Heart attack outcome predictions using FMM models

C. Canedo, A. Fernández-Santamónica, Y. Larriba, I. Fernández and C. Rueda

Department of Statistics and Operations Research, University of Valladolid, 47011, Valladolid, Spain

Abstract

As part of the PhysioNet Challenge 2023, our team FMMGroup_UVa presents an original approach for the prediction of the outcome of coma patients after a heart attack. Our methodology involves the integration of two types of EEG features extracted from 10-second epoch data, analyzed at various time intervals with patient clinical data. The first type is the FMM features created from the parameters of an FMM (Frequency Modulated Möbius) model fitted to the epoch data. The other type is indices taken from the literature that include spectral, entropy, and background measures. The mean feature values for patients are combined with clinical variables into classification models to obtain an outcome score in a given time block. Our best performance in the official challenge phase achieved a score of 0.45 at 72 hours on the hidden test. Additionally, we introduce an alternative proposal in this paper, displaying promising results in our laboratory. Unfortunately, this proposal was not ranked due to difficulties in submitting it on time without errors.

1. Introduction

The FMM approach is among the more interesting recent advancements in mathematical modeling techniques for the analysis of single-channel and multi-channel bioelectric signals. It has been successfully used to analyze electrocardiograms, neuronal spikes, or electroretinograms, among other biological signals [1,2]. FMM models for single channels are formulated as a sum of components defined in terms of four parameters. When multi-channel data is analyzed the individual components are connected, in such a way that there exists a set of parameters that are common to all channels, as well as other parameters that are channel-specific. As a result, the calculation of indices measuring channel-specific features and connectivity measures across channels is easily addressed. Moreover, the model parameters capture the morphology of the fundamental waves present in the signals and thus are easily interpretable and have great potential in tasks such as classification or prediction [2].

In addition, the FMM approach can effectively analyze signals regardless of the preprocessing steps applied, the type of recording device used, or the number of channels involved. These issues make the FMM approach versatile and adaptable to various data acquisition set-ups and conditions.

In this paper, we propose to derive helpful FMM features in the prediction outcomes of coma patients who suffered a heart attack [3–5]. This is the first time the FMM approach is used for EEG data analysis. For the aim of this paper, we propose extracting 10 components from each 10-second epoch. The rationale behind this strategy is rooted in the representation of slow waves, a well-documented phenomenon in the scientific literature. Slow waves, often referred to as delta waves, manifest as low-frequency, high-amplitude oscillations in the EEG. These distinctive electrical patterns are frequently associated with various sleep stages and periods of reduced neural activity. They reflect the synchronization of neural activity in large populations of neurons. Importantly, alterations in slow-wave activity can serve as indicators of brain injury or neurological disorders, as previously noted [6, 7]. Furthermore, 10-second epochs are commonly used as a standard unit of analysis in most studies. Moderately pronounced spiked components, indicative of slow waves, will be more prevalent in patients with favorable prognoses.

We create features using statistics from the FMM parameters that capture EEG core aspects for the prognosis, such as amplitude, complexity, connectivity, regularity, or spike counting. On the other hand, we also consider spectral, entropy, and background features [8–11]. Table 1 summarizes the features.

The patient average values for such features are combined with clinical variables included in classification models to obtain an outcome score in a given time block.

2. Methods

We assume that a raw digital EEG signal (time (seconds), voltage (microvolts)) is available for 18 bipolar channels. Occasionally, the voltage values are not the original ones but they have been scaled in some form. The challenge provided data over a 72-hour period. Our proposal can be described in three steps: the epoch selection, the creation of epoch-level features, and the development of patient classification models. Figure 1 illustrates the

Level	Type	Name	Statistics/Description	N^o
Epoch	FMM	α	Circular_$SD_l(\alpha_l)$ Watson test $\#\{\omega_l < 0.01\}$	2
		ω	$SD_l(\omega_l)$ $CV_l(\omega_l)$	3
		A	$ME_c(ME_l(A_{lc}))$ $CV_c(ME_l(A_{lc}))$	2
		R^2	$ME_c(ME_l(R^2_{lc}))$	1
		Binary rule	Features outside the percentile healthy range	1
Epoch	Entropy	Shannon	Individual values, Mean, SD	20
		Fractal dimension [11, 12]	Individual values, Mean, SD	20
	Spectral	$\delta, \theta, \alpha, \beta$ Frequency bands[8]	Individual values, Powers, Ratios.	126
	Background	BCI[9], BSAR[9]	Mean, SD	40
Patient	Clinical	Age, ROSC, Shockable rhythm	Individual values	3

Table 1: List of features by level and type. ME: median; SD: standard deviation and CV: variant coefficient. N^o : feature counts.

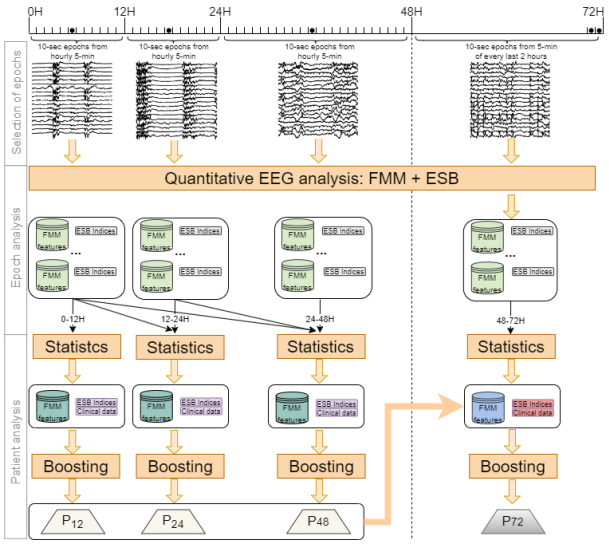


Figure 1: FMMGroup_UVa flowchart.

steps. Different strategies are designed within this general scheme. The fundamental keys of each step and the strategies considered are commented on below.

2.1. Selection of epochs

One pressing issue we have encountered is that the process of generating features is notably time-intensive. Furthermore, it is advisable to not select epochs of exceptionally low quality. Consequently, to remain competitive in the challenge, we found it necessary to fine-tune our algorithm to fit within the allotted computing time. A pivotal aspect of this adaptation involves the selection of epochs, which distinguishes our different solutions. One straightforward approach involved choosing a 5-minute window from the central portion of the last hour for each patient, and this yielded promising results over the course of 72 hours. Alternatively, we considered utilizing 5-minute segments from each hour for each patient. A solution combining both strategies produced the most favorable outcomes in our laboratory experiments. Moreover, distorted epochs can be identified and eliminated using a simplified version of the `EEG quality measure software` borrowed from the Challenge. The reason for using a simplified version is again to reduce the computational time.

2.2. FMM: model and features

The FMM model considered assumes that the voltage in a time point t for a given channel (c) is a combination of 10 parametrized components or waves (l), as follows:

$$\sum_{l=1}^{10} A_{lc} \cos\left(\beta_{lc} + 2\arctan\left(\omega_l \tan\left(\frac{t-\alpha_l}{2}\right)\right)\right).$$

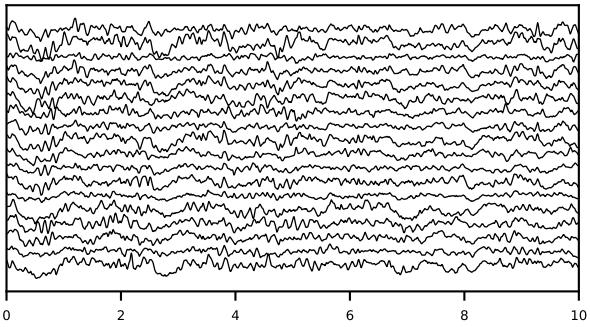
Where, $A_{lc} \in \Re^+$ is an amplitude parameter that measures the height of the waveform l in channel c, a proxy for the voltage or signal intensity. $\alpha_l \in [0, 2\pi)$ is a location parameter that identifies the time in [0,10] where the waveform l spikes. $\beta_{lc} \in [0, 2\pi)$ and $\omega_l \in [0, 1]$ describe the waveform shape. Specifically, ω_l is related to the width of the waveform l and to the frequency band. A spiked waveform is described with ω values close to 0, while ω close to 1 represents a wave lasting exactly 10 seconds.

The role of FMM parameters characterizing epochs is exemplified through the data presented in Figures 2 and 3. In particular, there are differences in the amplitude (A), width (ω), and wave locations (Watson statistic test), being these values higher for the case of less favorable prognosis (Figure 3).

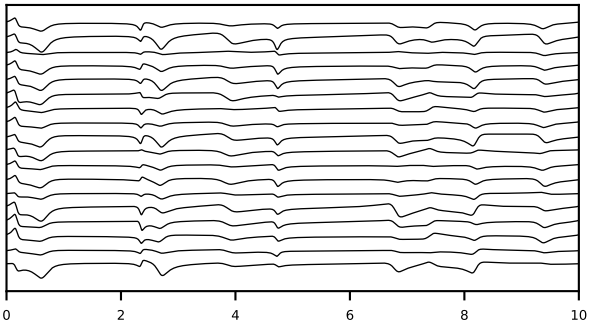
The observed voltages are used to estimate the model parameters for each epoch. We have adopted an approximation estimation algorithm different from that in the mentioned papers to reduce the computing time. In addition to the FMM parameters, we also consider the percentage of variance explained by the model, R^2_{lc}, for component l and channel c. The reader interested can see the papers on the topic [1, 2].

2.3. Patient classification models

A large collection of statistics from the estimates of parameters A, α, and ω have been calculated from values across channels and components for each epoch to create features. These include medians, percentiles, variability measures, and others. The FMM features that have proved

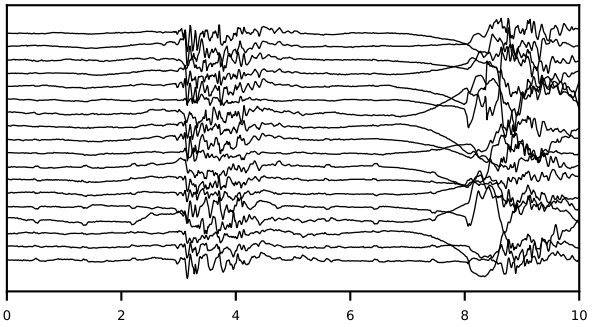


(a) EEG data

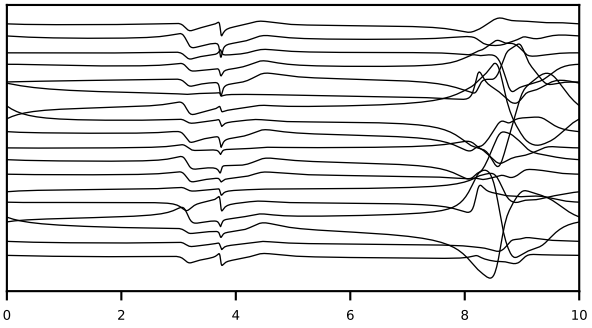


(b) FMM prediction

Figure 2: Typical 10-second epoch 18 channels of a patient with favorable prognosis.



(a) EEG data



(b) FMM prediction

Figure 3: Typical 10-second epoch 18 channels of a patient with non-favorable prognosis.

to be of relevance in classification models are listed in Table 1. Furthermore, we have also considered ESB (Entropy, Spectral, and Background) indices, see Table 1. For a given patient and time block, mean and standard deviation values, across epochs, of selected features have been calculated and have been included in classification models. We have considered logistic regression, gradient boosting, or random forest as classification models.

2.4. Strategies

Based on this scheme, we have explored various strategies that differ in terms of epoch selection, feature sets, and classification models. In this paper, we present two of them labeled as Strategy A and B.

Strategy A involves selecting all epochs from the central 5 minutes of the last hour of each patient. Feature selection comprises a preliminary binary epoch classifier using a set of FMM features, spectral features, as well as clinical variables. We constructed a random forest model using these selected features. On the other hand, in Strategy B (the recommended approach), the predictions obtained at 12, 24, and 48 hours are derived using the information of FMM and ESB indices of epochs available until that moment. We computed patient scores using boosting classification models that integrate FMM features and a selection of ESB indices, both detailed in Table 1. Additionally, predictions at 72 hours are derived by combining FMM features with a second set of selected ESB indices from the last two hours epochs within the 48-72 hour period. Boosting classification model incorporates these features in conjunction with clinical variables and predictions made at 12 hours, 24 hours, and 48 hours.

2.5. Results

We have conducted in our laboratory a comparative analysis of the two strategies which includes information from 603 patients following the Challenge rules. We have executed five experiments, each employing a 10-fold cross-validation approach.

The results for Strategy A have been computed for training, validation, and hidden dataset (see Table 2) along with our validation experiments (see Table 3). The true positive rate (TPR) for predicting a poor outcome given a false positive rate (FPR) of less than or equal to 0.05 at 72 hours after return of spontaneous circulation were 0.642 and 0.450

in the validation and hidden dataset, respectively. Our laboratory experiment agrees with the hidden dataset score. Regrettably, we did not arrive on time to achieve official Challenge scores for Strategy B. Nevertheless, we succeed in comparing both strategies in our laboratory experiments. The scores for Strategy B significantly outperform those of Strategy A. Theoretical ranks have been obtained for Strategy B: 1st, 14th, 13rd and 10th at 12, 24, 48 and 72 hours, respectively. These values suggest a greater potential for predicting outcomes with strategy B, particularly at 12 hours.

Table 2: TPR at 72 hours for Strategy A on the training, validation, and hidden test.

Strategy	Training	Validation	Hidden
A	0.877	0.642	0.450

Table 3: TPR mean (SD) for time block and strategy.

Strategy	Time block			
	12	24	48	72
A	0.236	0.226	0.324	0.442
	(0.031)	(0.021)	(0.035)	(0.033)
B	0.441	0.448	0.566	0.611
	(0.032)	(0.041)	(0.021)	(0.015)

3. Conclusions

Strategy A achieved scores of 0.642 and 0.450 on the validation and hidden tests, respectively. While we do not have the score for Strategy B. In our laboratory experiments, Strategy B outperforms Strategy A with scores of 0.442 and 0.611, respectively. In a theoretical ranking Strategy B at 12 and 72 hours would achieve 1st and 10th positions, respectively. Furthermore, this paper highlights the utility of the FMM approach in EEG data analysis, with a specific emphasis on demonstrating the predictive potential of these FMM features. Moreover, our proposal is original, marking the inaugural utilization of the FMM approach in EEG analysis. Once we have all the available data, a full comprehensive analysis will be conducted. This analysis will encompass the identification of anomalous epochs and the FMM analysis of all epochs with sufficient quality. With all this documentation in hand, we will once again compare the various proposals presented in this paper through validation experiments employing 10-fold cross-validation.

Furthermore, we will conduct a specific analysis of binary epoch classifiers based on FMM features and assess the predictive capability of other promising features, including those generated from β parameters, among others.

Acknowledgements

This work was supported in part by grant PID2019-106363RB-I00 given by the Spanish Ministry of Science, Innovation and Universities.

References

[1] Rueda C, Larriba Y, Lamela A. The hidden waves in the ECG uncovered revealing a sound automated interpretation method. Scientific reports 2021;11(1):1–11.

[2] Rueda C, Rodríguez-Collado A, Fernández I, Canedo C, Ugarte MD, Larriba Y. A unique cardiac electrocardiographic 3D model. Toward interpretable ai diagnosis. Iscience 2022;25(12):105617.

[3] Amorim E, Zheng WL, Ghassemi MM, Aghaeeaval M, Kandhare P, Karukonda V, Lee JW, Herman ST, Sivaraju A, Gaspard N, et al. The international cardiac arrest research (i-care) consortium electroencephalography database. Critical Care Medicine 2023 in press 2023;.

[4] Reyna MA, Amorim E, Sameni R, Weigle J, Elola A, Bahrami Rad A, Seyedi S, Kwon H, Zheng WL, Ghassemi M, et al. Predicting neurological recovery from coma after cardiac arrest: The george b. moody physionet challenge 2023. Computing in Cardiology 2023;50:1–4.

[5] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet. Circulation 2000;101(23):e215–e220.

[6] Jiang M, Su Y, Liu G, Huang H, Tian F. EEG pattern predicts awakening of comatose patients after cardiopulmonary resuscitation. Resuscitation 2020;151:33–38.

[7] Tian X, Li F. Application of Multimodal EEG in Coma Patients. Springer Singapore. ISBN 978-981-16-4493-1, 2022; 161–175.

[8] Admiraal M, Ramos L, Olabarriaga SD, Marquering H, Horn J, van Rootselaar A. Quantitative analysis of EEG reactivity for neurological prognostication after cardiac arrest. Clinical neurophysiology 2021;132(9):2240–2247.

[9] Ruijter BJ, Hofmeijer J, Tjepkema-Cloostermans MC, van Putten MJ. The prognostic value of discontinuous EEG patterns in postanoxic coma. Clinical neurophysiology 2018; 129(8):1534–1543.

[10] Zubler F, Tzovara A. Deep learning for EEG-based prognostication after cardiac arrest: from current research to future clinical applications. Frontiers in neurology 2023;14.

[11] Rubega M, Formaggio E, Molteni F, Guanziroli E, Di Marco R, Baracchini C, Ermani M, Ward NS, Masiero S, Del Felice A. EEG fractal analysis reflects brain impairment after stroke. Entropy 2021;23(5):592.

[12] Kannathal N, Choo ML, Acharya UR, Sadasivan P. Entropies for detection of epilepsy in eeg. Computer methods and programs in biomedicine 2005;80(3):187–194.

Address for correspondence:

Cristina Rueda

Paseo de Belén, 7; 47007 Valladolid, Spain.

cristina.rueda@uva.es