Clinical Neurophysiology 127 (2016) 3014–3024

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Improved multi-stage neonatal seizure detection using a heuristic classifier and a data-driven post-processor

A.H. Ansari [a,b,*], P.J. Cherian [c,d], A. Dereymaeker [e], V. Matic [a,b], K. Jansen [e,f], L. De Wispelaere [g], C. Dielman [h], J. Vervisch [e], R.M. Swarte [g], P. Govaert [g,h], G. Naulaers [e], M. De Vos [i], S. Van Huffel [a,b]

[a] *Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium*
[b] *iMinds Medical Information Technology, Leuven, Belgium*
[c] *Section of Clinical Neurophysiology, Department of Neurology, Erasmus MC, University Medical Center Rotterdam, The Netherlands*
[d] *Division of Neurology, Department of Medicine, McMaster University, Hamilton, Canada*
[e] *Department of Development and Regeneration, University Hospitals Leuven, Neonatal Intensive Care Unit, KU Leuven, Leuven, Belgium*
[f] *Department of Development and Regeneration, University Hospitals Leuven, Child Neurology, KU Leuven, Leuven, Belgium*
[g] *Section of Neonatology, Department of Pediatrics, Sophia Children's Hospital, Erasmus MC, University Medical Center Rotterdam, The Netherlands*
[h] *ZNA Koningin Paola Kinderziekenhuis, Antwerp, Belgium*
[i] *Institute of Biomedical Engineering, Department of Engineering, University of Oxford, Oxford, UK*

A R T I C L E   I N F O

Article history:
Accepted 6 June 2016
Available online 27 June 2016

Keywords:
Automated neonatal seizure detection
Hypoxic-ischemic encephalopathy
Machine learning
Support vector machines

H I G H L I G H T S

- An improved neonatal seizure detection method is discussed.
- A set of characteristic features of seizures are identified by data-driven methods.
- Described core characteristics of neonatal seizures can easily be used for other automated methods.

A B S T R A C T

Objective: After identifying the most seizure-relevant characteristics by a previously developed heuristic classifier, a data-driven post-processor using a novel set of features is applied to improve the performance.
Methods: The main characteristics of the outputs of the heuristic algorithm are extracted by five sets of features including synchronization, evolution, retention, segment, and signal features. Then, a support vector machine and a decision making layer remove the falsely detected segments.
Results: Four datasets including 71 neonates (1023 h, 3493 seizures) recorded in two different university hospitals, are used to train and test the algorithm without removing the dubious seizures. The heuristic method resulted in a false alarm rate of 3.81 per hour and good detection rate of 88% on the entire test databases. The post-processor, effectively reduces the false alarm rate by 34% while the good detection rate decreases by 2%.
Conclusion: This post-processing technique improves the performance of the heuristic algorithm. The structure of this post-processor is generic, improves our understanding of the core visually determined EEG features of neonatal seizures and is applicable for other neonatal seizure detectors.
Significance: The post-processor significantly decreases the false alarm rate at the expense of a small reduction of the good detection rate.

1. Introduction

Seizures are a common and distinctive sign of serious brain dysfunction in neonates (Volpe, 2008). The majority of neonatal seizures have an acute symptomatic basis and one of the most important causes is hypoxic ischemic encephalopathy (HIE) (Hahn and Olson, 2004; Cherian et al., 2011). Clinical presentation can be highly variable and manifestations of neonatal seizures can

* Corresponding author at: Department of Electrical Engineering (ESAT), KU Leuven, P.O. Box 2446, 3001 Leuven, Belgium.

E-mail addresses: amirhossein.ansari@kuleuven.be (A.H. Ansari), perumpij@mcmaster.ca (P.J. Cherian), anneleen.dereymaeker@uzleuven.be (A. Dereymaeker), Vmatic@singidunum.ac.rs (V. Matic), katrien.jansen@uzleuven.be (K. Jansen), a.dewispelaere@erasmusmc.nl (L. De Wispelaere), charlotte.dielman@zna.be (C. Dielman), jan.vervisch@uzleuven.be (J. Vervisch), r.swarte@erasmusmc.nl (R.M. Swarte), govaert@icloud.com (P. Govaert), gunnar.naulaers@uzleuven.be (G. Naulaers), maarten.devos@eng.ox.ac.uk (M. De Vos), Sabine.VanHuffel@esat.kuleuven.be (S. Van Huffel).

http://dx.doi.org/10.1016/j.clinph.2016.06.018
1388-2457/

be subtle, absent or resemble normal behavior. It is known that after treating with anticonvulsants, clinical seizures will change in subclinical seizures (Connell et al., 1989; Scher et al., 2003). Hence, clinical observation alone is ill-suited for their identification and monitoring (Bye and Flanagan, 1995; Rennie et al., 2004; Murray et al., 2008). Monitoring of the electroencephalogram (EEG) along with video is the gold standard for diagnosing and monitoring neonatal seizures (Rennie et al., 2004). However, most clinicians in NICUs opt to use amplitude integrated EEG [aEEG or cerebral function monitoring (CFM™)] instead, because of the ease of use and minimal need for support from EEG technology and clinical neurophysiology services (Gotman, 1990; Rennie et al., 2004). Since single channel aEEG often misses short, low-amplitude, or focal seizures (Eaton et al., 1994; Rennie et al., 2004), reliable automated neonatal seizure detection using continuous multi-channel EEG monitoring using 13–21 scalp electrodes has the potential to help clinical decision making in the NICUs and alleviate significantly the workload of the EEG interpreters.

In the literature, a few heuristic and model-based algorithms have been proposed to detect neonatal seizures. They typically consist of if-then rules and thresholds and are called heuristic because they are derived from mimicking the experts and their strategies used to detect neonatal seizures. Autocorrelation techniques (Liu et al., 1992), rhythmic discharge detection (Gotman et al., 1997), model-based EEG parameterization (Roessgen et al., 1998), modeling and complexity analysis (Celka and Colditz, 2002), wave-sequence analysis (Navakatikyan et al., 2006), pseudo-periodicity analysis (Stevenson et al., 2012) and atomic decomposition (Nagaraj et al., 2014) are some of the best known methods. Furthermore, an automated neonatal seizure detector mimicking a neonatal seizure expert was proposed in our group by Deburchgraeve et al. (2008) and was refined in Deburchgraeve (2010). In addition, artifact removal using different blind source separation techniques has been added to the detector to improve the performance (De Vos et al., 2011). Additionally, the performance of this method has been validated on an extensive dataset of asphyxiated neonates in the NICU of the Erasmus University Medical Center (EMC) Rotterdam (Cherian et al., 2011). The total good detection rate and positive predictive value (PPV) of this method, primarily reported to be 62% and 74% respectively, improved to 84% and 90% after removing four specific patients and some dubious seizures (Cherian et al., 2011). In this paper, this method is referred to as "heuristic" algorithm.

On the other hand, machine learning approaches have also been applied to train data-driven classifiers for this problem. The following methods have been considered: time–frequency based analysis and multi-layer perceptrons (MLPs) (Hassanpour et al., 2004), quantitative features and a linear discriminant classifier (Greene et al., 2008), support vector machine (SVM) based classifier (Temko et al., 2009), Bayesian classifier via Gaussian mixture models (Temko et al., 2009), adaptive multi-channel information fusion (Li and Jeremic, 2011), SVM classifier and Kalman filter (Bogaarts et al., 2014), and trend template analysis with SVM classifier tested on fetal lambs (Zwanenburg et al., 2015).

In addition, multi-stage classification composed of heuristic rules supplemented by a data-driven classifier was applied in Aarabi et al. (2007) and Mitra et al. (2009). In the former, a heuristic algorithm is used for artifact removal and EEG segmentation. Afterwards, the features are extracted from the segments and MLPs are applied as a classifier to identify the seizures. In the latter, conversely, MLPs 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 this study, we describe a method for improving a previously developed automated multi-channel EEG-based neonatal seizure detector, a so-called multi-stage classifier, as explained in Ansari et al. (2015). In the first stage, the heuristic algorithm mimicking an expert EEG reader detects the seizures. Then, in the second stage, a data-driven post-processor identifies the main characteristics of the detected segments such as evolution of spikes, synchronization of EEG and polygraphic signals, and other time–frequency domain features, in order to remove the falsely detected segments. An extensive test on three independent datasets exhibits the improved false alarm rate (FAR) in comparison to the original heuristic algorithm and its extensions.

2. Data description and methods

The used database composed of EEG-polygraphy recordings from 71 neonates acquired in the NICUs of Sophia Children's Hospital (part of the Erasmus University Medical Center Rotterdam, The Netherlands) (EMCR) and the NICU of the University Hospital of Leuven, Belgium (UZL). The polygraphic signals include electrocardiogram (ECG), electro-oculogram (EOG), chin or limb surface electromyogram (EMG), and abdominal respiratory movement signal (Resp.). All the neonates monitored in the EMCR ($n = 48$) had postasphyxial HIE whereas the included neonates in the UZL ($n = 23$) had different etiologies: HIE ($n = 6$), metabolic ($n = 5$), stroke ($n = 5$), genetic ($n = 2$), and others ($n = 5$). During the study period, all term neonates admitted to the NICUs with presumed postasphyxial HIE or with a high clinical suspicion of seizures underwent continuous EEG (cEEG) along with video for 24–48 h and magnetic resonance imaging (MRI). Inclusion criteria for asphyxia were either a five minute Apgar score below six or an umbilical artery pH <7.10 and clinical encephalopathy according to Sarnat score. When seizures were detected (either electro-clinical or electrographic) treatment with anti-epileptic drugs (AEDs) was initiated by protocol (Cherian et al., 2011). Newborns with heart malformation were excluded. All recordings were fully anonymized in their centers. The Erasmus MC medical ethics committee approved a study (2003–2007) to assess the utility of continuous EEG monitoring in neonates with postasphyxial hypoxic ischemic encephalopathy. Use of anonymized EEG data from this study, for analysis and research was subsequently approved. Furthermore, the study was approved by the medical ethics committee of UZ Leuven.

For this retrospective study, there was no preselection of data and no EEG recordings were excluded due to low-quality of EEG recordings, artifact contaminations, or expression of dubious seizures. Definite seizures were defined as paroxysmal EEG patterns with a change from ongoing background activity with repetitive spike-trains, oscillations or a mixture there of, with clear-cut onset and offset, lasting for at least 10 s. The dubious seizures are paroxysmal EEG events lasting for at least 10 s, composed of arrhythmic mixed oscillations or sharp waves of low amplitude (<30 μV) with irregular variation in amplitude, frequency and morphology (without well-defined evolution) (Cherian et al., 2011). Fig. 1 illustrates an arrhythmic dubious seizure with a low frequency and amplitude, with ill-defined onset and offset. In practice, the clinicians in the NICUs do not initiate treatment with anti-epileptic drugs (AEDs) when a dubious seizure is detected, unless this pattern frequently repeats itself or is accompanied by definite seizure patterns.

The database is partitioned into four datasets (DB1–DB4) according to their centers and durations of the scored EEG recordings. Some general characteristics of these datasets are mentioned in Table 1. DB1–DB3 were scored by a different rater, compared to DB4. DB1 has previously been used (about 8 h for each patient) for developing the heuristic algorithm (Deburchgraeve et al., 2008). This dataset was re-used here to develop and train the proposed data-driven post-processor (using the whole EEG recordings). The rest of the datasets have not been involved in the training phase in any way. In DB3 and DB4, only 2 h of each recording, which had at least one seizure observed by the rater, have been selected.