狄振博, 季薇. 基于语音的帕金森病亚型识别方案[J]. 智能计算机与应用,2026,16(1):66-73. DOI:10.20169/j.issn.2095-2163.24033102

基于语音的帕金森病亚型识别方案

狄振博, 季 薇
(南京邮电大学 通信与信息工程学院, 南京 210003)

摘 要: 帕金森病具有多个亚型,如吞咽困难、震颤和冻结步态等,这些亚型可以帮助医生更好地理解帕金森病的病理机制,方便医生及时调整治疗方案。因此,高效实现帕金森病亚型的识别具有重要的应用价值。由于帕金森病与发音障碍之间有一定的病理联系,并且语音的采集方便、快捷,本文提出了一种基于语音的帕金森病亚型识别方案,将帕金森病亚型识别归结为多标签分类问题,设计了一种多标签分类算法:首先,针对每个亚型选择出其特定特征;接着,从帕金森病语音数据集的特点出发,充分挖掘数据中的相关性信息,包括成对标签相关性信息、高阶标签相关性信息以及标签空间与特征空间相关性信息,进一步提升识别的准确率;最后,将所提算法与多个现有的多标签分类算法在现实帕金森病语音数据集上做对比试验。实验结果显示,所提方法在多个评估指标上表现最好,尤其在 $F1$ 分数($F1$ - score)和子集准确率(Subset accuracy)上,分别达到了 91.27%与 89.03%。

关键词: 帕金森病; 亚型识别; 语音信号; 多标签分类; 相关性信息
中图分类号: R742.5; TP181; TN911.7 **文献标志码**: A **文章编号**: 2095-2163(2026)01-0066-08

Speech-based recognition method for Parkinson's disease subtypes

DI Zhenbo, JI Wei
(School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Parkinson's disease has several subtypes, such as dysphagia, tremors, and freezing of gait, etc. These subtypes can help physicians better understand the pathological mechanisms of Parkinson's disease and facilitate timely adjustment of treatment plans. Therefore, efficiently achieving the identification of Parkinson's disease subtypes has important application value. Due to the pathological connection between Parkinson's disease and dysarthria, and the convenient and fast acquisition of speech, this paper proposes a speech-based recognition method for Parkinson's disease subtypes, attributes the recognition of Parkinson's disease subtypes to a multi-label classification problem, and designs a multi-label classification algorithm: first of all, the specific features of each subtype is selected; then, from the characteristics of the Parkinson's disease speech dataset, the relevance information in the data is fully mined, including pairwise label relevance information, higher-order label relevance information, as well as label-space and feature-space relevance information, to further improve the recognition accuracy; finally, the proposed algorithm is compared with several existing multi-label classification algorithms on the real Parkinson's disease speech dataset, and the experimental results show that the proposed method performs the best in several evaluation metrics, especially in the $F1$ score and Subset accuracy, which reach 91.27% and 89.03%, respectively.

Key words: Parkinson's disease; subtypes recognition; speech signal; multi-label classification; relevance information

0 引 言

帕金森病(Parkinson's Disease,PD)是全球第二大神经退行性疾病,全世界约有 850 多万人患有帕金森病[1]。该病常见于老年人群中,平均发病年龄为 60 岁左右,并随着年龄增长,进行性加重。随着老龄化社会的到来,帕金森病将会造成较大的经济与社会负担[2]。

帕金森病与发音障碍之间有一定的病理联系。

基金项目: 江苏省高校基础科学(自然科学)重大项目(21KJA520003)。
作者简介: 狄振博(1998—),男,硕士研究生,主要研究方向:机器学习与信号处理的交叉研究。
通信作者: 季 薇(1979—),女,博士,教授,主要研究方向:机器学习与信号处理的交叉研究、无线通信等。Email: 1221013728@njupt.edu.cn。
收稿日期: 2024-03-31

目前，已经有大量工作利用语音信号和机器学习技术完成帕金森病的二分类（患病与未患病）诊断[3-6]。然而，帕金森病具有多个亚型，如吞咽困难、震颤和冻结步态等。目前，对这些亚型的评估极度依赖人工评测，需要患者配合医生完成帕金森病综合评分量表 UPDRS（Unified Parkinson Disease Rating Scale，UPDRS），UPDRS 中涉及到的症状即为帕金森病亚型，整个过程不仅耗时耗力，而且评估结果会受患者病情起伏与医生主观判断的影响。因此，人工评测帕金森病亚型具有极大的局限性。

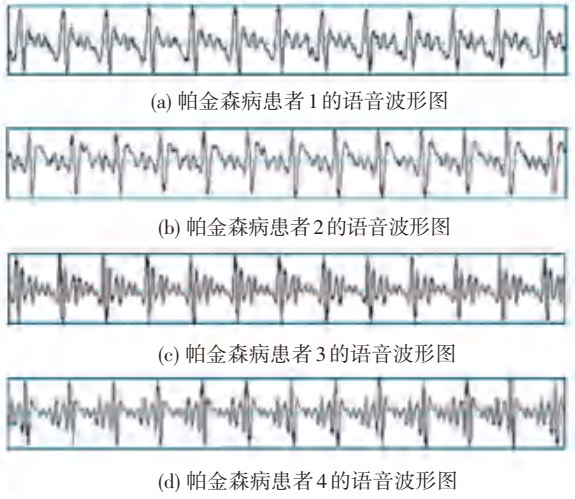
近年来，帕金森病亚型与语音的关系受到了广泛的关注。Shastry[7]提出了一种基于集成机器学习回归模型的帕金森病 UPDRS 运动评分预测模型，是以患者发出的元音 a 作为语料，从语料中提取特征并输入到预测模型中，最终取得了对 UPDRS 运动评分 92%的预测准确率。Tjaden[8]通过 Praat 软件分析了正常人与具有吞咽困难症状的帕金森病患者的语音，发现具有吞咽困难症状的患者语音的谐波噪声比更大，最大发声时间更短。El-Ziaat 等学者[9]设计了一种基于多特征融合的帕金森病患者冻结步态检测模型，分别从患者的不同语音中提取特征，并将这些特征进行融合，输入到卷积神经网络（Convolutional Neural Networks，CNN），结果显示，该模型检测冻结步态症状的准确率达到了 89%。Brown 等学者[10]通过对语音信号的分析发现震颤是帕金森病患者语音损伤最主要的症状。Burk 等学者[11]分析了正常人与具有震颤的帕金森病患者的元音语谱图，发现具有震颤患者的元音峰值突出指标较低。

基于以上研究，本文尝试利用语音信号同时完成多个亚型的识别，实现帕金森病亚型的高效识别。由于多个亚型可以同时存在，所以识别帕金森病亚型可以看作一项多标签分类任务（每个亚型代表一个标签）。目前，已经有大量多标签分类算法被提出[12-18]，然而，由于自采帕金森病语音数据集存在的一些特性，这些算法识别亚型的性能可能是次优的。

帕金森病每个亚型对应的病理机制不同，所以，就可能对应着不同的语音特征。因此，在利用多标签分类识别亚型时，挖掘每个亚型对应的特定特征则能进一步提升识别准确率。

此外，通过对自采帕金森病数据集的分析发现，某些亚型对应的样本数量较少。考虑到在多标签分类中，假如 2 个标签具有强相关性，可以利用样本数量较多的标签帮助样本数量较少的标签建模。另一方面，强相关的标签可能表现出强共现性，比如，当一个样本同时拥有“狮子”和“草原”两个标签时，那么该样本大概率拥有“非洲”的标签。使用高阶标签相关性，即每个标签受其它所有标签的影响程度，可以帮助捕获标签间的强共现性关系，有助于标签建模的相互促进。然而，在高阶标签相关性学习的过程中，可能会出现不必要的相关性信息，使得错误相关信息得到传播，从而降低多标签分类性能。

受年龄、性别以及患病程度的影响，帕金森病患者会表现出不同的语音特性。帕金森病患者的语音波形如图 1 所示。由图 1 可知，帕金森病患者 1 和 2 的语音波形相似，3 和 4 的语音波形相似。由于帕金森病数据集的特征是从患者的语音中提取的，所以，患者 1 和 2 对应语音样本的特征空间更相似，患者 3 和 4 对应语音样本的特征空间更相似。然而，在进行分类训练时，样本的标签空间也会起作用，如果在特征空间相似的样本在标签空间不相似，就会导致分类性能的下降。



(a) 帕金森病患者 1 的语音波形图

(b) 帕金森病患者 2 的语音波形图

(c) 帕金森病患者 3 的语音波形图

(d) 帕金森病患者 4 的语音波形图

图 1　帕金森病患者的语音波形图

Fig. 1　Speech waveforms in patients with Parkinson's disease

基于以上考虑，本文充分利用 3 种相关性信息，即二阶（2 个标签之间）标签相关性、高阶标签相关性以及特征空间与标签空间的相关性，设计了一种基于标签特定特征的多标签分类方法（LSMPFS），用于帕金森病亚型的识别。具体来说，本文首先引入 l_1 范数完成标签特定特征的生成；接着，引入二阶标签相关性和高阶标签相关性，并使用 l_1 范数对高阶标签关系矩阵进行稀疏化，将与每个标签不相关的标签权重值置为 0，从而去除高阶标签相关性学习的过程中出现的不必要的相关性；最后，通过一个正则化项约束在特征空间相似的样本在标签空间也相似。在帕金森

病数据集上与多个多标签分类算法做对比实验,结果显示,该算法在多个指标上表现最优。

1 算法设计

在介绍本文所提算法之前,为了方便表示,制定了一些符号规则:$\boldsymbol{R}=[\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots,\boldsymbol{r}_z]\in\mathbb{R}^{z\times m}$ 表示具有 z 个语音样本的训练集的特征矩阵,其中,$\boldsymbol{r}_i=[r_i^1,r_i^2,\cdots,r_i^m]\in\mathbb{R}^m$ 表示具有 m 维特征的第 i 个样本的特征向量。$\boldsymbol{P}=[\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_z]\in\mathbb{R}^{z\times k}$ 为训练集的标签矩阵,其中,$\boldsymbol{p}_i=[p_i^1,p_i^2,\cdots,p_i^k]$ 表示具有 k 维特征的第 i 个样本的标签向量。$t_j(j=1,2,\cdots,k)$ 表示第 j 个标签。当样本 $\boldsymbol{r}_i$ 具有标签 t_j 时,$p_i^j=1$;否则,$p_i^j=0$。

本文使用线性模型对目标函数建模,即:

$$\min_{U^1,c}\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{P}\|_F^2 \tag{1}$$

其中,$\boldsymbol{U}^1=[\boldsymbol{u}_1^1,\boldsymbol{u}_2^1,\cdots,\boldsymbol{u}_k^1]\in\mathbb{R}^{m\times k}$ 表示线性模型的权重矩阵;$\boldsymbol{u}_i^1=[u_{i1}^1,u_{i2}^1,\cdots,u_{iz}^1]^{\mathrm{T}}\in\mathbb{R}^{z\times 1}$ 表示第 i 个标签对应的特征权重向量;$\boldsymbol{u}_i^1$ 的第 j 列值 u_{ij}^1 决定了第 i 个特征对预测第 j 个标签的贡献度,u_{ij}^1 值越大,说明第 i 个特征对预测第 j 个标签的贡献越大;$\boldsymbol{c}=[c_1,c_2,\cdots,c_k]^{\mathrm{T}}\in\mathbb{R}^{k\times 1}$ 表示线性模型的偏差向量;$\boldsymbol{1}_z=[1,1,\cdots,1]^{\mathrm{T}}\in\mathbb{R}^{z\times 1}$ 表示全 1 向量。

1.1 标签特定特征生成

为了完成标签特定特征的生成,本文使用 l_1 范数对线性模型的权重矩阵 $\boldsymbol{U}^1$ 进行稀疏化:

$$\min_{U^1,c}\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{P}\|_F^2+\alpha_1\|\boldsymbol{U}^1\|_1 \tag{2}$$

其中,α_1 表示可调参数。图 2 给出了标签特定特征生成示意图,在 l_1 范数的稀疏作用下,权重矩阵 $\boldsymbol{U}^1$ 中多个值变为 0,这些 0 值在预测标签时无任何作用。与之相反,非零值在预测标签时起决定性作用,这些非零值对应的特征就是标签特定特征。在图 2 中,每个标签的特定特征用不同颜色覆盖,可以看出,标签 t_1 的特定特征为 f_1 和 f_6,标签 t_2 的特定特征为 f_1 和 f_3,标签 t_3 的特定特征为 f_2 和 f_5。

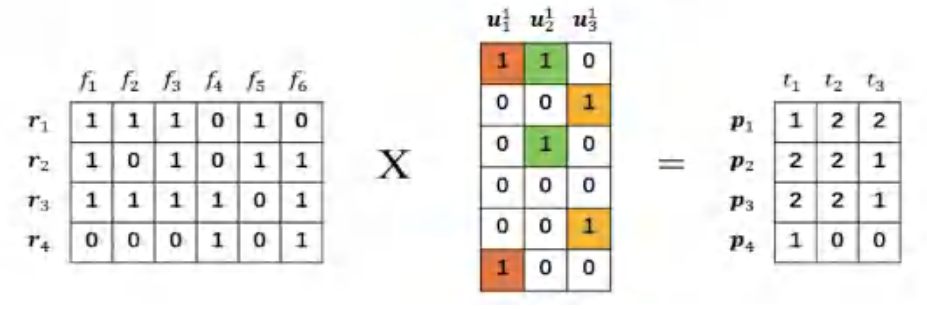


图 2 标签特定特征生成示意图

Fig. 2 Schematic diagram of label-specific feature generation

1.2 标签相关性建模

在建模二阶标签相关性时,本文基于一个假设:如果 2 个标签 t_i 与 t_j 具有强相关性,则对预测标签 t_i 起作用的特征大概率也对预测标签 t_j 起作用,即这 2 个标签将共享更多的标签特定特征,2 个标签对应的特征权重向量 $\boldsymbol{u}_i^1$ 与 $\boldsymbol{u}_j^1$ 将非常相似。本文使用余弦相似度计算标签之间的相似性,设标签 t_i 与 t_j 的相似性为 s_{ij},则标签相似性矩阵 $\boldsymbol{S}=[s_{ij}]_{k\times k}\in\mathbb{R}^{k\times k}$。

本文通过一个正则化项完成对二阶标签相关性的建模,该正则化项定义如下:

$$\sum_{i=1}^{k}\sum_{j=1}^{k}s_{ij}\|\boldsymbol{u}_i^1-\boldsymbol{u}_j^1\|=\mathrm{tr}(\boldsymbol{U}^1\boldsymbol{L}_1(\boldsymbol{U}^1)^{\mathrm{T}}) \tag{3}$$

其中,$\boldsymbol{L}_1=\boldsymbol{S}^*-\boldsymbol{S}\in\mathbb{R}^{k\times k}$ 表示 $\boldsymbol{S}$ 的拉普拉斯矩阵(Laplacian matrix)。$\boldsymbol{S}^*$ 是对角矩阵,其第 j 个对角元素 $s_{jj}^*=\sum_{i=1}^{k}s_{ji}$。

将上述正则化项加入式(2)定义的目标函数,得到新的目标函数:

$$\min_{U^1,c}\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{P}\|_F^2+\alpha_1\|\boldsymbol{U}^1\|_1+\frac{1}{2}\alpha_2\mathrm{tr}(\boldsymbol{U}^1\boldsymbol{L}_1(\boldsymbol{U}^1)^{\mathrm{T}}) \tag{4}$$

其中,α_2 表示可调参数。为了捕获标签间的强共现性关系,进一步提升帕金森病亚型识别的精度,本文引用高阶标签相关性:

$$\min_{U^1,U^2,c}\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{Y}\boldsymbol{U}^2-\boldsymbol{P}\|_F^2+\alpha_1\|\boldsymbol{U}^1\|_1+\frac{1}{2}\alpha_2\mathrm{tr}(\boldsymbol{U}^1\boldsymbol{L}_1(\boldsymbol{U}^1)^{\mathrm{T}}) \tag{5}$$

其中,$\boldsymbol{U}^2=[\boldsymbol{u}_1^2,\boldsymbol{u}_2^2,\cdots,\boldsymbol{u}_k^2]\in\mathbb{R}^{k\times k}$ 表示高阶标

签相关性矩阵，$\boldsymbol{u}_i^2=[u_{i1}^2,u_{i2}^2,\cdots,u_{ik}^2]^{\mathrm{T}}\in\mathbb{R}^{k\times 1}$ 表示第 i 个标签的标签权重向量。为了去除高阶标签相关性学习过程中出现的不必要的相关性信息，本文引入 l_1 范数对高阶标签相关性矩阵 $\boldsymbol{U}^2$ 进行稀疏化，将与每个标签不相关的标签权重值置为 0。进一步推得：

$$\min_{\boldsymbol{U}^1,\boldsymbol{U}^2,\boldsymbol{c}}\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{Y}\boldsymbol{U}^2-\boldsymbol{P}\|_F^2+\alpha_1\|\boldsymbol{U}^1\|_1+\frac{1}{2}\alpha_2\mathrm{tr}(\boldsymbol{U}^1\boldsymbol{L}_1(\boldsymbol{U}^1)^{\mathrm{T}})+\alpha_3\|\boldsymbol{U}^2\|_1 \tag{6}$$

其中，α_3 表示可调参数。

1.3　特征空间与标签空间约束

为了消除帕金森病患者的语音差异性对亚型识别的影响，本文设计一个正则化项，约束在特征空间相似的语音样本在标签空间也相似。由于帕金森病语音数据集在特征空间中可能存在噪声，使用余弦相似度评估语音样本间的相似性可能是不准确的。DPC 算法[19]能够在聚类的过程中剔除噪声样本，所以，本文采用 DPC 算法来评估语音样本间的相似性。对于语音样本 i 和 j，两者在特征空间的相似性定义如下：

$$h_{ij}=\begin{cases}1, & r_j\in\mathrm{DPC}(r_i)\\ 0, & \text{其它}\end{cases} \tag{7}$$

其中，$r_j\in\mathrm{DPC}(r_i)$ 表示语音样本 i 和 j 在同一聚类簇中。特征空间的相似性矩阵定义为 $\boldsymbol{H}=[h_{ij}]_{z\times z}\in\mathbb{R}^{z\times z}$。

本文采用预测标签空间来代替真实标签空间，定义上述正则化项：

$$\sum_{i=1}^{z}\sum_{j=1}^{z}h_{ij}\|(\boldsymbol{r}_i\boldsymbol{U}^1+\boldsymbol{c}^{\mathrm{T}})-(\boldsymbol{r}_j\boldsymbol{U}^1+\boldsymbol{c}^{\mathrm{T}})\|=\mathrm{tr}((\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{L}_2(\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}})) \tag{8}$$

其中，$\boldsymbol{L}_2=\boldsymbol{H}^*-\boldsymbol{H}\in\mathbb{R}^{k\times k}$ 表示 $\boldsymbol{H}$ 的拉普拉斯矩阵；$\boldsymbol{H}^*$ 是对角矩阵，其第 i 个对角元素 $h_{ii}^*=\sum_{j=1}^{z}h_{ij}$。

将该正则化项加入式(6)定义的目标函数，得到最终的目标函数，具体如下：

$$\min_{\boldsymbol{U}^1,\boldsymbol{U}^2,\boldsymbol{c}}\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{Y}\boldsymbol{U}^2-\boldsymbol{P}\|_F^2+\alpha_1\|\boldsymbol{U}^1\|_1+\frac{1}{2}\alpha_2\mathrm{tr}(\boldsymbol{U}^1\boldsymbol{L}_1(\boldsymbol{U}^1)^{\mathrm{T}})+\alpha_3\|\boldsymbol{U}^2\|_1+\frac{1}{2}\alpha_4\mathrm{tr}((\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{L}_2(\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}})) \tag{9}$$

其中，α_4 表示可调参数。

1.4　优化

本文采用交替优化的方式对式(9)定义的目标函数进行求解。

(1)固定 $\boldsymbol{c}$，对 $\boldsymbol{U}^1$ 和 $\boldsymbol{U}^2$ 进行优化。目标函数可以改写为：

$$\min_{\boldsymbol{U}^1,\boldsymbol{U}^2}f(\boldsymbol{U}^1,\boldsymbol{U}^2)+g(\boldsymbol{U}^1,\boldsymbol{U}^2) \tag{10}$$

其中，$f(\boldsymbol{U}^1,\boldsymbol{U}^2)$ 与 $g(\boldsymbol{U}^1,\boldsymbol{U}^2)$ 定义如下：

$$f(\boldsymbol{U}^1,\boldsymbol{U}^2)=\frac{1}{2}\|\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{P}\boldsymbol{U}^2-\boldsymbol{P}\|_F^2+\frac{1}{2}\alpha_2\mathrm{tr}(\boldsymbol{U}^1\boldsymbol{L}_1(\boldsymbol{U}^1)^{\mathrm{T}})+\frac{1}{2}\alpha_4\mathrm{tr}((\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{L}_2(\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}})) \tag{11}$$

$$g(\boldsymbol{U}^1,\boldsymbol{U}^2)=\alpha_1\|\boldsymbol{U}^1\|_1+\alpha_3\|\boldsymbol{U}^2\|_1 \tag{12}$$

可以看出，$f(\boldsymbol{U}^1,\boldsymbol{U}^2)$ 是式(10)定义的凸目标函数的光滑部分，且 $f(\boldsymbol{U}^1,\boldsymbol{U}^2)$ 是利普希茨连续(Lipschitz continuous)的，其对 $\boldsymbol{U}^1$ 和 $\boldsymbol{U}^2$ 的偏导分别是：

$$\nabla_{\boldsymbol{U}^1}f(\boldsymbol{U}^1,\boldsymbol{U}^2)=\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{P}\boldsymbol{U}^2-\boldsymbol{P})+\alpha_2\boldsymbol{U}^1\boldsymbol{L}_1+\alpha_4\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2(\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}) \tag{13}$$

$$\nabla_{\boldsymbol{U}^2}f(\boldsymbol{U}^1,\boldsymbol{U}^2)=\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{R}\boldsymbol{U}^1+\boldsymbol{1}_z\boldsymbol{c}^{\mathrm{T}}+\boldsymbol{P}\boldsymbol{U}^2-\boldsymbol{P}) \tag{14}$$

令 $\boldsymbol{U}_1=(\boldsymbol{U}_1^1,\boldsymbol{U}_1^2)$，$\boldsymbol{U}_2=(\boldsymbol{U}_2^1,\boldsymbol{U}_2^2)$，则有：

$$\|\nabla f(\boldsymbol{U}_1)-\nabla f(\boldsymbol{U}_2)\|_F^2\leqslant\begin{pmatrix}3\begin{pmatrix}\|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}\|_2^2+\|\alpha_2\boldsymbol{L}_1\|_2^2+\\ \|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{P}\|_2^2+\|\alpha_4\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2\boldsymbol{R}\|_2^2\end{pmatrix}+\\ 2\|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\|_2^2\end{pmatrix}\left\|\begin{matrix}\boldsymbol{U}_1^1-\boldsymbol{U}_2^1\\ \boldsymbol{U}_1^2-\boldsymbol{U}_2^2\end{matrix}\right\|_F^2 \tag{15}$$

即 $f(\boldsymbol{U}^1,\boldsymbol{U}^2)$ 的利普希茨常数(Lipschitz constant)为：

$$\boldsymbol{L}_f=\sqrt{3\begin{pmatrix}\|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}\|_2^2+\|\alpha_2\boldsymbol{L}_1\|_2^2+\\ \|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{P}\|_2^2+\|\alpha_4\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2\boldsymbol{R}\|_2^2\end{pmatrix}+2\|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\|_2^2} \tag{16}$$

相应地，本文采用基于近端梯度下降算法的迭代过程来求解式(10)。给定第 t 次迭代时的解 $\boldsymbol{U}^{1(t)}$，权重矩阵 $\boldsymbol{U}^1$ 下一次迭代解 $\boldsymbol{U}^{1(t+1)}$ 的计算方式如下：

$$\boldsymbol{Z}^{(t)}=\boldsymbol{U}^{1(t)}-\frac{1}{\boldsymbol{L}_f}\nabla_{\boldsymbol{U}^1}f(\boldsymbol{U}^{1(t)},\boldsymbol{U}^{2(t)}) \tag{17}$$

$$O_{ij}^{(t)}=\mathrm{sign}(Z_{ij}^{(t)})\cdot\max(|Z_{ij}^{(t)}|-\frac{\alpha_1}{L_f},0) \tag{18}$$

$$U_{ij}^{1(t+1)}=O_{ij}^{(t)}/\|\boldsymbol{o}_j\|_2 \tag{19}$$

其中，$\boldsymbol{Z}^{(t)}\in\mathbb{R}^{m\times k}$ 与 $\boldsymbol{O}^{(t)}\in\mathbb{R}^{m\times k}$ 均是计算

$\boldsymbol{U}^{1(t+1)}$ 的中间矩阵;$Z_{ij}^{(t)}$,$O_{ij}^{(t)}$ 与 $U_{ij}^{1(t+1)}$ 分别表示 $\boldsymbol{Z}^{(t)}$,$\boldsymbol{O}^{(t)}$ 与 $\boldsymbol{U}^{1(t+1)}$ 的第 i 行第 j 列值;$\boldsymbol{o}_j=[O_{1j}^{(t)},O_{2j}^{(t)},\cdots,O_{mj}^{(t)}]^{\mathrm{T}}$ 表示 $\boldsymbol{O}^{(t)}$ 的第 j 列向量;$\mathrm{sign}(Z_{ij}^{(t)})$ 为符号函数。对高阶标签相关性矩阵 $\boldsymbol{U}^2$ 的求解同 $\boldsymbol{U}^1$,给定第 t 次迭代时的解 $\boldsymbol{U}^{2(t)}$,高阶标签相关性矩阵 $\boldsymbol{U}^2$ 下一次迭代解 $\boldsymbol{U}^{2(t+1)}$ 的计算方式如下:

$$\boldsymbol{V}^{(t)}=\boldsymbol{U}^{2(t)}-\frac{1}{L_f}\nabla_{U^2}f(\boldsymbol{U}^{1(t)},\boldsymbol{U}^{2(t)}) \tag{20}$$

$$Q_{ij}^{(t)}=\mathrm{sign}(V_{ij}^{(t)})\cdot\max\left(|V_{ij}^{(t)}|-\frac{\alpha_3}{L_f},0\right) \tag{21}$$

$$U_{ij}^{2(t+1)}=Q_{ij}^{(t)}/\|\boldsymbol{q}_j\|_2 \tag{22}$$

其中,$\boldsymbol{V}^{(t)}\in\mathbb{R}^{k\times k}$ 与 $\boldsymbol{Q}^{(t)}\in\mathbb{R}^{k\times k}$ 均是计算 $\boldsymbol{U}^{2(t+1)}$ 的中间矩阵;$V_{ij}^{(t)}$,$Q_{ij}^{(t)}$ 与 $U_{ij}^{2(t+1)}$ 分别表示 $\boldsymbol{V}^{(t)}$,$\boldsymbol{Q}^{(t)}$ 与 $\boldsymbol{U}^{2(t+1)}$ 的第 i 行第 j 列值;$\boldsymbol{q}_j=[Q_{1j}^{(t)},Q_{2j}^{(t)},\cdots,Q_{kj}^{(t)}]^{\mathrm{T}}$ 表示 $\boldsymbol{Q}^{(t)}$ 的第 j 列向量;$\mathrm{sign}(V_{ij}^{(t)})$ 与 $\mathrm{sign}(Z_{ij}^{(t)})$ 的定义一致。

(2)固定 $\boldsymbol{U}^1$ 和 $\boldsymbol{U}^2$,对 c 进行优化。指定式(9)定义的目标函数为 $f(\boldsymbol{U}_1,\boldsymbol{U}_2,\boldsymbol{c})$,其对 $\boldsymbol{c}$ 的偏导为:

$$\frac{\partial f(\boldsymbol{U}^1,\boldsymbol{U}^2,\boldsymbol{c})}{\partial\boldsymbol{c}}=(\boldsymbol{U}^1)^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{1}_z+n\boldsymbol{c}+(\boldsymbol{U}^2)^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{1}_z-\boldsymbol{P}^{\mathrm{T}}\boldsymbol{1}_z+\alpha_4((\boldsymbol{U}^1)^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2\boldsymbol{1}_z+\boldsymbol{c}\boldsymbol{1}_z^{\mathrm{T}}\boldsymbol{L}_2\boldsymbol{1}_z) \tag{23}$$

令 $\frac{\partial f(\boldsymbol{U}^1,\boldsymbol{U}^2,\boldsymbol{c})}{\partial\boldsymbol{c}}=0$,可得:

$$\boldsymbol{c}=-\frac{1}{z+\alpha_4\boldsymbol{1}_z^{\mathrm{T}}\boldsymbol{L}_2\boldsymbol{1}_z}((\boldsymbol{U}^1)^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{1}_z+(\boldsymbol{U}^2)^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{1}_z-\boldsymbol{P}^{\mathrm{T}}\boldsymbol{1}_z+\alpha_4(\boldsymbol{U}^1)^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{L}_2\boldsymbol{1}_z) \tag{24}$$

其中,z 表示训练集中的语音样本数。

对于测试集中的语音样本,根据优化得到的参数 $\boldsymbol{U}^1$,$\boldsymbol{U}^2$ 和 $\boldsymbol{c}$ 完成标签预测。对待测样本 $\boldsymbol{x}_i$,其预测标签集 $\boldsymbol{y}_i$ 可通过下式计算得到:

$$\boldsymbol{y}_i=\{\mathbf{pre}_j\mid((\boldsymbol{x}_i\boldsymbol{u}_j^1+c_j)+(\boldsymbol{x}_i\boldsymbol{u}_j^1+c_j)\boldsymbol{u}_j^2)>0.5,\ 1\leqslant j\leqslant k\} \tag{25}$$

其中,$\mathbf{pre}_j$ 表示 $\boldsymbol{x}_i$ 的第 j 个预测标签,c_j 表示 $\boldsymbol{c}$ 的第 j 维值。

至此,研究给出 LSMPFS 算法优化流程具体如下。

算法 LSMPFS 算法

输入 训练集特征矩阵 $\boldsymbol{R}$ 与标签矩阵 $\boldsymbol{P}$,参数 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 与 η

输出 线性模型的权重矩阵 $\boldsymbol{U}^1$ 与偏差向量 $\boldsymbol{c}$,高阶标签相关性矩阵 $\boldsymbol{U}^2$

1. 初始化 $\boldsymbol{U}^1=(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}+\eta\boldsymbol{I})^{-1}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{P}$,$\boldsymbol{U}^2=(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}+\eta\boldsymbol{I})^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$;随机初始化 $\boldsymbol{c}$;

2. 重复:

$\boldsymbol{U}^{1(0)}=\boldsymbol{U}^1$,$\boldsymbol{U}^{2(0)}=\boldsymbol{U}^2$,$t=0$

重复:

3. 分别根据式(13)和式(14)计算 $\nabla_{U^1}f(\boldsymbol{U}^{1(t)},\boldsymbol{U}^{2(t)})$ 和 $\nabla_{U^2}f(\boldsymbol{U}^{1(t)},\boldsymbol{U}^{2(t)})$

4. 根据式(16)计算 $\boldsymbol{L}_f$

5. 根据式(17)~(19)更新 $\boldsymbol{U}^{1(t+1)}$;根据式(20)~(22)更新 $\boldsymbol{U}^{2(t+1)}$

6. $t=t+1$

直到收敛

7. $\boldsymbol{U}^1=\boldsymbol{U}^{1(t+1)}$,$\boldsymbol{U}^2=\boldsymbol{U}^{2(t+1)}$

8. 根据式(24)更新 $\boldsymbol{c}$

9. 直到收敛

2 实验

2.1 多标签帕金森病语音数据集介绍

本文采用的帕金森病语音数据集是课题组与医院合作采集的,所有采集均在安静的室内环境进行。采集完成后向医生询问受试者吞咽困难、震颤和冻结步态的情况(患有任意一种,对应标签置为 1,否则,置为 0),作为对应语音的标签集。

受试者包含 38 名男性和 32 名女性,具体情况见表 1。每个受试者都记录了 5~6 条语音,包括持续元音/a/、重复单音节/pa-ka-la/和一段固定的情景对话。经过处理后,共得到 414 条语音。

表 1 受试者信息统计

Table 1 Subject information statistics

性别	语音样本数	平均年龄/岁	平均患病时间/年
男	238	72.1	5.4
女	176	68.5	4.3

为了对采集到的语音数据进行表征,本文使用 Disvoice 工具[20]对每条语音进行特征提取,并将得到的特征拼接起来,得到一个 655 维的向量作为该条语音的特征向量。

2.2 实验设置

为了评估所提算法 LSMPFS 在帕金森亚型识别上的性能,本文采用常见的 6 种多标签分类性能评估指标,即 Hamming loss、Coverage、One error、Subset accuracy、$F1$ score 和 Average precision[21]。其中,Hamming loss, Coverage 与 One error 的值越低,代表分类性能越好,Subset accuracy、$F1$ score 和 Average precision 则相反,值越高,代表分类性能越好。在实验之前对数据集的所有特征做了归一化处理。

本文将 LSMPFS 与 7 种多标签分类方法进行对

比，包括 BR[12]、ML－kNN[13]、LIFT[14]、LLSF[15]、WRAP[16]、SENCE[17]与 RLFSCL[18]。LSMPFS 的参数设置：$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 5$，$\eta = 2$。其余 7 种多标签分类方法分别采用网格搜索的方式确定其在多标签帕金森病语音数据集上的最佳参数。采用 10 折交叉验证的方式统计实验结果。

2.3　实验结果与分析

图 3 给出了 LSMPFS、BR、ML－kNN、LIFT、LLSF、WRAP、SENCE 与 RLFSCL 在 6 种评估指标上的分类性能及标准差。从图 3 可以看出 LSMPFS 在大多数情况下表现最优，说明 LSMPFS 在识别帕金森病亚型方面具有很强的竞争力。当不考虑标准差时，8 种多标签分类算法在 6 种评估指标上的分类性能见表 2（加粗数据表示最优结果）。由表 2 可知，LSMPFS 在 Hamming loss、Coverage、Subset accuracy 和 *F*1 score 四个指标上均表现最优，分别达到了 0.121 2、0.500 3、0.890 3 与 0.912 7，而在 One error 和 Average precision 两个指标上表现次优，且仅与表现最优的 SENCE 分别相差 0.010 2 和 0.008 9。

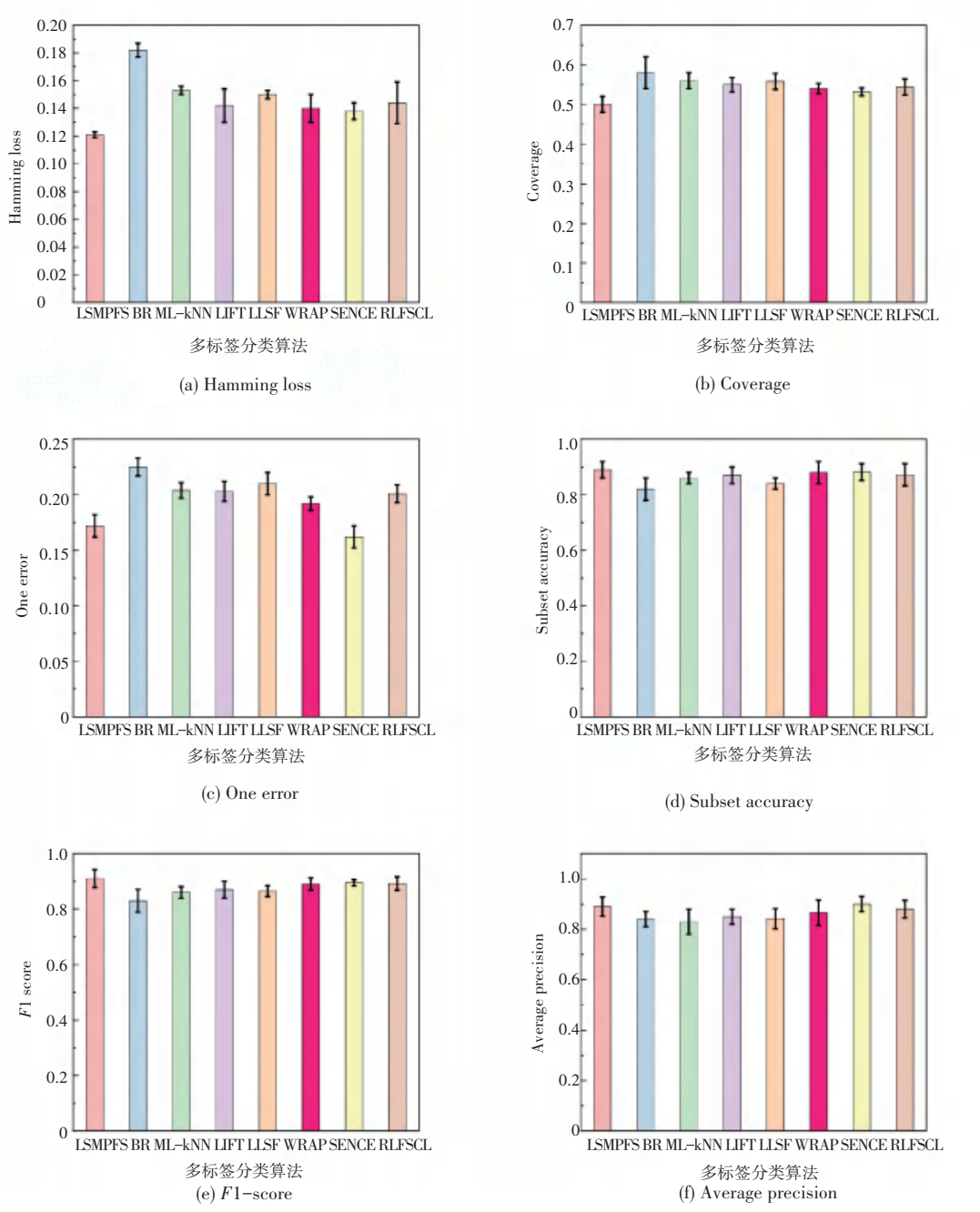


图 3　8 种多标签分类算法在 6 种评估指标上的分类性能及标准差

Fig. 3　Classification performance and standard deviation of eight multi-label classification algorithms on six evaluation metrics

表 2 8 种多标签分类算法在 6 种评估指标上的分类性能

Table 2 Classification performance of eight multi-label classification algorithms on six evaluation metrics

算法	Hamming loss	Coverage	One error	Subset accuracy	*F*1 score	Average precision
LSMPFS	**0.121 2**	**0.500 3**	0.172 4	**0.890 3**	**0.912 7**	0.891 1
BR	0.182 3	0.581 4	0.225 2	0.826 4	0.836 4	0.841 2
ML-kNN	0.153 2	0.562 1	0.204 7	0.863 1	0.862 3	0.832 4
LIFT	0.142 4	0.557 1	0.203 6	0.871 1	0.873 1	0.851 6
LLSF	0.153 2	0.558 2	0.215 3	0.843 1	0.865 4	0.842 3
WRAP	0.141 2	0.542 0	0.192 4	0.881 2	0.891 2	0.866 2
SENCE	0.138 7	0.532 5	**0.162 2**	0.882 8	0.895 2	**0.900 0**
RLFSCL	0.144 2	0.544 3	0.201 2	0.872 3	0.892 3	0.883 1

2.4 消融实验

本文所提 LSMPFS 算法含有 3 种相关性信息，即二阶标签相关性、高阶标签相关性以及特征空间与标签空间的相关性，为了方便表示，分别将这 3 种相关性命名为 c1、c2 与 c3。为进一步探究 LSMPFS 的分类性能，本节通过消融实验检测 3 种相关性信息对分类性能的影响。

实验结果见表 3(加粗数据表示最优结果)。由表 3 可知，缺失二阶标签相关性信息时(LSMPFS-without c1)，由于帕金森病数据集中某些标签的样本数量较少，在没有其它强相关标签帮助建模的情况下，LSMPFS 算法在 6 个评估指标上的性能均略有下降。缺失高阶标签相关性信息时(LSMPFS-without c2)，帕金森病数据集中标签的强共现性关系被忽略，导致 LSMPFS 算法在 6 个评估指标上的性能均有所下降。缺失特征空间与标签空间的相关性信息时(LSMPFS-without c3)，受帕金森病患者语音特性差异的影响，LSMPFS 算法在 6 个评估指标上的性能也都有所下降。

表 3 消融实验结果

Table 3 Results of ablation experiments

评估指标	LSMPFS	LSMPFS-without c1	LSMPFS-without c2	LSMPFS-without c3
Hamming loss	**0.121 2**	0.132 4	0.137 8	0.142 1
Coverage	**0.500 3**	0.511 8	0.520 4	0.522 1
One error	**0.172 4**	0.186 2	0.188 3	0.190 2
Subset accuracy	**0.890 3**	0.884 2	0.877 8	0.870 5
*F*1 score	**0.912 7**	0.900 3	0.899 2	0.891 3
Average precision	**0.891 1**	0.890 2	0.882 1	0.881 3

2.5 标签特定特征选择分析

表 4 给出了 LSMPFS 在帕金森病数据集上为每个标签选择出的特征的数量以及两两标签之间对应特征的相似性。相似性的定义如下：

$$J(t_i,t_j)=\frac{|FS_i \cap FS_j|}{|FS_i \cup FS_j|} \tag{26}$$

其中，FS_i 与 FS_j 分别表示第 i 和第 j 个标签对应的特定特征集合，分子表示两者的交集，分母表示两者的并集。从表 4 中可以看出，每个标签对应的特征的数量是不同的，并且两两标签之间对应特征的相似性均不为 1，说明 LSMPFS 具备为每个亚型选择特定特征的能力。其中，标签 t_1 与 t_2 对应特征的相似性最大，说明这 2 个标签具有较强的相关性，从而共享更多的特征。

表 4 每个标签对应的特征的数量及两两标签之间对应特征的相似性

Table 4 The number of features corresponding to each label and the similarity of the corresponding features between two labels

标签	t_1	t_2	t_3
特征数量	72	64	91
相似性	$J(t_1,t_2)=0.124$	$J(t_2,t_3)=0.033$	$J(t_3,t_4)=0.072$

3 结束语

本文设计了一种基于语音的帕金森病亚型识别方案。首先，针对每个帕金森病亚型选择出其特定

特征；接着，充分考虑帕金森病数据集中的相关性信息：二阶（2 个标签之间）标签相关性、高阶标签相关性以及特征空间与标签空间的相关性，进一步提升识别性能。实验结果表明，该方法在帕金森病亚型识别方面具有一定的优越性。此外，在实际使用时，该方案不仅能直接用于帕金森病亚型的自动识别，还能单独为每个亚型提供可靠的特定特征，实现亚型的个性化识别。然而，本文所提方案只考虑了亚型特定特征，未对所有亚型共享的特征进行讨论，这对研究帕金森病亚型之间的共性有一定的局限性。

参考文献

[1] ANDREW J L M D, HARDY J, REVESZ T F. Parkinson′s disease[J]. The Lancet, 2021, 373(9680): 2055-2066.

[2] LAU D L M, BRETELER M M. Epidemiology of Parkinson′s disease[J]. The Lancet Neurology, 2006, 5(6): 525-535.

[3] WARULE P, MISHRA S P, DEB S. Time-frequency analysis of speech signal using Chirplet transform for automatic diagnosis of Parkinson′s disease[J]. Biomedical Engineering Letters, 2023, 13(4): 613-623.

[4] OGAWA M, YANG Yiran. Residual - network - based deep learning for Parkinson′s disease classification using vocal datasets [C]//Proceedings of 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech). Piscataway, NJ: IEEE, 2021: 275-277.

[5] LITTLE M A, MCSHARRY P E, HUNTER E J, et al. Suitability of dysphonia measurements for telemonitoring of Parkinson′s disease[J]. IEEE Transactions on Biomedical Engineering, 2009, 56(4):1015-1022.

[6] TSANSA A, LITTLE M A, MCSHARRY P E, et al. Novel speech signal processing algorithms for high - accuracy classification of Parkinson′s disease[J]. IEEE Transactions on Biomedical Engineering, 2012, 59(5): 1264-1271.

[7] SHASTRY K. Ensemble machine learning regression model based predictive framework for Parkinson′s UPDRS motor score prediction from speech data[J]. International Journal of Speech Technology, 2023, 26(2): 433-457.

[8] TJADEN K. Speech and swallowing in Parkinson′s disease[J]. Topics in Geriatric Rehabilitation, 2008, 24(2): 115-126.

[9] EL-ZIAAT H, EL-BENDARY N, MOAWAD R. Using multi-feature fusion for detecting freezing of gait episodes in patients with Parkinson′s disease[C]//Proceedings of the International Conference on Innovative Trends in Communication and Computer Engineering. Piscataway, NJ: IEEE, 2020: 92-97.

[10] BROWN K A, SPENCER K A. The relationship between speech characteristics and motor subtypes of Parkinson′s disease[J]. American Journal of Speech-Language Pathology, 2020, 29(4): 2145-2154.

[11] BURK B R, WATTS C R. The effect of Parkinson disease tremor phenotype on cepstral peak prominence and transglottal airflow in vowels and speech[J]. Journal of Voice, 2019, 33(4): 580.

[12] SPYROMITROS E, TSOUMAKAS G, VLAHAVAS I. An empirical study of lazy multilabel classification algorithms [C]// Proceedings of Hellenic Conference on Artificial Intelligence. Cham: Springer, 2008: 401-406.

[13] ZHANG Minling, ZHOU Zhihua. ML-KNN: A lazy learning approach to multi-label learning[J]. Pattern Recognition, 2007, 40(7): 2038-2048.

[14] ZHANG Minling, WU Lei. Lift: Multi-label learning with label-specific features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(1): 107-120.

[15] HUANG Jun, LI Guorong, HUANG Qingming, et al. Learning label specific features for multi - label classification [C]// Proceedings of 2015 IEEE International Conference on Data Mining. Piscataway, NJ: IEEE, 2015: 181-190.

[16] YU Zebang, ZHANG Minling. Multi - label classification with label-specific feature generation: A wrapped approach[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(9): 5199-5210.

[17] WANG Y B, HANG J Y, ZHANG M L. Stable label-specific features generation for multi - label learning via mixture - based clustering ensemble [J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(7): 1248-1261.

[18] LIU Zhifeng, TANG Chuanjing, ABHADIOMHEN S E, et al. Robust label and feature space co - learning for multi - label classification[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(11): 11846-11859.

[19] ALEX R, ALESSANDRO L. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496.

[20] VÁSQUEZ-CORREA J C. DisVoice[EB/OL]. (2020-05-01). https://github.com/jcvasquezc/DisVoice.

[21] ZHANG Minling, ZHOU Zhihua. A review on multi - label learning algorithms[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(8): 1819-1837.