

基于光电容积脉搏波的心房颤动识别方法

张悦¹, 陈真诚¹, 梁永波², 朱健铭²

1. 桂林电子科技大学电子工程与自动化学院, 广西 桂林 541004; 2. 桂林电子科技大学生命与环境科学学院, 广西 桂林 541004

【摘要】为了能够简便快捷地检测心房颤动,提出一种基于光电容积脉搏波描记法(PPG)对心房颤动进行识别的方法。首先,将已确诊为心房颤动状态脉搏波与健康状态脉搏波数据进行对比分析;其次,基于分析结果,从脉搏波数据中提取与心房颤动相关的6类特征参数作为分类器的输入;最后,使用支持向量机(SVM)、BP神经网络、随机森林算法3种分类器建立心房颤动识别模型,其识别心房颤动的准确率分别可达89.1%、92.3%、95.2%。实验结果表明,基于PPG的心房颤动识别方法具有很高的识别准确率,尤其在使用随机森林算法作为分类器时,识别准确率达到最优。同时该检测方法简便快捷,是一种可以替代传统心电图检测识别心房颤动的方法,对心房颤动患者的长期观察监测具有临床价值。

【关键词】心房颤动;光电容积脉搏波;BP神经网络;随机森林;支持向量机

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2020)11-1416-05

Atrial fibrillation identification based on photoplethysmography

ZHANG Yue¹, CHEN Zhencheng¹, LIANG Yongbo², ZHU Jianming²

1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; 2. School of Life and Environmental Sciences, Guilin University of Electronic Technology, Guilin 541004, China

Abstract: A photoplethysmography (PPG)-based method for identifying atrial fibrillation is proposed to detect atrial fibrillation easily and quickly. Firstly, the pulse waves of atrial fibrillation and those in health status are compared and analyzed. Secondly, based on the analysis results, 6 types of characteristic parameters related to atrial fibrillation are extracted from PPG data as the input of classifiers. Finally, 3 classifiers, namely support vector machine, BP neural network and random forest algorithm, are used to establish the model for atrial fibrillation identification. The identification accuracies of the 3 models are 89.1%, 92.3% and 95.2%, respectively. The experimental results show that the atrial fibrillation identification method based on PPG has a high accuracy, especially when using random forest algorithm as the classifier. Meanwhile, the proposed detection method which is more convenient and rapid can replace the traditional ECG detection to identify atrial fibrillation, having clinical value for the long-term observation and monitoring of patients with atrial fibrillation.

Keywords: atrial fibrillation; photoplethysmography; BP neural network; random forest; support vector machine

前言

心房颤动(简称房颤)是一种常见的心脏疾病。房颤患者的体征包括脉律不齐、脉搏短绌、第一心音强弱不等、节律绝对不规整等^[1]。在传统的房颤诊断中,医生通过查看心电图(Electrocardiography, ECG)判断房颤的发生,但此诊断方式带有主观性,且耗时

耗力。Kora等^[2]基于ECG使用Cuckoo Search算法来实现自动识别房颤;王凯等^[3]基于ECG的自动心律失常检测模型能够有效监测心脏异常房颤信号的ECG信号特征。对于房颤监测的困难主要在于阵发性房颤,此类房颤的特点是突发突止、持续时间短,24 h动态心电图常作为首选的监测方法。李芳等^[4]通过分析总结房颤的心电散点图特征,分析24 h动态心电图来诊断房颤。武求花等^[5]运用Lorenz-RR散点图(Lorenz散点图)与时间-RR间期散点图(时间散点图)及逆向技术相结合的方法,快速识别不同发作时间的PAF,为临床提供更快更准确的心电诊断技术。

目前国内外对于房颤识别的研究大多基于

【收稿日期】2020-04-27

【基金项目】国家基金委重大仪器研制项目(61627807);国家自然科学基金(81873913);广西创新研究团队项目(2017GXNSF-GA198005)

【作者简介】张悦,硕士,研究方向:生物传感与仪器, E-mail: 3066952-98@qq.com

ECG,但ECG设备普遍有操作复杂、携带困难、价格较高、多通道等局限性。光电容积脉搏波描记法(PPG)是使用激光对待测流体或固体的体积进行检测与记录的方法,可以用来监测毛细血管中血液的体积变化^[6]。有研究表明,PPG信号与ECG信号有很高的相关性^[7]。PPG方法采集信息容易、方便携带,相较于ECG方法,PPG更加擅长长期监测。本文从脉搏波波形中提取房颤相关特征参数,使用机器学习算法构建心房颤动识别模型,以探寻基于脉搏波识别房颤的有效方法。

1 房颤状态脉搏波分析

将心房颤动状态脉搏波与健康状态脉搏波的波形进行对比分析,可发现二者间有明显不同。主要区别在于房颤发生时脉搏波的脉冲宽度(Pulse-Width, PW)、峰峰间隔(Peak-to-Peak Interval, PPI)、收缩期面积(Systolic Area, SA)会明显与相邻时间下同类特征的平均值有较大差别。其中每次房颤发生时,PPI会出现一次低于、一次高于平均PPI值的情况;PW和SA都会出现一次明显低于平均值的情况。

通过查阅文献,本文还选取了增强指数(Augmentation Index, AI)、脉搏波上升导数最大处至重搏波峰值的时间间隔(Half Rise to Dicrotic Notch, HRDN)、拐点面积比(Inflection Point Area, IPA)3种与房颤疾病相关的特征类型,这3类特征参数能作为受试者是否可能发生房颤的判断因素。

心房颤动的发生与高血压高度相关,心脏由于人体长期血压过高将造成心房电传导延迟、左心室肥厚、左心房功能异常、左心房扩大等现象^[8]。血压的增高会导致动脉的老化进程加速^[9]。动脉老化将引起动脉的僵硬度提升,脉搏波传导速度会随着动脉僵硬度的提升而加快,它的加速使得外周动脉到近端大动脉的反射波提前回传,最终使脉搏波波形变化^[10]。AI可以定义为 PT_2/PT_1 ,其中 PT_2 是收缩压后期振幅, PT_1 是收缩压早期振幅。已有研究表明,脉搏波特征中的增强指数AI就可以反映出动脉僵硬度的变化,且容积脉搏波的AI与主动脉AI相关^[11],故选用此类特征值。

有研究表明,脉压差的增大,将导致动脉牵拉的加重,血管壁弹性成分容易疲劳和断裂,这将造成内膜损伤,并很有可能进一步提高动脉僵硬和血管栓塞发生的可能性。内皮功能减退可能导致房颤的发生^[12]。因此,将能够反映脉压的HRDN列为房颤的判断因素之一^[13]。

房颤病人与正常人的外周血管总阻力(Total

Peripheral Resistance, TPR)有明显的区别^[14]。有研究发现IPA能够反映TPR指标。IPA定义为脉搏波中第二个峰与第一个峰的面积之比(SB/SA),与脉搏波反射强度有较强的相关性。动脉系统中,各部分间阻抗不匹配将出现脉搏波反射。已知,有九成左右的TPR存在于小动脉处,小动脉收缩将使得TPR变化,并造成阻抗不匹配,由此引发的脉搏波反射强度的变化将使IPA发生变化^[15]。本文所提房颤相关的各特征参数示意如图1所示。

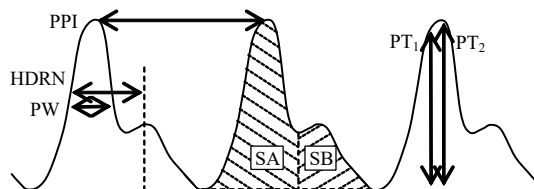


图1 脉搏波的特征参数

Fig.1 Characteristic parameters of pulse wave

2 实验过程

实验数据来源于美国麻省理工学院MIMIC公开数据库中房颤患者(通过ECG确诊)以及健康人的PPG数据。数据经筛选、分段、归一化等预处理,得到6 857例10 s脉搏波数据段,数据样本中房颤状态PPG样本共2 737例,非房颤状态PPG样本共4 120例。将样本数据分成房颤与非房颤两组并标记标签,提取PPG数据中与房颤相关的6类特征参数,使用BP神经网络、随机森林、支持向量机(SVM)作为分类器对数据进行房颤的识别分类。通过对比分析各个分类器的实验结果,探寻效果更佳的识别房颤方法。

2.1 数据处理

MIMIC数据库中下载的脉搏波数据经预处理后,得到归一化的PPG数据可视化样例(每段10 s,采样频率为125 Hz)如图2所示。将房颤状态的脉搏波数据标记为正类,非房颤状态的脉搏波数据标记为负类。

提取PPG波形中PW、PPI、SA、AI、HRDN、IPA共6类特征作为分类器的输入项进行房颤识别分类实验。

2.2 模型建立

本文使用BP神经网络算法、SVM算法、随机森林算法,以PW、PPI等6类特征参数作为输入,样本是否发生房颤现象作为输出,分别构建3种分类器,并将对分类效果进行评估。

BP神经网络属于多层前馈的神经网络,其含有

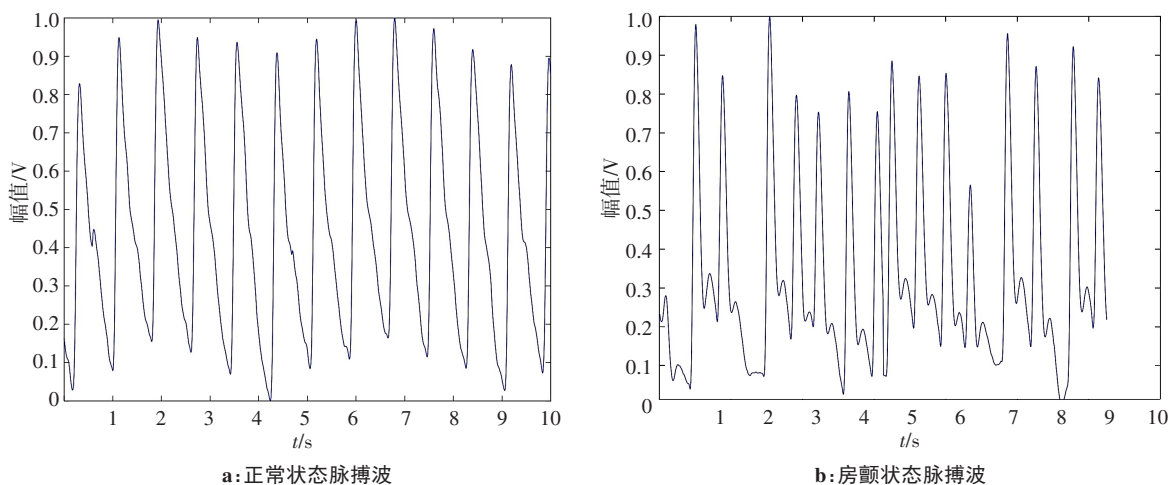


图2 不同状态下的脉搏波波形

Fig.2 Pulse waveforms in different statuses

输入层、一层及以上的隐含层和输出层。BP神经网络的输入信号前向传递,误差反向传播,若输出层无法得到期望输出,则转向反向传播,通过预测误差对网络权值和阈值进行修正,可以使预测输出向期望输出持续逼近^[16]。隐含层各神经元节点采用Sigmoid函数作为激励函数。实验中,将BP神经网络的隐含层设为两层,第一层含100个神经元,第二层含50个神经元;最大迭代次数epochs设为20 000次时,BP神经网络对于房颤脉搏波的识别率达到92.3%。

SVM是一种常用的机器学习方法,在统计样本少的情况下,能够处理二分类的问题。SVM的原理是找到一个最优分类超平面,其在能够达到分类要求和精度的基础上,还能够最大化分类面异侧的空白区域,可以完成线性可分数据的最优分类。

对于本文中关于发生房颤与否的二分类问题, $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ 是本次实验所使用的样本集。 $x_i \in RP, y_i \in \{-1, 1\}, i = 1, 2, \dots, n$; P 是各样本的维数; n 是各样本的个数。超平面方程是:

$$\omega \cdot x + b = 0 \quad (1)$$

其中, ω 是权向量, b 是阈值。经过归一化,若满足 $y_i(\omega \cdot x_i + b) \geq 1 (i = 1, 2, \dots, n)$,即可得出这个超平面对样本的分类已经完成^[17]。本文选用高斯径向基作为内核,其计算复杂度低,且参数设置简单,广泛运用于各个领域的研究。本实验中,SVM对于房颤识别分类的准确率达到89.1%。

将Bagging集成构建于决策树基学习器上,并在决策树训练中加入随机属性的选择,即可构成随机森林。决策树模仿树的结构进行判断决策,属于较普遍的机器学习算法。生成一棵决策树,包括决策

树生成、剪枝、特征选择三大内容。特征选择所表达的分裂标准将在很大程度上控制决策树的泛化误差,它是非常关键的一个部分。Bagging基于自助采样法,是一种经典的并行式集成学习方法。Bagging在训练集合中,将样本有放回地随机抽取,每一轮的分训练集的样本,都是从训练集中抽取的 N 个样本。一轮训练集中,某一训练样本被抽取0至 N 次。结果集的产生需要将每次抽取出的子集作为算法的输入,产生预测函数,再需要经过 T 轮循环。最后通过投票,解决预测函数的分类问题^[18]。随机森林算法的主要参数设置中,在考虑模型的准确度与稳定度时,还需兼顾模型效率,我们将树的个数 $n_estimators$ 设置为100;叶子节点上应有的最少样例数 $min_samples_leaf$ 设置为4;分裂内部节点需要的最少样例数 $min_samples_split$ 设置为8时,随机森林识别房颤的分类准确率达到95.2%。

2.3 模型评估

数据集中有70%的样本作为训练集;15%的样本作为验证集;15%的样本作为测试集。训练集作为构建预测模型的基础,训练样本的数量将影响机器学习算法模型的预测能力。测试集作为检验模型性能的分析样本。

本文选用准确率、灵敏度、召回率、ROC曲线等指标,作为分类器的性能指标进行评估,其中ROC曲线在二分类问题中是一种独立于分类器种类的分析方法。

混淆矩阵中,TP(True Positives)是将房颤脉搏波判定为房颤;FP(False Positives)是把非房颤脉搏波判定为房颤;FN(False Negatives)是把房颤脉搏波判定为非房颤;TN(True Negatives)是非房颤脉搏波判定为非房颤。

通过以下公式进行性能度量:

准确率: $Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$ (2)

精准率: $Precision = \frac{TP}{TP + FP}$ (3)

召回率: $Recall = \frac{TP}{TP + FN}$ (4)

在评价分类算法的指标中,准确率是最常见的衡量指标之一,分类器的准确率越高即表明该分类算法的效果越好。但由于本文的数据源中,房颤数据与非房颤数据的样本不平衡,还需引入精准率和召回率两个指标。如上文公式所示,精准率代表所有预测为正类的结果中,真正正类的比例。召回率代表所有真实正类中,预测正确的比例。同样,精准率和召回率的值越高,越能够表明该分类算法优质。本文还引入了F-Score指标,它将精准率和召回率进行加权调和平均,能够为精准率和召回率提供一个综合的评判标准^[19]。

$F-Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$ (5)

3 结果分析

基于实验结果及评价指标可得如表1所示的结果,随机森林模型的准确率最高,达到了95.2%,BP神经网络的准确率为92.3%,SVM方法的准确率最低只有89.1%。通过对比能够反映综合评价指标的F-Score可以看出,随机森林的F-Score为93.9%,高于另外两个分类模型。且随机森林模型的各项指标与其他模型相比,均保持了较高的水平。

表1 各分类器的性能指标

Tab.1 Performance indicators of each classifier

| 方法 | 准确率 | 精准率 | 召回率 | F-Score | AUC |
|--------|-------|-------|-------|---------|-------|
| 随机森林 | 0.952 | 0.947 | 0.931 | 0.939 | 0.986 |
| 支持向量机 | 0.891 | 0.916 | 0.792 | 0.849 | 0.938 |
| BP神经网络 | 0.923 | 0.895 | 0.922 | 0.908 | 0.966 |

在ROC空间中,ROC曲线越凸向左上方,表明该算法所能达到的分类效果越好;ROC曲线越靠近对角线,则分类器越趋向于随机分类器。笔者将3种分类器的分类效果进行比较,结果如图3所示。

Area Under the ROC Curve(AUC)能够准确反映分类器的性能,AUC值越大,则此分类器的性能越好。如图3所示,灰色虚线代表AUC=0.5,SVM的分类结果更靠近虚线,其AUC=0.938,随机森林的ROC曲线最接近左上角即(0,1)点的位置,其AUC值达到

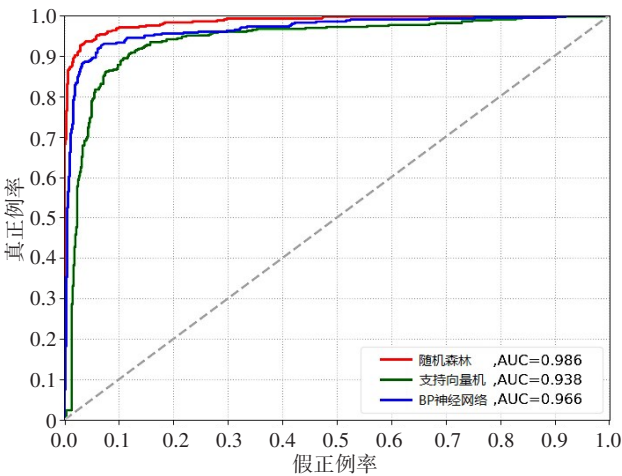


图3 接收者操作特征曲线
Fig.3 Receiver operating characteristic curve

了0.986。综合以上评价分析结果可得,随机森林模型对于房颤状态脉搏波的识别分类效果优于SVM和BP神经网络模型。

4 结论

本次研究结果表明,PPG信号可用于判定房颤的发生,是一种可行的替代ECG检测房颤的方法;基于本文选用的6类脉搏波特征参数,使用机器学习算法对脉搏波进行分类,能够有效区别出脉搏波的房颤状态与非房颤状态。本文还通过准确率、精准率、召回率、F-Score、AUC值等评价指标评估模型效果,结果表明,使用随机森林分类方法判别房颤发生的效果优于SVM和BP神经网络,其准确率达到了95.2%。

下一步工作中,可基于目前的研究,分析更多的特征点,尝试继续提升分类效果。由于本文所使用的临床数据来源于MIMIC数据库,缺乏相关校准记录等信息,之后可使用精确的传感器自主获取数据,以期得到更好的房颤识别分类效果。

【参考文献】

[1] 黄从新,张澍,黄德嘉,等. 心房颤动:目前的认识和治疗的建议-2018[J]. 中国心脏起搏与心电生理杂志, 2018, 32(4): 315-368.
HUANG C X, ZHANG H, HUANG D J, et al. Atrial fibrillation: current knowledge and treatment recommendations-2018[J]. Chinese Journal of Cardiac Pacing and Electrophysiology, 2018, 32(4): 315-368.

[2] KORA P, AYYEM V. ECG based atrial fibrillation detection using cuckoo search algorithm[J]. IJCA, 2017, 162(10): 37-42.

[3] 王凯,杨枢,李超. 一种基于ECG的多层共轭对称Hadamard特征变换的房颤异常信号分类方法[J]. 中国医学物理学杂志, 2019, 36(9): 1068-1073.
WANG K, YANG S, LI C. Classification method of atrial fibrillation abnormal signals based on ECG multi-layer conjugate symmetrical

- Hadamard feature transform [J]. Chinese Journal of Medical Physics, 2019, 36(9): 1068-1073.
- [4] 李芳, 戴淑婷, 韩敏, 等. 利用心电散点图快速识别心房颤动和心房扑动[J]. 临床医药实践, 2019, 28(8): 598-600.
- LI F, DAI S T, HAN M, et al. Quickly identify atrial fibrillation and atrial flutter using ECG scatter plots [J]. Clinical Medicine Practice, 2019, 28(8): 598-600.
- [5] 武求花, 龚春, 武洪林, 等. 心电散点图及逆向技术对阵发性房颤的快速诊断[J]. 中国老年学杂志, 2019, 39(2): 263-267.
- WU Q H, GONG C, WU H L, et al. Rapid diagnosis of paroxysmal atrial fibrillation by ECG scatter plot and reverse technique [J]. Chinese Journal of Gerontology, 2019, 39(2): 263-267.
- [6] 张天森, 朱兰·I·谢尔盖. 光体积描记术原理及应用[J]. 中国医疗器械杂志, 2016, 40(3): 186-190.
- ZHANG T M, ZHU LAN·I·SERGEY. Principle and application of photoplethysmography [J]. Chinese Journal of Medical Instrumentation, 2016, 40(3): 186-190.
- [7] 陈真诚, 于邦文, 梁永波, 等. 基于光电容积脉搏波信号与心电信号相关性的研究[J]. 生物医学工程研究, 2019, 38(3): 370-375.
- CHEN Z C, YU B W, LIANG Y B, et al. Research on the correlation between photoelectric volume pulse wave signal and ECG signal [J]. Research in Biomedical Engineering, 2019, 38(3): 370-375.
- [8] 李妍妍. 老年高血压患者房颤与脉搏波传导速度的关系[J]. 医学信息(上旬刊), 2010, 23(9): 3136-3137.
- LI Y Y. Relationship between atrial fibrillation and pulse wave velocity in elderly hypertensive patients [J]. Medical Information (1st Issue), 2010, 23(9): 3136-3137.
- [9] 施晓芸. 高血压加速并加重主动脉老化、钙化的发生发展过程及其可能机制[D]. 福州: 福建医科大学, 2016.
- SHI X Y. Hypertension accelerates and aggravates the occurrence and development of aortic aging and calcification and its possible mechanism [D]. Fuzhou: Fujian Medical University, 2016.
- [10] 王俊舟, 刘凯, 李龙心, 等. 长期持续性心房颤动患者静息心率控制与脉搏波传导速度的关系[J]. 华西医学, 2012, 27(11): 1606-1609.
- WANG J Z, LIU K, LI L X, et al. Relationship between resting heart rate control and pulse wave velocity in patients with long-term persistent atrial fibrillation [J]. West China Medical Journal, 2012, 27(11): 1606-1609.
- [11] 刘文彦, 王璐, 姚阳, 等. 脉搏波传导速度与反射波增强指数相关性分析[J]. 中国医疗设备, 2018, 33(8): 21-24.
- LIU W Y, WANG L, YAO Y, et al. Correlation analysis of pulse wave velocity and reflected wave enhancement index [J]. China Medical Equipment, 2018, 33(8): 21-24.
- [12] 李海珍. 高血压患者脉压与房颤的关系[J]. 中国卫生标准管理, 2015, 6(2): 11-12.
- LI H Z. The relationship between pulse pressure and atrial fibrillation in patients with hypertension [J]. Chinese Health Standard Management, 2015, 6(2): 11-12.
- [13] COX P, MADSEN C, RYAN K L, et al. Investigation of Photoplethysmogram morphology for the detection of hypovolemic states [C]. Conference Proceedings: Annual International Conference of the IEEE EMBS. 2008: 5486-5489.
- [14] 孙水英, 缪澄宇, 吉建江, 等. 风湿性心脏病并发房颤的血液动力学改变[J]. 青海医药杂志, 1992(1): 14-16.
- SUN S Y, MIAO C Y, JI J J, et al. Hemodynamic changes of rheumatic heart disease complicated by atrial fibrillation [J]. Qinghai Medical Journal, 1992(1): 14-16.
- [15] WANG L, PICKWELL-MACPHERSON E, LIANG Y P, et al. Noninvasive cardiac output estimation using a novel photoplethysmogram index [C]. Conference Proceedings: Annual International Conference of the IEEE EMBS. 2009: 1746-1749.
- [16] 宋世坤, 王书海. 基于BP神经网络的雪花形状分类研究[J]. 无线互联科技, 2018, 15(18): 109-110.
- SONG S K, WANG S H. Research on snowflake shape classification based on BP neural network [J]. Wireless Internet Technology, 2018, 15(18): 109-110.
- [17] 梁丽军, 刘子先, 王化强. 基于弹性网-SVM的疾病诊断关键特征识别[J]. 计算机应用研究, 2015, 32(5): 1301-1304.
- LIANG L J, LIU Z X, WANG H Q. Key features identification of disease diagnosis based on elastic net-SVM [J]. Journal of Computer Applications, 2015, 32(5): 1301-1304.
- [18] 吕红燕, 冯倩. 随机森林算法研究综述[J]. 河北省科学院学报, 2019, 36(3): 37-41.
- LÜ H Y, FENG Q. Review of random forest algorithm research [J]. Journal of the Hebei Academy of Sciences, 2019, 36(3): 37-41.
- [19] 谢炜芳. 基于机器视觉的金针菇分类识别方法研究[D]. 泉州: 华侨大学, 2019.
- XIE W F. Research on classification and recognition method of Flammulina velutipes based on machine vision [D]. Quanzhou: Huaqiao University, 2019.

(编辑: 陈丽霞)