



(12) 发明专利

(10) 授权公告号 CN 109284733 B

(45) 授权公告日 2021.02.02

(21) 申请号 201811197781.2

G07C 1/10 (2006.01)

(22) 申请日 2018.10.15

G06Q 30/02 (2012.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 109284733 A

(56) 对比文件

CN 108510000 A, 2018.09.07

US 2009222388 A1, 2009.09.03

(43) 申请公布日 2019.01.29

CN 107169421 A, 2017.09.15

(73) 专利权人 浙江工业大学
地址 310014 浙江省杭州市下城区潮王路
18号

芮挺等. 基于深度卷积神经网络的行人检测.《计算机工程与应用》.2015, 第162-166页.

Minsung Sung等. Vision based real-time fish detection using convolutional neural network.《IEEE》.2017, 第1-6页.

(72) 发明人 赵云波 林建武 李灏 宣琦

审查员 张倩

(74) 专利代理机构 杭州天正专利事务所有限公司 33201

代理人 王兵 黄美娟

(51) Int. Cl.

G06K 9/00 (2006.01)

G06N 3/04 (2006.01)

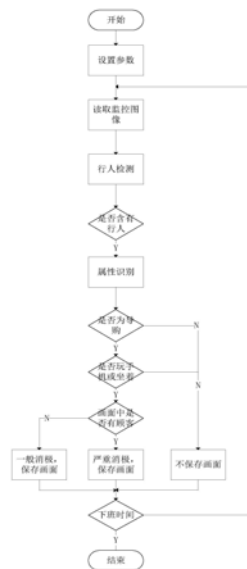
权利要求书2页 说明书7页 附图4页

(54) 发明名称

一种基于yolo和多任务卷积神经网络的导购消极行为监控方法

(57) 摘要

基于yolo和多任务卷积神经网络的导购消极行为监控方法,首先训练一个基于yolo的行人检测模型,用ImageNet和voc2007数据集预训练模型,再用监控场景图像微调模型;然后构建一个基于ResNet50的多任务卷积神经网络,用手工标注过的多标签图像数据训练该网络;接着用rtsp协议读取商场监控画面,用行人检测模型检测画面中的行人,而后再将行人图像输入到多任务卷积神经网络,识别行人是否为导购,是否闲坐,是否在玩手机,以此判断导购是否存在消极行为,并将“严重消极”和“一般消极”的导购画面保存在本地.最终实现用基于yolo的行人检测网络和多任务卷积神经网络对导购消极行为进行有效的监控和记录。



1. 一种基于yolo和多任务卷积神经网络的导购消极行为监控方法, 含有以下步骤:

(1) 训练基于yolo的行人检测模型;

步骤11: 构建基于yolo的行人检测模型;

使用跳层融合的方式, 特征提取阶段一使用13个卷积层和4个最大池化层, 阶段二使用7个卷积层, 在阶段一和阶段二中间有1个最大池化层, 调整阶段一的输出的特征图大小与阶段二输出的特征图大小一致; 然后将这两个调整过大小的特征图通过叠加的方式融合在一起, 成为阶段三的输出; 阶段三有两种模式, 模式一是分类网络, 该模式在预训练模型的时候用到, 具体而言就是一层 3×3 的卷积层, 再加一层全连接层, 全连接层神经元个数等于分类个数; 模式二是检测网络, 该模式是在加载模式一的预训练参数之后, 训练检测网络时用的, 具体而言就是一层 3×3 的卷积层, 再加一层 1×1 的卷积层, 卷积核的个数与检测类别相关, 具体数值为: anchors个数 \times (5+检测类别个数);

步骤12: 用ImageNet数据集预训练分类模型;

用ImageNet数据集去训练一个分类模型, 使训练好的分类模型参数作为检测模型中共有结构的初始化参数;

步骤13: 用voc2007数据集预训练检测模型;

由于检测模型的前几层结构与分类网络一致, 将步骤12中训练好的分类网络的参数作为检测网络中共有结构的初始化参数; voc2007数据集是常用的检测数据集, 共有20类的标注好的检测物体, 其中包括行人图像数据; 只取其中的行人图像数据进行训练, 对行人数据进行数据增强的操作, 调整图像大小为 448×448 , 用SGD优化器训练160个epoch, 初始学习率设置为0.0001;

步骤14: 用监控视角数据集微调模型;

选取BOT2018新零售技术挑战赛中的数据集中的数据集进行微调, 该数据集的行人图像采集自真实商场场景下的监控摄像头; 对该数据集的图像进行水平旋转、中心随机裁剪、HSV空间微调等数据增强的操作, 并将尺寸大小调整为 448×448 ;

加载步骤13训练好的模型, 用SGD优化器训练160个epoch, 初始学习率设置为0.001, 学习率随着训练次数的增多而减小, 0-5个epoch时学习率设置为0.001, 5-80个epoch时学习率设置为0.0001, 80-160个epoch时学习率设置为0.00001;

(2) 训练基于ResNet50的多任务卷积神经网络;

步骤21: 构建基于ResNet50的多任务卷积神经网络;

对步骤(1)中检测到的行人, 需要去识别行人的属性, 以此判断导购是否存在消极工作的行为, 数据集中标注的属性有: “顾客”或“导购”、“男性”或“女性”、“站立”或“坐着”、“玩手机”或“不玩手机”; 这些属性彼此之间无关联性, 视为不相关属性;

针对导购消极行为的识别, 基于ResNet50设计了一种多任务卷积神经网络;

具体地, 将原始ResNet50最后两层的全连接层和池化层去掉, 拼接上四个并行的全连接层, 每个全连接层的神经元个数为2, 分别代表着8种属性: “顾客”与“导购”、“男性”与“女性”、“站立”与“坐着”、“玩手机”与“不玩手机”, 在同一个全连接层上的两个属性是关联属性, 不在一个全连接层上的属性是不相关属性; 每个全连接后分别连一个softmax层; Softmax损失函数的计算公式为:

$$L_i = -\log\left(\frac{e^{x_k}}{\sum_{j=1}^2 e^{x_j}}\right) \quad (1)$$

最终的损失函数值为四个Softmax损失函数值相加,即:

$$\text{Loss} = L_1 + L_2 + L_3 + L_4 \quad (2)$$

步骤22:训练基于ResNet50的多任务卷积神经网络;

加载在ImageNet数据集上训练好的ResNet50中除最后两层外的参数作为多任务卷积神经网络的初始化参数;数据集采用BOT2018新零售技术挑战赛的标注数据,对该数据进行数据增强的操作,以获得更多的可用的训练数据,最后用Adam优化器进行训练,初始学习率设置为0.0005,训练40个epoch;

(3) 导购消极行为记录;

步骤31:读取监控画面;

在商场内,广泛分布的监控系统提供数据;在读取监控画面前,需要设置两个参数:上班时间区间、监控采样时间;设置上班时间区间,让本系统只关注于上班时间;设置监控采样时间,是控制读取监控画面的频率;

通过rtsp协议读取监控画面,通过计算机的系统时间控制读取画面;

步骤32:商场内行人检测;

加载步骤(1)中训练好的行人检测模型,读取步骤31中的监控图像,将图像进行归一化,并转为Tensor,之后将Tensor加载进行人检测模型中,行人检测模型会检测出行人在图像中上下左右四个坐标;

步骤33:识别行人属性;

加载步骤(2)中的多任务卷积神经网络,将步骤32中检测出的行人图像数据作为多任务卷积神经网络的输入数据,输出全连接层的值即为模型对该行人在某个属性上的置信度;若输出“导购”的置信度比“顾客”的置信度高,即识别该行人为“导购”,若输出“男性”的置信度比“女性”的置信度高,即识别该行人为“男性”,若输出“站立”的置信度比“坐”的置信度高,即识别该行人为“站立”,若输出“玩手机”的置信度比“不玩手机”的置信度高,即识别该行人为“玩手机”;

步骤34:记录导购消极行为画面;

在上班时间区间内,对步骤33中识别的行人属性进行导购消极行为的判断;首先,判断该行人的身份,是否属于“导购”,如果是导购,分析其姿势和工作状态,更进一步的,判断该导购所在的画面中,是否存在顾客,顾客在场的情况下,对导购有更严格的要求;当导购在玩手机或者坐着,并且此时画面中没有顾客,认为该导购是“一般消极”;当导购在玩手机或者坐着,而此时画面中有顾客,认为该导购是“严重消极”;具体导购消极行为的程度判定如下:当附近有顾客时,导购坐着或者玩手机的行为判定为“严重消极”,站着并且没有玩手机时判定为“积极”;当附近没有顾客时,导购坐着或者玩手机的行为判定为“一般消极”,站着并且没有玩手机时判定为“积极”;对于“严重消极”的导购,将其画面保存在文件夹1,对于“一般消极”的导购,将其画面保存在文件夹2,对于“积极”的导购,不保存其画面。

一种基于yolo和多任务卷积神经网络的导购消极行为监控方法

技术领域

[0001] 本发明涉及在新零售领域的导购消极行为监控方法。

背景技术

[0002] 随着人工成本的提高,在商场里,招聘更多的导购意味着成本的提升。然而有部分导购存在着消极工作的行为,比如“玩手机”、“附近有顾客时坐着”等消极行为,造成人力资源的浪费。为避免不必要的支出,商场对导购进行有效的考勤管理显得尤为重要。

[0003] 普通的考勤系统只能记录导购的上下班时间,不能自动分析导购在上班时间内是否存在消极工作的情况,也不能记录下导购消极工作时的画面。针对这一需求,本发明利用计算机视觉技术对商场内普遍存在的监控所采集的图像进行图像识别、分析。

[0004] 针对行人检测,已有的方法中,有利用方向梯度直方图作为行人检测的描述子,再用SVM进行分类,这种方法精度不是很高,容易误检。近年来深度卷积神经网络应用在行人检测方面,对行人检测的精度有很大提升,然而由于迁移学习中跨数据集的拟合问题,这种方法在监控视角下缺乏鲁棒性。

[0005] 针对属性识别,卷积神经网络在属性分类的精度上取得传统方法不能比拟的效果。近年来,VGG、ResNet、DenseNet等CNN模型框架已经得到广泛的应用。然而一个原始的ResNet只能对一个属性进行分类,多个属性则需要训练多个模型,这极大地增加了计算的负担。

[0006] 由此可见,对于识别、记录导购消极行为的监控系统,目前还没有完善的解决方案。

发明内容

[0007] 本发明要克服现有技术的上述缺点,提供一种基于yolo和多任务卷积神经网络的导购消极行为监控方法。

[0008] 为实现上述发明目的,本发明设计一种基于yolo和多任务卷积神经网络的导购消极行为监控系统。首先训练一个基于yolo的行人检测模型以及基于ResNe50的多任务卷积神经网络;进一步的,对于监控定时采样的图像,用基于yolo的检测模型检测行人;更进一步的,用一个基于ResNet50的多任务卷积神经网络去识别商场内导购的多种属性和行为,判断是否存在消极行为,并以图片的形式记录下导购做出消极行为的画面。在一定程度上解决了导购消极行为检测以及对上班情况进行自动考勤的问题。可应用在新零售场景中考勤系统、导购管理、店铺运营等方面。

[0009] 本发明解决其技术问题所采用的技术方案是:

[0010] 一种基于yolo和多任务卷积神经网络的导购消极行为监控方法,含有以下步骤:

[0011] 步骤1. 训练基于yolo的行人检测模型:构建基于yolo的行人检测模型,用ImageNet数据集预训练分类模型,用voc2007数据集预训练检测模型,用监控视角数据集微

调模型：

[0012] 步骤2. 训练基于ResNet50的多任务卷积神经网络：构建基于ResNet50的多任务卷积神经网络，训练基于ResNet50的多任务卷积神经网络；

[0013] 步骤3. 导购消极行为记录：读取监控画面，商场内行人检测，识别行人属性，记录导购消极行为画面；

[0014] 与现有技术相比，本发明技术方案的优点有：

[0015] (1) 本发明训练的行人检测模型能在商场监控视角下进行鲁棒的行人检测；

[0016] (2) 本发明训练的多任务卷积神经网络能同时识别行人多个属性，并保持高精度和鲁棒性；

[0017] (3) 本发明将考勤系统扩展到记录工作过程中的消极行为，而不仅仅是记录上下班的迟到早退，使考勤系统更为完善。

附图说明

[0018] 图1是本发明的yolo预训练分类模型示意图；

[0019] 图2是本发明的基于yolo的行人检测模型示意图；

[0020] 图3是本发明的基于ResNet50的多任务卷积神经网络示意图；

[0021] 图4是本发明方法的流程图；

具体实施方式

[0022] 为了便于本领域普通技术人员理解和实施本发明，下面结合附图和实施例对本发明作进一步的详细描述。

[0023] 实施例1：

[0024] 一种基于yolo和多任务卷积神经网络的导购消极行为监控系统，含有以下步骤：

[0025] (1) 训练基于yolo的行人检测模型；

[0026] 步骤11：构建基于yolo的行人检测模型；

[0027] 本发明借鉴yolo第二代的训练方式和网络结构，在此基础上，对其网络结构进行了改进，使模型在本发明的监控视角中更为鲁棒。具体地，原yolo-v2的网络结构中包含19个卷积层和5个最大池化层，本发明中使用跳层融合的方式，特征提取的阶段一使用13个卷积层和4个最大池化层，阶段二使用7个卷积层，在阶段一和阶段二中间有1个最大池化层，调整阶段一的输出的特征图大小与阶段二输出的特征图大小一致。然后将这两个调整过大小的特征图通过叠加的方式融合在一起，成为阶段三的输出。阶段三有两种模式，模式一是分类网络，该模式在预训练模型的时候用到，具体而言就是一层 3×3 的卷积层，再加一层全连接层，全连接层神经元个数等于分类个数；模式二是检测网络，该模式是在加载模式一的预训练参数之后，训练检测网络时用的，具体而言就是一层 3×3 的卷积层，再加一层 1×1 的卷积层，卷积核的个数与检测类别相关，具体数值为： $\text{anchors个数} \times (5 + \text{检测类别个数})$ 。

[0028] 针对模式一的分类网络，如图1所示，以下进行详细描述：

[0029] 阶段一：输入图像大小为 $448 \times 448 \times 3$ ，阶段一的第一层是一个卷积核大小为 $3 \times 3 \times 32$ 的卷积层，对该层进行批次归一化、ReLU非线性激活、 2×2 的最大池化操作；阶段一的第二层是一个卷积核大小为 $3 \times 3 \times 64$ 的卷积层，对该层进行批次归一化、ReLU非线性激活、

2×2 的最大池化操作;阶段一的第三层是一个卷积核大小为 $3 \times 3 \times 128$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第四层是一个卷积核大小为 $1 \times 1 \times 64$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第五层是一个卷积核大小为 $3 \times 3 \times 128$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第六层是一个卷积核大小为 $3 \times 3 \times 256$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第七层是一个卷积核大小为 $1 \times 1 \times 128$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第八层是一个卷积核大小为 $3 \times 3 \times 256$ 的卷积层,对该层进行批次归一化、ReLU非线性激活、 2×2 的最大池化操作操作;阶段一的第九层是一个卷积核大小为 $3 \times 3 \times 512$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第十层是一个卷积核大小为 $1 \times 1 \times 256$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第十一层是一个卷积核大小为 $3 \times 3 \times 512$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第十二层是一个卷积核大小为 $1 \times 1 \times 256$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的第十三层是一个卷积核大小为 $3 \times 3 \times 512$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段一的输出特征图记为output1。

[0030] 阶段二:先对阶段一输出的特征图进行 2×2 的最大池化操作,阶段二的第一层是一个卷积核大小为 $3 \times 3 \times 1024$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的第二层是一个卷积核大小为 $1 \times 1 \times 512$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的第三层是一个卷积核大小为 $3 \times 3 \times 1024$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的第四层是一个卷积核大小为 $1 \times 1 \times 512$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的第五层是一个卷积核大小为 $3 \times 3 \times 1024$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的第六层是一个卷积核大小为 $3 \times 3 \times 1024$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的第七层是一个卷积核大小为 $3 \times 3 \times 1024$ 的卷积层,对该层进行批次归一化、ReLU非线性激活操作;阶段二的输出特征图记为output2;

[0031] 阶段三:对阶段一的输出特征图output1进行 $1 \times 1 \times 64$ 的卷积,而后把它的尺寸调整为跟阶段二输出特征图output2一样的尺寸,调整后的特征图记为output1_1。将特征图output1_1与output2叠加,形成新的特征图output3。之后,对于融合的特征图output3进行 $3 \times 3 \times 1024$ 的卷积、批次归一化、ReLU非线性激活操作,最后,添加一层1000个神经元的全连接层,用softmax损失函数进行约束。

[0032] 这种跳层操作的好处在于output3既有深层卷积后得到的output2的精细特征,又有浅层卷积后得到的output1_1的基础特征,使得网络精度更高。

[0033] 针对模式二的检测网络,如图2所示,除最后两层之外,其余结构都与上述的分类模型相同;区别在于:检测网络中对阶段三中的融合特征output3进行 $3 \times 3 \times 1024$ 的卷积、批次归一化、ReLU非线性激活操作,而后去掉全连接层,替换为 $1 \times 1 \times 30$ 的卷积层最后用坐标损失、置信度损失、类别损失去约束模型。

[0034] 步骤12:用ImageNet数据集预训练分类模型;

[0035] 良好的初始化参数是模型收敛的重要一环,而检测数据集由于标注步骤繁琐,导致其每种类别的数据量不多。因此,我们用ImageNet数据集去训练一个分类模型,使训练好的分类模型参数作为检测模型中共有结构的初始化参数。

[0036] 首先对标准的1000分类的ImageNet数据集的图片进行随机裁剪、旋转和色调、饱和度和曝光度的偏移,以获得更多可用的训练数据,把这些数据调整为大小 224×224 的图像,训练160个epoch,使用SGD优化器,初始学习率设置为0.1,momentum设置为0.9,weight decay设置为0.0005。

[0037] 进一步的,用更大的尺寸(448×448)对网络进行微调,用0.001的学习率训练10个epoch即可。

[0038] 步骤13:用voc2007数据集预训练检测模型;

[0039] 由于检测模型的前几层结构与分类网络一致,我们将步骤12中训练好的分类网络的参数作为检测网络中共有结构的初始化参数。Voc2007数据集是常用的检测数据集,共有20类的标注好的检测物体,其中包括行人图像数据。我们只取其中的行人图像数据进行训练,对行人数据进行数据增强的操作,调整图像大小为 448×448 ,用SGD优化器训练160个epoch,初始学习率设置为0.0001;

[0040] 步骤14:用监控视角数据集微调模型;

[0041] 由于voc2007中的行人数据大部分都不是监控视角下的行人图像,因此将步骤13训练好的模型去检测商场监控画面中的行人,会有一些困难。因此我们选取BOT2018新零售技术挑战赛中的数据集进行微调,该数据集的行人图像采集自真实商场场景下的监控摄像头。对该数据集的图像进行水平旋转、中心随机裁剪、HSV空间微调等数据增强的操作,并将尺寸大小调整为 448×448 ;

[0042] 加载步骤13训练好的模型,用SGD优化器训练160个epoch,初始学习率设置为0.001,学习率随着训练次数的增多而减小,0-5个epoch时学习率设置为0.001,5-80个epoch时学习率设置为0.0001,80-160个epoch时学习率设置为0.00001。

[0043] (2) 训练基于ResNet50的多任务卷积神经网络

[0044] 步骤21:构建基于ResNet50的多任务卷积神经网络;

[0045] 对(1)中检测到的行人,我们需要去识别行人的属性,以此判断导购是否存在消极工作的行为,数据集中标注的属性有:“顾客”或“导购”、“男性”或“女性”、“站立”或“坐着”、“玩手机”或“不玩手机”。这些属性彼此之间无关联性,因此可视为不相关属性。

[0046] ResNet50是一个在分类性能表现优异的网络结构,然而一个原始的ResNet50直接去识别多个不相关属性的时候效果往往不佳,而针对每个属性都训练一个模型会导致占用额外的计算资源。因此,本发明针对导购消极行为的识别,基于ResNet50设计了一种多任务卷积神经网络,结构如图3所示。

[0047] 具体地,将原始ResNet50最后两层(全连接层和池化层)去掉,拼接上四个并行的全连接层,每个全连接层的神经元个数为2,分别代表着8种属性:“顾客”与“导购”、“男性”与“女性”、“站立”与“坐着”、“玩手机”与“不玩手机”,在同一个全连接层上的两个属性是关联属性,不在一个全连接层上的属性是不相关属性。每个全连接后分别连一个Softmax层。Softmax损失函数的计算公式为:

$$[0048] \quad L_i = -\log\left(\frac{e^{x_k}}{\sum_{j=1}^2 e^{x_j}}\right) \quad (1)$$

[0049] 最终的损失函数值为四个Softmax损失函数值相加,即:

[0050] $Loss=L_1+L_2+L_3+L_4$ (2)

[0051] 步骤22:训练基于ResNet50的多任务卷积神经网络;

[0052] 在卷积神经网络中,良好的初始参数对网络模型的收敛有重要作用,因此我们加载在ImageNet数据集上训练好的ResNet50中除最后两层外的参数作为本发明中多任务卷积神经网络的初始化参数。数据集采用BOT2018新零售技术挑战赛的标注数据,对该数据进行数据增强的操作,如图像水平翻转、中心随机裁剪、HSV空间增强等,以获得更多的可用的训练数据,最后用Adam优化算法进行训练,初始学习率设置为0.0005,训练40个epoch。

[0053] (3) 导购消极行为记录

[0054] 步骤31:读取监控画面;

[0055] 在商场内,广泛分布的监控系统为本发明提供数据,无需增加额外的设备。在读取监控画面前,我们需要设置两个参数:上班时间区间、监控采样时间。设置上班时间区间,可让本系统只关注于上班时间,减少额外的计算资源和误检,本发明的目的是为了检测导购在上班时间是否存在消极行为,因此下班时间导购的行为不在我们的考虑范围内;设置监控采样时间,是控制读取监控画面的频率,可减少额外的计算资源,没有必要每时每刻都去检测,采样时间设置的越小,识别次数越多,管理越严格,但计算负担比较大,采样时间设置的越大,识别次数越少,计算负担比较小,但管理比较松。本发明默认的监控采样时间是早上9点到晚上9点,采样时间是30秒采样1次。

[0056] 具体地,我们通过rtsp协议读取监控画面,通过计算机的系统时间与采样时间控制读取画面的频率。

[0057] 步骤32:商场内行人检测;

[0058] 加载步骤(1)中训练好的行人检测模型,读取步骤31中的监控图像,将图像进行归一化,并转为Tensor,之后将Tensor加载进行行人检测模型中,行人检测模型会检测出行人在图像中上下左右四个坐标,可在一张图像上检测出多个人;

[0059] 步骤33:识别行人属性;

[0060] 步骤32中检测出的行人,有可能是导购,也可能是顾客,我们要识别导购是否存在消极行为,本发明设计的多任务卷积神经网络可实现该功能。加载步骤(2)中的多任务卷积神经网络,将步骤32中检测出的行人图像数据作为多任务卷积神经网络的输入数据,输出全连接层的值即为模型对该行人在某个属性上的置信度。若输出“导购”的置信度比“顾客”的置信度高,即识别该行人为“导购”,若输出“男性”的置信度比“女性”的置信度高,即识别该行人为“男性”,若输出“站立”的置信度比“坐”的置信度高,即识别该行人为“站立”,若输出“玩手机”的置信度比“不玩手机”的置信度高,即识别该行人为“玩手机”。反之亦然。

[0061] 步骤34:记录导购消极行为画面;

[0062] 系统框图如图4所示。

[0063] 具体地,在上班时间区间内,我们对步骤33中识别的行人属性进行导购消极行为的判断。首先,判断该行人的身份,是否属于“导购”,如果是导购,我们分析其姿势(站着或坐着)和工作状态(是否玩手机),更进一步的,我们判断该导购所在的画面中,是否存在顾客,顾客在场的情况下,我们对导购有更严格的要求。比如,当导购在玩手机或者坐着,并且此时画面中没有顾客,我们认为该导购是“一般消极”;当导购在玩手机或者坐着,而此时画面中有顾客,我们认为该导购是“严重消极”。具体导购消极行为的程度判定如表1所示,对

于“严重消极”的导购,我们将其画面保存在文件夹1,对于“一般消极”的导购,我们将其画面保存在文件夹2,对于“积极”的导购,我们不保存其画面。店主可根据文件1和文件夹2的图像画面,对消极的导购做出相应的处罚。

[0064] 表1导购消极行为判定表

身份	姿势	是否玩手机	附近是否有顾客	判断
导购	站立	是	是	严重消极
导购	坐	是	是	严重消极
导购	坐	否	是	严重消极
[0065] 导购	站立	否	是	积极
导购	站立	是	否	一般消极
导购	坐	是	否	一般消极
导购	坐	否	否	一般消极
导购	站立	否	否	积极
顾客	/	/	/	/

[0066] 实施例2:

[0067] (1) 选取实验数据

[0068] 本发明使用BOT2018新零售技术挑战赛的数据集,数据采集自真实的商场场景,图像数据中的标签有:“顾客”与“导购”、“男性”与“女性”、“站立”与“坐”、“玩手机”与“不玩手机”。分为5个场景,共5000张图像,每张图像中包含数量不等的导购和顾客,本发明将这5000张图像按9:1的比例分为训练集和测试集,平均抽取。训练集中场景1有1980张,场景2有937张,场景3中有915张,场景4中有356张,场景5中有312张,共4500张;测试集中场景1有220张,场景2有105张,场景3中有101张,场景4中有40张,场景5中有34张。

[0069] (2) 实验结果

[0070] 按照实施例1中的步骤(2)训练基于ResNet50的多任务卷积神经网络,构造完模型之后,加载在ImageNet上训练过的ResNet50的参数,用Adam优化器在BOT商场数据集上训练40个epoch,初始学习率为0.0005,最终在测试集上的精度如表2所示:

[0071] 表2实验结果

	属性	精度
	性别（男/女）	93.7%
[0072]	身份（导购/顾客）	96.7%
	姿势（站立/坐）	99.4%
	玩手机（是/否）	94.6%

[0073] 本说明书实施例所述的内容仅仅是对发明构思的实现形式的列举,本发明的保护范围不应当被视为仅限于实施例所陈述的具体形式,本发明的保护范围也及于本领域技术人员根据本发明构思所能够想到的等同技术手段。

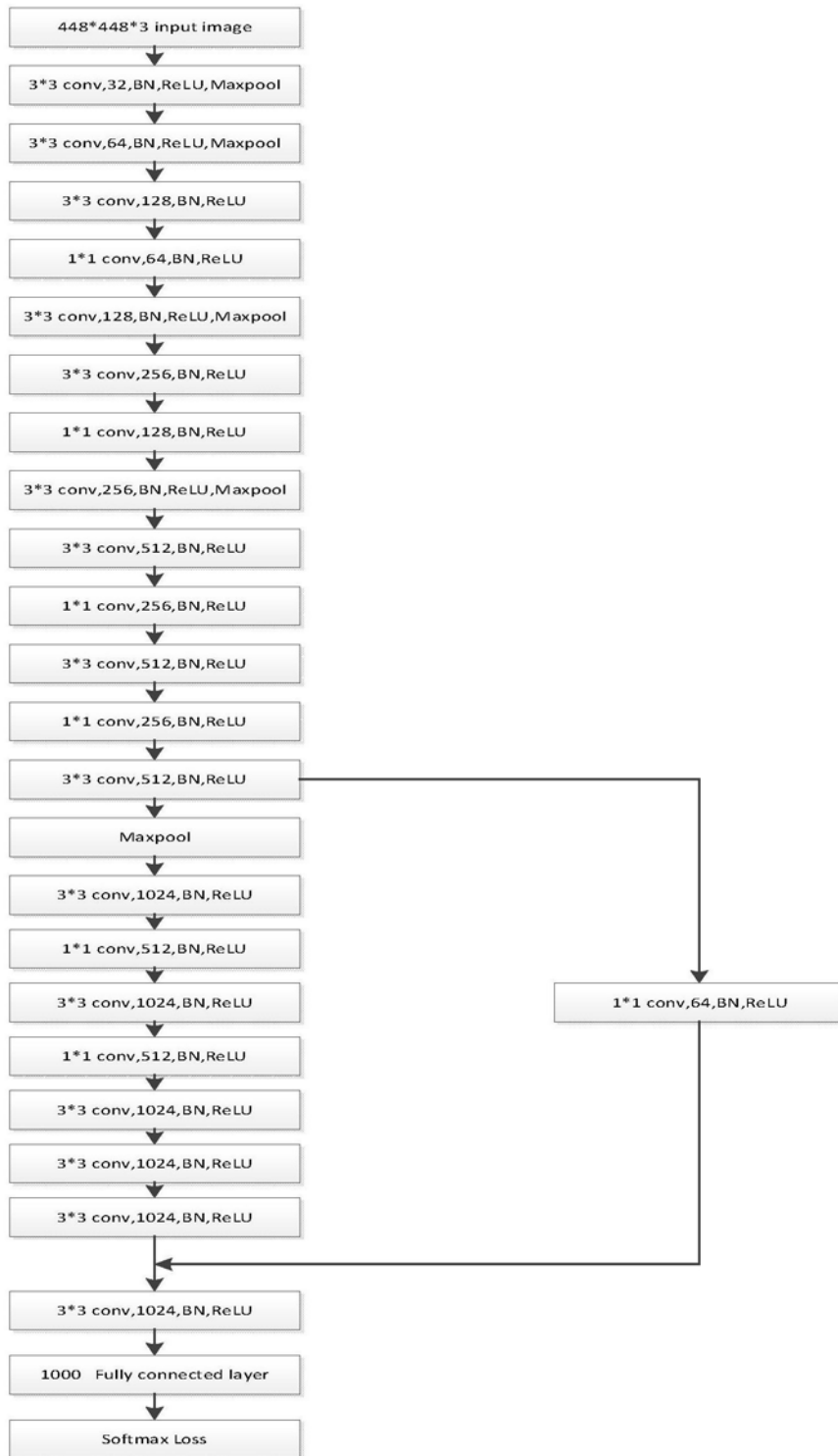


图1

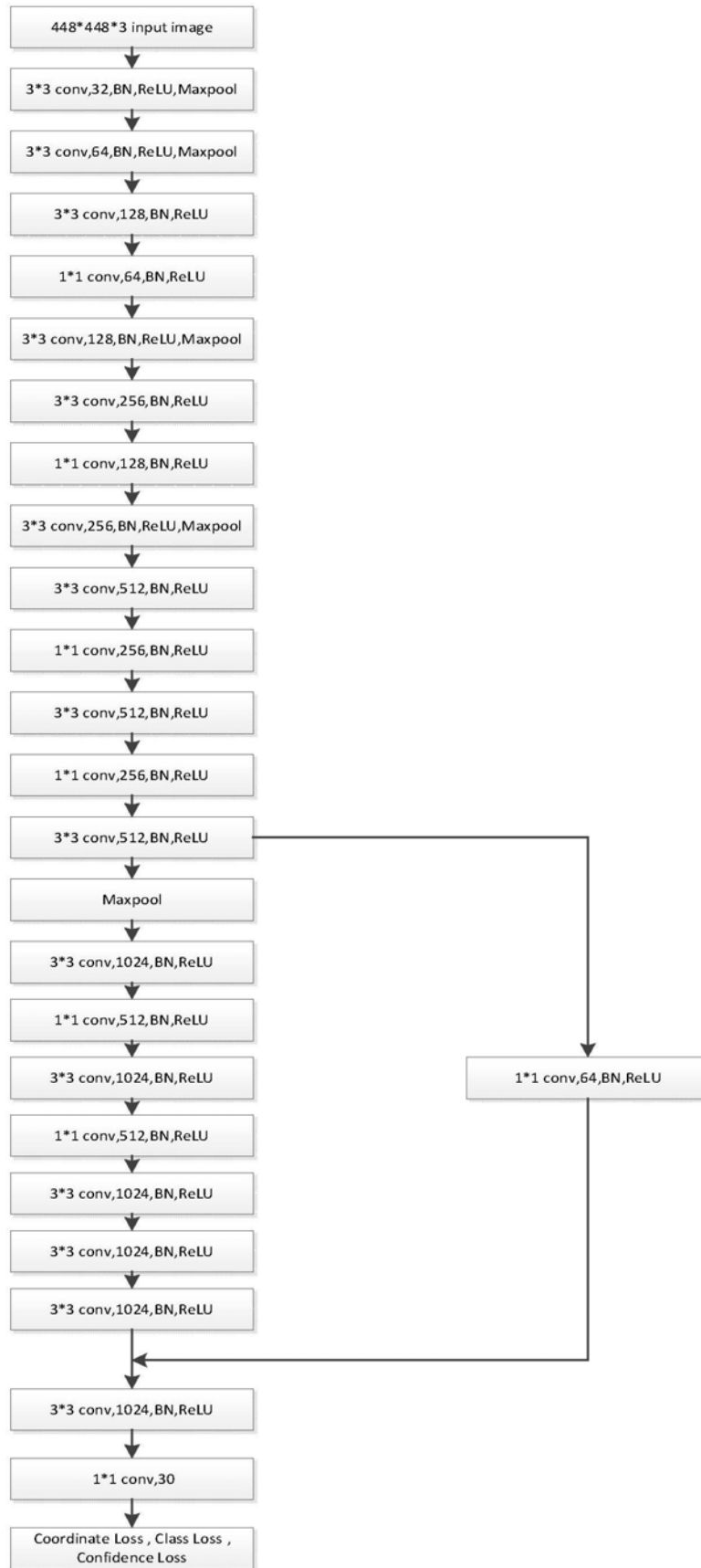


图2

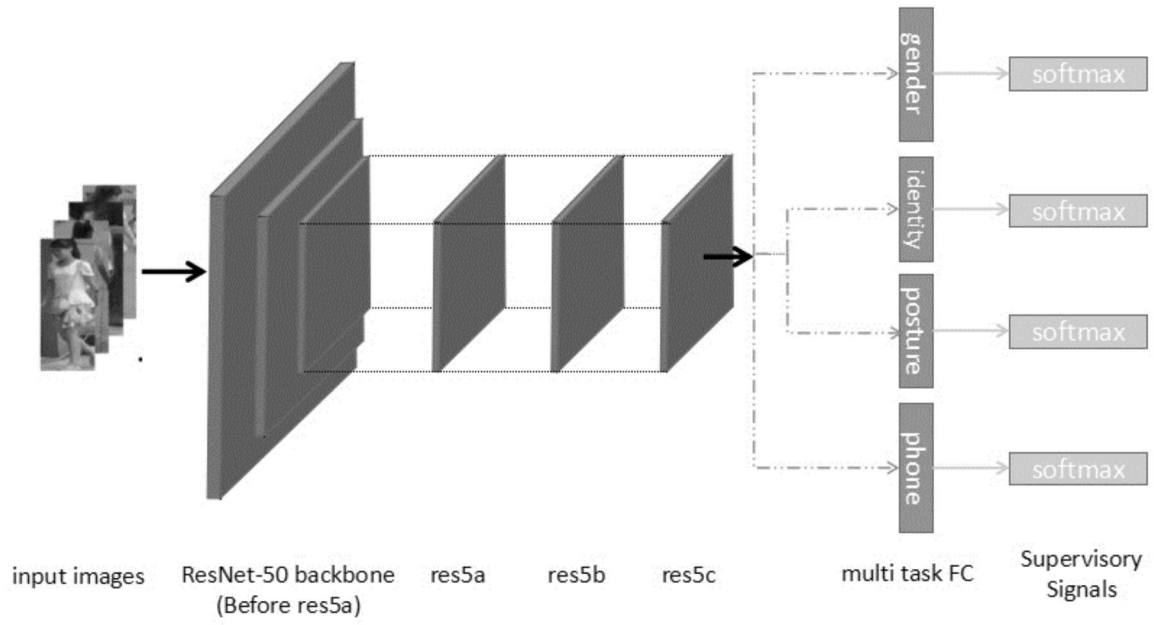


图3

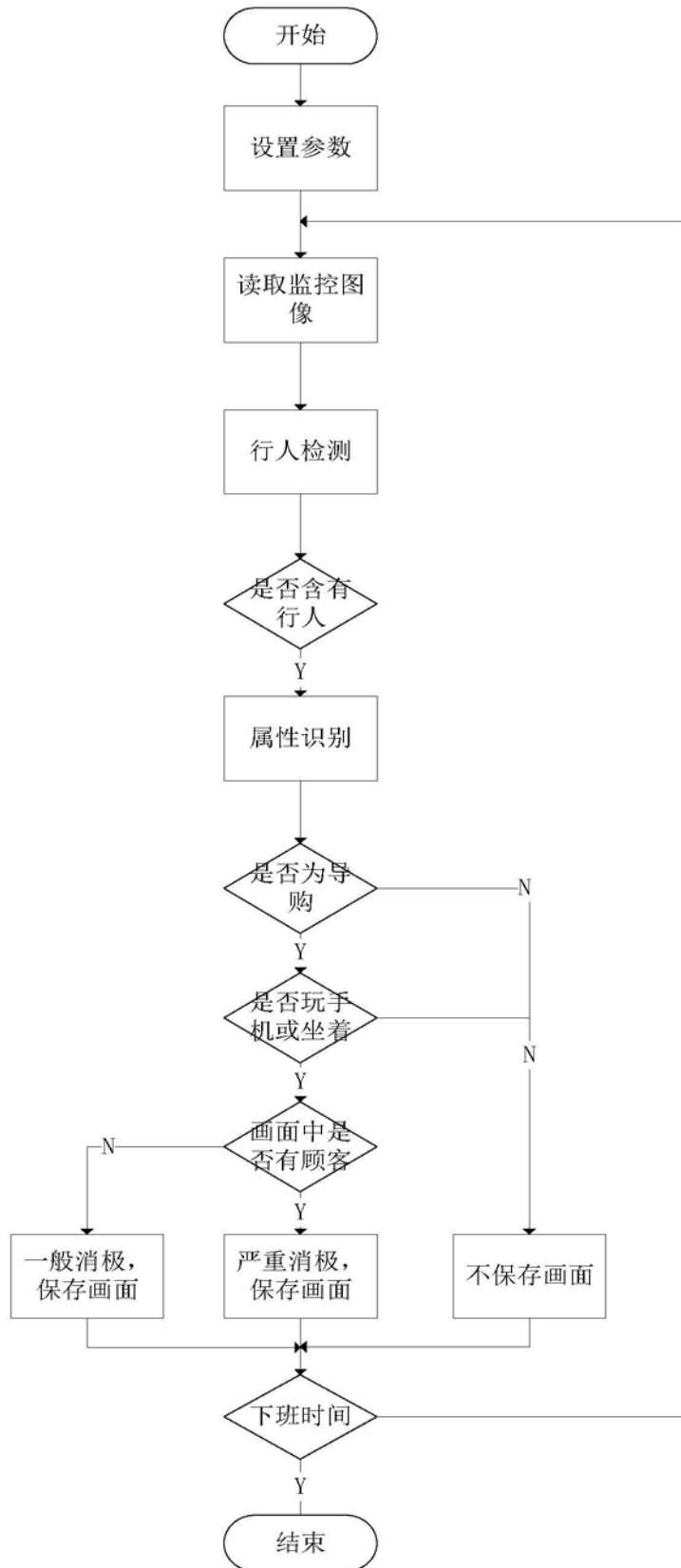


图4