Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 18-114

Supervised Transfer Learning for Personalized Heart Rate Based Epileptic Seizure Detection[1]

Thomas De Cooman, Carolina Varon, Wim Van Paesschen, Sabine Van Huffel[2]

September 2018
Submitted for publication in Clinical Autonomic Research

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *ftp://ftp.esat.kuleuven.be/pub/stadius/tdecooman/reports/SISTA_18_114_submit.pdf*

[2]T. De Cooman, C. Varon and S. Van Huffel are affiliated with KU Leuven, Dept. of Electrical Engineering (ESAT), Research group STADIUS, Kasteelpark Arenberg 10, 3001 Leuven, Belgium and imec, Leuven, Belgium. W. Van Paesschen is affiliated with Department of Neurology, UZ Leuven, Leuven, Belgium.

Acknowledgments: imec: Strategic Funding 2017, ICON: HBC.2016.0167 SeizeIT; European Research Council. The research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (n 339804). This paper reflects only the authors views and the Union is not liable for any use that may be made of the contained information. Thomas De Cooman is supported by an FWO SBO PhD grant. Carolina Varon is a postdoctoral fellow of the Research Foundation-Flanders (FWO).

Supervised Transfer Learning for Personalized Heart Rate Based Epileptic Seizure Detection

Thomas De Cooman[1,2], Carolina Varon[1,2], Wim Van Paesschen [3], Sabine Van Huffel [1,2]

[1] Department of Electrical Engineering (ESAT), STADIUS, KU Leuven, Leuven, Belgium
[2] imec, Leuven, Belgium
[3] Department of Neurology, UZ Leuven, Leuven, Belgium

Abstract

Seizure alarm systems can improve the quality of life of refractory epilepsy patients, but require seizure detection algorithms. State-of-the-art heart rate-based algorithms often use a patient-independent approach due to insufficient annotated patient data, which does not allow a robust personalization. Ictal heart rate changes are however patient-dependent and could benefit from personalized algorithms. In this study, we propose to personalize seizure detection by using supervised transfer learning, which allows to train a classifier with a limited amount of data by using a reference classifier. It is evaluated on 207 hours of data including 74 seizures from 6 patients. An optimal performance of 89.8% sensitivity was achieved with on average 1.1 false alarms per hour, which is 54% false alarms less than the reference patient-independent classifier by using a limited amount of patient data. This shows that transfer learning can be used for a fast and robust personalization of detection algorithms.

1. Introduction

Epilepsy is one of the most common neurological diseases, affecting around 1.5% of the people worldwide. Despite the availability of different treatments, 30% of the patients continue to have epileptic seizures. The quality of life of these refractory epilepsy patients is strongly affected by the unexpected occurrence of these seizures. A possible option to improve this is to use an automated alarm system. It automatically detects ongoing seizures and alarms the patient's caregiver about such an event [1]. They can then come over to the patient and provide them with the proper treatment.

Such a system requires an automated seizure detection algorithm. Although most detection algorithms in literature use electroencephalography (EEG), wearing a full EEG cap for long-term monitoring is not comfortable nor feasible for patients [2]. Therefore, more wearable sensors should be used for seizure detection. Previous studies already showed that heart rate changes occur in around 70% of the seizures, with a higher prevalence and magnitude in temporal lobe epilepsy and generalized seizures [3]. Most state-of-the-art heart rate-based seizure detection algorithms use patient-independent algorithms [4], but heart rate based algorithms could benefit from patient-specific models due to the patient-dependency of the heart rate characteristics [5]. In practice it is, however, difficult to get a fully patient-specific model due to the limited amount of available annotated patient data. Moreover, this type of data often does not contain a sufficient amount of seizures for a robust classifier.

Therefore a transfer learning approach is proposed here to efficiently train a patient-specific classifier for seizure detection. In transfer learning, a classification problem is solved with a limited amount of data by using a reference solution of a similar problem. In this study, a patient-specific classifier is trained using a limited amount of patient-specific data by using the patient-independent classifier as a reference model.

2. Data

The dataset used in this study consists of data acquired from 6 patients with temporal lobe epilepsy (TLE) who underwent presurgical evaluation at UZ Leuven. The dataset consists of 207 hours of continuous data, containing 74 focal impaired awareness seizures (see table 1). Seizures were annotated by clinical experts using the gold video-EEG standard. Only single-lead ECG is used in this study, with a sampling frequency of 250 Hz.

3. Methods

3.1. Preprocessing

The same preprocessing procedure used in [4] is applied according to the following steps. First, the heart rate is

Table 1. Overview of the dataset used in this study.

Patient	Age	# seizures	Length (h)
1	49	10	26.3
2	41	9	63.1
3	15	10	24.7
4	29	11	47.1
5	27	4	30.6
6	9	30	67.0
Total	[9-49]	74	258.7

extracted in real-time from the ECG using an R peak detection algorithm based on adaptive thresholding on the derivative signal. Next, the tachogram is filtered using a median filter with a filter order of 10 heart beats. Then, strong heart rate increases (HRI) are detected by heart rate slope analysis. If the slope exceeds a value of 1bpm/s, it is assumed that a strong HRI is occurring which is similar to an epileptic HRI. The start and end of this HRI is then found by looking back and forward when the slope becomes negative again.

Whenever such a strong HRI is detected, features are extracted from this HRI and the period before the HRI. In [4], it was shown that the peak heart rate (the maximal heart rate during the HRI) and the average pre-HRI heart rate are the most differentiating features. Only these 2 features are used for classification as they already led to the largest portion of performance. Adding more features also requires more patient data for the proposed transfer learning approach to robustly personalize the classifier.

3.2. Classification

Transfer learning is a machine learning method that solves a classification problem with a limited amount of data by using the solution of a similar problem. In this study, it is used to train a patient-specific heart rate based seizure detection classifier by using a reference patient-independent classifier. This way, the classifier can be personalized with a limited amount of patient-specific data.

3.2.1. Patient-independent classifier

The patient-independent (P-I) classifier is used as the reference classifier for the transfer learning approach discussed in section 3.2.2. It is trained without using any data from the patient that is tested using a leave-one-patient-out approach as in [4]. P-I classification is done here by using weighted support vector machines (SVM), which is found by solving the following optimization problem [6]

$$\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} c_i\xi_i \\ &s.t. \quad \begin{cases} y_i(w^T\varphi(x_i)+b) \geq 1-\xi_i \\ \xi_i \geq 0 \end{cases} \quad ,\forall i \in [1,N] \end{aligned} \tag{1}$$

with x_i the P-I training points, y_i the corresponding annotations, ξ_i the errors for x_i, N the number of data points, φ the non-linear transformation and w and b the unknown variables to obtain from the minimization problem. The values of c_i are defined as

$$c_i = \begin{cases} \gamma * \frac{(N^+ + N^-)}{2*N^-} & : y_i = -1, \text{seizure data} \\ \frac{N^+ + N^-}{2*N^+} & : y_i = +1, \text{non-seizure data} \end{cases} \tag{2}$$

wit N^+ and N^- the number of non-seizure and seizure data points in the dataset. The values of the hyperparameters C and γ and the applied RBF kernel parameter σ are taken as in [4].

3.2.2. Personalized classification

Transfer learning is applied to SVM in order to personalize the classifier to each patient. We therefore propose to solve the following problem for patient-specific data $\tilde{x}_k$ and annotations $\tilde{y}_k$, which is an SVM adaptation of the solution proposed for least-squares SVM in [7]:

$$\begin{aligned} &\min_{\tilde{w},\tilde{b},\tilde{\xi}} \frac{1}{2}||\tilde{w}-\beta w||^2 + D\sum_{k=1}^{M} \tilde{c_k}\tilde{\xi_k} \\ &s.t. \quad \begin{cases} \tilde{y_k}(\tilde{w}^T\varphi(\tilde{x}_k)+\tilde{b}) \geq 1-\tilde{\xi}_k \\ \tilde{\xi}_k \geq 0 \end{cases} \quad ,\forall k \in [1,M] \end{aligned} \tag{3}$$

with w the weight vector found from the P-I classifier from equation (1), defined as

$$w = \sum_{i=1}^{N} \alpha_i y_i \varphi(x_i) \tag{4}$$

by the original SVM problem with α_i the Langrange multipliers from (1). Parameters $\tilde{c}_k$ are similar as in (2), whereas hyperparameter D has a similar impact on the optimization problem as C in (1). The β parameter ($0 \leq \beta \leq 1$) indicates how similar the patient-specific $\tilde{w}$ should be with respect to the P-I w. The dual problem is then defined as

$$\begin{aligned} min_{\widetilde{\alpha}} =& \beta\sum_{i=1}^{N}\sum_{k=1}^{M} y_i\widetilde{y}_k\alpha_i\widetilde{\alpha}_k K(x_i,\widetilde{x}_k) - \sum_{k=1}^{M}\widetilde{\alpha}_k \\ &+ \frac{1}{2}\sum_{k=1}^{M}\sum_{l=1}^{M}\widetilde{y}_k\widetilde{y}_l\widetilde{\alpha}_k\widetilde{\alpha}_l K(\widetilde{x}_k,\widetilde{x}_l) \\ s.t. \quad & \begin{cases} \sum_{k=1}^{M}\widetilde{\alpha}_k\widetilde{y}_k = 0 \\ 0 \leq \widetilde{\alpha}_k \leq D * \tilde{c}_k \end{cases} \quad ,\forall k \in [1,M] \end{aligned} \tag{5}$$

with $K(x_i,x_k) = \varphi(x_i)^T\varphi(x_k)$ the used Gaussian kernel with the same kernel parameter σ as used in the reference P-I classifier. Note that it can be found that

$$\tilde{w} = \beta w + \sum_{k=1}^{M} \tilde{\alpha}_k\tilde{y}_k\varphi(\tilde{x}_k) \tag{6}$$

which indeed illustrates that the β parameter indicates the level of importance of the reference classifier with respect to the patient-specific information contained in $\tilde{x}_k$.

The classifier is trained and tested using 5-fold crosstesting on the available data per patient. The reference P-I classifier is trained using data from all patients except the patient under evaluation. The corresponding classification of the data point $\tilde{x}_n$ is then defined by

$$y(\tilde{x}_n) = \beta \sum_{k=1}^{N} \alpha_k y_k K(x_i, \tilde{x}_n) + \sum_{k=1}^{M} \tilde{\alpha}_k \tilde{y}_k K(\tilde{x}_k, \tilde{x}_n) + \tilde{b}. \tag{7}$$

A personalized alarm is only generated in this study if the HRI is classified as an epileptic HRI by both the P-I and personalized classifier. The main purpose of the personalized approach is to lower the false alarm rate while maintaining a similar sensitivity.

3.3. Performance evaluation

The mentioned seizure detection algorithms are evaluated using sensitivity (percentage of detected seizures) and false alarm rate (FAR, expressed in false positives per hour (FP/h)). A seizure is said to be detected if an alarm is generated between 30 s before and 60 s after the seizure onset. Also the positive predictive value (PPV) is used to combine the information of the two above mentioned metrics in one metric. The detection delay indicates the time difference between the seizure onset and the moment of detection by the seizure detection algorithm.

4. Results and discussion

The mentioned P-I approach reached an overall sensitivity of 93.9% and an average FAR of 2.6 FP/h. Figure 1 shows the impact of parameters D and β on the performance. Values for parameter D were mostly found to be optimal around 0.1. The best performance with a sensitivity higher than 90% is found for $\beta = 0.75$ and $D = 0.1$, with an overall sensitivity of 92.1% and 1.4 FP/h. The average detection delay was 11.7 s.

The β parameter indicates the importance of the reference classifier during training. In the above mentioned simulation, the same β value was chosen for all patients. The optimal β value however depends on the patient and the amount of patient-specific data. This is illustrated in Figure 2, which shows the impact of the number of seizures in the training set on the PPV for different β values (with optimal D value) for patient 6. Seizures used in training are randomly chosen from the dataset, and simulations were repeated 100 times for each result. If only 1 seizure is used, a performance decrease is found compared to the P-I approach (0 seizures used) for $\beta < 0.5$, which indicates that there is too few data to give too much

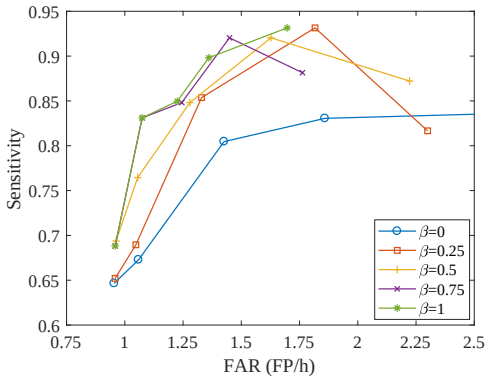


Figure 1. Impact of D and β on the performance of the proposed transfer learning approach. Parameter D is evaluated for logarithmic values between 0.01 and 100, with the values of 100 starting at the bottom left of the figure.

importance to only the patient-specific data. For higher β values, an increased performance is found, with $\beta = 1$ resulting in the best performance. When 5 seizures are used during training, all performances increase, with the biggest gain found in the lower β values. The best result is still found for $\beta = 1$. This however changes when 10 seizures are used. The results for all βs are quite similar, with the best performance obtained with $\beta = 0.75$. If all seizures available in the training are used (typically around 20-25 seizures for this patient), the optimal performance is found with $\beta = 0.25$, and the worst result is found with $\beta = 1$. When sufficient patient-specific data is available, the reference classifier is of less importance, and higher performance can be found by trying to fully optimize the classifier only on the patient-specific data. Overall, it can be seen that lower β values lead to stronger increases when the number of seizures increases, whereas for higher β values, the increase is limited after 5 available seizures. This illustrates that the optimal β value depends on the amount of available patient data, and should be set automatically together with parameter D in an hyperparameter optimization procedure during training.

Figure 2 illustrates that it will take around 10 seizures for patient 6. This breakpoint varies for each patient, depending on the average strength of ictal HRIs and the inter-seizure variability in ictal heart rate changes. For patients who typically only have nocturnal seizures, less patient-specific seizures might be required than for patients who experience seizures throughout the day. This is due to the fact that the ictal heart rate changes also depend on the current state of the patient (sleep/awake, resting/active, relaxed/stressed, ...). When optimal β values are chosen per patient, an overall sensitivity of 92.2% is found with on average 1.1 FP/h, which is a drop of 54% in FAR compared to the P-I approach with a similar sensitivity. The results

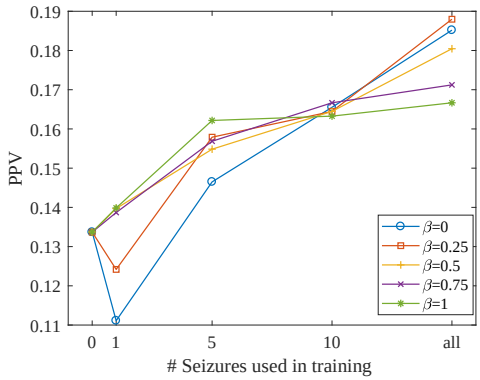


Figure 2. Impact of the amount of seizures in the training set on the average positive predictive value (PPV) for different β values for patient 6.

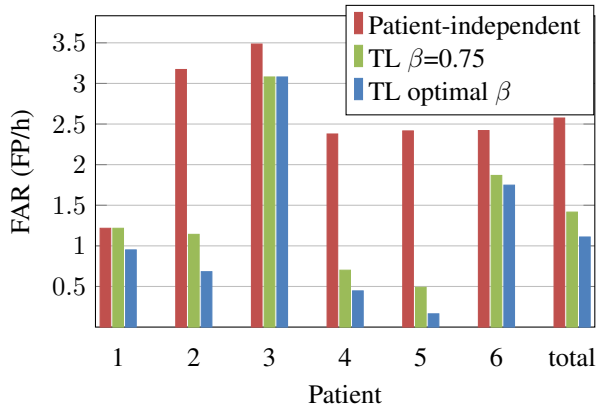


Figure 3. Average FAR per patient and in total for the P-I approach, personalized approach with fixed β and personalized approach with patient-specific optimal β.

per patient for both the personalized approach with fixed and optimal β values are compared with the P-I approach in Figure 3. It can be seen that for some patients a strong decrease in FAR can be found, whereas the added value of personalization is limited in some patients. This is the case in patients with rather small ictal HRIs, which make it hard to differentiate them from non-epileptic HRIs.

Personalization leads to an increased performance for heart rate based seizure detection. The performance is however still too low to be used in practice in an unimodal setting. The proposed methodology should be be part of a multimodal setting, including other modalities such as accelerometers, ear-EEG or electromyogram [1]. The applied methodology can also be applied on these modalities to personalize these algorithms in a fast and robust way.

5. Conclusion

Personalizing heart rate-based seizure detection leads to strong performance increases. The proposed transfer learning method allows to personalize the algorithm robustly and allows to easily adapt to situations with both limited and sufficient amount of patient data, resulting in a maximal FAR decrease of 54%. In future work, it will also be applied in seizure detection using other modalities.

Acknowledgements

Research supported by imec: Strategic Funding 2017, ICON: HBC.2016.0167 SeizeIT; European Research Council. The research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (n 339804). This paper reflects only the authors views and the Union is not liable for any use that may be made of the contained information. Thomas De Cooman is supported by an FWO SBO PhD grant. Carolina Varon is a postdoctoral fellow of the Research Foundation-Flanders (FWO).

References

[1] Van de Vel A, Cuppens K, Bonroy B, Milosevic M, Jansen K, Van Huffel S, Vanrumste B, Cras P, Lagae L, Ceulemans B. Non-eeg seizure detection systems and potential sudep prevention: State of the art: Review and update. Seizure 2016;41:141–153.

[2] Schulze-Bonhage A, Sales F, Wagner K, Teotonio R, Carius A, Schelle A, Ihle M. Views of patients with epilepsy on seizure prediction devices. Epilepsy behavior 2010; 18(4):388–396.

[3] Jansen K, Lagae L. Cardiac changes in epilepsy. Seizure 2010;19(8):455–460.

[4] De Cooman T, Varon C, Hunyadi B, Van Paesschen W, Lagae L, Van Huffel S. Online automated seizure detection in temporal lobe epilepsy patients using single-lead ecg. International Journal of Neural Systems 2017;1750022.

[5] De Cooman T, Kjaer TW, Van Huffel S, Sorensen HB. Adaptive heart rate-based epileptic seizure detection using real-time user feedback. Physiological Measurement 2018; 39(1):014005.

[6] Vapnik V, Guyon I, Hastie T. Support vector machines. Mach Learn 1995;20(3):273–297.

[7] Tommasi T, Orabona F, Caputo B. Safety in numbers: Learning categories from few examples with multi model knowledge transfer. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. ISSN 1063-6919, June 2010; 3081–3088.

Address for correspondence:

Thomas De Cooman
Kasteelpark Arenberg 10 box 2446
3001 Leuven, Belgium
thomas.decooman@esat.kuleuven.be