迈向客观的胃肠听诊：肠鸣音模式的自动分割与标注

Zahra Mansour[1, 3], **Verena Uslar**[2], **Dirk Weyhe**[2], **Danilo Hollosi**[3], **和 Nils Strodthoff**[1,*]

[1]奥尔登堡卡尔·冯·奥西茨基大学AI4Health部门、卫生服务研究系、医学与健康科学学院，26129 奥尔登堡，德国¨内脏外科大学诊所、医学与健康科学学院，奥尔登堡卡尔·冯·奥西茨基大学，26121 奥尔登堡，德国[2]弗劳恩霍夫IDMT，HSA研究所，26129 奥尔登堡，德国¨nils.strodthoff@uol.de

arXiv:2603.07215v1 [cs.SD] 7 Mar 2026

摘要

肠鸣音（BS）通常转瞬即逝且振幅较低，因此难以通过人工听诊准确检测。这导致临床评估存在显著差异。数字声学传感器能够采集高质量肠鸣音，并支持自动化信号分析，有望为临床医生提供关于肠道活动的客观、定量反馈。本研究介绍了一种使用可穿戴声学SonicGuard传感器进行肠鸣音分割与分类的自动化流程。

使用 SonicGuard 传感器记录了 83 名受试者的肠鸣音信号。临床专家对其中 40 名受试者的数据进行了人工标注，并用于训练自动标注算法；其余受试者的数据则用于进一步评估模型。研究开发了一种基于能量的事件检测算法，用于检测 BS事件。随后，使用预训练音频频谱图变换器（AST）模型将检测到的声音片段分类为 BS模式。模型性能分别针对健康个体和患者进行了评估。

最佳配置采用了两个专用模型，分别在健康受试者和患者的数据上训练；在健康组中的表现为（准确率：0.97，AUROC：0.98），在患者组中的表现为（准确率：0.96，AUROC：0.98）。自动标注方法将人工标注时间缩短了约 70%，专家审核显示，自动检测到的片段中需要更正的比例不到 12%。

所提出的自动分割与分类系统能够对肠道活动进行定量评估，为临床医生提供客观的诊断工具，有望改善胃肠功能的诊断，并支持大规模数据集的标注。

引言

肠鸣音（BS）是消化过程中腹壁平滑肌收缩以及气体和液体在肠道内移动所产生的零星声学事件。因此，肠鸣音反映了胃肠动力和整体消化功能[1]。腹部听诊最早于20世纪初被提出，作为评估胃肠活动的一种临床技术[2]，此后，临床医生一直致力于了解肠鸣音、胃肠生理学与疾病状态之间的关系[3]。在腹部检查中，医生通常将听诊器置于腹部四个象限，聆听数分钟，以判断是否存在肠鸣音，并据此判断是否有肠梗阻或术后肠麻痹。然而，这种评估具有主观性，且缺乏定量测量[1,4]。

肠鸣音的传统听诊受到肠鸣音信号本身间歇性特征的影响。与呈现节律性或连续模式的心音或肺音不同，肠鸣音以不可预测的间隔出现，因此短时间聆听时可能会漏听[5]。因此，临床医生往往需要在多个腹部象限聆听数分钟，才能检测到声学事件，这增加了临床工作量，也降低了繁忙医疗环境中的实用性[4]。此外，肠鸣音事件持续时间极短（从数毫秒到几秒），且能量较低，因此难以通过人耳可靠检测，使听诊具有很强的主观性[6]。这些特性降低了检测灵敏度和不同检查者之间的一致性，也凸显了对自动化定量方法的需求，以便在听诊之外提供客观且可重复的评估。

高保真数字听诊器和可穿戴声学传感器的出现，如今使连续录制肠道

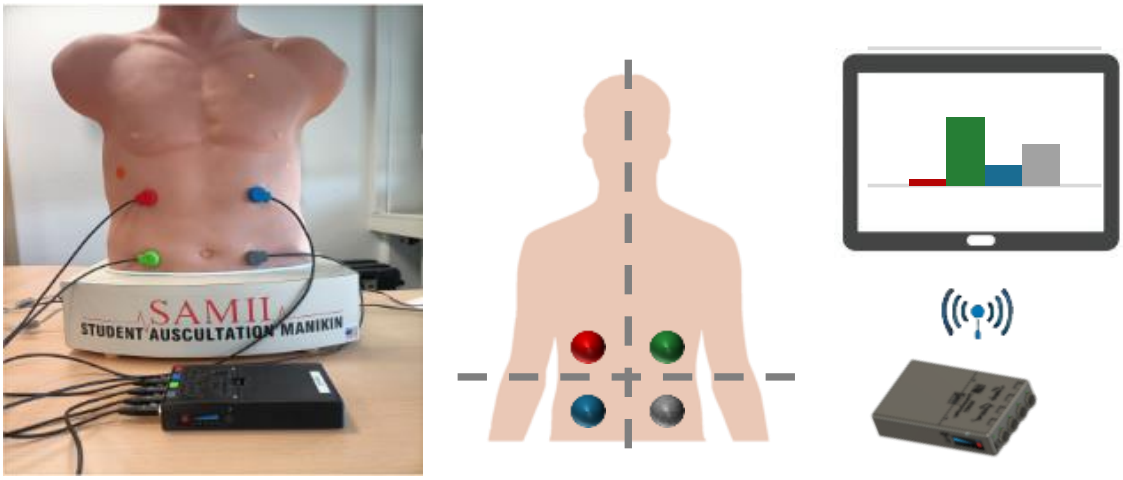

图 1. SonicGuard 可穿戴声学感知平台。在腹部四个象限（右上腹象限（RUQ）、左上腹象限（LUQ）、右下腹象限（RLQ）和左下腹象限（LLQ））放置四个传感器，以持续采集肠鸣音。采集到的信号会传输至处理平台，再由平台转发至配套移动应用程序进行自动分析并提供实时用户反馈。

声音成为可能，且录音质量足以支持计算分析[7,8]。这克服了传统听诊的局限，使研究人员能够应用自动化方法检测并表征肠鸣音活动。然而，由于大型标注数据集十分稀缺，该领域的进展仍受到限制，使肠鸣音分析比其他生理声学信号更具挑战性。自动化肠鸣音分析的研究已从基础信号处理发展到先进的机器学习和深度学习方法。早期研究主要聚焦于检测肠鸣音事件，使用滤波或时频表示从噪声中分离肠鸣音，例如基于小波的平稳滤波[9], 基于短时傅里叶变换的分析[10], 以及分形或非线性特征提取[11]。随着机器学习的发展，支持向量机和神经网络等经典模型被用于检测肠鸣音片段，并将其与伪迹区分开来[12]。近期研究转向深度学习：基于卷积神经网络的检测器如今能够识别连续录音中的肠鸣音事件，并已应用于行为或消化状态监测[13]一种较新的研究方向是肠鸣音模式分类，其目标不仅是检测肠鸣音事件，还要将其归类为具有生理学意义的类型。早期尝试包括小波神经网络[11]和基于频谱图的聚类[14]近期进展表明，研究人员采用预训练音频模型（如 Wav2Vec 和基于卷积神经网络的架构），以提高识别准确率并降低标注需求[6,15,16]。尽管取得了这些进展，将肠鸣音分类为多种临床相关模式的研究仍然有限[16]。正如上述综述的研究所指出的，现有研究主要侧重于检测肠鸣音片段，或将预定义片段分类为肠鸣音模式，参见例如[16]。然而，迄今尚无研究在单一统一框架内实现从肠鸣音听诊到模式定量分析的完整流程自动化。缺乏一个整合自动肠鸣音检测与模式级分类的端到端系统，构成了明显的研究空白。

在本研究中，我们基于此前开发的可穿戴声学传感器[8]，将其用于记录 80 多名受试者的肠鸣音（BS），其中包括健康个体和患者。利用这些录音，我们开发了一种自动分析框架，可检测肠鸣音活动并将事件归类为四种主要肠鸣音模式。我们使用专家审核的标注对算法进行评估，并将其性能与人工标注数据进行比较。研究结果凸显了此类自动分析在为临床医生提供有关胃肠活动的量化、客观反馈方面的潜力，从而支持他们在日常临床实践中做出更明智的决策。

材料与方法

数据集

录音方案 数据录音所依据的研究方案已获奥尔登堡大学医学伦理委员会批准（2022-056）。在参与研究前，研究人员向受试者说明了研究目的，并征得其同意，以匿名形式使用所收集的数据。在录音过程中，要求每位受试者采取仰卧位。随后，从四个腹部象限（右上腹象限（RUQ）、左上腹象限（LUQ）、右下腹象限（RLQ）、左下腹象限（LLQ））录制肠鸣音，具体如图 1所示，使用 SonicGuard

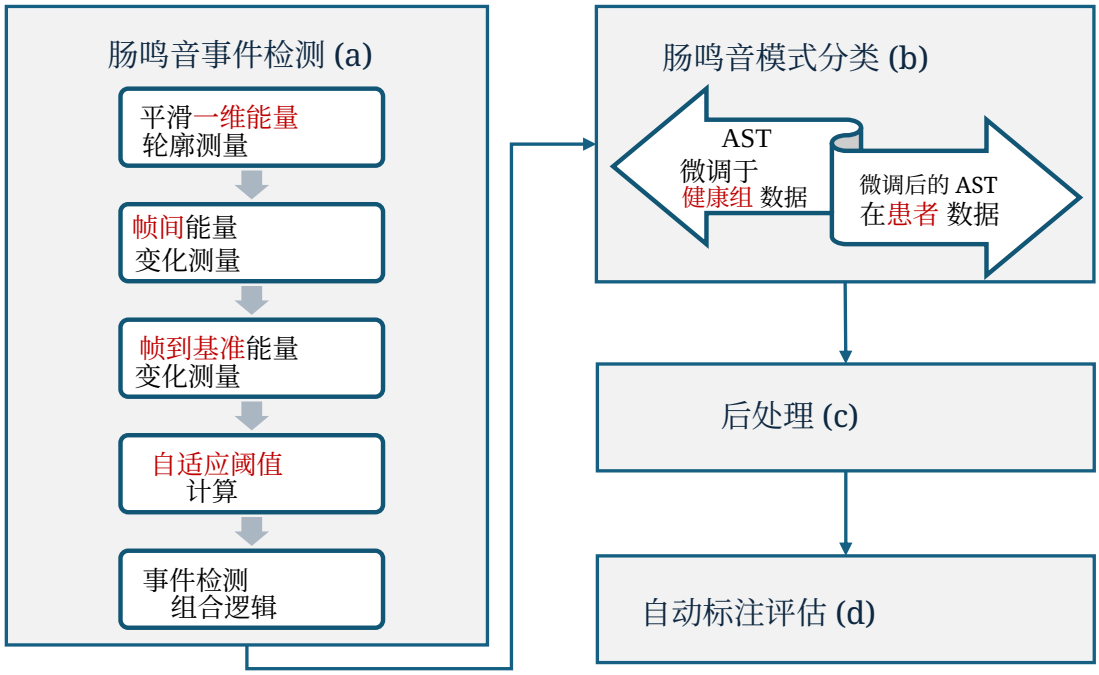

图 2. 自动标注算法的方法：首先通过测量能量变化检测肠鸣音事件，然后使用 AST模型进行肠鸣音模式分类，接着调整分段，最后将人工创建的标签与自动标签进行比较。

可穿戴多通道声学传感器[8]。每个象限录制7分钟的肠鸣音，每位受试者共录制28分钟。

受试者信息 在德国奥尔登堡的 PIUS医院，共采集了 84 名受试者的 40 多小时肠鸣音录音，其中包括 36 名确诊患有胃肠道疾病的患者和 48 名健康对照。受试者年龄为 20 至 83 岁，包括 40 名女性和 44 名男性，体重指数（BMI）范围为 18 至 40 kg/m^2。患者组包括结肠癌、克罗恩病和术后胃肠道疾病患者。

数据标注与肠鸣音模式 本研究分析的肠鸣音模式已在先前的研究中作了详细定义[16]。这些模式与腹部听诊和既往文献中常用的分类一致，涵盖了数据集中观察到的所有声音事件。各模式及其示例见图 3，总结如下：

单次爆发 (SB) 短促、孤立的脉冲（10–30 毫秒），可能由小肠收缩或肠腔内容物中液体的飞溅引起。

多次爆发 (MB) 由短暂且不规则的间隔分隔的多组单次爆发（总时长为 40–1500 毫秒），可能由液体在肠道内的移动产生。

连续随机声 (CRS) 连续、成簇的隆隆声（200–4000 毫秒），期间没有静音间隔，由含气肠腔内容物的移动，或肠内容物通过直径不一的肠道节段引起。

谐波声 (HS) 最不常见的模式，具有三至四个谐波频率成分（50–1500 毫秒），与狭窄有关。

来自40名受试者（18名患者和22名健康对照）的录音，由一名在医务人员监督下接受训练的成员使用 Audacity 软件，根据波形和声谱图特征以及听觉评估，手动标注为上述四种肠鸣音模式。其余43段录音用于评估自动标注算法。

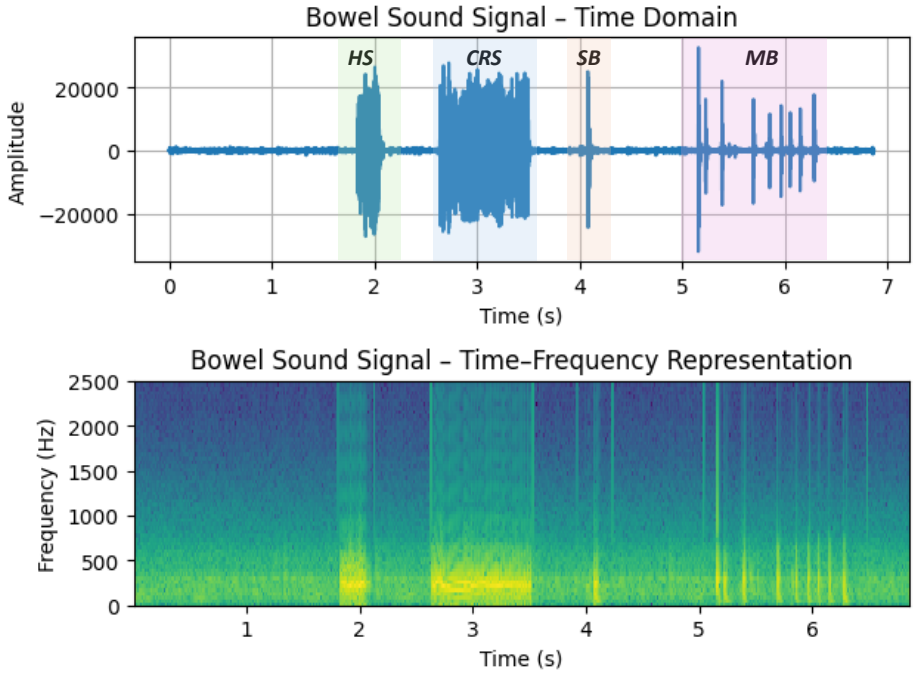

图 3. 本研究考察的四种肠鸣音模式在时域和频域中的表现，依次为谐波声（HS）、连续随机声（CRS）、单次爆发（SB）和多次爆发（MB）

方法

自动标注算法所采用的方法如图 2所示。下面，我们将详细介绍其各个主要组成部分。

*BS*事件检测

与心音等准周期性生物医学声学信号不同，肠鸣音（BS）具有显著的形态变异性，可分为上述四种不同模式。这些模式在时域和频谱特征上均存在差异，因此仅凭单一特征或判据进行事件检测颇具挑战。短时、脉冲式肠鸣音事件（SB 和 MB）难以通过传统的基于能量的方法检测，因为这类方法依赖相对于基线能量的变化；而仅依据帧级能量或 RMS 幅度的方法，可能无法充分捕捉时长较长、呈簇状的模式（CRS 和 HS）。为应对这些挑战，我们采用短时分析方法，使用时长为 1ms、互不重叠的帧，开发了一种定制的肠鸣音（BS）事件检测算法。

令$x_i[n]$表示i帧内的离散时间信号采样值，其中$n = 1,\ldots,N$，且N表示一帧中的采样点数。对于每一帧，提取用于表征帧内振幅和能量变化的时域特征，以识别候选事件的起始和结束边界。

帧级均方根（RMS）幅度和信号能量定义为

$$\mathrm{RMS}_i = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_i^2[n]}, \tag{1}$$

$$E_i = \sum_{n=1}^{N} x_i^2[n]. \tag{2}$$

为减少受试者间的差异并增强对背景噪声的鲁棒性，首先使用录音中频谱幅值的最大值作为全局参考，将能量表示转换为分贝刻度，从而确保整个信号的归一化保持一致。随后，对各频率区间求平均并进行时间平滑，得到帧级能量曲线。基线定义为平滑后能量曲线的中位数，代表用于归一化的典型非事件活动水平。归一化均方根和归一化能量计算如下

$$\mathrm{RMS}_i^{\mathrm{norm}} = \frac{\mathrm{RMS}_i}{\overline{\mathrm{RMS}}_{\mathrm{baseline}}}, \tag{3}$$

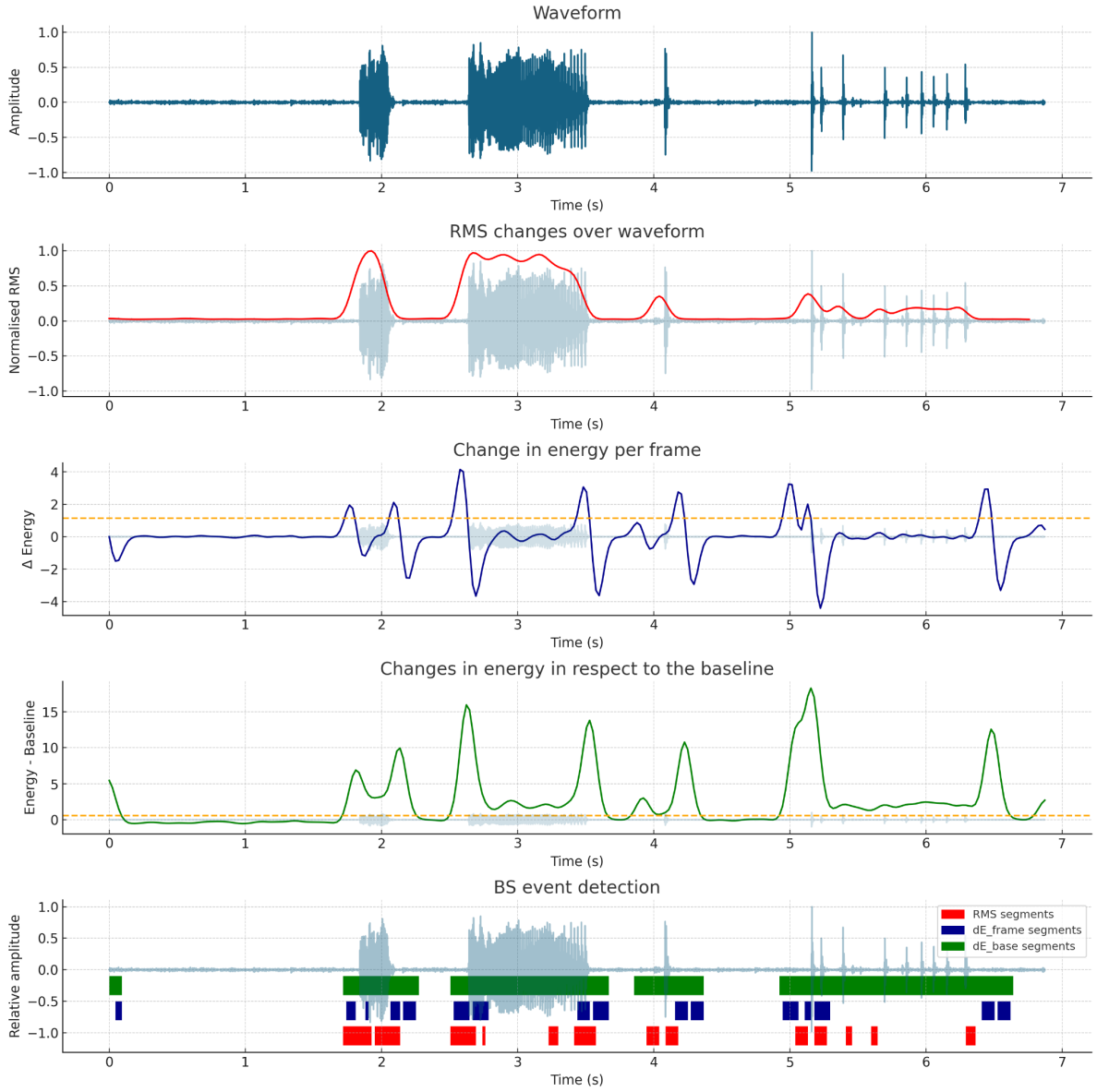

图 4. 所提出的肠鸣音（BS）事件检测特征示意图。第一行显示时域中的肠鸣音波形。第二行展示沿波形变化的均方根（RMS）振幅随时间的变化及其对应的自适应阈值。第三行描绘每个短帧内信号能量的变化及其自适应阈值。第四行显示帧能量相对于信号基线能量的变化及相应的自适应阈值。最后一行展示使用不同特征检测到的肠鸣音事件：基于RMS的检测（红色）、帧内能量变化（蓝色）以及相对于基线能量的能量变化（绿色）。

$$E_i^{\text{norm}} = \frac{E_i}{\overline{E}_{\text{baseline}}}, \tag{4}$$

其中，$\overline{\text{RMS}}_{\text{baseline}}$ 和 $\overline{E}_{\text{baseline}}$ 分别表示在基线段上计算得到的均方根和能量的平均值。

与此同时，通过分析连续帧之间归一化帧级能量的时间变化，来表征持续声学活动。BS检测采用基于RMS幅度和能量变化标准的联合判决策略。具体而言，当归一化RMS值和帧间能量差均超过预定义阈值时，即定义为新片段的起始点；该阈值设定为各自逐帧分布的中位数。为

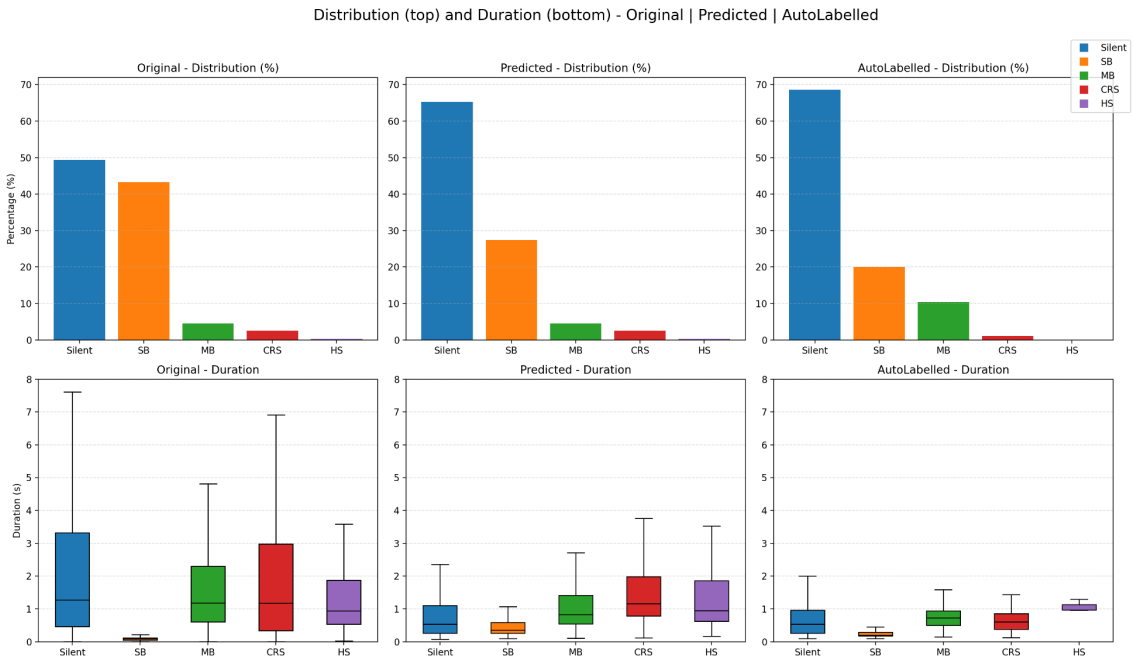

图5. 肠鸣音模式在分布（第一行）和时长（第二行）上的差异，比较对象包括：经人工标注且标题为原始的数据、同一数据若采用自动标注技术标注且标题为预测的结果，以及我们没有人工标签、而使用自动标注算法生成标签并标注为自动标注的患者组；所使用的模型仅在健康数据上训练，并同样在健康数据上测试。

确保时间连续性并防止错误分段（尤其是在帧内能量波动可能产生误导的簇状肠鸣音模式中），只要相对于基线的能量仍高于在整个记录中确定的同一基于中位数的阈值，就持续保留该片段。仅当三个参数（RMS、帧间能量差以及相对于基线的能量）全部低于该阈值时，才判定片段结束。在统一的检测框架中联合使用帧内振幅动态和帧间能量连续性，能够可靠地检测具有异质时间模式的BS事件。

肠鸣音模式分类

在检测到包含肠鸣音的信号片段后，采用专门的分类阶段，以在四种预定义模式中确定肠鸣音类型。近期研究结果[16]表明，预训练深度学习模型借助从大规模音频语料库中学习到的可迁移声学表征，在肠鸣音模式分类中表现更优。基于这一证据，本研究评估了两种先进的预训练模型：Wav2Vec模型和音频频谱图变换器（AST）。

音频频谱图变换器（AST）分类器[17]对二维对数梅尔频谱图进行处理，将每个声谱图划分为互不重叠的图块，这些图块经线性嵌入后，由带有自注意力机制的变换器架构处理，使每个时频图块能够关注声谱图中的所有其他图块，并通过位置嵌入保持时间顺序。该网络使用在AudioSet数据集上预训练的权重进行初始化[18],，从而在针对特定任务微调之前提取通用且稳健的音频表征。

此外，我们还评估了基于Wav2Vec的分类器在肠鸣音模式分类中的表现[19]。与AST不同，Wav2Vec直接处理原始音频波形，并通过自监督预训练学习分层潜在语音与音频表征。该模型无需依赖显式时频变换，即可捕捉细粒度时间信息和长上下文依赖关系，因此非常适合处理肠鸣音这类脉冲性、非平稳声学事件。我们使用检测到的肠鸣音片段对预训练的Wav2Vec编码器进行端到端微调，使其学习到的表征适应肠鸣音模式分类任务。对于每种预训练架构，均微调了三个队列特异性模型：(i) 仅健康组，(ii) 仅患者组，以及 (iii) 健康组+患者组混合模型。所有模型均采用相同的预处理方法（AST使用128频带对数梅尔频谱图，窗长为25 ms、跳长为10 ms；Wav2Vec使用原始波形作为输入），在适用情况下进行逐频带均值-方差归一化，并使用相同的训练集、验证集和测试集划分方案（按受试者分层，比例为70%、15%、

表1. Wav2Vec 2.0 与 AST 模型的性能比较，采用准确率（ACC）和曲线下面积（AUROC）进行评估。

训练集	Test	Wav2Vec 2.0		AST	
		ACC	AUROC	ACC	AUROC
健康组	**健康组**	0.82	0.74	**0.97**	**0.98**
	患者	0.77	0.74	0.76	0.82
	All	0.79	0.73	0.87	0.94
患者	健康组	0.67	0.61	0.91	0.90
	患者	0.75	0.69	**0.96**	**0.98**
	All	0.73	0.66	0.93	0.95
All	健康组	0.73	0.80	0.94	0.93
	患者	0.75	0.81	0.94	0.96
	All	0.77	0.83	**0.94**	**0.98**

和15%），以及匹配的超参数。这种受控设置确保性能差异反映的是队列相关和领域特异性效应，而非训练变异性。

我们使用受试者工作特征曲线下面积（AUROC）量化模型性能。AUROC是一项不受阈值影响的分类判别能力指标[20]。每个模型均在三个测试集（仅健康组、仅患者组和混合组）上进行评估，以考察域内性能、跨域泛化能力以及对队列偏移的稳健性。

后处理

完成分类后，流程会执行时间细化阶段。系统按时间顺序排列检测到的事件；当相邻事件之间的间隔超过最小时长阈值（100毫秒）时，可选择用默认标签（非BS片段）填补间隔。随后，系统执行后处理合并程序，将两个或多个时间相邻且具有相同预测标签的片段合并为一个片段。此步骤可缓解帧级波动造成的连续事件被人为拆分为较短片段的问题，从而生成更长且具有临床可解释性的序列，以表示一致的BS活动。

评估

本研究使用三个互补的数据组评估所提出的自动标注算法的性能，旨在考察其准确性、可用性和统计一致性。

首先，使用来自 12 名未曾见过的受试者的人工标注数据作为参考标准，对自动标注性能进行定量评估。根据事件计数、时间分布和事件持续时间，将自动生成的标签与人工标注进行比较，从而直接评估这两种标注方法之间的一致性。

其次，使用来自其余 35 名受试者的数据，在缺乏人工真实值的情况下，评估自动标注标签的生理合理性和一致性。将从自动标注的肠鸣音事件中提取的统计特征（如事件持续时间、事件间间隔和时间分布）与正常标注的肠鸣音数据所报告的相应特征进行比较，以验证生成的标签是否符合预期模式。

最后，使用来自 10名受试者的数据，评估自动标注系统在半自动标注工作流程中的实用性。对于这组数据，首先生成自动标签，然后由专家标注员进行审核。记录所需的标注时间以及对自动生成的标签进行的人工调整次数，以衡量标注效率和专家工作量的减少。

结果与讨论

BS事件检测

由于BS事件检测算法基于多种特征的组合，图4 展示了每种特征如何捕捉不同BS模式的不同方面。例如，基于RMS的特征能够有效检测肠鸣音事件的起始和结束，尤其适用于SB和MB等峰状模式。然而，对于簇状事件（MB和CRS）的连续部分，该特征难以检测，因为这些信号在较长时段内保持相对恒定的振幅，导致基于RMS的方法错误地判断事件已经终止。依赖于

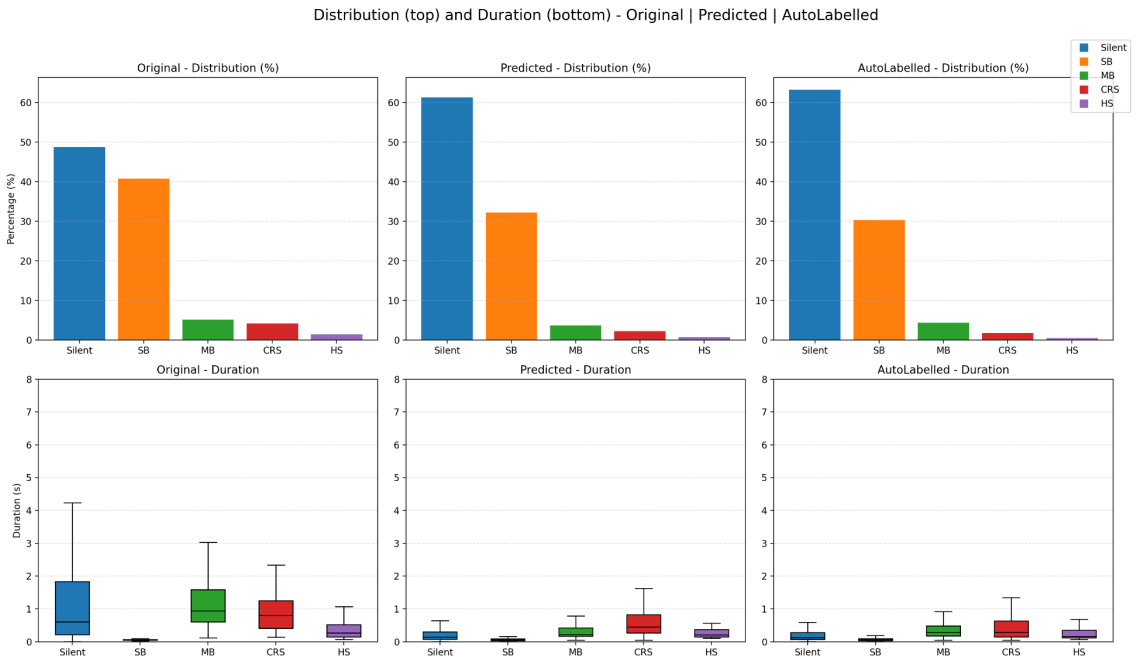

图 6. 真实标签（原始；左图）、同一组样本的模型预测结果（预测；中图）与自动标注样本（自动标注；右图）中肠鸣音模式的归一化计数（第一行）和时长（第二行）差异。

单帧内能量变化的算法也存在类似困难：它们或许能够识别事件边界，却无法捕捉持续事件的完整时间范围。相比之下，将基于RMS的检测、帧内能量变化以及相对于基线的能量变化等特征结合起来，不仅能够准确检测事件的起始和结束，还能可靠地将整个肠鸣音事件识别为一次连续发生。

肠鸣音模式分类

评估了两种用于肠鸣音（BS）模式分类的预训练模型：Wav2Vec 2.0 和音频频谱图变换器（AST）。每种模型均使用三个不同的训练子组（健康组、患者和全部）进行微调。分类性能以准确率（ACC）和ROC曲线下面积（AUROC）表示，汇总于表1。

对于两种架构，在每个训练子组中，模型在来自同一队列的测试数据上评估时均取得最高性能。具体而言，使用健康数据微调的模型在健康队列上测试时表现最佳，采用 AST 时尤为如此（ACC = 0.97，AUROC = 0.98）。同样，使用患者数据微调的模型在患者测试集上评估时表现最佳，采用 AST 时亦是如此（ACC = 0.96，AUROC = 0.98）。对于使用合并数据集微调的模型，在混合测试集上采用 AST 获得了最佳结果（ACC = 0.94，AUROC = 0.98）。

这些结果表明，从健康受试者和患者身上记录的肠鸣音在形态和结构上的差异会直接影响分类性能。这种对队列的依赖促使我们在自动标注框架中采用针对不同队列的分类器，并根据输入数据来自健康受试者还是患者来选择模型。另一方面，在合并队列上训练的 AST模型在两个子队列上仍表现良好。

在所有训练集和测试集子组中，AST 的表现始终优于 Wav2Vec 2.0。因此，AST 被选为后续实验的默认分类器。

肠鸣音自动标注框架

我们通过两个示例用例展示自动标注框架可能的应用场景。

用例 ***1***：自动提取肠鸣音模式特征

为评估所提出的肠鸣音自动标注框架的性能，我们区分了三个子组。原始(org)组包含来自40名受试者的人工标注肠鸣音标签，并作为参考标准。预测 (pred)组包含肠鸣音自动标注框架为同一批经过人工标注的录音自动生成的标签，以便直接比较人工标注与自动标注。自动组代表完全

表2. 自动生成与专家调整的肠鸣音标签按类别比较（10名受试者和全部象限的总体平均值）。

类别	Auto 数量	专家 数量	平均持续时间 自动（秒）	平均持续时间 专家（秒）
None	929.7	825.3	0.278	0.200
SB	400.4	314.9	0.084	0.101
MB	42.4	61.4	0.415	0.525
CRS	21.3	27.2	0.531	0.625
HS	7.3	3.0	0.740	0.307

从未进行人工标注的录音中获得的自动生成标签（43名受试者），用于评估该框架在未标注数据上的表现。结果分别针对健康受试者（图 5）和患者受试者（图 6）单独呈现，以突出不同队列在肠鸣音分布和时间特征方面的差异。健康受试者数据集的原始类别分布与患者队列略有不同，其中高音调声样本明显较少。这一现象符合预期，因为高音调声模式更常见于肠狭窄等病理状况（图 5）。占比最高的始终是 无 类别（原始：49%，预测：65%，自动：68%），其次是单次爆发（原始：43%，预测：28%，自动：20%）、混合肠鸣音（原始：5%，预测：5%，自动：10%）、CRS（原始：2.9%，预测：2.9%，自动：1.99%），占比最低的类别是高音调声（原始：0.1%，预测：0.1%，自动：0.01%）。与患者数据集类似，自动标注框架保留了类别占比的正确排序。

此外，与患者数据相比，健康受试者的肠鸣音模式通常时长更长，尤其是混合肠鸣音、CRS和高音调声事件。尽管自动标注技术总体上略微低估了肠鸣音事件的时长，但它成功捕捉到了这一事件持续时间的增加。对于患者数据集，所提出的肠鸣音自动标注框架正确捕捉了所有类别中肠鸣音模式的相对分布。出现频率最高的始终是无 类别（原始：49%，预测：61%，自动：63.5%），其次是单次爆发（原始：40%，预测：32%，自动：30%）、混合肠鸣音（原始：6%，预测：4%，自动：3.5%）、CRS（原始：4%，预测：2.5%，自动：2.6%），以及最罕见的高音调声（原始：1%，预测：0.5%，自动：0.4%）。重要的是，自动标注框架保留了各类别出现频率的正确排序。

自动标注算法还识别出了多个模式，其肠鸣音事件的持续时间范围与人工标注一致。尤其是持续时间最短的模式——单次爆发（最大时长 < 0.2 s）、CRS（最大时长 ≈ 2s）和高音调声（最大时长 $\approx$ 1s）——均被准确检测到。然而，该框架倾向于在混合肠鸣音事件实际结束前将其截断，导致该类别检测到的持续时间系统性偏短（原始：3 s；预测：0。9s；自动：1 s）。

在健康组和患者数据集中，人工标注与自动生成的标签之间差异最显著的是无 类别。这种差异源于该框架倾向于将肠鸣音事件之间极短的静默间隔明确标记出来，并将较长的静默时段分割为多个较短的无 片段。因此，与原始人工标注相比，自动标注数据中的无 类别分布增加，平均持续时间则有所缩短。

先前的评估表明，肠鸣音自动标注算法能够为医学专家提供肠鸣音信号的初步特征描述，并支持更定量的评估。借助该算法，临床医生可以快速测量胃肠动力偏离正常水平的程度，并监测其随时间的变化。这类定量信息对于追踪疾病进展、评估治疗反应，以及支持术后或重症监护评估尤其有用。总体而言，肠鸣音定量分析有助于以更标准化、客观的方式解读听诊结果，并促进其融入临床工作流程。

用例 *2*：人在回路的肠鸣音标注
肠鸣音自动标注框架用于标注 10名受试者（40 份记录）的肠鸣音，随后由标注专家审核自动生成的标签并进行调整，所需的调整表明专家干预有限。平均而言，约 12% 的检测到的事件被删除或合并，而平均事件时长的差异小于 15 毫秒，具体总结见表 2。类别分析证实，肠鸣音模式的相对分布得以保留，修正主要影响事件边界以及多次爆发事件的时长。这些结果表明，所提出的自动标注框架能够生成接近专家质量的标签，同时将人工标注时间大幅缩短 70%。

带标签的肠鸣音数据集的创建受到数据可用性有限和人工标注成本高昂的制约。人工标注需要对长时间录音进行帧级检查，因此是数据集开发中最耗时的步骤之一

构建。因而，将自动预标注与专家校正相结合，可简化大型数据集的标注，为扩大数据集开发规模提供高效策略。

结论

本研究提出了一种自动化流程，利用可穿戴声学传感器采集的录音进行肠鸣音分割和模式分类。该框架将基于能量的事件检测方法与预训练深度学习模型相结合，并采用队列特异性训练，以考虑健康受试者与患者之间的差异。

结果表明，所提出的算法支持多种用例，包括实现肠鸣音的自动定量分析，以及在数据集创建过程中协助人工标注员。这些互补的用例共同应对临床解读和构建高质量肠鸣音数据集这两方面的挑战，从而支持进一步的数据驱动分析，不再局限于预定义的事件特征，而是深入理解肠鸣音及其与疾病的关系。

参考文献

1. Breum, B. M., Rud, B., Kirkegaard, T. & Nordentoft, T. 腹部听诊对肠梗阻的诊断准确性。《世界胃肠病学杂志》21, 10018–10024, DOI: 10.3748/wjg.v21.i34.10018 (2015)。2. Cannon, W. B. 胃和肠道产生的节律性声音的听诊。《美国生理学杂志：法定内容版》14, 339–353, DOI: 10.1152/ajplegacy.1905.14.4.339 (1905)。3. Talley, N. J. & O'Connor, S. 《临床检查：体格诊断系统指南》（爱思唯尔健康科学，阿姆斯特丹，荷兰，2013年），第7版。4. Koenig, S. J., Loffroy, R., Keller, M. S. & Jaffe, T. A. 肠鸣音听诊：术后肠麻痹的标志？《外科研究杂志》171, 502–507, DOI: 10.1016/j.jss.2011.06.018 (2011)。5. Baronetto, A., Fischer, S., Neurath, M. F. & Amft, O. 可穿戴监测系统连续音频数据中高度不平衡的多尺度肠鸣音事件检测。JMIR AI, 《数字健康》1, e51118, DOI: 10.2196/51118 (2024年)。6. Yu, Y., Zhang, M., Xie, Z. & Liu, Q. 通过自注意力和自监督预训练增强肠鸣音识别。《公共科学图书馆·综合》19, e0311503, DOI: 10.1371/journal.pone.0311503 (2024年)。7. Choudry, M., Stead, T. S., Mangal, R. K., Ganti, L. & Mangal, R. 听诊器的历史与演变。Cureus 14 (2022年)。8. Mansour, Z., Uslar, V., Weyhe, D., Hollosi, D. & Strodthoff, N. Sonicguard传感器——一种用于长期监测腹部声音的多通道声学传感器：资格验证研究。《传感器》24, DOI: 10.3390/s24061843 (2024年)。9. Hadjileontiadis, L. J. & Panas, S. M. 利用四阶统计量从肺音中自适应去除心音。《IEEE生物医学工程汇刊》46, 1185–1195, DOI: 10.1109/10.802884 (1999)。10. Ficek, J. N., Bani Mohammad, M., Wong, S. H. & Laskowski, E. R. 健康志愿者肠鸣音的计算机化声学分析。《科学报告》11, 16160, DOI: 10.1038/s41598-021-95442-z (2021年)。11. Dimoulas, C. A., Kalliris, G. M., Chatzis, G. N. & Kaloumenos, N. S. 使用小波和神经网络进行肠鸣音模式分析。《IEEE生物医学工程汇刊》54, 1393–1402, DOI: 10.1109/TBME.2007.900816 (2007)。12. Yin, J., Chen, Z., Liu, L., Lu, S. & Wan, M. 结合小波变换与机器学习自动识别肠鸣音。《医学工程与技术杂志》42, 546–554, DOI: 10.1080/03091902.2018.1530439 (2018)。13. Wang, Z., Li, X. & Fang, D. 利用卷积神经网络深度学习实现胃肠音的实时检测。载于《IEEE生物信息学与生物医学国际会议论文集》（BIBM），3003–3010, DOI: 10.1109/BIBM55620.2022.9995841 (2022年)。14. Du, X., Chen, T., Yu, Y. & Hu, Q. 通过声谱图聚类进行无监督肠鸣音分类。《IEEE生物医学与健康信息学杂志》22, 1001–1011, DOI: 10.1109/JBHI.2017.2703878 (2018)。15. Baronetto, A., Di Mauro, D., Giordano, D., Aricò, P. & Mattia, D. 用于肠鸣音分析与分类的多尺度卷积神经网络架构。载于《第45届IEEE医学与生物学工程学会年度国际会议》（EMBC），5121–5125, DOI: 10.1109/EMBC53107.2024.10658056 (2024年)。16. Mansour, Z., Uslar, V. N., Weyhe, D., Hollosi, D. & Strodthoff, N. 肠鸣音模式分类的机器学习基准评测——从表格特征到预训练模型。《公共科学图书馆·综合》21, 1–16, DOI: 10.1371/journal.pone.0338911 (2026)。

17. Gong, Y., Chung, Y.-A. & Glass, J. AST：音频频谱图变换器。载于《*Interspeech*会议论文集》(2021年)。**18.** Gemmeke, J. F. & et al. AudioSet：用于音频事件的本体与人工标注数据集。载于《*ICASSP*会议论文集》(2017年)。**19.** Baevski, A., Zhou, Y., Mohamed, A. & Auli, M. Wav2Vec 2.0：语音表示自监督学习框架。载于《第*34*届神经信息处理系统国际会议论文集》（*NeurIPS*），12449–12460 (2020年)。**20.** McDermott, M. B. A., Zhang, H., Hyldig Hansen, L., Angelotti, G. & Gallifant, J. 类别不平衡情况下对AUROC和PR曲线下面积的深入探讨。*arXiv*预印本 *arXiv:2401.06091* DOI：10.48550/arXiv.2401.06091 (2024年)。

致谢

作者感谢奥尔登堡卡尔·冯·奥西茨基大学第六学院（研究池、移动健康潜力领域）以及下萨克森州科学与文化部（通过弗劳恩霍夫IDMT HSA研究所的互联健康项目）提供的校内经费支持。

数据可用性

四名受试者同意公开其数据。数据集可通过以下链接获取：`https://doi.org/10.6084/m9.figshare.28595741.v1`。我们方法的完整实现以及训练脚本，可在相应的代码仓库中获取：`https://github.com/AI4HealthUOL/Bowel-sound-auto-annotations`