Methodological Article

Isfahan Artificial Intelligence Event 2023: Reflux Detection Competition

Abstract

Background: Gastroesophageal reflux disease (GERD) is a prevalent digestive disorder that impacts millions of individuals globally. Multichannel intraluminal impedance-pH (MII-pH) monitoring represents a novel technique and currently stands as the gold standard for diagnosing GERD. Accurately characterizing reflux events from MII data are crucial for GERD diagnosis. Despite the initial introduction of clinical literature toward software advancements several years ago, the reliable extraction of reflux events from MII data continues to pose a significant challenge. Achieving success necessitates the seamless collaboration of two key components: a reflux definition criteria protocol established by gastrointestinal experts and a comprehensive analysis of MII data for reflux detection. **Method:** In an endeavor to address this challenge, our team assembled a dataset comprising 201 MII episodes. We meticulously crafted precise reflux episode definition criteria, establishing the gold standard and labels for MII data. **Result:** A variety of signal-analyzing methods should be explored. The first Isfahan Artificial Intelligence Competition in 2023 featured formal assessments of alternative methodologies across six distinct domains, including MII data evaluations. **Discussion:** This article outlines the datasets provided to participants and offers an overview of the competition results.

Keywords: *24-h monitoring, deep learning, Isfahan Artificial Intelligence Challenge, multichannel intraluminal impedance, reflux*

Submitted: 17-Jul-2024 **Revised:** 23-Aug-2024 **Accepted:** 18-Sep-2024 **Published:** 28-Feb-2025

Azra Rasouli Kenari[1], Ahmadreza Montazerolghaem[2], Zahra Zojaji[2], Mehdi Ghatee[3], Behnam Yousefimehr[3], Amin Rahmani[3], Mahdi Kalani[4], Farnoush Kiyanpour[4,5], Mohamad Kiani-Abari[2], Mohammad Yasin Fakhar[2], Safiyeh Rezaei[2], Mojtaba Tahernia[2], Mohammad Hossein Vafaie[1], Hamidreza Besharatnezhad[6], Vahid Rahimi Bafrani[7], Mohamad Taghi Tofighi[6], Peyman Adibi Sedeh[8], Maryam Soheilipour[8], Hossein Rabbani[1]

[1]Medical Image and Signal Processing Research Center, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran, [2]Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran, [3]Department of Mathematics and Computer Science, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, [4]Regenerative Medicine Research Center, Isfahan University of Medical Science, Isfahan, Iran, [5]Department of Bioinformatics, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran, [6]Department of Electrical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran, [7]School of Electrical engineering, Iran University of Science and Technology, Tehran, Iran, [8]Isfahan Gastroenterology and Hepatology Research Center, Department of Internal Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

Introduction

Gastroesophageal reflux disease (GERD) is a prevalent digestive disorder that impacts the lower esophageal sphincter (LES), the muscle ring between the esophagus and stomach.[1] According to the Montreal definition, GERD causes troublesome symptoms and complications arising from stomach contents refluxing into the esophagus.[2] In its resting state, the LES, in conjunction with the right crus of the diaphragm, forms a firm barrier against gastric reflux. The antireflux barrier, composed of the LES and the right crus of the diaphragm, maintains a requisite pressure differential between the abdominal and thoracic compartments to prevent stomach content regurgitation.[3] A GER occurs when the LES weakens or relaxes inappropriately, permitting stomach contents to flow back into the esophagus.[4]

While traditional diagnostic methods such as endoscopy and 24-h pH monitoring have been valuable in diagnosing and guiding treatment for GERD, there are cases where patients exhibit GERD symptoms despite normal results from these tests, leading to inadequate responses to antisecretory therapy. Multichannel intraluminal impedance-pH (MII-pH) monitoring, introduced as a new technique in 1991,[5] is now considered the gold standard for GERD diagnosis.[6] This method combines impedance and pH measurements to determine if symptoms are triggered by GER episodes.

MII-pH monitoring detects bolus movement within the esophagus without radiation,

Address for correspondence: *Prof. Peyman Adibi Sedeh, Isfahan Gastroenterology and Hepatology Research Center, Department of Internal Medicine, Isfahan University of Medical Sciences, Isfahan - 8174673461, Iran.*
E-mail: adibi@med.mui.ac.ir
Dr. Maryam Soheilipour, Isfahan Gastroenterology and Hepatology Research Center, Department of Internal Medicine, Isfahan University of Medical Sciences, Isfahan - 8174673461, Iran.
E-mail: maryamsoheilip@med.mui.ac.ir

How to cite this article: Kenari AR, Montazerolghaem A, Zojaji Z, Ghatee M, Yousefimehr B, Rahmani A, *et al.* Isfahan Artificial Intelligence Event 2023: Reflux detection competition. J Med Signals Sens 2025;15:6.

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_46_24

Quick Response Code:

allowing for the identification of both acid and nonacid reflux. By measuring impedance at multiple sites, the direction of bolus movement can be ascertained based on temporal differences in entry and exit. Impedance monitoring relies on electrical impedance measurements between closely spaced electrodes during bolus passage, providing information on bolus movement direction. Bolus movements can be retrograde or antegrade, corresponding to GER and swallow-related events, respectively. The pH sensor attached to the catheter records esophageal pH levels during impedance monitoring, classifying GER episodes as acid or nonacid.

MII measurement assesses bolus transit and physical state using disposable catheters equipped with impedance electrodes and pH sensors. The protocol typically involves a 24-h outpatient study. Different states of GER events – liquid, gaseous, or mixed – can be detected through MII signals.

In the absence of swallow or GER within the esophagus, the impedance is identified by the electrical conductivity of the inner wall, and it is relatively stable, which is known as the baseline impedance value. When the current goes through the air, it will experience almost infinitely high impedance. In contrast, when a well-conducting fluid, such as saliva or gastric juice, is between the electrodes, the impedance is low. Using these principles, intraluminal MII measurement can be used to study the transit of a bolus and also to determine the physical state of it.

To measure the impedance and study the pH of the esophagus, disposable catheters with an array of impedance ring electrodes and one or two ISFET pH electrodes are commercially available. Cylindrical metal electrodes are mounted on a thin plastic catheter. The usual protocol is to conduct a 24-h study in an outpatient setting. All types of physical states of GER events, whether liquid, gaseous, or mixed (liquid-gas), can be detected in the MII signals.[7-9] The difference in patterning of the electrical conductivity of liquid, gas, or mixed intraluminal content allows distinction among these luminal contents whereas the sequence of impedance changes in different segments allows recognition of flow in either antegrade or retrograde directions.[6]

Depending on its nature, a GER may appear as specific patterns in each of the pH and impedance signals, leading to changes in the time-space behavior of these signals. A GER is associated with some variables of clinical importance. For each event, the following variables can be determined based on the information provided by MII-pH monitoring: physical content, chemical composition, symptom association, duration, proximal extent, and position of the patient's body during reflux.[10] The pH of the refluxate can be determined on the basis of the pH signal,[7] which can be ignored in investigations for characterizing GER events.[1,2]

While MII-pH monitoring represents a significant advancement in GERD diagnostics, the accurate interpretation of impedance data is paramount to its effectiveness in clinical practice. Proper analysis of impedance measurements is essential for maximizing the utility of MII-pH monitoring and ensuring optimal patient outcomes in the management of GERD. Several clinical studies have highlighted the importance of thorough analysis and interpretation of impedance data.[4-6] The subject of automatic detection of GER events analyzing MII data has only been addressed in very few works.[1,2,11,12]

Deep learning has emerged as an extraordinary tool in the field of artificial intelligence (AI) and machine learning applications in the recent past, and its progress can be extended to almost every industry.[13-15] We hypothesize that the deep learning approaches may potentially lead to a better characterization performance.

The GER events are infrequent, taking a very short time; therefore, the problem of manipulating MII dataset can be considered the rare event problem. In fact, the rare event problem occurs when we encounter the unbalanced dataset. This problem is quite common in medical datasets where we are looking for special abnormality in a huge amount of data recorded from patient monitoring devices. In a typical rare event problem, the positively labeled samples (around 10% of the total data) are much fewer than negatively labeled data.[16] In such cases, deep learning methods, which try to learn the pattern of input data in successive trainable layers, have proven to act effectively.[17] Furthermore, the facility of using different architectures of layers makes these methods more flexible.

The Isfahan National Elite Foundation hosted a GER detection competition during the intra-abdominal infection 2023 events with the goal of advancing MII GER detector technology by impartially assessing a range of AI-based techniques. Participants were given a meticulously documented dataset of labeled MII episodes, referred to as the "training set," to develop their methods. Researchers subsequently submitted the outputs of their algorithms for evaluation using a separate test dataset. The labels of this dataset were concealed to ensure an unbiased assessment of performance, eliminating any influence from method selection or parameter tuning specific to the data.

In the upcoming sections of this article, we aim to explain and compare the methodologies employed by the top competitors.

Materials and Methods

Ranking of competition results

The results of the submitted contests were approached with caution, as they might not have offered a completely unbiased evaluation of excellence due to various influencing factors. To address these challenges, the final phase of the competition was conducted live and assessed based on five equally important criteria. These criteria were utilized to rank the top four teams in the competition.

The five criteria of equal weight were as follows:
1. Innovation of the proposed approach

2. Performance in the initial stage competition (based on the best submission and score)
3. Performance in the final submission (second phase conducted live at Abbasi Hotel, Isfahan, Iran) using the dataset
4. Quality and clarity of the final report
5. Quality and clarity of the presentation.

Each criterion was rated on a scale of 1, 2, or 3. The team that excelled in each criterion received three points, whereas the second-best team received two points. The final rankings were determined based on the total points accumulated across all five criteria during the judges' deliberations at the conclusion of the competition.

Performance metrics

To assess classification models and provide valuable insights into different aspects of model performance, positive predictive value (PPV), sensitivity, and F1-measure have been calculated along each channel separately.

PPV measures the proportion of true positive (TP) predictions out of all positive predictions made by the model. It is calculated as Eq. 1.

Sensitivity, also known as recall, quantifies the model's ability to correctly identify all actual positive instances in the dataset. It is calculated as Eq. 2. Sensitivity emphasizes the model's capability to capture all positive instances.

The F1-measure combines precision (PPV) and recall (sensitivity) into a single metric, providing a balanced assessment of a model's performance. It is calculated as Eq. 3. The F1-measure considers both false positives (FPs) and false negatives (FNs), offering a comprehensive evaluation of the model's precision and recall trade-off.

$$PPV(\%) = 100\frac{TP}{TP \quad FP} \qquad \text{Eq. 1}$$

$$Sen(\%) = 100\frac{TP}{TP + FN} \qquad \text{Eq. 2}$$

$$F1\_Score(\%) = 100\frac{2 \times TP}{2 \times TP + FP + FN} \qquad \text{Eq. 3}$$

The TP, true negative, FP, and FN are calculated using the output of the network for element-wise classification, and label annotations of each channel would be described by the confusion matrix in Table 1.

Dataset

This dataset was provided by the Noor and Hazrate Ali-Asghar Hospital, Isfahan University of Medical Sciences (Head: M. Soheilipour, M.D.), Govarsanji center, Isfahan University of Medical Sciences (Head: P. Adibi, M.D.), and the School of Advanced Technologies in Medicine (Head: H. Rabbani), Isfahan University of Medical Sciences.

Table 1: Confusion matrix for elementwise classification

Ground truth annotations	Output of network	
	0	**1**
0	TN	FP
1	FN	TP

FP – False positive; FN – False negative; TP – True positive; TN – True negative

Description of dataset

For patients with normal endoscopy who did not respond to antisecretory therapy, 24-h MII-pH monitoring was conducted and subsequently archived. Each study spanned approximately 24 h. The ambulatory 24-h esophageal MII-pH monitoring was executed using a mobile recording device (Ohmega Impedance ambulatory pH meter; MMS, Enschede, the Netherlands). A transnasal passage of the six-channel impedance – pH catheter was performed under topical anesthesia and affixed 5 cm above the LES to capture pH at 5 cm and impedance readings at 3, 5, 7, 9, 15, and 17 cm proximal to the LES. The six impedance sites (channels 1 through 6) were positioned between every two consecutive impedance electrodes. Data from the impedance channels and pH electrodes were stored on the portable data recorder, and digital data loggers (Ohmega R; MMS B. V., Enschede, the Netherlands) linked to the catheter. Subsequently, the data were transferred to a computer after the study. The MMS database software (MMS B.V.) facilitated the recording and conversion of raw data into CSV files. MII-pH data were recorded at a sampling rate of 50 Hz.

The pivotal concept in diagnosing GERD is characterizing GERs from MII data. The presence of various GER types can be identified by analyzing the MII signal, regardless of pH levels. To simplify and streamline the process and avoid multimodal processing, the pH data were disregarded. Consequently, the data vector provided for the challenge is of a single impedance type.

We selected the archived data of 26 patients with different age and sex for the experiment. Our objective was to extract informative and balanced data. However, a typical 24-h MII-pH study includes long durations of isoelectric intervals; therefore, for each of the 24-h MII-pH signal, a number of at least four episodes were selected including one GER event or more. Each MII episode lasts for 2 min that may contain swallows and at least one GER event.

The dataset included a total number of 202 episodes, including 208 GER events. To be scaled properly, all the episodes were normalized to the unit, by dividing the impedance value of each channel by its corresponding –norm. At the end, the binary mask of GER events was provided for all of the 202 episodes with the assistance of three experienced gastroenterologist experts in several meetings.

Figure 1 illustrates graphs of two samples extracted from the MII dataset, demonstrating two separate instances of

MII data from the six channels plotted against the time axis, accompanied by their respective labels.

The dataset was subdivided into training, testing, and hidden sets for the competition. The training data, along with their corresponding labels, were distributed to the teams. As the competition progressed to the second stage, the teams were provided with test MII data devoid of labels. Subsequently, eight submissions were received for the test labels, with some submissions showcasing notably high-quality results, 4 out of the 8 submissions achieving superior F1 scores. In the final phase of the competition, which took place live, participating teams were given access to the hidden episodes. The number of MII episodes, subjects, and GER events in each set is detailed in Table 2.

Notably, the MII data allocated to the training and testing sets originated from distinct individuals, whereas the hidden data were randomly sourced from any patient in the dataset.

Results and Discussion

Eight competitors submitted their results for evaluation on the dataset. The top four winning teams achieved the highest F-scores, specifically at channel 6, surpassing the 5th team by 8% and significantly outperforming the other leading competitors. Figure 2 presents box plots depicting the performance assessment of the competitors.

Observing the results, it is evident that all teams exhibited superior performance on channels 6 through 4. This observation aligns with the logical expectation, as channels 3-1 are closer to the throat where more air is present, leading to increased noise in the recorded data. Another reason is the rarity of GER events in proximal channels in comparison to distal channels. Consequently, analyzing data from proximal channels poses a greater challenge.

The final phase of the competition was conducted live, where the hidden data described in Table 2 were provided to the participants. The top four teams in the final assessment, along with a brief overview of their proposed methods, are outlined below:

1. The NORC team from Amirkabir University of Technology, Tehran, Iran, achieved an F1-measure of 72% for semantic segmentation of GER events across channel 6 of MII data. Their proposed architecture entails training six distinct convolutional neural network (CNN) models. Each model is designed to capture the relationship between all channels of MII data by considering both the MII data input from a channel and its corresponding label, along with the MII data from other channels. They introduced a two-dimensional CNN with a kernel size of 6 × 1 × 1 to model the correlation between different channels. Preprocessing of MII data involved normalization, moving average, and interval segmentation with overlap, tailored to each segment's labels. The averaging of predicted labels from overlapping segments was utilized, converting the probability signal into a predicted label signal of the same size as the input MII signal. Training encompassed 30 iterations with the utilization of varying weights (class weight) to compute backward propagation of errors, addressing the impact of imbalanced classes. For the detection of GER events, they employed a simple, low-parameter CNN model with a single convolution layer, a prudent choice given the limited dataset. The block diagram of their proposed method together with the graphical abstract is depicted in Figure 3
2. The RMRC team from Regeneration Medicine Research Center, Isfahan University of Medical Sciences, Isfahan, Iran, were the second-best competitors. They could achieve 74% of the F1-measure for semantic segmentation of GER events at channel 6. The methodological approach adopted in this study encompassed two main steps: diagnosis of reflux and signal segmentation (Step 1) followed by segmentation of channels 6-1 (Step 2). The block diagram of their proposed method together with the graphical abstract is depicted in Figure 4

   The U-Net architecture employed in this study comprised an initial convolutional layer, an encoder with down-convolution blocks, central convolution blocks, and a decoder with up-convolution blocks, ultimately yielding a single-channel output. Initially, the dataset was preprocessed using median filtering and resampling. Step 1 involved training a U-Net model using binary cross-entropy loss and the IoU metric. The model was trained on a subset of the data using a sixfold cross-validation approach. In addition, the best-performing model was selected based on its IoU score

   In Step 2, the segmentation task was extended to cover channels 6-1 individually. Signal intervals were extracted and cropped based on positive signal intervals from the previous step. Training for each channel involved utilizing the appropriate input signals and labels. Specifically, channels 6 and 5 served as inputs for diagnosing and segmenting channel 6, whereas subsequent channels utilized signals from channels 6 to the respective channel and the label of the previous channel

   Postprocessing steps were applied to further refine the segmentation results. The postprocessing techniques

Table 2: Multichannel intraluminal impedance dataset partitioning in each level of competition

	MII dataset		
	Train	Test	Hidden
Number episodes	147	27	28
Number subjects	18	8	-
Number GERs	148	27	33

MII – Multichannel intraluminal impedance;
GERs – Gastroesophageal refluxes

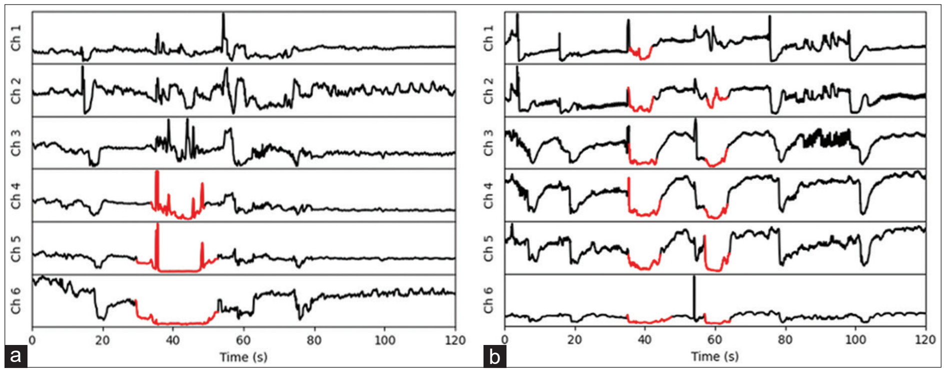

Figure 1: Sample multichannel intraluminal impedance episodes including swallows and gastroesophageal reflux disease (GER) events. All nonzero-labeled signal points representing GER events are highlighted in red. (a) There is one GER event with a proximal extent up to channel 4, and (b) Two GER events are observed, with the left GER event having a proximal extent to channel 1 and the right GER event showing a proximal extent to channel 2

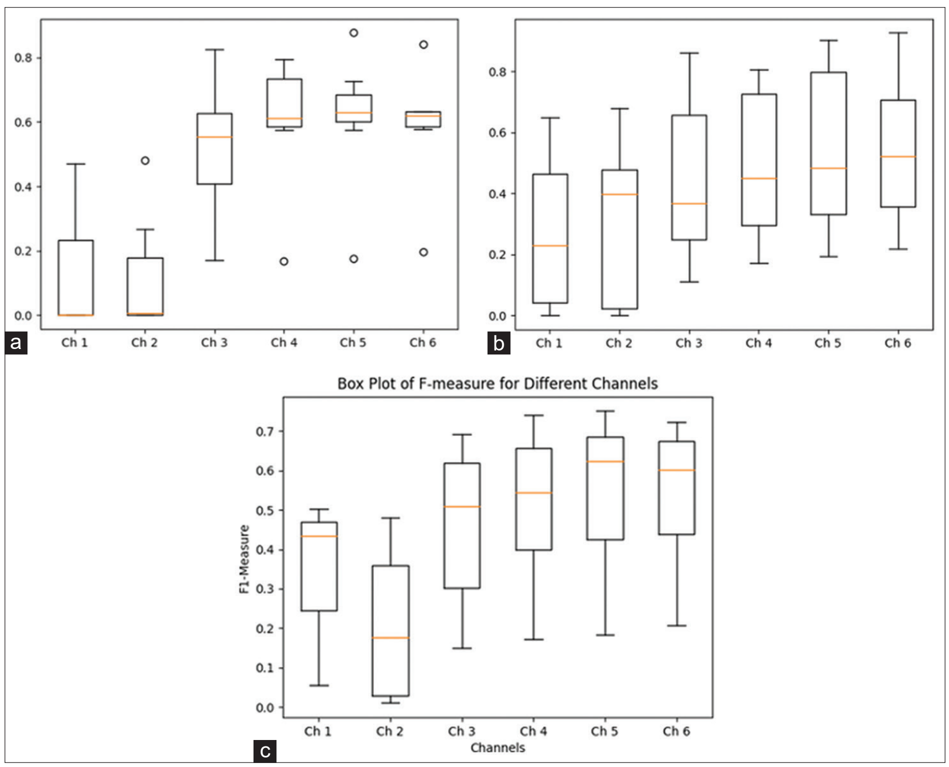

Figure 2: Average performance (mean ± standard deviation) across all teams during the initial phase of the competition. A box plot showcasing various statistics for the eight participating teams (excluding outliers resulting from division by zero) is presented. (a) Displays sensitivity, (b) Represents positive predictive value, and (c) Depicts F1-measure in relation to each channel of multichannel intraluminal impedance data

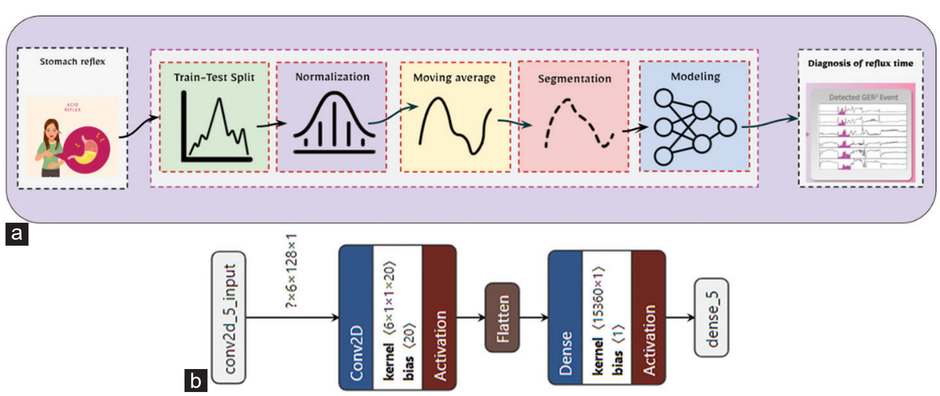

Figure 3: Proposed method of the 1st best competitor. (a) Graphical abstract, (b) Model

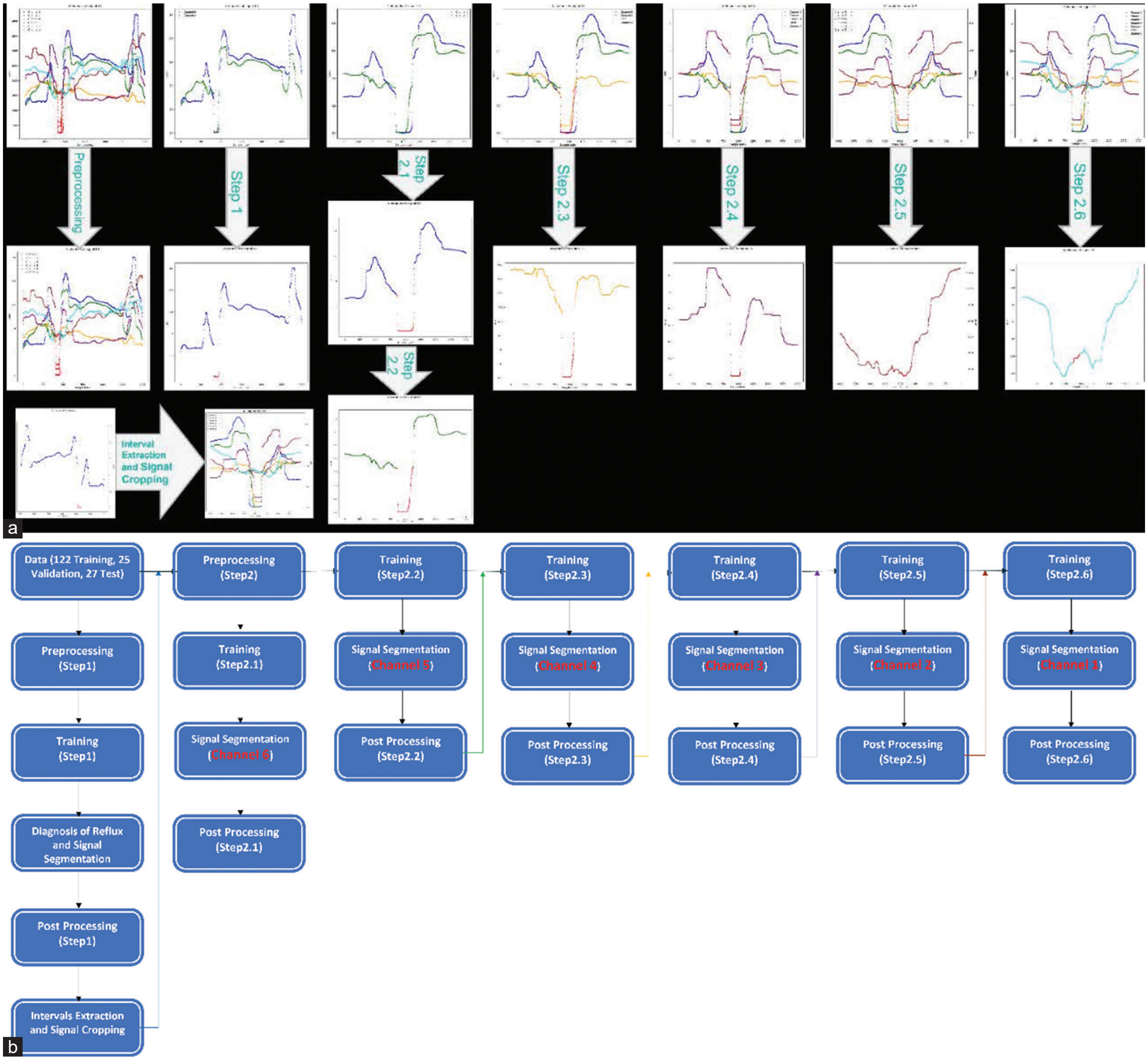

Figure 4: Proposed method of the 2nd best competitor. (a) Graphical abstract, (b) Model

helped improve the overall quality and reliability of the segmentation results, ensuring a more precise analysis of esophageal impedance data

3. The Hoopad team from Isfahan University, Iran, along with the Soorena team from Isfahan University of Medical Sciences, Iran, achieved the 3rd position in the competition. The Hoopad team attained an F1-measure of 61.75% for GER event detection on the 6th channel by utilizing features extracted through Short-Time Fourier Transform and a BiLSTM model. On the other hand, M. Vafaee and his group achieved an F-measure of 17.05% for GER event detection on channel 6 by employing a U-Net architecture that utilized one-dimensional convolutional blocks for processing MII data. Their network processed a single 6-channel raw data episode. They conducted a fivefold cross-validation with an 80/20 split between training and validation datasets. The architectural outline of their network is illustrated in Figure 5.

Conclusion and Outlook

Looking at the best winning teams reveals several interesting aspects.

- Emphasis on Channel Relationships and Cross-Channel Data Utilization
- Importance of Preprocessing Techniques for Enhancing Model Performance

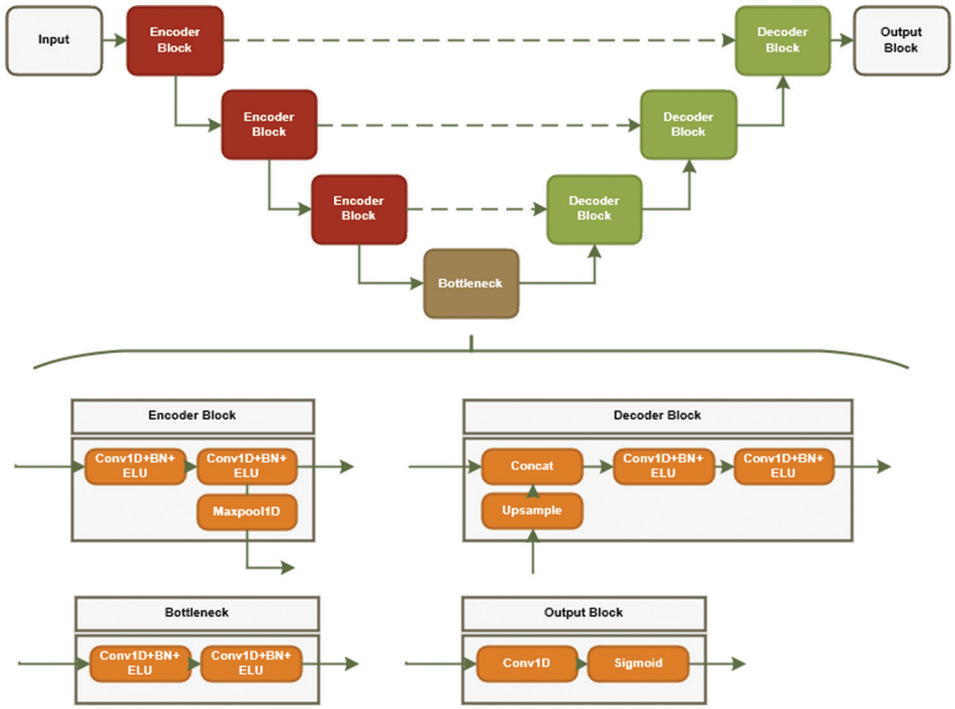

Figure 5: Proposed framework of the 3rd (M. Vafaee *et al.*) best competitors

- Utilization of Different Architectures (CNN, UNet, BiLSTM) for GER Event detection
- Detailed Preprocessing and Postprocessing Strategies for Improved Results
- Overall Focus on Precision and Quality of Semantic Segmentation Results.

The competition to detect GER events using esophageal impedance data showcased the innovative approaches and methodologies employed by the top teams.

While M. Ghatee *et al*. leveraged the use of CNNs with a focus on modeling the correlation between different channels of data, Y. Gheisari *et al*. adopted a U-Net model combined with postprocessing techniques for more refined segmentation results. Kiani's team adopted techniques such as Short-Time Fourier Transform and a BiLSTM model which underscored the importance of advanced feature extraction methods and model architectures in enhancing detection accuracy. Vafaee's team, despite a lower F-measure, provided valuable insights into the utilization of U-Net architecture and one-dimensional convolution blocks for processing MII data. Kiani's team achieved a significantly higher F1 measure (61.75%) compared to Vafaee's team (17.05%). This difference can be attributed to the use of a BiLSTM model, which is known for its ability to capture long-range dependencies in time series data. Kiani's team also employed the Short-Time Fourier Transform, effectively extracting features from the MII signal. The observed disparity in performance underscores the critical role of feature engineering and model selection in achieving effective GER event detection from multichannel MII data.

The competition not only highlighted the potential of machine learning in medical research but also underscored the importance of precise analysis in diagnosing medical conditions such as GER. Moving forward, further research and collaboration in this field could pave the way for enhanced diagnostic tools and treatment strategies for patients with GER-related issues.

Financial support and sponsorship

This work was supported by the Isfahan Elite Foundation (IEF), which sponsored the Isfahan Artificial Intelligence Event 2023 (IAI2023). The IEF organized the event and provided financial support for the 10 challenges, including Challenge I: Reflux Detection Competition. Several winners received prizes from the IEF.

Conflicts of interest

The authors declare the following potential conflicts of interest:

- ZZ and AM were the organizers of the Isfahan AI (IAI) 2023 competitions on behalf of Isfahan Elite Foundation (IEF), which included 10 challenges.
- PAS, MS, HR and ARK served as scientific committee members for Challenge I: Reflux Detection Competition. They were responsible for evaluating the methodologies and results of all participant teams.
- MG, BY, AR, MK, FK, MKA, MYF, SR, MT, MHV, HB, VRB, MTT are members of the winning teams in this challenge.

None of the organizers and scientific committee members (PAS, MS, HR, ARK, ZZ and AM) contributed to the development of the methods used by the participating teams. The final decision regarding the winners was made by the policy council members based on the following criteria:

- Technical contribution in developed models by teams,
- The results on initial and final test data of each team,
- The submitted reports and teams' presentations.

The authors have disclosed these relationships to ensure transparency and maintain the integrity of the research.

Acknowledgements

We extend our appreciation to the policy council members, Prof. Behrouz Minaei Bidgoli from Iran University of Science and Technology, Prof. Mohammad Hassan Moradi from Amirkabir University of Technology, Prof. Heshaam Faili from University of Tehran, Prof. Emad Fatemizadeh from Sharif University of Technology, and Prof. Hossein Rabbani from Isfahan University of Medical Sciences, for making the final decisions regarding the winners of all 10 challenges.

Finally, we thank the anonymous reviewers for their valuable feedback and contributions to the improvement of this paper.

References

1. Rasouli A, Rabbani H, Kermani S, Raisi M, Soheilipour M, Adibi P. A multichannel intraluminal impedance gastroesophageal reflux characterization algorithm based on sparse representation. IEEE J Biomed Health Inform 2021;25:3576-86.
2. Rasouli A, Rabbani H, Raisi M, Soheilipour M, Adibi P. Liquid gastroesophageal reflux characterization by investigating multichannel intraluminal impedance-pH monitoring data. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:4636-9.
3. Sung HJ, Cho YK, Moon SJ, Kim JS, Lim CH, Park JM, *et al.* Role of acid and weakly acidic reflux in gastroesophageal reflux disease off proton pump inhibitor therapy. J Neurogastroenterol Motil 2012;18:291-7.
4. Roman S, Bruley des Varannes S, Pouderoux P, Chaput U, Mion F, Galmiche JP, *et al.* Ambulatory 24-h oesophageal impedance-pH recordings: Reliability of automatic analysis for gastro-oesophageal reflux assessment. Neurogastroenterol Motil 2006;18:978-86.
5. Gyawali CP, Kahrilas PJ, Savarino E, Zerbib F, Mion F, Smout AJ, *et al.* Modern diagnosis of GERD: The Lyon consensus. Gut 2018;67:1351-62.
6. Sifrim D, Castell D, Dent J, Kahrilas PJ. Gastro-oesophageal reflux monitoring: Review and consensus report on detection and definitions of acid, non-acid, and gas reflux. Gut 2004;53:1024-31.
7. Bredenoord AJ, Smout A, Tack J. A Guide to Gastrointestinal Motility Disorders. Cham, Switzerland: Springer; 2016.
8. Ghoshal UC. Evaluation of Gastrointestinal Motility and Its Disorders. Cham, Switzerland: Springer; 2016.
9. Tutuian R, Castell DO. Esophageal Multichannel Intraluminal Impedance Testing. In: Agrawal A, editor. Functional and Motility Disorders of the Gastrointestinal Tract. New York: Springer; 2014. p. 1-14.
10. Lawenko RM, Lee YY. Evaluation of gastroesophageal reflux disease using the bravo capsule pH system. J Neurogastroenterol Motil 2016;22:25-30.
11. Wong MW, Rogers BD, Liu MX, Lei WY, Liu TT, Yi CH, *et al.* Application of artificial intelligence in measuring novel pH-impedance metrics for optimal diagnosis of GERD. Diagnostics (Basel) 2023;13:960.
12. Zhou MJ, Zikos T, Goel K, Goel K, Gu A, Re C, *et al.* Development and validation of a machine learning system to identify reflux events in esophageal 24-hour pH/impedance studies. Clin Transl Gastroenterol 2023;14:e00634.
13. Chollet F. Deep Learning With Python. New York, NY: Manning Publications; 2017.
14. Bruno A, Ardizzone E, Vitabile S, Midiri M. A novel solution based on scale invariant feature transform descriptors and deep learning for the detection of suspicious regions in mammogram images. J Med Signals Sens 2020;10:158-73.
15. Amini N, Mostaar A. Deep learning approach for fusion of magnetic resonance imaging-positron emission tomography image based on extract image features using pretrained network (VGG19). J Med Signals Sens 2022;12:25-31.
16. Weiss GM, Hirsh H. Learning to predict extremely rare events. In: AAAI Workshop on Learning from Imbalanced Data Sets. Menlo Park, CA: AAAI Press; 2000. p. 64-8.
17. Liu T, Abd-Elrahman A, Morton J, Wilhelm VL. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. GIsci Remote Sens 2018;55:243-64.