

# Neonatal Seizures Detection using Stationary Wavelet Transform and Deep Neural Networks: Preliminary Results

Lorenzo Frassinetti  
Department of Information Engineering  
University of Florence  
Department of Medical Biotechnologies  
University of Siena  
Florence, Italy  
lorenzo.frassinetti@student.unisi.it

Daniele Ermini  
Department of Information Engineering  
University of Florence  
Florence, Italy  
daniele.ermi@stud.unifi.it

Rachele Fabbri  
Department of Information Engineering  
University of Florence  
Florence, Italy  
rachele.fabbri@stud.unifi.it

Claudia Manfredi  
Department of Information Engineering  
University of Florence  
Florence, Italy  
claudia.mandredi@unifi.it

**Abstract**— The increasing use of Electroencephalography (EEG) in the field of pediatric neurology allows more accurate and precise diagnosis of several cerebral pathologies, mainly in Neonatal Intensive Care Units (NICUs), where it represents the gold-standard for the diagnosis of neonatal epileptic seizures. However, EEG interpretation is time consuming and requires highly specialized staff. For this reason, in the last years there was a growing interest in the development of systems for automatic and fast detection of neonatal epileptic seizures. To this aim, we propose here hybrid systems that combines techniques related to the Stationary Wavelet Transform (SWT) as a support to deep-learning algorithms such as Convolutional Neural Networks and Fully Convolutional Networks. The proposed methods are validated on a public dataset of NICUs seizures recorded at the Helsinki University Hospital. Results are encouraging both in terms of Area Under the receiver-operating Curve, AUC (81%), Good Detection Rate, GDR (77%) and False Detection per hour, FD/h (1.6). Actually, the SWT step increases the performance of the proposed methods of about 5% for the AUC as compared to considering the raw EEG time-series only. These results, though preliminary, represent a significant step forward in solving the problem of neonatal seizure detection.

**Keywords**—Epilepsy, SWT, NICU, Seizure Detection, Neural Networks

## I. INTRODUCTION

One of the most challenging issues in Neonatal Intensive Care Units (NICUs) concerns the early neonatal seizures detection (NSD). Epilepsy, especially in the preterm newborn, represents one of the most common brain syndromes [1]. Its aetiology may be due to various causes and several comorbidities [2]. Its early detection, especially in the first 48 hours after birth, could be essential for a timely and effective administration of the appropriate therapy and thus a possible better prognosis. Moreover, it could help in controlling the adverse effects of the syndrome on brain health and patient's neurodevelopment. Currently, for a correct diagnosis, the identification of epilepsy in newborn is tricky and time-consuming, requiring expert and high qualified staff [3]. This is mainly due to the high incidence in this population of the so-called subtle seizures [4]. These seizures exhibit a limited or even absent motor activity, thus without the classical clinical signs as in the tonic-clonic seizures of the adult [5]. Therefore, in clinical practice the gold-standard for the NSD

is represented by the electroencephalogram (EEG), both amplitude-EEG (aEEG) and continuous (cEEG). The aEEG is useful for neonatologists as it provides a first assessment about patient's brain health. The cEEG allows a complete and exhaustive understanding of the pathology, also in those cases that cannot be detected with aEEG [6]. Unfortunately, the analysis of cEEG signals is time-consuming, especially for long-monitoring recordings and even for expert staff. For these reasons, recently in the literature there was a growing interest in investigating methods that could allow an automatic detection of neonatal seizures [7]. For newborns this is in fact still an open problem, mainly due to the high variability of the electrophysiological pathological patterns, also related to different gestational ages [8].

Recently, Temko et al. [9] and Tapani et al. [10] developed machine-learning algorithms that exhibit high performance (Area Under the receiver-operating Curve AUC 0.97 - 0.98). They represent the state-of-art for the development of automatic NSD methods based on machine-learning techniques. The main drawback of these approaches concerns complex hand-crafted features that may not be optimal in some scenarios [11]. To get over such limitations, deep learning methods for EEG applications were widely investigated in the last years, and are becoming increasingly popular [12]. Ansari et al. [13] evaluated deep neural networks for multi-channel EEG systems to solve the NSD problem on a private dataset. In addition, they developed a hybrid algorithm where deep learning techniques are used as a support to features extraction and then as input to several classifiers (Support Vector Machine, Random Forest etc.). These techniques enhance the time characteristics inside the EEG signals. Finally, for the NSD Temko and O'Shea evaluated deep-learning algorithms for a single-channel EEG system [14]. Moreover in [15] they proposed a first multi-channel EEG analysis with a Fully Convolutional Network (FCN) with high performance (AUC 98.5% and AUC90 88.2%) trained on a private dataset. These results could represent the state-of-art for solving the NSD problem with deep-learning algorithms.

In this paper, we propose hybrid techniques that allow better information balance both in space and time domain. To this aim, we propose a combination of Stationary Wavelet Transform (SWT) [16] applied to the raw EEG signal as the input to Convolutional Neural Networks (CNN) and FCN.

The SWT was added with the aim of eliminating unnecessary frequency information for the NSD problem [10] and performing a sort of data-augmentation [17] on the EEG signals, representing each single derivation through its levels of decomposition and approximation. The use of Time-Frequency representation for EEG signals in deep-learning applications was already successfully tested by Yuan et al. [18] for another seizure detection problem.

All the methods proposed here were trained and tested on a public dataset of neonatal EEG signals collected in NICU at the Helsinki University Hospital [19]. The sampling frequency is 256 Hz. In particular, the dataset consists of 79 multi-channel EEG recordings from at-term neonates independently evaluated by three experts: 39 had seizures and 22 were diagnosed seizure free with unanimous consent of the three experts. To date, this dataset represents one of the most complete public datasets of neonatal seizures. As stated by the authors in [19], it is also the only dataset with multi-expert clinical annotations. Therefore it can be used as a reference set of neonatal seizures, especially for validation and reproducibility assessment of algorithms related to the NSD. Indeed, one of the possible limitations of proposed algorithms could be their evaluation on private datasets only. Furthermore, we highlight the fact the multi-expert annotations may affect the quality and the impact of the information provided to the algorithms, especially in case of disagreement between the experts.

This paper is organized as follows: in Section II the proposed methods for the SWT to generate the EEG input matrix, the CNN and FCN architectures used are described. Results concerning the proposed techniques applied to the dataset in [19] are presented in Section III. Section IV is devoted to the discussion about the encouraging results obtained, and the possible impact of these hybrid techniques (SWT + CNN or FCN) to the solution of the NSD problem.

## II. METHODS AND MATERIAL

The proposed methods are implemented under MATLAB computing environment (Deep Learning Toolbox, version 2019b [20]), OS Windows 10, 64 bit. Processor: AMD Ryzen 5 2600 Six-Core, CPU 3.40 GHz, RAM 16 Gb, GPU NVIDIA GeForce GT 1030.

### A. Pre-processing and SWT-EEG generation

In this work the same bipolar configuration as in [15] were used: F4-C4, C4-O2, F3-C3, C3-O1, T4-C4, C4-Cz, Cz-C3, C3-T3.

One of the most critical points about the analysis of EEG signals is the choice of the appropriate time-window to evaluate the information inside it. This is due to the non-stationarity of the signals [21] and the presence of artifacts that affect the recordings. In the literature, different settings of the window length have been proposed. In [15] windows with duration of 8s were used, in analogy to the clinical definition of epileptic seizures. In [13] Ansari et al. argued that some neonatal seizures with a very low electrophysiological activity cannot be found using 8s windows, therefore they used a window length of 90s to detect them. Instead, in [18] Yuan et al. used 3s windows to compute the EEG spectrogram, applying the Short Time Fourier Transform. This short duration helps with taking into account the quasi-stationarity of the signal.

In this work, time windows of 4s, 8s and 16s duration and 50% overlap were considered. Each window was filtered with a pass-band filter (0.5-32Hz) and a Notch Filter (50Hz). The high-pass cutoff (0.5Hz) in the pass-band filter was selected following the considerations presented in [9, 10]. The low-pass cutoff (32Hz) was selected according to [10] and based on the subsequent downsampling steps of the method (described in the following paragraphs).

To each window (4s, 8s and 16s) the Stationary Wavelet Transform (SWT) with 6 levels of decomposition was applied. The mother wavelet function is the Daubechies 5 shown in Fig. 1 [22]. Fig. 2 shows an example of decomposition for a 8s window (only one EEG channel is displayed). Six levels of decomposition and one approximation level allow expanding one single derivation  $1 \times n$ , where  $n$  is the number of samples in the time series, into a  $7 \times n$  matrix where each row represents the corresponding level of decomposition and approximation (Fig.2).

According to [23], for each derivation the decomposition levels d6, d5, d4 and the approximation level a6 were taken into account. This choice was made as they approximately correspond to the conventional EEG rhythms for a sampling frequency of 256Hz [23]. Thus, the other decomposition levels were removed, obtaining a  $4 \times n$  matrix. As an example, for an 8s window with sampling frequency 256 Hz, a  $4 \times 2048$  matrix is obtained for each derivation. The SWT operation could thus be seen as a sort of data-augmentation or transformation of the input signals. The eight derivations were concatenated together, generating a  $32 \times 2048$  matrix. Then a down-sampling step was performed to have windows with 256 samples for each row. This  $32 \times 256$  matrix was used as

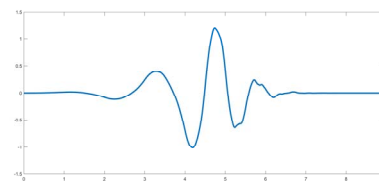


Fig. 1 . Db5 Profile.

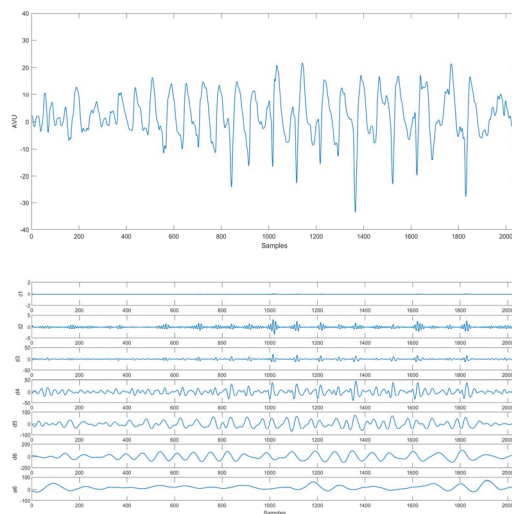


Fig. 2. Upper plot: example of single channel EEG showing a seizure  
Lower plot: its corresponding 6-levels SWT. Window size: 8s.

input for the proposed neural networks. The down-sampling step was added both to increase the balance between the information across the matrix dimensions (conforming to the remaining decomposition levels in the final matrix) and to obtain a matrix with standard lengths for each window considered.

To assess possible redundancy of the SWT operation, two different tests were performed. First, an EEG matrix 8x256 was built without using the SWT operation, that is keeping the original time series. In other words, after filtering only the concatenation and the down-sampling steps were replicated for the original time series. This test aims at evaluating whether for a neural network it is better to keep the original EEG signal or to use a modified version of it, and if the processed versions could be somewhat harmful to correctly detect seizures. For the second test an EEG matrix 32x256 was built, but in this case each derivation was replicated for each decomposition and approximation level considered. More precisely, each original time series was replicated 4 times and then concatenated with the others. This test is intended to mimic the same data-augmentation on the data performed by the SWT, establishing whether an increase in performance could be due to the concatenation step only and not to the application of the SWT.

In this paper, we refer to the first test as Test\_1 and to the second one as Test\_2.

### B. Convolutional Neural Networks and Fully Convolutional Networks

For all the case studied (windows 4s, 8s and 16s) a preliminary Convolutional Neural Network (CNN) architecture was built, whose parameters are summarized in Table I. Further details about the layers can be found in [20, 24]. Based on it, a Fully Convolutional Network (FCN) architecture was developed, where the densely connected layers in the Classification Blocks (Table I, Classification Blocks FCN) were replaced by convolutional and pooling layers. For Test\_1 the architecture described in Table I was adapted to allow an 8x256 matrix as input layer. The layer devoted to assess if an input matrix corresponds to a seizure or not is the Classification Layer: according to [19], a value equal to 1 means that a seizure is detected, while a value equal to 0 means no seizure. In addition it gives a score between 0 and 1 for each class. The parameters used in the training phase of network are the following: Solver stochastic gradient descent with momentum [15, 25] with learning rate 0.001 and momentum 0.9, Mini-Batch Size 128, Validation Patience 8 [20]. We also investigated other activation functions such as Rectified Linear Unit (ReLU) and LeakyReLU [26], but without satisfying results and lower performance than the hyperbolic tangent (Tanh) layer. Therefore they were not included in the Results section. The architectures and the parameters used in the training phase were obtained by a trial-and-error process, thus they might not represent the optimal configuration. In the data set the seizure set was much lower than the non-seizure one. Considering only the consensus seizures, using the multi-expert labels provided in [19], there were 39259 seizures labels and 363566 non-seizures labels, where a label represents the corresponding time instant (in seconds) of the EEG exam (at sampling frequency 256 Hz). Thus, for example considering the 8-seconds windows and the 50% overlap, we obtained 10239 seizures windows and 90342 non-seizures windows (with an imbalance ratio of 8.82).

TABLE I. ARCHITECTURES PROPERTIES

Feature Extraction Blocks (CNN and FCN)	
Input Layer 32x256 rescale symmetric [-1,1]	
Convolution 2d Layer, filter size (2,4), 32 filters, Stride (2,2)	
Batch Normalization Layer	
Max Pooling Layer, pool size (1,2), Stride (1,2)	
Tanh	
Convolution 2d Layer filter size (2,4), 64 filters, Stride (2,2)	
Batch Normalization Layer	
Max Pooling Layer, pool size (1,2), Stride (1,2)	
Tanh	
Convolution 2d Layer filter size (1,4), 128 filters, Stride (1,1)	
Batch Normalization Layer	
Max Pooling Layer, pool size (1,2), Stride (1,2)	
Tanh	
Classification Blocks	
CNN	FCN
Fully Connected Layer (1024x1024)	Convolution 2d Layer, filter size (1,3), 2 filters, Stride (1,2)
Dropout (0.5)	Average Pooling, pool size (1,2)
Fully Connected Layer (2x1024)	Max Pooling, pool size (8,1)
Softmax Layer	Softmax Layer
Classification Layer (Cross-Entropy)	Classification Layer (Cross-Entropy)

Therefore, a random down-sampling for each patient was applied to the non-seizures set to obtain approximately a ratio 1:1 between the two classes. In other words, the non-seizures windows in excess were excluded randomly a priori before each training phase of the networks. To avoid over-fitting, during the training phase an internal k-fold cross-validation was applied, using 5 patients randomly selected in the original training set. This operation was performed during the training phase only, to get a stop criterion according to the Validation Patience parameter previously mentioned. To correctly quantify the performance of the methods, a Leave-One-Out approach was used.

The reasons of this choice, crucial for the correct assessment of the performance for a neonatal seizure detector, are discussed in the following sub-section.

### C. Performance Evaluation

When a new automatic detector is proposed with the aim of solving the NSD problem, the most critical point is the choice of the correct validation procedure. As demonstrated in [27] and reaffirmed in [10], classical validation approaches such as k-fold cross-validation tend to overestimate the algorithm performances of a patient-independent system. Thus, we performed a Leave-One-Out validation for our methods. A single patient during the training process was iteratively excluded and used as test set. This procedure ensures the independence between the training and the test set and allows a true evaluation of the generalization and accuracy of the proposed methods.

Furthermore, as in the used dataset the annotations are provided by three experts, we have considered as ground truth the combination ‘and’ of them, i.e. when all the three experts agree. Thus, in our experiments epileptic patients are only those for which there is a full consensus among the three clinicians (39 patients). The last point taken into account concerns the metrics used to evaluate and compare the methods. As in [27] the following epoch-based metrics are applied: Accuracy (Acc); Sensitivity (Sen); Specificity (Spe), F1score (F1); Matthews Correlation Coefficients (MCC); Area Under the receiver-operating Curve (AUC), based on

False Positive Rate and True Positive Rate) and AUC concatenated ( $AUC_{cc}$ ), considering all the patients as a single recording [10, 15]. Moreover, according to [27, 28] we reported also the event-based metrics: Good Detection Rate (GDR) and False Detection per hour (FD/h) with moving-average filters of different length, greater than or equal to the used windows. This filter averages the scores obtained after the classification step made by the neural networks (Classification Blocks in Table I). If the scores were lower than a predefined threshold the outputs were considered as not-seizures, otherwise as seizures. The window parameters (length of the filters and thresholds used) were optimized in order to obtain the best event-based metrics for each window considered (i.e. at the same time the highest GDR and the lowest FD/h). To choose the best methods, we considered as primarily metrics the parameters AUC and  $AUC_{cc}$  [10, 15, 27]. However, all the metrics are considered and shown here, for a complete view about the performance of the tests.

### III. RESULTS

Table II shows the average epoch-based performances obtained for each window using the CNN architecture. In Table III, the epoch-based metrics obtained using the FCN architecture are reported. This comparison is useful to ascertain which one, among the two architectures, could be considered the best one, and how the performances change according to the selected time window. Next, only the test with the highest  $AUC_{cc}$ , that corresponds to 8s windows, is taken into account. To assess if the SWT operation could effectively increase the performance, Table IV reports the comparison between the best case, Test\_1 and Test\_2. The ROC curves related to the parameters  $AUC_{cc}$  are shown in Fig.3. Furthermore, in Table IV, epoch-based and event-based metrics are shown. For the GDR and FD/h metrics the parameter Latency-Time is reported. Latency-Time represents the time duration of the moving-average filters used for each test. The Latency-Time can be considered as an event-based metric that provides a measure of the system's time-response, necessary to obtain such performances. Therefore, high values of the Latency-Time are not suitable for online seizure detection, but only for offline application [29, 30]. Results were obtained after a complete Leave-One-Out validation process: each patient was iteratively excluded and the networks trained with the remaining ones, using the internal k-fold cross-validation as stopping criterion (see subsection Performance Evaluation). After training, the networks were used as seizure detectors for all the test patients' epochs.

Then, the performances were averaged across all patients. These procedures were repeated for all the windows (4s, 8s and 16s). As the EEG signals were splitted into several windows, to complete one single validation process the proposed approaches required different computational time depending on the window considered: the shorter the window the larger the training set.

TABLE II. RESULTS CNN

Test	CNN						
	AUC (%)	$AUC_{cc}$ (%)	Acc (%)	Sen (%)	Spe (%)	F1 (%)	MCC
4s	77	76	83	39	92	40	0.28
8s	77	79	79	51	81	41	0.27
16s	75	77	85	37	94	49	0.30

TABLE III. RESULTS FCN

Test	FCN						
	AUC (%)	$AUC_{cc}$ (%)	Acc (%)	Sen (%)	Spe (%)	F1 (%)	MCC
4s	79	83	84	47	89	42	0.29
8s	81	87	82	63	83	48	0.35
16s	79	84	80	57	82	45	0.31

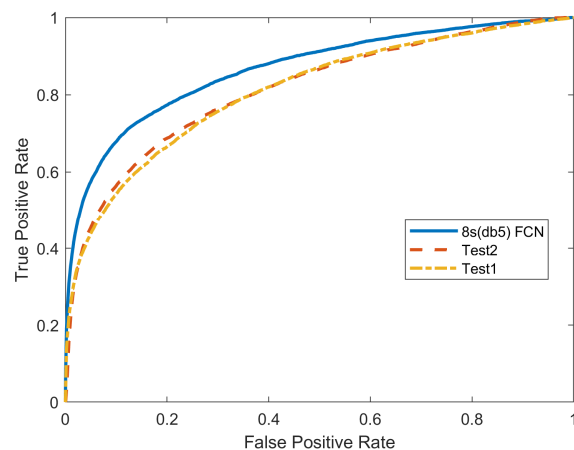


Fig. 3 . ROC curves of  $AUC_{cc}$  values reported in Table IV.

With the available hardware and software resources it took more than 20h for the 4s window, about 12h for the 8s window and less than 6h for the 16s window. Because of the random selection of the patients in the internal cross-validation, we investigated if this approach could significantly alter the performance. To this aim, we repeated all the described procedures three times with the five patients used, without finding significant differences in the performances (less than 1-2% in Accuracy). An exhaustive research about the optimal number of patients for the internal cross-validation was not performed, but it will be considered in future developments of the methods.

Anyway, we tried also with three patients and ten patients. In both cases the performances were lower than with five patients. Specifically, the 3-fold networks exhibited an overfitting trend while the 10-fold networks showed a worse fitting on the training data and very poor performances on the test patients.

TABLE IV. COMPARISON BETWEEN SWT-FCN, TEST\_1 AND TEST\_2

Test	Metrics									
	EPOCH-BASED					EVENT-BASED				
	AUC (%)	$AUC_{cc}$ (%)	ACC (%)	SEN (%)	SPE (%)	F1score (%)	MCC	GDR (%)	FD/h (h <sup>-1</sup> )	Latency-Time (s)
8s(db5) (FCN)	81	87	82	63	83	48	0.35	78	1.6	32
Test_1	76	81	78	55	80	41	0.26	75	1.9	40
Test_2	75	81	81	50	84	40	0.27	77	2	32

#### IV. DISCUSSION AND CONCLUSION

Results show a better performance of the proposed FCN with respect to the original CNN for all the parameters considered, in agreement with [15]. The results presented in Table IV show that the SWT is a valid support to reach a higher performance with respect to the original time series (Test\_1 and Test\_2). The SWT operation seems capable to better detect the discriminative information inside a seizure sample. Indeed it gives the best results in terms of Sensitivity (Table IV). Instead, the tests show the same ability to detect a non-seizure sample (Specificity). This may suggest that the original time series cannot provide a useful representation of the information for the networks to discriminate between a seizure sample and a non-seizure sample. This could be due to the fact that without any pre-processing some noisy or redundant information prevail that might significantly alter the classification rules produced by the networks during the training step. Another possible explanation is that separating the frequency information in their decomposition levels with SWT provides more details about higher frequencies (about 4-16 Hz for levels d4 and d5). Instead, when the original time series were considered, the lower frequencies (about 0-4 Hz for levels d6 and a6) might mask the information present in the higher frequencies, given their higher amplitude. In our opinion the separation operated by the SWT could effectively balance the information across frequencies and mitigate the effect of non significant information.

Among the considered time windows (4s, 8s and 16s), the best choice could be the 8s window that corresponds to the best  $AUC_{cc}$  value, although also the 4s one should be taken into account, thanks to its high values in Specificity and Accuracy. For the 16s window we found that the performances are lower than with the 8s one (CNN and FCN). The reason could be that, increasing the scale of observation, the useful details for a correct discrimination between seizure and not seizure become a negligible part of the available information.

The latency times obtained (Table IV) suggest that the systems can provide a reply to the user almost two times per minute (32, 40 seconds). This timing could provide an almost constant feedback about the brain condition of the newborn in NICU, and thus a valid support for an early detection of possible seizures events and timely application of the most suitable therapy for each patient.

The performance obtained for the 8s window are close to those presented in [15], both for epoch-based metrics ( $AUC$  83%) and event-based metrics (GDR 77% and FD/h 0.90). However, they are still lower than those obtained by O'Shea et al. in [15]. One reason could be that the training dataset used in [15] is larger than the dataset used here. Indeed, the authors used recordings with an overall duration of 834h with 1389 seizure events. Instead, the overall duration of the recordings in the dataset used here amounts to 112h and there are only 342 seizure events. As stated in [15] the deep-learning performance strictly depend on the available training data set; therefore, the possibility to have a large data set may represent one of the basic requirements for the success of the methods proposed in this field.

The performances shown in Table IV are the average among all the 39 patients for which all the three experts agree about seizures. Thanks to the Leave-One-Out validation, we have found that 5 patients among them show performances, in

terms of Sensitivity, well below the average of the observed population (about 18% for 8s windows). This result could be related to the seizure patterns of these patients that could be quite different from those of the other patients in the dataset. Therefore, the proposed system cannot correctly detect them.

We remark that we performed also the training and validation steps on this data set. Therefore, it could be considered as a starting point for future works concerning the same methodologies. Based on the information presented in Table IV, the results are still to be considered preliminary. The methods need further refinement before being comparable to the already existing machine-learning methods applied to the neonatal seizure detection problem. However, promising results are obtained, especially for the event-based metrics, highlighting that our methods might be used in the next future for a clinical validation even for different datasets. In addition, SWT operations could really increase the performances of deep-learning methods when applied to highly irregular time series such as newborn EEG.

Future work will be devoted to the neural network architectures used and to the choice of the mother-wavelet function for the SWT. Furthermore, to enlarge the case studies and solve the unbalancing problem between classes, we will explore the use of the Generative Adversarial Networks (GAN) [31, 32].

In conclusion, the work presented here is one of the first approaches to the application and testing of deep-learning methods on a public dataset. It provides first indications about the use of the Stationary Wavelet Transform in combination with deep-learning algorithms to solve the NSD problem, using a public dataset with multi-expert labels. The encouraging results show the possibility of successfully using these hybrid techniques. Moreover, they pave the way for a novel approach to perform data-augmentation for EEG time-series.

#### REFERENCES

- [1] M-P. Thibeault-Eybalin, Anne Lortie and Lionel Carmat "Neonatal Seizures: Do They Damage the Brain?" *Pediatr Neurol* 2009; 40: pp. 175-180.
- [2] SK. Kang and SD. Kadam "Neonatal Seizures: Impact on Neurodevelopmental Outcomes" *Front. Pediatr.* 2015 3:101. doi:10.3389/fped.2015.00101.
- [3] H. Ansari, P. J. Cherian, A. C. Dorado, K. Jansen, A. Dereymaeker, L. De Wispelaere et al. "Weighted Performance Metrics for Automatic Neonatal Seizure Detection Using Multiscored EEG Data" *IEEE Journal of Biomedical and Health Informatics*, 2018 Jul;22(4):1114-1123. doi: 10.1109/JBHI.2017.2750769.
- [4] AR. Hart, EL. Pilling and JJ. Alix "Neonatal seizures—part 1: Not everything that jerks, stiffens and shakes is a fit" *Archives of Disease in Childhood - Education and Practice* 2015;100: pp. 170-175.
- [5] R. Gursahani and N. Gupta "The adolescent or adult with generalized tonic-clonic seizures" *Ann Indian Acad Neurol.* 2012 Apr-Jun; 15(2): 81-88. doi: 10.4103/0972-2327.94988.
- [6] R.A. Shellhaas, A.I. Soaita and R.R. Clancy "Sensitivity of Amplitude-Integrated Electroencephalography for Neonatal Seizure Detection" *Pediatrics* October 2007, 120 (4) pp. 770-777 doi:10.1542/peds.2007-0514.
- [7] J.S. Malak, H. Zeerat, F.S. Nayeri, R. Safdari and A.D. Shahraki "Neonatal intensive care decision support systems using artificial intelligence techniques: a systematic review" *Artif Intell Rev* 2019 52: 2685 doi:10.1007/s10462-018-9635-1
- [8] R.O. Llyod, J.M. O'Toole, E. Pavlidis, P.M. Filan and G. B. Boylan "Electrographic Seizures during the Early Postnatal Period in Preterm Infants" *The Journal of Pediatrics*, Volume 187, pp. 18-25 2017.
- [9] A. Temko, A.K. Sarkar, G.B. Boylan, S. Mathieson, W.P. Marnane and G. Lightbody "Toward a Personalized Real-Time Diagnosis in

- Neonatal Seizure Detection” *IEEE J Transl Eng Health Med.* 2017 Sep 11;5:2800414. doi: 10.1109/JTEHM.2017.2737992.
- [10] K.T. Tapani, S. Vanhatalo and N.J. Stevenson “Time-Varying EEG Correlations Improve Automated Neonatal Seizure Detection.” *Int J Neural Syst.* 2019 May;29(4):1850030. doi: 10.1142/S0129065718500302. Epub 2018 Jun 24.
- [11] A. Emami, N. Kunii, T. Matsuo, T. Shinozaki, K. Kawai and H. Takahashi “Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images” *NeuroImage: Clinical* 22 (2019) 101684 doi:10.1016/j.nicl.2019.101684.
- [12] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H Falk, and J. Faubert “Deep learning-based electroencephalography analysis: a systematic review” *Journal of Neural Engineering*, Volume 16, Number 5 2019.
- [13] A. H. Ansari, P. J. Cherian, A. Caicedo, G. Naulaers, M. De Vos and S. Van Huffel, “Neonatal seizure detection using deep convolutional neural networks,” *International Journal of Neural Systems*, vol. 28, p. 1850011, 2018.
- [14] A. O’Shea, G. Lightbody, G. Boylan and A. Temko, “Neonatal seizure detection using convolutional neural networks,” *IEEE International Workshop on Machine Learning for Signal Processing*, pp. 1 -6, 2017.
- [15] A. O’Shea, G. Lightbody, G. Boylan and A. Temko “Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture” *Neural Networks*, Volume 123, 2020, pp. 12-25, ISSN 0893-6080.
- [16] G.P Nason and B.W. Silverman (1995) *The Stationary Wavelet Transform and some Statistical Applications*. In: Antoniadis A., Oppenheim G. (eds) *Wavelets and Statistics*. Lecture Notes in Statistics, vol 103. Springer, New York, NY.
- [17] Shorten and T. M. Khoshgoftaar “A survey on Image Data Augmentation for Deep Learning” *J Big Data* (2019) 6:60. doi: 10.1186/s40537-019-0197-0.
- [18] Y. Yuan, G. Xun, K. Jia and A. Zhang “A Multi-View Deep Learning Framework for EEG Seizure Detection” *IEEE Journal of Biomedical and Health Informatics*, Vol. 23, No. 1, January 2019.
- [19] N. J Stevenson., K. Tapani, L. Lauronen and S. Vanhatalo “A dataset of neonatal EEG recordings with seizure annotations.” *Sci. Data.* 6:190039 <https://doi.org/10.1038/sdata.2019.39> (2019).
- [20] MATLAB and Statistics and Machine Learning Toolbox Release 2019b. The MathWorks, Inc., Natick, Massachusetts, United States.
- [21] V. Rasoulzadeh, E. C. Erkus, T. A. Yogurt, I. Ulusoy and S. Aykan Zergeroglu “A comparative stationarity analysis of EEG signals” *Ann Oper Res* (2017) 258: 133-157, doi: 10.1007/s10479-016-2187-3.
- [22] G. Ekim ; N. Ikizler and A. Atasoy “The effects of different wavelet degrees on epileptic seizure detection from EEG signals” *Proceedings of the IEEE International Conference on INnovations in Intelligent SysTems and Applications, INISTA*; pp. 316-321, 2017.
- [23] P. S. Baboukani, G. Azemi, B. Boashash. P. Colditz and A. Omidvarnia “A novel multivariate phase synchrony measure: Application to multichannel newborn EEG analysis” *Digital Signal Processing* 84 (2019) pp. 59–68.
- [24] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*. MIT Press, 2016
- [25] N. Qian “On the momentum term in gradient descent learning algorithms” *Neural networks*, 12(1):145–151, 1999
- [26] A. L. Maas, Y.H. Awni, and A.Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." In *Proc. ICML*, vol. 30, no. 1. 2013.
- [27] A. Temko, E. Thomas, W. Marnane, G. Lightbody and G.B. Boylan “Performance assessment for EEG-based neonatal seizure detectors” *Clinical Neurophysiology* 122 (2011) pp. 474–482.
- [28] J. Gotman, D. Flanagan, J. Zhang and B. Rosenblatt “Automatic seizure detection in the newborn: methods and initial evaluation.” *Electroencephalogr Clin Neurophysiol* 1997;103(3): pp. 256–62.
- [29] L. Frassinetti, C. Barba, F. Melani, F. Piras, R. Guerrini and C. Manfredi “Automatic detection and sonification of nonmotor generalized onset epileptic seizures: Preliminary results” *Brain Research* 2019;1721:146341 doi: 10.1016/j.brainres.2019.146341.
- [30] L. Frassinetti, R. Guerrini, C. Barba, F. Melani, F. Piras and C. Manfredi “Sonification techniques applied to EEG signals of nonmotor generalized onset epileptic seizures” *11th International Workshop, Models and Analysis of Vocal Emissions for Biomedical Applications*, December, 17-19, 2019, ISSN 2704-5846, pp. 257-260. doi: 10.36253/978-88-6453-961-4.
- [31] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair et al. “Generative adversarial nets” *Proceedings of the Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.
- [32] K. G. Hartmann, R. T. Schirrmeyer and T. Ball “EEG-GAN: Generative adversarial networks for electroencephalographic (EEG) brain signals” 2018 arXiv:1806.01875.