3D CNN as an Approach to Predict the Cerebral Performance of Comatose Patients

Rafael Teodoro Ors-Quixal[1], Elisa Ramírez-Candela [1], Samuel Ruipérez-Campillo [1,2], Francisco Castells-Ramón[1], José Millet[1,3]

[1] Universitat Politècnica de València, Valencia, Spain
[2] Swiss Federal Institute of Technology (ETH), Zurich, Switzerland
[3] Centro de Investigación Biomédica en Red Enfermedades Cardiovascular, Madrid, Spain

Abstract

Many patients remain in a comatose state after initially surviving a resuscitation following a cardiac arrest. The prognosis in this state carries the decision of life support withdrawal, thus needing an objective and deterministic guideline. The objective of this study, is to assist this decision by providing a model able to predict the cerebral performance category (CPC) of comatose patients following cardiac arrest from their electroencephalographic (EEG) signal. To achieve this, binary classifiers built with 3D Convolutional Neural Networks (CNNs) followed by Dense Neural Networks (DNN) are used in combination with a "divide and conquer" strategy, thus enabling the automatic extraction of features from the tensors of EEG signals, taking into consideration the spatial relation of the signals according to the electrodes' distribution on the scalp. This work was submitted under the team name "BioITACA_UPV" to "Predicting Neurological Recovery from Coma After Cardiac Arrest: The George B. Moody PhysioNet Challenge 2023", and while the team did not score in the official phase, results obtained from a held-out subset of the training set demonstrate the capability of the model to classify by CPC from short segments of 5 seconds to long recordings of EEG data. Results show an average accuracy of 0.76 between the CPC classifiers and capability to discern between a good or bad outcome prognosis.

1. Introduction

After surviving a cardiac arrest resuscitation, some patients may remain in a comatose state due to the brain damage produced by the generalized hypoxia. Protocols dictate that these patients should be kept under life support assistance until death or the physician's prognosis on the likelyhood of the patient to recover consciousness [1]. This prognosis results in keeping the life support system or its withdrawal. As can be seen, a false positive and its consequent life support withdrawal provokes the death of a patient who would have recovered consciousness in the following days. Some scales were implemented to avoid the subjectivity of the prognosis, such as the Glasgow Coma Scale or the Cerebral Performance Category (CPC) scale, composed by: CPC 1: Good neurological function and independency. CPC 2: Moderate neurological disability and dependency. CPC 3: Severe neurological disability. CPC 4: Unresponsive wakefulness syndrome. CPC 5: Dead.

Despite this, even after the use of some statistical or machine learning models, literature shows the prevalence of wrong prognosis [2]. Typically these machine learning models automatically analyze the signals, extracting features channel by channel with 1D CNN or by group of channels with 2D CNN [3, 4]. The problem with this approach is the channels' lack or reduced spatial resolution, respective to the distribution of the electrodes across the scalp, is lost.

Therefore, this study aims to evaluate a machine learning model that can extract features from the channels, keeping the relative distribution of the electrodes in space. To achieve this, a 3D CNN is employed to process continuous EEG data fragments. This model is evaluated on a database of 609 patients from EEUU and Europe.

2. Materials and Methods

2.1. Dataset

The data used in this study is extracted from the International Cardiac Arrest REsearch consortium (I-CARE) database [5], available at PhysioNet [6]. This dataset comprises populational data (sex, age...) and different signals such as electroencephalography and electrocardiography recordings from patients who were in a comatose state following a cardiac arrest.

 ISSN: 2325-887X DOI: 10.22489/CinC.2023.364

These EEG signals were continuously recorded, which means approximately 58,000 hours of collected EEG data. For each patient, the length of the recordings may vary from hours to days, depending on their clinical evolution. The patients' neurological outcomes, which were assessed 3-6 months after the hospital discharge, are labeled within the CPC scale, ranging from 1 to 5, and as good outcome (CPC 1 and 2) or bad outcome (CPC 3, 4, and 5). From a total of 609 patients given, the class distribution is as follows:

Table 1. CPC distribution across the given database

CPC	Nº Patients	Percentage (%)
CPC 1	181	29.72
CPC 2	44	7.22
CPC 3	22	3.61
CPC 4	9	1.48
CPC 5	353	57.97

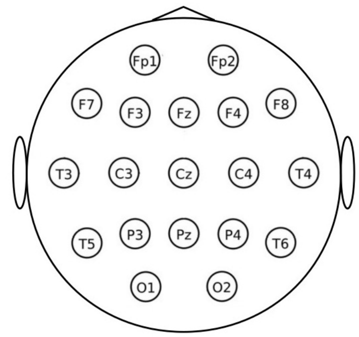

Figure 1. International 10-20 system electrodes distribution. Only the channels employed are shown in the graphic.

Ø	Fp1	Ø	Fp2	Ø
F7	F3	Fz	F4	F8
T3	C3	Cz	C4	T4
T5	P3	Pz	P4	T6
Ø	O1	Ø	O2	Ø

Figure 2. Template of channels following the international 10-20 system. The void symbols indicate an array of zeros employed to keep the orthoedral shape of the tensor.

2.2. Data Preprocessing

Due to the dynamic nature of the EEG data recorded following the cardiac arrest and the volume of data load, only the last registers of each patient were retrieved. Furthermore, in order to decrease the differences between the amount of available data from each CPC, only the last two registers were selected for CPCs 1, 2, 3 and 5, and the last 4 for CPC 4. Considering the hypothesis that as time passes after the cardiac arrest, the EEG signals become more characteristic for each CPC, increasing the amount of recordings used from past states of the patients may confuse the model, but due to the low availability of the CPC 4 data, this measure was taken in order to avoid techniques of data augmentation, that could lead to overfitting.

Once the registers were ready, the signal was filtered at 50 Hz with both a Notch and a low pass filters since part of the data was acquired in Europe. Then, in order to reduce and standardize the frequency of samples per second, which originally variated from 200 to 2048 Hz, the signals were resampled to 100 Hz in order to respect the Nyquist theorem. After the resampling, a segment of the first 30,000 samples (per channel) is taken in order to establish a common length for all registers. Then, the signals were reduced to the minimum amount of channels shown in the database, a total of 19 channels configured under the international 10-20 system and ordered into a tensor designed to keep the same relative spatial distribution of the electrodes across the scalp that was used during the recording. The template used to reorder the channels and the international 10-20 system are shown in Figures 1, 2.

To conclude with the preprocessing, a "divide et vinces" strategy is followed in order to facilitate parallelization, diminish the trained space of the models, and increase the number of data on hand. The 30.0000 samples (per channel) segments are then divided into multiple fragments of 500 samples (5 seconds of the record) as can be seen in Figure 3.

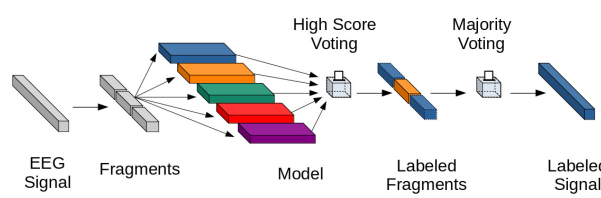

Figure 3. Partition architecture of the data.

2.3. Model Design

The model used is based on a group of submodels with the same architecture. In order to design it, different architectures were tested dynamically, choosing always power of two parameters [7]. The best architecture found is employed by every submodel and follows a sequential architecture composed by two initial 3D CNN layers. Its output is flattened and introduced into a set of three dense neural layers whose exit ends in the final layer, to give the binary prediction of class belonging with a softmax activation function. Despite this last layer and the flatten, each other layer uses a ReLu activation function, and works in a decreasing order of complexity as shown in Figure 4.

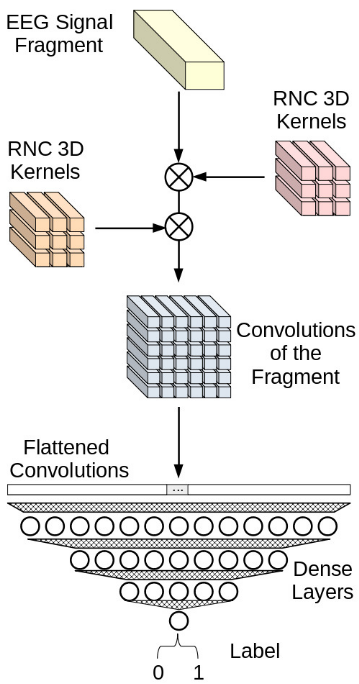

Figure 4. Architecture of the submodels designed in the study.

Once the segments are classified by each submodel, the final CPC category is given in function of how all the segments from the signal's fragment have been labeled. This is done by majority voting with the segments labels, and in case of a draw between classes, the better outcome class is prioritized, resulting: CPC1 $\prec$ CPC2 $\prec$ CPC3 $\prec$ CPC4 $\prec$ CPC5. The same preference applies in the higher score voting used to determinate the label of each segment from the signal.

2.4. Evaluation

The submodels are designed to classify five-seconds fragments from a total of five minutes of continuous EEG signal recordings within the CPC scale. Therefore, in order to evaluate their effectiveness the accuracy, F1-Score and the AUC-ROC values are calculated for both the fragments and segments (patients) classification.

The data was partitioned helding-out patients from the training set into the groups of train, test and validation patients. Given the dependence of the volume of data according to CPC 4, ($\sim$1:40 with respect to CPC5), and with the aim of increasing statistical significance, the models have been evaluated with a k-fold of 10, selecting different patients for each group in every fold.

3. Results

The evaluation of the submodels shows similar performances across CPCs, employing the same architecture parameters as seen in Figure 5 and Tables 2 and 3.

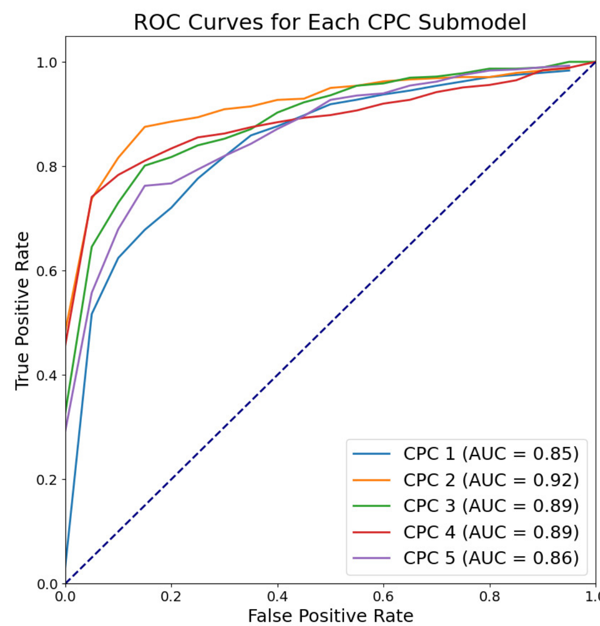

Figure 5. Averaged 10-fold AUC-ROC results for each CPC submodel.

Table 2. Metrics evaluated for each submodel.

	Accuracy	F1-Score	AUC
CPC 1	0.79	0.75	0.85
CPC 2	0.79	0.86	0.92
CPC 3	0.75	0.80	0.89
CPC 4	0.72	0.85	0.89
CPC 5	0.75	0.80	0.86

Despite these good results, the code submitted for evaluation during the official phase did not score. Therefore, we have neither score nor rank. Nevertheless testing the code with a group of held-out patients from the training set, the model archieved a challenge score [2] of 0.833.

Table 3. Statistical results for each CPC submodel with a support of 1200 segments for each class. True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN).

CPC	TP (%)	FP (%)	TN (%)	FN (%)
CPC 1	0.99	0.01	0.64	0.36
CPC 2	0.99	0.01	0.91	0.09
CPC 3	0.99	0.01	0.73	0.27
CPC 4	0.99	0.01	0.83	0.17
CPC 5	0.99	0.01	0.60	0.40

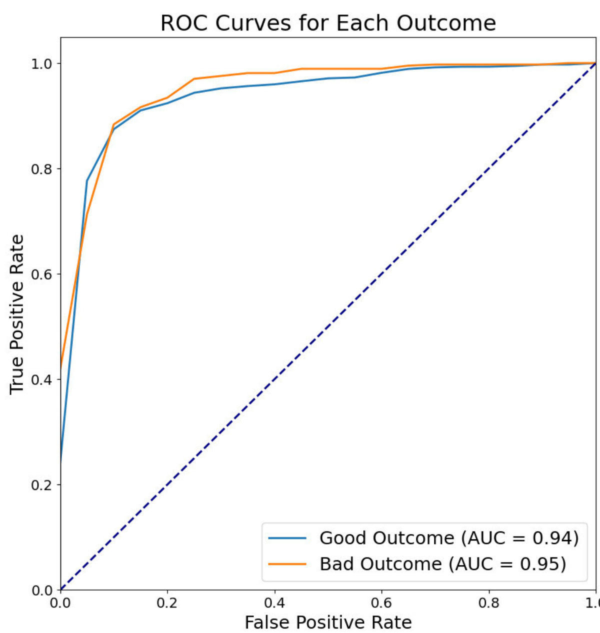

Figure 6. Averaged 10-fold AUC-ROC results for each outcome group.

4. Conclusion

Despite the fact that a lower true positive rate is observed in CPC 1 and 5 submodels, as can be observed in Figure 6 and Table 3, taking a look at Figure 6, most of the submodels missclassifications stay in their respective outcome. Which means in one hand, that the majority of CPC1 misslabels of its class segments are labeled into CPC 2, and in the other hand, that the CPC 5 submodel is more conservative with its inclusion criteria, which benefits the patient in this extreme case.

The obtained results indicate that 3D CNN + DNN based models show capability in the field of comatose patients EEG classification into the CPC scale. This study shows how combining the “divide and conquer” strategy with binary classifiers a good result can be obtained, even in complex multi-class classification problems.

Acknowledgments

This work was supported by PID2019-109547RB I00 (National Research Program, Ministry of Science and Innovation, Government of Spain) and CIBERCV CB16/11/00486 (Instituto de Salud Carlos III).

References

[1] Sandroni C, Cariou A, Cavallaro F, Cronberg T, Friberg H, Hoedemaekers C, Horn J, Nolan JP, Rossetti AO, Soar J. Prognostication in comatose survivors of cardiac arrest: An advisory statement from the european resuscitation council and the european society of intensive care medicine. Intensive Care Medicine 12 2014;40:1816–1831. ISSN 0342-4642.

[2] Reyna, MA*, Amorim, E*, Sameni S, Weigle J, Elola A, Bahrami Rad A, Seyedi S, Kwon H, Zheng W, Ghassemi M, van Putten M, Hofmeijer J, Gaspard N, Sivaraju A, Herman S, Lee J, Westover, MB**, Clifford, GD**. Predicting neurological recovery from coma after cardiac arrest: The George B. Moody PhysioNet Challenge 2023. Computing in Cardiology 2023 ;50:1-4.

[3] Tjepkema-Cloostermans MC, da Silva Lourenço C, Ruijter BJ, Tromp SC, Drost G, Kornips FHM, Beishuizen A, Bosch FH, Hofmeijer J, van Putten MJAM. Outcome prediction in postanoxic coma with deep learning*. Critical Care Medicine 10 2019;47:1424–1432. ISSN 0090-3493.

[4] Xu Z, Wang J, Wang R, Zhang Z, Yang S. Deep convolutional neural network for detection of disorders of consciousness. IEEE. ISBN 978-9-8815-6390-3, 7 2020; 7084–7089.

[5] Amorim E, Zheng WL, Ghassemi MM, Aghaeeaval M, Kandhare P, Karukonda V, Lee JW, Herman ST, Sivaraju A, Gaspard N, Hofmeijer J, Van Putten MJAM, Sameni R, Reyna MA, Clifford GD, Westover MB. The international cardiac arrest research consortium electroencephalography database. Critical Care Medicine October 2023;ISSN 0090-3493.

[6] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Circulation June 2000;101(23). ISSN 0009-7322, 1524-4539.

[7] Su Y, Zhang Z, Li X, Zhang B, Ma H. The multiscale 3D convolutional network for emotion recognition based on electroencephalogram. Frontiers in Neuroscience 8 2022;16. ISSN 1662-453X.

Address for correspondence:

Rafael Teodoro Ors Quixal: raorqui@inf.upv.es

ITACA Institute, Universitat Politècnica de València, 8G, 1st, Camino de Vera S/N, 46022, Valencia, Spain