



(12)发明专利

(10)授权公告号 CN 106485215 B

(45)授权公告日 2020.03.06

(21)申请号 201610861396.8

(22)申请日 2016.09.29

(65)同一申请的已公布的文献号
申请公布号 CN 106485215 A

(43)申请公布日 2017.03.08

(73)专利权人 西交利物浦大学
地址 215123 江苏省苏州市工业园区独墅湖高等教育区仁爱路111号

(72)发明人 张百灵 夏翌彰 钱荣强 颜诗洋

(74)专利代理机构 苏州创元专利商标事务所有
限公司 32103

代理人 范晴 丁浩秋

(51)Int.Cl.
G06K 9/00(2006.01) (续)

(56)对比文件
CN 105740823 A,2016.07.06,
CN 105868689 A,2016.08.17,
Yizhang Xia 等.Face Occlusion
Detection Using Deep Convolutional Neural
Networks.《The 2015 11th International
Conference on Natural Computation and the
2015 12th International Conference on

Fuzzy Systems and Knowledge Discovery》
.2015,第1节第6段,第3节,第4.1节,第4.2节,第
4.3节,图6-图8.

Yizhang Xia 等.Face Occlusion
Detection Using Deep Convolutional Neural
Networks.《The 2015 11th International
Conference on Natural Computation and the
2015 12th International Conference on
Fuzzy Systems and Knowledge Discovery》
.2015,第1节第6段,第3节,第4.1节,第4.2节,第
4.3节,图6-图8.

邵蔚元 等.多任务学习及卷积神经网络在
人脸识别中的应用.《计算机工程与应用》.2016,
第52卷(第13期),32-37,88. (续)

审查员 崔蓝芳

权利要求书2页 说明书9页 附图4页

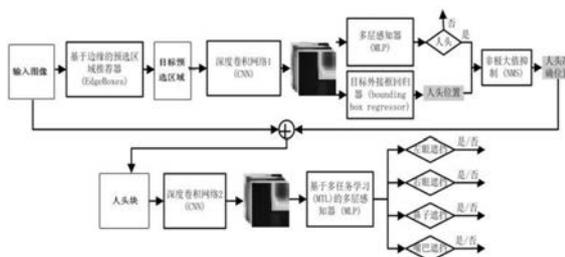
(54)发明名称

基于深度卷积神经网络的人脸遮挡检测方法

(57)摘要

本发明公开了一种基于深度卷积神经网络的人脸遮挡检测方法,包括:对输入图像进行分块,得到目标预选区域;构建第一深度卷积神经网络,通过训练包括第一深度卷积网络和与其连接的第一多层感知器的第一深度卷积神经网络得到所需参数,提取目标预选区域的特征并进行分类;根据提取的特征,通过第二多层感知器预测人头位置;将分类类别是人头的可信度和预测得到的人头位置通过非极大值抑制过滤去除重叠的重复检测框;联合原图分割得到人头块,构

建基于多任务学习策略的第二深度卷积神经网络,判断该人头块的左眼、右眼、鼻子和嘴巴是否被遮挡。该方法能准确检测遮挡的人脸,并且判断其具体的遮挡部位,主要用于自动取款机前摄像机视频的犯罪预警。



CN 106485215 B

[接上页]

(51) Int.Cl.

G06K 9/62(2006.01)

(56)对比文件

陈耀丹 等.基于卷积神经网络的人脸识别方法.《东北师大学报(自然科学版)》.2016,第48卷(第2期),70-76.

Ross Girshick 等.Region-Based Convolutional Networks for Accurate

Object Detection and Segmentation.《IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE》.2016,第38卷(第1期),第7.3节.

Yizhang Xia 等.Face Occlusion Detection Based on Multi-task Convolution Neural Network.《2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)》.2015,375-379.

1. 一种基于深度卷积神经网络的人脸遮挡检测方法,其特点在于,包括以下步骤:

S01:对输入图像进行分块,得到目标预选区域;

S02:构建第一深度卷积神经网络,通过训练第一深度卷积神经网络得到所需参数,提取目标预选区域的特征并进行分类,所述第一深度卷积神经网络包括第一深度卷积网络和与其连接的第一多层感知器;根据链式法则逐步反向传播误差,更新各层的权重参数值;反向传播表达式如下:

$$\nabla w = -\gamma * \frac{\partial E}{\partial w} \quad (6)$$

式中 ∇w 为加在原权重上的权重变化值, γ 为学习速率, E 为误差, $\frac{\partial E}{\partial w}$ 为计算误差对权重 w 的偏导数;

所述第一深度卷积神经网络包括依次连接的卷积层、激活层、池化层、全连层、输出层和代价函数层;所述全连层为多层感知器中的一层;

所述卷积层表达式:

$$y = bc + wc * x \quad (1)$$

式中 x 和 y 分别为输入和输出, wc 为深度卷积网络的权重参数, bc 为卷积神经网络权重的偏置参数;

所述激活层表达式:

$$y = f_{\text{ReLU}}(x) = \max(x, 0) \quad (2)$$

式中 x 和 y 分别为输入和输出;

所述池化层表达式:

$$y = f_{\text{max}}(x) = \max\{x_{s,s}\}$$

式中,输出 y 由输入图像 $x_{s,s}$ 上大小为 $s \times s$ 的区域池化而成;

所述全连层表达式:

$$y = f(x) = bf + wf * x \quad (3)$$

式中, wf 是全连层中的权重参数, bf 是偏置参数;

所述输出层表达式:

$$y^{io} = f(x^{io}) = \frac{\exp(x^{io})}{\sum_{io=1}^K \exp(x^{io})} \quad (4)$$

式中, x^{io} 是第 io 类的输入, K 是总类数, y^{io} 是第 io 类的输出得分;

所述代价函数层表达式:

$$Y_{\text{loss}} = - \sum_{j=1}^K [Y_G^j \log(Y^j) + (1 - Y_G^j) \log(1 - Y^j)] \quad (5)$$

式中, Y_{loss} 是代价值, Y^j 和 Y_G^j 分别是实际输出和真实标签第 j 个神经元的值;

S03:根据步骤S02提取的特征,构建第二多层感知器,通过第二多层感知器预测人头位置;

S04:将分类类别是人头的可信度和预测得到的目标位置通过非极大值抑制过滤去除重叠的重复检测框;

S05:联合原图分割得到人头块,构建基于多任务学习策略的第二深度卷积神经网络,所述第二深度卷积神经网络包括第二深度卷积网络和四个并联的第三多层感知器,所述四个第三多层感知器分别用于判断该人头块的左眼、右眼、鼻子和嘴巴是否被遮挡;

所述第二深度卷积神经网络的构建方法包括:

构建第二深度卷积神经网络;

运用知识迁移策略,载入步骤S02中最终收敛的第一深度卷积神经网络的参数至第二深度卷积神经网络的相应层中,具体层数为卷积层1至池化层3;

将分割得到的人头块加载入此第二深度卷积神经网络中,训练至收敛;

一个多层感知器完成一个任务,每个任务间分享自身的同异,多任务学习方法表达式:

$$L_{\text{joint}} = \sum_{it=1}^{N_t} \theta_{it} L_{it} \quad (11)$$

式中 L_{joint} 表示所有任务的代价总和值, θ_{it} 和 L_{it} 表示任务 it 的权重和代价值, N_t 为任务的总数,为4。

2.根据权利要求1所述的基于深度卷积神经网络的人脸遮挡检测方法,其特征在于,所述步骤S01中通过基于边缘的预选区域推荐器对输入图像进行分块,通过计算滑窗窗口内边缘个数,打分并排序,过滤低分的候选框;所述基于边缘的预选区域推荐器的参数 α 为0.65, β 为0.75,其中, α 为区域推荐器扫框的密度, β 为非最大值抑制的阈值。

3.根据权利要求1所述的基于深度卷积神经网络的人脸遮挡检测方法,其特征在于,所述步骤S02中训练第一深度卷积神经网络得到所需参数,包括,第一深度卷积神经网络通过通用的数据库训练至收敛;然后载入人脸遮挡数据,继续训练至收敛;通过前传方法得到训练误差,通过反向传播方法更新各层的权重参数值。

4.根据权利要求1所述的基于深度卷积神经网络的人脸遮挡检测方法,其特征在于,所述步骤S03包括,将人头的位置转换成一个连续值;训练一个包含四个输出的第二多层感知器直至收敛,分别预测人头最小外接框中心点坐标 (x_p, y_p) 和宽高 (wid, hei) ,变换表达式如下:

$$\widehat{x_p} = \frac{G_{x_p} - P_{x_p}}{P_{wid}} \quad (7)$$

$$\widehat{y_p} = \frac{G_{y_p} - P_{y_p}}{P_{hei}} \quad (8)$$

$$\widehat{wid} = \log(G_{wid} / P_{wid}) \quad (9)$$

$$\widehat{hei} = \log(G_{hei} / P_{hei}) \quad (10)$$

式中 $\widehat{*}$ 是回归器的目标值,其中 $*$ 是 x_p, y_p, wid, hei 之一。

5.根据权利要求1所述的基于深度卷积神经网络的人脸遮挡检测方法,其特征在于,所述步骤S04中非极大值抑制的面积重叠比例阈值为0.3。

基于深度卷积神经网络的人脸遮挡检测方法

技术领域

[0001] 本发明涉及一种人脸遮挡检测方法,具体地涉及一种基于深度卷积神经网络的人脸遮挡检测方法。

背景技术

[0002] 自从20世纪70年代自动取款机(Automatic Teller Machine(ATM))被广泛引入,其一直是犯罪分子的目标。例如,骗子使用各种手段获取用户的卡号和密码。实时自动报警系统是解决此问题的最直接技术。因为,监控摄像机几乎被安装在所有的自动取款机上。然而,视频需要人工二十四小时监管,但是人的疲劳和分心将不可避免。因此,政商们急需一种用于自动取款机的人脸遮挡检测方法。

[0003] 面遮挡检测已经研究了数年,也提出了一些方法,其中许多旨在加强自动取款机的安全性。但是其特征表达方法不能满足自动取款机场景下的复杂的情况。直到深度学习被提出。大量研究者将其应用于各种机器视觉问题,比如检测,分类和分割。深度卷积神经网络模型(deep convolutional neural network),作为深度学习最为成功的模型之一,其集特征提取与分类模型为一体,在有监督的机器学习任务上,具有良好的表现。

[0004] 最近,目标检测最成功的方法是利用众所周知的滑动窗口(sliding window)模式。但是为了准确检测大小变化较大的目标时,此方法将导致后续分类器(classifier)计算量剧增。基于边缘的预选区域推荐器(EdgeBoxes)是一种主流的解决方法,其不仅计算快,推荐的区域少,而且几乎能确保所需目标一定在推荐结果中。

[0005] 为了克服拟合深度学习模型时训练数据不足的问题,知识迁移(knowledge transfer)技术的提出很好得缓解了此问题。

[0006] 多任务学习(multi-task learning)策略共享任务间的信息,其不仅能分享任务间的信息,以达到同时提高每个任务的性能。

[0007] 本发明因此而来。

发明内容

[0008] 针对上述存在的缺陷,本发明的目的是提出了一种基于深度卷积神经网络的人脸遮挡检测方法。该方法能实现基于自动取款机视频场景下的人脸遮挡检测,并且提供具体的遮挡部位。同时,该方法载入通用的预训练模型减少了训练数据,使用基于边缘的预选区域推荐器降低了分类器在复杂场景下的计算复杂度,通过深度卷积神经网络提取了更具类别差异的特征,借用多任务学习策略共享多任务间的差异信息提高了每个任务的性能。

[0009] 本发明的技术方案是:

[0010] 一种基于深度卷积神经网络的人脸遮挡检测方法,包括以下步骤:

[0011] S01:对输入图像进行分块,得到目标预选区域;

[0012] S02:构建第一深度卷积神经网络,通过训练第一深度卷积神经网络得到所需参数,提取目标预选区域的特征并进行分类,所述第一深度卷积神经网络包括第一深度卷积

网络和与其连接的第一多层感知器；

[0013] S03:根据步骤S02提取的特征,构建第二多层感知器,通过第二多层感知器预测人头位置;

[0014] S04:将分类类别是人头的可信度和预测得到的目标位置通过非极大值抑制过滤去除重叠的重复检测框;

[0015] S05:联合原图分割得到人头块,构建基于多任务学习策略的第二深度卷积神经网络,所述第二深度卷积神经网络包括第二深度卷积网络和四个并联的第三多层感知器,所述四个第三多层感知器分别用于判断该人头块的左眼、右眼、鼻子和嘴巴是否被遮挡。

[0016] 优选的,所述步骤S01中通过基于边缘的预选区域推荐器对输入图像进行分块,通过计算滑窗窗口内边缘个数,打分并排序,过滤低分的候选框;所述基于边缘的预选区域推荐器的参数 α 为0.65, β 为0.75,其中, α 为区域推荐器扫框的密度, β 为非最大值抑制的阈值。

[0017] 优选的,所述步骤S02中训练第一深度卷积神经网络得到所需参数包括,第一深度卷积神经网络通过通用的数据库训练至收敛;然后载入人脸遮挡数据,继续训练至收敛;通过前传方法得到训练误差,通过反向传播方法更新各层的权重参数值。

[0018] 优选的,所述第一深度卷积神经网络包括依次连接的卷积层、激活层、池化层、全连层、输出层和代价函数层;所述全连层为多层感知器中的一层;

[0019] 所述卷积层表达式:

$$[0020] \quad y = bc + wc * x \quad (1)$$

[0021] 式中 x 和 y 分别为输入和输出, wc 为深度卷积网络的权重参数, bc 为卷积神经网络权重的偏置参数;

[0022] 所述激活层表达式:

$$[0023] \quad y = f_{\text{ReLU}}(x) = \max(x, 0) \quad (2)$$

[0024] 式中 x 和 y 分别为输入和输出;

[0025] 所述池化层表达式:

$$[0026] \quad y = f_{\max}(x) = \max\{x_{s,s}\}$$

[0027] 式中,输出 y 由输入图像 $x_{s,s}$ 上大小为 $s \times s$ 的区域池化而成;

[0028] 所述全连层表达式:

$$[0029] \quad y = f(x) = bf + wf * x \quad (3)$$

[0030] 式中, wf 是全连层中的权重参数, bf 是偏置参数;

[0031] 所述输出层表达式:

$$[0032] \quad y^{i0} = f(x^{i0}) = \frac{\exp(x^{i0})}{\sum_{i0=1}^K \exp(x^{i0})} \quad (4)$$

[0033] 式中, x^{i0} 是第 $i0$ 类的输入, K 是总类数, y^{i0} 是第 $i0$ 类的输出得分;

[0034] 所述代价函数层表达式:

$$[0035] \quad Y_{\text{loss}} = - \sum_{j=1}^K [Y_G^j \log(Y^j) + (1 - Y_G^j) \log(1 - Y^j)] \quad (5)$$

[0036] 式中, Y_{loss} 是代价值, Y^j 和 Y_G^j 分别是实际输出和真实标签第 j 个神经元的值。

[0037] 优选的,根据链式法则逐步反向传播误差,更新各层的权重参数值;反向传播表达式如下:

$$[0038] \quad \nabla \mathbf{w} = -\gamma \frac{\partial E}{\partial \mathbf{w}} \quad (6)$$

[0039] 式中 $\nabla \mathbf{w}$ 为加在原权重上的权重变化值, γ 为学习速率, E 为误差, $\frac{\partial E}{\partial \mathbf{w}}$ 为计算误差对权重 w 的偏导数。

[0040] 优选的,所述步骤S03包括,将人头的位置转换成一个连续值;训练一个包含四个输出的第二多层感知器直至收敛,分别预测人头最小外接框中心点坐标 (x_p, y_p) 和宽高 (wid, hei) ,变换表达式如下:

$$[0041] \quad \hat{x}_p = \frac{G_{xp} - P_{xp}}{P_{wid}} \quad (7)$$

$$[0042] \quad \hat{y}_p = \frac{G_{yp} - P_{yp}}{P_{hei}} \quad (8)$$

$$[0043] \quad \hat{wid} = \log(G_{wid} / P_{wid}) \quad (9)$$

$$[0044] \quad \hat{hei} = \log(G_{hei} / P_{hei}) \quad (10)$$

[0045] 式中 $\hat{*}$ 是回归器的目标值,其中 $*$ 是 x 、 y 、 w 、 h 之一。

[0046] 优选的,所述步骤S04中非极大值抑制的面积重叠比例阈值为0.3。

[0047] 优选的,所述步骤S05包括:

[0048] 构建第二深度卷积神经网络;

[0049] 运用知识迁移策略,载入步骤S02中最终收敛的第一深度卷积神经网络的参数至第二深度卷积神经网络的相应层中;

[0050] 将分割得到的人头块加载入此第二深度卷积神经网络中,训练至收敛;

[0051] 一个多层感知器完成一个任务,每个任务间分享自身的同异,多任务学习方法表达式:

$$[0052] \quad L_{\text{joint}} = \sum_{it=1}^{N_t} \theta_{it} L_{it} \quad (11)$$

[0053] 式中 L_{joint} 表示所有任务的代价总和值, θ_{it} 和 L_{it} 表示任务 it 的权重和代价值, N_t 为任务的总数,为4。

[0054] 与现有技术相比,本发明的优点是:

[0055] 该方法能实现基于自动取款机视频场景下的人脸遮挡检测,能准确检测遮挡的人脸,并且判断其具体的遮挡部位,主要用于自动取款机前摄像机视频的犯罪预警。

[0056] 该方法载入通用的预训练模型减少了训练数据,使用基于边缘的预选区域推荐器降低了分类器在复杂场景下的计算复杂度,通过深度卷积神经网络提取了更具类别差异的特征,借用多任务学习策略共享多任务间的差异信息提高了每个任务的性能。

附图说明

[0057] 下面结合附图及实施例对本发明作进一步描述:

- [0058] 图1为本发明基于深度卷积神经网络的人脸遮挡检测方法的流程图；
- [0059] 图2为人脸遮挡检测数据库；
- [0060] 图3为深度卷积神经网络模型(网络1)的结构图；
- [0061] 图4为基于多任务学习和深度卷积神经网络(网络2)的结构图；
- [0062] 图5为基于边缘的目标区域推荐器在人脸遮挡检测数据库中三组 α 和 β 值的测试性能；
- [0063] 图6为基于边缘的目标区域推荐器在人脸遮挡检测数据库中K的测试召回率；
- [0064] 图7为人头检测结果示例图。

具体实施方式

[0065] 为使本发明的目的、技术方案和优点更加清楚了,下面结合具体实施方式并参照附图,对本发明进一步详细说明。应该理解,这些描述只是示例性的,而并非要限制本发明的范围。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本发明的概念。

[0066] 如图1所示,本发明基于深度卷积神经网络的人脸遮挡检测方法,包括以下步骤:

[0067] 步骤1基于边缘的预选区域推荐方法(EdgeBoxes)分块输入图像

[0068] 对图2所示的输入图像采用基于边缘的预选区域推荐方法进行分块。其分块出来的区域几乎一定包含后续分类器需要的目标区域。基于边缘的预选区域推荐方法是一个不需要学习的方法,其结合滑窗,通过计算窗口内边缘个数进行打分排序来过滤掉低分的候选框。

[0069] 步骤2基于深度卷积神经网络(depp convolutional neural networks)的特征提取和分类

[0070] 深度卷积神经网络模型用于提取由步骤1推荐的区域的特征,同时,一个多层感知器紧连在卷积网络后。深度卷积网络和多层感知器连成一个整体,深度卷积神经网络抽取目标预选区域的特征,多层感知器依据以上所得特征区分预选区域是否为所需目标。但是此模型需要就所处理的不同问题学习出所需参数。

[0071] 首先构建一个随机初始化的深度卷积网络和一个多层感知器命名为深度卷积网络1。为了缓解训练样本不足问题,根据知识迁移策略,这个随机初始的模型被一个通用的数据库训练至收敛。然后,载入人脸遮挡数据,继续训练此模型直至模型收敛。图3为深度卷积神经网络模型的结构图,其逐层依次连接。以下将以公式形式说明网络中的每层,其中包含卷积层、激活层、池化层、全连层、输出层和代价函数层。

[0072] 1) 卷积层表达式:

$$[0073] \quad y = bc + wc * x \quad (1)$$

[0074] 式中 x 和 y 分别为输入和输出, wc 为深度卷积网络的权重参数, bc 为卷积神经网络权重的偏置参数。

[0075] 2) 激活层表达式:

$$[0076] \quad y = f_{\text{ReLU}}(x) = \max(x, 0) \quad (2)$$

[0077] 式中 x 和 y 分别为输入和输出。

[0078] 3) 池化层表达式:

[0079] 池化层表达式:

$$[0080] \quad y = f_{\max}(x) = \max\{x_{s,s}\}$$

[0081] 式中,输出 y 由输入图像 $x_{s,s}$ 上大小为 $s \times s$ 的区域池化而成。

[0082] 4) 全连层表达式:

$$[0083] \quad y = f(x) = bf + wf * x \quad (3)$$

[0084] 式中, wf 是全连层中的权重参数, bf 是偏置参数;由公式可知,全连层位为步骤S02中提到的多层感知器中的一层。

[0085] 5) 输出层表达式:

$$[0086] \quad y^{io} = f(x^{io}) = \frac{\exp(x^{io})}{\sum_{io=1}^K \exp(x^{io})} \quad (4)$$

[0087] 式中, x^{io} 是第 io 类的输入, K 是总类数, y^{io} 是第 io 类的输出得分;

[0088] 6) 代价函数层表达式:

$$[0089] \quad Y_{\text{loss}} = - \sum_{j=1}^K [Y_G^j \log(Y^j) + (1 - Y_G^j) \log(1 - Y^j)] \quad (5)$$

[0090] 式中, Y_{loss} 是代价值。 Y^j 和 Y_G^j 分别是实际输出和真实标签第 j 个神经元的值。

[0091] 图3描述了此过程步骤2,卷积层1(Convolution 1)到池化层(Maxpooling 3)用于抽取特征,全连层1(FC 1)到全连层4(FC 4)用于分类;其中4层全连层构成一个层数为4的多层感知器。

[0092] 为了训练模型,不仅需要以上的前传方法(feed forward)得到训练误差,还需要反向传播方法(back propagation)更新各层的权重参数值。本发明根据链式法则(chain rule),逐步反向传播误差,更新各层的权重参数值。反向传播表达式如下:

$$[0093] \quad \nabla w = -\gamma * \frac{\partial E}{\partial w} \quad (6)$$

[0094] 式中 ∇w 为加在原权重上的权重变化值, γ 为学习速率, E 为误差, $\frac{\partial E}{\partial w}$ 为计算误差对权重 w 的偏导数。

[0095] 步骤3目标位置预测(bounding box prediction)

[0096] 目标位置预测利用步骤2中提取的特征,运用多层感知器预测目标位置。大多由基于边缘的区域推荐器分块的区域都很靠近目标最小外接框。此问题可以当做回归(regression)问题来解决,将目标的位置转换成一个连续值;训练一个包含四个输出的多层感知器,分别预测目标最小外接框中心点坐标(x_p, y_p)和宽高(wid, hei),变换表达式如下:

$$[0097] \quad \widehat{x_p} = \frac{G_{x_p} - P_{x_p}}{P_{wid}} \quad (7)$$

$$[0098] \quad \widehat{y_p} = \frac{G_{y_p} - P_{y_p}}{P_{hei}} \quad (8)$$

$$[0099] \quad \widehat{wid} = \log(G_{wid} / P_{wid}) \quad (9)$$

$$[0100] \quad \widehat{h_{ei}} = \log(G_{hei}/P_{hei}) \quad (10)$$

[0101] 式中 $\hat{h_{ei}}$ 是回归器 (regressor) 的目标值, 其中 $*$ 是 x, y, w, h 之一。

[0102] 训练此回归器的方式与步骤2中的前传方法和反向传播方法相同, 因此, 训练方式的表达式请参照步骤2。

[0103] 步骤4非极大值抑制 (NMS) 去除重叠和重复的检测结果

[0104] 至此, 基于边缘的区域推荐器分块的区域都被计算了一个目标置信度和位置。但是对于每一个目标, 有许多重叠了的重复区域。非极大值抑制根据置信度由高到底排序, 两两比较临近的区域重叠比例 (Intersection over Union of area (IoU)), 去除重叠比例大于阈值的重复检测结果。最后, 为每个目标得到唯一的、精确的位置。

[0105] 步骤5基于多任务学习和深度卷积神经网络的遮挡位置分类

[0106] 依据步骤4得到的目标位置, 联合原图, 分割得到目标块。然后, 先构建一个随机初始化的深度卷积网络和四个多层感知器, 命名为深度卷积网络2, 四个多层感知器分别用于判断此人的左眼, 右眼, 鼻子和嘴巴是否被遮挡。接下来, 运用知识迁移策略, 载入步骤2中最终收敛的深度卷积神经网络1的参数到深度卷积神经网络2的相应层中, 具体层数为卷积层1至池化层3。下一步, 将以上分割得到的目标块加载入此网络中, 训练此网络直至收敛。图4为深度卷积神经网络的结构图, 其逐层依次连接。整个网络包含卷积层、激活层、池化层、全连层、输出层和代价函数层, 网络中每层的公式形式说明与步骤2相同, 详细公式请参照步骤2。

[0107] 如图4, 输入的是整张人头照片, 输出为四个并联的多层感知器, 一个多层感知器完成一个任务, 每个任务间将分享自身的同异, 最终达到每个任务都得到提升的效果。每个多层感知器利用同一个深度卷积神经网络提取的特征, 这就是本发明使用的多任务学习方法。以下是多任务学习方法表达式:

$$[0108] \quad L_{\text{joint}} = \sum_{it=1}^{Nt} \theta_{it} L_{it} \quad (11)$$

[0109] 式中 L_{joint} 表示所有任务的代价总和值, θ_{it} 和 L_{it} 表示任务 it 的权重和代价值, Nt 为任务的总数, 为4。

[0110] 此步骤中使用的深度卷积神经网络与步骤2中相同, 训练过程请参照步骤2。步骤2和步骤5中使用相同的网络结构不仅避免了重新设计一个网络的重复工作, 也便于运用知识迁移策略, 将收敛了的深度卷积神经网络1的参数载入随机初始化的深度卷积神经网络2中。但是, 步骤2和步骤5中的深度卷积神经网络结构可以不同。

[0111] 实施例:

[0112] 步骤1: 目标预选区域分块:

[0113] 首先将图4所示数据库中人上半身图片随机分成8:1:1的三份, 分别用于训练, 验证和测试。

[0114] 对640X640大小图片的人上身图片进行基于边缘的预选框分块。

[0115] 基于边缘的目标区域推荐器中两个重要的参数是 α 和 β , α 控制区域推荐器扫框 (sliding window) 的密度, β 是其非最大值抑制的阈值。这两个数值决定了目标区域推荐器的性能。为了均衡计算时间复杂度和计算空间复杂度, 本具体实施例中 α 和 β 分别为0.65和

0.75。图5展示了基于边缘的目标区域推荐器在图2所示的数据库中三组 α 和 β 值的测试性能。纵轴为召回率 (Recall)，召回率就是所有目标在推荐区域与手工标定的目标地点的重叠率大于阈值时有多少比例被推荐器分块出来了，理想的目标区域推荐器应该使得其为1，即所有目标的所在区域都被目标区域推荐器分块出来了。

[0116] 基于边缘的目标区域推荐器为每个分块出来的区域计算一个可能是目标的值，从大到小排序所有块。然后选取前 K_s 个区域，送入后续处理。如图6所示，为了均衡计算的时间和空间复杂度，本发明中 K_s 被设置成500。

[0117] 步骤2:深度卷积神经网络提取特征和分类:

[0118] 首先构建一个深度卷积神经网络，命名为深度卷积网络1，网络的详细参数见表1。然后载入ILSCVR12 (Imagenet Large Scale Visual Recognition 2012) 训练之至模型收敛。

[0119] 初始学习速率为0.01，若连续五次发生验证集错误率上升的情况，则把模型参数还原到五个循环之前，并把学习速度除以2，然后继续训练。梯度下降方法采用随机梯度下降法 (stochastic gradient descent)，批大小为128。

[0120] 随后，将步骤1中计算的数据继续训练，验证和测试网络。训练时设置的参数与上述一致。

[0121]

深度卷积神经网络1详细参数					
	层数	操作类型	核心数	核心大小	特征大小
特征抽取层	1	输入层		100 X 100 图片	
	2	卷积层	8	5 X 5	96 X 96 X 8
	3	最大池化层		2 X 2	48 X 48 X 8
	4	卷积层	16	5 X 5	44 X 44 X 16
	5	最大池化层		2 X 2	22 X 22 X 16
	6	卷积层	32	5 X 5	18 X 18 X 32
	7	最大池化层		2 X 2	9 X 9 X 32
多层感知器层	8	全连层	1024 个神经元		
	9	全连层	128 个神经元		
	10	全连层	16 个神经元		
	11	全连层	2 个神经元		

[0122] 表1:深度卷积神经网络1详细参数

[0123] 步骤3:目标位置预测 (bounding box prediction)

[0124] 保存深度卷积网络1中第7层的输出和步骤1中的相应预选框的位置 (x, y, w, h)，载入步骤2中确认是人头区域的特征，根据表达式7, 8, 9, 10转换预选框位置，训练表2所示回归器直至收敛。

[0125]

回归器	
输入层	2592 个神经元
隐层	512 个神经元
输出层	4 个神经元

[0126] 表2:回归器详细参数

[0127] 步骤4非极大值抑制 (NMS) 去除重叠和重复的检测结果

[0128] 将分类器确认是人头的分类可信度和回归得到的位置输入非极大值抑制，设置面积重叠比例 (IoU) 阈值为0.3，过滤重叠的重复检测框。

[0129] 步骤1,2,3,4连在一起,从人上半身图像检测到人头图像。图7展示了部分检测结果。其中黑色框为人工标定的位置,灰色框为检测结果。

[0130] 表3展示了本发明中的方法与知名的梯度直方图(HOG)结合支持向量机(SVM)的检测方法的性能比较,其表明了此方法是有效的。

	准确率
[0131] 梯度直方图(HOG)和支持向量机(SVM)	72.41%
本专利发明中的方法	85.61%

[0132] 表3:人头检测方法性能比较

[0133] 步骤5基于多任务学习和深度卷积神经网络的遮挡位置分类

[0134] 将以上步骤得到的人头图像随机分成8:1:1的三部分,分别用于训练,验证和测试后续步骤中的模型。

[0135] 首先构建一个随机初始化的深度神经网络,命名为深度卷积网络2,详细参数见表4。然后,将步骤2中收敛了的深度卷积神经网络1的参数载入随机初始化的深度卷积神经网络2中,具体层数为卷积层1至池化层3。然后载入步骤5中准备好的人头图像训练深度卷积神经网络2直至模型收敛。

[0136] 初始学习速率为0.01,若连续五次发生验证集错误率上升的情况,则把模型参数还原到五个循环之前,并把学习速度除以2,然后继续训练。梯度下降方法采用随机梯度下降法(stochastic gradient descent),批大小为128。设置表达式11中的 θ 为0.25,即每个任务对中的总代价的权重是相同的。

[0137] 随后,将步骤5中计算的验证和测试数据用于验证和测试网络。

深度卷积神经网络2详细参数						
	层数	操作类型	核心数	核心大小	特征大小	
[0138] 特征抽取层	1	输入层		100 X 100 图片		
	2	卷积层	8	5 X 5	96 X 96 X 8	
	3	最大池化层		2 X 2	48 X 48 X 8	
	4	卷积层	16	5 X 5	44 X 44 X 16	
	5	最大池化层		2 X 2	22 X 22 X 16	
	6	卷积层	32	5 X 5	18 X 18 X 32	
	7	最大池化层		2 X 2	9 X 9 X 32	
	8	共享层	1024 个神经元			
微调参数			左眼	右眼	鼻子	嘴巴
	9	全连层	128 个神经元	128 个神经元	128 个神经元	128 个神经元
	10	全连层	16 个神经元	16 个神经元	16 个神经元	16 个神经元
	11	全连层	2 个神经元	2 个神经元	2 个神经元	2 个神经元

[0139] 表4:深度卷积神经网络2详细参数

[0140] 表5展示了本方法的检测结果,其充分说明了本方法的有效性。

[0141]	左眼	右眼	鼻子	嘴巴	总计
准确率	98.15%	99.07%	98.15%	99.07%	94.55%

[0142] 表5:基于多任务学习的深度卷积神经网络分类结果

[0143] 表6比较了本发明中的方法与知名的基于海尔特特征(Haar)的维奥拉-琼斯(Viola-

Jones) 人脸检测器的准确率,其进一步表明了本发明中方法的有效性。

	准确率	
[0144]	基于海尔特特征的维奥拉-琼斯人脸检测器	21.20%
	本专利发明中的方法	94.55%

[0145] 表6:人脸遮挡分类方法性能比较。

[0146] 应当理解的是,本发明的上述具体实施方式仅仅用于示例性说明或解释本发明的原理,而不构成对本发明的限制。因此,在不偏离本发明的精神和范围的情况下所做的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。此外,本发明所附权利要求旨在涵盖落入所附权利要求范围和边界、或者这种范围和边界的等同形式内的全部变化和修改例。

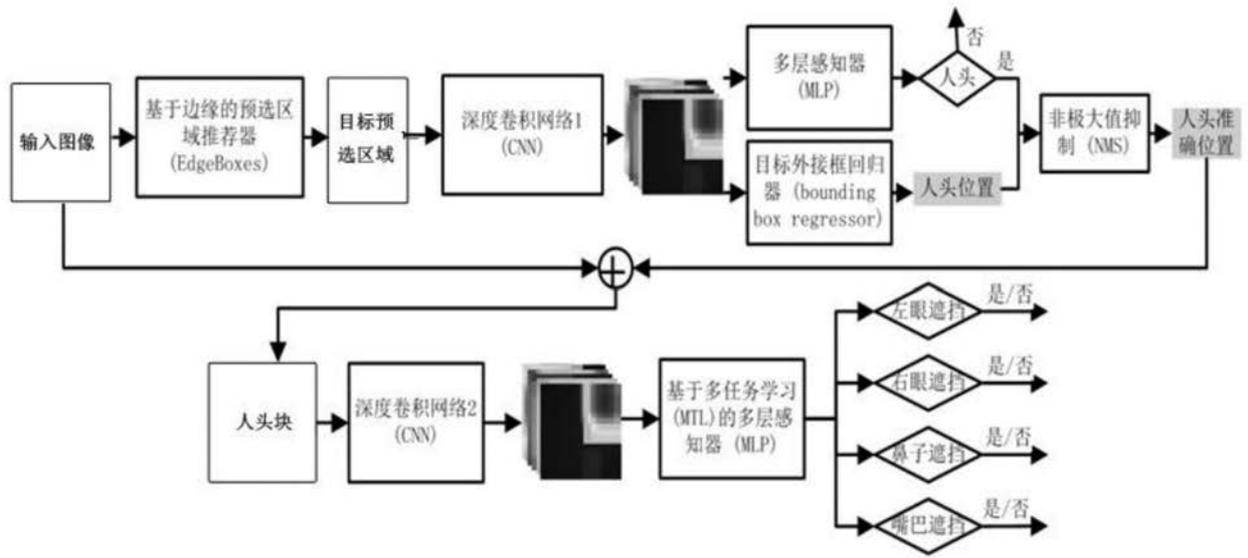


图1



图2

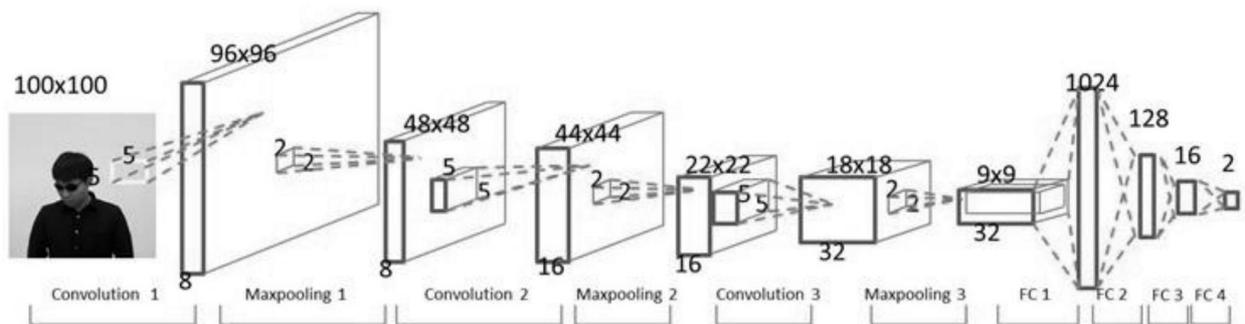


图3

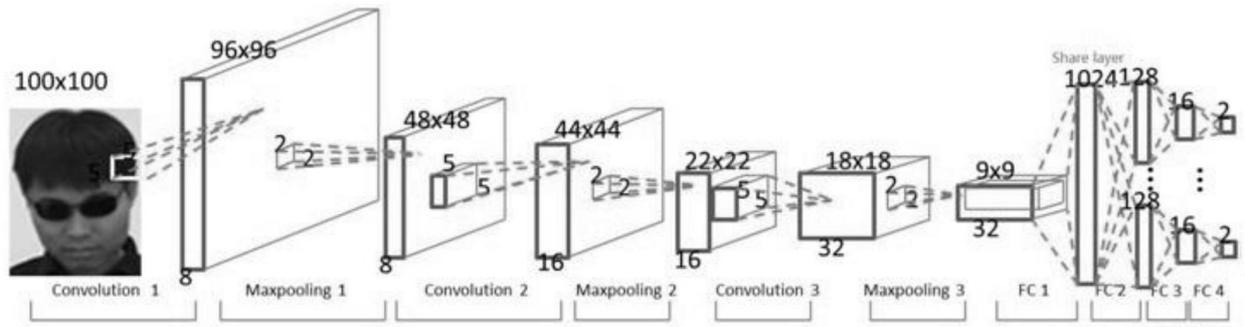


图4

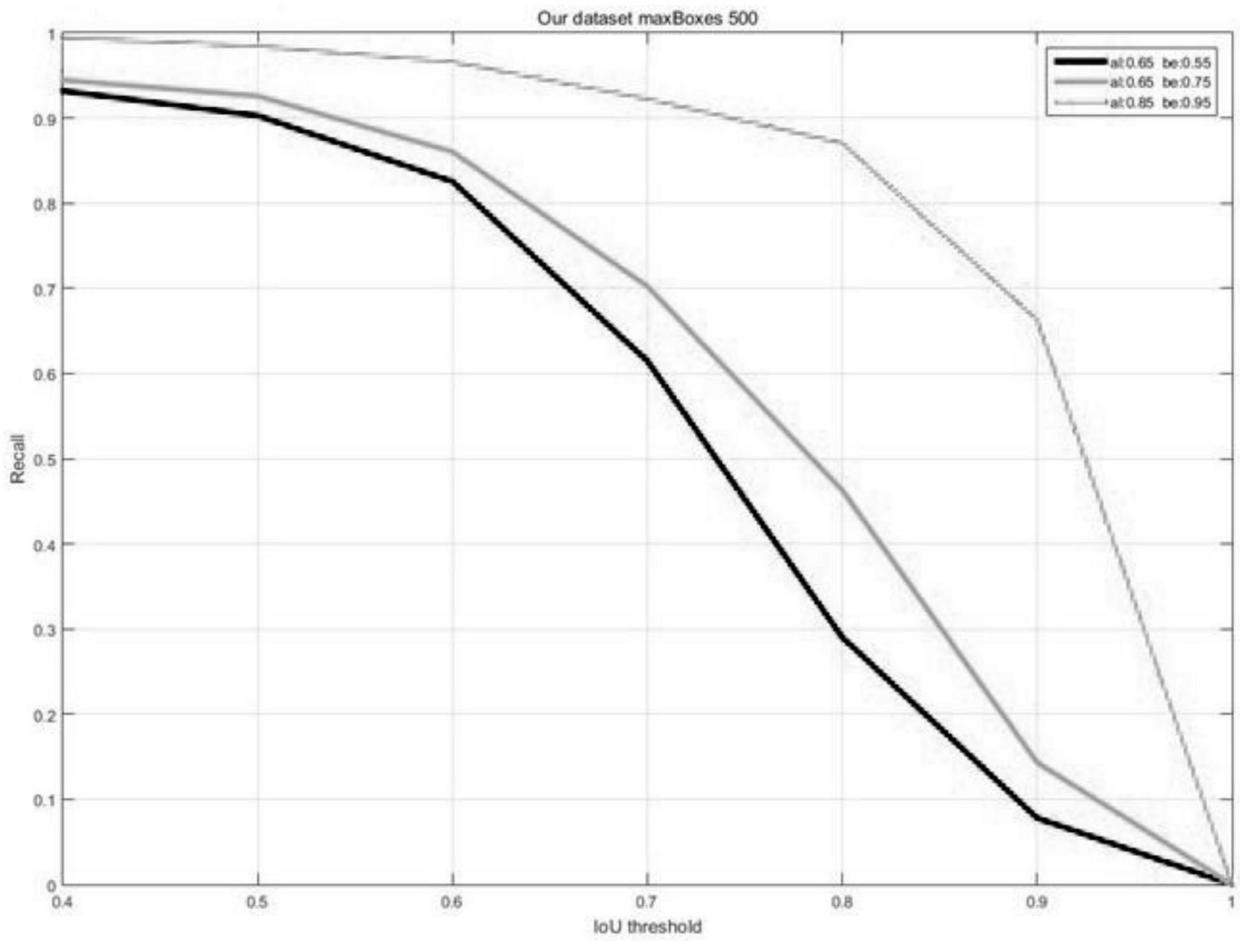


图5

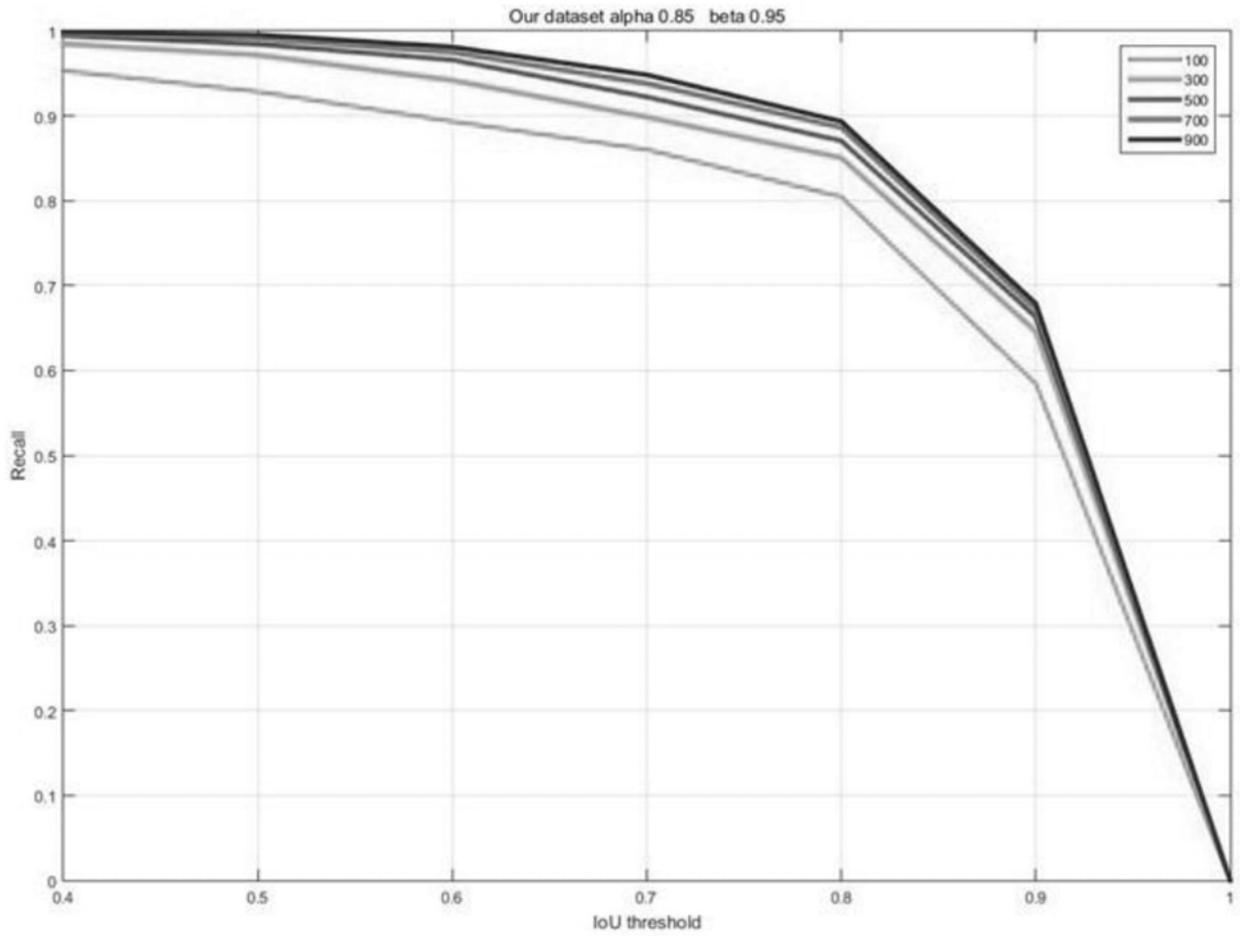


图6



图7