



(12) 发明专利

(10) 授权公告号 CN 113192322 B

(45) 授权公告日 2022. 11. 25

(21) 申请号 202110294815.5

G06K 9/62 (2022.01)

(22) 申请日 2021.03.19

G06F 9/50 (2006.01)

(65) 同一申请的已公布的文献号
申请公布号 CN 113192322 A

(56) 对比文件

CN 103238136 A, 2013.08.07

CN 106793045 A, 2017.05.31

(43) 申请公布日 2021.07.30

CN 107341041 A, 2017.11.10

(73) 专利权人 东北大学
地址 110819 辽宁省沈阳市和平区文化路3号巷11号

US 2015089505 A1, 2015.03.26

吕相文等. 云计算环境下多GPU资源调度机制研究.《小型微型计算机系统》.2016, (第04期),

(72) 发明人 郭军 高赫妍 蒙龙 闫泽宇
张斌

审查员 杨钰

(74) 专利代理机构 沈阳东大知识产权代理有限公司 21109
专利代理师 李珉

(51) Int. Cl.
G08G 1/01 (2006.01)

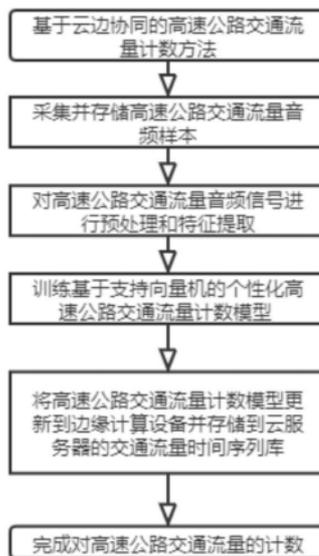
权利要求书3页 说明书9页 附图3页

(54) 发明名称

一种基于云边协同的高速公路交通流量计数方法

(57) 摘要

本发明提供一种基于云边协同的高速公路交通流量计数方法,涉及边缘计算技术领域。本方法中边缘计算设备负责采集高速公路交通流量音频样本,云服务器对交通流量音频进行预处理和特征提取,训练交通流量计数支持向量机模型,为避免资源浪费和任务等待时间过长,自适应调整计算资源,最后将高速公路交通流量计数模型更新到边缘计算设备上,由于边缘计算设备的存储空间和计算能力有限,边缘计算设备不能独立训练基于音频的高速公路交通流量计数模型,充分考虑了边缘计算设备的存储和计算能力,采用云边协同的方式缓解云服务器对于图像处理计算量的计算压力和网络传输带宽压力大的问题,有效解决了有限边缘计算能力下高速公路交通流量计数问题。



1. 一种基于云边协同的高速公路交通流量计数方法,其特征在于,包括以下步骤:

步骤1:边缘计算设备采集高速公路交通流量音频样本,然后将音频样本传送到云服务器并存储在交通流量音频样本库中,每一个边缘计算设备生成自己的音频样本库;

步骤2:云服务器根据每个边缘计算设备的高速公路交通流量音频样本库,对音频信号进行预处理和特征提取;

步骤2.1:对高速公路交通流量音频信号进行预处理;

步骤2.2:对预处理后的高速公路交通流量音频信号的每一帧数据进行快速傅里叶变换Fast Fourier Transformation,即FFT,将高速公路交通流量音频信号时域分析转换成频域分析,得到频域信号;

步骤2.3:将高速公路交通流量音频信号中每一帧频域信号通过梅尔滤波器,将频率标度转化为梅尔频率标度,公式如下:

$$mel(f) = 2595 \times \log_{10} \left(1 + \frac{f}{700} \right) \quad (5)$$

其中, $mel(f)$ 表示梅尔频率标度, f 表示频率标度;

步骤2.4:计算高速公路交通流量音频信号每一帧滤波后的能量和并取对数;

$$y = \lg \left(\sum_{i=1}^n x_i \right)$$

其中, y 表示计算结果, x_i 表示高速公路交通流量音频信号的第*i*帧滤波后的能量

步骤2.5:对步骤2.4计算结果进一步做离散余弦变换Discrete Cosine Transform,即DCT,得到梅尔倒谱系数,得到高速公路交通流量音频信号的特征和标签,生成高速公路交通流量训练样本数组;

步骤3:训练基于支持向量机Support Vector Machine,即SVM的个性化高速公路交通流量计数模型,所述个性化高速公路交通流量计数模型为对所述训练样本数组进行训练得到的模型;

步骤3.1将高速公路交通流量分成0,1,2,...,k,一共k+1个类别,将经过步骤2.5生成的高速公路交通流量音频计算模型样本集作为高速公路交通流量分类器模型的输入;

步骤3.2:将高速公路交通流量的每一个类别作为一个二分类问题,所述二分类问题指在包含两类事项的比较研究中,按两个标志所作的分类的问题,属于该类别的高速公路交通流量音频信号特征作为该类别的正样本,其余类别的高速公路交通流量音频信号特征作为负样本,训练k+1个二分类器,具体方法如下:

步骤3.2.1:在高速公路交通流量训练样本集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in \{-1, +1\}$ 构成的特征空间中,找到最大间隔划分超平面,将不同类别的样本分开,确定高速公路交通流量分类器模型,具体公式如下:

$$f(x) = \omega^T x + b \quad (6)$$

其中, ω 和**b**为模型参数, x 为高速公路交通流量训练样本集, (x_m, y_m) 为高速公路交通流量训练样本集中的第*m*个样本;

步骤3.2.2:将在高速公路交通流量音频信号特征构成的特征空间寻找最优分类超平面的问题转化成如下公式所示的二次规划问题,所述二次规划是非线性规划中的数学

规划问题:

$$\min_{\omega, b, \xi_i} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (7)$$

$$\begin{aligned} \text{s.t.} \quad & y_i (\omega^T x_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i=1, 2, \dots, N \end{aligned} \quad (8)$$

式中, ξ_i 为第 i 个样本的松弛系数, 满足约束条件的高速公路交通流量音频信号特征样本 $\xi=0$, 其中违反最大间隔规则的高速公路交通流量音频信号特征样本 $\xi>0$; C 为惩罚系数, C 越大对违反约束样本的惩罚力度越大; N 为高速公路交通流量训练样本集的样本数;

所述SVM的原始目标, 即间隔最大化, 并将其表示为约束最优化问题;

步骤3.2.3: 使用拉格朗日乘子得到拉格朗日函数, 如下公式所示:

$$L(\omega, b, \alpha, \xi, \mu) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i + \sum_{i=1}^N \alpha_i (1 - \xi_i - y_i (\omega^T x_i + b)) - \sum_{i=1}^N \mu_i \xi_i \quad (9)$$

其中, α_i 和 μ_i 表示为拉格朗日乘子, 且 $\alpha_i \geq 0, \mu_i \geq 0$;

步骤3.2.4: 通过拉格朗日对偶将SVM的优化问题转化为等价对偶问题, 对偶化得到表达式如下所示:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i \quad (10)$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, i=1, 2, \dots, N \end{aligned} \quad (11)$$

其中, $K(x_i, x_j)$ 表示为核函数, α_i 表示为拉格朗日乘子;

步骤3.2.5: 当公式(10)最小时对应的 α 向量的值为 α^* 向量, 则 b 的估计值 b^* 的表达式如下:

$$b^* = y_j - \sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) \quad (12)$$

步骤3.2.6: 得到高速公路交通流量音频信号分类决策函数 $g(x)$, 公式如下:

$$g(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) + b^* \right) \quad (13)$$

其中, sign 表示为符号函数;

步骤3.3: 由于每个交通流量音频样本库的样本累积速率不同, 根据交通流量音频样本库中样本累计达到阈值的时间先后顺序生成模型训练任务队列, 云服务器自适应调整计算资源训练高速公路交通流量计数模型;

步骤3.3.1: 定义每一个高速公路交通流量音频样本 sample 包含边缘计算设备编号 edge_id 、交通流量音频文件名 filename 和交通流量值 vehicle_count ;

步骤3.3.2: 根据每个边缘计算设备对应的高速公路交通流量音频样本库中样本数量达到阈值的先后顺序, 生成模型训练任务并加入到模型训练任务队列中;

步骤3.3.2.1: 任务调度器不断从任务队列中提取任务, 申请计算资源, 资源分配器分

配虚拟机来完成模型训练任务：

当任务队列中的模型训练任务数变大时，任务调度器申请的计算资源也随之增大；

当任务队列中的模型训练任务数变小时，任务调度器申请的计算资源也随之减少；

步骤3.3.2.2:为了避免任务等待时间过长和计算资源浪费，随着任务数量的变化，申请的计算资源数量做出调整，满足如下公式：

$$\frac{n}{m_{\min}} \leq k_1 \quad (14)$$

$$m_{\min} = \left\lceil \frac{n}{k_1} \right\rceil \quad (15)$$

其中，n表示为任务队列中任务数量， m_{\min} 表示为模型训练时虚拟机数量最少值， k_1 表示为平均每台虚拟机分配的训练任务阈值；

步骤3.3.3:云服务器自适应调整计算资源训练高速公路交通流量计数模型；

若当前当前正在占用的虚拟机数量 vm_{count} 大于模型训练时虚拟机数量最少值 m_{\min} ，则关闭 $vm_{\text{count}} - m_{\min}$ 个虚拟机；

若当前当前正在占用的虚拟机数量 vm_{count} 小于模型训练时虚拟机数量最少值 m_{\min} ，则申请 $vm_{\text{count}} - m_{\min}$ 个虚拟机；

步骤4:将高速公路交通流量计数模型更新到边缘计算设备上，边缘计算设备将高速公路交通流量计数结果存储到云服务器的交通流量时间序列库中，由此完成了对高速公路交通流量的计数。

2.根据权利要求1所述的一种基于云边协同的高速公路交通流量计数方法，其特征在于，步骤2.1具体包括以下步骤：

步骤2.1.1:通过一阶数字滤波器对高速公路交通流量音频信号的高频部分进行预加重，让该段音频信号的频谱波动区间相对减小，如下公式所示：

$$H(z) = 1 - \alpha z^{-1} \quad (1)$$

$$y(t) = x(t) - \alpha x(t-1) \quad (2)$$

其中， $H(z)$ 表示传递函数， α 表示预加重系数， $0.9 < \alpha < 1.0$ ， z 表示高速公路交通流量音频信号的高频部分， $x(t)$ 表示t时刻的高速公路交通流量音频信号采样值， $y(t)$ 表示预加重后的音频信号；

步骤2.1.2:采用可移动的有限长度窗口进行加权的方法对步骤2.1.1中得到的预加重后的高速公路交通流量音频信号进行分帧；

步骤2.1.3:采用汉明窗函数对经步骤2.1.2处理后得到的高速公路交通流量音频信号进行加窗，加窗后的高速公路交通流量音频信号表示如下：

$$w(x) = \begin{cases} 0.54 - 0.46 \cos \left[\frac{2\pi x}{N-1} \right], & 0 \leq x \leq N-1 \\ 0, & \text{其他} \end{cases} \quad (3)$$

$$s_w(x) = s(x) * w(x) \quad (4)$$

其中， $w(x)$ 表示汉明窗函数， N 为正整数， $s_w(x)$ 表示加窗后的高速公路交通流量音频信号， $s(x)$ 表示原始高速公路交通流量音频信号。

一种基于云边协同的高速公路交通流量计数方法

技术领域

[0001] 本发明涉及边缘计算技术领域,尤其涉及一种基于云边协同的高速公路交通流量计数方法。

背景技术

[0002] 近年来随着物联网技术的不断创新,智能交通系统也在不断发展和完善。智能交通被认为是一个涵盖路线优化、停车、路灯、交通流量、事故预防/检测、道路异常和基础设施应用的总括术语。当前智能交通系统中用于高速公路交通流量监测的常用技术为视频监测技术,该技术通过摄像头拍摄的视频,对高速公路上的车辆进行检测识别和跟踪,分析交通流量状态。但是使用视频检测的方法仍存在以下不足:

[0003] (1)摄像头受外界环境影响较大。摄像头在光照条件好时检测效果好,当光线较暗时,摄像头拍摄的视频可能会模糊,特别是出现遮挡和异常天气时,视频检测的效果大大降低甚至有时会导致检测失效,同时摄像头的拍摄范围有限。

[0004] (2)图像处理计算量大。摄像头拍摄的视频是一帧一帧图像组成,运用视频检测技术和图像处理技术对采集的数据进行分析,得到交通信息。计算机处理大量交通图片时,计算机的计算压力增大,而高性能的计算机和图像处理的硬件设备成本较大。

[0005] (3)网络传输带宽压力大。网络带宽的增长速度远不及数据的增长速度,信息采集设备将采集的交通信息不经过任何处理直接传送到服务器,由于信息采集的连续性,信息传送也是连续不断的,会导致网络信息传输的压力增大。

[0006] 现在也存在一些通过音频分析交通流量状态的方法,通过车辆碰撞的声音检测交通异常是最直接的检测方法,但是事故地点距离采集音频的设备较远时,音频设备可能无法采集到碰撞声音从而判断失误。同时,由于每个采集高速公路交通流量音频样本的边缘计算设备所处的环境不同,每一个边缘计算设备需要训练单独的高速公路交通流量计数模型,但是由于边缘计算设备的存储空间和计算能力有限,不能在边缘计算设备上训练高速公路交通流量计数模型,所以前人方法中对于高速公路交通流量统计的效果并不理想。

发明内容

[0007] 本发明要解决的技术问题是针对上述现有技术的不足,一种基于云边协同的高速公路交通流量计数方法,解决有限边缘计算能力下高速公路交通流量计数问题。

[0008] 为解决上述技术问题,本发明所采取的技术方案是:

[0009] 一种基于云边协同的高速公路交通流量计数方法,包括以下步骤:

[0010] 步骤1:边缘计算设备采集高速公路交通流量音频样本,然后将音频样本传送到云服务器并存储在交通流量音频样本库中,每一个边缘计算设备生成自己的音频样本库;

[0011] 步骤2:云服务器根据每个边缘计算设备的高速公路交通流量音频样本库,对音频信号进行预处理和特征提取,生成高速公路交通流量训练样本;

[0012] 步骤2.1:对高速公路交通流量音频信号进行预处理;

[0013] 步骤2.1.1:通过一阶数字滤波器对高速公路交通流量音频信号的高频部分进行预加重,让该段音频信号的频谱波动区间相对减小,如下公式所示:

$$[0014] \quad H(z) = 1 - \alpha z^{-1} \quad (1)$$

$$[0015] \quad y(t) = x(t) - \alpha x(t-1) \quad (2)$$

[0016] 其中, $H(z)$ 表示传递函数, α 表示预加重系数, $0.9 < \alpha < 1.0$, z 表示高速公路交通流量音频信号的高频部分, $x(t)$ 表示 t 时刻的高速公路交通流量音频信号采样值, $y(t)$ 表示预加重后的音频信号;

[0017] 步骤2.1.2:采用可移动的有限长度窗口进行加权的方法对步骤2.1.1中得到的预加重后的高速公路交通流量音频信号进行分帧;

[0018] 步骤2.1.3:采用汉明窗函数对经步骤2.1.2处理后得到的高速公路交通流量音频信号进行加窗,加窗后的高速公路交通流量音频信号表示如下:

$$[0019] \quad w(x) = \begin{cases} 0.54 - 0.46 \cos\left[\frac{2\pi x}{N-1}\right], & 0 \leq x \leq N-1 \\ 0, & \text{其他} \end{cases} \quad (3)$$

$$[0020] \quad s_w(x) = s(x) * w(x) \quad (4)$$

[0021] 其中, $w(x)$ 表示汉明窗函数, N 为正整数, $s_w(x)$ 表示加窗后的高速公路交通流量音频信号, $s(x)$ 表示原始高速公路交通流量音频信号;

[0022] 步骤2.2:对预处理后的高速公路交通流量音频信号的每一帧数据进行快速傅里叶变换(Fast Fourier Transformation,FFT),将高速公路交通流量音频信号时域分析转换成频域分析,得到频域信号;

[0023] 步骤2.3:将高速公路交通流量音频信号中每一帧频域信号通过梅尔滤波器,将频率标度转化为梅尔频率标度,公式如下:

$$[0024] \quad mel(f) = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right) \quad (5)$$

[0025] 其中, $mel(f)$ 表示梅尔频率标度, f 表示频率标度;

[0026] 步骤2.4:计算高速公路交通流量音频信号每一帧滤波后的能量并取对数;

$$[0027] \quad y = \lg\left(\sum_{i=1}^n x_i\right)$$

[0028] 其中, y 表示计算结果, x_i 表示高速公路交通流量音频信号的第 i 帧滤波后的能量;

[0029] 步骤2.5:对步骤2.4计算结果进一步做离散余弦变换(Discrete Cosine Transform,DCT),得到梅尔倒谱系数,得到高速公路交通流量音频信号的特征和标签,生成高速公路交通流量训练样本数组;

[0030] 步骤3:训练基于支持向量机(Support Vector Machine,SVM)的个性化高速公路交通流量计数模型,所述个性化高速公路交通流量计数模型为对所述训练样本数组进行训练得到的模型;

[0031] 步骤3.1将高速公路交通流量分成 $0, 1, 2, \dots, k$,一共 $k+1$ 个类别,将经过步骤2.5生成的高速公路交通流量音频计算模型样本集作为高速公路交通流量分类器模型的输入;

[0032] 步骤3.2:将高速公路交通流量的每一个类别作为一个二分类问题,所述二分类问题指在包含两类事项的比较研究中,按两个标志所作的分类的问题,属于该类别的高速公

路交通流量音频信号特征作为该类别的正样本,其余类别的高速公路交通流量音频信号特征作为负样本,训练k+1个二类分类器,具体方法如下:

[0033] 步骤3.2.1:在高速公路交通流量训练样本集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in \{-1, +1\}$ 构成的特征空间中,找到最大间隔划分超平面,将不同类别的样本分开,确定高速公路交通流量分类器模型,具体公式如下:

$$[0034] \quad f(x) = \omega^T x + b \quad (6)$$

[0035] 其中, ω 和 b 为模型参数, x 为高速公路交通流量训练样本集, (x_m, y_m) 为高速公路交通流量训练样本集中的第 m 个样本;

[0036] 步骤3.2.2:将在高速公路交通流量音频信号特征构成的特征空间寻找最优分类超平面的问题转化成一个如下公式所示的二次规划问题,所述二次规划是非线性规划中的数学规划问题:

$$[0037] \quad \min_{\omega, b, \xi_i} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (7)$$

$$[0038] \quad \begin{aligned} \text{s.t.} \quad & y_i (\omega^T x_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \dots, N \end{aligned} \quad (8)$$

[0039] 式中, ξ_i 为第 i 个样本的松弛系数,满足约束条件的高速公路交通流量音频信号特征样本 $\xi = 0$,其中违反最大间隔规则的高速公路交通流量音频信号特征样本 $\xi > 0$; C 为惩罚系数, C 越大对违反约束样本的惩罚力度越大; N 为高速公路交通流量训练样本集的样本数;

[0040] 所述SVM的原始目标,即间隔最大化,并将其表示为约束最优化问题;

[0041] 步骤3.2.3:使用拉格朗日乘子得到拉格朗日函数,如下公式所示:

$$[0042] \quad L(\omega, b, \alpha, \xi, \mu) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i + \sum_{i=1}^N \alpha_i (1 - \xi_i - y_i (\omega^T x_i + b)) - \sum_{i=1}^N \mu_i \xi_i \quad (9)$$

[0043] 其中, α_i 和 μ_i 表示为拉格朗日乘子,且 $\alpha_i \geq 0, \mu_i \geq 0$;

[0044] 步骤3.2.4:通过拉格朗日对偶将SVM的优化问题转化为等价对偶问题,对偶化得到表达式如下所示:

$$[0045] \quad \min_{\alpha} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i \quad (10)$$

$$[0046] \quad \begin{aligned} \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, 2, \dots, N \end{aligned} \quad (11)$$

[0047] 其中, $K(x_i, x_j)$ 表示为核函数, α_i 表示为拉格朗日乘子;

[0048] 步骤3.2.5:当公式(10)最小时对应的 α 向量的值为 α^* 向量,则 b 的估计值 b^* 的表达式如下:

$$[0049] \quad b^* = y_j - \sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) \quad (12)$$

[0050] 步骤3.2.6:得到高速公路交通流量音频信号分类决策函数 $g(x)$, 公式如下:

$$[0051] \quad g(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) + b^* \right) \quad (13)$$

[0052] 其中, sign表示为符号函数;

[0053] 步骤3.3: 由于每个交通流量音频样本库的样本累积速率不同, 根据交通流量音频样本库中样本累计达到阈值的时间先后顺序生成模型训练任务队列, 云服务器自适应调整计算资源训练高速公路交通流量计数模型;

[0054] 步骤3.3.1: 定义每一个高速公路交通流量音频样本sample包含边缘计算设备编号edge_id、交通流量音频文件名filename和交通流量值vehicle_count;

[0055] 步骤3.3.2: 根据每个边缘计算设备对应的高速公路交通流量音频样本库中样本数量达到阈值的先后顺序, 生成模型训练任务并加入到模型训练任务队列中;

[0056] 步骤3.3.2.1: 任务调度器不断从任务队列中提取任务, 申请计算资源, 资源分配器分配虚拟机来完成模型训练任务;

[0057] 当任务队列中的模型训练任务数变大时, 任务调度器申请的计算资源也随之增大;

[0058] 当任务队列中的模型训练任务数变小时, 任务调度器申请的计算资源也随之减少;

[0059] 步骤3.3.2.2: 为了避免任务等待时间过长和计算资源浪费, 随着任务数量的变化, 申请的计算资源数量做出调整, 满足如下公式:

$$[0060] \quad \frac{n}{m_{\min}} \leq k_1 \quad (14)$$

$$[0061] \quad m_{\min} = \left\lceil \frac{n}{k_1} \right\rceil \quad (15)$$

[0062] 其中, n表示为任务队列中任务数量, m_{\min} 表示为模型训练时虚拟机数量最少值, k_1 表示为平均每台虚拟机分配的训练任务阈值;

[0063] 步骤3.3.3: 云服务器自适应调整计算资源训练高速公路交通流量计数模型;

[0064] 若当前当前正在占用的虚拟机数量 vm_{count} 大于模型训练时虚拟机数量最少值 m_{\min} , 则关闭 $vm_{\text{count}} - m_{\min}$ 个虚拟机;

[0065] 若当前当前正在占用的虚拟机数量 vm_{count} 小于模型训练时虚拟机数量最少值 m_{\min} , 则申请 $vm_{\text{count}} - m_{\min}$ 个虚拟机;

[0066] 步骤4: 将高速公路交通流量计数模型更新到边缘计算设备上, 边缘计算设备将高速公路交通流量计数结果存储到云服务器的交通流量时间序列库中, 由此完成了对高速公路交通流量的计数。

[0067] 采用上述技术方案所产生的有益效果在于:

[0068] 本发明提供了一种基于云边协同的高速公路交通流量计数方法, 边缘计算设备负责采集高速公路交通流量音频样本, 云服务器对交通流量音频进行预处理和特征提取, 训练交通流量计数支持向量机模型, 为避免资源浪费和任务等待时间过长, 自适应调整计算资源, 最后将高速公路交通流量计数模型更新到边缘计算设备上, 由于边缘计算设备的存储空间和计算能力有限, 边缘计算设备不能独立训练基于音频的高速公路交通流量计数模

型,本发明充分考虑了边缘计算设备的存储和计算能力,采用云边协同的方式缓解云服务器对于图像处理计算量的计算压力和网络传输带宽压力大的问题,有效解决了有限边缘计算能力下高速公路交通流量计数问题。

附图说明

- [0069] 图1为本发明实施例提供的总体流程图;
 [0070] 图2为本发明实施例提供的原始音频波形图;
 [0071] 图3为本发明实施例提供的原始音频频谱图;
 [0072] 图4为本发明实施例提供的预加重音频波形图;
 [0073] 图5为本发明实施例提供的预加重音频频谱图;
 [0074] 图6为本发明实施例提供的模型训练任务数随时间变化图;
 [0075] 图7为本发明实施例提供的计算资源随时间变化图。

具体实施方式

[0076] 下面结合附图和实施例,对本发明的具体实施方式作进一步详细描述。以下实施例用于说明本发明,但不用来限制本发明的范围。

[0077] 一种基于云边协同的高速公路交通流量计数方法,如图1所示,包括以下步骤:

[0078] 步骤1:边缘计算设备采集高速公路交通流量音频样本,然后将音频样本传送到云服务器并存储在交通流量音频样本库中,每一个边缘计算设备生成自己的音频样本库;

[0079] 本实施例以python为编程语言,以HPZ820工作站作为云服务器,HPZ820工作站的配置参数如表1所示。

[0080] 表1 HPZ820服务器配置参数表

性能指标	性能参数
CPU 类型	四核至强 E5-2603
CPU 主频	1.8GHz
内存类型	DDR3-1333 ECC
内存容量	8GB
硬盘类型	15000 转 SAS 硬盘
硬盘容量	300GB
带宽	1000Mbps

[0082] 边缘计算设备将高速公路交通音频样本传送到云服务器高速公路交通流量音频样本库中,云服务器为边缘计算设备训练高速公路交通流量支持向量机模型。以其中一个边缘计算设备的训练样本库为例,样本标签分为0辆车,1辆车,2辆车,3辆车,训练样本集组成如表2所示。

[0083] 表2训练样本集组成

	样本标签	样本数量 (个)
	0 辆车	151
[0084]	1 辆车	80
	2 辆车	58
	3 辆车	55

[0085] 步骤2:云服务器根据每个边缘计算设备的高速公路交通流量音频样本库,对音频信号进行预处理和特征提取;

[0086] 步骤2.1:对高速公路交通流量音频信号进行预处理;

[0087] 步骤2.1.1:通过一阶数字滤波器对高速公路交通流量音频信号的高频部分进行预加重,让该段音频信号的频谱波动区间相对减小,如下公式所示:

$$[0088] \quad H(z) = 1 - \alpha z^{-1} \quad (1)$$

$$[0089] \quad y(t) = x(t) - \alpha x(t-1) \quad (2)$$

[0090] 其中, $H(z)$ 表示传递函数, α 表示预加重系数, $0.9 < \alpha < 1.0$, z 表示高速公路交通流量音频信号的高频部分, $x(t)$ 表示 t 时刻的高速公路交通流量音频信号采样值, $y(t)$ 表示预加重后的音频信号;

[0091] 步骤2.1.2:采用可移动的有限长度窗口进行加权的方法对步骤2.1.1中得到的预加重后的高速公路交通流量音频信号进行分帧;

[0092] 步骤2.1.3:采用汉明窗函数对经步骤2.1.2处理后得到的高速公路交通流量音频信号进行加窗,加窗后的高速公路交通流量音频信号表示如下:

$$[0093] \quad w(x) = \begin{cases} 0.54 - 0.46 \cos\left[\frac{2\pi x}{N-1}\right], & 0 \leq x \leq N-1 \\ 0, & \text{其他} \end{cases} \quad (3)$$

$$[0094] \quad s_w(x) = s(x) * w(x) \quad (4)$$

[0095] 其中, $w(x)$ 表示汉明窗函数, N 为正整数, $s_w(x)$ 表示加窗后的高速公路交通流量音频信号, $s(x)$ 表示原始高速公路交通流量音频信号;

[0096] 声音预处理是音频信号分析的基础,本实施例中对声音进行预处理前后的波形图和频谱图对比图如图2-图5所示。

[0097] 由图2-图5原始音频和预处理音频对比图可以看出,声音经过预处理后高频部分的幅度得到了提升。

[0098] 步骤2.2:对预处理后的高速公路交通流量音频信号的每一帧数据进行快速傅里叶变换(Fast Fourier Transformation,FFT),将高速公路交通流量音频信号时域分析转换成频域分析,得到频域信号;

[0099] 步骤2.3:将高速公路交通流量音频信号中每一帧频域信号通过梅尔滤波器,将频率标度转化为梅尔频率标度,公式如下:

$$[0100] \quad mel(f) = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right) \quad (5)$$

[0101] 其中, $mel(f)$ 表示梅尔频率标度, f 表示频率标度;

[0102] 步骤2.4:计算高速公路交通流量音频信号每一帧滤波后的能量和并取对数;

$$[0103] \quad y = \lg \left(\sum_{i=1}^n x_i \right)$$

[0104] 其中, y 表示计算结果, x_i 表示高速公路交通流量音频信号的第*i*帧滤波后的能量

[0105] 步骤2.5:对步骤2.4计算结果进一步做离散余弦变换(Discrete Cosine Transform,DCT),得到梅尔倒谱系数,得到高速公路交通流量音频信号的特征和标签,生成高速公路交通流量训练样本数组;

[0106] 步骤3:训练基于支持向量机(Support Vector Machine,SVM)的个性化高速公路交通流量计数模型,所述个性化高速公路交通流量计数模型为对所述训练样本数组进行训练得到的模型;

[0107] 步骤3.1将高速公路交通流量分成 $0,1,2,\dots,k$,一共 $k+1$ 个类别,将经过步骤2.5生成的高速公路交通流量音频计算模型样本集作为高速公路交通流量分类器模型的输入;

[0108] 步骤3.2:将高速公路交通流量的每一个类别作为一个二分类问题,所述二分类问题指在包含两类事项的比较研究中,按两个标志所作的分类的问题,属于该类别的高速公路交通流量音频信号特征作为该类别的正样本,其余类别的高速公路交通流量音频信号特征作为负样本,训练 $k+1$ 个二类分类器,具体方法如下:

[0109] 步骤3.2.1:在高速公路交通流量训练样本集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in \{-1, +1\}$ 构成的特征空间中,找到最大间隔划分超平面,将不同类别的样本分开,确定高速公路交通流量分类器模型,具体公式如下:

$$[0110] \quad f(x) = \omega^T x + b \quad (6)$$

[0111] 其中, ω 和 b 为模型参数, x 为高速公路交通流量训练样本集, (x_m, y_m) 为高速公路交通流量训练样本集中的第*m*个样本;

[0112] 步骤3.2.2:将在高速公路交通流量音频信号特征构成的特征空间寻找最优分类超平面的问题转化成如下公式所示的二次规划问题,所述二次规划是非线性规划中的数学规划问题,在很多方面都有应用,如投资组合、约束最小二乘问题的求解、序列二次规划在非线性优化问题中应用等:

$$[0113] \quad \min_{\omega, b, \xi_i} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (7)$$

$$[0114] \quad \begin{aligned} s.t. \quad & y_i (\omega^T x_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \dots, N \end{aligned} \quad (8)$$

[0115] 式中, ξ_i 为第*i*个样本的松弛系数,满足约束条件的高速公路交通流量音频信号特征样本 $\xi = 0$,其中违反最大间隔规则的高速公路交通流量音频信号特征样本 $\xi > 0$;C为惩罚系数,C越大对违反约束样本的惩罚力度越大;N为高速公路交通流量训练样本集的样本数;

[0116] 所述SVM的原始目标,即间隔最大化,并将其表示为约束最优化问题;

[0117] 步骤3.2.3:使用拉格朗日乘子得到拉格朗日函数,如下公式所示:

$$[0118] \quad L(\omega, b, \alpha, \xi, \mu) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i + \sum_{i=1}^N \alpha_i (1 - \xi_i - y_i (\omega^T x_i + b)) - \sum_{i=1}^N \mu_i \xi_i \quad (9)$$

[0119] 其中, α_i 和 μ_i 表示为拉格朗日乘子, 且 $\alpha_i \geq 0, \mu_i \geq 0$;

[0120] 步骤3.2.4: 通过拉格朗日对偶将SVM的优化问题转化为等价对偶问题, 对偶化得到表达式如下所示:

$$[0121] \quad \min_{\alpha} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^N \alpha_i \quad (10)$$

$$[0122] \quad s.t. \quad \sum_{i=1}^N \alpha_i y_i = 0 \quad (11)$$

$$0 \leq \alpha_i \leq C, i = 1, 2, \dots, N$$

[0123] 其中, $K(x_i, x_j)$ 表示为核函数, α_i 表示为拉格朗日乘子;

[0124] 步骤3.2.5: 当公式(10)最小时对应的 α 向量的值为 α^* 向量, 则b的估计值 b^* 的表达式如下:

$$[0125] \quad b^* = y_j - \sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) \quad (12)$$

[0126] 步骤3.2.6: 得到高速公路交通流量音频信号分类决策函数 $g(x)$, 公式如下:

$$[0127] \quad g(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) + b^* \right) \quad (13)$$

[0128] 其中, sign表示为符号函数;

[0129] 步骤3.3: 由于每个交通流量音频样本库的样本累积速率不同, 根据交通流量音频样本库中样本累计达到阈值的时间先后顺序生成模型训练任务队列, 云服务器自适应调整计算资源训练高速公路交通流量计数模型;

[0130] 步骤3.3.1: 定义每一个高速公路交通流量音频样本sample包含边缘计算设备编号edge_id、交通流量音频文件名filename和交通流量值vehicle_count;

[0131] 步骤3.3.2: 根据每个边缘计算设备对应的高速公路交通流量音频样本库中样本数量达到阈值的先后顺序, 生成模型训练任务并加入到模型训练任务队列中;

[0132] 步骤3.3.2.1: 任务调度器不断从任务队列中提取任务, 申请计算资源, 资源分配器分配虚拟机来完成模型训练任务;

[0133] 当任务队列中的模型训练任务数变大时, 任务调度器申请的计算资源也随之增大;

[0134] 当任务队列中的模型训练任务数变小时, 任务调度器申请的计算资源也随之减少;

[0135] 步骤3.3.2.2: 为了避免任务等待时间过长和计算资源浪费, 随着任务数量的变化, 申请的计算资源数量做出调整, 满足如下公式:

$$[0136] \quad \frac{n}{m_{\min}} \leq k_1 \quad (14)$$

$$[0137] \quad m_{\min} = \left\lceil \frac{n}{k_1} \right\rceil \quad (15)$$

[0138] 其中, n表示为任务队列中任务数量, m_{\min} 表示为模型训练时虚拟机数量最少值, k_1 表示为平均每台虚拟机分配的训练任务阈值;

[0139] 步骤3.3.3:云服务器自适应调整计算资源训练高速公路交通流量计数模型;

[0140] 若当前当前正在占用的虚拟机数量 vm_{count} 大于模型训练时虚拟机数量最少值 m_{min} ,则关闭 $vm_{count} - m_{min}$ 个虚拟机;

[0141] 若当前当前正在占用的虚拟机数量 vm_{count} 小于模型训练时虚拟机数量最少值 m_{min} ,则申请 $vm_{count} - m_{min}$ 个虚拟机;

[0142] 用高速公路交通流量音频样本训练基于音频的高速公路交通流量计数支持向量机模型。高速公路交通流量音频样本库中共有344个交通流量音频样本,其中0辆车音频样本151个,1辆车音频样本80个,2辆车音频样本58个,3辆车音频样本55个。其中训练集为样本总数70%,240个,测试集样本数量104个。每个类型的交通流量计数结果准确率如表3所示。

[0143] 表3模型训练分类准确率

	交通流量类型	分类准确率
	0 辆车	0.94
[0144]	1 辆车	0.75
	2 辆车	0.64
	3 辆车	0.60

[0145] 由表3可以得出,无车时的分类精度最大,交通流量值越小,模型的分类准确越高。这是由于确定时间内车辆越多,车辆密度越大,车辆声音就会越混乱不容易区分具体的交通流量大小。

[0146] 步骤4:将高速公路交通流量计数模型更新到边缘计算设备上,边缘计算设备将高速公路交通流量计数结果存储到云服务器的交通流量时间序列库中,由此完成了对高速公路交通流量的计数。

[0147] 云服务器上根据每个边缘计算设备对应的高速公路交通流量音频样本库中样本累计达到阈值的先后顺序生成模型训练任务队列,根据任务队列中任务数的多少自适应调整分配的计算资源,设模型训练任务队列中的任务数为 n ,则模型训练任务数随时间变化如图6所示。设计算资源数为 m ,则分配计算资源数随时间变化如图7所示。

[0148] 由图6和图7得出,随着时间增加,训练模型任务数和计算资源数实现增加后减少。由于开始时边缘计算设备都在不断积累高速公路交通流量音频样本,当样本库中样本数量达到一定阈值时加入到模型训练任务队列中,随着训练任务越来越多申请更多的计算资源训练高速公路交通流量模型。当所有的边缘计算设备都采集完高速公路交通流量音频样本后,模型训练任务队列中的任务不再增加,伴随着高速公路交通流量计数模型的训练,任务数不断减少,为了避免资源浪费同时也在不断关闭计算资源,所以计算资源也在不断减少。

[0149] 最后应说明的是:以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明权利要求所限定的范围。

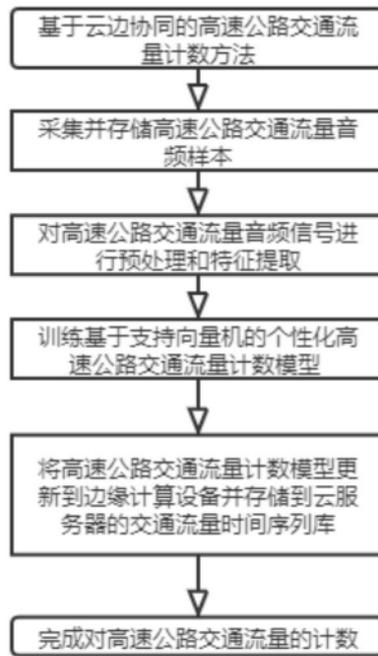


图1

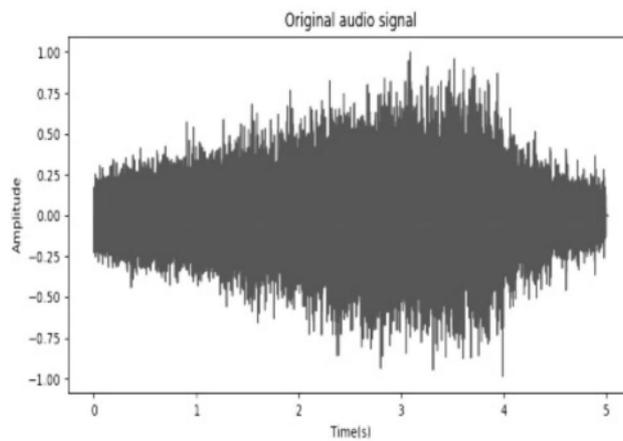


图2

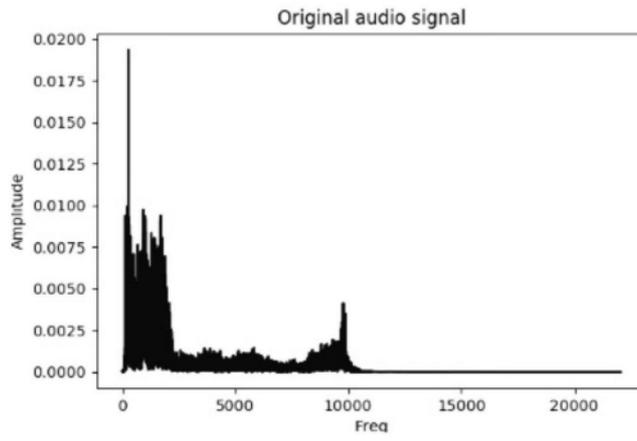


图3

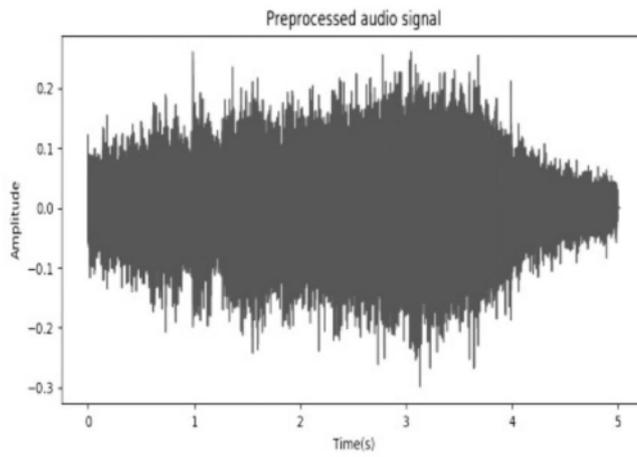


图4

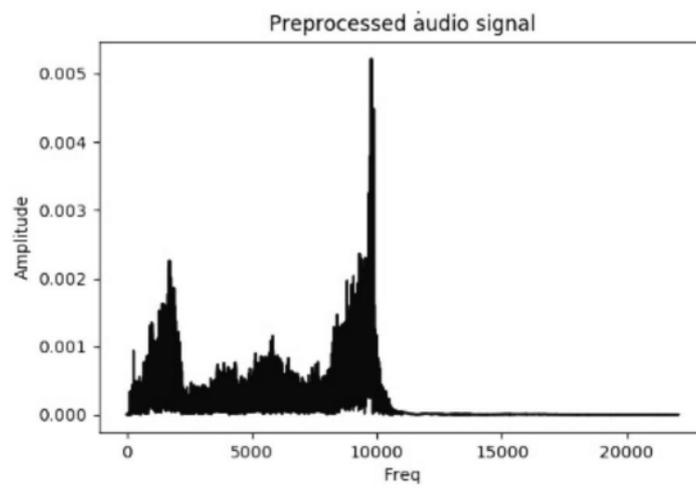


图5

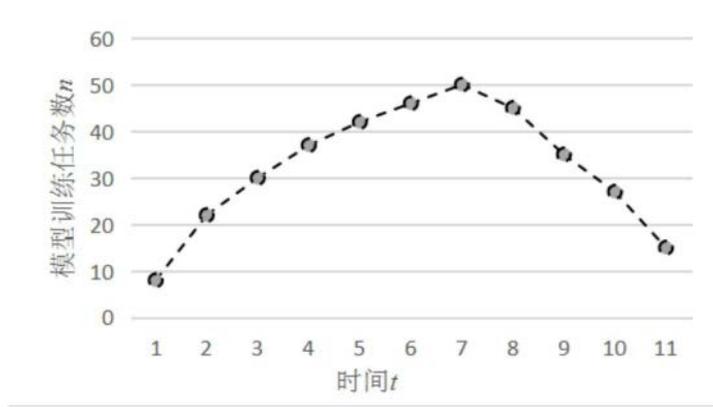


图6

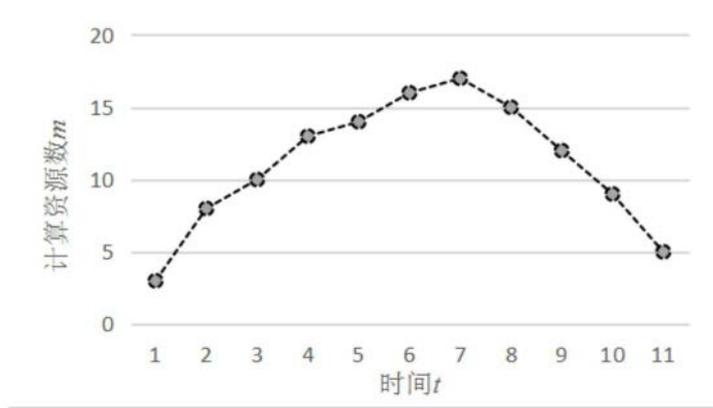


图7