Robustness of Time Frequency Distribution based Features for Automated Neonatal EEG Seizure Detection

S. B. Nagaraj[1], N. J. Stevenson[2], W. P. Marnane[1], G. B. Boylan[2] and G. Lightbody[1]

Abstract— **In this paper we examined the robustness of a feature-set based on time-frequency distributions (TFDs) for neonatal EEG seizure detection. This feature-set was originally proposed in literature for neonatal seizure detection using a support vector machine (SVM). We tested the performance of this feature-set with a smoothed Wigner-Ville distribution and modified B distribution as the underlying TFDs. The seizure detection system using time-frequency signal and image processing features from the TFD of the EEG signal using modified B distribution was able to achieve a median receiver operator characteristic area of 0.96 (IQR 0.91–0.98) tested on a large clinical dataset of 826 h of EEG data from 18 full-term newborns with 1389 seizures. The mean AUC was 0.93.**

I. INTRODUCTION

Neonatal seizures occur at the rate of 1-3/1000 births and they are most commonly caused by conditions such as hypoxic ischemic encephalopathy (HIE), stroke and infection [1]. In the neonatal intensive care unit (NICU) the electroencephalogram (EEG) is a useful tool for seizure detection [2], [3]. The visual interpretation of the EEG is a time-consuming task and the availability of experienced annotators is limited. These limitations have driven the development of automated methods of neonatal EEG seizure detection.

The EEG signal can be classified into two states: non-seizure which appears to be a noisy random signal and seizure states that appears like repeated sharp waves or periodic rhythmic spike [2]. Several methods have been proposed for the detection of neonatal seizures; these methods, however, do not achieve a level of performance similar to the inter-observer agreement seen between human experts [4].

The non-stationarity nature of the neonatal EEG has led to the application of non stationary signal processing and segmented analysis for seizure detection problem [5]. An interesting approach was outlined by Boashash et al. in which features obtained from the time-frequency distribution (TFD) were used to detect neonatal seizures [6]. TFDs inherently deal with nonstationarity by projecting a signal onto a joint-time frequency domain. This method was tested on a small dataset of artefact free neonatal EEG (50 segments of seizure and non-seizure with individual segments of length 12.8s) with no apparent validation of results.

A significant problem with the development of a neonatal seizure detection algorithm (NSDA) is the lack of commonly available datasets to allow for the comparison of proposed methods. This problem, in conjunction with, a lack of data on inter-observer variability between human experts hinders the progress of NSDA development.

In this paper we apply the NSDA proposed by Boashash et al. (TFD feature-set classified by a support vector machine (SVM)) to a large dataset of neonatal EEG that better represents the application of a NSDA in the NICU (a NSDA as a surrogate neurophysiologist for long duration assessment of the neonatal EEG). The usefulness of such a process is twofold in that it provides information on the robustness of this method when applied to a real world setting and permits the comparison between this method and other methods developed on the same dataset ([7], [8], [9]).

II. METHODOLOGY

A. Dataset

A dataset of EEG recordings from 18 newborns obtained from the NICU, Cork University Maternity Hospital (CUMH), Cork, Ireland was used in this work. The combined duration of recordings was 826h and contains 1389 seizures with a mean duration of 194s (median 249s, IQR 96-356s). The EEG was recorded using a Viasys NicOne EEG system, with a sampling frequency of 256 Hz. Eight EEG channels in bipolar montage were used to annotate the data: F4-C4, C4-O2, F3-C3, C3-O1, T4-C4, C4-Cz, Cz-C3, and C3-T3. The EEG recordings were not altered to remove various artifacts which are common in the real-world NICU recordings. With the assistance of video EEG, seizures were annotated by 2 experienced neonatal electroencephalographers. The EEG was down-sampled from 256 to 32 Hz with an anti-aliasing filter set at 12.8 Hz. A single pole highpass filter with a cutoff of 0.5Hz was also used to filter the data and then segmented into 8s epochs with a 4s overlap. This study had full ethical approval from the CUMH and the University College Cork. Informed parental consent was obtained to record EEGs and it was anonymized at the time of recording.

B. Time Frequency (TF) distribution

Non-stationary signals can be analyzed and classified using a TFD based on the Wigner-Ville distribution (WVD). The WVD is contaminated by cross-terms generated by the bilinear nature of the signal kernel. These cross-terms can be

This material is based upon works supported by Science Foundation Ireland under grant No. SFI/10/IN.1/B3036 and Science Foundation Ireland Research Centres award No. 12/RC/2272.

[1]S. B. Nagaraj, W. P. Marnane and G. Lightbody, are with the SFI INFANT research centre and the Department of Electrical Engineering, University College Cork, Ireland. sunil@eleceng.ucc.ie, l.marnane@ucc.ie, g.lightbody@ucc.ie

[2]N. J. Stevenson and G. B. Boylan are with the SFI INFANT research centre and the Department of Pediatrics and Child Health, University College Cork, Ireland. n.stevenson@ucc.ie

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

minimized with an appropriate smoothing window (a 2D low pass filter) [11]. The quadratic time-frequency distribution (QTFD) of a given analytic signal $z[n]$ obtained from the real discrete time signal $x[n]$, $n = 0, 1, \ldots, N-1$, is defined as:

$$\psi[n,k] = 2\gamma[n,k] \underset{n,k}{**} \sum_{m=1}^{N} z[n+m]z^*[n-m]e^{\frac{-j2\pi km}{N}} \quad (1)$$

where $\gamma[n,k]$ is the kernel of the TFD and $\psi[n,k]$ is an $N \times N$ matrix. We use the modified B distribution (MBD) and smoothed WVD (SPWVD) in this work as it has been shown that these TFDs provide superior performance for adult and neonatal seizure detection.

For the SPWVD $\gamma[n,k]$ is a 2D Hanning window of duration (1.2 s) and bandwidth (0.13 Hz). For the modified B distribution $\gamma[n,k] = IFT_{l\rightarrow n} n \cosh[l]^{-2b}, b = 0.01$. Figure 1 shows a sample 8s seizure and non-seizure epoch and their corresponding TFDs using SPWVD.

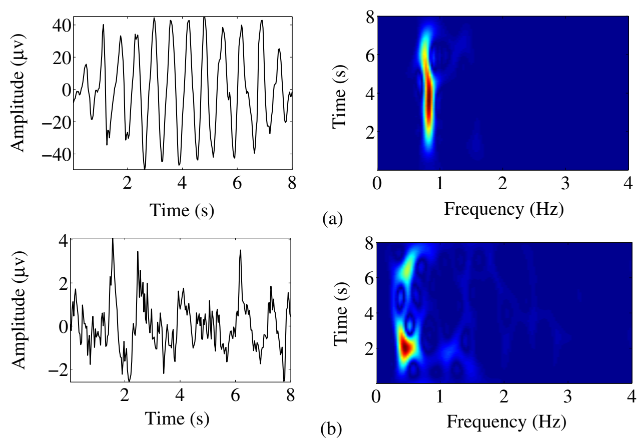

Fig. 1. Example of an 8s seizure and non-seizure epoch used in this paper and their corresponding TFDs using Hanning window are shown in (a) and (b) respectively.

C. TF Feature Extraction

The TFD features proposed by Boashash et al. were directly used for our work [6]. These TFD features were initially used by Boashash et al. for multichannel-based newborn seizure detection.

1) Instantaneous frequency (IF): The instantaneous frequency (IF) of $x[n]$ with a sampling frequency f_s can be obtained from the peak of the TFD of $x[n]$ as

$$f_i[n] = \frac{f_s}{2M} \arg\max_k \{\psi[n,k]\}. \quad (2)$$

From this we get features 1 and 2 as,

$$F_1 = \frac{1}{N} \sum_{n=1}^{N} f_i[n], \quad (3)$$

and

$$F_2 = \Delta f_i[n] = \max(f_i[n]) - \min(f_i[n]) \quad (4)$$

respectively.

2) Singular value decomposition (SVD) based features: Several features were generated from the SVD of the TFD. SVD is based on the definition of a TFD as,

$$\psi = \mathbf{USV}^H \quad (5)$$

where, in this case, **U**, **S** and **V** are $N \times N$ matrices, and the diagonal values of **S** contain the singular values of the decomposition. The maximum (F_3) and the variance (F_4) of the diagonal entries (singular values) of **S** are taken as features 3 and 4, respectively. A TF complexity measure is also estimated from **S** and is defined as,

$$F_5 = -\sum_{i=1}^{N} \bar{S}_i \log \bar{S}_i. \quad (6)$$

3) Energy concentration measure given by

$$F_6 = \left(\sum_{n=1}^{N} \sum_{k=1}^{M} |\psi[n,k]|^{\frac{1}{2}} \right)^2. \quad (7)$$

4) Sub-band energy features: These two features capture the sub-band energies of $x[n]$ corresponding to $\delta = 0 - 4Hz$ and $\theta = 4 - 8Hz$ respectively which are defined as:

$$F_7 = \sum_{n=1}^{N} \sum_{k=1}^{M_\delta} \psi[n,k] \quad (8)$$

$$F_8 = \sum_{n=1}^{N} \sum_{k=M_\delta}^{2M_\delta} \psi[n,k] \quad (9)$$

where $M_\delta = \lfloor 8M/f_s \rfloor$ and $\lfloor . \rfloor$ is the floor operator.

In addition to the above mentioned features, Boashash et al. also proposed a method to extract several image processing features based on visual descriptors from the TFD viewed as an image [6]. The TF image (TFI) is based on a binary interpretation of the TFD which is obtained by applying threshold using the Otsu's method on gray image such that [12]:

$$\begin{aligned} \psi_b[n,k] &= 1, \text{ for } \psi[n,k] > \text{threshold, and} \quad (10) \\ &= 0, \text{ for } \psi[n,k] \leq \text{threshold.} \end{aligned}$$

Figure 3 shows an example of TFDs (using SPWVD) for 8s seizure and non-seizure epochs and their corresponding binary segmented images. The moment of order (p,q) and the central moments of $\psi_b[n,k]$ can be expressed as [6]:

$$m_{pq} = \sum_n \sum_k n^p k^q \psi_b[n,k], \text{ and} \quad (11)$$

$$\eta_{pq} = \sum_n \sum_k (n-\bar{n})^p (k-\bar{k})^q \psi_b[n,k] \quad (12)$$

where $p, q = 0, 1, 2, \ldots$. Five features were estimated from $\psi_b[n,k]$: $F_9 = \eta_{00}$, $F_{10} = (m_{30} + m_{12})^2 + (m_{03} + m_{21})^2$, $F_{11} = F_9^2/F_{10}$, $F_{12} = \frac{m_{10}}{m_{00}}$, and $F_{13} = \frac{m_{01}}{m_{00}}$. Some of these features were obtained directly by using MATLAB function *regionprops*.

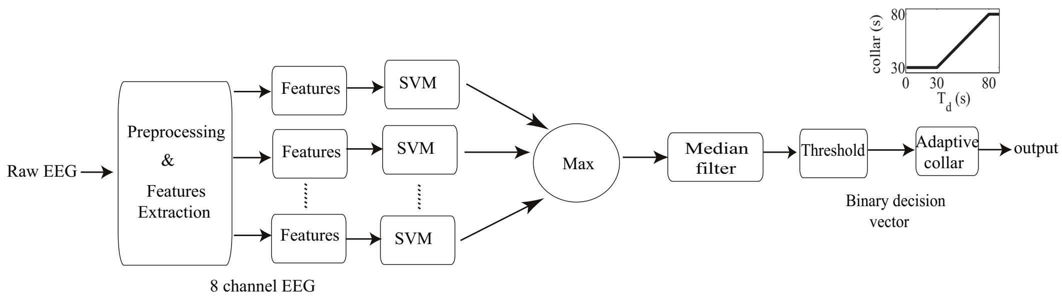

Fig. 2. Architecture of neonatal seizure detection system. The input signal is downsampled to 32Hz with an anti aliasing filter set at 12.8Hz, and then segmented into 8s epochs with 4s overlap. The TF features extracted from the TFD of each epoch were then fed to a support vector machine (SVM) classifier with a radial basis function kernel. The outputs of the SVM were then converted to probability-like values. The maximum of the median probabilities across all channels was computed to represent the final support of a seizure. The sharp transients in the SVM output are then suppressed using a median filter of 12s in duration. An adaptive collar (ac) that is related to the duration of the detected seizure is then applied to the binary output to extend the detection such that $collar = 30$s for $T_d < 30$, $collar = T_d$ for $30 \leq T_d \leq 80$s, and $collar = 80$s for $T_d > 80$s where T_d is the duration of the detected seizure (in seconds). An automated seizure annotation is then obtained.

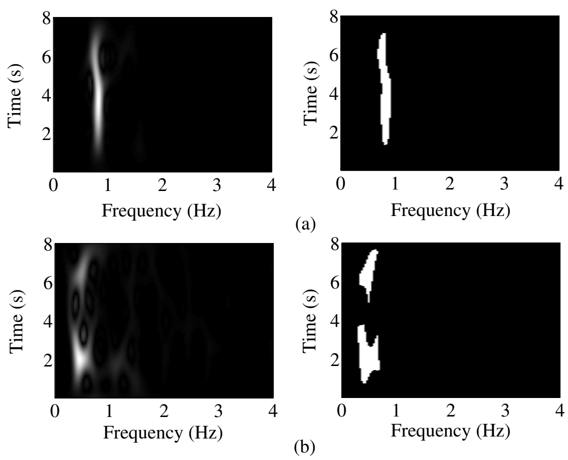

Fig. 3. Example TFDs (using SPWVD) of 8s seizure and non-seizure epochs and their corresponding binary segmented images using Otsu's method are shown in (a) and (b) respectively.

D. Automated neonatal seizure detection system architecture

Figure 2 shows the automatic neonatal seizure detection system. The features are fed into an SVM classifier with radial basis function (rbf) kernel. The maximum of the probabilities across all channels were computed to represent the final support of a seizure. Sharp transients in the support are suppressed using a median filter of 12s in duration.An initial decision is generated using a simple threshold of the SVM output. The adaptive collar technique was then applied in which every seizure decision was extended proportional to the duration of the detected seizure on either side. An automated annotation of the seizure is then obtained.

E. Training and Testing

The performance of the proposed algorithm was estimated using a leave-one-out (LOO) cross validation as it provides almost an unbiased estimation of the true generalization error. In this validation method 17 patient's data were used in training and the left-out patient's data was used for testing. This process was repeated until data from each patient was used for testing (18 different combinations of test/train sets). The mean and median value across all 18 test folds were then obtained. To select suitable model parameters for the SVM, nested cross-validation model selection on the training data was performed. Probability-like values were then obtained from the SVM. A subset of 5 minutes of seizure and 50 minutes of non-seizure were selected from 17 neonates at each step (a total of 85 minutes of seizure and 850 minutes of non-seizure data at each training iteration). The features extracted during training were then fed to an SVM classifier with rbf kernel and then tested on the full recording of the remaining neonate to generate an automated annotation of seizure. All features were converted to z-scores before classification. Time (sensitivity and specificity/AUC) and event (seizure detection rate (SDR) and false detections per hr (FD/h)) based metrics were used to assess the level of agreement between the human and automated annotation of seizure. More details about these metrics can be found in Temko et al. [7].

III. Results and Discussion

The performance of each TF feature for the detection of EEG seizures is shown in Fig. 4. The F_3 and F_9 features provide the highest discrimination between seizure and non-seizure. The results obtained using proposed NSDA for individual recordings is shown in Fig. 5. The performance of the NSDA was 0.89 and 0.91 with TF features based on the SPWVD and MDB, respectively. This performance was improved by using both signal and imaging features (AUC increased from 0.91 to 0.93 using the MBD based feature-set). The TF features set based on the MBD provided slightly superior performance than the SPWVD. This was also noted in Boashash et al. (SPWVD=0.93 and MBD=0.96) [6]. Table I shows time and event based metrics at select thresholds for the TF feature-sets generated with the SPWVD and MBD. These values are significantly lower than outlined in Boashash et al. (0.96 vs 0.91) [6]. This was expected as the large unedited dataset used here includes a large proportion of confounding EEG, such as artefact, that may not be available in smaller selected datasets.

The performance of a NSDA based on this TF feature-set is less than the methods of Temko et al. and Thomas et al. which are reported on the same datasets (AUCs of 0.96, 0.93) [7], [8]. A more useful comparison at a clinically relevant threshold of 0.1 FD/h shows approximately a 50% reduction in seizure detection rate (52.1% vs 26.5%) as can be seen from Fig. 6 and table I. It must be noted that the TF feature-set is significantly smaller than the feature-set used in these methods (13 vs 55 features) and may not be able to represent the variety of seizure and non-seizure EEG seen in our larger dataset.

Nevertheless, it was surprising that the TF feature-set which was developed on such a small dataset of neonatal EEG has relatively good performance. This suggests the TF feature-set is robust and its incorporation in the larger feature-set used in Temko et al. [7] and Thomas et al. [8] may further improve NSDA performance.

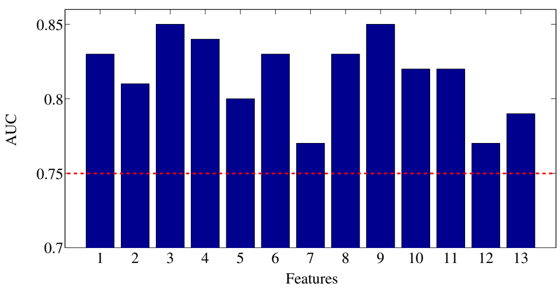

Fig. 4. Performance of individual features using MBD and simple thresholding method.

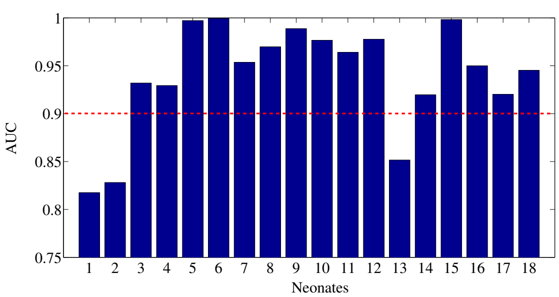

Fig. 5. Performance of the proposed NSDA for each neonate using MBD. We can clearly see that the AUCs for all neonates are above 0.90 (marked as red dotted line) except for neonates 1,2 and 13.

IV. CONCLUSIONS

The robustness of a NSDA based on a recently proposed TF feature-set classified by an SVM was investigated. While the NSDA showed reduced performance on a more realistic set of neonatal EEG, the TF features provided a high level of detection performance.

REFERENCES

[1] J. J. Volpe, *Neurology of the Newborn*, 5th ed., Saunders Co., 2008.

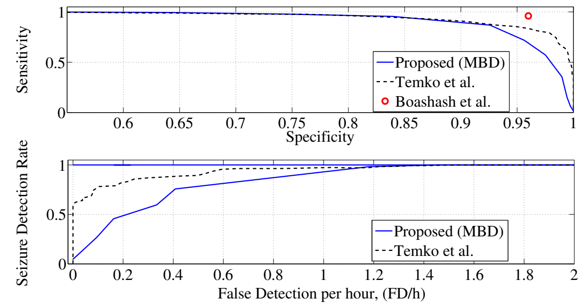

Fig. 6. Comparison of the performance of the proposed NSDA (median values) with several methods currently disclosed in the literature using MBD. The results reported by Boashash et al. [6] on their testing data are included for comparison.

TABLE I

TESTING RESULTS (MEAN, MEDIAN, IQR)% USING THE PROPOSED NSDA. THE PERFORMANCE USING MBD OUTPERFORMS SPWVD.

	FD/h	SDR	Sensitivity	Specificity
	0.01	6.8, 7.2	4.2, 5.1	99.8, 100
		(6.1–8.8)	(3.5–5.8)	(99.6–100)
	0.1	26.5, 28.3	13.8, 16.6	98.1, 99.3
MBD		(22.6–30.4)	(10.8–20.3)	(97.2–99.8)
	1.0	90.8, 92.5	81.7, 83.6	91.4, 94.8
		(86.6–94.3)	(74.6–89.3)	(86.5–97.2)
	0.01	6.1, 6.4	3.6, 4.0	99.6, 99.8
		(5.8–8.6)	(3.2–5.4)	(99.5–100)
	0.1	23.2, 25.1	12.6, 14.2	97.3, 98.5
SPWVD		(21.4–28.7)	(11.5–18.7)	(96.8–99.6)
	1.0	89.5, 91.4	78.3, 80.2	90.6, 93.4
		(85.3–93.2)	(73.5–88.2)	(85.8–96.7)

[2] E. Niedermeyer and F. H. Lopes da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and related Fields*, 5th ed., Lippincott Williams and Wilkins, 2004.

[3] G. B. Boylan, N. J. Stevenson, and S. Vanhatalo, “Monitoring neonatal seizures”, *Semin. Fetal Neonatal Med.*, vol. 18, pp. 202–208, 2013.

[4] S. B. Wilson, M. L. Scheuer, C. Plummer, B. Young and S. Pacia, “Seizure detection: correlation of human experts”, *Clin. Neurophysiol.*, pp. 2156–2164, vol. 114, 2003.

[5] P. Celka, and P. Colditz, “A computer-aided detection of EEG seizures in infants: a singular spectrum approach and performance comparison”, *IEEE Trans. on Biomed. Eng.*, vol. 49, pp. 455–462, 2002.

[6] B. Boashash, L. Boubchir, and G. Azemi, “A methodology for time-frequency image processing applied to the classification of non-stationary multichannel signals using instantaneous frequency descriptors with application to newborn EEG signals”, *EURASIP J. Adv. Signal Process*, no. 117, 2012.

[7] A. Temko, E. Thomas, G. Boylan, W. Marnane, and G. Lightbody, “EEG-based neonatal seizure detection with support vector machines”, *Clin. Neurophysiol.*, vol. 122, no. 3, pp. 464–473, 2011.

[8] E. M. Thomas, A. Temko, G. Lightbody, W. P. Marnane and G. B. Boylan, “Gaussian mixture models for classification of neonatal seizures using EEG”, *Physiol. Meas.*, vol. 31, pp. 1047–1064, 2010.

[9] N. J. Stevenson, J. M. OToole, L. J. Rankine, G. B. Boylan and B. Boashash, “A nonparametric feature for neonatal EEG seizure detection based on a representation of pseudo-periodicity”, *Med. Eng. and Phys.*, vol. 34, no. 4, pp. 437–446, 2012.

[10] M. Scher, M. Sun, D. Steppe, R. Guthrie, and R. Sclabassi, “Comparison of EEG spectral and correlation measures between healthy term and preterm infants”, *Pediatr. Neurol.*, vol. 10, no. 2, pp. 104–108, 1994.

[11] B. Boashash, *Time Frequency Signal Analysis and Processing–A Comprehensive Reference*, Elsevier, 2003.

[12] N. Otsu, “A threshold selection method from gray-level histograms”, *IEEE Trans. on Systems, Man and Cybernetics*, vol. 9, pp. 62–66, 1979.