



(12) 发明专利申请

(10) 申请公布号 CN 112699769 A

(43) 申请公布日 2021.04.23

(21) 申请号 202011564655.3

(22) 申请日 2020.12.25

(71) 申请人 北京竞业达数码科技股份有限公司

地址 102308 北京市门头沟区石龙工业区

雅安路6号院1号楼C座8层805

(72) 发明人 巩凯强 朱全福

(74) 专利代理机构 北京市盛峰律师事务所

11337

代理人 于国强

(51) Int. Cl.

G06K 9/00 (2006.01)

G06T 7/66 (2017.01)

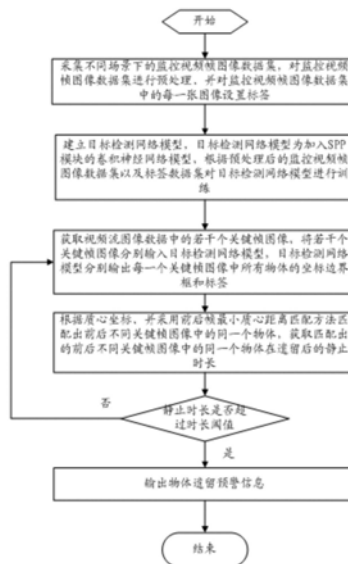
权利要求书2页 说明书5页 附图1页

(54) 发明名称

一种安防监控中遗留物品的检测方法及系统

(57) 摘要

本发明涉及图像处理领域,公开了一种安防监控中遗留物品的检测方法及系统,包括对监控视频帧图像数据集进行预处理;建立加入SPP模块的卷积神经网络模型;将若干个关键帧图像输入目标检测网络模型,目标检测网络模型输出关键帧图像中所有物体的坐标边界框和标签;根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,判断同一个物体在遗留后的静止时长是否超过时长阈值,输出物体遗留预警信息。本发明在物体遗留较长时间后更具有鲁棒性和稳定性;本发明不仅提升了小物品的召回率,还通过基于质心最小距离快速匹配同一个物体,匹配准确率高,保证了物品遗留后及时发出预警信息。



1. 一种安防监控中遗留物品的检测方法,其特征在于,包括以下步骤:

S1) 采集不同场景下的监控视频帧图像数据集,对所述监控视频帧图像数据集进行预处理、并对所述监控视频帧图像数据集中的每一张图像设置标签类别,获得预处理后的监控视频帧图像数据集以及与所述预处理后的监控视频帧图像数据集相对应的标签数据集;

S2) 建立目标检测网络模型,所述目标检测网络模型为加入SPP模块的卷积神经网络模型,根据所述预处理后的监控视频帧图像数据集以及所述标签数据集对所述目标检测网络模型进行训练,获得训练完成的目标检测网络模型;

S3) 采集待检测遗留物品的视频流图像数据,按时间顺序获取所述视频流图像数据中的若干个关键帧图像,将所述若干个关键帧图像按时间先后顺序分别输入所述训练完成的目标检测网络模型,所述训练完成的目标检测网络模型分别输出每一个关键帧图像中所有物体的坐标边界框和标签;

S4) 根据所述每一个关键帧图像中所有物体的坐标边界框计算每一个关键帧图像所有物体的质心坐标,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,并获取前后不同关键帧图像中的同一个物体在遗留后的静止时长,设置时长阈值,判断所述匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长是否超过时长阈值,若否,则返回步骤S3);若是,则输出物体遗留预警信息。

2. 根据权利要求1所述的安防监控中遗留物品的检测方法,其特征在于,在步骤S1)中,对所述监控视频帧图像数据集进行预处理、并对所述监控视频帧图像数据集中的每一张图像设置标签类别,包括对所述监控图像数据集中的每一张图像数据进行增强预处理,所述增强预处理包括图像裁剪、图像翻转和Mosaic数据增强方式;所述标签包括行人、纸箱、书包和行李箱。

3. 根据权利要求1或2所述的安防监控中遗留物品的检测方法,其特征在于,在步骤S2)中,所述加入SPP模块的卷积神经网络模型包括输入层、中间层和输出层;所述中间层包括若干层卷积层、若干层批量归一化层、若干层非线性激活函数层以及SPP模块;所述输出层输出每一张图像中的物体坐标边界框、置信度以及标签。

4. 根据权利要求3所述的安防监控中遗留物品的检测方法,其特征在于,在步骤S3)中,按时间顺序获取所述视频流图像数据中的若干个关键帧图像,包括设置预设帧数,按时间顺序在所述视频流图像数据中每隔预设帧数采集关键帧图像。

5. 根据权利要求4所述的安防监控中遗留物品的检测方法,其特征在于,在步骤S4)中,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后关键帧图像中的同一个物体,包括在不同关键帧图像间以物体的质心坐标之间的最小距离为特征进行匹配关联,通过所述匹配关联获取位于不同关键帧图像中的同一个物体。

6. 根据权利要求5所述的安防监控中遗留物品的检测方法,其特征在于,在步骤S4)中,获取前后不同关键帧图像中的同一个物体在遗留后的静止时长,还包括设置质心坐标变化阈值,计算 $\omega_j = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$ 获得前后不同关键帧图像中的第j个同一个物体在前后不同关键帧图像中的质心坐标差值 ω_j ,所述第j个同一个物体在第i+1个关键帧图像中的质心坐标为 (x_{i+1}, y_{i+1}) ,所述第j个同一个物体在第i个关键帧图像中的质心坐标为 (x_i, y_i) ,当第j个同一个物体在前后不同关键帧图像中的质心坐标差值 ω_j 小于所述质心坐

标变化阈值时,将所述第j个同一个物体判定为遗留物体、并获取所述第j个同一个物体在遗留后的静止时长。

7.一种安防监控中遗留物品的检测系统,适用于如权利要求1至6任一项所述的安防监控中遗留物品的检测方法,其特征在于,包括监控视频帧图像采集模块、目标检测网络模型建立模块和遗留物品检测模块;

所述监控视频帧图像采集模块,用于采集不同场景下的监控视频帧图像数据集以及采集待检测遗留物品的视频流图像数据,对所述监控视频帧图像数据集进行预处理、并对所述监控视频帧图像数据集中的每一张图像设置标签类别,获得预处理后的监控视频帧图像数据集以及与所述预处理后的监控视频帧图像数据集相对应的标签数据集;按时间顺序获取所述视频流图像数据中的若干个关键帧图像;

所述目标检测网络模型建立模块,用于建立目标检测网络模型,所述目标检测网络模型为加入SPP模块的卷积神经网络模型,根据所述监控视频帧图像采集模块中预处理后的监控视频帧图像数据集以及标签数据集对所述目标检测网络模型进行训练,获得训练完成的目标检测网络模型;

所述遗留物品检测模块,用于将所述监控视频帧图像采集模块采集的若干个关键帧图像按时间先后顺序分别输入所述训练完成的目标检测网络模型,所述训练完成的目标检测网络模型分别输出每一个关键帧图像中所有物体的坐标边界框和标签;根据所述每一个关键帧图像中所有物体的坐标边界框计算每一个关键帧图像所有物体的质心坐标,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,获取匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长,并根据所述静止时长输出物体遗留预警信息。

一种安防监控中遗留物品的检测方法及系统

技术领域

[0001] 本发明涉及图像处理领域,具体地涉及一种安防监控中遗留物品的检测方法及系统。

背景技术

[0002] 随着网络监控摄像头的普及和广泛使用,遗留物检测在安全防范领域中应用广泛,例如,在银行、地铁、火车站、商场等场所,对出现的遗留物品进行监控并报警,在很大程度上提高了视频监控系统的实用性和有效性。尽管现在关于遗留物检测有很多研究,但是依然存在一些未解决的问题,还没有适用于各种监控场景下的遗留物检测算法,同时复杂环境下遗留物检测的准确度和实时性存在改善空间,帮助建立有效的地铁监控机制。目前遗留物检测主要有以下几种方法:a.基于传统图像检测物体方法;b.基于深度学习的物体检测方法。

[0003] 基于传统图像检测物体方法,大多采取的是基于混合高斯模型,比如国家专利公开文献CN111914670A,公开了“一种遗留物品检测方法、装置、系统及存储介质”,该发明包括:进行背景模型建立;获取检测图像中的运动区域,在运动区域提取第一特征点;获取检测图像的前一帧图像,并在前一帧图像中与运动区域重合的区域中提取第二特征点。该发明通过对视频图像背景建模,提取出图像中前景物体,对静止的前景物体进行计时,若计时器达到阈值时进行标记为遗留物品并报警。然而使用背景建模方法检测存在以下两个问题:一是运动较慢的前景物体会被错误的处理为背景;二是静止的前景物体逐渐会被背景吸收,对于静止较长时间的物体可能造成漏检。

[0004] 基于深度学习的检测方法,比如国家专利公开文献CN107527009B,公开了“一种基于YOLO目标检测的遗留物检测方法”,该发明通过YOLO算法对目标实时检测,得到每帧图像数据中的目标类别和其对应的具体坐标,再经过目标类别和两个坐标的重叠度对检测到的目标分类,对可疑目标跟踪计时后进行背景搬移物判断得到准确的遗留物。该发明虽然能达到实时的效果,但对于小目标物体(如包)的检测效果较差。

[0005] 现有遗留物品的检测方法存在的主要问题有:传统图像检测物体技术中前景物体的召回率低以及较长时间遗留的物体漏检。现有基于深度神经网络的检测技术对于场景中中小物体的召回率低,对目标较小的包、行李箱等检测效果较差。

发明内容

[0006] 本发明提供一种安防监控中遗留物品的检测方法及系统,从而解决现有技术的上述问题。

[0007] 第一方面,本发明提供了一种安防监控中遗留物品的检测方法及系统,包括以下步骤:

[0008] S1) 采集不同场景下的监控视频帧图像数据集,对监控视频帧图像数据集进行预处理、并对监控视频帧图像数据集中的每一张图像设置标签,获得预处理后的监控视频帧

图像数据集以及与预处理后的监控视频帧图像数据集相对应的标签数据集；

[0009] S2) 建立目标检测网络模型,目标检测网络模型为加入空间金字塔池化结构(Spatial Pyramid Pooling,SPP)模块的卷积神经网络模型,根据预处理后的监控视频帧图像数据集以及标签数据集对目标检测网络模型进行训练,获得训练完成的目标检测网络模型;

[0010] S3) 采集待检测遗留物品的视频流图像数据,按时间顺序获取视频流图像数据中的若干个关键帧图像,将若干个关键帧图像按时间先后顺序分别输入训练完成的目标检测网络模型,训练完成的目标检测网络模型分别输出每一个关键帧图像中所有物体的坐标边界框和标签;

[0011] S4) 根据每一个关键帧图像中所有物体的坐标边界框计算每一个关键帧图像所有物体的质心坐标,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,获取匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长,设置时长阈值,判断匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长是否超过时长阈值,若否,则返回步骤S3);若是,则输出物体遗留预警信息。

[0012] 进一步的,在步骤S1)中,对监控视频帧图像数据集进行预处理、并对监控视频帧图像数据集中的每一张图像设置标签,包括对监控图像数据集中的每一张图像数据进行增强预处理,增强预处理包括图像裁剪、图像翻转和Mosaic数据增强方式;标签包括行人、书包、纸箱和行李箱。

[0013] 进一步的,在步骤S2)中,加入SPP模块的卷积神经网络模型包括输入层、中间层和输出层;中间层包括若干层卷积层、若干层批量归一化层、若干层非线性激活函数层以及SPP模块;所述输出层输出每一张图像中的物体坐标边界框、置信度以及标签。

[0014] 进一步的,在步骤S3)中,按时间顺序获取视频流图像数据中的若干个关键帧图像,包括设置预设帧数,按时间顺序在视频流图像数据中每隔预设帧数采集关键帧图像。

[0015] 进一步的,在步骤S4)中,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后关键帧图像中的同一个物体,包括在不同关键帧图像间以物体的质心坐标之间的最小距离为特征进行匹配关联,通过匹配关联获取位于不同关键帧图像中的同一个物体。

[0016] 进一步的,在步骤S4)中,获取前后不同关键帧图像中的同一个物体在遗留后的静止时长,还包括设置质心坐标变化阈值,计算 $\omega_j = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$ 获得前后不同关键帧图像中的第j个同一个物体在前后不同关键帧图像中的质心坐标差值 ω_j , $j \leq m$,m为前后不同关键帧图像中匹配为同一个物体的总数,第j个同一个物体在第i+1个关键帧图像中的质心坐标为 (x_{i+1}, y_{i+1}) ,第j个同一个物体在第i个关键帧图像中的质心坐标为 (x_i, y_i) ,当第j个同一个物体在前后不同关键帧图像中的质心坐标差值 ω_j 小于质心坐标变化阈值时,将第j个同一个物体判定为遗留物体、并获取第j个同一个物体在遗留后的静止时长。

[0017] 第二方面,本发明提供了一种安防监控中遗留物品的检测系统,,包括监控视频帧图像采集模块、目标检测网络模型建立模块和遗留物品检测模块;

[0018] 监控视频帧图像采集模块,用于采集不同场景下的监控视频帧图像数据集以及采集待检测遗留物品的视频流图像数据,对监控视频帧图像数据集进行预处理、并对监控视频帧图像数据集中的每一张图像设置标签类别,获得预处理后的监控视频帧图像数据集以

及与预处理后的监控视频帧图像数据集相对应的标签数据集;按时间顺序获取视频流图像数据中的若干个关键帧图像;

[0019] 目标检测网络模型建立模块,用于建立目标检测网络模型,目标检测网络模型为加入SPP模块的卷积神经网络模型,根据监控视频帧图像采集模块中预处理后的监控视频帧图像数据集以及标签数据集对目标检测网络模型进行训练,获得训练完成的目标检测网络模型;

[0020] 遗留物品检测模块,用于将监控视频帧图像采集模块采集的若干个关键帧图像按时间先后顺序分别输入训练完成的目标检测网络模型,训练完成的目标检测网络模型分别输出每一个关键帧图像中所有物体的坐标边界框和标签;根据每一个关键帧图像中所有物体的坐标边界框计算每一个关键帧图像所有物体的质心坐标,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,获取匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长,并根据静止时长输出物体遗留预警信息。

[0021] 本发明的有益效果是:

[0022] 与现有图像检测物体方法相比,本发明提出了一种基于深度学习的物品遗留方法,利用计算机视觉技术对物品与行人检测,提供了一种长时间稳定的检测器,比传统的背景建模方法相比,本发明在物体遗留较长时间后被背景吸收的情况下更具有鲁棒性和稳定性。

[0023] 与基于Darknet53的检测方法相比,本发明通过对图像进行数据增强、在预测特征图前面融入SPP模块,提升了小物品的召回率。

[0024] 对于物品遗留事件,本发明通过基于质心最小距离匹配不同帧的每个物体,该方法在检测器输出的边界框信息快速匹配同一个物体,匹配准确率高,保证了物品遗留后系统能及时发出预警信息。

附图说明

[0025] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0026] 图1为本实施例一提供的安防监控中遗留物品的检测方法流程示意图。

[0027] 图2为本实施例一提供的在不同关键帧图像间物体匹配关联示意图。

具体实施方式

[0028] 为了使本发明的目的、技术方案及优点更加清楚明白,以下结合附图,对本发明进行进一步详细说明。应当理解,此处所描述的具体实施方式仅仅用以解释本发明,并不用于限定本发明。

[0029] 实施例一,一种安防监控中遗留物品的检测方法及系统,如图1所示,包括以下步骤:

[0030] S1) 采集不同场景下的监控视频帧图像数据集,对监控视频帧图像数据集进行预处理、并对监控视频帧图像数据集中的每一张图像设置标签,获得预处理后的监控视频帧

图像数据集以及预处理后的监控视频帧图像数据集相对应的标签数据集。其中,对监控视频帧图像数据集进行预处理、并对监控视频帧图像数据集中的每一张图像设置标签,包括对监控图像数据集中的每一张图像数据进行增强预处理,增强预处理包括图像裁剪、图像翻转和Mosaic数据增强方式;标签包括行人、书包、纸箱和行李箱。

[0031] S2) 建立目标检测网络模型,目标检测网络模型为加入SPP模块的卷积神经网络模型,根据预处理后的监控视频帧图像数据集以及标签数据集对目标检测网络模型进行训练,获得训练完成的目标检测网络模型;加入SPP模块的卷积神经网络模型包括输入层、中间层和输出层;中间层包括若干层卷积层、若干层批量归一化层、若干层非线性激活函数层以及SPP模块;所述输出层输出每一张图像中的物体坐标边界框、置信度以及标签。

[0032] S3) 采集待检测遗留物品的视频流图像数据,按时间顺序获取视频流图像数据中的若干个关键帧图像,将若干个关键帧图像按时间先后顺序分别输入训练完成的目标检测网络模型,训练完成的目标检测网络模型分别输出每一个关键帧图像中所有物体的坐标边界框和标签。

[0033] 在步骤S3)中,按时间顺序获取视频流图像数据中的若干个关键帧图像,包括设置预设帧数为4,按时间顺序在视频流图像数据中每隔4帧采集关键帧图像。

[0034] S4) 根据每一个关键帧图像中所有物体的坐标边界框计算每一个关键帧图像所有物体的质心坐标,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,获取匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长,设置时长阈值,判断匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长是否超过时长阈值,若否,则返回步骤S3);若是,则输出物体遗留预警信息。

[0035] 在步骤S4)中,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后关键帧图像中的同一个物体,包括在不同关键帧图像间以物体的质心坐标之间的最小距离为特征进行匹配关联,通过匹配关联获取位于不同关键帧图像中的同一个物体。如图2所示,上一帧关键帧图像 L_{i-1} 检测出的两个物品分别为物品 A_{i-1} 和物品 B_{i-1} ,当前帧关键帧图像 L_i 检测出的两个物品分别为物品 A_i 和物品 B_i ,通过质心坐标计算得到物品 B_{i-1} 与物品 B_i 之间的质心距离为0.1,物品 B_{i-1} 与物品 A_i 之间的质心距离为0.7,物品 A_{i-1} 与物品 A_i 之间的质心距离为0.8,物品 A_{i-1} 与物品 B_i 之间的质心距离为0.3。本实施例中,物品 B_{i-1} 与物品 B_i 之间的质心距离最小,因此,将物品 B_{i-1} 与物品 B_i 认为是前后不同关键帧图像中的同一个物体。

[0036] 在步骤S4)中,获取前后不同关键帧图像中的同一个物体在遗留后的静止时长,还包括设置质心坐标变化阈值,计算 $\omega_j = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$ 获得前后不同关键帧图像中的第j个同一个物体在前后不同关键帧图像中的质心坐标差值 ω_j ,所述第j个同一个物体在第i+1个关键帧图像中的质心坐标为 (x_{i+1}, y_{i+1}) ,所述第j个同一个物体在第i个关键帧图像中的质心坐标为 (x_i, y_i) ,当第j个同一个物体在前后不同关键帧图像中的质心坐标差值 ω_j 小于质心坐标变化阈值时,将第j个同一个物体判定为遗留物体、并获取第j个同一个物体在遗留后的静止时长。

[0037] 第二方面,本发明提供了一种安防监控中遗留物品的检测系统,包括监控视频帧图像采集模块、目标检测网络模型建立模块和遗留物品检测模块;

[0038] 监控视频帧图像采集模块,用于采集不同场景下的监控视频帧图像数据集以及采

集待检测遗留物品的视频流图像数据,对监控视频帧图像数据集进行预处理、并对监控视频帧图像数据集中的每一张图像设置标签类别,获得预处理后的监控视频帧图像数据集以及预处理后的监控视频帧图像数据集相对应的标签数据集;按时间顺序获取视频流图像数据中的若干个关键帧图像;

[0039] 目标检测网络模型建立模块,用于建立目标检测网络模型,目标检测网络模型为加入SPP模块的卷积神经网络模型,根据监控视频帧图像采集模块中预处理后的监控视频帧图像数据集以及标签数据集对目标检测网络模型进行训练,获得训练完成的目标检测网络模型;

[0040] 遗留物品检测模块,用于将监控视频帧图像采集模块采集的若干个关键帧图像按时间先后顺序分别输入训练完成的目标检测网络模型,训练完成的目标检测网络模型分别输出每一个关键帧图像中所有物体的坐标边界框和标签;根据每一个关键帧图像中所有物体的坐标边界框计算每一个关键帧图像所有物体的质心坐标,根据质心坐标、并采用前后帧最小质心距离匹配方法匹配出前后不同关键帧图像中的同一个物体,获取匹配出的前后不同关键帧图像中的同一个物体在遗留后的静止时长,并根据静止时长输出物体遗留预警信息。

[0041] 通过采用本发明公开的上述技术方案,得到了如下有益的效果:

[0042] 与现有图像检测物体方法相比,本发明提出了一种基于深度学习的物品遗留方法,利用计算机视觉技术对物品与行人检测,提供了一种长时间稳定的检测器,比传统的背景建模方法相比,本发明在物体遗留较长时间后更具有鲁棒性和稳定性。

[0043] 与基于Darknet53的检测方法相比,本发明通过对图像进行数据增强、在预测特征图前面融入SPP模块,提升了小物品的召回率。

[0044] 对于物品遗留事件,本发明通过基于质心最小距离匹配不同帧的每个物体,该方法在检测器输出的边界框信息快速匹配同一个物体,匹配准确率高,保证了物品遗留后及时发出预警信息。

[0045] 以上仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视本发明的保护范围。

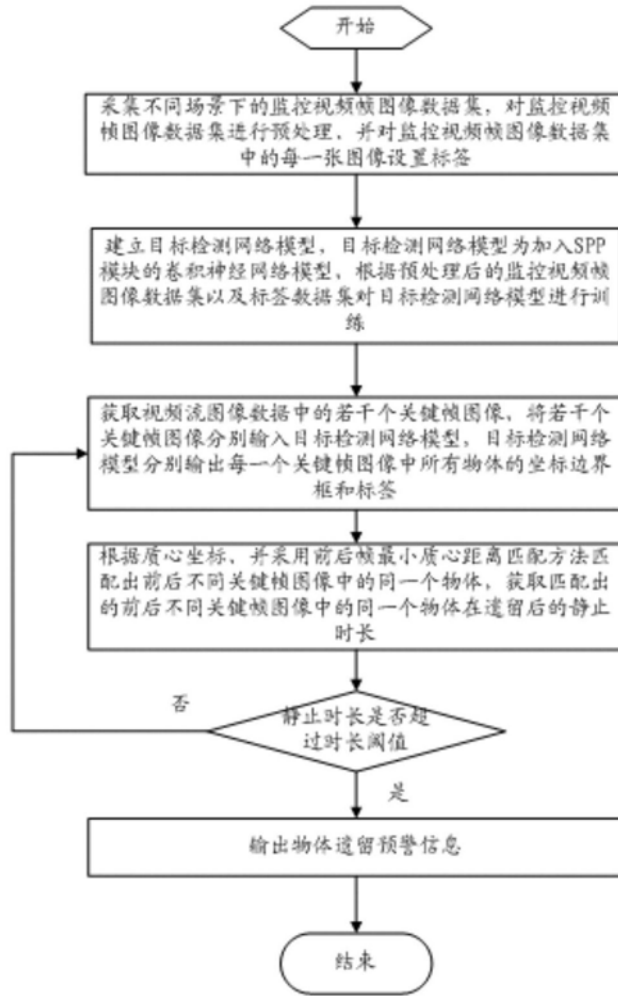


图1

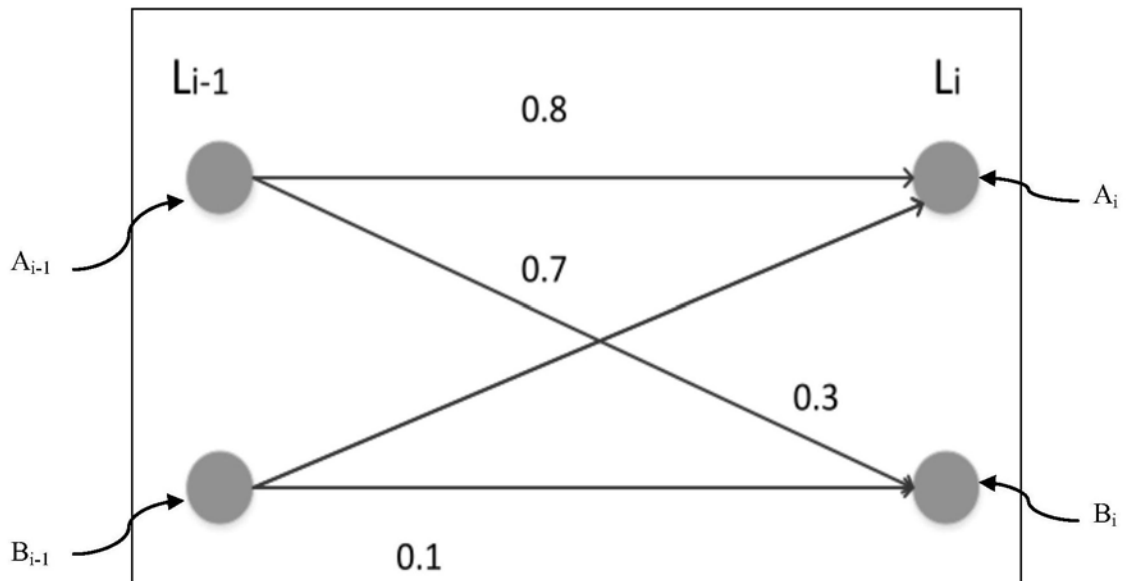


图2