KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Thomas De Cooman, Kaat Vandecasteele, Carolina Varon, Borbala Hunyadi, Evy Cleeren, Wim Van Paesschen and Sabine Van Huffel (2020), **Personalizing Heart Rate-Based Seizure Detection Using Supervised SVM Transfer Learning** Frontiers in Neurology, vol. ?, 2020, ??-??.
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	TBD
Journal homepage	https://www.frontiersin.org/journals/neurology
Author contact	tdc@esat.kuleuven.be kaat.vandecasteele@esat.kuleuven.be +32 (0) 16 32 73 60
Abstract	Objective: Automated seizure detection is a key aspect of wearable seizure warning systems. As a result, the quality of life of refractory epilepsy patients could be improved. Most state-of-the-art algorithms for heart rate-based seizure detection use a so-called patient-independent approach, which do not take into account patient-specific data during algorithm training. Although such systems are easy to use in practice, they lead to many false detections as the ictal heart rate changes are patient-dependent. In practice, only a limited amount of accurately annotated

patient data is typically available, which makes it difficult to create fully patient-specific algorithms.
Methods: In this context, this study proposes for the first time a new transfer learning approach
that allows to personalize heart rate-based seizure detection by using only a couple of days of data per patient. The algorithm was evaluated on 2172 hours of single-lead ECG data from 24 temporal lobe epilepsy patients including 227 focal impaired awareness seizures.
Results: The proposed personalized approach resulted in an overall sensitivity of 71% with 1.9 false detections per hour. This is an average decrease in false detection rate of 37% compared to the reference patient-independent algorithm using only a limited amount of personal seizure
data. The proposed transfer learning approach adapts faster and more robustly to patient-specific characteristics than other alternatives for personalization in the literature.
Conclusion: The proposed method allows an easy implementable solution to personalize heart rate-based seizure detection, which can improve the quality of life of refractory epilepsy patients when used as part of a multimodal seizure detection system.

IR	url in Lirias TBD

(article begins on next page)

frontiers

Personalizing Heart Rate-Based Seizure Detection Using Supervised SVM Transfer Learning

Thomas De Cooman [1,*], **Kaat Vandecasteele** [1,*], **Carolina Varon** [1,2], **Borbála Hunyadi** [2], **Evy Cleeren** [3], **Wim Van Paesschen** [3] and **Sabine Van Huffel** [1]

[1] *STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium*
[2] *Department of Microelectronics, TU Delft, Delft, Netherlands*
[3] *Department of Neurosciences, KU Leuven, University Hospitals Leuven, Leuven, Belgium*

Correspondence*:
Thomas De Cooman
tdc@esat.kuleuven.be

Kaat Vandecasteele
kaat.vandecasteele@esat.kuleuven.be

Word count: 7080 words

ABSTRACT

Objective: Automated seizure detection is a key aspect of wearable seizure warning systems. As a result, the quality of life of refractory epilepsy patients could be improved. Most state-of-the-art algorithms for heart rate-based seizure detection use a so-called patient-independent approach, which do not take into account patient-specific data during algorithm training. Although such systems are easy to use in practice, they lead to many false detections as the ictal heart rate changes are patient-dependent. In practice, only a limited amount of accurately annotated patient data is typically available, which makes it difficult to create fully patient-specific algorithms.
Methods: In this context, this study proposes for the first time a new transfer learning approach that allows to personalize heart rate-based seizure detection by using only a couple of days of data per patient. The algorithm was evaluated on 2172 hours of single-lead ECG data from 24 temporal lobe epilepsy patients including 227 focal impaired awareness seizures.
Results: The proposed personalized approach resulted in an overall sensitivity of 71% with 1.9 false detections per hour. This is an average decrease in false detection rate of 37% compared to the reference patient-independent algorithm using only a limited amount of personal seizure data. The proposed transfer learning approach adapts faster and more robustly to patient-specific characteristics than other alternatives for personalization in the literature.
Conclusion: The proposed method allows an easy implementable solution to personalize heart rate-based seizure detection, which can improve the quality of life of refractory epilepsy patients when used as part of a multimodal seizure detection system.

Keywords: epilepsy, transfer learning, seizure detection, personalization, heart rate analysis, SVM

1 INTRODUCTION

Epilepsy is one of the most common neurological disorders, which affects around 1% of the population worldwide (Forsgren et al., 2005). Anti-epileptic drugs provide adequate treatment for about 70% of epilepsy patients (French, 2007). The remaining 30% of the patients continue to have seizures, which drastically affects their quality of life. This can be improved by an automated warning system that alarms the parents or caregivers when the patient experiences a seizure. By using such a system, the patients and relatives feel more at ease knowing someone will be able to help the patient out when a seizure would occur. Quick intervention can then lead to a decrease in injuries and avoid (post-)ictal complications, including sudden unexplained death in epilepsy (SUDEP). In addition, a seizure diary, automatically generated from the alarms, can be used for a follow-up of the disease and evaluation of the treatment. A seizure diary, kept by the patients or their families, has proven to be unreliable, which leads to bad treatment follow-up (Fisher et al., 2012). An automatically generated seizure diary could lead to a more objective seizure count and an improved treatment selection.

The key element of such a warning system is the automated seizure detection algorithm. In the literature, these algorithms are typically based on full electroencephalography (EEG). EEG recordings mostly require wet electrodes on the scalp, which is uncomfortable for a long-term monitoring solution (Schulze-Bonhage et al., 2010). More easily obtainable biomedical signals used to detect epileptic seizures include accelerometry (ACC), electromyography (EMG), electrocardiography (ECG) and galvanic skin response (Van de Vel et al., 2016). The most suitable modality or combination of modalities depends on the seizure type. ECG-based seizure detection, for instance, is ideal for the detection of focal impaired awareness seizures (FIAS) arising from the temporal lobe, as they are not accompanied by typical motor components, but they are associated with ictal tachycardia (Zijlmans et al., 2002; Jansen and Lagae, 2010; Leutmezer et al., 2003). Therefore, this study focuses on patients with temporal lobe epilepsy, as for this type of focal seizures, ECG-based seizure detection is of most added value compared to other wearable modalities. It should however be noted that ECG-based seizure detection algorithms can be used for a wider range of seizure types, including focal seizures with non-temporal seizure onsets and generalized tonic-clonic seizures (De Cooman et al., 2018). It is also a very important modality for long-term monitoring applications as it allows to assess the patient's general health status (e.g. sleep and general heart conditions).

Most ECG-based seizure detection systems from the literature are based on patient-independent models (De Cooman et al., 2017a; Osorio, 2014; Ungureanu et al., 2014; Massé et al., 2013). For this type of models, no patient-specific data is required, making them directly usable in practice. However, due to the large inter-patient differences in heart rate characteristics, performance is too low for practical use.

In order to increase the performance, models can be adapted to the patient heart rate characteristics (De Cooman et al., 2018; van Elmpt et al., 2006). Different options are possible. A first option is the manual setting of some parameter thresholds per patient (van Elmpt et al., 2006; Boon et al., 2015). This requires manual screening of previous patient data, and it only works well if the parameters are easily understandable. Simple thresholding approaches are however too simple to grasp the large complexity of the problem. Automated personalization is therefore advised, but it normally requires a lot of patient-specific data in order to find a robust algorithm for a specific patient (Cogan et al., 2016). Often, only a limited amount of accurately annotated patient-specific data is available, hence a lot of complex approaches are not useful for making fully patient-specific classifiers. Heuristic automated algorithms allow a low-complex and fast personalization, but might lead to suboptimal results (De Cooman et al., 2018).

A more robust and optimal solution can be found through transfer learning (Pan and Yang, 2009; Weiss et al., 2016). In transfer learning, the solution to a classification problem is found by using the solution

from a similar problem (”source task”) as a starting point (see Figure 1). This way, less data for the new problem (”target task”) is required in order to get a robust solution as a part of the knowledge is already contained in the reference model. In this paper, a patient-specific heart rate-based seizure detection model is trained through transfer learning by using a patient-independent classifier as a reference model. Therefore, only a limited amount of patient-specific heart rate data is required in order to get a robust patient-specific model, obtained with a relative limited complexity. This paper proposes a new enhanced transfer learning solution that is also able to deal with the large class imbalance and allows to give more importance to seizure samples during classifier training for each individual.

The novelty of this study is threefold. First, a new transfer learning procedure is proposed. An existing transfer learning method is enhanced in order to deal with class imbalance and allows to give increased importance to sensitivity rather than specificity. The latter is a crucial aspect for automated seizure detection. Secondly, the proposed transfer learning approach is applied in order to efficiently personalize heart rate-based seizure detection with a limited amount of patient data. To the best of our knowledge, it is the first time transfer learning is used to personalize automated non-EEG-based seizure detection. Unimodal seizure detection systems based on heart rate typically lead to inaccurate results that are insufficient to be used in practice (van Westrhenen et al., 2018). In this context, the proposed methodology offers a novel solution that can be used as part of a multimodal system, which allows to adapt the detection system to each patient. As a result, a more accurate and usable solution could be achieved (De Cooman et al., 2018). These multimodal systems are currently getting close to sufficiently high accuracy for practical usage, but are typically restricted to a certain seizure type (Leijten et al., 2018). The added value of using heart rate in a multimodal system is that it is the ictally most activated modality besides EEG, which allows detection of a wide range of seizure types. Improving the unimodal heart rate-based seizure detection using the proposed method will also improve the multimodal seizure detection, closing the gap to practical usage for a wide range of seizure types. Finally, an in-depth analysis is performed in order to indicate the added value of the proposed transfer learning approach compared to the state-of-the-art literature of heart rate-based seizure detection.

2 METHODS AND MATERIALS

2.1 Data Acquisition

The dataset contains recordings from refractory epilepsy patients, who underwent presurgical evaluation at the University Hospitals Leuven (UZ Leuven), Belgium and had at least five FIAS originating from the temporal lobe during the evaluation. The patients were recorded with 10-20 scalp EEG with 1 bipolar ECG channel (lead II) with a sampling frequency of 250 Hz in a fully wired system. The data was recorded continuously and contains both day and night data of the patients within a hospital room. The single-lead ECG signal was continuously unreadable due to noise during 4.6 (patient 13), 8.8 (patient 16) and 7.4 (patient 18) hours. Those segments were removed from the analysis. The remaining data consists of 24 patients with 2172 hours of data. In total, 228 seizures were recorded (see Table 1).

A clinical expert annotated the seizure onsets and offsets with the use of video-EEG data, without considering the ECG. Afterwards, a neurologist validated the annotations. The seizure duration was defined as the time between EEG seizure onset and offset. However, during 30% of the seizures, the offsets could not be determined. The ethical committee of the UZ Leuven approved the study (approval number S59662). All patients signed the informed consent for their participation in this study.

2.2 Preprocessing

Single-lead ECG was used as input for the proposed seizure detection algorithm. First, the heart rate was extracted from the ECG using an approach that used a real-time R peak detection algorithm. It detected the R peaks based on the derivative signal and an adaptive threshold T_t that changed based on the maximal derivatives P_t of the previously detected R peaks ($T_t = 0.9 * T_{t-1} + 0.1 * P_t$). Next, strong heart rate increases (HRI) caused by sympathetic activations were detected by automated slope analysis on the tachogram. HRI extraction was performed on a filtered tachogram, using a median filter with an order of 15 heart beats. If a heart rate gradient was larger than 1 bpm/s, a strong HRI was assumed. The beginning and end of this HRI was found by analyzing when this heart rate gradient became negative again. The HRI was then assumed to be a strong HRI if thresholds on the length of the HRI ($> 8s$), the achieved peak heart rate during the HRI (>60bpm) and the (percentual) increase in heart rate during the HRI were exceeded (>10bpm absolute heart rate increase, >25% percentual increase). These thresholds are based on the findings presented in De Cooman et al. (2017a).

2.3 Feature extraction

Features were extracted whenever such a strong HRI was detected. In De Cooman et al. (2017a), it was shown that four features extracted from the HRI led to optimal patient-independent results: the peak heart rate, the heart rate at the start of the HRI, the baseline heart rate (extracted from the minute before the HRI) and the standard deviation of the baseline heart rate period. As the primary goal of this study was to lead to optimal patient-specific results with a limited amount of patient-specific data, only the first two features were used in this study. The reason for this was that most of the performance of the system was already accomplished by those two features. Adding more features to the system requires more training data for robust personalization through transfer learning. Choosing these two features led to an optimal balance between performance and limited requirement of patient-specific data. These features were then classified with either the patient-independent (PI) classifier or the patient-specific (PS) transfer learning classifier.

2.4 Patient-independent classification

Let $\{x_i, y_i\}_{i=1}^{N}$ be the training data points extracted from patients different than the ones used for testing the algorithm, with $x_i \in \mathbb{R}^2$ the data samples and $y_i \in \{-1, +1\}$ the corresponding labels. Let class -1 correspond to seizure samples and class +1 to non-seizure samples. Support vector machines (SVM) will map data points to a higher dimensional space using a (non-)linear transformation $\varphi(x)$, so that the data points can be separated in this space by the hyperplane $w^T\varphi(x) + b$, with w the unknown weight vector and b an unknown constant.

The solution for weighted SVM can be found by solving the following optimization problem

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

with ξ_i the error of the model on x_i and C a tunable hyperparameter. A modification of the typical SVM is used here to remove the class imbalance from the dataset (Akbani et al., 2004). The values of c_i are defined as

$$c_i = \begin{cases} \gamma\frac{(N^+ + N^-)}{2N^-} & : y_i = -1 \\ \frac{N^+ + N^-}{2N^+} & : y_i = +1 \end{cases} . \tag{2}$$

The parameter γ gives more importance to the correct classification of seizure samples compared to non-seizure samples during classifier training. The Lagrangian of (1) becomes

$$\mathcal{L}(w,b) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} c_i\xi_i - \sum_{i=1}^{N}\nu_i\xi_i - \sum_{i=1}^{N}\alpha_i\left(y_i\left(w^T\varphi\left(x_i\right)+b\right)-1+\xi_i\right) \tag{3}$$

with $\alpha_i, \nu_i \geq 0$ the Lagrange multipliers, leading to the dual problem

$$\begin{aligned} &\min_{\alpha} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K\left(x_i, x_j\right) - \sum_{i=1}^{N}\alpha_i \\ &s.t. \quad \begin{cases} \sum_{i=1}^{N}\alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C c_i \quad ,\forall i \in [1,N] \end{cases} \end{aligned} \tag{4}$$

with $K(x_i, x_j) = \varphi(x_i)^T\varphi(x_j)$ the kernel function. The values of the hyperparameter C, γ and the Gaussian kernel parameter σ are taken as reported to be optimal in De Cooman et al. (2017a). The classifier is trained using the leave-one-patient-out crossvalidation (LOPO-CV) approach, training the classifier on all patients except the one used for evaluating the algorithm.

2.5 Personalized classification through transfer learning

The goal of this study is to personalize the seizure detection classification in order to get an optimal patient performance with a limited amount of patient-specific data. One solution for this is to use transfer learning (TL), which allows to train a new classifier for a problem with a limited amount of data by using a reference classifier that solves a similar problem. The previously trained patient-independent classifier discussed in section 2.4 is used as the reference classifier. This way, the classifier can be personalized with a limited amount of patient-specific data by using the knowledge already incorporated in the patient-independent classifier. An overview of the proposed procedure for personalizing the heart rate-based seizure detection is given in Figure 2.

The newly proposed transfer learning approach is based on the concept described in Yang et al. (2007). It states that the weight vector of the new SVM solution should be sufficiently similar compared to the weight vector of the reference SVM solution, while also minimizing the misclassification error on the data of the new problem. In this use case, it means that the weight vector of the new patient-specific and reference patient-independent classifier should be sufficiently similar. The minimization problem proposed in Yang et al. (2007) is enhanced here in order to also be able to deal with class imbalance and the increased importance of the sensitivity in seizure detection algorithms. The following minimization problem to create an SVM classifier for the patient-specific data $\{\widetilde{x}_k, \widetilde{y}_k\}_{k=1}^{M}$ is proposed

$$\begin{aligned} &\min_{\tilde{w},\tilde{b},\tilde{\xi}} \frac{1}{2}||\tilde{w}-w||^2 + D\sum_{k=1}^{M}\tilde{c_k}\tilde{\xi_k} \\ &s.t. \quad \begin{cases} \tilde{y_k}(\tilde{w}^T\varphi(\tilde{x}_k)+\tilde{b}) \geq 1-\tilde{\xi}_k \\ \tilde{\xi}_k \geq 0 \end{cases} \quad ,\forall k \in [1,M] \end{aligned} \tag{5}$$

with $\tilde{\xi}_k$ the error of the model on data point $\tilde{x}_k$ and w the weight vector obtained from the patient-independent classifier trained using (1), defined as

$$w = \sum_{i=1}^{N} \alpha_i y_i \varphi(x_i) \tag{6}$$

by the original SVM optimization problem. The same transformation function φ and corresponding kernel K used in the reference classifier are used here. This minimization problem contains weights $\tilde{c}_k$ for each $\tilde{x}_k$, which allows to also take care of the class imbalance and sensitivity importance in the optimization. These weights are defined as

$$\tilde{c_k} = \begin{cases} \widetilde{\gamma}\frac{(M^+ + M^-)}{2M^-} & : \tilde{y}_k = -1 \\ \frac{M^+ + M^-}{2M^+} & : \tilde{y}_k = +1 \end{cases}, \tag{7}$$

with M^+ and M^- indicating the number of patient-specific non-seizure and seizure data points. The introduction of these weights $\tilde{c}_k$ are crucial as they ensure that the personalized classifier is sufficiently stable and targets a sufficiently large sensitivity, which is required for real-life seizure detection systems. Hyperparameter D allows to balance between minimizing the errors for the patient-specific data points and minimizing the difference compared to the reference patient-independent classifier (defined by w). Parameter $\widetilde{\gamma}$ is experimentally initialized to a value of 1.5. The initial value for hyperparameter D depends on the amount of seizures available in the training set, and is set to 0.1 for patients with less than 10 seizures and set to 100 for patients with more than 10 seizures. These values are based on the findings reported in De Cooman et al. (2017b). Nevertheless, these parameters should ideally be optimized per patient, based on their validation performance, but it is very challenging to get robust hyperparameter optimization results due to the low amount of patient training data. Therefore, a more heuristic method is applied, which lowers the values of $\widetilde{\gamma}$ (linear decrease of 0.25) and D (exponential decrease by factor 1/10) if the resulting classifier leads to 50% more false detections on the patient training data than the available patient-independent approach. This reduction is repeated until the false detection rate (FDR) is dropped below this threshold or minimal values for $\widetilde{\gamma}$ (=0.25) and D (=0.01) are reached. This procedure is required to avoid that the personalized approach would overtrain on a limited amount of abnormally small ictal heart rate increases, which could lead to a drastic increase in FDR compared to the reference patient-independent classifier.

An optimal solution for (5) is found for the saddle point in the Lagrangian $\mathcal{L}$

$$\begin{aligned} &\max_{\tilde{\alpha},\tilde{\beta}} \min_{\tilde{w},\tilde{b},\tilde{\xi}} \mathcal{L}(\widetilde{w},\tilde{b},\widetilde{\xi};\widetilde{\alpha},\widetilde{\beta}) \\ &= \max_{\tilde{\alpha},\tilde{\beta}} \min_{\tilde{w},\tilde{b},\tilde{\xi}} \frac{1}{2}||\widetilde{w} - w||^2 + D\sum_{k=1}^{M} \widetilde{c}_k\widetilde{\xi}_k - \sum_{k=1}^{M} \widetilde{\beta}_k\widetilde{\xi}_k - \sum_{k=1}^{M} \widetilde{\alpha}_k \left(\tilde{y}_k \left(\widetilde{w}^T\varphi(\widetilde{x}_k) + \tilde{b}\right) - 1 + \widetilde{\xi}_k\right) \end{aligned} \tag{8}$$

with $\widetilde{\alpha}_k$, $\widetilde{\beta}_k \geq 0$ the Lagrangian multipliers. This leads to

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial\tilde{w}} = 0 & \rightarrow \tilde{w} = w + \sum_{k=1}^{M} \widetilde{\alpha}_k\tilde{y}_k\varphi\left(\widetilde{x}_k\right) \\ \frac{\partial\mathcal{L}}{\partial\tilde{b}} = 0 & \rightarrow \sum_{k=1}^{M} \tilde{y}_k\widetilde{\alpha}_k = 0 \\ \frac{\partial\mathcal{L}}{\partial\tilde{\xi}_k} = 0 & \rightarrow \widetilde{\alpha}_k + \widetilde{\beta}_k = D\tilde{c}_k, \forall k \in [1, M] \end{cases} \tag{9}$$

so that the dual problem of (5) is defined as

$$\min_{\widetilde{\alpha}} \sum_{i=1}^{N}\sum_{k=1}^{M} y_i\widetilde{y}_k\alpha_i\widetilde{\alpha}_k K\left(x_i,\widetilde{x}_k\right) - \sum_{k=1}^{M}\widetilde{\alpha}_k + \frac{1}{2}\sum_{k=1}^{M}\sum_{l=1}^{M}\widetilde{y}_k\widetilde{y}_l\widetilde{\alpha}_k\widetilde{\alpha}_l K\left(\widetilde{x}_k,\widetilde{x}_l\right)$$
$$s.t. \quad \begin{cases} \sum_{k=1}^{M}\widetilde{\alpha}_k\widetilde{y}_k = 0 \\ 0 \leq \widetilde{\alpha}_k \leq D\tilde{c}_k \quad ,\forall k \in [1,M] \end{cases} . \tag{10}$$

Note that

$$\widetilde{w} = w + \sum_{k=1}^{M}\tilde{\alpha}_k\tilde{y}_k\varphi(\tilde{x}_k) \tag{11}$$

indicates that the patient-specific $\tilde{w}$ is a combination of patient-independent and patient-specific information. A new data point $\tilde{x}_n$ is then classified using

$$y(\tilde{x}_n) = sign\left(\sum_{k=1}^{M}\tilde{\alpha}_k\tilde{y}_k K(\tilde{x}_k,\tilde{x}_n) + \tilde{b} + \sum_{i=1}^{N}\alpha_i y_i K(x_i,\tilde{x}_n)\right) \quad . \tag{12}$$

The TL classifier is trained and tested using a 5-fold crosstesting scheme, in which 4 folds of patient-specific data are used for training and 1 for testing. This is then repeated 5 times so that each fold is used once as test set.

2.6 Alternative automatic personalization solutions

The proposed transfer learning approach is also compared to two different alternatives for personalization. The first alternative includes a fully PS approach which is trained with only PS data using the SVM classifier defined by (1) as in De Cooman et al. (2018). The other alternative is a so-called mixed model, in which both PI and PS data are used for training an SVM classifier defined by (1), but adapting the values of c_i in (2) into c_i^{MIX}:

$$c_i^{MIX} = \begin{cases} s_i\gamma\frac{N^+ + N^- + M^+ + M^-}{2(N^- + M^-)} & : y_i = -1 \\ s_i\frac{N^+ + N^- + M^+ + M^-}{2(N^+ + M^+)} & : y_i = +1 \end{cases} \tag{13}$$

with

$$s_i = \begin{cases} 4 & : i \in PSdata \\ 1 & : i \notin PSdata \end{cases} \tag{14}$$

such that misclassification of PS data is more critical during training than misclassification of non-PS data. The value 4 is chosen as recommended in De Cooman et al. (2018).

2.7 Algorithm evaluation

In order to compare the different seizure detection algorithms, four performance metrics were used. The first two metrics correspond to the the sensitivity (Se, percentage of detected seizures) and false detection rate (FDR, expressed in false positives/hour, FP/h). A seizure is detected if a detection is generated between 1 minute prior to the seizure onset and 2 minutes after the seizure onset. All other detection were classified as false detections. False detections within 1 min of each other were counted as one false detection. In order to combine the Se and FDR in one metric, the F_β-score with $\beta = 3$ is defined as:

$$F_\beta = \frac{(1+\beta^2)TP}{(1+\beta^2)TP + \beta^2 FN + FP} \tag{15}$$

with TP, FN and FP the number of true positives (detected seizures), false negatives (missed seizures) and false positives (false detections). The F_3-score is chosen for this application because it gives more importance to Se compared to FDR. As last metric, the detection delay was determined, which indicates the time difference between the moment of detection and the EEG seizure onset. Average measures over the entire dataset can be expressed as patient average performance (Pat.-Av.), which is the average of the performance of each patient, or overall average (Tot.-Av.), computed on the total number of seizures or recording duration. The first average measure is used, unless specifically mentioned. To prove the significant differences between the algorithms, paired t-tests were performed. All results were obtained retrospectively in a simulation which replicated a real-time setting.

2.8 Impact of number of seizures in training

One of the advantages of transfer learning is that it allows to train a new classifier with a relative limited amount of data by using a reference classifier (Pan and Yang, 2009). In this simulation, the number of required training seizures needed to gain sufficient added value in seizure detection performance was evaluated. Instead of using the full training set, only a few seizures (0-4) from the patient-specific training set were used for training (using the crosstesting scheme described in section 2.5). A model trained with zero seizures is equivalent to a patient-independent model. These training seizures were chosen randomly in 100 simulations for each selected amount of training seizures. This simulation was done for the proposed transfer learning approach and the alternatives for personalization.

2.9 Comparison with the literature

In order to compare the proposed algorithms against the state-of-the-art literature, three algorithms (Jeppesen et al., 2015; Osorio, 2014; van Elmpt et al., 2006) were also implemented and evaluated on the same dataset. The algorithms were implemented based on the corresponding publications. Both Osorio (2014) and van Elmpt et al. (2006) were based on an algorithm that requires two moving windows: one short window indicating the current heart rate behavior and one longer window indicating the reference/baseline heart rate. Both approaches were described to be patient-independent algorithms, but no preferred threshold values were described in the papers. Therefore, optimal patient-independent threshold values leading to the highest F_3-score were automatically chosen using a LOPO-CV procedure on each training set. The algorithm described in Jeppesen et al. (2015) was originally designed to be a patient-specific algorithm (based on non-seizure patient data), but also a patient-independent variation of the algorithm is evaluated here. The used threshold on the parameter value is also optimized in a LOPO-CV procedure.

Also new patient-specific versions of these algorithms from the literature were constructed in an automated fashion. This is done using 5-fold crosstraining, where the threshold values are automatically chosen based on the optimal F_3-score found in the 4 folds of training. Also the test mentioned in Section 2.8 is performed for the different state-of-the-art algorithms to compare how fast these can adapt to a patient's characteristics.

3 RESULTS

The preprocessing procedure discussed in Section 2.2 identified 84.7% of the seizures, which gives an indication of the amount of seizures with ictal heart rate increases. Table 2 gives an overview of the results of the patient-independent (PI), fully patient-specific (PS), mixed (MIX) and transfer learning (TL) approaches. The PI algorithm results in an average Se of 75% with 3.0 FP/hour and an F_3-score of 0.22. By adapting the model to the patient characteristics with the TL algorithm, a similar Se is observed (71%)

with 37% less false positives (1.9 FP/hour). The average F_3-score is increased to 0.30. Figure 3 shows the results of the proposed TL approach and the reference PI approach for each patient. It illustrates that personalization allows to strongly increase the performance for most patients.

The alternative mixed approach results in a similar Se, but with on average 0.7 FP/h more than the proposed TL approach. The fully PS approach results in a decreased Se, with a slightly increased FDR (2.2 FP/h) compared to the TL approach. The TL approach did not only result in a decreased average FDR, but also decreased FDR variability over the different patients compared to the personalization alternatives (see Figure 4).

By performing a two-sided paired t-test, the sensitivity of the PI and MIX algorithm were found to be not statistically different from the TL algorithm (p=0.29 (PI vs. TL) and p=0.82 (MIX vs. TL)). However, the FDR of the PI and MIX algorithm are different compared to the TL algorithm ($p<0.001$ (PI vs. TL), p=0.03 (MIX vs. TL)). The fully PS approach has a lower Se than the TL algorithm (p=0.01), whereas the FDRs are not significantly different (p=0.29). The F_3-score of the PI, fully PS and MIX algorithm are lower than the TL algorithm ($p<0.001$ (PI vs TL), p=0.01 (PS vs TL), $p<0.01$ (MIX vs TL)). This shows that the proposed transfer learning method is indeed statistically better than the other evaluated approaches for personalization. The average detection delay for the proposed transfer learning approach was 21 s.

Table 2 also shows the results of both the patient-independent and personalized versions of the algorithms from the literature. The algorithms from van Elmpt et al. (2006); Jeppesen et al. (2015) result in clearly lower sensitivity with a comparative FDR as the proposed patient-independent approach. The algorithm from Osorio (2014) leads to nearly the same average result as the proposed patient-independent approach, with a slightly lower sensitivity, FDR and F_3-score. Both algorithms from Osorio (2014); Jeppesen et al. (2015) show an increase in performance by personalizing the thresholds, but result in a lower performance than the proposed TL approach. The personalized version of van Elmpt et al. (2006) leads to a strongly increased sensitivity, but also the FDR increases.

As explained in Section 2.8, the influence of the number of training seizures on the TL algorithm performance was investigated. Figure 5A shows the impact on the F_3-score performance in function of the number of seizures available during training. The F_3-score for the TL approach strongly improves compared to the PI performance if only one patient-specific seizure is available in the training set. When two seizures are available, the variation between the results decreases, while the average performance increases. The performance further increases by including additional seizures to the training set, while still maintaining a similar variation in results. Figure 5A also shows the effect of the number of seizures in the training set for the other personalization methods, which results in clearly lower performances for each number of available seizures.

The same simulation results for the state-of-the-art algorithms are shown in Figure 5B. The algorithms from Osorio (2014); van Elmpt et al. (2006) lead to a decent increase with one available patient-specific seizure, but only increases in accuracy slowly by adding more seizures. This increase is stronger for Jeppesen et al. (2015) by adding more seizures, although the results with four available seizures are still worse than the proposed patient-independent algorithm. Although the algorithm from Osorio (2014) had a similar patient-independent F_3-score, the proposed personalized TL approach outperforms the personalized version of Osorio (2014).

4 DISCUSSION

4.1 Performance comparison of the PI and TL approach

Table 2 and Figure 4 show that the mean and median sensitivity of the PI and TL approaches are similar, whereas the FDR decreases and the F_3-score increases. By looking at the patients individually (Figure 3), it can be observed that the TL approach clearly reduced the FDR for some patients without decreasing the sensitivity (e.g. patients 2, 5 and 20). The transfer learning model adapts well to the patient-specific heart rate characteristics. In these patients, the ictal heart rate changes are often very stereotypical, showing little intra-patient variability, which leads to a strong decrease in FDR. Figure 6 illustrates the impact of the number of seizures on the performance for some patients. For patient 5, the personalized approach already gets most performance increase by only including 1 patient-specific seizure. There is also limited variability in the results of the different simulations, showing that the model is accurate and robust for this patient, and little intra-patient variability is found in the ictal HRIs. For patient 2, also a fast personalization can be obtained, but more variability is found between the results of the simulations. This is due to a larger intra-patient ictal variability, so that the selection of seizures in training has a bigger impact on the performance. This impact however reduces if more patient-specific seizures are added to the model, leading to a more robust personalized model. Patient 15 had a lot of seizures during the recording, but these seizures showed a large intra-patient variability in terms of ictal HRIs. This intra-patient variation was mostly caused by the variation of seizure duration, which were typically to short to be strongly differentiable from non-epileptic HRIs. The ictal HRIs are also harder to differentiate from non-ictal HRIs compared to patients 2 and 5, leading to a slower learning curve than those from patients 2 and 5. In general, the steepness of the learning curve and amount of seizures before convergence depend on both ictal and inter-ictal heart rate behavior of each patient.

For some patients, however, the sensitivity dropped slightly. This is due to the fact that the model adapts to the patient characteristics. However, the heart rate characteristics during some seizures were atypical (different from other seizures) for that patient. Less severe seizures were typically accompanied with smaller heart rate increases. Those seizures were sometimes not detected with the TL model. An example of a smaller ictal heart rate increase compared to normal ictal heart rate behavior for a particular patient is illustrated in Figure 7. The proposed TL approach adapts to the majority of seizures, and therefore might lead to a missed detection of these atypical seizures. However, this small decrease in sensitivity is often accompanied with a strong decrease in FDR for those patients. The reverse also occurred in some patients, where borderline seizures that were missed by the PI approach are detected by the proposed TL approach (e.g. patient 24). The results above show that, although personalization in general allows to improve the performance, it still has difficulty to counter unpredictable intra-patient variability in seizure behavior. If five seizures within a patient are stereotypical and used for training, an atypical sixth seizure will not be detected accurately using supervised personalization. This is however typically not the case for stronger seizure type (e.g. focal to bilateral seizures), as they are typically easier to detect with heart rate-based seizure detection, even if the algorithm is not specifically trained for this seizure type.

This increase in performance was, however, not achieved for all patients, for example in patient 1. For those patients, rather atypical ictal HR increases are observed, or the HRIs are too weak in magnitude, making it difficult to differentiate them from non-epileptic heart rate activity. If the seizure activity is too similar to non-seizure activity for that patient, the model cannot be improved by means of personalization. Other ECG-based features or features from other modalities might have to be included then in order to achieve better performance for these patients.

Only 84.7% of seizures had ictal HRIs in the analyzed dataset, which is a similar percentage as in the literature (Zijlmans et al., 2002; Jansen and Lagae, 2010; Leutmezer et al., 2003). Personalizing the algorithm will not help to detect these seizures without ictal heart rate changes. Also false detections or missed seizures caused by too strong ECG noise cannot be avoided by the personalization as they can occur both ictally and inter-ictally (Vandecasteele et al., 2017). Other approaches should be used to further improve the performance, such as improved noise removal techniques.

An example of the obtained classifier boundaries for patient 11 are illustrated in Figure 8. The seizure and non-seizure data points are shown together with the PI, PS and TL boundary for two different values of hyperparameter D (0.01 and 1), with a fixed $\widetilde{\gamma}$ value of 1.5. The PI boundary gives a good general indication, but lacks adaptation to the patient characteristics. The TL approach adapts to these characteristics with a limited amount of available patient-specific data. By choosing a low value for D (i.e. 0.01) the TL model will be similar to the PI model because a relative low weight is given to the errors obtained for the patient-specific data compared to the first term in (5), which quantifies how different the new model is to the original one. With a higher value of D the model adapts more to the patient-specific data (and the corresponding errors $\widetilde{\xi}_k$) and shows less similarity to the PI boundary. The PS boundary is less optimal, compared to the TL solution. It can be seen from the TL boundary that it still contains information gathered in the PI classifier (especially for low values of D), indicating the added value of this approach. This way, the FDR is strongly decreased, without affecting the sensitivity.

4.2 Comparison of alternative personalization solutions

Different alternatives for personalization were implemented and evaluated in order to compare with the proposed transfer learning approach (see Table 2 and Figure 4). The naive approach only uses patient-specific data points for training a normal SVM classifier. Despite all patients had at least 5 seizures, this was still an insufficient amount of data for most patients in order to get a robust patient-specific classifier. In Cogan et al. (2016), it was mentioned that at least 6-8 seizures were required in order to make a personal algorithm, and the reason for this was to better include the inter-seizure variability of autonomic changes within a patient. Due to the relative low amount of patient-specific data, it often occurred that the classifier was overtrained on a limited amount of seizure data, not taking into account potential fluctuations between different seizures from a patient. This led to a strong decrease in sensitivity, although the FDR was not so much higher than the proposed TL approach. The TL approach is able to better take this inter-seizure variation into account by holding on to the knowledge described by the reference PI classifier.

The mixed model (MIX) uses a mixture of patient-specific data with data from other patients in the training set of a standard SVM training procedure. It produced more robust results than the fully patient-specific approach without strong negative outliers. However, it generated, on average, around 0.5 FP/h more than the proposed TL approach. Applying transfer learning to the reference classifier allows to better take over the information of the PI classifier and translate it to the patient-specific model, whereas the mixed approach would try to create a new model without this prior model knowledge. The proposed transfer learning method not only leads to a better performance, but is also trained faster with an optimization problem which contains less data to analyze.

4.3 Impact of number of seizures on personalization

Transfer learning allows to train a personalized classifier for heart rate-based seizure detection. It is however important to have an idea of how much data is actually needed for this. In seizure detection, the amount of seizures is often the restricting factor as some patients have a low seizure frequency. In Section 2.8, a simulation study was described to evaluate the impact of the number of seizures in training on the personalized performance (see Figure 5).

The proposed TL approach already leads to a strong increase in F_3 performance by only including 1 seizure in the training set. This shows that with only 1 seizure, the algorithm can already be personalized. There is a lot of variation in the results, which is caused by the heterogeneity of the seizures in the training set. There is also a lot of variation in ictal heart rate changes between different seizures within a patient, so if an atypical seizure is used in training, this will lead to suboptimal results for that patient. By using 2 seizures during training, the F_3 performance increases slightly, and also the variability between results of different simulations decreases. However, from 3 seizures onwards, the variability is no longer strongly decreased with the proposed method. The average performance increases further up until 4 available seizures, and is expected to increase further if more seizures are included in training (Cogan et al., 2016). From a certain amount of seizures, it is however expected that the fully PS approach would lead to a better performance than the proposed TL method. Nevertheless, it is currently not known how many seizures are required for this with the proposed procedure.

Figure 5 also shows the effect of the number of seizures on the alternative personalization approaches. It can be seen that by only having 1 patient-specific seizure in the training data, the TL performance strongly increased, whereas this increase is less evident for the other approaches. For the PS approach, the median performance is increased, but a large portion of the results were actually worse than for the reference PI approach. This is due to the fact that the PS method is often strongly overfitting on data from 1 seizure, which is not a robust way for training a classifier. The average F_3 scores for the fully PS approach increase by adding more seizures, and the variation on performance decreases. The results however remain lower than those from the proposed TL approach when using 4 seizures. The mixed model is more robust than the fully PS approach for a limited amount of seizures, but it slightly loses its added value when more seizures are added to the training set.

4.4 Comparison to the literature

Different algorithms from the literature were also implemented both patient-independently and patient-specifically and evaluated on the dataset described in Section 2.1. Table 2 shows that the patient-independent version of the proposed algorithm clearly outperforms the algorithms from van Elmpt et al. (2006); Jeppesen et al. (2015), but has a similar result as Osorio (2014). A simplified patient-independent algorithm from De Cooman et al. (2017a) is used here as described in section 2.3, which has shown to outperform the literature. Due to the simplification, the added value of this algorithm over Osorio (2014) is reduced, but its added value compared to van Elmpt et al. (2006) and Jeppesen et al. (2015) remained similar.

However, when patient-specific alternatives for these three algorithms were made using the automatic procedure described in Section 2.9, the proposed personalization procedure clearly outperforms all these algorithms (see Figure 5B and Table 2). The algorithm from Osorio (2014) shows a big increase by personalizing the algorithm, even for only one available seizure, but the added value is much smaller compared to the proposed transfer learning approach. Smaller increase in performance is found in the algorithm of van Elmpt et al. (2006), with only a very slow learning curve with increasing number of annotated seizures per patient. A much steeper learning curve is found for the algorithm from Jeppesen et al. (2015). Although the algorithm has the lowest patient-independent performance, the performance increases very fast with increasing number of seizures. The algorithm was originally meant for patient-specific evaluation. However, even with four seizures available per patient, it performs worse than the proposed personalized transfer learning method. The proposed personalization method using transfer learning thus not only performs more accurately in general, it also allows a much faster training if only a limited amount of patient-specific seizures are available. This is crucial in practice as some patients have a very low seizure

frequency, and thus the personalized algorithm can reach much faster a desired level of accuracy than state-of-the-art algorithms.

4.5 Limitations of the study

There are however some limitations to the performed study that have to be taken into account. First of all, the data is recorded in the hospital, where the patients were restricted to move within their room. This leads to a limited activity of the patient, which can lead to an underestimation of the amount of false detections in practice. However, it is compared to state-of-the-art algorithms from the literature on the same dataset, and has proven to outperform these on this dataset. However, currently no study in the literature has shown results of such heart rate-based algorithms for full day-and-night monitoring in a real home environment.

Furthermore, during the presurgical evaluation, drug treatment can be altered or completely removed. This can influence the results in two ways. First, certain drugs can alter the heart rate variability of the patient, which might lead to different (stronger) ictal and inter-ictal heart rate changes during the presurgical evaluation compared to the home situation. In some patients, indeed small changes in heart rate variability could be found during the first and last day of monitoring. However, the ictal heart rate changes between the first and last day of monitoring were found to be limited compared to the variability that has to be taken into account for the circadian fluctuation of the heart rate features. Therefore, this did not lead to extra missed seizures or a large increase in false detection rate in this study. A second influence is the fact that stronger seizures can occur during the presurgical evaluation compared to the typical home situation. However, only a couple of focal to bilateral seizures were found in our dataset, and the large majority of seizures were perceived to be typical seizures for that patient in their home environment. Other drugs not applied in this study could however have stronger influence on the heart rate variability and the ictal changes. This should be further investigated in future work, as it could influence the usability of data from the presurgical evaluation as training data for such algorithms when used in a home environment.

No dedicated wearable ECG derived device is used in this study. Currently, a large portion of false detections and a small percentage of missed seizures is caused due to poor ECG quality, largely due to the wiring. A previous study has shown that better performance can be obtained by using a wearable ECG device rather than the standard wired hospital ECG (Vandecasteele et al., 2017). This would however be the case for all evaluated algorithms in this study, including those from the literature tested in this study on the discussed dataset.

4.6 General discussion and future work

The proposed method allows a fully automated personalization of heart rate-based seizure detection. In literature, personalization was often reached after adjusting thresholds manually after visual inspection of the data (van Elmpt et al., 2006; Boon et al., 2015). This is however a very costly and non-scalable solution. In Jeppesen et al. (2015), the personalization was done automatically by adjusting the threshold per patient based on data from a non-epileptic segment of 30 minutes. Although this is a more scalable solution, only 30 minutes of non-seizure data does not contain sufficient information to grasp the full complexity of the heart rate variability of that patient (e.g. the circadian rhythm). The proposed solution does take into account much more complexity in a fully automated way, which allows a better implementation in practice.

Despite the increased performance caused by personalization through transfer learning, still a too high FDR is obtained for usage in practice. Personalization allows to solve some of the issues leading to a too high FDR, but is not able to solve all issues related to heart rate-based seizure detection. Some seizures do not contain ictal heart rate changes, and still a large portion of non-epileptic heart rate changes cannot be differentiated from epileptic HRIs with the current techniques. The proposed method should, however, be used as part of a multimodal algorithm, where it is combined with another modality. Such multimodal

combination has shown to lead to FDR decreases with factors 5 to 10 compared to unimodal performance (De Cooman et al., 2018). Similar to the seizure detection algorithms, also the used modalities should be chosen for each individual patient based on its typical ictal changes and seizure type. Accelerometers and EMG sensors could lead to an increased performance for the detection of motor seizures (De Cooman et al., 2018; Van Andel et al., 2017). For the detection of non-motor focal seizures, behind-the-ear EEG could be used (Gu et al., 2017). The advantage of heart rate-based seizure detection over other modalities is that ECG is often monitored as well during video-EEG monitoring in the hospital, which allows to get accurately annotated heart rate data to personalize the algorithm. It is also ictally the most activated modality (apart from full EEG), so it is ideal to increase the detection performance of a wide range of seizure types. If a late integration approach is used for combining information from different modalities, the proposed personalized method can be easily integrated and further improve the multimodal performance with a similar accuracy increase as in an unimodal setting.

The proposed transfer learning approach is a supervised approach, which means that annotated data was required. In practice, these annotations can be made in the hospital during, for example, presurgical evaluation, but they could also be made by the patient or their caregivers/family. Extra procedures should then be added to avoid a too big impact of incorrectly annotated data as patients are not always aware about whether they actually had a seizure or not (Fisher et al., 2012; De Cooman et al., 2018). Ideally, an unsupervised approach could be used (De Cooman et al., 2018, 2017b; Jeppesen et al., 2015), which indicates that epileptic heart rate activity can be seen as an outlier to normal heart rate activity. This is however only the case during the night (Delamont and Walker, 2011) or in certain severe seizure types, which makes this approach only sufficiently successful for nocturnal monitoring of severe seizures. Supervised approaches are thus still required for personalizing full day monitoring applications. Due to the supervised approach, still at least one patient-specific seizure is required in order to adapt to the patient characteristics. In patients who have a very low seizure frequency, this might still be a problem for a fast personalization. For these patients, it is then advised to have a large pool of patients, and seizures from patients with similar HRV parameters as the test patient. Then, these seizures could be added to the patient-specific data in order to be able to apply this method.

5 CONCLUSION

Transfer learning allows to personalize heart rate-based seizure detection in a fast and robust way by using only a limited amount of annotated patient-specific data. The false detection rate dropped by 37% compared to the patient-independent approach while maintaining a similar sensitivity. The novel automated personalization approach proposed in this study outperforms the state-of-the-art patient-independent algorithms while also being less prone to overfitting than manual state-of-the-art patient-specific approaches. The proposed method can be used as part of a multimodal algorithm in order to increase the performance and make real-time epileptic seizure warning systems clinically feasible.

AUTHOR CONTRIBUTIONS

TDC, KV, CV, BH, EC, WVP and SVH designed the experiments and wrote and revised the manuscript. WVP and EC collected the data for the experiments. TDC and KV carried out all experiments as shared first authors.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the International Conference on Harmonization guidelines on Good Clinical Practice with written informed consent from all subjects. All

subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Medical Ethics Committee of the University Hospitals KU Leuven.

FUNDING

Bijzonder Onderzoeksfonds KU Leuven (BOF): Prevalentie van epilepsie en slaapstoornissen in de ziekte van Alzheimer (C24/18/097). Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (FWO): Hercules Foundation (AKUL 043) 'Flanders BCI Lab - High-End, Modular EEG Equipment for Brain Computer Interfacing'. Agentschap Innoveren en Ondernemen (VLAIO): 150466: OSA+, O&O HBC 2016 0184 eWatch. imec funds 2017. Flemish Government: Flanders AI Research Program. EIT: 19263 – SeizeIT2: Discreter Personalized Epileptic Seizure Detection Device. This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

DATA AVAILABILITY STATEMENT

The datasets for this study will not be made publicly available because the patients and healthy participants in the study have not given their approval for the data to be made public.

REFERENCES

Akbani, R., Kwek, S., and Japkowicz, N. (2004). Applying support vector machines to imbalanced datasets. In *Machine Learning: ECML 2004* (Springer). 39–50

Boon, P., Vonck, K., van Rijckevorsel, K., El Tahry, R., Elger, C. E., Mullatti, N., et al. (2015). A prospective, multicenter study of cardiac-based seizure detection to activate vagus nerve stimulation. *Seizure* 32, 52–61

Cogan, D., Heydarzadeh, M., and Nourani, M. (2016). Personalization of noneeg-based seizure detection systems. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the* (IEEE), 6349–6352

De Cooman, T., Kjær, T. W., Van Huffel, S., and Sorensen, H. B. (2018). Adaptive heart rate-based epileptic seizure detection using real-time user feedback. *Physiological measurement* 39, 014005

De Cooman, T., Varon, C., Hunyadi, B., Van Paesschen, W., Lagae, L., and Van Huffel, S. (2017a). Online automated seizure detection in temporal lobe epilepsy patients using single-lead ecg. *International journal of neural systems* 27, 1750022

De Cooman, T., Varon, C., Van de Vel, A., Ceulemans, B., Lagae, L., and Van Huffel, S. (2017b). Semi-supervised one-class transfer learning for heart rate based epileptic seizure detection. In *Proceedings of the 44th Computing in Cardiology Conference (CinC 2017)*

De Cooman, T., Varon, C., Van de Vel, A., Ceulemans, B., Lagae, L., and Van Huffel, S. (2018). Comparison and combination of electrocardiogram, electromyogram and accelerometry for tonic-clonic seizure detection in children. In *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*. 438–441. doi:10.1109/BHI.2018.8333462

De Cooman, T., Varon, C., Van de Vel, A., Jansen, K., Ceulemans, B., Lagae, L., et al. (2018). Adaptive nocturnal seizure detection using heart rate and low-complexity novelty detection. *Seizure* 59, 48–53

Delamont, R. S. and Walker, M. C. (2011). Pre-ictal autonomic changes. *Epilepsy research* 97, 267–272

Fisher, R. S., Blum, D. E., DiVentura, B., Vannest, J., Hixson, J. D., Moss, R., et al. (2012). Seizure diaries for clinical research and practice: limitations and future prospects. *Epilepsy Behav.* 24, 304–310. doi:10.1016/j.yebeh.2012.04.128

Forsgren, L., Beghi, E., Oun, A., and Sillanpää, M. (2005). The epidemiology of epilepsy in europe–a systematic review. *European Journal of Neurology* 12, 245–253

French, J. A. (2007). Refractory epilepsy: Clinical overview. *Epilepsia* 48, 3–7. doi:10.1111/j.1528-1167.2007.00992.x

Gu, Y., Cleeren, E., Dan, J., Claes, K., Van Paesschen, W., Van Huffel, S., et al. (2017). Comparison between scalp eeg and behind-the-ear eeg for development of a wearable seizure detection system for patients with focal epilepsy. *Sensors* 18, 29

Jansen, K. and Lagae, L. (2010). Cardiac changes in epilepsy. *Seizure* 19, 455–460. doi:10.1016/j.seizure.2010.07.008

Jeppesen, J., Beniczky, S., Johansen, P., Sidenius, P., and Fuglsang-Frederiksen, A. (2015). Detection of epileptic seizures with a modified heart rate variability algorithm based on lorenz plot. *Seizure* 24, 1–7

Leijten, F. S., Consortium, D. T., van Andel, J., Ungureanu, C., Arends, J., Tan, F., et al. (2018). Multimodal seizure detection: A review. *Epilepsia* 59, 42–47

Leutmezer, F., Schernthaner, C., Lurger, S., Pötzelberger, K., and Baumgartner, C. (2003). Electrocardiographic changes at the onset of epileptic seizures. *Epilepsia* 44, 348–354. doi:10.1046/j.1528-1157.2003.34702.x

Massé, F., van Bussel, M., Serteyn, A., Arends, J., and Penders, J. (2013). Miniaturized wireless ECG monitor for real-time detection of epileptic seizures. *ACM Trans. Embed. Comput. Syst.* 12, 1. doi:10.1145/2485984.2485990

Osorio, I. (2014). Automated seizure detection using ekg. *International Journal of Neural Systems* 24, 1450001. doi:10.1142/S0129065714500014. PMID: 24475899

Pan, S. J. and Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 1345–1359

Schulze-Bonhage, A., Sales, F., Wagner, K., Teotonio, R., Carius, A., Schelle, A., et al. (2010). Views of patients with epilepsy on seizure prediction devices. *Epilepsy & behavior* 18, 388–396

Ungureanu, C., Bui, V., Roosmalen, W., Aarts, R., Arends, J., Verhoeven, R., et al. (2014). A wearable monitoring system for nocturnal epileptic seizures. *Med. Inf. Commun. Technol. 8th Int. Symp.* , 3–7

Van Andel, J., Ungureanu, C., Arends, J., Tan, F., Van Dijk, J., Petkov, G., et al. (2017). Multimodal, automated detection of nocturnal motor seizures at home: Is a reliable seizure detector feasible? *Epilepsia Open* 2, 424–431. doi:10.1002/epi4.12076

Van de Vel, A., Cuppens, K., Bonroy, B., Milosevic, M., Jansen, K., Van Huffel, S., et al. (2016). Non-eeg seizure detection systems and potential sudep prevention: State of the art: Review and update. *Seizure* 41, 141–153

van Elmpt, W., Nijsen, T., Griep, P., and Arends, J. B. A. M. (2006). A model of heart rate changes to detect seizures in severe epilepsy. *Seizure* doi:10.1016/j.seizure.2006.03.005

van Westrhenen, A., De Cooman, T., Lazeron, R. H. C., Van Huffel, S., and Thijs, R. D. (2018). Ictal autonomic changes as a tool for seizure detection: a systematic review. *Clinical Autonomic Research* doi:10.1007/s10286-018-0568-1

Vandecasteele, K., De Cooman, T., Gu, Y., Cleeren, E., Claes, K., Van Paesschen, W., et al. (2017). Automated epileptic seizure detection based on wearable ecg and ppg in a hospital environment. *Sensors* 17, 2338

Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *Journal of Big Data* 3, 9. doi:10.1186/s40537-016-0043-6

Yang, J., Yan, R., and Hauptmann, A. G. (2007). Adapting svm classifiers to data with shifted distributions. In *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)* (IEEE), 69–76

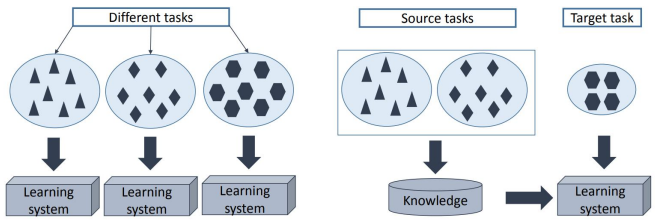
Zijlmans, M., Flanagan, D., and Gotman, J. (2002). Heart rate changes and ECG abnormalities during epileptic seizures: Prevalence and definition of an objective clinical sign. *Epilepsia* 43, 847–854.

doi:10.1046/j.1528-1157.2002.37801.x

FIGURE CAPTIONS

Table 1. An overview of the dataset (RD = Recording Duration, SD = Seizure Duration, L: Left, R: Right, B = Bihemispheric, UC: Unclear, T = Temporal, O:Occipital, F: Frontal)

Patient	# Seizures	RD (h)	Hemisphere	Origin	Age	Gender	Mean SD (s)	Range SD (s)
1	10	26	B	T	49	M	31	[24 39]
2	9	63	L	F-T	41	F	13	[13 13]
3	13	71	R	T	27	M	71	[28 96]
4	10	25	B	T	18	M	19	[14 26]
5	11	47	R	T	29	F	50	[40 60]
6	7	148	L	T	26	M	63	[32 116]
7	30	67	R	F-T	19	M	50	[17 90]
8	11	114	L	T	38	M	39	[19 75]
9	8	64	L	T (7), O (1)	28	M	23	[11 31]
10	6	111	EEG not readable		35	M	126	[69 183]
11	6	100	B (5), R (1)	T	67	F	26	[21 31]
12	9	91	R	F-T	24	F	47	[33 85]
13	8	109	R	T	32	M	46	[25 61]
14	5	100	L	T	19	F	56	[29 83]
15	13	110	EEG not readable		49	M	16	[8 30]
16	5	96	L	T	45	M	64	[11 95]
17	7	102	UC (4), R (3)	UC (4), T (3)	18	F	33	[33 33]
18	5	84	L (3), R (2)	T	62	M	125	[89 187]
19	15	113	UC		40	F	23	[6 52]
20	5	113	L	T	41	F	74	[58 83]
21	8	103	L (3), R (5)	T (3), F-T (5)	43	M	67	[29 99]
22	12	115	L (4), R (3), B (5)	T	35	M	27	[17 38]
23	6	101	R	T	23	F	80	[51 108]
24	8	99	R	T (1), O-T (5)	24	F	42	[17 84]
Total	227	2172			[18-67]	14M/10F	50 ± 30	[6 187]



(a) Traditional machine learning (b) Transfer learning

Figure 1. Illustration of the learning procedure using (a) traditional machine learning and (b) transfer learning.

Table 2. Results for the patient-independent (PI), fully patient-specific (PS), mixed (Mix) and transfer learning (TL) approach. Both patient average (P-Av) ± standard deviation and overall average (T-Av) are shown in the table.

	Se (%)		FDR (FP/h)		F_3-score	
	P-Av	T-Av	P-Av	T-Av	P-Av	T-Av
Proposed method						
PI	$\mathbf{75 \pm 22}$	**76**	3.0 ± 1.3	3.0	0.22 ± 0.14	0.20
PS	58 ± 27	59	2.2 ± 1.7	2.3	0.24 ± 0.20	0.19
MIX	72 ± 24	74	2.6 ± 1.5	2.5	0.25 ± 0.17	0.22
TL	71 ± 27	73	$\mathbf{1.9 \pm 1.1}$	**2.0**	$\mathbf{0.30 \pm 0.22}$	**0.26**
Jepessen (Jeppesen et al., 2015)						
PI	51 ± 32	52	2.8 ± 1.6	3.0	0.16 ± 0.17	0.13
PS	46 ± 23	49	2.0 ± 1.5	1.9	0.23 ± 0.21	0.18
Osorio (Osorio, 2014)						
PI	71 ± 25	72	2.8 ± 1.2	3.0	0.22 ± 0.15	0.19
PS	75 ± 23	76	2.3 ± 0.8	2.3	0.27 ± 0.18	0.24
Vanelmpt (van Elmpt et al., 2006)						
PI	59 ± 35	61	3.4 ± 2.4	3.5	0.17 ± 0.15	0.14
PS	76 ± 20	77	4.9 ± 3.4	5.2	0.19 ± 0.17	0.13

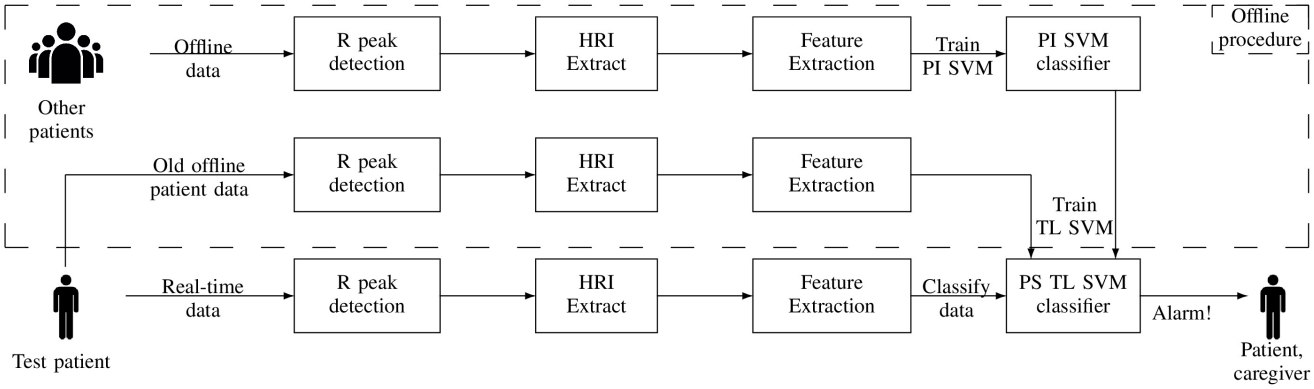


Figure 2. Overview of the proposed transfer learning approach for personalized heart rate-based seizure detection. HRI: heart rate increase; PI SVM: patient-independent support vector machine; PS TL SVM: patient-specific transfer learning support vector machine.

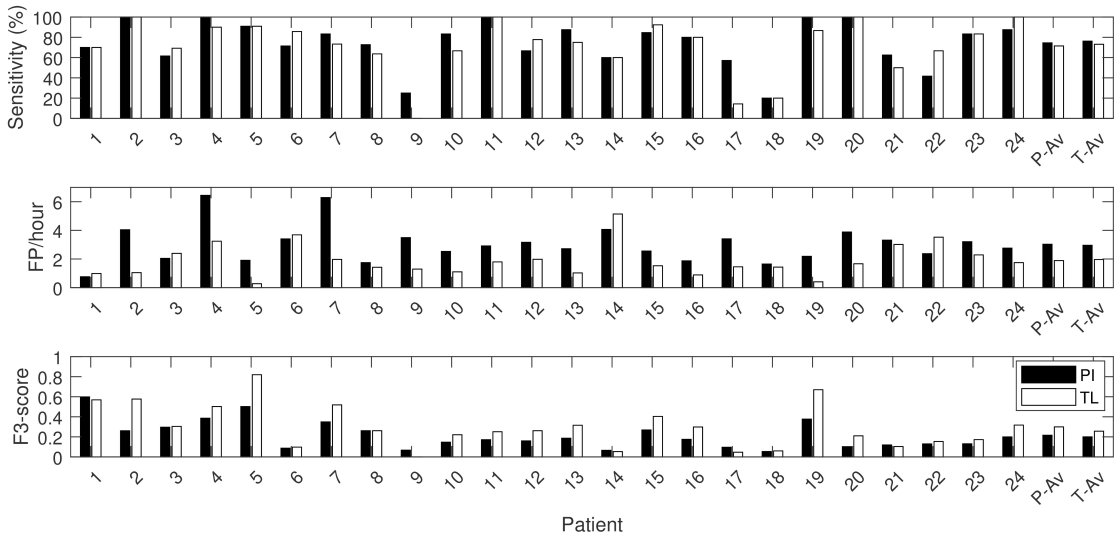


Figure 3. Sensitivity, FDR and F_3-score per patient with patient average (P-Av) and overall average (T-Av) performances for the patient-independent (PI) and transfer learning (TL) algorithm.

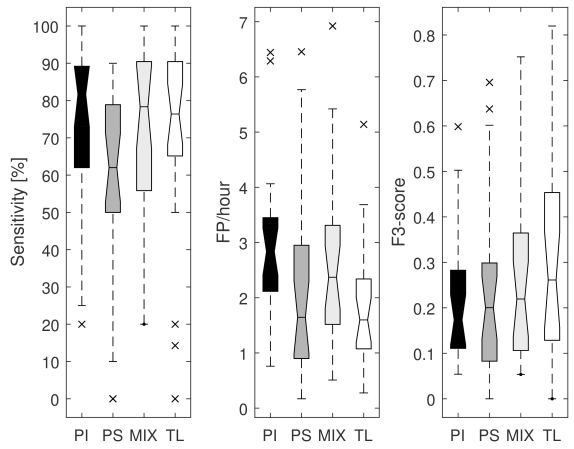


Figure 4. Boxplots of sensitivity, FDR and F_3-score for the patient-independent (PI), fully patient-specific (PS), mixed (MIX) and transfer learning (TL) algorithm.

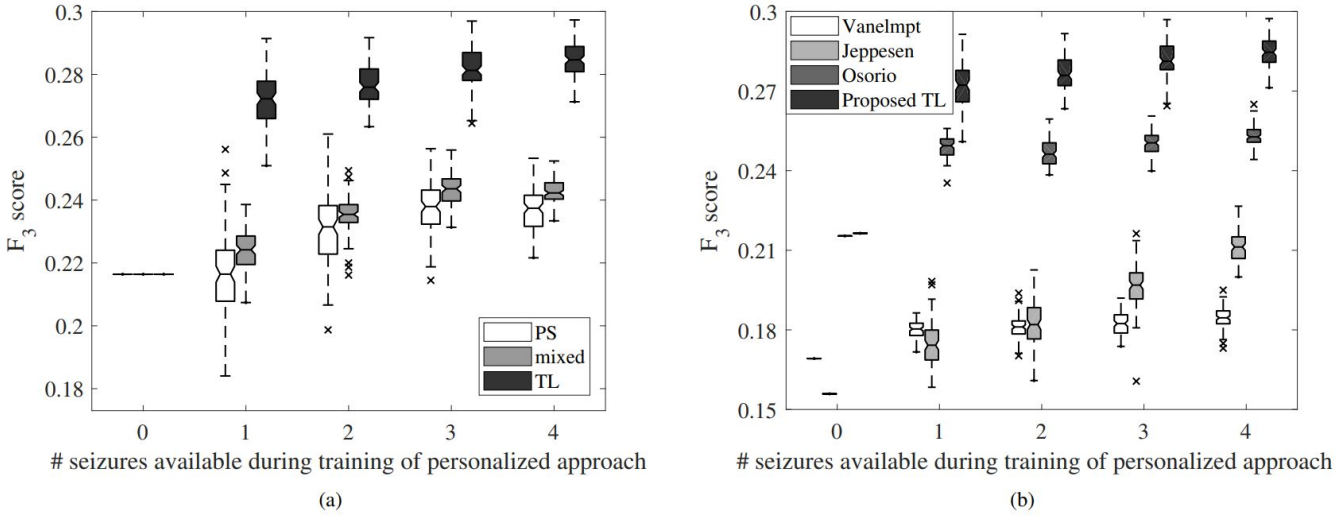


Figure 5. Impact of the number of seizures on the F_3-score performance for different alternatives. (a) Effect for the different proposed personalization approaches, including the full patient-specific (PS) approach, mixed approach (mixed) and transfer learning (TL) approach. In case 0 seizures are available for training, the performance of the patient-independent algorithm is shown. (b) Effect for the different algorithms implemented from the literature, compared to the proposed transfer learning approach.

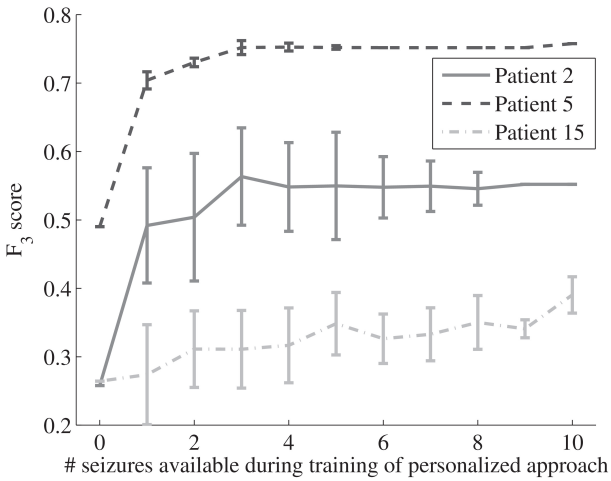


Figure 6. Impact of the number of seizures on the average F_3-score performance (including standard deviations for the performed simulation test) for the proposed personalized transfer learning approach for patients 2, 5 and 15.

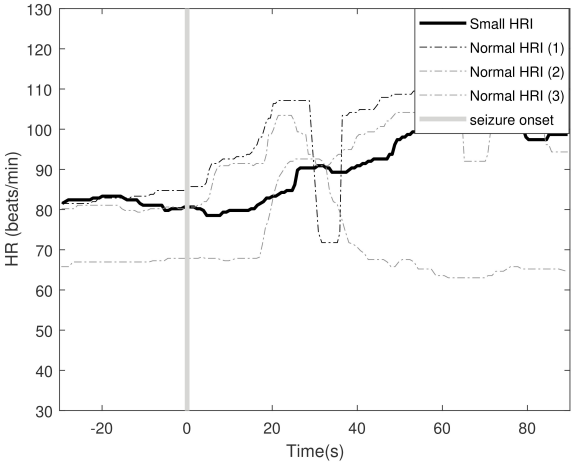


Figure 7. Example of an ictal small heart rate increase (HRI) and 3 normal ictal heart rate increases for patient 8. The vertical line indicates the seizure onset of the different seizures. The shown tachogram signals are filtered using the median filter discussed in section 2.2.

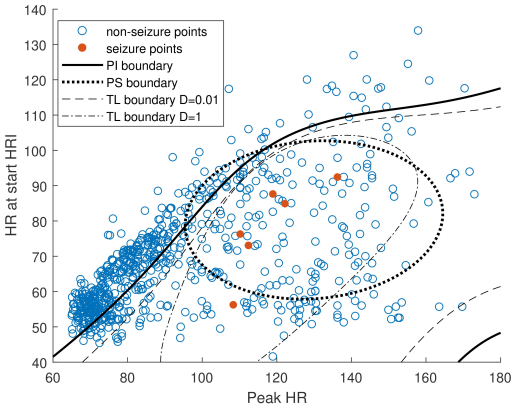


Figure 8. Visualization of SVM boundary of the patient-independent (PI), fully patient-specific (PS) and personalized transfer learning (TL) classifier for different D values (0.01 and 1).