

基于改进反向传播算法的声音识别及健康检测技术

田昊旻¹, 马祎航²

(1. 中国医科大学附属盛京医院, 沈阳 110141;

2. 沈阳工学院 机械工程与自动化学院, 沈阳 113122)

摘要: 随着计算机技术的发展, 声音识别与健康检测成为现代医学诊断的重要手段之一; 通过对新生儿声音的分析, 可以早期发现和诊断多种健康问题; 研究提出一种基于改进反向传播神经网络的声音识别模型, 通过声音实现对新生儿的健康状况分析; 该模型通过小波变换对声音数据进行预处理, 随后结合粒子群优化算法和反向传播神经网络设计检测模型; 通过引入粒子群优化算法对反向传播算法进行改进, 提高了模型的局部搜索能力和收敛速度; 实验结果表明, 在数据集为 1 000 时, 小波去噪模型的信噪比为 0.97, 结构信息损失率为 0.18, 交并比为 0.96; 针对不同类型的声音, 改进反向传播神经网络模型识别的准确率分别为 0.87、0.83、0.97、0.88, 均方根误差值为 0.09、0.07、0.05、0.07; 结果表明, 所提出的声音识别与健康检测模型能够有效提高声音数据的识别精度和检测效率, 有助于新生儿健康状态的评估。

关键词: 健康检测; 声音识别; 粒子群优化算法; 小波去噪; 反向传播

Voice Recognition and Health Detection Technology Based on Improved Backpropagation Algorithm

TIAN Haomin¹, MA Yihang²

(1. China Medical University Affiliated with Shengjing Hospital, Shenyang 110141, China;

2. School of Mechanical Engineering and Automation, Shenyang Institute of Technology, Shenyang 113122, China)

Abstract: With the development of computer technology, voice recognition and health monitoring have become important means of modern medical diagnosis. By analyzing the sounds of newborn babies, various health problems can be detected and diagnosed early. A voice recognition model based on improved back propagation neural network is proposed, the health status of newborns is analyzed through sound. This model is used to preprocess sound data through wavelet transform, and then combined with particle swarm optimization algorithm and backpropagation neural network to design the detection model. The back propagation algorithm is improved by introducing particle swarm optimization algorithm, enhancing the local search ability and convergence speed of the model. Experimental results show that with a dataset of 1 000, the signal-to-noise ratio of the wavelet denoising model is 0.97, the structural information loss rate is 0.18, and the intersection to union ratio is 0.96. The recognition accuracy of the improved back propagation neural network model is 0.87, 0.83, 0.97, and 0.88 for different types of sounds, with the root mean square errors of 0.09, 0.07, 0.05, and 0.07, respectively. The results indicate that the proposed voice recognition and health detection model can effectively improve the recognition accuracy and detection efficiency of voice data, which is helpful for evaluating the health status of newborns.

Keywords: health testing; voice recognition; particle swarm optimization algorithm; wavelet denoising; back propagation

0 引言

在数字信号处理领域, 声音识别和健康检测技术得到了迅速的发展, 特别是在医疗保健和智能设备方面, 声音信号的处理与分析尤为重要。声音信号作为一种非稳态信号, 其特征复杂多变, 传统的信号识别方法如粒子群优化算法 (PSO, particle swarm optimization) 存在明显不足^[1-2]。文献 [3] 为了提升多飞行器编队执行任务的效率和性能, 提出了一种改进的粒子群优化算法。该算法考虑

了任务能力约束、威胁代价、航程代价和任务收益。研究表明, 所提算法能有效解决局部收敛问题, 显著提高任务分配的收益和效率^[3]。文献 [4] 为了优化永磁球形电机的驱动策略, 提出了一种改进的粒子群优化算法。该算法通过自适应动态惯性权重和自适应学习因子提高计算效率, 实时计算永磁球形电机期望转矩对应的驱动电流。研究表明, 改进的粒子群优化算法在保持精度的同时, 相比传统粒子群优化算法具有更快的计算速度, 并通过实验验证了其有效性^[4]。文献 [5] 为了优化连续纤维增强聚

收稿日期: 2024-07-20; 修回日期: 2024-08-08。

基金项目: 辽宁省科技计划联合计划 (2023JH2/101700066)。

作者简介: 田昊旻 (1986-), 女, 大学本科, 护师。

引用格式: 田昊旻, 马祎航. 基于改进反向传播算法的声音识别及健康检测技术[J]. 计算机测量与控制, 2024, 32(11): 87-94.

合物的结构设计和纤维分布以实现更好的物理性能,提出了一种基于粒子群优化算法与独立连续映射相结合的方法。研究表明,该方法能有效降低局部最优的可能性,实现具有清晰宏观拓扑结构和微观纤维分布的连续纤维增强聚合物结构设计,为碳纤维布结构设计和纤维铺设路径规划提供了指导^[5]。

许多学者在声音识别方面也有过研究,文献[6]等人了解决变压器与中压开关柜传统绝缘状态检测依赖人工的问题,提出了一种基于可听声声音识别的方法。通过构建声谱图数据集,结合改进的卷积神经网络和注意力机制,实现了对放电故障的在线检测。研究表明,所设计的网络模型识别准确率高达 99.2%^[6]。文献[7]为了检测供水管道泄漏,提出了一种基于卷积神经网络的供水管道泄漏声音识别方法。该方法通过实时采集声音信息、提取梅尔谱特征图并使用改进的 ShuffleNet V2 网络进行训练与识别。研究表明,该方法对供水管道泄漏声音识别效果最佳,参数量低,运算资源占用少^[7]。文献[8]为了改善脑卒中后构音障碍的评估效果,提出了一种基于梅尔频谱特征和一维卷积网络与双 LSTM 网络的混合识别模型。研究表明,该模型在音节级和说话人级的识别准确率分别为 82.1%和 97.4%,显著优于传统方法,有助于提高脑卒中后构音障碍的治疗效率和诊断准确性^[8]。

综上所述,许多学者对深度学习和声音识别进行了研究,但是在通过深度学习对声音进行识别时,并未注意到声音与健康的联系,且采用的都是基础算法,并未针对算法的缺陷对其进行改进。PSO 受到群体智能行为的启发,通过个体和群体的信息交流来寻找最优解,但在局部搜索能力上有所不足。BP 算法常用于训练神经网络,通过梯度下降法优化参数,能够有效提高模型精度。因此,为探究新生儿声音与健康之间的联系,研究提出了一种基于小波去噪和改进 BP 的声音识别与健康检测模型。首先,通过小波去噪对声音数据进行预处理;然后,引入 PSO 与反向传播算法(BP, back propagation)算法相结合,优化声音识别模型。研究的创新点在于小波去噪方法对声音数据进行预处理,并通过 PSO 算法结合 BP 算法优化模型参数,从而提高声音识别和健康检测的精度和效率。旨在实际应用中能够显著提高声音识别和健康检测的性能。

1 基于改进 BP 的声音识别与健康检测模型研究

1.1 基于小波去噪的声音数据预处理模型研究

对数字信号来说,采用时域和频域对信号进行分析,通过傅里叶变换,可以将两者联系起来。傅里叶变换是一种数学变换,用于将信号从时域转换到频域。该方法将原始信号表示为不同频率的正弦波的叠加,便于分析信号的频谱特性。傅里叶变换的基本思想是将一个复杂的信号分解成多个不同频率的正弦波的组合^[9-10]。傅里叶变换可以将一个时域信号转换为频域信号,表达为各个频率成分的叠加。通过傅里叶变换,可以从时域信号中提取频率信息,分析信号的频谱特性。傅里叶变换的逆变换则可以将频域

信号还原为时域信号,这使得频域分析的结果可以直接应用于实际信号处理。通过在频域对信号进行滤波,然后使用逆傅里叶变换将其转换回时域,可以实现对信号的噪声抑制和信号增强。傅里叶变换的表达式如式(1)所示:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (1)$$

式中, $X(f)$ 是频域信号的傅里叶变换结果,表示频率的幅值和相位信息。 $x(t)$ 表示时域信号, $e^{-j2\pi ft}$ 表示复指数函数,其中 j 为虚数单位。傅里叶逆变换用于将频域信号转换回时域信号,其表达式如式(2)所示:

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} df \quad (2)$$

式中, f 表示频率。虽然傅里叶变换提供了信号的全局频率信息,但丧失了时间局部化的信息。即傅里叶变换不能描述在信号的某一特定时间点上频率成分的变化,这对于分析非平稳信号时尤其不利。在对新生婴儿的声音进行识别时,其信号一般是非稳态信号,因此,采用傅里叶变换会有一定的局限性。研究选择短时傅里叶变换(STFT, short-time fourier transform)与小波变换(WT, wavelet transform)对傅里叶变换进行改进。STFT是一种解决傅里叶变换在处理非平稳信号时不足的方法,STFT通过将信号分割成多个短时间片段,并对每个片段进行傅里叶变换,从而提供信号的时间和频率局部化信息^[11-12]。其表达式如式(3)所示:

$$S(\omega, \tau) = \int_{-\infty}^{\infty} f(t)g^*(t - \tau)e^{-j\omega t} dt \quad (3)$$

式中, $f(t)$ 表示未进行处理的信号, $g(t)$ 表示窗函数,*表示共轭, τ 表示时刻, ω 表示频率, S 表示信号成分比重。但是在STFT中,时间分辨率和频率分辨率之间存在权衡。窗口越短,时间分辨率越高,但频率分辨率越低;反之亦然。若STFT使用固定长度的窗口函数,对于不同频率成分的适应性较差。两者无法共同满足实际要求。而小波变换能够很好地解决STFT的不足。小波变换是一种能够有效处理非平稳信号的方法,通过使用可变的窗口函数,提供了时间和频率的局部化分析^[13-14]。其核心思想是使用一组被称为小波的基函数对信号进行分解。小波函数具有局部化特性,即在时间和频率两个域中都具有有限的支撑,小波函数在时域上具有局部化特性,即它在时域上的能量集中在一个有限的区域内。使得小波函数在时域上是有限长度或快速衰减的。这种局部化特性使得小波变换能够很好地捕捉信号的局部特征,尤其适用于分析具有突变或不连续点的信号。小波函数在频域上也具有局部化特性,即在频域上的能量集中在某一频带内。小波变换通过对信号进行多尺度分析,可以同时提供信号在不同尺度下的局部特征。这种频域局部化特性使得小波变换能够有效地表示信号的频率成分,并且能够根据需要在不同的尺度上进行精细或粗略的分析^[15-16]。其表达式如式(4)所示:

$$\phi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (4)$$

式中, $\psi(t)$ 表示一个基本小波函数, a 表示缩放参数, 控制小波的伸缩。 b 是平移参数, 控制小波在时间轴上的位置。 对于一个连续时间信号, 其连续小波变换表达式如式 (5) 所示:

$$W(a, b) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) dt \quad (5)$$

式中, $W(a, b)$ 表示信号 $x(t)$ 在尺度和位置上的变换系数。 通过改变尺度和位置, 可以得到信号在不同尺度和位置上的局部特征, 在实际的应用中, 需要对小波函数进行离散化处理来减少计算量, 离散小波基函数表达式如式 (6) 所示:

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) \quad (6)$$

式中, j 表示尺度参数, k 表示平移参数, 2^{-j} 表示小波函数的缩放因子, 使得函数在时间轴上缩放。 Mallat 算法基于小波变换理论, 通过一系列的低通滤波器和高通滤波器, 将信号在不同尺度下进行分解, 从而实现信号的多尺度分析。 与傅里叶变换不同, 小波变换可以提供更好的时域和频域局部化, 因此在处理非平稳信号时具有优势^[17-18]。 Mallat 算法是离散小波变换的高效实现方法, 该方法利用滤波器组进行信号的逐级分解和重构, Mallat 算法分解和重构的示意图如图 1 所示。

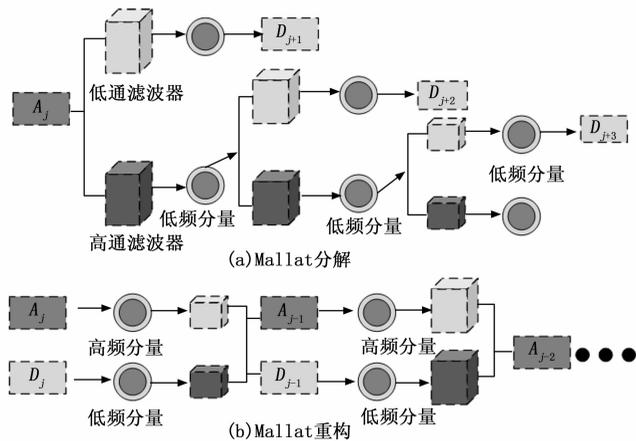


图 1 Mallat 算法分解和重构示意图

图 1 (a) 表示 Mallat 分解算法, 图 1 (b) 表示 Mallat 重构算法。 由图 1 (a) 可知, 首输入信息数据, 用低通滤波器和高通滤波器分别对信号进行卷积。 对卷积结果进行下采样, 得到第一级的低频分量和高频分量。 使用第一级分解得到的低频分量作为下一级的输入信号, 重复第一级分解的过程, 直到达到所需的分解层数^[19-20]。 由图 1 (b) 可知, 使用上采样和滤波器对高频和低频分量进行重构, 逐级将重构得到的信号和下一层的高频分量进行组合, 直到重构出原始信号。 以某噪声为例, 对该数据进行小波分解, 其过程如图 2 所示。

由图 2 可知, 多层小波分解是指将信号进行多次小波变换, 每次变换得到的低频部分再次进行小波变换, 这种多层次的分解可以提供更丰富的信号信息, 并允许对信号

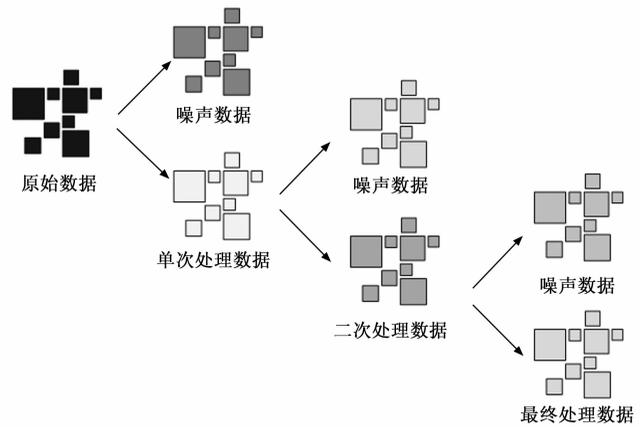


图 2 多层小波分解示意图

进行更细致的分析。 对于某信号的小波分解, 可以采用适合的分解层数。 综上所述, 小波阈值去噪的流程主要有三步, 分别是小波分解、阈值量化和信号重构。 将信号分解到小波域, 得到多尺度的近似系数和细节系数, 然后选择适当的阈值和阈值函数, 对细节系数进行阈值处理, 去除噪声。 最终将处理后的系数进行小波逆变换, 重构出去噪后的信号。

1.2 基于改进 BP 的声音识别检测模型设计

在数据预处理完成后, 将数据输入到设计的模型中, 对数据进行识别。 研究采用 BP 构建模型, 并引进 PSO 对模型进行改进。 PSO 是一种群体智能算法, 受到鸟群或鱼群等生物群体行为的启发而提出。 在 PSO 中, 通过模拟个体在搜索空间中的迭代移动, 寻找最优解^[21]。 这些粒子在搜索过程中通过个体最优和群体最优两个方面的信息交流来调整自身的移动方向和速度, 以达到优化目标。 具体的更新公式如式 (7) 所示:

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (p_{id}^k - X_{id}^k) \quad (7)$$

式中, V_{id} 表示粒子的速度, ω 表示惯性权重, c 表示加速度因子, k 表示迭代次数, r 为分布于 $[0, 1]$ 之间的随机数。 粒子位置的表达式如式 (8) 所示:

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (8)$$

式中, X_{id} 表示第 i 个粒子的位置。 PSO 的基本原理如下, 先随机生成一群粒子, 并给定每个粒子的位置和速度。 然后计算每个粒子的适应度, 即每个粒子的目标函数值。 然后对于每个粒子, 更新其个体最优解, 根据所有粒子的个体最优解, 更新全局最优解, 接着根据个体最优和群体最优, 更新粒子的速度和位置, 直到满足条件为止^[22-23]。 在此过程中, 通常会使用惯性权重来平衡局部搜索和全局搜索的影响, 其表达式如式 (9) 所示:

$$\omega(t+1) = \omega_{\max} - \frac{k}{K} \cdot (\omega_{\max} - \omega_{\min}) \quad (9)$$

式中, ω_{\max} 表示惯性权重的最大值, ω_{\min} 表示惯性权重的最小值。 K 表示总迭代次数。 PSO 算法在全局搜索方面表现优秀, 但在局部搜索能力方面表现不足, 且 PSO 算法需要较多的迭代次数才能达到收敛。 因此引入 BP 算法相结合, 利

用梯度信息对参数进行更新，可以更快地调整粒子的位置，使得算法更快地收敛到最优解附近。BP 算法是一种用于训练人工神经网络的常见算法之一。该算法是一种基于梯度下降的优化算法，通过反向传播误差来更新神经网络的权重，以最小化损失函数，从而使神经网络能够学习到输入和输出之间的映射关系。BP 神经网络算法流程如图 3 所示。

由图 3 可知，首先初始化网络参数，然后通过前向传播计算输出并评估损失，接着利用反向传播算法计算损失函数对每个参数的梯度，进而更新网络参数以减小损失。这个过程持续进行直到达到预设的停止条件，如达到最大迭代次数或损失收敛。在训练过程中，采用优化算法来调整学习率以平衡收敛速度和精度。训练完成后，需要通过独立的验证集或测试集来评估模型的性能，确保其具有良好的泛化能力。最后，根据评估结果可能需要进行调优，包括调整网络结构、超参数或数据预处理等，以进一步提升模型性能^[24-25]。在前向传播的过程中，通过输入数据和当前的网络参数计算网络的输出值，其表达式如式 (10) 所示：

$$z_j = \sum_{i=1}^n w_{ij} x_i + b_j \tag{10}$$

式中， z_j 表示第 j 个神经元的带权输入， x_i 是输入数据中的第 i 个特征， w_{ij} 表示连接第 i 个输入特征和第 j 个神经元的权重， b_j 表示第 j 个神经元的偏置项^[26]。误差通常使用损失函数来度量，其表达式如式 (11) 所示：

$$E = \frac{1}{E} \sum_{i=1}^N L(y_i, \hat{y}_i) \tag{11}$$

式中， E 表示总体误差， N 表示样本数量， y_i 表示真实标签， \hat{y}_i 表示模型的预测值， $L(\cdot)$ 表示损失函数。在反向传播中，计算误差对权重和偏置的梯度，以便用于参数的更新，其表达式如式 (12) 所示：

$$\begin{cases} \delta_j = \frac{\partial E}{\partial z_j} = \frac{\partial E}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \\ \frac{\partial E}{\partial w_{ij}} = x_i \delta_j \\ \frac{\partial E}{\partial b_j} = \delta_j \end{cases} \tag{12}$$

式中， δ_j 表示第 j 个神经元的误差项， $\frac{\partial E}{\partial w_{ij}}$ 表示损失函数对连接权重的偏导数， $\frac{\partial E}{\partial b_j}$ 表示损失函数对偏置项的偏导数。通过使用梯度下降法或其变种来更新网络的权重和偏置，其表达式如式 (13) 所示：

$$\begin{cases} w_{ij} = w_{ij} - \alpha \frac{\partial E}{\partial w_{ij}} \\ b_j = b_j - \alpha \frac{\partial E}{\partial b_j} \end{cases} \tag{13}$$

式中， α 表示学习率，控制参数更新的步长。BP 神经网络是一种基于人工神经元相互连接的网络模型，用于解决监督学习问题。该算法由输入层、隐藏层和输出层组成，其中隐藏层可以包含多个层级。每个神经元都与前一层的所有神经元连接，并且具有一个权重值和一个偏置值。输入层用于接收输入数据的节点，每个输入特征对应一个节点。输入层的节点只传递数据给下一层的节点，不接收来自其他节点的信息。隐藏层位于输入层和输出层之间的一层或多层节点。每个隐藏层节点接收来自上一层节点的加权输入，并通过激活函数产生输出，然后将输出传递给下一层的节点。隐藏层节点之间不直接相连，它们之间的连接只存在于相邻的两层之间。输出层负责产生网络的输出。与隐藏层类似，输出层的每个节点接收来自上一层节点的加权输入，并经过激活函数后产生输出值。输出层的输出代表着模型的预测结果。改进后的 BP 神经网络也存在一些缺陷，如收敛速度慢，容易陷入局部最优等，因此采用动量法对其进行改进。

动量法是基于梯度下降算法的一种改进方法，能够解决梯度下降算法收敛速度慢、易陷入局部最优解等问题。动量法的基本思想是在参数更新过程中引入“动量”概念，使得参数更新的方向不仅取决于当前的梯度方向，还考虑了历史梯度方向的影响。这样可以加速收敛并减少参数更新的震荡，有助于跳出局部最优解。其可以将动量系数设置为变量，该方法可以使得误差曲线将变得更加平滑。基于 BP-PSO 的声音识别检测模型结构如图 4 所示。

由图 4 可知，先数据输入并进行预处理，接着进行 PSO

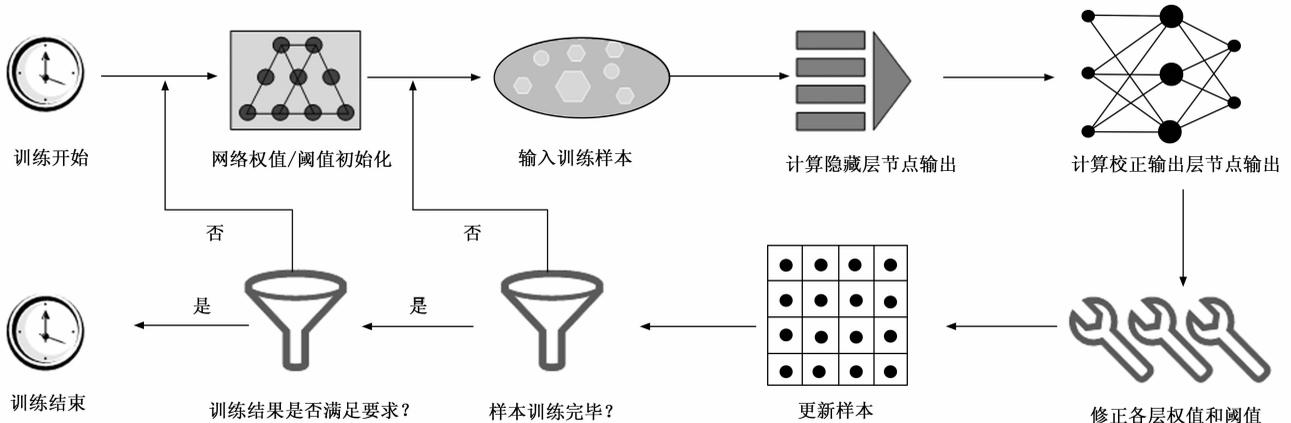


图 3 BP 神经网络训练流程图

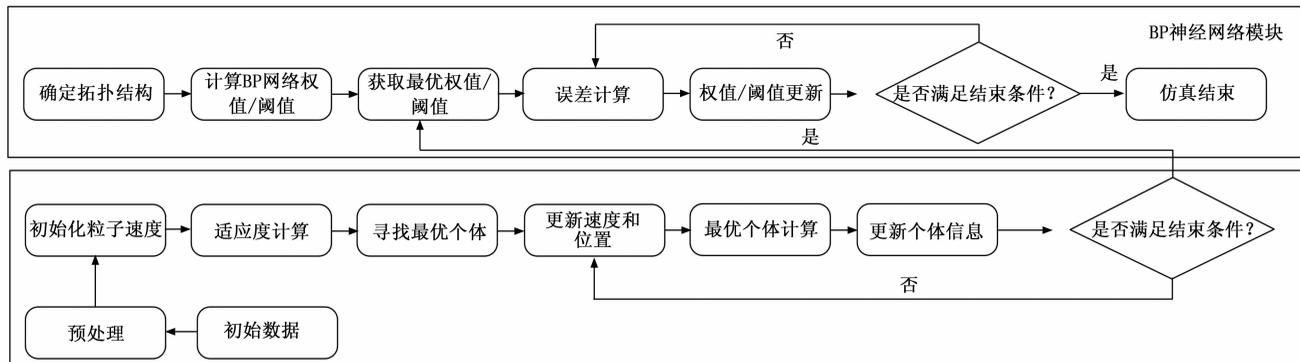


图 4 基于 BP-PSO 的声音识别检测模型结构

模型初始化粒子和速度, 计算每个粒子的适应度值, 寻找个体极值和群体极值, 并通过更新速度和位置来优化粒子的位置。该过程循环进行, 直到满足终止条件。随后, 进入 BP 神经网络部分, 确定网络拓扑结构, 并计算 BP 神经网络的权值和阈值长度。利用从 PSO 模型部分获得的最优权值和阈值, 通过 BP 神经网络计算误差, 并根据误差更新权值和阈值。这一过程也不断迭代, 直至满足终止条件。最终, 进行仿真诊断, 得到结果。

2 基于改进反向传播算法的声音识别与健康检测模型分析

2.1 基于小波去噪的声音数据预处理模型性能分析

为了验证小波去噪模型的性能, 研究搭建了合适的实验平台。选择使用配备 Windows 10 操作系统和 8 GB 内存的计算机, 并搭载了 Intel Core i5-6300CPU, 以充分满足实验的计算需求。研究采用 Kaggle Baby Cry Classification 数据集, 该数据集包含多种不同情境下的婴儿声音, 数据集包含大量的音频文件, 这些文件按照不同的声音类型分类, 且每个音频文件都有相应的标签文件, 用于指示声音的类型。更大的数据集通常可以提高模型的性能, 可以提供更多的样本和特征, 从而帮助模型更好地学习和泛化。因此研究从中选取 1 200 条语音数据, 按照 5 : 1 的比例将其分为训练集和验证集。引入傅里叶变换与短时傅里叶变换与此次采用的小波去噪方法进行对比, 结果如图 5 所示。

表示不同去噪模型的信息量损失率比较, 图 5 (c) 表示不同去噪模型的交并比比较。由图 5 (a) 可知, 随着数据集的增大, 各个去噪模型的信噪比在逐渐增大, 在数据集为 1 000 时, 傅里叶变换去噪模型、短时傅里叶变换去噪模型、小波去噪模型的信噪比分别为 0.81、0.89、0.97。由图 5 (b) 可知, 随着数据集的增大, 各个去噪模型的结构信息损失率在逐渐降低, 在数据集为 1 000 时, 傅里叶变换去噪模型、短时傅里叶变换去噪模型、小波去噪模型的结构信息损失率分别为 0.43、0.30、0.18。由图 5 (c) 可知, 随着数据集的增大, 各个去噪模型的交并比在逐渐增大, 在数据集为 1 000 时, 傅里叶变换去噪模型、短时傅里叶变换去噪模型、小波去噪模型的交并比分别为 0.78、0.91、0.96。这是由于小波变换在时域和频域上都具有更好的局部性。在信号中存在局部突变或者短时突发性的噪声时, 小波变换能够更准确地定位和分析这些局部特征, 从而更有效地去除噪声。且更适用于处理非平稳信号。在非平稳信号的分析 and 处理过程中, 小波变换能够提供更好的时域和频域分辨率, 更好地捕捉信号的瞬时特征。实验结果表明, 所提出的小波去噪模型能够较好地对原始数据进行去噪, 数据去噪后的信噪比, 结构信息损失率和交并比都表现出色。选取两段音频, 对两段音频进行去噪处理后的音频进行比较, 结果如图 6 所示。

图 5 (a) 表示不同去噪模型的信噪比较, 图 5 (b)

图 6 (a) 表示采用不同去噪模型对 A 段音频进行去噪后的音频对比, 图 6 (b) 表示采用不同去噪模型对 B 段音

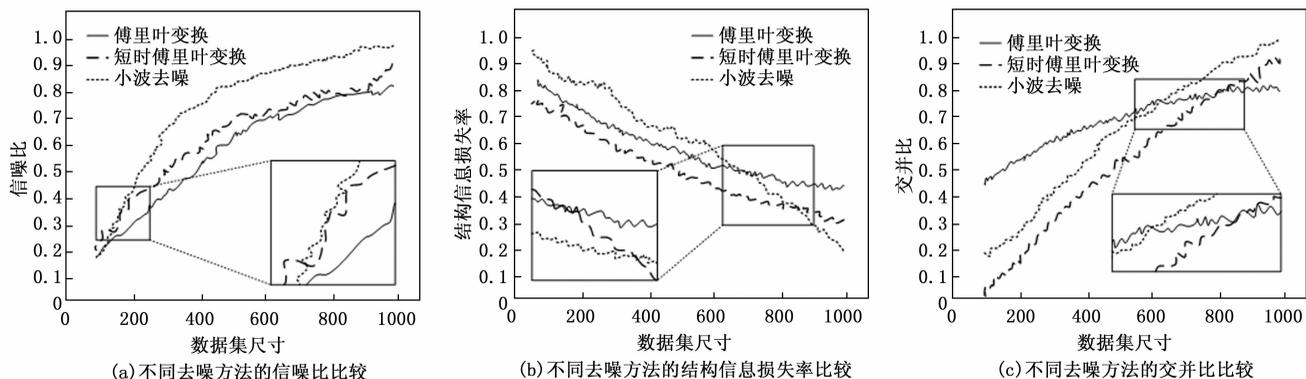


图 5 各个去噪模型的性能比较

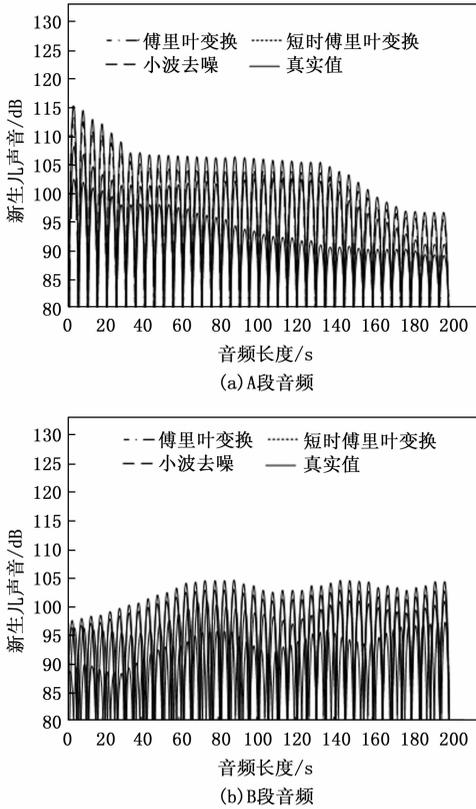


图 6 两段音频的去噪效果分析

频进行去噪后的音频对比。由图 6 可知，新生儿的声
音分贝基本在 110 dB 左右，在进行预处理后，其分贝有所
下降，这是由于去除了部分噪音，使得整体的数据量变小，
从而使得分贝降低，在三种模型中，降低最少的是小波去
噪，这是由于小波去噪能够在对数据中的噪声进行剔除的
同时，最大程度保证原始数据的完整性。在三种方法中，
傅里叶变换去噪处理后的数据，其拟合程度与真实值最低。
出现该情况的原因是，傅里叶变换对噪音的去除效果较差，
只能大量压缩音阶来对音频数据进行降噪。综上所述，小
波变换的多尺度分析特性使得它能够同时提供粗糙和细节
信息，从而更好地反映信号的整体结构和局部细节。这种
逼近能力使得小波去噪方法在去除噪声的同时，尽可能
地保留信号的重要特征。且小波变换具有自适应性，可以
根据信号的特性选择合适的小波基函数。相比之下，傅里
叶变换和短时傅里叶变换的基函数是固定的，不适用于
所有类型的信号。实验结果表明，所提出的小波变换去
噪方法在三种方法中表现出色，能够保存原有数据的情
况下，去除噪声。

2.2 声音识别与健康检测模型性能分析

为了验证所提出声音识别健康检测模型的性能，数
据集采用经过预处理后的数据，按照 5 : 1 的比例划分为
训练集和验证集。研究引入粒子群优化算法，遗传算法
和蚁群优化算法与此次的 BP-PSO 模型进行对比，对
各个模型的性能进行分析，结果如图 7 所示。

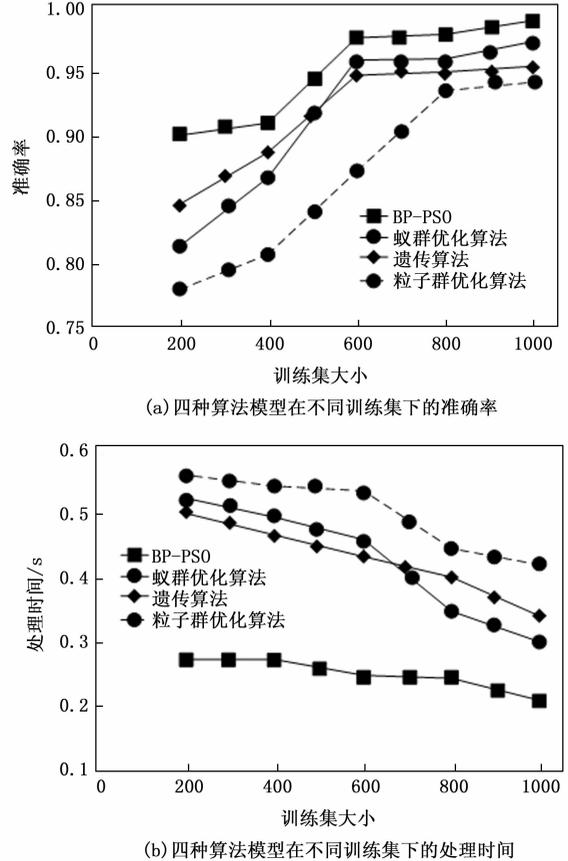


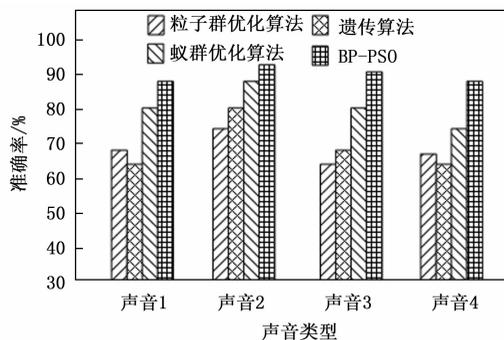
图 7 各个模型的准确率和处理时间对比

图 7 (a) 表示四种算法模型在不同训练集下的准确
率，图 7 (b) 表示四种算法模型在不同训练集下的处理
时间。由图 7 (a) 可知，随着训练集的增大，各个模型
的准确率也在增大，四种模型中，所提出的 BP-PSO
模型有着最为出色的模型性能。由图 7 (b) 可知，随
着训练集的增大，各个模型的处理时间也在减小，四
种模型中，所提出的 BP-PSO 模型有着最为出色的
处理效率。实验结果表明，提出的 BP-PSO 模型在
拥有较高准确率的同时也有较高的效率。对各个不同
类型的新生儿声音来进行识别，结果如图 8 所示。

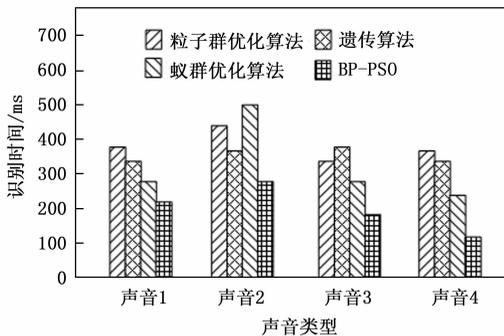
图 8 中，声音 1 到声音 4 分别表示正常声音、弱
声音、强声音、短促声音。图 8 (a) 表示不同类型
的新生儿声音下，各个模型的识别准确率，图 8 (b)
表示不同类型的新生儿声音下，各个模型的识别时间。
由图 8 (a) 可知，在各个模型中，所提出的 BP-
PSO 模型的性能表现最为出色，其对声音 1 到声音
4 的识别准确率分别为 88.2%、92.3%、90.1%、
89.1%。由图 8 (b) 可知，对于不同类型的声音，
对声音 2 的识别时间较长，在四种模型中，所提出
的 BP-PSO 模型对声音 1 到声音 4 的识别时间分
别为 231 ms、287 ms、199 ms、121 ms。实验
结果表明，所提出的 BP-PSO 模型不仅在识别率
上有着较为出色的性能，对识别的处理时间也较短。
对模型的综合性能进行分析，结果如表 1 所示。

表 1 各个模型的综合性能对比

声音类型	粒子群优化算法		遗传算法		蚁群优化算法		BP-PSO	
	准确率	RMSE	准确率	RMSE	准确率	RMSE	准确率	RMSE
声音 1	0.74	0.31	0.79	0.15	0.81	0.13	0.87	0.09
声音 2	0.67	0.29	0.72	0.13	0.77	0.11	0.83	0.07
声音 3	0.81	0.27	0.86	0.11	0.91	0.09	0.97	0.05
声音 4	0.72	0.26	0.77	0.15	0.82	0.11	0.88	0.07



(a) 不同模型的准确率比较



(b) 不同模型的识别时间比较

图 8 不同类型新生儿声音检测性能

表 1 中的数据同样源于 Kaggle Baby Cry Classification 数据集, 且选取 1 200 条语音数据, 按 5 : 1 比例划分为训练集和验证集。首先, 对声音数据进行小波去噪预处理, 去除背景噪音, 保留信号特征。接着, 将预处理后的数据用于训练和验证改进的 BP-PSO 模型。数据划分后, 利用训练集训练模型, 再用验证集评估模型性能, 比较准确率和均方根误差等指标。由表 1 可知, 在四种模型中, 粒子群优化算法的性能表现较差, 而 BP-PSO 模型的性能表现较为出色。其中, 粒子群优化算法对声音 1 到声音 4 的准确率分别为 0.74、0.67、0.81、0.72, RMSE 值为 0.31、0.29、0.27、0.26 表现出较差的性能。BP-PSO 模型对声音 1 到声音 4 的准确率分别为 0.87、0.83、0.97、0.88, RMSE 值为 0.09、0.07、0.05、0.07。这是由于 BP 神经网络的梯度下降法和 PSO 的全局搜索能力, 使得模型能够在初期利用 PSO 进行全局搜索, 在后期利用 BP 算法进行精细的局部搜索, 从而提高了搜索的精度和收敛速度。实验结果表明, 所提出的 BP-PSO 模型有着较为出色的综合性能, 满足实际应用需求。但在更大规模数据集上的性能仍需验证。后续研究应当扩展数据集规模以验证模型性能, 进一步优化模型参数

和结构提升适应性, 以及升级硬件配置提高计算效率。

3 结束语

声音识别和健康检测在医疗领域的重要性日益突出, 但传统的傅里叶变换方法在处理非稳态信号时存在局限性。为了解决这一问题, 研究提出了一种基于改进 BP 算法的声音识别与健康检测模型。实验结果表明, 在数据集为 1 000 时, 傅里叶变换去噪模型、短时傅里叶变换去噪模型、小波去噪模型的信噪比分别为 0.81、0.89、0.97, 结构信息损失率分别为 0.43、0.30、0.18, 交并比分别为 0.78、0.91、0.96。新生儿的语音噪音分贝基本在 110 dB 左右, 在进行预处理后, 其分贝有所下降, 这是由于去除了部分噪音, 使得整体的数据量变小, 从而使得分贝降低, 在三种模型中, 降低最少的是小波去噪, 这是由于小波去噪能够在对数据中的噪声进行剔除的同时, 最大程度保证原始数据的完整性。所提出的 BP-PSO 模型对声音 1 到声音 4 的识别准确率分别为 88.2%、92.3%、90.1%、89.1%。对于不同类型的声音, 对声音 2 的识别时间较长, 在四种模型中, 所提出的 BP-PSO 模型对声音 1 到声音 4 的识别时间分别为 231 ms、287 ms、199 ms、121 ms。BP-PSO 模型对声音 1 到声音 4 的准确率分别为 0.87、0.83、0.97、0.88, RMSE 值为 0.09、0.07、0.05、0.07。研究结果表明, 所提出的 BP-PSO 模型有着较为出色的综合性能。但是研究仍然存在不足, 如模型在处理更大规模数据集时的性能有待进一步验证。未来的研究将继续优化算法, 提升其在复杂环境下的适应性和鲁棒性。

参考文献:

- [1] 韩雪, 慕昱, 盛佳敏. CCPSO 优化支持向量机的鸟声识别技术研究 [J]. 声学技术, 2023, 42 (1): 118-126.
- [2] 胡琰琪, 李艳. 基于 PSO-ELM 的外宣翻译文本语料识别研究 [J]. 现代科学仪器, 2022, 39 (2): 194-197.
- [3] 王磊, 徐超, 李森, 等. 多飞行器协同任务分配的改进粒子群优化算法 [J]. 兵工学报, 2023, 44 (8): 2224-2232.
- [4] 周嗣理, 李国丽, 王群京, 等. 基于改进粒子群优化算法的永磁球形电机驱动策略研究 [J]. 电工技术学报, 2023, 38 (1): 166-176.
- [5] YE H, DONG Y, YANG J, et al. Concurrent optimization method of principal stress orientation interpolated continuous fiber angle (PSO-CFAO) and structural topology [J]. Composite Structures, 2023, 325 (10): 1-12.
- [6] 汪志成, 王泽旺, 朱梦帆, 等. 基于卷积神经网络的局部放电

- 声音识别研究 [J]. 电子测量技术, 2023, 46 (20): 148 - 155.
- [7] 杨智伦, 朱铮涛, 陈树雄, 等. 改进 CNN 的供水管道泄漏声音识别 [J]. 国外电子测量技术, 2023, 42 (1): 153 - 158.
- [8] YE W, JIANG Z, LI Q, et al. A hybrid model for pathological voice recognition of post-stroke dysarthria by using 1DCNN and double-LSTM networks [J]. Applied Acoustics, 2022, 197 (8): 108934.1 - 108934.10.
- [9] EZALDEEN H, BISOY S, MISRA R, et al. Semantics-aware context-based learner modelling using normalized PSO for personalized e-learning [J]. J. Web Eng, 2022, 21 (4): 1187 - 1224.
- [10] SU Z, YANG J, LI P, et al. A precise method of color space conversion in the digital printing process based on PSO-DBN [J]. Textile Research Journal, 2022, 92 (10): 1673 - 1681.
- [11] PAPADIMITRAKIS M, KAPNOPOULOS A, TSAVARTZIDIS S, et al. A cooperative PSO algorithm for Volt-VAR optimization in smart distribution grids [J]. Electric Power Systems Research, 2022, 212 (10): 1 - 9.
- [12] 蔡改贫, 宋佳, 罗小燕, 等. 基于相空间重构和 PSO-K-means 的球磨机负荷状态识别方法 [J]. 科学技术与工程, 2023, 23 (10): 4126 - 4134.
- [13] 于力涵, 洪儒, 吴宇伦, 等. 基于 IKGC-PSO 算法的无人机三维路径规划系统 [J]. 计算机测量与控制, 2023, 31 (8): 259 - 266.
- [14] 刘杰, 付雪娇, 孙兴伟. 基于 PSO-DBN 的风电机组齿轮箱运行状态识别 [J]. 传感技术学报, 2023, 36 (3): 434 - 440.
- [15] 翟永杰, 彭雅妮, 杨旭, 等. 融合 MFCC 和 IMFCC 特征的电厂设备声音识别算法 [J]. 现代电子技术, 2022, 45 (8): 6 - 12.
- [16] 陈振林, 张馨怡, 艾华. 红外目标模拟器校准装置构建和校准技术研究 [J]. 计算机测量与控制, 2023, 31 (1): 237 - 245.
- [17] PRIYADARSHI R, GUPTA B. 2-D coverage optimization in obstacle-based FOI in WSN using modified PSO [J]. Journal of Supercomputing, 2023, 79 (5): 4847 - 4869.
- [18] AGUIAR A L S, SOUSA F B C, MELO Y V L D. Optical distribution network design using PSO [J]. IEEE Communications Letters, 2023, 27 (1): 239 - 242.
- [19] WANG D, FAN X, GUO Y, et al. Quality prediction and control of thin-walled shell injection molding based on GWO-PSO, ACO-BP, and NSGA-II [J]. Journal of Polymer Engineering, 2022, 42 (9): 86 - 94.
- [20] MA C, YAO C, SONG E Z, et al. Prediction and optimization of dual-fuel marine engine emissions and performance using combined ANN with PSO algorithms [J]. International Journal of Engine Research, 2022, 23 (4): 560 - 576.
- [21] OZGOCER T, DAGLI E N, CEYLAN M R, et al. Analysis of long-term antibody response in COVID-19 patients by symptoms grade, gender, age, BMI, and medication [J]. Journal of Medical Virology, 2022, 94 (4): 1412 - 1418.
- [22] WU M, ZHANG F, RUI X. An energy-aware approach for resources allocating in the internet of things using a forest optimization algorithm [J]. Circuit World, 2023, 49 (3): 269 - 280.
- [23] CHAUHAN D K S, VUNDAVILLI P R. Forward kinematics of the stewart parallel manipulator using machine learning [J]. International Journal of Computational Methods, 2022, 19 (8): 21 - 42.
- [24] GAO K, DO D M, CHU S, et al. Robust topology optimization of structures under uncertain propagation of imprecise stochastic-based uncertain field [J]. Thin-Walled Structures, 2022, 175 (7): 109238.1 - 109238.17.
- [25] HAN G, LI H, WANG G, et al. Prediction and control of profile for silicon steel strip in the whole tandem cold rolling based on PSO-BP algorithm [J]. Journal of Manufacturing Processes, 2024, 120 (2): 250 - 259.
- [26] SHARMA S, BHARTI R. Two phased Yarn scheduling with novel binary hypergraphed PSO [J]. The Journal of Supercomputing, 2023, 79 (18): 20712 - 207.
- [13] 黄得女. 基于 Robot 测试框架的关键字驱动的研究与应用 [J]. 计算机系统应用, 2010 (4): 162 - 163.
- [14] 邹明. 组合测试用例生成技术研究 [D]. 重庆: 重庆大学, 2012.
- [15] 刘丽丽, 杜永锋. 通信设备技术指标自动测试与故障诊断系统设计 [J]. 软件, 2023, 44 (5): 52 - 56.
- [16] 陈冬严, 邵杰明, 王东刚, 等. 精通自动化测试框架设计 [M]. 北京: 人民邮电出版社, 2016.
- [17] 郭萌. 基于 Python 的路由器产品自动化测试方案设计与实现 [J]. 信息记录材料, 2022 (10): 144 - 145.
- [18] 储婷婷. 基于 B/S 架构的网络化可配置综合测试平台实现与应用 [J]. 计算机测量与控制, 2017 (7): 32 - 33.
- [19] 窦钠. 航天器数据管理系统软件的自动化测试系统设计 [J]. 航天器工程, 2018 (1): 143 - 144.
- [20] 张松. 基于网口的自动化测试软件的设计和实现 [J]. 信息与电脑 (理论版), 2022 (17): 148 - 151.

(上接第 86 页)

- [6] 陈佐, 张怀相, 方景龙. 嵌入式软件自动化测试技术 [J]. 计算机工程与设计, 2018, 39 (10): 3125 - 3131.
- [7] 张成俊. 以太网转 RS232 转换器设计 [J]. 仪表技术与传感器, 2013 (6): 35 - 36.
- [8] 唐祎祿. RS485-CAN 网关设计与实现 [J]. 信息技术, 2011 (9): 62 - 67.
- [9] 杨国勋. WebDriver 自动化测试研究与应用 [J]. 科技传播, 2019 (2): 95 - 96.
- [10] 崔岩. 软件稳定性测试研究及实践 [J]. 信息技术与标准化, 2013 (7): 61 - 62.
- [11] 马萌. 一种基于软件测试用例库的设计方法 [J]. 计算机与网络, 2021 (15): 48 - 49.
- [12] 接卉. 一种关键字驱动的自动化测试框架 [J]. 计算机应用研究, 2009 (3): 927 - 928.