

基于加法网络 and 自注意力机制的 OSA 检测方法

郑雅纯 郑和裕 林美娜

(广东工业大学自动化学院, 广东 广州 510006)

摘要: 为实现准确率高且计算复杂度低的阻塞性睡眠呼吸暂停 (OSA) 检测, 根据血氧饱和度 (SpO_2) 提出一种基于加法网络和自注意力机制的 OSA 检测方法。该方法引入加法网络用于 SpO_2 的特征提取, 以降低计算复杂度; 采用自注意力机制调整采样点间的特征权重, 以提高检测精度。在数据集 Apnea-ECG 上实现了敏感性 94.56%、特异性 96.91% 的检测性能, 且浮点运算数量为 2.88 K FLOPs, 比卷积操作减小了 40%。实验结果表明: 该方法在获得较高检测精度的同时, 可以保持较低的计算复杂度。

关键词: 血氧饱和度; 阻塞性睡眠呼吸暂停检测; 加法网络; 自注意力机制

中图分类号: TP391

文献标识码: A

文章编号: 1674-2605(2021)06-0008-06

DOI: 10.3969/j.issn.1674-2605.2021.06.008

0 引言

阻塞性睡眠呼吸暂停 (obstructive sleep apnea, OSA) 表现为睡眠期间患者的呼吸气流减少或呼吸受阻^[1]。研究表明, 长期 OSA 会增加人们患高血压、冠心病、中风等疾病的风险^[2-4]。2019 年, 柳叶刀呼吸医学杂志报告显示: 全球 30~69 岁人群中约 9.36 亿人患有 OSA^[5]。因此, 尽早发现和治疗 OSA 具有重要意义。血氧饱和度 (oxygen saturation, SpO_2) 是表征 OSA 事件的参考信号之一^[6]。其采集设备脉搏血氧计具有廉价、简单和非侵入式等特点^[1]。因此, 本文采用 SpO_2 作为 OSA 检测的参考信号。

过去十几年间, 研究人员提出了许多基于 SpO_2 经验特征的 OSA 检测方法。如 BURGOS 等采用 2 种时域特征并结合决策树实现 OSA 检测^[7]。MOSTAFA 等利用遗传算法、最小冗余最大相关性算法等特征选择算法和人工神经网络 (artificial neural network, ANN)、支持向量机 (support vector machine, SVM) 等分类器实现 OSA 检测^[8-9]。BEHAR 等利用数据特征、患者信息, 并结合逻辑回归算法检测 OSA^[10]。上述方法虽然有效实现 OSA 检测, 但严重依赖经验参数, 且未能很好地提取 SpO_2 潜在的复杂特征。近年

来, 不少研究人员将卷积神经网络应用于 OSA 检测。如 CHAW 等构建一个具有 10 个卷积层, 32 614 个参数的深度卷积神经网络检测 OSA^[11]; VAQU-ERIZOVILLAR 等构建基于卷积神经网络和贝叶斯优化方法的网络模型预测儿童 OSA 的严重程度^[12]; MOSTAFA 等开发具有自动网络结构及超参数选择能力的一维卷积网络 (one dimensional convolutional neural network, CNN1D) 检测 OSA^[1]。上述方法都取得了显著成果, 但存在以下问题尚未讨论: 1) 卷积网络运算量大, 运算复杂度高, 不利于在移动设备, 如手表、手机上部署; 2) SpO_2 采集过程中, 由于被测者肢体运动产生的大量运动伪迹造成 SpO_2 的非平稳变化, 影响 OSA 的检测精度^[13]。

为减少伪迹干扰, 提高检测性能, 同时降低模型运算复杂度, 本文提出一种基于加法网络 (AdderNet) 和自注意力机制的 OSA 检测方法。同时, 构建并训练可以准确识别含有呼吸暂停事件的 SpO_2 片段的网络模型, 从而实现基于 SpO_2 的 OSA 检测。其中, 加法网络采用 $\mathcal{L}_1$ 距离运算代替卷积操作, 提取 SpO_2 的波形特征, 以降低计算复杂度^[14]; 自注意力机制用于调整不同时间采样点特征的权重, 以适应伪迹干扰带来 SpO_2 的非平稳变化。

1 检测方法

1.1 问题描述

设数据集 $X = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ ，其中 $\mathbf{x}_i$ 表示第 i 个输入， $y_i \in \{1, 0\}$ 为 $\mathbf{x}_i$ 的标签，正样本的标签 $y_i = 1$ ，负样本的标签 $y_i = 0$ ，则输出 o_i 的定义为

$$o_i = F(\mathbf{x}_i; \theta) \quad (1)$$

式中， $F(\cdot)$ 为输入到输出的映射； θ 为网络参数。

针对训练集正负样本比例失衡的问题，引入 AUC 近似函数 $J(\cdot)$ 作为损失函数，如式(2)所示^[15]。

$$\min J(\theta) = \sum_{i=1}^{n^+} \sum_{j=1}^{n^-} \gamma(o_i^+, o_j^-) \quad (2)$$

式中， o^+ 为 n^+ 个正样本输出； o^- 为 n^- 个负样本输出。 $\gamma(o_i^+, o_j^-)$ 定义为

$$\gamma(o_i^+, o_j^-) = \begin{cases} \left[-(o_i^+ - o_j^- - 0.2) \right]^3, & o_i^+ - o_j^- < 0.2 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

对于给定 o_i ，预测标签 $\hat{y}_i$ 计算公式为

$$\hat{y}_i = \begin{cases} 1, & \text{if } o_i \geq th \\ 0, & \text{if } o_i < th \end{cases} \quad (4)$$

式中， th 通过最大化约登指数获得^[16]。

1.2 基于加法网络和自注意力机制的 OSA 检测方法

1.2.1 特征提取

为降低网络模型的计算复杂度，用加法网络代替卷积操作，实现 SpO₂ 形态特征的提取。加法网络是由 CHEN 等提出的一种仅含加法计算的神经网络，其加法层利用只具有加法的距离度量函数 $\mathcal{L}_1$ 距离，代替卷积神经网络中的卷积计算，以实现更高的速度和更低的能耗^[14]。

为了实现对采样点间 $\mathcal{L}_1$ 距离特征权重的调整，将加法核 F 与输入 $\mathbf{x}$ 对应元素之差的绝对值向量作为 $\mathcal{L}_1$ 距离加法层的输出，即特征向量 $\mathbf{y}$ ，如式(5)所示。

$$\mathbf{y} = (|x_1 - F_1|, |x_2 - F_2|, \dots, |x_n - F_n|) \quad (5)$$

1.2.2 特征权重调整

为适应伪迹干扰带来的 SpO₂ 非平稳变化，提高检测精度，利用自注意力机制生成含有相邻样本点间时序关系的掩码，以自适应调整特征权重。自注意力机制是由 VASWANI 等提出的一种注意力机制，广泛应用于文本识别、机器翻译等领域^[17]。

自注意力机制的数学计算过程如下：

- 1) 输入向量 $\mathbf{x}$ 被重新编码为特征向量 $\mathbf{a}$ ，即 $\mathbf{a} = \mathbf{w}^a \mathbf{x}$ ；
- 2) 根据输入向量 $\mathbf{x}$ 进行位置编码，生成位置向量 $\mathbf{e}$ ，即 $\mathbf{e} = \mathbf{w}^e \mathbf{x}$ ；
- 3) 根据特征向量 $\mathbf{a}$ 与位置向量 $\mathbf{e}$ 的和产生匹配项 $\mathbf{q}$ 、被匹配项 $\mathbf{k}$ 、特征项 $\mathbf{v}$ ，即 $\mathbf{q} = \mathbf{w}^q (\mathbf{a} + \mathbf{e})$ ， $\mathbf{k} = \mathbf{w}^k (\mathbf{a} + \mathbf{e})$ ， $\mathbf{v} = \mathbf{w}^v (\mathbf{a} + \mathbf{e})$ ；
- 4) 将 $\mathbf{q}$ 与 $\mathbf{k}$ 相乘得到匹配项与被匹配项的相似度 $\mathbf{k}^T \mathbf{q}$ ；
- 5) 将 $\mathbf{k}^T \mathbf{q}$ 与 $\mathbf{v}$ 相乘得到自注意力特征 $\mathbf{b}$ ，即 $\mathbf{b} = \mathbf{v} \mathbf{k}^T \mathbf{q}$ 。

上述， $\mathbf{w}^a$ ， $\mathbf{w}^e$ ， $\mathbf{w}^q$ 、 $\mathbf{w}^k$ 、 $\mathbf{w}^v$ 均为网络模型训练过程中需要进行学习的参数矩阵。

1.2.3 基于加法网络和自注意力机制的网络模型

基于加法网络和自注意力机制的网络模型 (neural network based on AdderNet and Self-attention, A-S-NN) 主要包括输入层、加法层、自注意力机制、全连接层及 Sigmoid 激活函数，如图 1 所示。

基于加法网络和自注意力机制的网络的具体数学计算如下：

- 1) 输入 $\mathbf{x}$ 经自注意力机制产生掩码 $\mathbf{b}$ 。

$$\mathbf{b} = \mathbf{v} \mathbf{k}^T \mathbf{q} \quad (6)$$

式中， $\mathbf{v} = \mathbf{w}^v (\mathbf{a} + \mathbf{e})$ ， $\mathbf{a} = \mathbf{w}^a \mathbf{x}$ ， $\mathbf{e}$ 为位置信息； $\mathbf{k} = \mathbf{w}^k (\mathbf{a} + \mathbf{e})$ ； $\mathbf{q} = \mathbf{w}^q (\mathbf{a} + \mathbf{e})$ 。

- 2) 加法层提取输入信号 $\mathbf{x}$ 到加法核 F 的 $\mathcal{L}_1$ 距离特征向量 $\mathbf{y}$ ，如式(5)所示。

- 3) 特征向量 $\mathbf{y}$ 经过批归一化和 Sigmoid 函数激活得到特征向量 $\mathbf{y}'$ 。

- 4) 掩码 $\mathbf{b}$ 与特征向量 $\mathbf{y}'$ 的对应元素相乘得到加

权 $\mathcal{L}_1$ 距离特征向量 y'' ：

$$y'' = [b_1 y'_1, b_2 y'_2, \dots, b_n y'_n] \quad (7) \quad \text{函数输出预测值 } o_i。$$

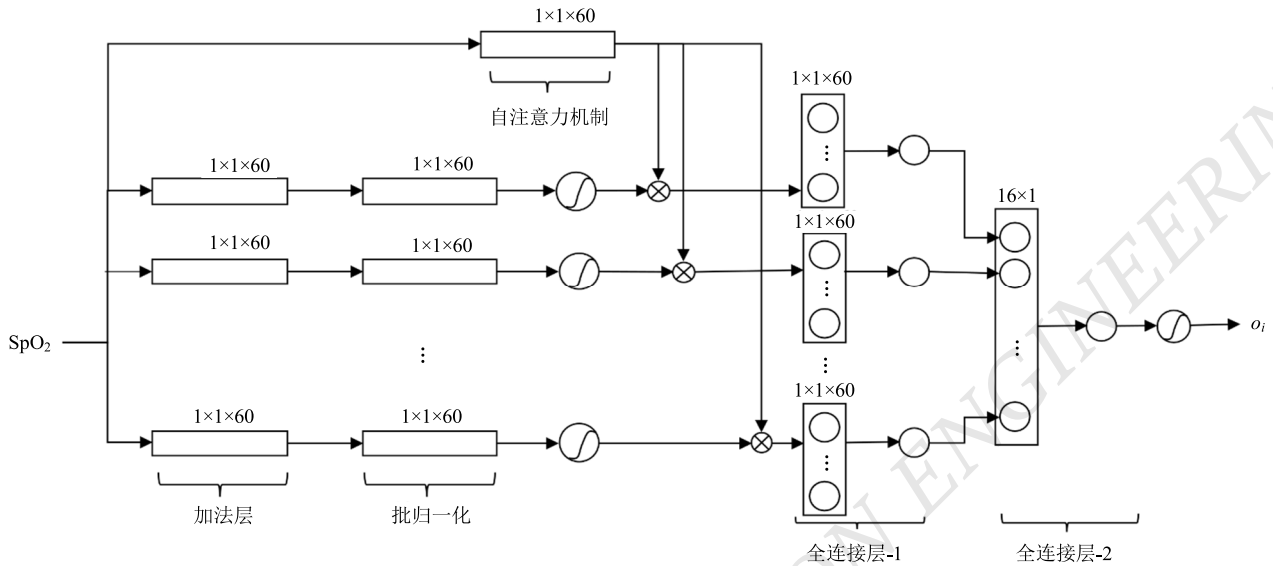


图 1 A-S-NN 网络结构图

2 数据处理

2.1 数据集

本文采用来自 Physionet 的公开数据集呼吸暂停-心电图数据库 (Apnea-ECG database, AED) 中的 SpO_2 数据进行实验^[18-19]。AED 中有 8 个含有 SpO_2 数据的记录，这些记录持续时间为 7~10h。医生给出每分钟的注释，若该分钟出现呼吸暂停则标注为 A，否则标注为 N。 SpO_2 的采样频率为 100 Hz。

2.2 数据预处理与数据划分

为适应网络输入，去除异常数据，对原始 SpO_2 数据进行预处理：

- 1) 将采样频率降为 1 Hz；
- 2) 将数据切为 1 min 的片段；
- 3) 剔除含有异常值的片段，即 SpO_2 低于 50% 的片段^[20]。

根据受试者不同，将 8 个记录划分为四折交叉验证对结果进行评估，具体划分如表 1 所示。

表 1 四折交叉验证的数据划分

序号	记录	N	A
1	a01, b01	2 790	959
2	a02, c01	2 760	989
3	a03, c02	2 751	998
4	a04, c03	2 946	803

3 实验结果与分析

3.1 实验环境

实验采用的深度学习框架为 Pytorch，服务器 CPU 为 Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz，内存为 16 GB。

3.2 参数设置

A-S-NN 网络的参数设置如表 2 所示。

表 2 A-S-NN 网络参数设置

网络结构	参数	数量/个
	核个数	16
加法层	核宽	60
	核步长	1
自注意力机制	窗宽	10

模型训练采用的优化器为 Adam, 迭代步数为 500, 学习率 (learning rate, LR) 为

$$LR = 0.005 \times [1 + \cos(epoch) / 100\pi] \quad (8)$$

式中, $epoch$ 为当前已迭代的步数。

3.3 评估指标

利用灵敏度 (sensitivity, Se)、特异性 (specificity, Sp), 平衡错误率 (balanced error rate, Ber) 来评估模型的检测性能。

$$Se = T_p / N_p \quad (9)$$

$$Sp = T_n / N_n \quad (10)$$

$$Ber = 1 - (Se + Sp) / 2 \quad (11)$$

式中, T_p 为正样本被预测为正样本的数量; N_p 为所有正样本的数量; T_n 为负样本被正确预测为负样本的数量; N_n 为所有负样本的数量。

同时, 使用浮点运算量 (amount of floating point arithmetics, FLOPs)、乘加次数 (amount of multiply-

adds, MAdd) 和内存使用来评估模型所需的算力及其他计算资源, 这 3 个指标均由 torchstat 工具包计算得出。

3.4 实验对比与分析

为验证本文采用的加法层及自注意力机制的有效性, 构建以下 4 种网络模型:

1) 卷积网络 (neural network based on convolution, C-NN);

2) 卷积及自注意力机制网络 (neural network based on CNN and Seft-attention, C-S-NN);

3) 加法网络 (neural network based on AdderNet, A-NN);

4) 基于加法网络和自注意力机制的网络 (neural network based on AdderNet and Seft-attention, A-S-NN)。

卷积层与加法层的参数设置一致, 如表 2 所示。4 种网络模型的 OSA 检测性能对比如表 3 所示, 其中 Se 、 Sp 、 Ber 为跨被试四折交叉验证的平均结果。

表 3 4 种网络模型的 OSA 检测性能对比

网络模型	$Se/\%$	$Sp/\%$	$Ber/\%$	浮点运算量	乘加次数	内存使用/kB
C-NN	93.57	96.93	4.75	4.80K FLOPs	7.68K MAdd	19.24
A-NN	93.73	97.01	4.63	2.88K FLOPs	5.76K MAdd	15.13
C-S-NN	94.24	97.18	4.29	4.80K FLOPs	8.27K MAdd	19.24
A-S-NN	94.56	96.91	4.27	2.88K FLOPs	6.35K MAdd	15.13

由表 3 可以看出:

1) C-S-NN 和 A-S-NN 的检测性能分别高于 C-NN 和 A-NN, 说明自注意力机制能有效减小运动伪迹的干扰, 提高检测性能;

2) A-S-NN 与 A-NN 的浮点运算量、乘加次数及内存使用均小于 C-S-NN 与 C-NN, 说明加法层可有效减少计算量;

3) A-S-NN 与 A-NN 的浮点运算量和内存使用相等, 表明自注意力机制几乎不增加模型的运算复杂度。

A-S-NN 与其他 OSA 检测方法的性能对比如表 4 所示。

表 4 A-S-NN 与其他 OSA 检测方法的性能对比

方法	是否跨被试测试	$Se/\%$	$Sp/\%$
ANN	否	96.50	98.50
SVM	是	84.57	97.28
CNN1D	是	92.04	95.78
A-S-NN	是	94.56	96.91

由表 4 可以看出:

1) ANN 方法取得最好的检测性能, 但其采用非跨被试方法, 即被测者的数据同时存在于训练集与测试集, 但在实际检测时, 需要先采集新患者数据并重新学习模型参数, 显然是不合理的;

2) A-S-NN、CNN1D 取得了比 SVM 更好的检测性能, 说明深度学习方法更能从数据中学习得到有效信息;

3) 与 CNN1D 相比, A-S-NN 的 Se 提高了 2.7%, Sp 提高了 1.2%。

4 结语

为减少 SpO_2 中的运动伪迹干扰, 提高 OSA 检测性能, 同时降低模型运算复杂度, 本文提出一种基于加法网络和自注意力机制的 OSA 检测方法, 并在公开数据集 AED 上获得较高的 OSA 检测精度 (Se : 94.56%, Sp : 96.91%)。首先, 与 CNN1D 相比, Se 、 Sp 分别提高了 2.7%和 1.2%; 然后, 与 A-NN 相比, Se 提高了 0.9%, 说明应用自注意力机制可有效提高模型的检测性能; 最后, A-S-NN 的浮点运算量仅为 2.88K FLOPs, 比 C-S-NN 小了约 40%, 说明加法层可有效降低计算复杂度。因此, 本文提出的 A-S-NN 能够在保持较低计算复杂度的同时, 实现较高的 OSA 检测性能, 适合部署在算力及其他计算资源配置较低的设备上。

参考文献

- [1] MOSTAFA S S, MENDONÇA F, RAVELO-GARCIA A G, et al. Multi-objective hyperparameter optimization of convolutional neural network for obstructive sleep apnea detection[J]. IEEE Access, 2020, 8: 129586-129599.
- [2] GROTE L, PLOCH T, HEITMANN J, et al. Sleep-related breathing disorder is an independent risk factor for systemic hypertension[J]. Journal of Hypertension, 2000, 18(6): 679-685.
- [3] MOOE T, FRANKLIN K A, HOLMSTROM K, et al. Sleep-disordered breathing and coronary artery disease: long-term prognosis[J]. American journal of Respiratory and Critical Care Medicine, 2001, 164(10): 1910-1913.
- [4] MOHSENIN V. Sleep-related breathing disorders and risk of stroke[J]. Stroke, 2001, 32(6): 1271-1276.
- [5] BENJAFIELD A, AYAS N, EASTWOOD P, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis[J]. The Lancet Respiratory Medicine, 2019, 7(8): 687-698.
- [6] QUAN S F, GILLIN J C, LITTNER M R, et al. Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research. editorials[J]. Sleep (New York, NY), 1999, 22(5): 667-689.
- [7] BURGOS A, GONI A, ILLARRAMENDI A, et al. Real-time detection of apneas on a PDA[J]. IEEE Transactions on Information Technology in Biomedicine, 2009, 14(4): 995-1002.
- [8] MOSTAFA S S, CARVALHO J P, MORGADO-DIAS F, et al. Optimization of sleep apnea detection using SpO_2 and ANN[C]//2017 XXVI International Conference On Information, Communication and Automation Technologies (ICAT). IEEE, 2017: 1-6.
- [9] MOSTAFA S S, MORGADO-DIAS F, RAVELO-GARCÍA A G. Comparison of SFS and mRMR for oximetry feature selection in obstructive sleep apnea detection[J]. Neural Computing and Applications, 2018: 1-21.
- [10] BEHAR J A, PALMIUS N, LI Q, et al. Feasibility of single channel oximetry for mass screening of obstructive sleep apnea[J]. EClinicalMedicine, 2019, 11: 81-88.
- [11] CHAW H T, KAMOLPHIWONG S, WONGSRITRANG K. Sleep apnea detection using deep learning[J]. Tehnički Glasnik, 2019, 13(4): 261-266.
- [12] VAQUERIZO-VILLAR F, ALVAREZ D, KHEIRANDIS H -GOZAL L, et al. A convolutional neural network architecture to enhance oximetry ability to diagnose pediatric obstructive sleep apnea[J]. IEEE Journal of Biomedical and Health Informatics, 2021 (99): 1-11.
- [13] 闫佳运. 可穿戴式睡眠监护设备中若干关键技术研究[D]. 北京: 北京工业大学, 2016.
- [14] CHEN H, WANG Y, XU C, et al. AdderNet: Do we really need multiplications in deep learning?[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1468-1477.
- [15] YAN L, DODIER R H, MOZER M, et al. Optimizing classifier performance via an approximation to the Wilcoxon-Mann-Whitney statistic[C]//Proceedings of the 20th International Conference on Machine Learning (ICML- 03), Washington DC, 2003: 848-855.
- [16] RUOPP M D, PERKINS N J, WHITCOMB B W, et al. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection[J]. Biometrical Journal: Journal of Mathematical Methods in Biosciences, 2008, 50(3): 419-430.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv preprint arXiv:1706.03762, 2017.

- [18] GOLDBERGER A L, AMARAL L A N, GLASS L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals[J]. Circulation, 2000,101(23):215-220.
- [19] PENZEL T, MOODY G B, MARK R G, et al. The apnea-ECG database[C]//Computers in Cardiology 2000, IEEE, 2000: 255-258.
- [20] XIE B, MINN H. Real-time sleep apnea detection by classifier combination[J]. IEEE Transactions on Information Technology in Biomedicine, 2012,16(3):469-477.

OSA Detection Method Based on Additive Network and Self-attention Mechanism

Zheng Yachun Zheng Heyu Lin Meina

(School of Automation, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: In order to achieve obstructive sleep apnea (OSA) detection with high accuracy and low computational complexity, according to blood oxygen saturation (SpO_2) proposes an OSA detection method based on additive network and self-attention mechanism. In this method, additive network is introduced for SpO_2 feature extraction to reduce the computational complexity; self-attention mechanism is used to adjust the feature weight between sampling points to improve the detection accuracy. The detection performance of sensitivity 94.56%, specificity 96.91% and floating-point operation are realized on data set Apnea-ECG. The number is 2.88 K flops, which is 40% less than the convolution operation. The experimental results show that this method can obtain high detection accuracy and maintain low computational complexity.

Key words: oxygen saturation; obstructive sleep apnea detection; AdderNet; Self-attention mechanism

作者简介:

郑雅纯, 女, 1996 年生, 硕士研究生, 主要研究方向: 模式识别, 机器学习, 生物信号处理。E-mail: zheng_yachun@163.com

郑和裕, 男, 1996 年生, 硕士研究生, 主要研究方向: 模式识别, 机器学习, 生物信号处理。E-mail: zheng_hy1209@qq.com

林美娜, 女, 1997 年生, 硕士研究生, 主要研究方向: 模式识别, 生物信号处理。E-mail: meina.lin@mail.gdut.edu.cn

(上接第 35 页)

Abstract: Aiming at the problem of low accuracy of bayberry identification and location in orchard environment due to complex illumination and different fruit states, a method of bayberry fruit identification and location based on local sliding window technology is proposed. Firstly, adaptive histogram equalization with limited contrast is used to improve the image illumination; Secondly, the CbCr color difference method was proposed, which combined with maximum inter class variance method, morphological operation and regional growth strategy to extract the potential region of fruit; Thirdly, according to the foreground region extraction results, the local sliding window technology is used to obtain the area to be detected; Finally, a support vector machine based on local texture features is established, combined with non maximum suppression to filter out the interference such as low matching background, so as to improve the accuracy of bayberry fruit identification and location. Experiments show that this method can accurately identify and locate bayberry fruit in orchard environment. The accuracy and recall are 92.51% and 90.82% respectively. The overall time-consuming of the algorithm is only 0.123 s, taking into account the real-time and accuracy.

Key words: orchard natural environment; identification and localization of bayberry; CLAHE; local sliding window technique; textural features; SVM

作者简介:

吴亮生, 男, 1988 年生, 大学本科, 工程师, 主要研究方向: 机器视觉、人工智能。

雷欢, 男, 1987 年生, 硕士研究生, 工程师, 主要研究方向: 机器视觉、人工智能。

陈再励(通信作者), 男, 1989 年生, 硕士研究生, 工程师, 主要研究方向: 机器视觉、人工智能。E-mail: zl.chen@giim.ac.cn

马敬奇, 男, 1988 年生, 硕士研究生, 工程师, 主要研究方向: 机器视觉、人工智能。