4 096 个通道,第三个全连接层有 1 000 个通道,所有全连接层的配置均相同,经过全连接层后,Softmax 层输出目标分类得分的概率。

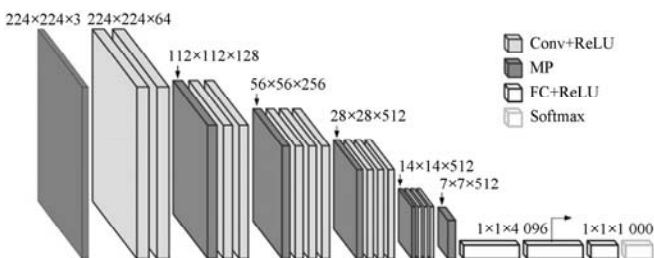


图 2 基于 VGG16 的人脸特征提取模型

综上所述,该模型的输入是一幅 224×224 大小的包含单个人脸或者多个人脸的 RGB 图像,输出是一组目标类别的分类得分概率。当该模型被用来进行人脸特征提取时,从第 7 个全连接层输出一个 4 096 维的特征向量。

2 基于深度字典学习和特征重建的遮挡人脸检测方法

2.1 总体框架

基于深度字典学习和特征重建的遮挡人脸检测模

型如图 3 所示,它由通用特征提取模块、人脸特征重建模块和人脸精确回归模块三部分组成。

通用特征提取模块生成人脸候选区域并用特征描述符对人脸候选区域进行描述。人脸特征重建模块首先建立了一个由典型性的人脸和非人脸构成的深度检索字典,接着采用局部保留投影的方法将人脸候选区域的特征描述符转化为一个基于相似性的特征向量。人脸精确回归模块将重建后的特征向量送入到一个深层神经网络并同时进行人脸/非人脸分类和人脸边界框位置回归,以进一步提高对遮挡人脸的检测精度。

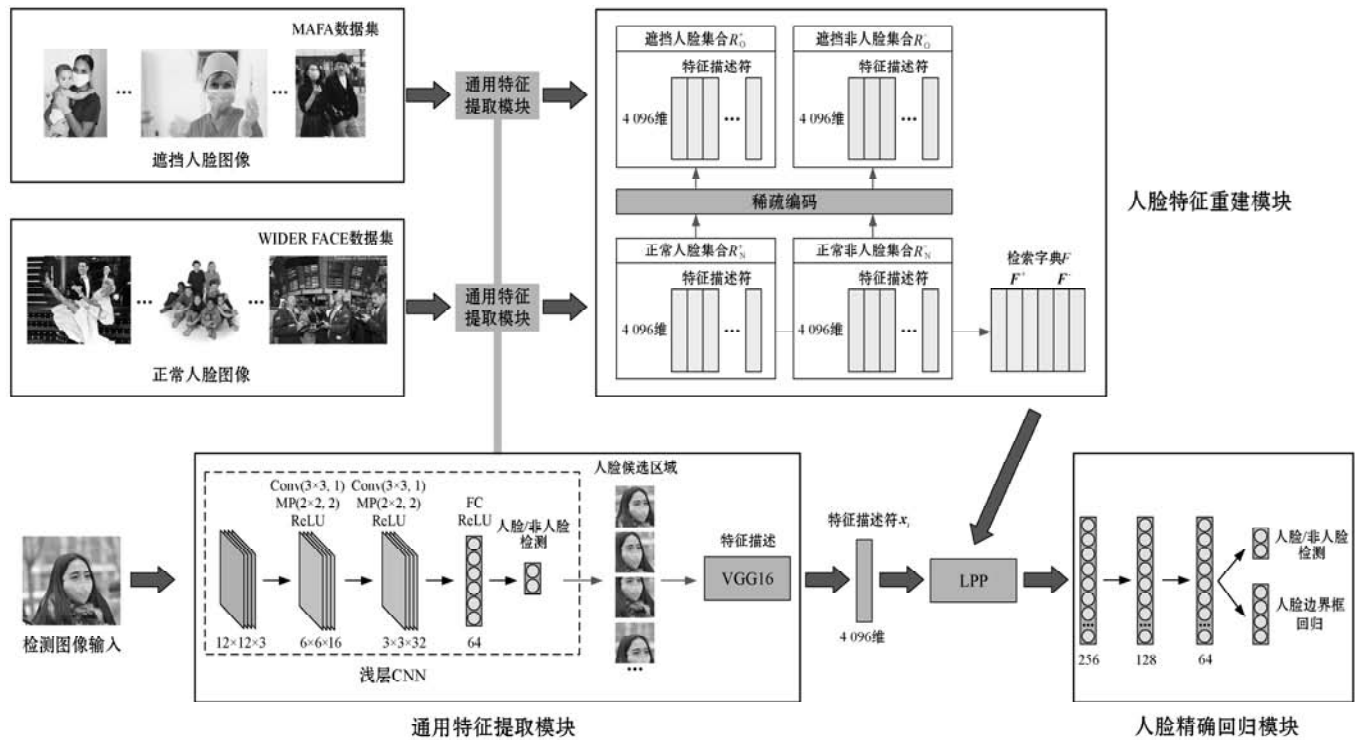


图3 基于深度字典学习和特征重建的遮挡人脸检测模型

2.2 通用特征提取模块

通用特征提取模块首先使用一个浅层 CNN 完成人脸候选区域的快速生成,然后使用预训练好的 VGG16 网络对人脸候选区域进行特征描述。

该浅层 CNN 的网络结构如图 3 左下角所示,它由 2 个卷积层、2 个最大池化层、3 个 ReLU 层和 1 个全连接层组成。通过这个浅层的 CNN,可以滤除掉大部分置信度较低的人脸候选区域,考虑到浅层 CNN 对遮挡人脸的检测效果较差,在最后一层通过设置一个较低的阈值以生成大量的人脸候选区域。

接着,将浅层 CNN 生成的每一个人脸候选区域 P_i 送到预训练好的 VGG16 网络,从第 7 个全连接层生成一个 4 096 维的向量,对其进行归一化后得到特征描述符。

2.3 人脸特征重建模块

人脸特征重建模块在 x_i 中恢复缺失的人脸信息,并抑制由遮挡区域产生的噪声。本文通过建立检索字典的方法在外部数据库中检索由典型性的人脸和非人脸组成的特征子空间,并使用它们来恢复人脸候选区域的特征描述符 x_i 。

首先建立 x_i 的检索字典 F^+ 和 F^- , F^+ 由最具代表性的人脸特征描述符集合组成, F^- 由最具代表性的非人脸特征描述符集合组成。为了构造检索字典 F^+ 和 F^- ,本文采用通用特征提取模块的浅层 CNN 对包含遮挡人脸的 MAFA 数据集和包含正常人脸的 WIDER FACE 数据集分别进行检测,得到正常人脸集合 R_N^+ 、正常非人脸集合 R_N^- 、遮挡人脸集合 R_0^+ 和遮挡非

人脸集合 R_0^+ , 并采用预训练好的 VGG16 网络提取集合中每个人脸候选区域的特征描述符。

检索字典 F^+ 和 F^- 中的列分别由来自 R_N^+ 和 R_N^- 的特征描述符组成。 F^+ 在 R_0^+ 中对遮挡人脸进行稀疏编码时误差最小, 在 R_0^- 中对遮挡非人脸进行稀疏编码时误差最大, 因此, 可以通过求解式(1)得到 F^+ 。

$$\min_{F^+} \frac{1}{|R_0^+|} \sum_{p_n^+ \in R_0^+} \|x_n^+ - F^+ \alpha\|_2^2 - \frac{1}{|R_0^-|} \sum_{p_n^- \in R_0^-} \|x_n^- - F^+ \beta\|_2^2 \quad (1)$$

$$\text{s. t. } \|\alpha\|_0 = \|\beta\|_0 = 1, F_i^+ \in \{x_N^+ \mid p_N^+ \in R_N^+\}$$

式中: p_n^+ 、 p_n^+ 和 p_n^- 分别表示人脸集合 R_N^+ 、 R_0^+ 和 R_0^- 中的 1 个候选区域; x_n^+ 、 x_n^+ 和 x_n^- 分别表示 p_n^+ 、 p_n^+ 和 p_n^- 的 4 096 维特征描述符; α 和 β 是用字典 F^+ 编码 x_n^+ 和 x_n^- 的稀疏系数向量, F_i^+ 表示 F^+ 的第 i 列。在 $\|\alpha\|_0 = \|\beta\|_0 = 1$ 的约束下, 该稀疏编码过程等价于建立 x_i 在 R_N^+ 中的最近邻人脸特征描述符集合 F^+ 。

F^- 在 R_0^+ 中对遮挡人脸进行稀疏编码时误差最大, 在 R_0^- 中对遮挡非人脸进行稀疏编码时误差最小, 因此, 可以通过求解式(2)得到 F^- 。

$$\max_{F^-} \frac{1}{|R_0^+|} \sum_{p_n^+ \in R_0^+} \|x_n^+ - F^- \alpha\|_2^2 - \frac{1}{|R_0^-|} \sum_{p_n^- \in R_0^-} \|x_n^- - F^- \beta\|_2^2 \quad (2)$$

$$\text{s. t. } \|\alpha\|_0 = \|\beta\|_0 = 1, F_i^- \in \{x_N^- \mid p_N^- \in R_N^-\}$$

式中: p_n^- 表示正常非人脸集合 R_N^- 中的 1 个候选区域; x_n^- 表示候选区域 p_n^- 的 4 096 维特征描述符; F_i^- 表示 F^- 的第 i 列, 在 $\|\alpha\|_0 = \|\beta\|_0 = 1$ 的约束下, 该稀疏编码过程等价于建立 x_i 在 R_N^- 中的最近邻非人脸特征描述符集合 F^- 。

合并 F^+ 和 F^- 得到检索字典 $F = [F^+, F^-]$, F 包含了典型的人脸特征描述符集合和非人脸特征描述符集合。接着, 采用局部保留投影 (Locality Preserving Projections, LPP) 的方法, 利用最近邻搜索, 在检索字典中寻找加权求和后与人脸候选区域特征描述符最相似的特征描述符子集, 该子集的权重集合就是基于相似性的特征向量 v_i 。 v_i 恢复了缺失的人脸信息, 并抑制了由遮挡造成的噪声。

$$v_i = \arg \min_v \|x_i - Fv\| \quad (3)$$

2.4 人脸精确回归模块

人脸精确回归模块将重建后的特征向量 v_i 送入到同时执行分类和回归任务的深层神经网络中。如图 3 右下角所示, 该网络仅由 3 个全连接层组成, 其中分类

分支判定所框选目标是否为人脸, 回归分支输出人脸候选框的精确位置坐标。

3 实验

为了验证本文方法的精确性和有效性, 选择 MAFA 遮挡人脸检测基准数据集^[21] 和 WIDER FACE 人脸检测基准数据集^[2] 进行实验。实验中使用 MAFA 训练集和 WIDER FACE 训练集进行训练, 使用 MAFA 测试集进行测试。本文共选择了 4 种主流人脸检测方法进行对比, 分别是 HPM^[8]、MT^[14]、HR^[1] 和 Pyramid-Box^[16]。各对比方法运行的硬件环境均为配备了 i7 8700K CPU 和 Nvidia GTX Titan X GPU 的计算机。

3.1 数据准备和参数设置

MAFA 遮挡人脸数据集在 30 811 幅图像中标记了 35 806 个遮挡人脸。该数据集中的遮挡人脸有各种不同的方向和遮挡级别, 每个人脸至少有一部分被遮罩遮挡。在标记时, 每幅图像至少包含了一个被遮罩遮挡的人脸, 并且每个遮挡人脸具有六种属性: 人脸位置、眼睛位置、遮罩位置、人脸方向、遮挡级别和遮罩类型。MAFA 训练集包含了 25 876 幅图像, 其中有 29 452 个遮挡人脸, MAFA 测试集包含了剩余的 4 935 幅图像, 其中有 6 354 个遮挡人脸。

WIDER FACE 人脸数据集在 32 203 幅图像中标记了 393 703 个人脸, 所标人脸存在不同程度的模糊、光照差异、尺度变化和姿态变化。该数据集具有 61 种事件类别, 每种事件类别中, 随机选择 40%、10%、50% 的数据作为训练、验证和测试子集。由于 MAFA 训练集包含的遮挡人脸数量远小于 WIDER FACE 训练集包含的人脸数量, 为了平衡期间, 本文仅在 WIDER FACE 训练集中选择 1/3 的数据进行实验。

在训练时, 本文网络使用反向传播的随机梯度下降优化, 最大迭代次数设置为 10^5 , 前 6.5×10^4 次迭代, 学习率设置为 10^{-3} , 后 3.5×10^4 次迭代, 学习率设置为 10^{-4} , 批大小设置为 16, 权重衰减设置为 0.000 5, 动量设置为 0.9, 所有卷积层参数使用 MSRA 方法初始化。

3.2 MAFA 数据集测试结果分析

本文方法与 HPM、MT、HR 和 PyramidBox 等方法的平均精度 (Average Precision, AP) 对比结果如表 1 所示, 本文方法与 HPM、MT、HR 和 PyramidBox 等方法在 MAFA 测试集的部分对比结果如图 4 所示。

表 1 本文方法与 HPM、MT、HR 和 PyramidBox 等方法的平均精度对比结果(%)

属性	HPM ^[8]	MT ^[14]	HR ^[1]	PyramidBox ^[16]	本文方法
Left	1.29	6.89	10.20	13.80	20.10
Left-front	26.60	31.90	33.40	38.20	63.90
Front	64.40	62.20	69.70	71.30	82.60
Right-front	18.90	20.20	23.80	27.60	58.20
Right	0.93	1.94	5.75	8.87	18.70
Weak	58.50	56.20	63.30	63.80	80.10
Medium	34.80	45.60	39.70	52.30	72.30
Heavy	5.31	5.24	10.10	12.40	27.50
Simple	54.70	51.60	59.30	58.40	78.80
Complex	46.10	48.20	51.70	55.70	75.50
Body	23.40	30.40	28.40	37.90	67.40
Hybrid	6.00	6.48	11.00	13.20	29.20
All	60.00	60.80	65.30	67.90	80.20

由表 1 可知,本文方法在 MAFA 测试集的平均精度明显高于其他 4 种主流人脸检测方法。本文方法的平均精度达到了 80.2%,比第二名的平均精度高出了约 12.3 个百分点。分析不同人脸方向、遮挡程度和遮罩类型下的实验结果可以发现,随着人脸偏转角度的增加,所有方法对 MAFA 测试集的平均精度会明显下降。并且遮挡程度的提高,也会严重影响各人脸检测方法的检测性能。此外,简单的遮罩也会干扰人脸检测方法的检测性能,并且随着遮罩趋于复杂,各人脸检测方法在 MAFA 测试集的平均精度均下降明显。尽管存在上述困难,本文方法在人脸方向、遮挡程度和遮罩类型共 12 个属性下的遮挡人脸检测效果仍优于其他 4 种人脸检测方法,这证明了本文方法的有效性和精确性。



图 4 本文方法与 HPM、MT、HR 和 PyramidBox 等方法在 MAFA 测试集的部分对比结果

为进一步研究本文方法与 HPM、MT、HR 和 PyramidBox 等方法的遮挡人脸检测性能,对图 4 中具有代表性的对比结果进行分析。由图 4 可知,对于遮挡程度较大的人脸,主流的人脸检测方法甚至不能进行识别,而本文方法均检测出了遮挡人脸。分析这些检测结果可知,特征稳健性是本文方法优于其他对比方法的主要原因。本文方法在人脸特征重建阶段通过参考由典型性的人脸和非人脸构成的特征描述符集合,重建了遮挡人脸的特征描述符,并恢复了遮挡区域的人脸固有特征。通过这种方式,典型性的人脸和非人脸构成的特征描述符集合为识别真实人脸并精确定位其位置提供了有效的特征信息。而其他 4 种人脸检测方法均依赖于从包含遮挡的区域内直接提取局部特征,当人脸遮挡程度较大时,这些特征的稳健性较差,进而降低了对遮挡人脸的检测精度。

3.3 速度和精度对比

本文方法旨在低时间损耗下完成精确的遮挡人脸检测。为了定量比较本文方法与对比方法的精度和速度,并验证引入特征重建的有效性,本文选择 MAFA 测试集进行实验。本文方法与 HPM、MT、HR 和 PyramidBox 等方法在 100 次误报时的精度和速度对比如表 2 所示。

表 2 本文方法与 HPM、MT、HR 和 PyramidBox 等方法在 100 次误报时的速度和精度对比

方法	精度	速度/(帧·s ⁻¹)
HPM ^[8]	61.3	17
MT ^[14]	62.4	13
HR ^[1]	67.1	8
PyramidBox ^[16]	69.7	24
本文算法(无特征重建)	70.6	34
本文算法	81.5	26

可以看出,本文方法的运行速度高于其他 4 种对比方法,这主要受益于本文方法使用了一个浅层 CNN 生成人脸候选区域并用预训练好的 VGG16 网络对人脸候选区域进行特征描述。此外,本文方法在人脸特征重建阶段通过参考由典型性的人脸和非人脸构成的特征描述符集合,重建遮挡人脸的特征描述符,并恢复遮挡区域的人脸固有特征。这一稳健性的特征也使本文方法的检测精度高于其他 4 种对比方法,证明本文方法具有较高的精确性和实时性。此外,在未引入特征重建时,本文方法的检测速度较高,但是精度较低,引入特征重建后,检测速度有所下降,精度得到显著提高,这证明特征重建的引入有助于检测精度的提高。

4 结 语

本文提出一种融合深度字典学习和特征重建的遮挡人脸检测方法。本文方法在初步生成人脸候选区域特征描述符集合的基础上,利用稀疏编码建立一个由典型性的人脸和非人脸构成的深度检索字典。接着,采用局部保留投影的方法,利用该检索字典将检测人脸候选区域的特征描述符重建为一个基于相似性的特征向量,进而在深层神经网络中进行人脸/非人脸分类和人脸边界框位置回归。在 MAFA 测试集上的实验结果表明,本文方法具有比当前主流人脸检测方法更高的检测精度。本文方法可作为遮挡人脸检测的基础框架,对推动遮挡人脸检测发展具有一定意义。下一步,计划设计一种更科学的人脸特征重建模块,以增强本文方法对人脸遮挡检测的稳健性。

参 考 文 献

- [1] Hu P, Ramanan D. Finding tiny faces[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1522 - 1530.
- [2] 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.
- [3] Shi B, Bai X, Liu W, et al. Face alignment with deep regression[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(1): 183 - 194.
- [4] 甘俊英, 翟懿奎, 黄聿, 等. 基于双激活层深度卷积特征的人脸美丽预测研究[J]. 电子学报, 2019, 47(3): 636 - 642.
- [5] Villamizar M, Sanfeliu A, Moreno-Noguer F. Online learning and detection of faces with low human supervision[J]. The Visual Computer, 2019, 35(3): 349 - 370.
- [6] 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.
- [7] Viola P, Jones M J. Robust real-time face detection[J]. International Journal of Computer Vision, 2004, 57(2): 137 - 154.
- [8] Ghiasi G, Fowlkes C C. Occlusion coherence: Localizing occluded faces with a hierarchical deformable part model[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1899 - 1906.
- [9] 李小薪, 梁荣华. 有遮挡人脸识别综述: 从子空间回归到深度学习[J]. 计算机学报, 2018, 41(1): 177 - 207.
- [10] Xiao Y, Cao D, Gao L. Face detection based on occlusion area detection and recovery[J]. Multimedia Tools and Applications, 2019, 1(2): 1 - 16.
- [11] Zhang J, Wu X, Hoi S C H, et al. Feature agglomeration networks for single stage face detection[J]. Neurocomputing, 2020, 380: 180 - 189.
- [12] Ranjan R, Bansal A, Zheng J, et al. A fast and accurate system for face detection, identification, and verification [J]. IEEE Transactions on Biometrics, Behavior, and Identity Science, 2019, 1(2): 82 - 96.
- [13] Wang L, Yu X, Bourlai T, et al. A coupled encoder-decoder network for joint face detection and landmark localization[J]. Image and Vision Computing, 2019, 87: 37 - 46.
- [14] Zhang S, Wen L, Shi H, et al. Single-shot scale-aware network for real-time face detection[J]. International Journal of Computer Vision, 2019, 127(6-7): 537 - 559.
- [15] Deng J, Trigeorgis G, Zhou Y, et al. Joint multi-view face alignment in the wild[J]. IEEE Transactions on Image Processing, 2019, 28(7): 3636 - 3648.
- [16] Zeng D, Liu H, Zhao F, et al. Proposal pyramid networks for fast face detection[J]. Information Sciences, 2019, 495: 136 - 149.
- [17] Li C Y, Ge S M, Hua Y Y, et al. Occluded face recognition by identity-preserving inpainting[C]//The 3rd International Symposium on Artificial Intelligence and Robotics, 2018: 427 - 437.
- [18] Dang L M, Hassan S I, Im S, et al. Face image manipulation detection based on a convolutional neural network[J]. Expert Systems with Applications, 2019, 129: 156 - 168.
- [19] Zhao Z Q, Zheng P, Xu S, et al. Object detection with deep learning: A review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212 - 3232.
- [20] Mutneja V, Singh S. Modified Viola-Jones algorithm with GPU accelerated training and parallelized skin color filtering-based face detection[J]. Journal of Real-Time Image Processing, 2019, 16(5): 1573 - 1593.
- [21] Ge S, Li J, Ye Q, et al. Detecting masked faces in the wild with LLE-CNNs[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2682 - 2690.

(上接第 180 页)

- [13] Zacarias I, Schwarzrock J, Gaspary L P, et al. Enhancing mobile military surveillance based on video streaming by employing software defined networks[J]. Wireless Communications & Mobile Computing, 2018, 2018: 1 - 12.
- [14] Mininet: An instant virtual network on your laptop (or other pc) [EB/OL]. [2019-05-20]. <http://mininet.org/>.
- [15] Fontes R, Rothenberg C E. Mininet-WiFi: A platform for hybrid physical-virtual software-defined wireless networking research[C]//Proceedings of the 2016 ACM Conference on Special Interest Group on Data Communication, 2016: 607 - 608.