An Analysis of Cavitation in Sonothrombolysis through Convolutional Neural Networks

Patricia A S Guenkawa[1], Sergio S Furuie[1]

[1] Escola Politécnica, Universidade de São Paulo, São Paulo, Brazil

Abstract

Myocardial infarction is one of the main causes of morbidity and mortality worldwide. Among the possible treatments for blood flow obstruction, an emerging technique is named sonothrombolysis. To reach satisfactory results, the event allied to the technique (cavitation of microbubbles) needs to be controlled, so harm to the patient can be avoided. In view of that, this study aimed to detect and classify the phenomenon during sonothrombolysis therapy through the use of artificial intelligence.

The signals were generated considering an 8×8 transducers' matrix, and an automatic and uncomplicated classifier method was proposed, based on the Continuous Wavelet Transform tool and Convolutional Neural Network (CNN) approach. The method made use of a pre-trained CNN, called AlexNet, operating a database of 2,800 waves for training, testing, and validation. The evaluation of the statistics included both the detection using broad and narrow bands, the noise level applied, and the database size.

For the case of narrowband receivers, the results of the study indicated an accuracy of around 95.7%. The result demonstrates that the use of artificial intelligence could be an approach to explore the detection of cavitation for therapies applying ultrasound.

1. Introduction

Approximately 1.5 million cases of myocardial infarction (MI) occur annually in the United States; the yearly incidence rate is close to 600 cases per 100,000 people [1]. Ultrasonic energy has been assessed extensively as a method to promote thrombolysis [2], and when the amount of energy is sufficient along with the use of microbubbles (MBs), the activity can restore the blood flow of vessels, and such a procedure is called sonothrombolysis. Sonothrombolysis is a novel therapy that has as its main benefits the noninvasive and non-ionizing features [3].

By the application of acoustic emissions directed to the thrombus location, the intention is to mechanically break up the occlusive blood clot through the cavitation of MBs. A very important caution to be avoided is the trigger of bubble collapse in not wanted spots, which could result in damage to healthy tissues. Therefore, the detection of the type of cavitation that is occurring and the location, is crucial for methods that combine acoustic waves and MBs, and consistent sonification is necessary for the method's safety and efficacy. That said, this work aimed to develop a feedback mechanism of detection based on artificial intelligence (AI) for guidance, expecting that the computational simulations can predict and classify the cavitation phenomenon that involves the sonothrombolysis technique.

Currently, some of the feedback methods applied in the field are the passive cavitation detection (PCD) technique [4], passive cavitation imaging (PCI) [5], and active cavitation detection (ACD) [6]. As an alternative, we investigated the possibility of employing AI as part of a detector approach, since AI has been an important tool that is extensively applied in medical sciences: diagnosing, remotely treating patients, and cancer detection based on CT scans are examples of its significant impact in healthcare [7]. This study explored the supervised learning category in AI methods, more specifically a Deep Learning (DL) proposal.

2. Materials and method

Figure 1 shows how the method is organized. First, the signal's preprocessing, followed by the feature extraction, and finally the classification.

2.1. Dataset

The database was assembled considering a defined focus region. It was regarded as an 8x8 matrix of ultrasound transducers, which were designed as emitters and receivers, characterized by a 10.15mm diameter, and resonance frequency of $f_0 = 250$kHz [8]. The ap-

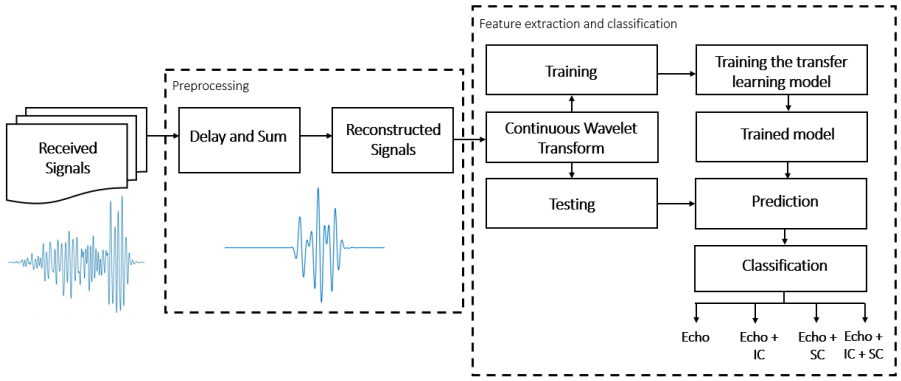

Figure 1. Flowchart with the general steps of the process.

proximate heart's dimensions are 120mm, 85mm, and 60mm, for length, width, and thickness, respectively [9]. Therefore, to ensure scanning the entire target, the volume of simulation selected is $120 \times 120 \times 120$ mm^3, where 874 foci are distributed in this region. Figure 2 shows the matrix array illustration.

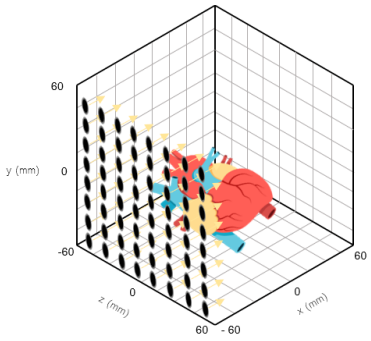

Figure 2. Normals at transducers' centers and the target region.

The signals were generated using the Kwave toolbox [10] for MATLAB®, where features of the acoustic medium can be settled, including non-linearities, attenuations, and the matrix array topology. Due to the different distances along each source's centers and the focal location, it was applied the delay-and-sum approach [11] in the received signals to reconstruct them constructively. In addition, the types of sources examined are echo, stable cavitation (SC), and inertial cavitation (IC). Based on [12], the SC produces ultra-harmonic waves, thus for detection purposes, we examined the spectral band (width of bands equal to $0.125 \times f_0$) around $1.5 \times f_0$. In contrast, since IC occurs when MBs collapse violently, and the literature correlates it to the fundamental concept of Kelvin impulse [13] (behavior of a jet impact as a consequence of a collapsing bubble), we represent the response of the source as an impulse, and the component frequencies are spread across the frequency spectrum, hence, we inspect bands between harmonics and ultra-harmonics ($1.75 \times f_0$ and $2.25 \times f_0$).

The dataset comprises a range of pairs (S, I) in which the values S and I are multipliers of the predefined 100 kPa amplitude for SC and IC, respectively, resulting in 2,800 waves for training, testing, and validation. Besides that, the noise level added follows the dynamic range of the signals, and we set values of 2%, 5%, and a level between 2 and 5% randomly (RN).

2.2. Continuous wavelet transform (CWT) scalogram

Since the signals analyzed are structured of different frequency elements and are nonstationary, each signal was transformed to the time-frequency domain to facilitate feature extraction. The spatial or temporal representation alone does not represent well corresponding data, so CWT [14] is a suitable way to inspect such source types. The time-frequency spectrogram maps the signal differences through the expansion and contraction of the CWT window. The CWT for a signal $x(t)$ is defined as

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \cdot \Psi \left(\frac{t-b}{a} \right) dt \qquad (1)$$

where a and b are the scales and time shifts of a reference wavelet $\Psi(t)$, respectively.

For our problem, each signal is composed of 2,257 temporal samples, and the data is converted into time-frequency spectrograms in which no information overlaps, thus overfitting problems are avoided.

The chosen wavelet mother $\Psi(t)$ was the Wavelet Morlet (Gabor) as a parameter of the CWT. We ob-

tain the absolute wavelet coefficient values from the signals through the CWT filterbank from Matlab, rescaling those coefficients to the interval [0,1]. The second step is converting them into RGB images by applying the Jet colormap with 256 colors (compatible with AlexNet inputs). To exemplify the process, Figure 3 shows the time-domain for signals with different features and the respective scalograms.

The scales a are used to create the wavelet bandpass filters and depending on the chosen parameters as input for the CWT, those values change. They are discretized considering the number of wavelet filters per octave, which is 12 for this work, resulting in 141,000 coefficients for each signal. The larger the number of voices per octave, the finer the discretization of a, thus, the selection of the value considers the amount of computation required since it increases with its increment. The number of scales is calculated considering the energy spread of the wavelet in frequency and time [15].

With the scalograms set, the database composed of the constructed images was divided into 70% for training and 30% for testing and validation.

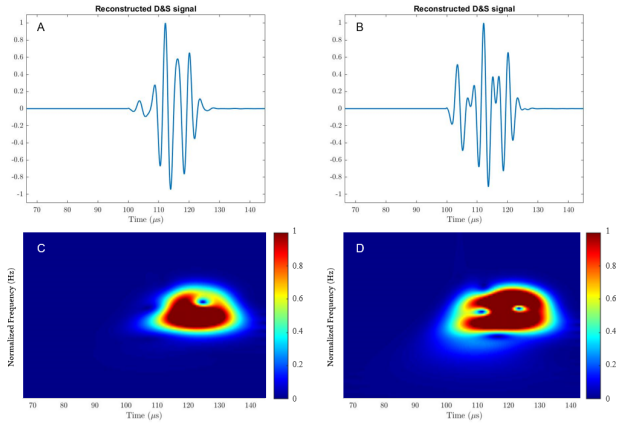

Figure 3. Signals and RGB images using the Jet colormap. (A) Example of an estimated temporal source signal with echo, SC, and noise. (B) Example of an estimated temporal source signal with echo, SC, IC, and noise. (C) RGB image for signal (A). (D) RGB image for signal (B).

2.3. Transfer Learning

AlexNet is a CNN known as a leading architecture for any object-detection task. The model is composed of 8 layers (5 convolutional layers and 3 fully connected layers) whereby the first layer is used to input a filtered image with dimensions of 227 × 227 × 3 respectively for width, height, and depth (red, green, blue), the last fully connected layer connects 1,000 neurons, and the rest of the layers work as feature extractors.

In this study, we took advantage of the transfer learning approach, where there is the reuse of a pretrained model (AlexNet) on a new problem and no need to create a CNN model from scratch. To operate the model for our problem, the last three layers of AlexNet were replaced: "Fully Connected", "SoftMax", and "Classification". Other network parameters are stochastic gradient descent with momentum (SGDM) as the optimizer, minibatch of size 32, and maximum epoch of 40 with a learning rate of 1e-4. CNN is implemented using the Deep Learning Toolbox Model for AlexNet Network [16] and trained on the NVIDIA GeForce RTX 970 graphical processing unit.

3. Results and discussion

The evaluation of our method's performance is obtained by three metrics: accuracy, precision, and F1-score, which are applied to analyze the results of transferred AlexNet.

Table 1. Performance's results for the different sources, comparing narrow and broadband receivers, and distinct levels of noise (2%, 5%, and RN)

	Narrowband			Broadband
Measure	RN	2%	5%	RN
Accuracy	95.7 ± 1.0	97.1 ± 0.6	96.1 ± 0.4	98.8 ± 0.4
Precision	96.0 ± 1.1	97.3 ± 0.5	96.3 ± 0.9	98.7 ± 1.0
F1-score	95.9 ± 1.7	97.2 ± 0.7	96.2 ± 0.5	98.7 ± 0.5

Table 1 shows the results both considering narrow and broadband receivers. The bandwidth of 100% of central frequency for broadband receivers yielded the best results (accuracy of 98.8%), which was expected among the cases proposed. However, the scenario applying narrowband receivers also produced acceptable detection indicators, with values of 96% for precision, 95.9% for F1-score, and 95.7% for accuracy for RN level. Those results suggest that it is possible to detect the phenomenon using the same set of ultrasound transducers or alternative broadband receivers.

A confusion matrix is also presented to assess the procedure's performance for the RN case (Figure 4). As we can observe, it referred to the 4 types of sources. The columns correspond to the target class (real classification) and the rows meet the output obtained

(method's classification). The approach's global accuracy acquired was 95.7% (last cell). We can observe that the mislabelling concentration occurs between *Echo + IC* and *Echo + IC + SC signals* (cell(2,4)).

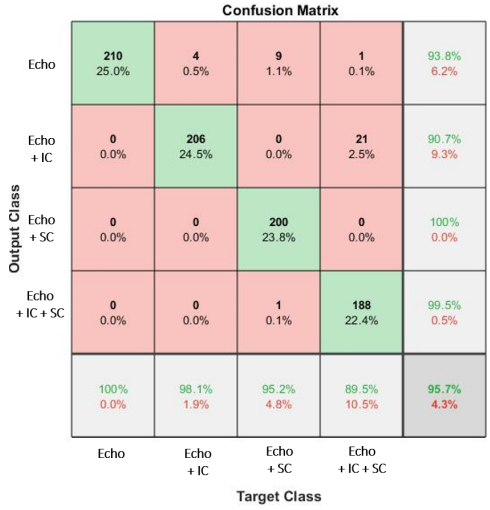

Figure 4. Confusion Matrix for narrowband receiver case - RN rate (840 waves for testing and training - 30% of the database).

In addition, to verify the influence of the database size in the classification, we tested smaller data as input of the CNN, and we attested a significant decrease in accuracy. For a database of 400 signals, for example, we obtained an accuracy of 68.3% for broadband receivers.

4. Conclusion

In this work, we developed a classification method for the cavitation phenomenon based on CWT and CNN. The procedure yields good results for the metrics explored, especially considering the few epochs applied (40), however, it is important to notice that the mislabelling errors present themselves, especially in signals with the occurrence of IC, which is the source we mainly want to supervise.

Acknowledgments

We would like to acknowledge the financial support from the Conselho Nacional de Desenvolvimento Cientifico e Tecnológico (CNPq) (130511/2021-8) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001.

References

[1] Zafari A, Abdou M. Myocardial infarction, May 2019. URL `https://emedicine.medscape.com/article/155919-overview#a6`.

[2] Trübestein G, Engel C, Etzel F, Sobbe A, Cremer H, Stumpff U. Thrombolysis by ultrasound. Clinical science and molecular medicine 01 1977;3:697s–698s.

[3] Medel R, Crowley R, McKisic M, Dumont A, Kassell N. Sonothrombolysis: An emerging modality for the management of stroke. Neurosurgery 11 2009;65:979–93; discussion 993.

[4] Liao AH, Wang CH, Weng PY, Lin YC, Wang H, Chen HK, Liu HL, Chuang HC, Shih CP. Ultrasound-induced microbubble cavitation via a transcanal or transcranial approach facilitates inner ear drug delivery. JCI Insight 01 2020;5.

[5] Yang Y, Zhang X, Ye D, Laforest R, Williamson J, Liu Y, Chen H. Cavitation dose painting for focused ultrasound-induced blood-brain barrier disruption. Scientific Reports 02 2019;9.

[6] Wu P, Wang X, Lin W. Acoustic characterization of cavitation intensity: A review. Ultrasonics Sonochemistry 12 2021;82:105878.

[7] Basu K, Sinha R, Ong A, Basu T. Artificial intelligence: How is it changing medical sciences and its future? Indian Journal of Dermatology 05 2020;65:365.

[8] International A. Ltd, material 850. URL `https://www.americanpiezo.com/`.

[9] Driscoll P. Gray's anatomy, 39th edition. Emergency Medicine Journal 06 2006;23:492–492.

[10] Treeby B, Cox BT. k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wavefields. Journal of biomedical optics 03 2010;15:021314.

[11] Haworth K, Bader K, Rich K, Holland C, Mast TD. Quantitative frequency-domain passive cavitation imaging. IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control 10 2016;PP:1–1.

[12] Chen X, Wang J, Pacella J, Villanueva F. Microbubble behavior during long tone-burst ultrasound excitation. Journal of the Acoustical Society of America 04 2015;137:2253–2253.

[13] Blake J, Leppinen D, Wang Q. Cavitation and bubble dynamics: The kelvin impulse and its applications. Interface Focus 10 2015;5:20150017.

[14] Mallat S. A Wavelet Tour of Signal Processing. 01 2009.

[15] MATLAB. Continuous 1-d wavelet transform, 2016.

[16] MATLAB. Deep learning toolbox model for alexnet network, 2020.

Address for correspondence:

Patricia Akemi Sekini Guenkawa
Avenida Professor Luciano Gualberto, Travessa 3, nº 158 – São Paulo - SP, Brazil
patricia.guenkawa@usp.br