ENEBI 2022 – VII Encontro Nacional de Engenharia Biomecânica
1 a 3 de junho de 2022

Epilepsy Seizure Detection Using Time and Frequency Domain EEG Signal via Convolutional Neural Network

Natalia Espinosa, UFMG - LabBio, nespinosa@ufmg.br, orcid:0000-0003-0945-4896
Clarissa Loures, UFMG, clarissalloures@ufmg.br, orcid: 0000-0002-6866-0051
Rudolf Huebner, UFMG - LabBio, rudolf@ufmg.br, orcid: 0000-0003-2613-304X

Abstract. *This work aims to develop an epilepsy detection system based on Convolutional Neural Networks (CNNs). The CNN model consists of 8-layers and receives three input signal types. Moreover, to evaluate the model performance is applied two training cases. In case 1, it was created one training set per patient. In case 2, one training set was built containing different seizure types. Hence, in case 1, the recall and precision achieved with both inputs in the time and frequency domain were: 44.1% / 43.8%, and 44.4% / 45.1%, respectively; whereas with the input in the time-frequency domain, the recall and precision were 81.6% and 92.5%, respectively. On the other hand, in case 2, evaluating the test session of the same patients used to train the model, the recall was 82% and precision was 86%. Demonstrating that the CNN model learns each person's brain-behavior, generalizing the knowledge, and detecting any seizure type.*

Keywords: *Epilepsy; Seizure detection; Convolutional Neural Network; Short Time Fourier Transform; Electroencephalogram.*

1. INTRODUCTION

Epilepsy is a common neurological disorder produced by transient and unexpected electrical disturbances in the brain characterized by unprovoked and recurrent seizures (Gautam et al., 2015). About 50 million people in the world suffer from this disease. Besides, the patient sometimes develops drug-resistant, making it not possible to treat, causing a strong economic, family, and personal impact on the patient (WHO, 2019). Electroencephalogram (EEG) is widely used to diagnose epilepsy since it is a noninvasive and low-cost study. However, its main limitation is that it takes a long time to analyze it by professionals specialized in EEG (Schirrmeister et al., 2017). For this reason, the automated epilepsy detection challenge begins.

Recently, Convolutional Neural Networks (CNNs) have been used to identify abnormalities in EEG recordings. Therefore, most researchers process the EEG signal as an image to use this technique because their main application is image recognition. In (Schirrmeister et al., 2017) reported the CNNs use to distinguish pathological from normal EEG using recordings from Temple University Hospital (TUH) database. The authors in (Zhou et al., 2018) use CNNs to detect the ictal, preictal, and interictal periods from EEG on the intracranial Freiburg Hospital and scalp Boston Children's Hospital–MIT (CHB-MIT) databases. The study by (Acharya et al., 2018) suggested a 13-layer deep convolutional neural network for epilepsy diagnosis applied to databases from Freiburg, CHB-MIT, and the American Epilepsy Society Seizure Prediction Challenge (Kaggle), this last is an EEG dataset from dogs. (Alhussein et al., 2019) proposed a convolutional neural network already trained (AlexNet) for pathology detection on the TUH database.

Hence, this work aims to develop an automated epilepsy detection system based on Convolutional Neural Networks. The CNN model consists of 8-layers and makes a binary classification where 1 represents ictal activity, no matter the epilepsy type that the patient presented. The algorithm performance is analyzed with two training cases. In addition, the CNN model is fed with three independent inputs, in the time domain,in the frequency domain via Fast Fourier Transform (FFT), and in time-frequency via Short Time Fourier Transform (STFT).

2. METHODS

The proposed method based on Convolutional networks works on 21 EEG derivations. The preprocessed EEG signal is filtered into three frequency bands, obtaining an input signal in the time domain. Then, the input signal in the frequency domain is got via FFT and the input in the time-frequency domain via STFT. Thus, three independent inputs are received by the CNN model. Next, the CNN hyperparameters are determined, and the model is trained using two training cases to evaluate the performance of the detection. So, in *case 1* was created one training set per patient, i.e., we have *eight training sets*, while, in *case 2* was built *one training set*, which contains several sessions with different seizure types. Finally, the model was tested on the test set that consists of one session per patient did not include in the training sets.

2.1 Database Description

The Temple University Hospital EEG (TUH) Corpus is the database used in this work that consists of EEG recordings and reports written by the attending neurologist of each patient. Thus, it is provided 1423 total sessions from 642 patients, whose 451 are seizure sessions from 241 patients, presenting 3055 total seizures, i.e., about 30% of the data contain seizure events (Ferrell et al., 2020; Shah et al., 2018).

To validate the CNN model, we have chosen eight seizure patients from the TUH Seizure Corpus database, each one with several sessions, obtaining a total of 74.1 hours of EEG recordings. The experimental group consisted of three women and five men between 24 to 62 years old, who presented different seizure types. As shown in Table 1 the chosen database consist of 40 complex partial seizures *(cpsz)*, 79 focal seizures *(fnsz)*, 20 generalized seizures *(gnsz)*, 17 tonic seizures *(tnsz)*, and 51 combined seizures *(fnsz/gnsz)*. The picked EEGs have a resolution of 16 bits and a sampling frequency of 256 Hz; besides, an analogical filter of 128 Hz low pass has been applied in EEG to remove artifacts. Additionally, for electrode placement, the international 10-20 system was used with 21 channels, applying a Temporal Central Parasagittal (TCP) montage (Ferrell et al., 2020).

Table 1. Summary of eight Epileptic Seizure Patients chosen from Temple University Hospital Seizure Corpus database

Patient	Age	Sex	Seizure Type	Training set		Test set	
				# Seizure	Time[h]	# Seizure	Time[s]
1	60	F	cpsz	19	10.96	1	601
2	43	M	fnsz	37	10.18	1	601
3	52	M	cpsz	18	4.50	2	514
4	24	F	tnsz	16	4.99	1	365
5	59	M	gnsz	19	5.37	1	601
6	62	M	fnsz/gnsz	23	14.30	1	702
7	39	F	fnsz/gnsz	24	10.09	3	699
8	59	M	fnsz	40	13.66	1	601

Furthermore, this experimental group (Table 1) was divided into two sets. The training set involves many sessions for each patient, whereas the test set contains only one session per patient. Thus, the training process has more data.

2.2 Signal Processing

To prepare the signal or preprocess it, the first recorded minute is deleted to remove any instrumentation artifacts in EEG. Next, the signal is split into 1-second size rectangular windows without overlap; yet, this procedure to split the signal is not used to obtain the Short-Time Fourier Transform (STFT), as shown later. In addition, the 60 Hz line power noise is rejected via a notch digital filter using a 5-th order Butterworth filter with cut frequency $\omega_{c_1} =$ 59 Hz and $\omega_{c_2} =$ 61 Hz. On the other hand, the signal is filtered in three frequency bands, 1 to 7 Hz, 8 to 30 Hz, and 30 to 100 Hz that include the standard brain waves delta, theta, alpha, beta, and gamma. In these frequency bands, brain activity is evaluated by patterns such as epileptic spikes and continuous background (van Drongelen, 2006).

Finally, we get the input in the time domain filtered in three frequency bands. Then, the filtered EEG signal is converted to the frequency domain via FFT and input in the time-frequency domain using STFT implemented with a Hann window, with 50% overlapping, and 256 samples per second are applied to the preprocessed EEG. Thus, three independent inputs are obtained. Figure 1 shows a representation of these inputs for each band-limited, there is a matrix of 21 rows corresponding to the number of EEG channels, for 1 second in the time domain, 256 samples per second in the frequency domain via FFT, and 1 second and 256 [Hz] in the time-frequency domain via STFT. These matrices represent the input image used to train the CNN network.

2.3 Convolutional Neural Networks

The Convolutional Neural Networks are widely used for image recognition, but recently, the challenge is to use them for EEG pattern recognition. CNNs employ the convolution math operation. Their architecture consists of an input layer that typically is an image set, convolutional layer, pooling layer, fully connected layer, and an output layer. In the convolutional layer are defined several filters (kernels) convolved with the input, using Eq. (1), to obtain the output known as feature maps (Goodfellow et al., 2016).

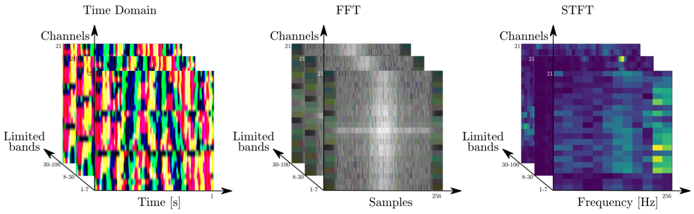

Figure 1. Representation of three input types received by CNN model; in time domain filtering in three frequency bands, in frequency domain via FFT and in time-frequency domain via STFT

$$y(t) = x * h, \tag{1}$$

where $y(t)$ is the feature map, x is the input signal and h is the filter.

Next, the previous convolutional layer output is downsampling in the Pooling layer. The most used pooling operation is max-pooling, which gets the maximum output within a rectangular region (Goodfellow et al., 2016). Then, in the fully connected layers occur the classification learned by the network. Finally, the output layer returns the network output.

Moreover, to build our CNN model, the activation functions Rectified Linear Unit (relu) and sigmoid are used. The relu is commonly used after convolutional layers to add nonlinearity and provide robustness to noise in the input data (Daoud and Bayoumi, 2019), and it is defined by Eq. (2).

$$f(x) = \begin{cases} x & if \quad x > 0 \\ 0 & if \quad x < 0 \end{cases}, \tag{2}$$

where x is the input signal and the negative values are eliminated.

Therefore, the sigmoid activation function Eq. (3) is applied to the output layer to estimate the probability of a binary event, such as normal vs epilepsy activity, since the output is a probability value between 0 or 1 (Goodfellow et al., 2016).

$$\rho(x) = \frac{1}{1 + e^{-x}}, \tag{3}$$

where x is the input signal and $\rho(x)$ is the probability of the input being an ictal event.

2.4 CNN model implementation

As seen in Figure 2, the CNN model architecture consists of two 2D convolutional layers, with 128 and 64 filters, respectively, both applying a 3x3 kernel size to reduce the computational cost and the number of weights for backpropagation. Furthermore, a stride of one in x and y directions is applied. Additionally, the nonlinearity is increased by the relu function to break up the linearity that might be imposed when the image is convolved, since an image is naturally nonlinear. Next, each convolutional layer is followed by a max-pooling layer over a 2x2 region and a stride of two for downsampling the image in half. Then, the preceding output is flatted, converting the 2D structure into a vector to enter into fully connected layers. Thus, it is created three fully connected layers with 64, 32, and 1 neuron, where in the firsts two layers are applied the relu function and in the last layer is used Sigmoid activation function to evaluate the probability of the binary class is an epileptic activity. The model was implemented with `keras` and `tensorflow` libraries in Phyton language.

The model is optimized according to the binary cross-entropy loss, applying the Adam optimizer, which begins with a 0.0001 learning rate and is adjusted throughout the training. Additionally, the accuracy of the validation set is evaluated while the model is trained. Then, it is possible to evaluate the learning process via analyzing the learning curves.

To train the model, several EEG sessions are concatenated after them have been processed as explained in the signal processing section, so these sessions are appended in a size array of 21 channels for the total time (the time sum of all sessions appended). The same process is made on the event files of each session. Thus, we obtain an array which contains several EEG recordings and one event vector to train the network. Besides, the training set is randomly split into two subsets, where 70% of EEG data is for the network training, and the last 30% of the EGG signal is used for model validation. On the other hand, it is used 20 epochs to train the model and 32 training samples per each iteration, known as batch size.

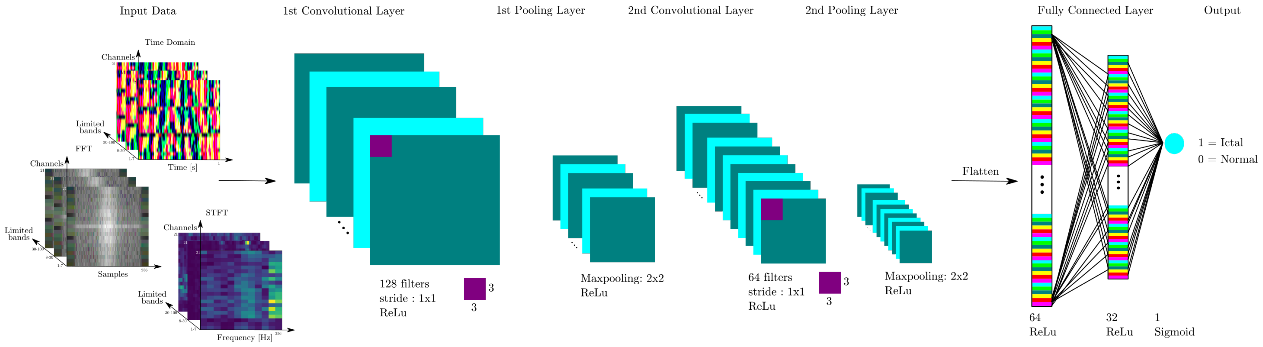

Figure 2. CNN model architecture, which consists of two convolutional layers, each one followed by a max-pooling layer. Then, the output is flatted and passed by three fully connected layers. Finally, the network classify the signal into 1 = ictal activity or 0 = seizure-free periods. The model receives three inputs one at time; in the time domain, frequency domain, and time-frequency domain.

3. RESULTS

This section shows the performance results of the proposed ictal activity detection algorithm based on CNN, using eight EEG recordings from the TUH database. Since this work aims the epilepsy seizure detection, all seizure events are classified as ictal activity. Thus, the network model classifies the EEG into two classes, 1 for the ictal event and 0 for the seizure-free period. The CNN model used 21 EEG channels and three bands limited to build the input 3D matrix, obtaining three input signals in the time domain, in the frequency domain, and time-frequency. The model performance was evaluated on two training cases; case 1 with eight training sets, one per patient, and case 2 with one training set, which contain all seizure types present in the database.

The results obtained for each input type for training case 1 are presented below. Table 2 on section Time Domain shows the epilepsy seizure detection results on the time domain signal for each analyzed patient. Despite to get acceptable results in five patients, it is not possible to detect any seizure event in patients 2, 6, and 8. Hence, the gotten results used this input were: accuracy of 57.6%, recall of 44.1%, precision of 43.8% and F-score of 43%, finally the average error of 42.4% is too high. Then the section frequency domain from Table 2 exhibits the epilepsy detection results on the frequency domain signal for each patient. Yet, as the same as the last results, in the same three patients it was not possible to detect ictal activity. However, the accuracy, recall, precision and F-score increase a little, obtaining 58.2%, 44.4%, 45.1%, and 44.4% respectively. Besides, the error is maintained high at 41.8%. Therefore, in CNN model fed with the input in the time and frequency domain did not apply the training case 2, due to the model tendency not change.

Finally, Table 2 in section time-frequency domain presents the performance results of the CNN model with time-frequency domain input obtained via STFT. The main success of this method was that the model was able to detect ictal activity in all patients, now, it is possible to see a great difference in the global results compared to previous methods, achieving an accuracy of 95.3%, recall of 81.6%, precision of 92.5%, and F-score of 85.7%, besides, a low error of 4.7% was obtained.

Table 2. CNN model performance using the three input types for training case 1

	Time Domain					Frequency Domain (FFT)					Time-Frequency Domain (STFT)				
Patient	**Acc.**	**recall**	**prec.**	**error**	**F-score**	**Acc.**	**recall**	**prec.**	**error**	**F-score**	**Acc.**	**recall**	**prec.**	**error**	**F-score**
1	88,4	29,0	41,9	11,6	34,3	95,1	64,5	85,1	4,9	73,4	96,5	71,0	97,8	3,5	82,2
2	-	-	-	100	-	-	-	-	100	-	99,1	90,9	100,0	0,9	95,2
3	87,5	93,4	93,0	12,5	93,2	89,8	97,2	92,2	10,2	94,6	91,2	100,0	91,2	8,8	95,4
4	94,2	55,6	43,5	5,8	48,8	98,6	94,4	81,0	1,4	87,2	97,1	83,3	71,4	2,9	76,9
5	98,7	77,8	77,8	1,3	77,8	95,8	5,6	11,1	4,2	7,4	99,3	84,2	94,1	0,7	88,9
6	-	-	-	100	-	-	-	-	100	-	96,1	75,0	100,0	3,9	85,7
7	91,8	96,8	94,2	8,2	95,5	86,3	93,1	91,8	13,7	92,4	93,6	98,8	94,3	6,4	96,5
8	-	-	-	100	-	-	-	-	100	-	89,9	50,0	90,9	10,1	64,5
Average	57,6	44,1	43,8	42,4	43,7	58,2	44,4	45,1	41,8	44,4	95,3	81,6	92,5	4,7	85,7

Since the results of training case 1 were much better than the others methods previously presented, we proceed to test training case 2. The training set for case 2 contains the sessions from patients 1, 4, and 6, involving all seizure types present on the whole database. Note that in case 1 we have one CNN model per patient, trained with the sessions of each one and tested with the session that did not participate in the training process, i.e., the test set shown in Table 1; whereas in case 2 we have one single CNN model trained with the sessions of patients 1, 4, and 6, then, the model was tested with the test set. Thus, Table 3 shows the performance of this final test, obtaining an accuracy, recall, precision, and F-score of 78.1%, 76.9%, 66.2%, and 65.4%, respectively, and the error increase to 21.9%. However, if we evaluate only the patients who used to train the model (1, 4, and 6), the performance was the accuracy of 96.8%, recall of 82%, precision of 86%, and F-score of 83.8%, moreover, the error of 3.2% is the lowest of all methods reported in this research. Finally, note that the test set in either of the cases was not included to train the CNN model, i.e., the test set was unknown for the model trained.

Table 3. CNN model performance using a Time-Frequency Domain input (STFT) for training case 2

Patient	accuracy	recall	precision	error	F-score
1	96,3	75,8	90,4	3,7	82,5
2	85,2	81,8	39,1	14,8	52,9
3	85,3	88,7	94,8	14,7	91,7
4	97,7	83,3	78,9	2,3	81,1
5	33,9	57,9	3,0	66,1	5,8
6	96,3	87,0	88,8	3,7	87,9
7	61,4	56,5	99,7	38,6	72,1
8	68,3	84,0	35,0	31,7	49,4
Average	78,1	76,9	66,2	21,9	65,4
1, 4, 6[a]	96,8	82,0	86,0	3,2	83,8

[a] Average between performances results of patients 1, 4 and 6.

Hence, Table 4 is presented the summary results of both training cases with each input type used. Such as seen in training case 1, the best performance was for the CNN model which received input in the time-frequency domain acquired via STFT, due to using this method, the ictal activity was detected in the whole experimental group, achieving accuracy, recall, and precision over 80% and an error less than 5%, which is acceptable. In training case 2, since it was possible to apply only in the last method, it was made a comparison between the global results and the results achieved if we evaluate only the patients who used in the training process, where the best performance was obtained for the last situation (average between patients 1,4, and 6), obtaining accuracy, recall, and precision over 82% and an error less than 3.5%. This proves that the CNN model developed has better results when is created an individual model, i.e., one model per patient (case 1), or when the future data of the trained model for some people correspond to these same people. In other words, the CNN model learned the "brain behavior" of each person.

Table 4. Compare performance of the two training cases

Case	Input Type	accuracy	recall	precision	error	F-score
case 1	time domain (filtering 3 limited bands)[a]	57,6	44,1	43,8	42,4	43,7
	frequency domain (FFT)[a]	58,2	44,4	45,1	41,8	44,4
	time-frequency (STFT)[a]	95,3	81,6	92,5	4,7	85,7
case 2	time-frequency (STFT)[a]	78,1	76,9	66,2	21,9	65,4
	time-frequency (STFT)[b]	96,8	82,0	86,0	3,2	83,8

[a] Average of the eight patients performance results
[b] Average between performance results of patients 1, 4 and 6

4. CONCLUSION

This study has assessed an automatic epilepsy seizure detection system based on Convolutional Neural Networks, applying three input types: in the time domain via filtering of three frequency bands, in the frequency domain via FFT, and in time-frequency via STFT. The model standardized with 21 EGG channels used eight EEGs from Temple Hospital

University (THU) Corpus database. Moreover, to evaluate the model performance was applied two training cases. In case 1, it was created one training set per patient. In case 2, one training set was built, which contains several sessions with different seizure types. Finally, both cases were tested on the same test set. The CNN model met the objective of epilepsy seizure detection with all input types, regardless of the seizure type.

Hence, the best performance achieved was for CNN model with input in time-frequency domain creating a model per person. Achieving an accuracy of 95.3%, recall of 81.6%, the precision of 92.5%, F1-score of 85.7%, and low error of 4.7%. On the other hand, training case 2 was applied only to the CNN model with input in time-frequency domain acquired via STFT because only this method was able to detect the ictal activity in the whole patients tested. The model was trained with training sessions for patients 1, 4, and 6. Obtaining an accuracy, recall, precision, F1-score, and error of: 78.1%, 76.9%, 66.2%, 65.4%, and 21.9% respectively. However, if we evaluate only the test session of the same patients used to train the model, the performance was better, achieving an accuracy of 96.8%, recall of 82%, the precision of 86%, and F1-score of 83.8%, and the error of 3.2% which was the lowest of all methods reported in this research. Given these points, the best performance was achieved by the CNN model with the input in the time-frequency domain created for each patient, which proves that the model learned each person's brain behavior.

5. REFERENCES

Shreya Gautam, Shreety Sriya, and Tanuj Chauhan. Focal and non-focal epilepsy detection using EEG signals via empirical mode decomposition. In *2015 International Conference on Signal Processing and Communication, ICSC 2015*, pages 452–455. Institute of Electrical and Electronics Engineers Inc., Jul. 2015. ISBN 9781479967612. doi: 10.1109/ICSPCom.2015.7150696.

Epilepsia, *World Health Organization,* 2019. [Online]. Available: https://www.who.int/es/news-room/fact-sheets/detail/epilepsy. [Accessed: 21-June-2019].

R. Schirrmeister, L. Gemein, K. Eggensperger, F. Hutter, and T. Ball. Deep learning with convolutional neural networks for decoding and visualization of EEG pathology. *2017 IEEE Signal Processing in Medicine and Biology Symposium, SPMB 2017 - Proceedings*, 2018-Janua:1–7, 2017. doi: 10.1109/SPMB.2017.8257015.

Mengni Zhou, Cheng Tian, Rui Cao, Bin Wang, Yan Niu, Ting Hu, Hao Guo, and Jie Xiang. Epileptic seizure detection based on EEG signals and CNN. *Frontiers in Neuroinformatics*, 12(December):1–14, 2018. ISSN 16625196. doi: 10.3389/fninf.2018.00095.

U. Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Hojjat Adeli. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine*, 2018. ISSN 18790534. doi: 10.1016/j.compbiomed.2017.09.017.

Musaed Alhussein, Ghulam Muhammad, and M. Shamim Hossain. EEG Pathology detection based on deep learning. *IEEE Access*, 7:27781–27788, 2019. ISSN 21693536. doi: 10.1109/ACCESS.2019.2901672.

Sean Ferrell, Vineetha Mathew, Matthew Refford, Vincent Tchiong, Tameem Ahsan, Iyad Obeid, and Joseph Picone. Electrode Location and Channel Labels. Technical report, The Neural Engineering Data Consortium College of Engineering, Temple University, Philadelphia, Pennsylvania, 2020. URL `https://www.isip.piconepress.com/publications/reports/2020/tuh_eeg/electrodes/`.

Vinit Shah, Eva von Weltin, Silvia Lopez, James Riley McHugh, Lillian Veloso, Meysam Golmohammadi, Iyad Obeid, and Joseph Picone. The temple university hospital seizure detection corpus. *Frontiers in Neuroinformatics*, 12 (November):1–6, 2018. ISSN 16625196. doi: 10.3389/fninf.2018.00083.

Win van Drongelen. *Signal Processing for Neuroscientists*. Elsevier Academic Press, 2006. ISBN 13:978-0-12-370867-0.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

Hisham Daoud and Magdy A. Bayoumi. Efficient Epileptic Seizure Prediction Based on Deep Learning. *IEEE Transactions on Biomedical Circuits and Systems*, 13(5):804–813, 2019. ISSN 19409990. doi: 10.1109/TBCAS.2019.2929053.

6. ACKNOWLEDGEMENT(s)

The authors thank the Universidade Federal de Minas Gerais (UFMG) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for all financial support via project number 830195/1999-8, process 133009/2020-3. We would also like gratefuly to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to allow us web access to www.periodicos.capes.gov.br.

7. INFORMATION RESPONSIBILITY

The authors are solely responsible for the information included in this work.