Improvement of an automated neonatal seizure detector using a post-processing technique

A. H. Ansari [1], V. Matic [1], M. De Vos [2], G. Naulaers [3], P. J. Cherian [4], and S. Van Huffel [1]

Abstract— **Visual recognition of neonatal seizures during continuous EEG monitoring in neonatal intensive care units (NICUs) is labor-intensive, has low inter-rater agreement and requires special expertise that is not available around the clock. Development of an accurate automated seizure detection system with a low false alarm rate will support clinical decision making and alleviate significantly the workload. However, this is an ongoing difficult challenge for engineers as the neonatal EEG signal is non-stationary and often includes complex patterns of seizures and artifacts. In this study, we show an improvement of our previously developed neonatal seizure detector (developed using heuristic if-then rules). In order to improve the detection accuracy, mean phase coherence as a new feature is used to characterize artifacts and also support vector machine is applied to perform the post-processing step to remove false detections. As a result, the false alarm rate drops 42% (from 2.6 h^{-1} to 1.5 h^{-1}), whereas the good detection rate reduces only by 4%.**

I. Introduction

Since neonatal seizures are one of the most important signs of acute brain dysfunction [1] and need immediate medical attention [2], their early detection is vital in the neonatal intensive care units (NICUs) [3]. Analysis of electroencephalography (EEG) signals is the optimum way for identification, classification, and quantification of seizures [4]. However, EEG also expresses non-seizure activity such as sharp waves and sometimes the signal is contaminated by rhythmic appearing artifacts [5]. Although human experts can recognize seizures among such artifacts with their knowledge and experience, it is labor-intensive, expensive and has only modest inter-rater agreement. Thus, an accurate automated seizure detector is helpful in the NICUs [6]–[8].

In general, two groups of automated seizure detection algorithms have been proposed: heuristic models based on if-then rules (white-box models) [9]–[13] and data-driven models based on machine learning techniques (black-box models) [14]–[17]. The former ones are transparent and the output is interpretable because they partially mimic the seizure detection procedure of a human expert. On the other hand, the latter ones are more flexible and able to be retrained on several datasets.

Furthermore, a multi-stage classifier has been proposed in [18]. First, an artificial neural network (ANN) and a clustering technique are used to detect and cluster seizures (stage I, II) and then a heuristic model is applied to remove artifacts (stage III). In our method outlined in this paper, a similar procedure including a heuristic model and a data-driven model is used in order to improve the performance of our previously developed heuristic model. However, the order of the stages is inverted. First, a heuristic model mimicking a human expert detects seizures and then a pre-trained classifier reduces false alarms with the aid of machine learning techniques. In this way, although a data-driven model is used, the overall classifier is still transparent and the outputs are rather interpretable.

The reason for using such a multi-stage procedure is that some detections of the heuristic classifier are false because of thresholds and inflexibility of the heuristic model. Therefore, in this paper, a support vector machine (SVM) has been trained and is used to remove some of the false detections. The SVM uses time and frequency-based features concatenated with some features of the outputs of the heuristic model. In addition, mean phase coherence (MPC) is calculated between EEG and polygraphic signals in order to characterize the potentially existing artifacts.

II. Methods

A. The heuristic model

The heuristic algorithm is described in [6] in detail. Briefly, it consists of two independent, parallel detectors: spike-train detector and oscillatory seizure detector. In the former one, first, Teager-Kaiser nonlinear energy (TKE) as the only feature of the signal is extracted. Then, based on this feature and two pre-defined thresholds, sequential spikes are recognized in each EEG channel. If the cross correlations of sequential spikes satisfy some secondary thresholds, that period of the signal is marked as a spike-train seizure.

* A.H.A. V.M. and S.V.H. are supported by: Research Council KUL: GOA/10/09 MaNet, CoE PFV/10/002 (OPTEC); Flemish Government: FWO: projects: G.0427.10N (Integrated EEG-fMRI), G.0108.11 (Compressed Sensing) G.0869.12N (Tumor imaging) G.0A5513N (Deep brain stimulation); IWT: projects: TBM 080658-MRI (EEG-fMRI), TBM 110697-NeoGuard; iMinds Medical Information Technologies SBO 2015, ICON: NXT_Sleep; Belgian Federal Science Policy Office: IUAP P7/19/ (DYSCO, 'Dynamical systems, control and optimization', 2012-2017); Belgian Foreign Affairs- Development Cooperation: VLIR UOS programs (2013-2019); EU: EU MC ITN TRANSACT 2012, #316679, ERASMUS EQR: Community service engineer , #539642-LLP-1-2013, INTERREG IVB NWE programme #RECAP 209G; European Research Council: ERC Advanced Grant, #339804 BIOTENSORS

[1] A. H. Ansari, V. Matic, and S. Van Huffel are with the KU Leuven, Department of Electrical Engineering-ESAT, STADIUS, and iMinds Medical IT, Leuven, Belgium (e-mail: amir.hosseinansari@esat.kuleuven.be)

[2] M. De Vos, Institute of Biomedical Engineering, Department of Engineering, University of Oxford, Oxford, UK

[3] G. Naulaers, Neonatal Intensive Care Unit, University Hospital Gasthuisberg, Leuven, Belgium

[4] P. J. Cherian, Section of Clinical Neurophysiology, Department of Neurology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

In the oscillatory detector, the EEG signal is decomposed using the discrete wavelet transform (DWT) to characterize the increase of low-frequency activities. Then, autocorrelation analysis is used to quantify continuous repetitive low-frequency activities in EEG. Finally, those repetitive activities that exceed an adaptive threshold are defined as oscillatory seizures.

To design this algorithm, the visual seizure detection method of an expert clinical neurophysiologist was simplified and mathematically formulated. Then, the thresholds and parameters were tuned based on trial and error in order to minimize the difference between the expert's scores and automated detections.

B. Dataset

The used datasets (training and test) were recorded at the Sophia Children's Hospital (part of the Erasmus University Medical Center Rotterdam, The Netherlands). For training the SVM, 17 neonates previously used in [13] (for designing and tuning the heuristic model) are reused. Each neonate has on average 27 hours of recorded EEG-polygraphy signal including electrocardiogram (ECG), electrooculogram (EOG), chin or limb surface electromyogram (EMG), and abdominal respiratory movement signal (Resp.). Moreover, for testing the algorithms, 18 neonates recorded similarly in the same center are used. The data were filtered between 1 and 20 Hz. Next, 20 bipolar channels were formed based on the full-montage using 17 scalp electrodes placed according to the 10-20 International System [19]. In this work, dubious seizures and artifacts are not removed manually. TABLE I. lists the duration and the number of seizures of each neonate (test dataset).

C. post-processing method

The following steps indicate the procedure of the proposed post-processing technique.

1. The heuristic model operates on each channel of the EEG signals and defines some segments as seizure.
2. Each segment is split into 8s epochs with 4s overlap by a rectangular window. Then, 64 features of each epoch are extracted.
3. Epochs are classified by a pre-trained SVM classifier into "truly detected" and "falsely detected" groups.
4. The segments including more than 50% "falsely detected epochs" are known as artifact (Majority voting of epochs) and are removed.
5. This procedure is employed on all channels separately. If at least one channel has a seizure segment, that moment is marked as seizure ("OR" operator).

The SVM used radial basis function (RBF) as kernel and the hyper-parameters are optimized by leave-one-out (LOO) cross-validation on the training dataset. The test dataset was not used for designing the heuristic model, training the SVM, or tuning its hyper-parameters.

D. Feature extraction

In total, 64 features are measured in three groups. Note that these features are measured for each channel separately.

I) *Segment Features:* The first five features are extracted from the detected segments (outputs of the heuristic model) 1) Length of segment in sec, 2) number of channels which detected this segment, 3) seizure type (spike-train, oscillatory, or mixed), 4,5) (x, y) coordinates of the channel on the brain using a simplified two-dimensional grid with origin on channel C_z.

II) *Signal Features:* The second group consists of 55 features including frequency domain (28), time domain (23), and information theory (4) as presented in [20]. In our work, no analysis is performed to find the best set of the *signal features*.

III) *MPC Features:* The remaining four features are extracted by mean phase coherence (MPC) between the EEG and polygraphic signals: ECG, EOG, EMG, and Resp. These four features can characterize the phase synchronization of EEG with the mentioned polygraphic signals in order to quantify their correspondence to artifacts. The MPC is calculated by:

$$R = \left(\left[\frac{1}{N} \sum_j \sin \phi(j\Delta t) \right]^2 + \left[\frac{1}{N} \sum_j \cos \phi(j\Delta t) \right]^2 \right)^{\frac{1}{2}}, \quad (1)$$

where N is the number of samples, $j \in [0, N-1]$, and $\phi(t)$ is the phase difference between two signals as:

$$\phi(t) = \phi_a(t) - \phi_b(t) = \arctan \frac{\widetilde{s_a}(t) s_b(t) - s_a(t) \widetilde{s_b}(t)}{\widetilde{s_a}(t) \widetilde{s_b}(t) + s_a(t) s_b(t)}, \quad (2)$$

with $s_x(t)$ and $\widetilde{s_x}(t)$ denoting the signal and its Hilbert transform. The latter one is measured by:

$$\widetilde{s_x}(t) = -i\, \mathcal{F}^{-1}\{\mathcal{F}\{s_x(t)\} sign(\omega)\}, \quad (3)$$

where $\mathcal{F}\{.\}$ and $\mathcal{F}^{-1}\{.\}$ are the Fourier transform and its inverse respectively [21].

Before using the extracted features, all of them are statistically normalized with mean and standard deviation of the features of the training dataset. Figure 1. gives the schematic overview of the post-processing and feature extraction strategy.

III. Results and Discussion

In order to measure the performance of the proposed post-processor compared to the original reference model (II. A), good detection rate (GDR %), false alarm rate (FAR), and positive predictive value (PPV %) are measured by (4-6).

$$GDR = \frac{TP}{TP+FN} \times 100 \quad (4)$$

$$FAR = \frac{FP}{total\ time\ (in\ hour)} \quad (5)$$

$$PPV = \frac{TP}{TP+FP} \times 100 \quad (6)$$

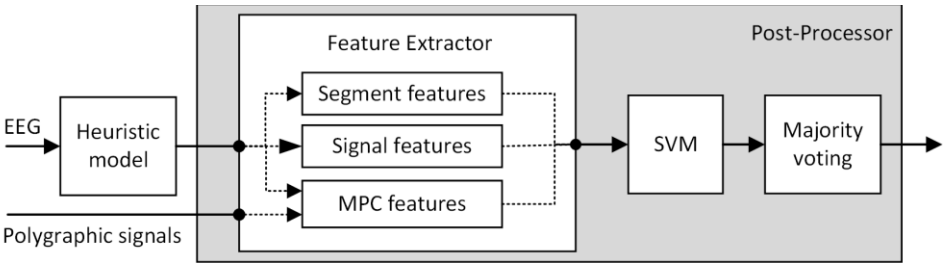


Figure 1. Diagram of the post-processing technique. The input of the post-processor is the detected segments of Heuristic model in addition to polygraphic signals.

TP is number of truly detected seizures, FN is the number of missed seizures, and FP is the number of false alarms [13], [20].

When the heuristic method is applied on the test dataset, expectedly some nonseizure segments are also falsely detected. Then, the post-processor recognizes and removes some of those falsely detected segments. Therefore, the GDR will not increase at all. The ideal performance of the post-processor is reached when the GDR and the FAR equal α and 0 respectively (α is the GDR of the original method). In practice, the post-processor also removes some true detections undesirably and, therefore, the GDR decreases as well as the FAR. In this case, one can only say that the performance is improved when the decrease of FAR is significantly larger than the decrease of GDR.

TABLE I. depicts the list of neonates in addition to the GDR and FAR (per hour) of the original method (heuristic detector without post-processing) as well as of the proposed post-processor on the test dataset. Patient no. 1, 4, 9, 14, and 15 have long-duration artifacts (electrode and/or respiration) which are effectively improved by the post-processor. However, FARs of no. 2 (with many slow oscillations) and no. 3 (with ECG artifact) are not decreased. The reason is that the training dataset includes few ECG artifacts and such oscillations. Therefore, it was not rich enough for training these patterns. Moreover, in number 8, many falsely detected movement artifacts are successfully removed. However, the GDR is dropped by the post-processor. The reason is the existence of many dubious seizures having a similar pattern with artifacts and the SVM is not able to distinguish between them.

Figure 2. shows the receiver operating characteristic (ROC) curve of the proposed technique (dashed line) on the test dataset in addition to the performance of the original method (A). When the threshold is one, clearly none of the events are removed. Consequently, the performance is the same as the original (A). However, decreasing the threshold results in more events being removed and the FAR and GDR reduced in turn. (B) is a selected point of the ROC curve with a trade-off between the GDR and the FAR. The presented data in TABLE I. (*Proposed* column) correspond to (B) in Fig 2. Comparison of (A) and (B) demonstrates that the post-processor decreases the FAR more than the GDR (42% vs. 4%) which means the performance is improved on average.

Figure 3. shows the histogram of the number of detected seizures (A) and false alarms (B) as a function of duration. It is clear that the proposed method removed false alarms more than truly detected seizures in all bars. Hence, not only the long-duration artifacts and status epilepticus are recognized, but also the short events are better classifiable by this technique. Furthermore, comparison of the first bar of (A) and (B) shows that the PPV of the original method for short events (< 30s) is really poor (38%) while it is improved (to 50%) by the post-processor. The overall PPV also increased with 12%, demonstrating the increased reliability of the detector. Furthermore, the total length of false alarm reduces with 15%. However, the trade-off is a decrease of 5% in the detected seizure burden.

TABLE I. PERFORMANCE OF THE METHODS FOR EACH NEONATE

N	*Length (hour)*	*Num. of Seiz.*	*Original*		*Proposed*	
			GDR %	*FAR*	*GDR %*	*FAR*
1	21	115	97	2.7	92	1.0
2	24	53	74	2.5	74	2.4
3	21	272	92	4.5	92	4.3
4	28	12	100	2.9	83	0.8
5	24	18	22	0.2	22	0.0
6	24	49	88	1.0	88	1.0
7	24	23	74	1.7	74	1.1
8	25	93	62	6.5	35	2.3
9	23	33	88	2.8	79	1.3
10	25	13	62	0.0	62	0.0
11	24	60	43	0.5	43	0.5
12	24	174	49	1.2	49	1.0
13	24	95	61	0.2	60	0.2
14	36	200	74	4.2	70	2.8
15	53	243	78	7.4	74	4.1
16	24	8	88	0.2	88	0.1
17	28	98	84	0.5	84	0.4
18	38	198	80	1.4	73	0.7

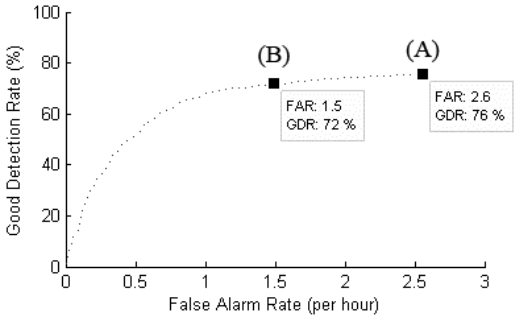


Figure 2. ROC of the seizure detector using the post-processor. (A) is the result of the heuristic detector without using the post-processor. The tiny dots are depicting the ROC of the post-processing method and (B) is a selected point with a rational trade-off between GDR and FAR.

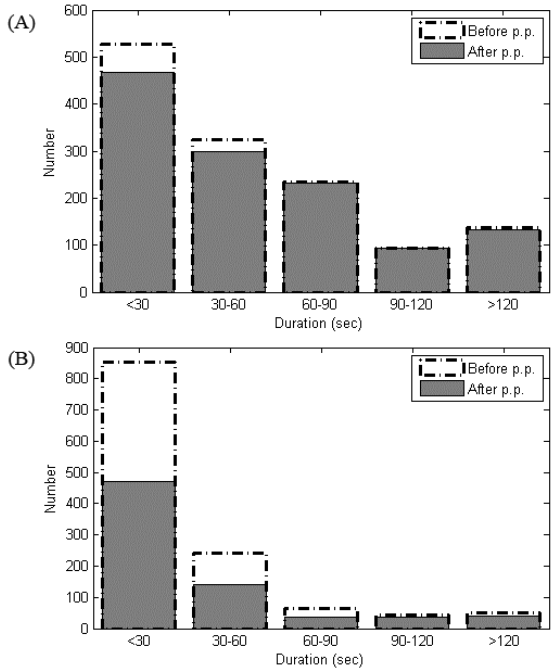


Figure 3. Histogram of the detected seizures (A) and false alarms (B) before post-processing (p.p., dashed bars) and afterwards (grey bars)

IV. Conclusion

In this work, support vector machine was used to reduce the false alarm rate of a heuristic detector. Furthermore, mean phase coherence was applied in order to characterize different types of artifacts in the feature extraction process. These changes successfully improved the overall performance such that the false alarm rate drops 42% while the good detection rate decreases only 4%. However, the disadvantage of this method is that the whole signal of a segment (a seizure or an artifact) should be available to be post-processed. Hence, the technique is applied after the offset of each event. Thus, it is not a fully online technique.

References

[1] J. J. Volpe, *Neurology of the Newborn*, 5th ed. Philadelphia: Saunder WB, 2008.

[2] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan, "A comparison of quantitative EEG features for neonatal seizure detection," *Clin. Neurophysiol.*, vol. 119, no. 6, pp. 1248–1261, Jun. 2008.

[3] L. S. Smit, R. J. Vermeulen, W. P. F. Fetter, R. L. M. Strijers, and C. J. Stam, "Neonatal Seizure Monitoring Using Non-Linear EEG Analysis," *Neuropediatrics*, vol. 35, no. 6, pp. 329–335, Nov. 2004.

[4] J. M. Rennie, G. Chorley, G. B. Boylan, R. Pressler, Y. Nguyen, and R. Hooper, "Non-expert use of the cerebral function monitor for neonatal seizure detection," *Arch. Dis. Child. - Fetal Neonatal Ed.*, vol. 89, no. 1, pp. F37–F40, Jan. 2004.

[5] G. Coppola, P. Plouin, C. Chiron, O. Robain, and O. Dulac, "Migrating Partial Seizures in Infancy: A Malignant Disorder with Developmental Arrest," *Epilepsia*, vol. 36, no. 10, pp. 1017–1024, Oct. 1995.

[6] W. Deburchgraeve, "Development of an automated neonatal EEG seizure monitor," PhD Thesis, Katholieke Universiteit of Leuven, Leuven, 2010.

[7] P. J. Cherian, W. Deburchgraeve, R. M. Swarte, M. De Vos, P. Govaert, S. Van Huffel, and G. H. Visser, "Validation of a new automated neonatal seizure detection system: A clinician's perspective," *Clin. Neurophysiol.*, vol. 122, no. 8, pp. 1490–1499, Aug. 2011.

[8] M. De Vos, W. Deburchgraeve, P. J. Cherian, V. Matic, R. M. Swarte, P. Govaert, G. H. Visser, and S. Van Huffel, "Automated artifact removal as preprocessing refines neonatal seizure detection," *Clin. Neurophysiol.*, vol. 122, no. 12, pp. 2345–2354, Dec. 2011.

[9] A. Liu, J. S. Hahn, G. P. Heldt, and R. W. Coen, "Detection of neonatal seizures through computerized EEG analysis," *Electroencephalogr. Clin. Neurophysiol.*, vol. 82, no. 1, pp. 30–37, Jan. 1992.

[10] J. Gotman, D. Flanagan, B. Rosenblatt, A. Bye, and E. M. Mizrahi, "Evaluation of an automatic seizure detection method for the newborn EEG," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 3, pp. 363–369, Sep. 1997.

[11] P. Celka and P. Colditz, "A computer-aided detection of EEG seizures in infants: a singular-spectrum approach and performance comparison," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 5, pp. 455–462, May 2002.

[12] M. A. Navakatikyan, P. B. Colditz, C. J. Burke, T. E. Inder, J. Richmond, and C. E. Williams, "Seizure detection algorithm for neonates based on wave-sequence analysis," *Clin. Neurophysiol.*, vol. 117, no. 6, pp. 1190–1203, Jun. 2006.

[13] W. Deburchgraeve, P. J. Cherian, M. De Vos, R. M. Swarte, J. H. Blok, G. H. Visser, P. Govaert, and S. Van Huffel, "Automated neonatal seizure detection mimicking a human observer reading EEG," *Clin. Neurophysiol.*, vol. 119, no. 11, pp. 2447–2454, Nov. 2008.

[14] M. Roessgen, A. Zoubir, and B. Boashash, "Seizure detection of newborn EEG using a model-based approach," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 6, pp. 673–685, Jun. 1998.

[15] P. Zarjam, M. Mesbah, and B. Boashash, "Detection of newborn EEG seizure using optimal features based on discrete wavelet transform," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*, 2003, vol. 2, pp. II–265–8 vol.2.

[16] A. Aarabi, R. Grebe, and F. Wallois, "A multistage knowledge-based system for EEG seizure detection in newborn infants," *Clin. Neurophysiol.*, vol. 118, no. 12, pp. 2781–2797, Dec. 2007.

[17] A. Temko, E. Thomas, G. Boylan, W. Marnane, and G. Lightbody, "An SVM-based system and its performance for detection of seizures in neonates," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009. EMBC 2009*, 2009, pp. 2643–2646.

[18] J. Mitra, J. R. Glover, P. Y. Ktonas, A. Thitai Kumar, A. Mukherjee, N. B. Karayiannis, J. D. Frost, R. A. Hrachovy, and E. M. Mizrahi, "A Multi-stage System for the Automated Detection of Epileptic Seizures in Neonatal EEG," *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 26, no. 4, pp. 218–226, Aug. 2009.

[19] P. J. Cherian, R. M. Swarte, and G. H. Visser, "Technical standards for recording and interpretation of neonatal electroencephalogram in clinical practice," *Ann. Indian Acad. Neurol.*, vol. 12, no. 1, pp. 58–70, 2009.

[20] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with Support Vector Machines," *Clin. Neurophysiol.*, vol. 122, no. 3, pp. 464–473, Mar. 2011.

[21] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Phys. Nonlinear Phenom.*, vol. 144, no. 3–4, pp. 358–369, Oct. 2000.