

(19)中华人民共和国国家知识产权局



(12)发明专利申请

(10)申请公布号 CN 109034119 A

(43)申请公布日 2018. 12. 18

(21)申请号 201810980287.7

(22)申请日 2018.08.27

(71)申请人 苏州广目信息技术有限公司

地址 215155 江苏省苏州市相城区望亭镇
迎湖村御亭路666号

(72)发明人 储超群 刘景峰

(74)专利代理机构 广州市红荔专利代理有限公司
44214

代理人 关家强

(51)Int.Cl.

G06K 9/00(2006.01)

G06N 3/04(2006.01)

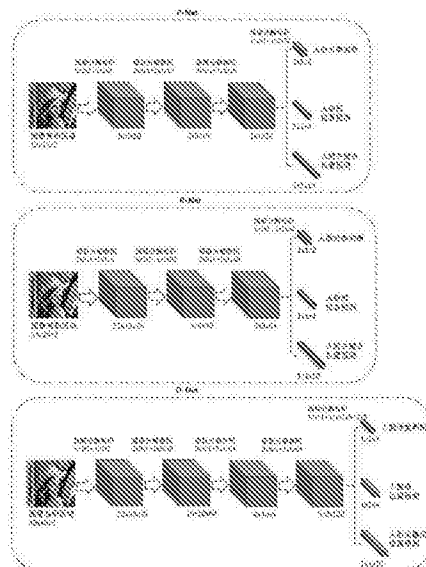
权利要求书2页 说明书7页 附图3页

(54)发明名称

一种基于优化的全卷积神经网络的人脸检测方法

(57)摘要

本发明提出了一种基于优化的全卷积神经网络的人脸检测方法,通过对现有MTCNN模型中的全连接层全部替换为卷积层,并利用级联分解卷积运算代替原有的标准卷积运算,相较于现有MTCNN模型,本发明的运算速度大大提高;此外,通过迭代贪心压缩运算,对模型的参数进行有效的压缩,在保证精度的前提下,降低了参数的规模,减少了模型参数占用的存储空间,增大了模型的通用性和使用范围。



1. 一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,包括如下步骤:

S1. 建立全卷积神经网络模型;

S2. 准备训练数据;

S3. 对所述全卷积神经网络模型进行训练;

S4. 使用训练完成后的全卷积神经网络模型对待测图像进行人脸检测;

其中,所述全卷积神经网络模型为改进的MTCNN模型,所述全卷积神经网络模型包括P-Net网络层,R-Net改进网络层以及O-Net改进网络层;且所述P-Net网络层通过三层卷积层运算后得到人脸区域的候选窗口和边界框的回归向量;所述R-Net改进网络层是指将R-Net网络层中的全连接层替换为卷积层;所述O-Net改进网络层是指将O-Net网络层中的全连接层替换为卷积层;所述P-Net网络层,R-Net改进网络层以及O-Net改进网络层中各卷积层的卷积运算采用级联分解卷积运算,所述级联分解卷积运算包括逐通道卷积运算和逐点卷积运算。

2. 根据权利要求1所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,所述级联分解卷积运算是针对各网络层的各卷积层进行平面和通道上的分解计算,依次进行卷积计算。

3. 根据权利要求1所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤3中具体包括如下步骤:

S31. 利用步骤S2中准备的训练数据,对步骤S1建立的全卷积神经网络模型中P-Net网络层的参数依次进行迭代训练;

S32. 利用迭代贪心压缩算法分别对步骤S31迭代训练完成的参数进行压缩修剪,得到实用的P-Net网络层;

S33. 按照步骤S31,S32的训练方法,利用实用的P-Net网络层对后续的R-Net改进网络层和O-Net改进网络层进行级联训练,得到训练好的全卷积神经网络模型。

4. 根据权利要求3所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤S31的迭代训练采用TensorFlow框架,包括:

S311. 确定全卷积神经网络模型的目标函数;

S312. 使用反向传播和随机梯度下降的方式对模型参数进行迭代训练;

其中,所述全卷积神经网络模型的目标函数为:

$\min(\sum_{i=1}^3(a_i \cdot Loss^i))$; a_i 为P-Net网络层,R-Net改进网络层以及O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测对应的权重值; $Loss^i$ 为P-Net网络层,R-Net改进网络层以及O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测对应的损失函数;

$Loss^1$ 为人脸分类预测的损失参数,可以表示为 $Loss^1 = -(\log(p) + (1-X)(1-\log(p)))$;

$Loss^2$ 为人脸框的位置预测的损失函数,可以表示为: $Loss^2 = \|Y_0 - Y\|_2^2$;

$Loss^3$ 为人脸关键点的位置预测的损失函数,可以表示为: $Loss^3 = \|Z_0 - Z\|_2^2$,其中, p 表示P-Net网络层,R-Net改进网络层以及O-Net改进网络层的输出; X 为训练样本的标注信息;1表示是人脸,0表示不是人脸; Y_0 为四维数值向量,表示各网络层对于人脸位置的预测值; Y 为四维数值向量,表示训练样本对应的标注人脸位置信息; Z_0 为十维数值变量,表

示各网络层对于人脸五个关键点位置的实际预测值; Z 为十维数值变量,表示训练样本对应的标注人脸关键点信息。

5. 根据权利要求4所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,所述P-Net网络层和R-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测的权重系数 a_i 均分别为1.0,0.5,0.5;所述O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测的权重系数 a_i 分别为0.5,1.0,1.0。

6. 根据权利要求3所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤S32中的迭代贪心压缩算法的流程如下:

- S321. 从单个网络层的第一个特征图开始压缩;
- S322. 通过置0的方式,得到当前卷积层的各个特征图对应的测试精度损失;
- S323. 将精度损失最小的特征图对应的卷积核删除;
- S324. 对应修改当前卷积层的前后层的大小,并对该网络层进行重新训练至收敛;
- S325. 判断是否达到终止条件,若是则对本网络层的压缩结束;若否,则转到下一步;
- S326. 判断是否达到最后一个卷积层,若是,则重新开始步骤S322;若否,则将后一卷积层设置为当前层,转至S322。

7. 根据权利要求6所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤S325中的终止条件为:所述全卷积神经网络模型的精度损失达到设定的不可接受的阈值,或者压缩比例达到设定的期望值。

8. 根据权利要求1所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤S4包括如下具体步骤:

- S41. 构建图像金字塔;
- S42. 级联的人脸检测,将图像金字塔内的图片依次输入到训练后的全卷积神经网络模型中的P-Net网络层中,并汇总不同尺度图片对应的人脸候选框;
- S43. 将P-Net网络层的输出作为R-Net改进网络层的输入,得到优化后的人脸候选框;
- S44. 将R-Net改进网络层的输出作为O-Net改进网络层的输入,得到最终的人脸检测结果。

9. 根据权利要求8所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤S41的具体步骤包括:

- S411. 确定图像金字塔的基础缩放尺度,所述基础缩放尺度为 $M/12$;其中 M 为图片中需要检测到的最小人脸大小 $M*M$;
- S412. 构建图像尺度金字塔:将所述基础缩放尺度作为初始值,按照缩放后的图片的面积是前一层的 $1/2$ 的迭代策略,直到缩放后的图片的较短边的长度小于12停止迭代;
- S423. 使用构建好的图像尺度金字塔作为缩放比例,将待测图片缩放到不同的大小,已建立该待测图片的图像金字塔。

10. 根据权利要求8所述的一种基于优化的全卷积神经网络的人脸检测方法,其特征在于,步骤S42,S43和S44中,采用NMS算法实现去重。

一种基于优化的全卷积神经网络的人脸检测方法

技术领域

[0001] 本发明涉及人脸识别技术领域,尤其是涉及一种基于优化的全卷积神经网络的人脸识别方法。

背景技术

[0002] 人脸检测这一任务在计算机视觉相关的研究领域中是一个十分具有代表性的物体检测任务,历经几十年的发展,整体的检测效果也日趋成熟。以深度学习技术的出现为分界点,人脸检测算法可以大致分为基于传统机器学习方法的人脸检测算法和基于深度学习模型的人脸检测方法。基于传统机器学习方法的人脸检测算法在速度上已经可以满足大部分应用的实际要求,然而其检测精度还不能很好地满足工业界的需要,应用的场景和最终的分类效果也不够好。深度学习的出现,在一定程度上很好地弥补了上述传统机器学习方法中的缺陷。

[0003] MTCNN(多任务级联卷积网络)模型是目前在精度和速度上表现最好的深度模型之一,采用了级联网络的设计,同时配合了多任务学习的方式进行网络训练,由于采用多个小型网络的设计也使得整体的运算速度比较快,经过测试,目前算法在普通的个人计算机上运行,对一张大小为640*480的图片进行人脸检测,耗时约为90ms。

[0004] 但在当下实际的应用中,对于人脸检测算法的需求越来越多的是来自于前端,包括嵌入式设备、移动手机等等,这就意味着算法的运行环境会更加苛刻,运算资源和存储资源也会更加有限。在这些实际的应用环境限制之下,目前深度学习的人脸检测方法都显得不太实用,需要更进一步的改进,主要的问题总结如下:

[0005] (1)深度模型的运算过程仍然显得有些复杂,在运算环境较差的前端上,算法的运算速度难以达到实时或可以满足应用需要的速度。因此,需要在不降低检测精度的同时,对已有的深度学习人脸检测算法进行速度上的优化;

[0006] (2)深度学习模型的参数较多,使得模型的参数文件需要占用较多的存储空间,这对于前端设备而言也会带来更高的应用门槛,从实用性的角度出发,需要对目前的深度模型进行参数压缩,降低模型带来的存储压力。

发明内容

[0007] 为解决以上问题,本发明提出了一种基于优化的全卷积神经网络的人脸检测方法。

[0008] 本发明的主要内容包括:

[0009] 一种基于优化的全卷积神经网络的人脸检测方法,包括如下步骤:

[0010] S1.建立全卷积神经网络模型;

[0011] S2.准备训练数据;

[0012] S3.对所述全卷积神经网络模型进行训练;

[0013] S4.使用训练完成后的全卷积神经网络模型对待测图像进行人脸检测;

[0014] 其中,所述全卷积神经网络模型为改进的MTCNN模型,所述全卷积神经网络模型包括P-Net网络层,R-Net改进网络层以及O-Net改进网络层;且所述P-Net网络层通过三层卷积层运算后得到人脸区域的候选窗口和边界框的回归向量;所述R-Net改进网络层是指将R-Net网络层中的全连接层替换为卷积层;所述O-Net改进网络层是指将O-Net网络层中的全连接层替换为卷积层;所述P-Net网络层,R-Net改进网络层以及O-Net改进网络层中各卷积层的卷积运算采用级联分解卷积运算,所述级联分解卷积运算包括逐通道卷积运算和逐点卷积运算。

[0015] 优选的,所述级联分解卷积运算是指依次进行逐通道卷积和逐点卷积运算是指对各网络层的各卷积层进行平面和通道上的分解计算,依次进行卷积计算。

[0016] 优选的,步骤3中具体包括如下步骤:

[0017] S31.利用步骤S2中准备的训练数据,对步骤S1建立的全卷积神经网络模型中P-Net网络层的参数依次进行迭代训练;

[0018] S32.利用迭代贪心压缩算法分别对步骤S31迭代训练完成的参数进行压缩修剪,得到实用的P-Net网络层;

[0019] S33.按照步骤S31,S32的训练方法,利用实用的P-Net网络层对后续的R-Net改进网络层和O-Net改进网络层进行级联训练,得到训练好的全卷积神经网络模型。

[0020] 优选的,步骤S31的迭代训练采用TensorFlow框架,包括:

[0021] S311.确定全卷积神经网络模型的目标函数;

[0022] S312.使用反向传播和随机梯度下降的方式对模型参数进行迭代训练;

[0023] 其中,所述全卷积神经网络模型的目标函数为:

[0024] $\min(\sum_{i=1}^3(a_i \cdot \text{Loss}^i))$; a_i 为P-Net网络层,R-Net改进网络层以及O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测对应的权重值; Loss^1 为P-Net网络层,R-Net改进网络层以及O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测对应的损失函数;

[0025] Loss^1 为人脸分类预测的损失参数,可以表示为 $\text{Loss}^1 = -(\log(p) + (1-X)(1-\log(p)))$;

[0026] Loss^2 为人脸框的位置预测的损失函数,可以表示为: $\text{Loss}^2 = \|Y_0 - Y\|_2^2$;

[0027] Loss^3 为人脸关键点的位置预测的损失函数,可以表示为: $\text{Loss}^3 = \|Z_0 - Z\|_2^2$,其中, p 表示P-Net网络层,R-Net改进网络层以及O-Net改进网络层的输出; X 为训练样本的标注信息;1表示是人脸,0表示不是人脸; Y_0 为四维数值向量,表示各网络层对于人脸位置的实际预测值; Y 为四维数值向量,表示训练样本对应的标注人脸位置信息; Z_0 为十维数值变量,表示各网络层对于人脸五个关键点位置的实际预测值; Z 为十维数值变量,表示训练样本对应的标注人脸关键点信息。

[0028] 优选的,所述P-Net网络层和R-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测的权重系数 a_i 均分别为1.0,0.5,0.5;所述O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测的权重系数 a_i 分别为0.5,1.0,1.0。

[0029] 优选的,步骤S32中的迭代贪心压缩算法的流程如下:

- [0030] S321.从单个网络层的第一个特征图开始压缩;
- [0031] S322.通过置0的方式,得到当前卷积层的各个特征图对应的测试精度损失;
- [0032] S323.将精度损失最小的特征图对应的卷积核删除;
- [0033] S324.对应修改当前卷积层的前后层的大小,并对该网络层进行重新训练至收敛;
- [0034] S325.判断是否达到终止条件,若是则对本网络层的压缩结束;若否,则转到下一步;
- [0035] S326.判断是否达到最后一个卷积层,若是,则重新开始步骤S322;若否,则将后一卷积层设置为当前层,转至S322。
- [0036] 优选的,步骤S325中的终止条件为:所述全卷积神经网络模型的精度损失达到设定的不可接受的阈值,或者压缩比例达到设定的期望值。
- [0037] 优选的,步骤S4包括如下具体步骤:
- [0038] S41.构建图像金字塔;
- [0039] S42.级联的人脸检测,将图像金字塔内的图片依次输入到训练后的全卷积神经网络模型中的P-Net网络层中,并汇总不同尺度图片对应的人脸候选框;
- [0040] S43.将P-Net网络层的输出作为R-Net改进网络层的输入,得到优化后的人脸候选框;
- [0041] S44.将R-Net改进网络层的输出作为O-Net改进网络层的输入,得到最终的人脸检测结果。
- [0042] 优选的,步骤S41的具体步骤包括:
- [0043] S411.确定图像金字塔的基础缩放尺度,所述基础缩放尺度为 $M/12$;其中M为图片中需要检测到的最小人脸大小 $M*M$;
- [0044] S412.构建图像尺度金字塔:将所述基础缩放尺度作为初始值,按照缩放后的图片的面积是前一层的 $1/2$ 的迭代策略,直到缩放后的图片的较短边的长度小于12停止迭代;
- [0045] S423.使用构建好的图像尺度金字塔作为缩放比例,将待测图片缩放到不同的大小,已建立该待测图片的图像金字塔。
- [0046] 优选的,步骤S42,S43和S44中,采用NMS算法实现去重。
- [0047] 与现有技术相比,本发明的有益效果在于:
- [0048] (1)通过卷积运算优化的方式,利用级联分解卷积代替原有的标准卷积,以及对原有模型的全卷积结构改造,模型的运算速度得到了有效的提升。相比于原来的MTCNN模型,在同样的计算环境下,算法运算速度达到了原有模型的3倍,目前可以支持在普通的智能手机上实时进行人脸检测。
- [0049] (2)通过提出的迭代贪心压缩算法,对原有的模型进行了有效地参数压缩。相比于原来的MTCNN模型,在精度损失不超过5%的情况下,参数规模为原有参数规模的 $1/3$,在精度损失不超过10%的情况下,参数规模为原有规模的 $1/5$,在确保检测精度的情况下,目前在智能手机端运行的模型参数大小可以控制在500KB以内。

附图说明

- [0050] 图1为本发明全卷积神经网络的结构示意图;
- [0051] 图2为现有标准卷积示意图;

- [0052] 图3为本发明逐通道卷积示意图；
[0053] 图4为本发明逐点卷积示意图；
[0054] 图5为迭代贪心压缩算法的流程图。

具体实施方式

- [0055] 以下结合附图对本发明所保护的技术方案做具体说明。
- [0056] 请参阅图1至图5。本发明提出了一种基于优化的全卷积神经网络的人脸检测方法，是对现有MTCNN模型的改进，主要包括训练阶段和使用阶段两个部分；主要包括如下步骤：
- [0057] S1. 建立全卷积神经网络模型；
- [0058] 本发明的全卷积神经网络模型是在现有MTCNN模型基础上的改进，具体地，将MTCNN模型中的全连接层全部替换为卷积层，即将现有MTCNN模型中的R-Net网络层的全连接层替换为卷积层，构成R-Net改进网络层，将现有MTCNN模型中的O-Net网络层的全连接层替换为卷积层，构成O-Net改进网络层，而P-Net网络层与MTCNN模型中的P-Net相同。
- [0059] S2. 准备训练数据；
- [0060] 同其他的人脸检测模型一样，首先需要准备大量的训练图片用于训练网络参数，训练图片中都必须包含至少一个人脸，同时需要获取到这些人脸在图片中的位置信息，同时考虑到模型多任务学习设计中的人脸关键点检测任务，因此还需要知道这些人脸对应的关键点位置信息，需要包含5个关键点：左眼、右眼、鼻子、左嘴角、右嘴角。由于在实际场景中，人脸所出现的场景是十分多样的，人脸本身的表情等变化也十分丰富，因此在准备训练数据时，需要尽可能地包含多种不同场景下、不同人脸变化的图片，通过对这些复杂多变的训练数据进行拟合学习，训练出的模型才能在更加复杂的真实场景中表现得更加鲁棒。准备好足够的训练数据之后，就可以按照后续的步骤进行模型的训练。
- [0061] 在其中一个实施例中，训练数据可以来源于Wider_face, CelebA等数据库。
- [0062] S3. 对所述全卷积神经网络模型进行训练；
- [0063] 在本实施例中，对本发明的全卷积神经网络模型进行训练采用TensorFlow框架，为加快模型各网络层的运算速度，对各网络层的卷积运算采用级联分解卷积运算，所述级联分解卷积运算包括逐通道卷积运算和逐点卷积运算，在本实施例中，参照MobileNet模型（请参阅Howard AG, Zhu M, Chen B, et al. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications[J]. arXiv preprint arXiv:1704.04861, 2017.）中深度可分解卷积的设计，将现有MTCNN模型中的标准卷积进行平面和通道上的分解计算，依次进行卷积计算。
- [0064] 在标准的卷积运算中，面对多个特征图的输入，对应的卷积核会有多个多通道的卷积核，如图2所示，一次卷积层的前馈计算需要的计算量是 $S*S*L*K*T$ ；计算量加大，如图3和图4可知，分解后的逐通道卷积运算和逐点卷积运算的计算量分别为 $S*S*L*T$ 和 $S*S*L*K$ ，故采用级联分解卷积运算的计算量为 $(S*S*L*T+ S*S*L*K)$ ，通过比较本发明的级联分解卷积运算和现有的标准卷积运算的计算量可知，两者的比值为 $(1/K+1/T^2)$ ，而通常卷积核的个数K值较大，故现有的标准卷积的计算量约为本发明级联分解卷积运算计算量的 T^2 ，由此可知，采用本发明的级联分解卷积运算能够有效的降低卷积部分的计算量。

[0065] 由图1可知,本发明的模型的网络结构,其中,P-Net网络层的第一层卷积层的逐通道卷积运算采用1个3*3的卷积核,而逐点卷积运算采用3个1*1卷积核;第二层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和10个1*1卷积核;第三层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和16个1*1卷积核;而R-Net改进网络层的第一层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和3个1*1卷积核;第二层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和28个1*1卷积核;第三层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和48个1*1卷积核;所述O-Net改进网络层的第一层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和3个1*1卷积核;第二层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和32个1*1卷积核;第三层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和64个1*1卷积核;第四层卷积层的逐通道卷积运算和逐点卷积运算分别采用1个3*3的卷积核和64个1*1卷积核。

[0066] 由以上可知,虽然各网络层的结构略有不同,但三者的预测任务均由人脸的分类预测、人脸的位置预测以及人脸关键点的位置预测三个子任务组合而成;其中,所述人脸的分类预测是指判断对应区域中是否含有人脸;所述人脸的位置预测是指若含有人脸的话,则预测对应人脸在图像中的具体位置;所述人脸关键点的位置预测是指若含有人脸的话,则预测人脸中的五个关键点在图像中的具体位置;故将以上三个子任务作为本网络的预测目标。

[0067] 具体地,本步骤包括如下具体步骤:

[0068] S31.利用步骤S2中准备的训练数据,对步骤S1建立的全卷积神经网络模型中P-Net网络层的参数依次进行迭代训练;

[0069] S32.利用迭代贪心压缩算法分别对步骤S31迭代训练完成的参数进行压缩修剪,得到实用的P-Net网络层;

[0070] S33.按照步骤S31,S32的训练方法,利用实用的P-Net网络层对后续的R-Net改进网络层和O-Net改进网络层进行级联训练,得到训练好的全卷积神经网络模型。

[0071] 由于本发明提出的全卷积神经网络模型的三个网络层为级联结构,即P-Net网络层的输出作为R-Net改进网络层的输入,而R-Net改进网络层的输出作为O-Net网络层的输入;故在对本发明的模型进行训练时,训练过程也采用级联的方式进行,即先完成P-Net网络层的训练,得到实用的P-Net网络层,然后再对后续的R-Net改进网络层进行训练,即需要使用训练好的P-Net网络层对训练数据进行计算,并将输出结果整理好,作为R-Net的训练样本;对于O-Net改进网络层的训练亦如此。

[0072] 对于步骤S31,包括如下更具体的步骤:

[0073] S311.确定全卷积神经网络模型的目标函数;

[0074] S312.使用反向传播和随机梯度下降的方式对模型参数进行迭代训练;

[0075] 其中,所述全卷积神经网络模型的目标函数为:

[0076]
$$\min(\sum_{i=1}^3(a_i \cdot Loss^i));$$
 a_i 为P-Net网络层,R-Net改进网络层以及O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测对应的权重值; $Loss^i$ 为P-Net网络层,R-Net改进网络层以及O-Net改进网络层的人脸分类检测、人脸框的位置预测以及人脸关键点的位置预测对应的损失函数;

[0077] $Loss^1$ 为人脸分类预测的损失参数,可以表示为 $Loss^1 = -(\log(p) + (1-X)(1-\log(p)))$;

[0078] $Loss^2$ 为人脸框的位置预测的损失函数,可以表示为: $Loss^2 = \|Y_0 - Y\|_2^2$;

[0079] $Loss^3$ 为人脸关键点的位置预测的损失函数,可以表示为: $Loss^3 = \|Z_0 - Z\|_2^2$,其中, p 表示P-Net网络层,R-Net改进网络层以及O-Net改进网络层的输出; X 为训练样本的标注信息;1表示是人脸,0表示不是人脸; Y_0 为四维数值向量,表示各网络层对于人脸位置的实际预测值; Y 为四维数值向量,表示训练样本对应的标注人脸位置信息; Z_0 为十维数值变量,表示各网络层对于人脸五个关键点位置的实际预测值; Z 为十维数值变量,表示训练样本对应的标注人脸关键点信息。

[0080] 由于各个网络层的训练侧重点不同,故各个网络层训练时的权重值的偏重不同,如P-Net网络层和R-Net改进网络层的更需要有效地甄别出非人脸的区域,故其权重值更偏向于人脸分类预测任务,在其中一个实施例中,其在三个子任务的权重值 a_i 均分别为1.0,0.5,0.5;而O-Net改进网络层的作为最终的输出网络层,对于检测位置的精度要求更高,故其更偏向于人脸框的位置预测以及人脸关键点的位置预测,在其中一个实施例中,其在三个子任务的权重值 a_i 分别为0.5,1.0,1.0。

[0081] 步骤32用于完成对各网络层中参数的修剪,采用迭代贪心压缩算法进行运算,如图5所示,具体包括如下步骤:

[0082] S321.从单个网络层的第一个特征图开始压缩;

[0083] S322.通过置0的方式,得到当前卷积层的各个特征图对应的测试精度损失;

[0084] S323.将精度损失最小的特征图对应的卷积核删除;

[0085] S324.对应修改当前卷积层的前后层的大小,并对该网络层进行重新训练至收敛;

[0086] S325.判断是否达到终止条件,若是则对本网络层的压缩结束;若否,则转到下一步;

[0087] S326.判断是否达到最后一个卷积层,若是,则重新开始步骤S322;若否,则将后一卷积层设置为当前层,转至S322。

[0088] 优选的,步骤S325中的终止条件为:所述全卷积神经网络模型的精度损失达到设定的不可接受的阈值,或者压缩比例达到设定的期望值。具体的阈值数值可以根据实际应用的需要来自由设定,但一般要考虑客观的情况,需要遵循压缩比例过大会导致模型精度出现明显下降这一实际的规律。

[0089] 以上,首先通过对P-Net网络层进行训练,并对其参数进行修剪,已建立实用的P-Net网络层,随后,利用实用的P-Net网络层的输出作为R-Net改进网络层的输入,重复步骤S31,和S32,对R-Net改进网络层进行训练,并完成其参数的修剪;最后将R-Net改进网络层的输出作为O-Net改进网络层的输入,对其进行训练,并完成其参数的修剪,从而实现对整个模型的训练;且经过以上训练后,三个网络层的训练都顺利收敛。

[0090] S4.使用训练完成后的全卷积神经网络模型对待测图像进行人脸检测;

[0091] 当完成对模型的全部训练,建立实用的全卷积神经网络模型后,即进入本模型的使用阶段,具体包括如下步骤:

[0092] S41.构建图像金字塔;

[0093] 为了尽可能地提升检测的召回率,在将图片输入到网络之前,需要对图片进行不

同尺度上的伸缩变换,本发明采用建立图像金字塔的方式,以使得网络可以对不同大小的人脸都能有效检出,构建图像金字塔的步骤包括如下步骤:

[0094] S411. 确定图像金字塔的基础缩放尺度,所述基础缩放尺度为 $M/12$;其中 M 为图片中需要检测到的最小人脸大小 $M*M$;由于P-Net网络层的输入大小为 $12*12$,故将基础缩放尺度设定为 $M/12$ 。

[0095] S412. 构建图像尺度金字塔:将所述基础缩放尺度作为初始值,按照缩放后的图片的面积是前一层的 $1/2$ 的迭代策略,直到缩放后的图片的较短边的长度小于12停止迭代;

[0096] S423. 使用构建好的图像尺度金字塔作为缩放比例,将待测图片缩放到不同的大小,已建立该待测图片的图像金字塔。

[0097] S42. 级联的人脸检测,将图像金字塔内的图片依次输入到训练后的全卷积神经网络模型中的P-Net网络层中,并汇总不同尺度图片对应的人脸候选框,由于不同尺度的图片突出了不同尺度的人脸,因此需要将每个尺度图片得到的输出人脸候选框进行汇总,在其中一个实施例中,同时使用NMS算法消除所有候选框中的重复框。

[0098] S43. 将P-Net网络层的输出作为R-Net改进网络层的输入,得到优化后的人脸候选框;

[0099] R-Net改进网络层的主要作用是进一步甄别出P-Net网络层的输出结果中并不包含人脸的候选框,以及尽可能对P-Net网络层中的输出的比较粗糙的人脸框位置进行更精细的调整,缩小人脸框的范围,同样需要对输出结果进行汇总,在其中一个实施例中,同样采用NMS算法实现去重,以得到一些相对比较可能的人脸候选框的结果,作为后续O-Net改进网络层的输入。

[0100] S44. 将R-Net改进网络层的输出作为O-Net改进网络层的输入,得到最终的人脸检测结果;本步骤是对R-Net改进网络层的输出进行最终的分类判断,以及最终的人脸位置的精细化测定,同时还需要对每个输出的人脸框给出对应的五个人脸关键点的位置信息,在其中一个实施例中,同样采用NMS算法实现去重,从而得到最终的人脸检测结果。

[0101] 以上所述仅为本发明的实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

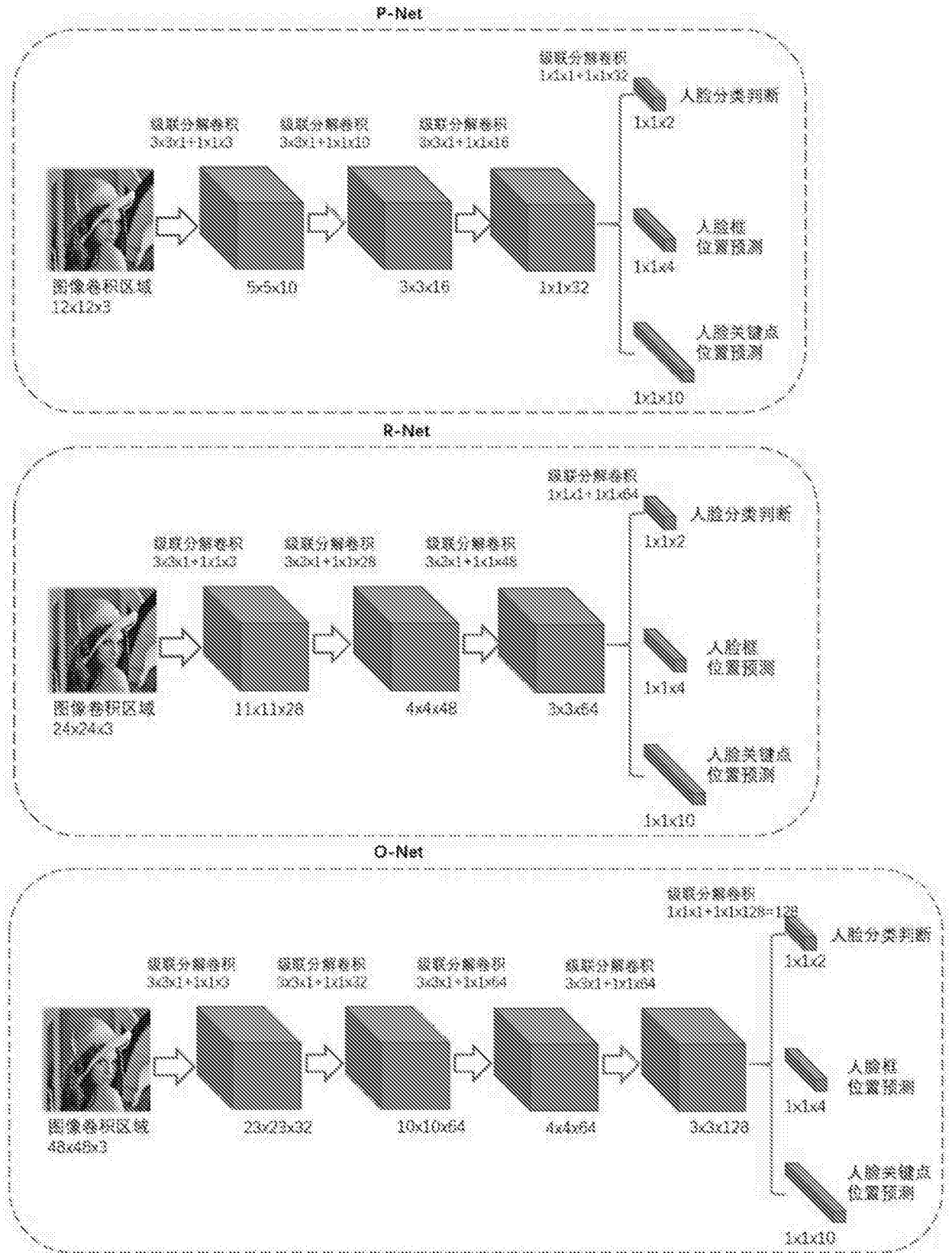


图1

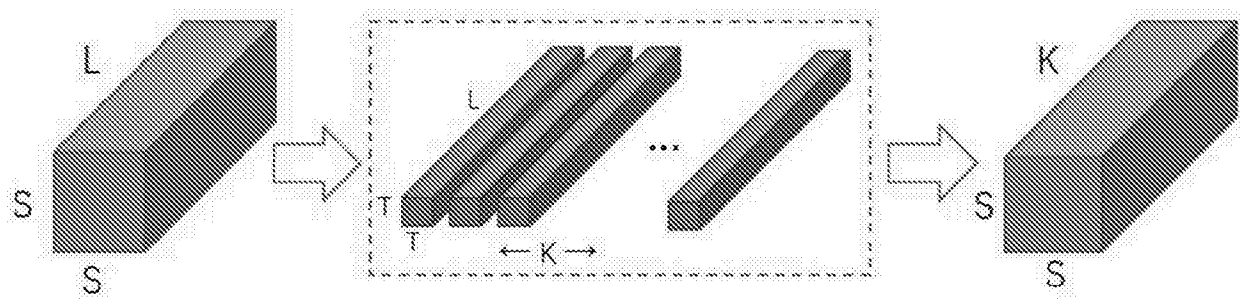


图2

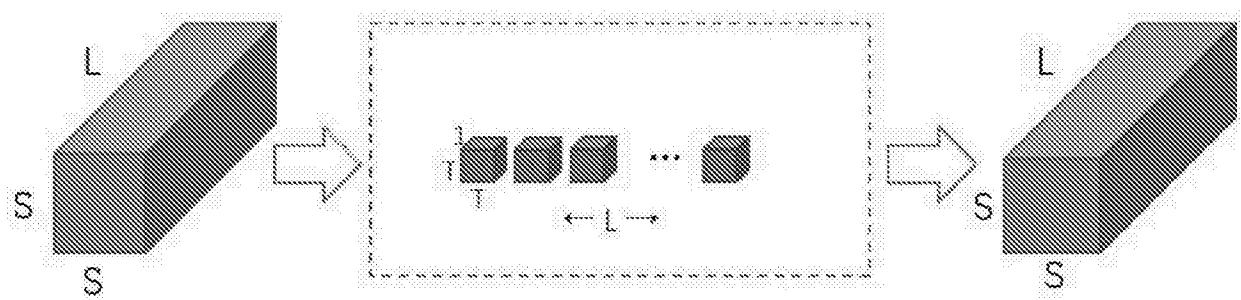


图3

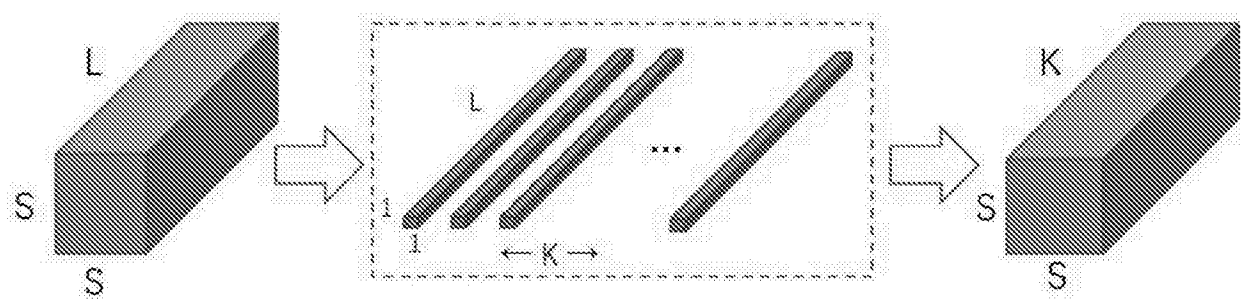


图4

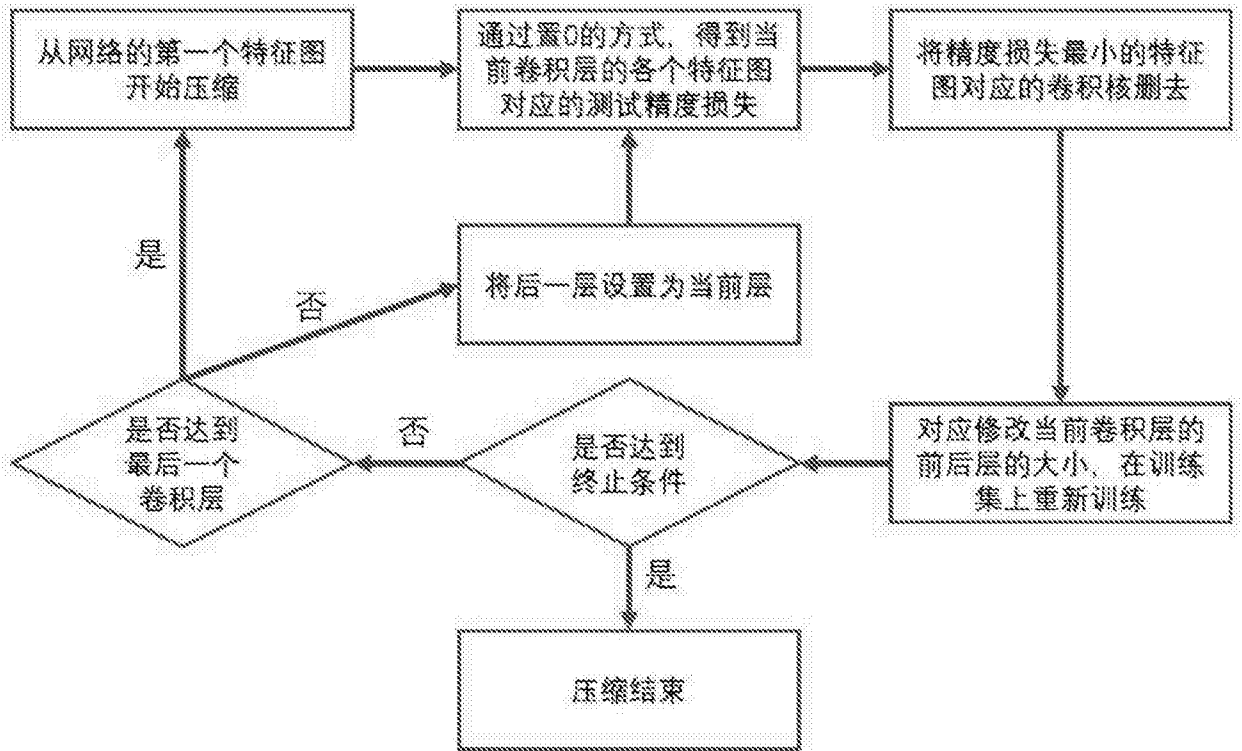


图5