Non-invasive Detection of Bowel Sounds in Real-life Settings Using Spectrogram Zeros and Autoencoding

I. Bilionis, G. Apostolidis, V. Charisis, C. Liatsos, L. Hadjileontiadis

Abstract— **Gastrointestinal (GI) diseases are amongst the most painful and dangerous clinical cases, due to inefficient recognition of symptoms and thus, lack of early-diagnostic tools. The analysis of bowel sounds (BS) has been fundamental for GI diseases, however their long-term recordings require technical and clinical resources along with the patient´s motionless concurrence throughout the auscultation procedure. In this study, an end-to-end non-invasive solution is proposed to detect BS in real-life settings utilizing a smart-belt apparatus along with advanced signal processing and deep neural network algorithms. Thus, high rate of BS identification and separation from other domestic and urban sounds have been achieved over the realization of an experiment where BS recordings were collected and analyzed out of 10 student volunteers.**

Clinical relevance— **Precise separation of BS from both stationary and non-stationary background noise allows for reliable long-term remote monitoring of bowel function; hence, timely diagnosis or prognosis of GI diseases.**

I. INTRODUCTION

Gastroenterology defines bowel sounds (BS) as the sounds detected in the abdominal region of human body and generated by the movement of gases and liquids during intestinal peristalsis. Gastrointestinal physiology has been studied for decades because of the significant inherited information related to various gastrointestinal (GI) diseases. To date, GI diagnostic tools consist of invasive medical procedures such as computerized topography scan, manometry, biopsy and endoscopy secondary effects of which pose high risk of bleeding, infection, anaesthetic complications, physical discomfort and psychological distress.

Hence, BS investigation has introduced promising non-invasive solutions through auscultation of the abdominal area in combination with signal processing techniques. First and foremost, Cannon [1] auscultated, examined methodically the intestinal area and linked each type of sound to the corresponding GI process. In addition, in favor of the rapid technological advances, the scientific interest has been focused on spectral analysis of BS [2], [3] by applying complex signal processing methods, high order statistics [4]-[8] and by implementing machine learning models [9]-[13] in order to extract the most clinically important, information. Accordingly, other authors linked BS with specific GI diseases such as pyloric stenosis [14], irritable bowel syndrome, Crohn's and ileus [15]-[19]. However, in order for these diseases to be diagnosed effectively, a long-term monitoring of intestinal functionality is required [10] while a silent background increases BS reliability.

This study proposes an end-to-end non-invasive tool for BS detection in adverse background environments based on multi-channel recordings of the abdominal region, and is composed of three fundamental parts, (a) stationary noise suppression, (b) artifacts' elimination taking into account the correlation between the recording channels, (c) comparison of extracted sounds with non-stationary signals derived from domestic and urban environments by an autoencoder-based deep learning model.

II. METHODOLOGY

This survey's methodology is based on four fundamental parts where each one filters different unwanted signal components. At first, part A contains BS-oriented pre-processing of the raw recording channels, part B removes the stationary background noise via a time-frequency signal processing technique, then part C detects BS by investigating the correlation of signals from different channel and finally, part D separates BS from environmental non-stationary noise through a deep autoencoder deep neural network.

A. Pre-processing of Raw Recordings

The first phase consists of a 16^{th} order Butterworth high pass filter with cut-off frequency at 90Hz in order to avoid interference with heart sounds. In addition, short or weak sound-segments with duration less than 15ms or amplitude less than 0.005mV, respectively, are removed. As final pre-processing step, consecutive sound segments that are separated by less than 10ms of silence are considered a single event.

B. Stationary Noise Removal

This stage aims at removing stationary background noise. Although this is a common task in signal processing, in this work a more advanced technique is applied being adapted to BS signal characteristics. This technique is a time-frequency (TF) filtering based on spectrogram zeros proposed by Flandrin in [20]. Specifically, the zero values of the spectrogram form triangles which can be clustered to different regions in the TF plane where each cluster indicates a component of the

*This work has received funding from the European Union's Horizon 2020 Research and Innovation Action programme under grant agreement No 817732.

I. Bilionis, G. Apostolidis and V. Charisis are with School of Electrical & Computer Eng., Aristotle University of Thessaloniki, Greece

C. Liatsos is with Department of Gastroenterology, Pathology, and Surgery, 401 Army General Hospital of Athens, Athens, Greece

L. Hadjileontiadis is with Department of Electrical & Computer Eng., Khalifa University, UAE and with School of Electrical & Computer Eng., Aristotle University of Thessaloniki, Greece (emails leontios.hadjileontiadis@ku.ac.ae, leontios@auth.gr)

signal. Moreover, Flandrin proved that in presence of white Gaussian noise the distribution of edge lengths of Delaunay triangles appears to be normal, differentiating it from the corresponding distribution of a signal.

The stationary noise removal procedure begins with the computation of the signal's short-time Fourier transform (STFT) and extracting the locations of the zero values. As follows, triangles are formed from the zero points' coordinates applying Delaunay triangulation. Next, depending on each triangle's largest edge length value, the corresponding area is considered to represent either a signal component or white noise. Neighboring "signal" triangles are clustered together forming an individual component. These areas of clusters define masks which when multiplied with the initial STFT result the TF representation of the respective component. Lastly, the signal component reconstructed in time domain applying inverse STFT to masked STFT representation. The different stages of the stationary noise removal based on spectrograms zeros procedure is illustrated in Fig.1.

C. Multi-channel Bowel Sound Detection (MBSD)

Except for stationary noise, there can be non-stationary noise produced by external sources e.g. urban noises, motion artifacts and inferences with other biomedical sounds. Given that one microphone of the smart-belt is unattached recording the environment (see Section III), non-stationary external noisy sources mostly have higher amplitude in the unattached channel than the attached ones. Hence, this property is utilized in order to detect and remove unwanted external signal components. The MBSD procedure implements the extraction of the highest correlated sounds amongst data channels, proposing, thus, an inter-channel type of signal detection which is organized as follows.

1) Identification of two similar, in terms of the cross-correlation measure, sound-segments recorded by distinct microphones and consideration of that pair as siblings-sounds. By examining the similarity of all segments from different channels, the final set of siblings is created.
2) Elimination of every subset of siblings in which the cross-correlation of its two component sounds is inferior to an empirical threshold value of 0.2.
3) Deletion of noise-correlated segments by observation of cross-correlation values between siblings and noise channel. Likewise, siblings-sounds are removed if (a) a sound event takes place at the same time in the noise channel, (b) is highly correlated at with a sibling pair and (c) is twice more powerful than these, respectively.
4) Single channel construction through concatenation of the cleanest and loudest sound of each pair of siblings.

According to the aforementioned procedure, multi-channel BS are processed in such a way that the most correlated sounds, which indicate intestinal motility, are selected. Thereby, a single-channel output may enable clinicians to validate and label the extracted sounds effortlessly.

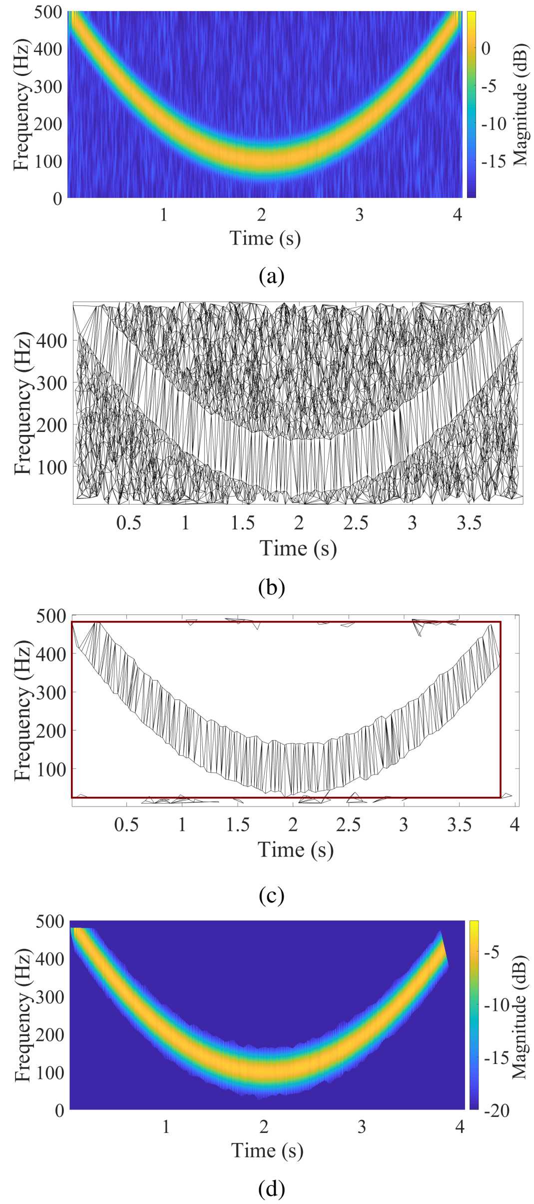


Fig. 1: Stationary noise removal based on the spectogram zeros applied on a noisy quadratic chirp: (a) STFT magnitude, (b) Delaunay triangles whose vertices coincide with zero values, (c) selected domain with large edge length, (d) masked STFT.

D. Bowel Sound Detection

The last part of methodology is dedicated to differentiating the previous extracted sounds from other non-stationary signals with similar to BS TF features, i.e. duration, frequency range, spectral centroid, generated either by an internal, e.g. respiratory sounds, cardiac signals and/or external mechanism e.g. urban noise, human speech, domestic sounds. Thus, this work employs an autoencoder deep neural network approach in order for the unwanted components to be distinguished from those containing BS.

The autoencoders are artificial neural networks (ANN)

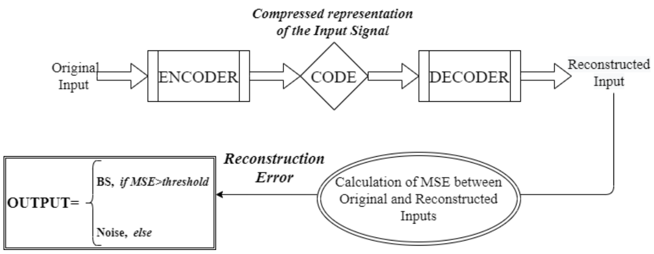


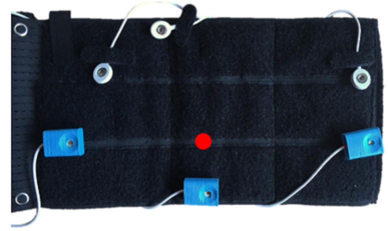
Fig. 2: Schematic representation of the data flow and involved processes.

that are trained to accomplish optimized reconstruction of the input data to its output. These ANN models are met in different architectures and they have numerous applications such as dimensionality reduction, denoising, information retrieval and anomaly detection [21]-[23], among others. An autoencoder, typically, is composed of three layers, (a) input layer or encoder, (b) hidden layer or code and (c) the output layer or decoder whose dimensions coincide with the encoder's ones while the size of the code depends on the application and the available data. This paper, implements an autoencoder for detection with equal size of encoder, code and decoder layers, which is created with MATLAB. The training set consists of the filtered BS dataset (positive label class) only whereas the validation set consists of both BS and non-BS sounds (see Section III). The objective of this deep learning method is to clarify whether a signal is BS or not, depending on the reconstruction error of the autoencoder upon its attempt to copy the input to its output. The mean squared error (MSE) function has been used as reconstruction error and the optimal threshold defines the classification rule between the BS and non-BS cases as can be seen in Fig.2. The model's evaluation is accomplished by ROC (Receiver Operating Characteristic) analysis which demonstrates the succession rate and the capabilities of the proposed solution.

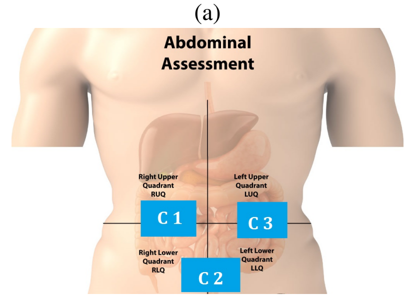
III. Experimental Data

This section describes the methodology's application to a real-life BS detection problem utilizing two datasets, depicting the BS and non-BS cases, respectively. The first dataset is comprised of BS data that were collected from ten healthy volunteers (aged 20-25 years old) and recorded (one 10-minute-long recording per participant) in an anechoic chamber of Aristotle University of Thessaloniki. The sampling frequency was set to 1 kHz since literature indicates maximum frequencies of the abdominal lower than 500Hz [3],[6],[14].

The BS were captured via a prototype smart-belt constructed by PLUX wireless biosignals S.A. (Fig.3) and is equipped with four microphones that three of them were attached to the abdominal area (data channels) of each participant, while the fourth one was unattached, monitoring the background environment (noise channel). The volunteers applied a fasting protocol by not consuming any food in less than 6 hours and not drinking any liquids in less than 2 hours



(a)



(b)

Fig. 3: Instrumentation for abdominal auscultation. a) Smart-belt with four embedded microphones. b) Attachment area of the corresponding microphone (https://nursecepts.com).

before the experiment. After the recording, in order to get ground-truth, the recorded signals were examined by a gastroenterologist who annotated the areas that contain medical information. In addition, the non-BS case was synthesized by non-stationary signals recorded in both domestic and urban environments (human speech, traffic, nature, television, kithcen, domestic animals, etc.) and obtained by public data web repositories (kaggle and VoxForge) [24]-[26] that were resampled to 1kHz as for the data to be consistent with the smart-belt's sampling frequency. In both BS and non-BS cases, each group of data consists of 2440 sound signals with 5 seconds maximum duration. Overall, a 4880-size dataset has been utilized for this study as the combination of two subsets (BS and non-BS). Table 1 shows the total dataset's division into training, validation and testing sets along with the corresponding sample sizes in order to fit the autoencoder model.

Datasets	BS	Non-BS
Training	1952	0
Testing	0	1952
Validation	488	488
Total	2440	2440

TABLE I: Data Splitting

IV. Results and Discussion

This section presents the experimental results produced by applying the proposed methodology to the aforementioned dataset. At first, Fig.4 and Fig.5 show the stationary noise removal through spectrogram´s zeros technique applied to a recorded raw signal [20], whereas Fig.6 reveals the generation of siblings-sounds (MBSD's output), taking into account the noise channel. The MBSD algorithm´s performance was

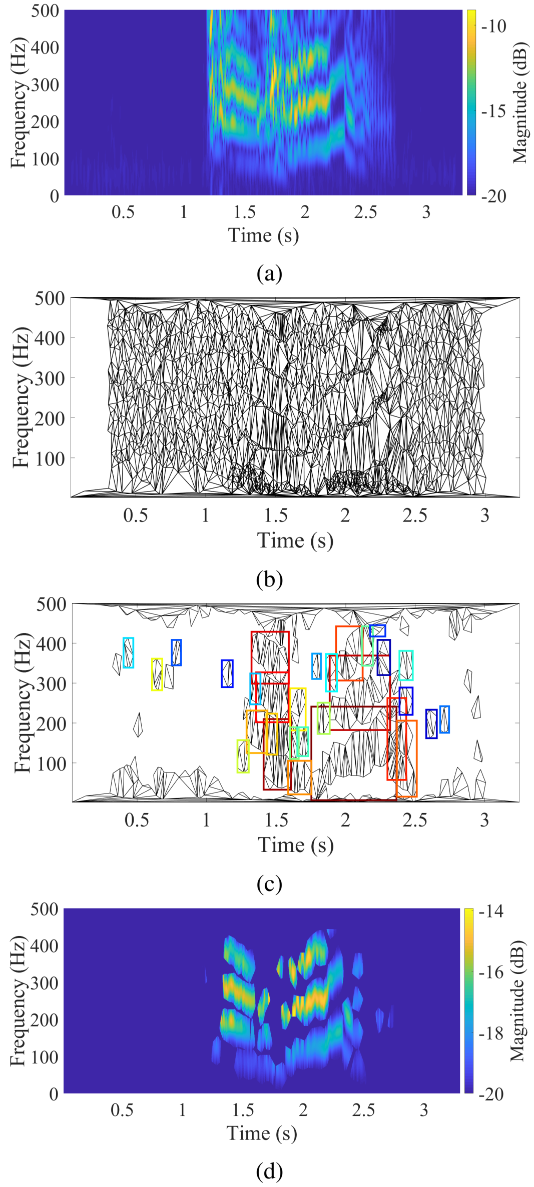


Fig. 4: Noise removal based on the spectogram zeros applied to a single BS.

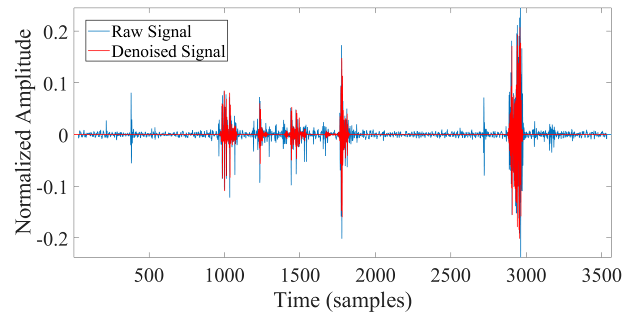


Fig. 5: BS identification and separation from the recorded noisy signal applied to multiple BS segments.

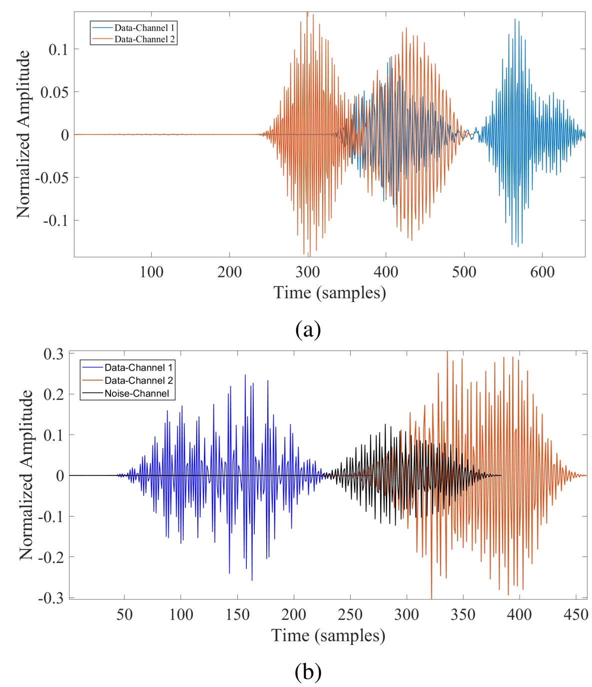


Fig. 6: (a) Generation of brother-segments consisted of high-correlated sound waves that were recorded by different data channels (orange-channel 1, blue-channel 2) with a time latency as it takes for the bowel sounds to be propagated through the intestinal area. (b) Example of brother-segments deletion by the MBSD algorithm due to high-correlation with the noise channel.

evaluated by means of Sensitivity=50%, Precision=91.9%, F1-Score=64.8% and Success Rate=100% (BS detected by MBSD / Total number of annotated BS), based on the clinical annotations, indicating a significantly high percentage of BS detection. It's important to highlight that Sensitivity is affected by the False Negative values, therefore low values of this metric indicate the presence of numerous signals that were initially classified by MBSD as Bowel Sounds but not included in the medical annotations, due to low power or short duration.

Thereupon, following the methodology's parts A and B, BS data have been extracted creating, thus, the BS case of the deep learning model. Specifically, 1952 signals out of the total 2440 BS sounds were used for training and 488 for validation. Likewise, 1952 non-BS data samples were used for testing, while the remaining 488 to model's validation. Fig.7 show the distribution of the reconstruction error in both BS and non-BS cases revealing an obvious source separation of the two cases.

In addition, ROC analysis has been used in order to select the optimal discrimination threshold that maximizes the performance of the autoencoder. The best threshold is observed at MSE=0.0075 where the AUC (Area Under the ROC Curve) measure is calculated to 0.9991. Thus, by applying the best observed threshold (MSE=0.0075), effective discrimination (i.e., Sensitivity = 99.2%, Precision = 99.6%, F1-Score = 99.4%, Accuracy = 99.4%) has been achieved between BS and the non-stationary noise.

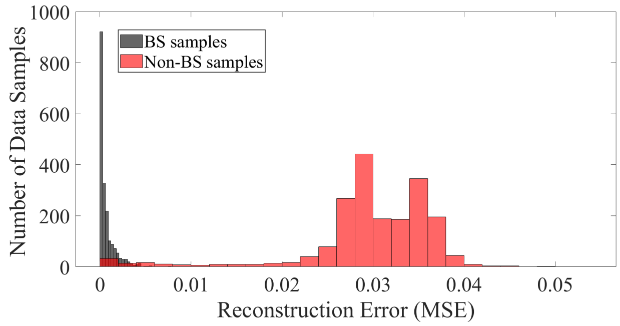


Fig. 7: Distribution of the reconstruction error (in terms of MSE) in BS (black color) and non-BS (other non-stationary noise samples with red color) cases.

V. CONCLUSIONS

This work proposes a multi-stage signal processing pipeline that achieves effective and efficient identification of BS, even in noisy settings, such as home and urban environments. Towards this direction, a prototype abdominal sound capturing device, namely, smart-belt, has been used. The novel architecture of the proposed approach in combination with the rather promising experimental results paves the way for an abdominal auscultation tool that enables user-friendly long-term home-based bowel sound monitoring.

Future work includes optimization of denoising and autoencoder's training performance towards augmented BS detection behavior. Furthermore, further research will focus on utilizing the proposed algorithm to assess gastrointestinal activity in the context of diseases with relevant symptoms, such as Crohn's and Parkinson's disease.

ACKNOWLEDGMENT

The authors thank the volunteers for their participation in the experiment.

REFERENCES

[1] W. Cannon, "Auscultation of the rhythmic sounds produced by the stomach and intestines", Am. J. Physiol., vol. 14, pp. 339-353, 1905.
[2] Horn, G. E., and J. M. Mynors. "Recording the bowel sounds." Medical and biological engineering 4.2 (1966): 205-208.
[3] Yoshino, Hajime, et al. "Clinical application of spectral analysis of bowel sounds in intestinal obstruction." Diseases of the colon & rectum 33.9 (1990): 753-757.
[4] Hadjileontiadis, Leontios J. "Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding-Part I: Methodology." IEEE Transactions on Biomedical Engineering 52.6 (2005): 1143-1148.
[5] Hadjileontiadis, Leontios J. "Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding-Part II: Application results." IEEE transactions on biomedical engineering 52.6 (2005): 1050-1064.
[6] Hadjileontiadis, Leontios J., and Ioannis T. Rekanos. "Detection of explosive lung and bowel sounds by means of fractal dimension." IEEE Signal Processing Letters 10.10 (2003): 311-314.
[7] Dimoulas, Charalampos, et al. "Novel wavelet domain Wiener filtering de-noising techniques: application to bowel sounds captured by means of abdominal surface vibrations." Biomedical signal processing and control 1.3 (2006): 177-218.
[8] Ranta, R., et al. "A complete toolbox for abdominal sounds signal processing and analysis." 3rd European medical and biological engineering conference IFMBE-EMBEC. Vol. 5. 2005.
[9] Ranta, R., Louis-Dorr, V., Heinrich, C. H., Wolf, D., & Guillemin, F. (2004, September). Principal component analysis and interpretation of bowel sounds. In The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Vol. 1, pp. 227-230). IEEE.
[10] Dimoulas, C., Kalliris, G., Papanikolaou, G., Petridis, V., & Kalampakas, A. (2008). Bowel-sound pattern analysis using wavelets and neural networks with application to long-term, unsupervised, gastrointestinal motility monitoring. Expert Systems with Applications, 34(1), 26-41.
[11] Kim, Keo-Sik, Jeong-Hwan Seo, and Chul-Gyu Song. "Non-invasive algorithm for bowel motility estimation using a back-propagation neural network model of bowel sounds." Biomedical engineering online 10.1 (2011): 69.
[12] Ulusar, Umit D. "Recovery of gastrointestinal tract motility detection using Naive Bayesian and minimum statistics." Computers in biology and medicine 51 (2014): 223-228.
[13] Liu, J., Yin, Y., Jiang, H., Kan, H., Zhang, Z., Chen, P., ... & Wang, Z. (2018, October). Bowel Sound Detection Based on MFCC Feature and LSTM Neural Network. In 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS) (pp. 1-4). IEEE.
[14] Tomomasa T, Takahashi A, Nako Y, Kaneko H, Tabata M, Tsuchida Y, et al. Analysis of gastrointestinal sounds in infants with pyloric stenosis before and after pyloromyotomy. Pediatrics. 1999;104:e60.
[15] Craine BL, Silpa M, O'Toole CJ. Computerized auscultation applied to irritable bowel syndrome. Dig Dis Sci. 1999;44:1887–92.
[16] Craine BL, Silpa ML, O'Toole CJ. Enterotachogram analysis to distinguish irritable bowel syndrome from Crohn's disease. Dig Dis Sci. 2001;46:1974–9.
[17] Craine BL, Silpa ML. Use of a computerized GI sound analysis system. Am J Gastroenterol. 2003;98:944.
[18] Spiegel BM, Kaneshiro M, Russell MM, Lin A, Patel A, Tashjian VC, et al. Validation of an acoustic gastrointestinal surveillance biosensor for postoperative ileus. J Gastrointest Surg. 2014;18:1795–803.
[19] Andrisha-Jade Inderjeeth, Katherine Webberley, Josephine Muir, and BarryMarshall. The potential of computerised analysis of bowel sounds for diagno-sis of gastrointestinal conditions: A systematic review.Systematic Reviews,7, 08 2018.
[20] Flandrin, Patrick. "Time–frequency filtering based on spectrogram zeros." IEEE Signal Processing Letters 22.11 (2015): 2137-2141.
[21] Zhou, Chong, and Randy C. Paffenroth. "Anomaly detection with robust deep autoencoders." Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017.
[22] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016. http://www.deeplearningbook.org.
[23] Guillaume Alain and Yoshua Bengio. What regularized auto-encoders learn from the data-generating distribution. The Journal of Machine Learning Research, 15(1):3563–3593, 2014.
[24] Peter Foster, Siddharth Sigtia, Sacha Krstulovic, Jon Barker, and Mark D.Plumbley. Chime-home: A dataset for sound source recognition in a do-mestic environment. In 2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), pages 1–5, 2015.
[25] Robert S Bolia, W Todd Nelson, Mark A Ericson, and Brian D Simpson. A speech corpus for multitalker communications research. The Journal of theAcoustical Society of America, 107(2):1065–1066, 2000.
[26] Jean-Rémy Gloaguen, Arnaud Can, Mathieu Lagrange, and Jean-François Petiot. Realistic urban sound mixture dataset, March 2018.