Hermite Based Parametric Representation of Magnetohydrodynamic Effect for the Generation of Synthetic ECG Signals during Magnetic Resonance Imaging

Pierre G Aublin [1] Jacques Felblinger[1,2] Julien Oster [1,2]

1. IADI, Inserm U1254, Université de Lorraine
2. CIC-IT 1433, Université de Lorraine, Inserm, CHRU de Nancy

Abstract

Aim. *ECG signals during Magnetic Resonance Imaging (MRI) are often distorted by a strong magnetohydrodynamic (MHD) artefact. This effect is induced by charged particles in blood flowing under the MRI static magnetic field. Models of this artefact are hard to assess due to the absence of ground-truth measurements for this phenomenon. While large annotated ECG databases have been publicly released, there are only few small databases containing ECG signals acquired during MRI. This scarcity of data hinders the devel-opment of reliable automated ECG in MRI analysis solutions. We proposed a model to generate synthetic MHD artefacts to augment a dataset of standard ECG and to train deep learning models more robust to this artefact.* ***Methods.*** *An open database of ECG in MRI was used to extract a median MHD tem-plate over a small subject population. These templates were decomposed on a basis of Hermite functions to represent the MHD effect by a set of 29 parameters. A Gaussian mixture model was fitted on these coefficients, which allows MHD artefacts to be generated by sampling this probability distribu-tion. The model was assessed on a heartbeat classification task on an in-house database of ECG signals acquired in a 1.5T MRI scanner during standard clinical examination. A convolutional neural network (CNN) trained on the MIT-BIH arrhythmia (MITAR) database without pretraining was compared with models pretrained on the CinC 2021 database using the proposed MHD specific data augmentation.* ***Results.*** *The randomly initialized CNN, and the proposed augmentation obtained average F1 scores of 0.21, and 0.44 respectively on the in-house MRI database.* ***Conclusion.*** *The proposed MHD artefact generator can be used to effectively augment ECG data and learn a representation more robust to MRI environment distortions.*

1. Introduction

Electrocardiogram (ECG) is a well-known modality aiming at recording the electrical activity of the heart [1], whereas Magnetic Resonance Imaging (MRI) is an imaging modality that can depict functions or tissue characterization of this organ. ECG is mainly acquired during MRI exams to: (i) monitor the patient during the acquisition, (ii) synchronize the image acquisition with the cardiac activity and movement to ensure image quality during cardiac MRI [2]. ECG is highly distorted inside the MRI scanner due its electromagnetic environment and its three main characteristics: (i) RF pulses, which can induces burns; (ii) varying magnetic fields (so called gradients), which induces high amplitude voltage on the ECG electrodes; and (iii) the strong static magnetic field inside the MRI bore, which interacts with charged particles flowing inside the blood. This induces an extra voltage, so called magnetohydrodynamic (MHD) effect. This results in distorted signals that can impair the diagnosis use of the ECG recordings [2]. To our knowledge advanced analysis of the ECG signal during MRI, heartbeat and/or rhythm classification, is not yet available due to the MHD effect. MHD noise suppression on ECG remains challenging as there is no access to any ground-truth of the ECG signal inside the MRI scanner [2, 3].

A solution to circumvent this issue would be to develop features robust to this MRI specific noise [4]. While large ECG databases have recently been made publicly available [5], databases of ECG signals acquired in MRI remain small [6, 7]. It is therefore important to design a realistic MHD model that could be used to synthesize large databases of ECG with MRI specific noise, that could be used to build (deep learning) models for advanced analysis of ECG signals acquired in MRI.

This study proposes a synthetic model of the MHD effect, based on the sampling of density probability from a parametric representation of the MHD signal using an existing database of ECG signals acquired in MRI [7].

2. Methods

2.1. MHD Generator

Krug et al. released a database of ECG signals acquired in an MRI scanner [7]. This dataset contains 12-lead ECG signals, sampled at 1024Hz, recorded from patients outside and inside MRI scanners at various static magnetic fields. We focused on recordings performed on 3T and 7T scanners. Only patients for which signals were acquired while they were lying inside the MRI bore in head-first and feet-first supine position were kept, resulting in the inclusion of 10 pairs of ECG recordings (5 at 3T and 5 at 7T).

The pipeline is described on Figure 1. The ECGs cardiac cycles were segmented around the R peak based on the QRS position annotations. The RR intervals were mapped to a cyclic phase φ within the interval $[-\pi, \pi]$ (the R peak phase value was assigned the value $-\pi/3$). This phase shift ensured the the cardiac cycle to be approximatively centered around the T wave of the ECG ($\varphi \simeq 0$ at the T wave). Then an average heartbeat template for each lead in each recording was built (step 1). Inverting the position of the patient into the MRI bore (head-first or feet-first) inverts the polarity of the MHD artefact while not affecting the polarity of the ECG signal [3]. Taking the difference between the template of the heartbeat in head-first and feet-first thus roughly cancels the ECG contribution and provides an average MHD effect template (step 2).

$$h_i(\varphi) = \frac{(-1)^i}{\sqrt{2^i \sqrt{\pi}.i!}} \frac{d^i}{d\varphi^i}(e^{\frac{-\varphi^2}{2}}) \tag{1}$$

Hermite functions (eq. (1)) form a basis of orthogonal functions on $L^2(\mathbb{R})$ that can be used to decompose ECG waveforms [8]. In the third step, the average MHD template M could thus be parameterized by φ as:

$$M(\varphi) = \sum_{i=0}^{n} a_i h_i(\varphi). \tag{2}$$

We set the order of the decomposition to $N = 28$, the proposed procedure resulted in the estimation of 120 (12 leads multiplied by 10 subjects) sets of 29 Hermite coefficients a_i. A mixture of 12 Gaussians was then fitted on this training set (step4). Generating a new synthetic MHD template for a single-lead recording then simply consists in randomly sampling from this Gaussian mixture model, which will provide virtually an infinite numbers of the 29 Hermite coefficients (step 5) required for realistic MHD morphology.

2.2. Evaluation

In order to assess the usefulness of the proposed MHD model, a data (MHD) augmentation technique was used for the training of a heartbeat classifier in two consectuive steps. First, a convolutional neural network (CNN) was pretrained to build a representation of ECG heartbeats that would be invariant to artefacts generated by the proposed MHD model. This representation was then inputed to a simple linear model for heartbeat classification.

All ECG recordings were prefiltered and resampled to 250 Hz as in [4]. For a given QRS position, a 1.2s window was extracted with the analyzed QRS positioned at 2/3 of the window. The model architecture consisted in an 11-layer CNN followed by a global max pooling layer leading to 256 features [9]. A projection head and a predictor were added following [10] for the pretraining. These were removed and replaced by a single linear layer followed by a softmax activation function for the heartbeat classification task. All the neural networks were trained using a stochastic gradient descent (SGD) and a cosine annealing scheduler with warm restart [11].The maximum learning rate was set to 10^{-3}, with the scheduler hyperparameters set as in [10].

The Computing in Cardiology Challenge 2021 [5] database (CinC2021) was used for representation learning. Data from INCART [12] were removed resulting in a training set containing 88179 ECG recordings. The CNN was trained following a siamese scheme [10] to build in an unsupervised way a representation of ECG heartbeats. For each ECG recording, after preprocessing, a QRS detector [13] was run on lead II to build a cardiac phase φ and extract the 1.2 s windows. For each window two leads were randomly selected to build the two views needed for training the siamese network. Two MHD artefacts were then generated and added to both ECG views. The pretraining was done following the SimSiam approach [10].The model minimizing the loss over 100 epochs was kept as the heartbeat representation model. This model will be denoted $leadview + MHD$. A second model denoted $leadview$ was trained with only the random lead selection.

A linear evaluation [10] of the the ECG representation provided by $leadview + MHD$ was then performed. A heartbeat classifier was trained on the MITAR database [14] to detect ventricular (V') ectopic beats. The split DS1/DS2 proposed by Llamedo et al. [15] was used to ensure patient stratification between training and test set. DS1 was further split into 50% train, and 50% validation sets, stratified by beat class. DS2 was kept for model assessment. During training, the model with the best validation F1 score over 10 epochs was kept for evaluation. As a control, a randomly initialized CNN ($Randinit$) was trained from scratch with the same loss, optimizer and scheduler over 50 epochs. For $leadview$, $leadview + MHD$ and $Randinit$, 10 models were trained to enable statistical comparisons.

Final assesment on ECG acquired during MRI was per-

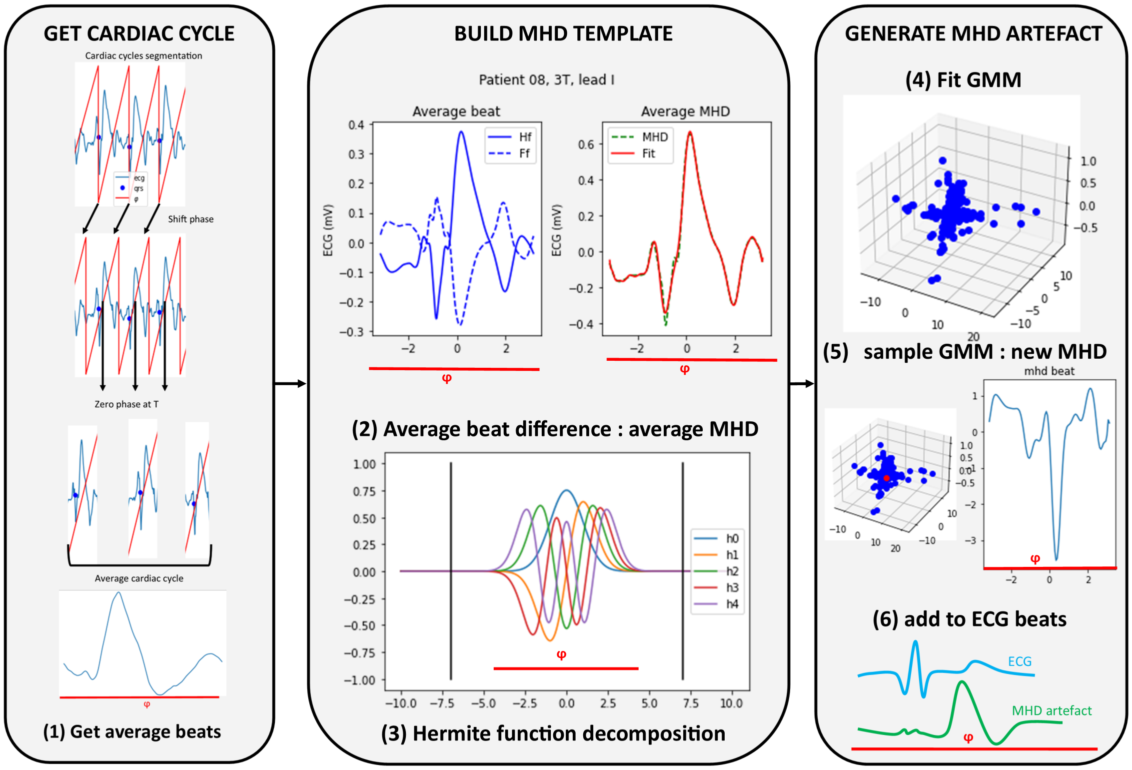

Figure 1. Pipeline to build the MHD generator.

Split	o-	i-	s-	all
N	5209	5506	12902	23617
V'	298	220	802	1320

Table 1. Beat distribution on the MRI database for the different patient configurations.

Dataset	Randinit	lead view	lead view+MHD
DS2	**0.88**±0.03	0.59±0.10	0.63±0.06
MRI			
o-	0.27±0.15	0.48±0.27	**0.62**±0.23
i-	0.21±0.08	0.30±0.17	**0.40**±0.15
s-	0.18±0.10	0.30±0.13	**0.40**±0.10
all	0.21±0.10	0.33±0.16	**0.44**±0.13

Table 2. F1 scores (mean and standard deviation) for the different training configurations on DS2 and MRI databases

formed on an in-house [4, 6] database. It contained signals acquired on a 1.5T MRI scanner, either outside the scanner (o-), inside the scanner without image acquisition (i-), during MRI pulse sequences (s-). The heartbeat classes distribution is summarized in Table 1.

3. Results

F1 scores obtained on DS2 and MRI databases are assembled in Table 2. $Randinit$ model, that is without pretraining, outperformed both finetuned models on DS2, but $Randinit$ results were not generalizable with significantly lower F1 performance on the MRI database. Comparison of both pretrained models showed that simple $leadview+MHD$ augmentation performed slightly better on DS2 than the simple $leadview$ but no statistically significant difference was found between both augmentations when using Wilcoxon test. On the in-house MRI database, $leadview$ was outperformed by the most MHD specific augmentation $leadview + MHD$ in all circumstances (inside, outside the MRI bore, and during sequences).

An average increase of 0.21 on the F1 score was observed when using the proposed data augmentation. A Wilcoxon test showed a difference between $Randinit$ and $leadview + MHD$ (p=0.02). Surprisingly, there is no difference in the self supervised learning approaches between the in bore (i-) and during the image acquisition (s-).

4. Discussion

In this work, we proposed a parametric representation of MHD artefacts based on Hermite function. Using a publicly available database we learnt the probability distribution of these Hermite parameters, which allows us to generate (by sampling) an infinite number synthetic MHD artefacts. We demonstrated that this approach can be used for data augmentation purpose and/or generate big databases of ECG data corrupted by synthetic MHD artefacts.

As demonstrated by the performance of the proposed data augmentation technique, the proposed MHD artefact generator can improve the robustness of the heartbeat classifier to ECG acquired in the MRI environment. Overall performance on the MRI database is relatively good, but there is a significant drop between ECG acquired outside the MRI bore and ECG acquired inside (i- and s-). This highlights the need for further improvements in the representation learning (more augmentations) but also maybe in the training of the classifier (using ECG in MRI for the training of linear classification?). As the focus of this study was on assessing the MHD artefact generator, neither denoising of the gradient artefacts [6], nor data augmentations related to the gradient noise were applied.

Comparison between i- and s- conditions for the self supervised approaches suggests the ECG representations may be already robust to gradient artefacts. This should be tested via introduction of gradient preprocessing or pretraining including data augmentations related to this artefact. The poor performance of $leadview$ on the MRI database suggests that MHD specific augmentation is useful, and seems to be indicating the generated MHD models are realistic enough to improve ECG in MRI analysis. Further improvements are required for the applicability of the solution in clinical practice, as current level performance do not allow for accurate and robust detection of ventricular heartbeats on ECG signals acquired in MRI.

Acknowledgments

This project was supported by a grant from the ERA-CVD Joint Translational Call 2019, MEIDIC-VTACH (ANR-19-ECVD-0004). We are grateful to France Life Imaging for financially supporting PA's attendance to the conference.

References

[1] Kligfield P. The Centennial of the Einthoven Electrocardiogram. Journal of Electrocardiology 2002;35(4, Part B):123–129. ISSN 0022-0736.

[2] Oster J, Clifford G. Acquisition of Electrocardiogram Signals During Magnetic Resonance Imaging. Physiological Measurement 2017;38(119).

[3] Oster J, et. al. Comparison of three Artificial Models of the Magnetohydrodynamic Effect on the Electrocardiogram. Computer Methods in Biomechanics and Biomedical Engineering April 2014;18(13):1400 – 1417. Publisher: Taylor & Francis.

[4] Aublin P, et. al. Automated Detection of Ventricular Heartbeats from Electrocardiogram (ECG) Acquired During Magnetic Resonance Imaging (MRI). Computing in Cardiology 2022;49:1–4.

[5] Reyna MA, et. al. Issues in the Automated Classification of Multilead ECGs using Heterogeneous Labels and Populations. Physiological Measurement August 2022; 43(8):084001. Publisher: IOP Publishing.

[6] Oster J, et. al. Nonlinear Bayesian Filtering for Denoising of Electrocardiograms Acquired in a Magnetic Resonance Environment. IEEE Transactions on Biomedical Engineering 2010;57(7):1628–1638.

[7] Krug J, et. al. A Database of Electrocardiogram Signals Acquired in Different Magnetic Resonance Imaging Scanners. 2017 Computing in Cardiology CinC 2017;1–4.

[8] Böck C, Kovács P, Laguna P, Meier J, Huemer M. ECG Beat Representation and Delineation by Means of Variable Projection. IEEE Transactions on Biomedical Engineering 2021;68(10):2997–3008.

[9] Aublin PG, et. al. Predict Alone, Decide Together: Cardiac Abnormality Detection based on Single Lead Classifier Voting. Physiological Measurement 2022;.

[10] Chen X, He K. Exploring Simple Siamese Representation Learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021; 15750–15758.

[11] Loshchilov I, Hutter F. SGDR: Stochastic Gradient Descent with Warm Restarts, May 2017. URL `http://arxiv.org/abs/1608.03983`. ArXiv:1608.03983 [cs, math].

[12] Tihonenko V, et. al. St Petersburg INCART 12-lead Arrhythmia Database. PhysioBank PhysioToolkit and PhysioNet 2008;.

[13] Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm. IEEE Transactions on Biomedical Engineering 1985; BME-32(3):230–236.

[14] Moody GB, Mark RG. The Impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine May 2001;20(3):45–50. ISSN 1937-4186.

[15] Chazal Pd, et. al. Automatic Classification of Heartbeats using ECG Morphology and Heartbeat Interval Features. IEEE Transactions on Biomedical Engineering 2004; 51(7):1196–1206.

Address for correspondence:

Julien Oster,
Bâtiment Recherche, Rue du Morvan, 54511 Vandœuvre-lès-Nancy, France
julien.oster@inserm.fr