Seizure Prediction using Serial-Parallel Block Concatenated Adaptive Filters

Saisri Padmaja Jonnalagedda, Rishav Rej, Srihari Shankar, Srinath Narayanan
University of California, San Diego
{sjonnala, rrej, s5shanka, srn001}@ucsd.edu

Abstract

Present state-of-the-art seizure prediction approaches use machine learning techniques and neural networks, which includes K-Means clustering, Discriminant Analysis, Radial Basis Function Neural Networks (RBFNN) and Backpropagation Neural Networks for efficient results, but most of them are not real-time approaches. We propose to implement an alternate real-time approach based on adaptive filtering of EEG time-series signal with a novel 'Serial-Parallel Block Concatenation' method which predicts multiple signal samples ahead of time, and hence allows real-time processing. We also propose a novel Probability Density Function (PDF) based method to estimate signal datapoints for the adaptive filter updates. When deployed across various adaptive filters like Kalman, Particle and RLS (Recursive Least Squares), our concatenation approach used along with PDF method gave more accurate real-time predictions ahead of time than a standard adaptive filter.

1. Introduction

Seizures are abnormal electrical activity in the brain, which can be classified as a very high or very low entropy activity. Consequently, the subject can either go into a dazed state or a state of uncontrollable fits or convulsions. They need quick medical attention and are presently there are no complete end-to-end treatment methods. Most adult seizures occur in a certain part of the brain and may or may not spread throughout and are called focal seizures. A condition of such frequent seizures is called epilepsy.

A seizure is divided into three parts: pre-ictal, ictal and post-ictal stages. The onset of seizure till it subsides is defined as the ictal state; the state preceding and post ceding the ictal stage are defined as pre-ictal and post-ictal states respectively. A subject takes some time to recover after a seizure. Thus both the ictal and post-ictal states require medical attention and treatment. Currently, epilepsy affects about 50 million people worldwide, out of which 80% stay in developing nations. The United States alone, currently spends $15.5 billion towards epilepsy treatment. Hence, there has been a lot of research focusing on treating seizures.

Electroencephalogram (EEG) is one of the most common methods for brain monitoring by measuring electrical activity of the brain. EEG can be used to monitor and detect seizures. Seizures, if accurately predicted, can help in reducing the recovery time as well as reduce the cost of treatment. A review by Ramgopal et al. [7] gives a highly exhaustive study of seizure detection methods and products available, which gives insight into the state of the art detection methods as well as emphasizes the importance of seizure prediction. EEG is used most frequently on account of its low noise and millisecond time-frame detection.

2. Prior work

There is a lot of research directed towards studying seizures and there have been some successful approaches in detecting seizures. Some of the early work dates back to 1980s when Gotman published an automatic seizure detection method and an improvement over it in 1990 [1] using background wave characterization and thresholding which had about 75% accuracy on the whole; whereas 41% seizures were not notified to the caregivers. His research was extended to infants by Gotman and team employing a three-way approach of spectral analysis of overlapping epochs, spike detection and low energy spectral evaluation in 1997 [2].

Ghosh-Dastidar et al. proposed use of classifiers and neural networks to classify EEG data into seizures [3]. The classifiers used were k-means clustering, Discriminant Analysis, Radial Basis Function Neural Networks (RBFNN) and Backpropagation Neural Networks. The overall accuracies of these methods typically ranged from about 60% to 90%. The data was decomposed using wavelet transforms. The method proposed by Tzallas et al. in [4] used spectrograms (STFT) of the EEG data to extract features and apply ANNs on it. There are many more approaches using wavelet transforms [5], PCA [6], etc.

However, these methods detect and classify seizures that undoubtedly help in understanding seizures, but do not pro-

vide the advantages that we would hope by predicting a seizure in advance. There has also been a lot of research interest in predicting seizures. There have been statistical, algorithmic and proof of concept based prediction methods proposed. Most methods take into account a direct variation in the time-series characteristics. Lehnertz et al. and Aschenbrenner-Scheibe et al., used the correlation dimension approach for prediction [8–10] whereas Martinerie et al., used correlation density [11]. With these approaches, they achieved 90-95% sensitivity and a 3-12 min in advance prediction of seizures. The use of similarity index was proposed widely in a lot of research by Le Van Quyen et al., Navarro et al., De Clercq et al., and Winterhalder et al. [12–18]. The similarity index methods claim very high sensitivity (96% typically) and a large time in advance prediction of seizures (some up to 80 minutes).

Phase synchronization between EEG data was proposed by Mormann et al., Le Van Quyen et al., and Schelter et al., [19–23]. These methods claim very high results in some cases but the reported results are variable (70-90% for sensitivity and 4-220 mins for prediction). Instead of similarity measures, dissimilarity measures were used by Hively and Protopopescu [24–26]. Dynamical entrainment was also attempted by a few researchers like Iasemidis and Chaovalitwongse [27–30]. Accumulated energy of the data is another measure that occurs in literature, used by Litt, Maiwald et al., Gigola et al., Esteller et al. and Harrison et al. [31–35]. The results for these also vary highly based on the papers, the best being a 90% sensitivity and few minutes ahead of time prediction. Some measures like phase clustering [36], Kolmogorov entropy [37] and Lerner density [38] also sporadically appear in literature. However, rigorous testing of these methods has rendered all of these methods no better than chance [39]. Studies revealed that the signal does not change just before ictal stage and is not localized. This widens the scope of this area.

Latest approaches in seizure prediction include machine learning and neural networks [40] that result in 71% sensitivity. SVM has been reported to give very good results in many articles, initially using a time-varying AR model (some studies also report 100% sensitivity and low FPR) [41, 42]. Cook et al., have proposed a long term advisory seizure prediction system [43] recently in 2013. It has been published as a first-in-man study and reports sensitivities in the range of 65-100% depending on various conditions. Some of the methods described above use Kalman Filter to implement the AR time varying models or even the SVM classifiers [41–44]. A study of Kalman Filter in these applications has been summarized by [45]. In this project, we use Kalman Filter to predict seizures using energy profiles and compare the results to an RLS and Particle filters. We use no machine learning or neural network based algorithms for prediction or classification and instead use a novel 'Serial-Parallel Block Concatenation of Adaptive Filtering' approach.

3. Our Approach

3.1. Proposed method

The breakdown of the proposed solution is as follows:

- Phase 1: Use data points 1 to M to predict $(M+1)^{\text{th}}$ data point using an adaptation of predictive filter (In our case, Kalman Filter).
- Phase 2 : Implement Serial-Parallel Block Concatenation of these filters to predict future points.
- Phase 3 : Attempt to normalize the seizure predicted signals to subside the seizure.

3.1.1 Phase 1

In Phase 1 (Figure 1), we use the first M data points to predict the $(M+1)^{\text{th}}$ data point. This is a standard prediction procedure and the focus here was the "Proof of Concept" which has been described in Section 3. From the predictions on various data sets of M points to test the robustness of our filter design, we safely concluded that the adaptation we use is a highly accurate one.

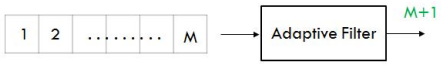

Figure 1. Standard Adaptive Filter

3.1.2 Phase 2

In Phase 2, Figures 2 and 3 implement the 'Serial-Parallel Block Concatenation' approach of the Adaptive Filter in Phase 1 to predict more data points, further ahead in time. This phase is also called 'Filter Block Processing' as the whole processing of serial and parallel filters that we implement happen after the $(i+M)^{\text{th}}$ sample is obtained and before the $(i+M+1)^{\text{th}}$ sample is obtained, where i is the instance at which we start to predict. This method helps us achieve a prediction of $N+k$ samples ahead in one single block processing, i.e, in one sample interval. This is described in detail in the next section.

3.1.3 Phase 3

If the prediction in Phase 2 classifies a signal as a seizure, we intend to subside the seizure by using a nerve stimulation technique to bring the brain back to normal state. One of these techniques is the 'Vagus Nerve Stimulation Procedure' [47]. This involves a device implantation that sends electrical impulses to the Vagus nerve. In this method, a device is implanted under the skin connecting to the Vagus nerve, which when activated sends electrical signals along

the Vagus nerve to the brain stem, which in turn sends signals to certain areas in the brain, thus calming the focal points.

Since Phase 2 of the project is intensive on adaptive filtering, we focused on implementing particle and RLS filters in addition to Kalman filter for further testing and assessment.

3.2. Novel Approaches

3.2.1 Serial-Parallel Block Concatenation Approach

The underlying principle in the **Serial Concatenation** is as follows:

1. The first filter is the standard filter from Phase 1 which gives us the $(M+1)^{\text{th}}$ prediction using the first M data samples.
2. Using this $(M+1)^{\text{th}}$ prediction that we got in the previous step, we now predict the $M+2$ data sample.
3. We append serial filters in this manner and a bound on the number of serial filters can be obtained with the following expression:

$$\textit{Number of serial filters} \times \textit{Filter processing time} < \textit{Sample time} \tag{1}$$

4. When we append blocks of filters serially in this manner, we predict data with a lower confidence level than the previous filter, as we go ahead in the series.

While Serial Concatenation includes delay due to filter processing, Parallel Concatenation gives us the prediction without any delay.

The underlying principle in **Parallel Concatenation** is as follows :

1. The first filter is the standard filter from Phase 1 which gives us the $(M+1)^{\text{th}}$ prediction using the first M data samples.
2. The second filter in parallel uses 2 to M data samples and the $(M+1)^{\text{th}}$ data sample from PDF method and so on.
3. We append parallel rows in this manner, however every parallel row predicts data with a lower confidence level than the previous one.

This way we achieve a prediction of the next $N+k$ data samples even before the next actual data sample is observed. We tested the block processing rigorously with different adaptive filters such as Kalman, RLS and Particle filters; with Kalman Filter being our primary focus. Figure 3 shows the 'Serial-Parallel Block Concatenation' approach.

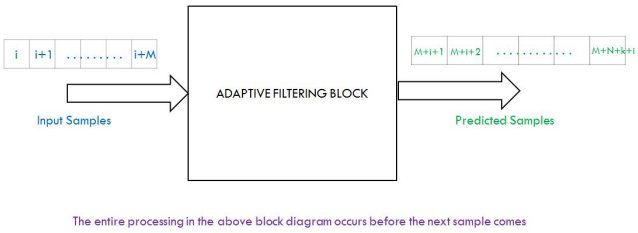

Figure 2. Adaptive Filter Block Processing

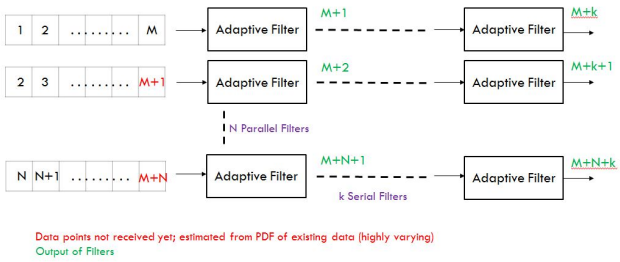

Figure 3. Serial-Parallel Block Concatenation Approach

3.2.2 PDF method

In order to the implement the 'Filter Block processing' explained in Phase 2, to predict the $(i+M+1)^{th}$ sample before the occurrence of $(i+M)^{th}$ sample, we need to generate the signal's pdf distribution, which helps us to predict $M+N+k$ samples in a single block of processing, using serial-parallel concatenation of Kalman filter blocks. The PDF basically establishes a near-accurate relationship with the test data, to obtain $x(M+1), x(M+2), x(M+3)..x(M+n)$ from the data $x(1), x(2)..x(M)$, much ahead of the occurrence of it. It achieves this by using prior knowledge of data collected (in this case, the dataset) and generating a conditional pdf distribution of each of the points $x(M+1)$ to $x(M+n)$ given $x(1)$ to $x(M)$. A number of criteria could be followed to bound the generation of the PDF function. We use the the 'Minimum Mean Error' approach and the 'Minimum Mean Square Error' approach.

Algorithm:

Step 1: The input test data $x(1), x(2), x(3)..x(M)$, where M is equal to the number of filter taps, is fed as input.

Step 2: The error_min is initialized as ∞.

Step 3: From all the training data sets available, a pdf distribution is created.

Step 4: Since the neurotic signals are bivariate, a separate $n \times L$ matrix is formed for neuron X and neuron Y, where n is the number of data present is each file and L is the total number of files.

Step 5: Once the PDF directory is created, consisting of

all the data sets arranged in a matrix form, the test data is compared with each of the files.

Step 6: The error is defined by the difference between the corresponding signals at the same instance of time.

Step 7: For the Minimum mean square approach, the sum of these squares is averaged out across the M samples to compute the average error. For the Minimum mean error approach the normal sum is taken instead of the squared sum.

Step 8: If the average error is less than error_min, the error_min is replaced, and the best fit is established.

Step 9: This process is repeated to generate the $x\_bestfit$ and $y\_bestfit$, where $x\_bestfit$ and $y\_bestfit$ contains the data points generated by the PDF function.

Step 10: The $x\_bestfit$ and $y\_bestfit$ is then passed over to predict the samples using serial-parallel concatenated adaptive filter block processing unit.

We observe that both the Minimum mean error and the Minimum mean square error approach give nearly the same estimate for the given set of data and hence we selected the Minimum mean square error approach for our paper. Even though there are other methods to generate the best-fit data points, the Minimum mean square approach yields the most accuracy, and hence is adopted in this paper.

3.3. Adaptive filter models

We use our 'Serial-Parallel Block Concatenation' approach on the following adaptive filters, and observe the results. The filters are :

1. Kalman filter
2. RLS filter
3. Particle filter

3.3.1 Kalman Filter

The Kalman Filter (KF) is a popular adaptive filter that uses state parameter models for estimation and prediction. It also takes into account, the measurements over time, statistical noise and sensor noise to generate estimates that are more accurate than a single measurement/source of data. The Kalman Filter uses both the observations and measurements to come up with the best estimate. The two most important applications of KF are smoothing and prediction. Mathematically, the KF is a set of five equations that include 'prediction' and 'update' steps.

The effectiveness of KF comes from the fact that it does not assume errors to be Gaussian. The KF requires a clear definition of the system model, initial conditions and error sources. The Kalman Gain (K) is crucial in determining how effectively it performs. It performs a trade-off between the observations and predictions by weighing the observations or predictions based on their respective noise covariance. In this project, we will be using KF to predict signal states of EEG data, and in turn use the predicted data to classify a seizure with a given level of confidence.

In this project, we have employed an adaptation of the Kalman filter, since KF requires an exact description of state model and it is not accurately available in literature for the brain EEG data. Thus, we used a filter that is heavily based on the KF in its framework and update, yet makes use of the knowledge we have of the EEG data. The derivation goes as follows:

$$\begin{aligned}
&\mathbf{x} = \{x(0), x(1), \ldots, x(M-1)\} - \text{Data frame of } M \text{ points} \\
&\mathbf{w} = \{w(0), w(1), 2 \ldots, w(M-1)\} - \text{Filter tap weights} \\
&\mathbf{S} = \mathbb{E}\left((\mathbf{w} - \mathbb{E}(\mathbf{w})(\mathbf{w} - \mathbb{E}(\mathbf{w}))^T\right) - \text{Covariance matrix}
\end{aligned} \tag{2}$$

Predict equations:

$$\begin{aligned}
\mathbf{y} &= \mathbf{w}\mathbf{x}^T \\
\mathbf{K} &= \frac{\mathbf{S}\mathbf{x}^T}{\text{var}(\mathbf{w}) + \mathbf{x}\mathbf{S}\mathbf{x}^T}
\end{aligned} \tag{3}$$

Update equations:

$$\begin{aligned}
\mathbf{w}_{new}^T &= \mathbf{w}^T + \mathbf{K}(\mathbf{y}_{est} - \mathbf{y}) \\
\mathbf{S}_{new} &= (\mathbf{I} - \mathbf{K}\mathbf{x})\mathbf{S}
\end{aligned} \tag{4}$$

These equations justify why the filtering is termed as an adaptation of KF. We have two equations in the predict stage to update the estimate as we have no exact state model and since we are predicting ahead by more than one sample, we have no current measurement data. Combining them, a simplified KF is obtained. As the update step is no longer with respect to the measurement data, the y_{est} here is obtained from a novel PDF approach. This parameter replaces the measurement obtained, and is defined as the estimate state. We use a method that takes into account a set of data obtained in the past and obtains a best estimate of what the measurement could be. W and S are updated in every iteration. The convergence is defined by a lower bound error between y_{est} and y_{actual} or by the maximum number of iterations in case the error bound is not met, as a way of safe-exit. This process predicts the signal state, one time step ahead. We now can use the block concatenation approach to predict successively more points and subsequently, seizures.

3.3.2 Recursive Least Square (RLS) Filter

We tested the RLS filter with our proposed algorithm to compare it with the performance of the Kalman Filter. To use the RLS filter for EEG data, the following equations

were used:

$$\begin{aligned}
&\mathbf{X} = \{x(0), x(1), \ldots, x(M-1)\} - \text{Data frame of } m \text{ points} \\
&\mathbf{W} = \{w(0), w(1), \ldots, w(M-1)\} - \text{Filter tap weights} \\
&\mathbf{P} = \mathbf{R}^{-1} = \mathbf{L}^T \left((\mathbf{X} - \mathbb{E}(\mathbf{X})(\mathbf{X} - \mathbb{E}(\mathbf{X}))^T) \right) \\
&\quad \text{where L is} [\lambda^m, \lambda^{m-1}..1] \\
&\quad \text{Error} = y_{est} \\
&\quad \mathbf{g} = \frac{\mathbf{P_{old}} * \mathbf{X}}{\lambda + \mathbf{X^T} * \mathbf{P_{old}} * \mathbf{X}} \\
&\quad \mathbf{P_{new}} = \frac{\mathbf{I} - g * \mathbf{X^T} * \mathbf{P_{old}}}{\lambda} \\
&\quad \mathbf{W_{new}} = \mathbf{W} + y_{est} * g
\end{aligned} \tag{5}$$

The above process is one iteration and W and P are updated after every such iteration. Just like the Kalman Filter, the RLS has bounds on error value and number of iterations. For the results stated in this paper, value of λ is 0.99.

3.3.3 Particle Filter

Kalman filters work well for linear predictive problems. However any non-linear dependence of the data would not be well modeled. To adapt to the non-linearity, there are several methods proposed in the literature, such as unscented Kalman filters [49], particle filters [50], dynamically re-weighted filter and Bayesian filters [51]. The extension of the Kalman filtering technique works by mapping the nonlinear space to a linear space via a kernel transformation, and hence might not be effective. Here, we venture into particle filters, which are not closed form, but rather an empirical filtering process on a Monte-Carlo simulation of the signal space. They are a class of evolution-based genetic algorithms, that intuitively derive state transformation from genetic variation, mutation-selection and the natural selection phenomena. As with the Kalman filtering problem, our idea is to find the most probable state of the signal at a given time.

The main work flow of the algorithm is borrowed from the Kalman pipelined architecture. The particle filter is given as a signal space-hidden space observational model with the observations being $Y : \{y_1, y_2, y_3..y_n\}$ and the hidden space as $X : \{x_1, x_2, x_3...x_n\}$. Our goal is to predict the state of X accurately from the sequence of Y values. Generally the model is proposed as,

$$\begin{aligned}
X_k &= G(X_k - 1) + N_k \\
Y_k &= H(X_k) + V_k
\end{aligned} \tag{6}$$

where at the k^{th} iteration, G is the state space model, H is the observation-measurement model, N is the noise sequence in the observation, V is the noise sequence in the measurement, and they are mutually independent and uncorrelated. The Monte-Carlo implementation of the particle filter, basically depends on identifying the conditional probability distribution, $P(X|Y)$ as a simulation rather than as a closed form solution.

Innovation
In our implementation, the EEG signal is a time-series signal. X_k and Y_k are modeled as the current and past EEG signal inputs. The signal model is assumed to be autoregressive (AR), although there could be a shift in the stationarity of the signal. Thus, we would need to dynamically model the EEG signal and update our AR parameters. We introduce an adaptation of the vanilla particle filter by dynamically formulating the signal parameters and adapting it. Therefore, the particle filter equations become:

$$\begin{aligned}
X_k &= W^T * Y_k + N_k \\
Y_k &= X_k + V_k
\end{aligned} \tag{7}$$

As described, the particle filter works by iteratively forming a candidate subspace, and identifying the best possible candidate from it. Each candidate point is assigned a probability of being picked based on the conditional probability distribution, as

$$W_{(t,k)} \sim \mathcal{N}(W_k, V_{1(k)}) \; \forall t \in \{1, P\} \tag{8}$$

where P denotes the total number of simulated points.

As seen above, the weights are sampled with a Gaussian around the ideal local optimum weight, with a variance of $V_{1(k)}$. The more probable a point is, the more likely it is to be picked, and hence carries a higher weight. This is accomplished by the following pseudo code:

1. Pick a threshold : $p \in (0, 1)$
2. Sample a point based on the distribution.
3. Generate a random number $r \in (0, 1)$
4. If $r > p$, accept the point
5. Else reject it and continue

In this way, we would generate new filter weights around the local optimum which better models the signal space, rather than just following the AR model. This also ensures that we do not need to impose a strictly linear AR model on the signal space. This flexibility allows us to describe the EEG signal well. Similar to the pipelined Kalman filtering, we employ a pipelined architecture.

4. Results and Design Methodologies

The data we used is from the dataset described in [46]. There were 5 folders (Z, O, N, F, S) of 100 EEG time-series data files each. The data is collected at 173.6 Hz. 'Z' and 'O' folders are EEG data of relaxed people with eyes closed and opened respectively. 'N' and 'F' are EEG data after a

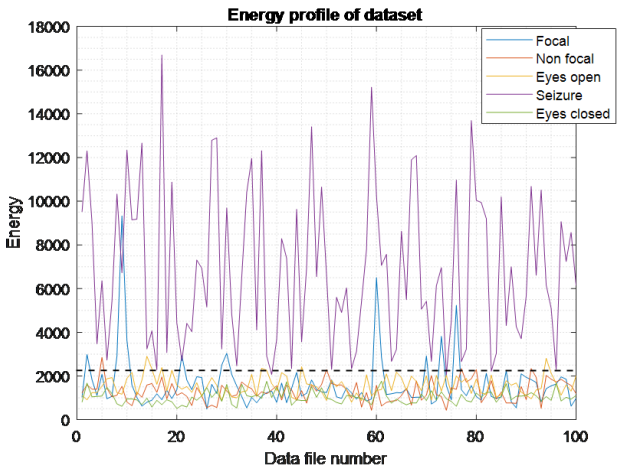

Figure 4. Energy profile of the dataset

seizure in the non-focal and focal areas of the brain respectively. The 100 files in 'S' folder are all seizure data in the ictal phase. The EEG placement on the scalp followed the US standard. The details of data collection, preprocessing of data and study on non-linearity and non-stationarity of the data are mentioned in [46]. For the classification, the energy profile of the data is assessed in Figure 4. For the 'S' folder data, the energy is clearly higher than the other parts. By analyzing the data (as shown in figure), a threshold at 2200 was set. This energy was obtained by sum of data squared (which is equal to sum of FFT squared, but the transform domain is computationally more intensive and by applying Parseval's theorem, they are both equal). Using this as the classification measure, the results for the data base are listed in this section.

To quantify the performance of the prediction process, using serial-parallel concatenated Kalman filters, three factors need to be analyzed. They are:

- The optimal Serial Length
- The optimal Parallel Length
- The optimal Filter Tap Size

The performance as a function of these factors in turn depends upon the accuracy parameters. The accuracy is measured in terms of two parameters:

- True Negative Rate (True seizure which is not predicted)
- False Positive Rate (Non-seizure being predicted as seizure)

Out of the two parameters, the TNR is the more critical error, since we expect a true seizure to always be detected. Nevertheless, we have bounded the false positive rate to be less than 5%. With the increase in the available data set, the prediction of the samples which could be obtained much before their occurrence using filter block processing also increases. Thus, the performance of the prediction process is assessed for the three above mentioned factors by analyzing the accuracy for each of them.

4.1. Assessment of the optimal serial length

The first step is to assess the optimal serial length with low TNR and FPR, giving more importance to TNR. Table 2 and Figure 5 show that as the number of serial blocks increase, the TNR improves, while FPR deteriorates. This increase in TNR as FPR decreases could be because we go beyond the time-frame of stationarity. Even though the crossing is observed at around 450 as the serial length, the TNR at that point is close to 5%. Since, the importance on the TNR is higher; we assume the optimal serial length to be 500, in which case, the FPR is 5%, but the TNR is just 1%.

4.2. Assessment of variation in the parallel length

Now that the optimal serial length is obtained, we need to identify the variation in the TNR and FPR parameters as a function of the parallel steps. From Table 1 and Figure 6, it is understood that for increase in parallel concatenated set, the FPR increases by 0.25% from parallel length 1 to 10, and then remains constant. On the other hand, the TNR does not change as a function of the parallel implementation. Hence, it is understood that the parallel does not alter the TNR and has very low influence on FPR. With the increase in the data set, we expect increase in FPR with an improvement in TNR.

4.3. Assessment of variation in the filter tap size (M)

The third factor which plays a crucial role in estimating the performance of the prediction is the filter tap size (M). Each of the filter blocks have a pre-defined number of filter taps. In order to establish the relationship between the accuracy of prediction as a function of the number of filter taps, we perform the experiment across different filter tap sizes and calculate FPR/TNR values. Further, an additional filtering processes (RLS and Particle), is compared with the Kalman, to appreciate the performance of Kalman for the same number of filter taps in both.

The Tables 3, 4 and 5, and Figures 7 and 8 establishes the performance comparison between the RLS, Particle and the Kalman filters as a function of the number of filter taps. It is understood that RLS fails when the filter size exceeds 30. Even when the filter size is less than 30, the accuracy parameters (TNR and FPR) is much better in the case of Kalman as compared to that of RLS and is comparable with particle filter. Hence, Kalman is a clear winner in this case and hence it is preferred for the prediction process.

Parallel \ Series	100		200		300		400		500	
	FPR	TNR	FPR	TNR	FPR	TNR	FPR	TNR	FPR	TNR
1	0.25	34	0.5	19	1	12	3.5	7	5	1
10	0.5	34	0.75	19	1.25	12	3.75	7	5.25	1
20	0.5	34	0.75	19	1.25	12	3.75	7	5.25	1
30	0.5	34	0.75	19	1.25	12	3.75	7	5.25	1
40	0.5	34	0.75	19	1.25	12	3.75	7	5.25	1

Table 1. Table of TNR and FPR variation in both Serial and Parallel Concatenation

Serial Length	FPR (Percentage)	TNR (Percentage)
100	0.25	34
200	0.5	19
300	1	12
400	3.5	7
500	5	1

Table 2. Filter Tap ($M = 30$). Parallel Length - 1, Number of Samples considered for FPR - 400 (4-Non-focal files containing 100 each), Number of Samples considered for TNR - 100 (1-Focal file containing 100)

Serial Length	FPR%	TNR%
100	0.4	34
300	1.5	14
400	5.25	2
Filter size	FPR%	TNR%
1	21	45
10	9	4
30	5.25	2

Table 3. Variation by serial length and filter size for Particle Filter

Filter Tap Size (m)	FPR	TNR
10	5.5	2
20	5	2
30	5	1
40	5.5	1
50	5.75	1
100	9.25	0

Table 4. Kalman Filter. Constant parameters: Serial Length - 500, Parallel Length - 1

Filter Tap Size (m)	FPR	TNR
10	2.5	28
20	3.5	9
30	0	100
40	0	97
50	0	98
100	0	100

Table 5. RLS Filter. Constant parameters: Serial Length - 500, Parallel Length - 1

Now that Kalman is preferred, let us understand the variation of the accuracy parameters as a function of the filter tap size. As the filter tap size increases there is a decrease in the FPR till $M = 30$. When $M > 30$, the FPR increases exponentially. The TNR on the other hand, decreases gradually as the filter size increases, and reaches 0 when $M = 100$. However, when M=100, the FPR is high (9.25%). The TNR remains at 1% between $M = 30$ to $M = 100$. Hence the optimal filter tap size to attain a good TNR and FPR is at $M = 30$. Thus, the optimal filter tap size is obtained as 30.

For the particle filter, the initialization plays a major role in convergence, and although we can expect a good convergence by 1000 iterations, it would not be guaranteed. But, even when the output does not converge, the next run of the particle filter creates a new conditional probability distribution and hence it would not affect the final seizure prediction significantly. We design the particle filter with 1000 particles with 1000 iterations in each run. The error bound is kept as 0.1 for convergence. Table 3 show the variation of FPR and TNR with serial length and filter size respectively for the particle filter. We can see that higher serial lengths result in lesser TNR for the particle filter. The filter lengths similarly, showed improved performance with longer filters. This proves the point that the intrinsic signal is not a Hidden Markov Model, but rather is a AR signal, with past dependencies.

However, the major drawback of using particle filters for our approach, is that the dimension of the state space is decided by the length of the filter. Hence, as we increase the dimension of the filter, the curse of dimensionality forces the algorithm to perform poorly and under-fit due to lack of data which is evident from Table 3.

5. Conclusion

In this paper, Seizure detection and prediction were implemented and tested using three major adaptive filtering techniques, namely, serial-parallel concatenated Kalman filter, RLS filter and Particle Filter. The innovative PDF method adopting mean square error was implemented to predict additional $N + k$ states, where k is limited based on the efficiency of the accuracy parameters. On extensive testing of the Kalman filter, the accuracy parameters improved

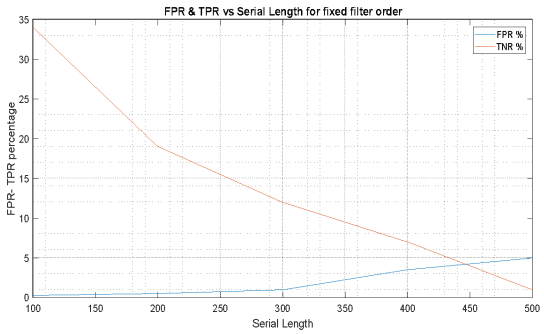

Figure 5. FPR and TPR vs Serial Length for fixed order Kalman filter

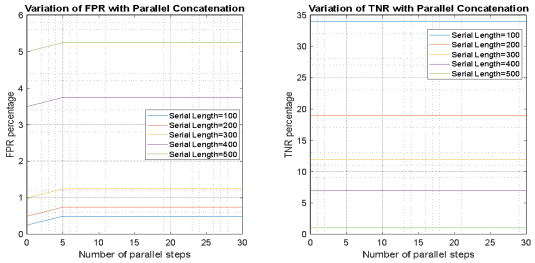

Figure 6. Variation of FPR(left) and TNR(right) for Kalman Filter with parallel block length

with serial implementation while the parallel concatenation had no significant improvement. But with no substantial degradation on parallel end we can now successfully predict $N + k$ states ahead. The threshold on the number of serial sections, the tap size per filter and the number of parallel concatenations were determined for improved performance. The performance of the Kalman filter was compared against the RLS and Particle filters. RLS had a poor performance when compared to Kalman with a fatal error at tap size equal to 30. In order to compensate for the non linear dependencies between the data, the prediction was implemented and tested using Particle filter. Even though the results are comparable with the Kalman filter, it was obtained at the cost of higher computational complexity. Thus, we see that the 'Serial-Parallel Block Concatenation' Kalman filtering approach used along with PDF method gives more accurate real-time predictions of seizure ahead of time ($N + k$ steps ahead) than a standard adaptive filter using EEG signals.

6. Future Scope

In this paper, we worked with a patient in a controlled environment to observe the aforementioned results. We can expand the scope by taking in neural readings which include high entropy physical and check the robustness of our 'Serial-Parallel Block Concatenation' approach. We can also improve the algorithm to work with scalp EEG, rather than brain EEG data for portability. There is also further

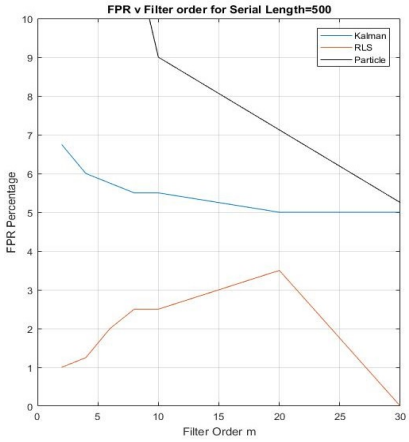

Figure 7. Variation of FPR with change in filter order for Kalman, RLS and Particle Filters

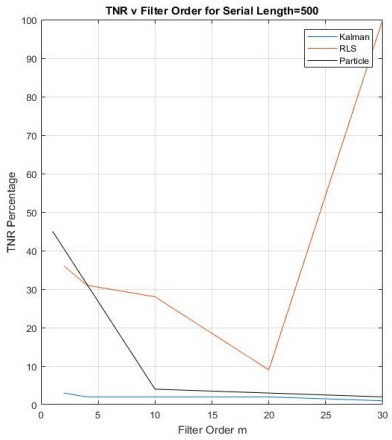

Figure 8. Variation of TNR with change in filter order for Kalman, RLS and Particle Filters

scope in devising a more accurate principle for the 'PDF Method' which can give better estimate than the underlying principle of 'Minimum Mean Square Approach'. The 'Serial-Parallel Block Concatenation' Approach has been tested across various adaptive filters alone and its performance has not been analyzed against state of the art approaches that include neural nets and SVM approaches as this was out of the scope of the course. This can be a definite research problem to look at their comparison.

References

[1] J. Gotman, "Automatic seizure detection: improvements and evaluation," *Electroencephalography and clinical Neurophysiology*, vol. 76, no. 4, pp. 317–324, 1990.

[2] J. Gotman, D. Flanagan, J. Zhang, and B. Rosenblatt, "Automatic seizure detection in the newborn: methods and initial evaluation," *Electroencephalography and clinical neurophysiology*, vol. 103, no. 3, pp. 356–362, 1997.

[3] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band

wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE transactions on biomedical engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.

[4] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in eegs using time-frequency analysis," *IEEE transactions on information technology in biomedicine*, vol. 13, no. 5, pp. 703–710, 2009.

[5] Y. U. Khan and J. Gotman, "Wavelet based automatic seizure detection in intra-cerebral electroencephalogram," *Clinical Neurophysiology*, vol. 114, no. 5, pp. 898–908, 2003.

[6] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, pp. 512–518, 2008.

[7] S. Ramgopal, S. Thome-Souza, M. Jackson, N. E. Kadish, I. S. Fernandez, J. Klehm, W. Bosl, C. Reinsberger, S. Schachter, and T. Loddenkemper, "Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy," *Epilepsy and behavior*, vol. 37, pp. 291–307, 2014.

[8] K. Lehnertz and E. Ce., "Can epileptic seizures be predicted?," *Evidence from non-linear time series analysis of brain electrical activity*, vol. 1998, p. 80.

[9] K. Lehnertz, R. G. Andrzejak, J. Arnhold, T. Kreuz, F. Mormann, and C. Rieke, "et al," *Non-linear EEG analysis in epilepsy: its possible use for inter-ictal focus localization, seizure anticipation, and prevention*, vol. 2001, p. 18.

[10] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. U. Voss, J. Timmer, and A. Schulze-Bonhage, "How well can epileptic seizures be predicted?," *An evaluation of a nonlinear method*, vol. 2003, p. 126.

[11] J. Martinerie and C. Adam, "Le van quyen m, baulac m, clemenceau s, renault b, et al," *Epileptic seizures can be anticipated by non-linear analysis*, vol. 1998, p. 4.

[12] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, "Anticipating epileptic seizure in real time by a non-linear analysis of similarity between eeg recordings," *Neuroreport*, vol. 1999, p. 10.

[13] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clemenceau, and F. Varela, "Spatio-temporal characterization of non-linear changes in intra-cranial activities prior to human temporal lobe seizures," *Eur J Neurosci*, vol. 2000, p. 12.

[14] M. Le Van Quyen, J. Martinerie, V. Navarro, P. Boon, M. D.Have, and C. Adam, "et al," *Anticipation of epileptic seizures from standard EEG recordings*, vol. 357, p. 183.

[15] V. Navarro and J. Martinerie, "Le van quyen m, clemenceau s, adam c, et al," *Seizure anticipation in human neo-cortical partial epilepsy*, vol. 2002, p. 125.

[16] W. De Clercq, P. Lemmerling, S. Van Huffel, and W. Van Paesschen, "Anticipation of epileptic seizures from standard eeg recordings," *Lancet*, vol. 2003, p. 361.

[17] M. Winterhalder, T. Maiwald, H. U. Voss, R. Aschenbrenner-Scheibe, J. Timmer, and A. Schulze-Bonhage, "The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods," *Epilepsy Behaviour*, vol. 2003, p. 4.

[18] V. Navarro and J. Martinerie, "Le van quyen m, baulac m, dubeau f, gotman j," *Seizure anticipation: do mathematical measures correlate with video-EEG evaluation*, vol. 2005, p. 46.

[19] F. Mormann, K. Lehnertz, P. David, and E. Ce., "Mean phase coherence as a measure for phase synchronization and its application to the eeg of epilepsy patients," *Physica D*, vol. 2000, p. 144.

[20] M. Le Van Quyen, J. Martinerie, V. Navarro, Baulac, and V. F. M., "Characterizing neuro-dynamic changes before seizures," *J Clin Neurophysiol*, vol. 18, p. 191.

[21] F. Mormann, T. Kreuz, R. G. Andrzejak, P. David, K. Lehnertz, and E. Ce., "Epileptic seizures are preceded by a decrease in synchronization," *Epilepsy Res*, vol. 53, p. 173.

[22] M. Le Van Quyen, J. Soss, V. Navarro, R. Robertson, M. Chavez, and M. Baulac, "et al," *Preictal state identification by synchronization changes in long-term intra-cranial EEG recordings*, vol. 2005, p. 116.

[23] B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, and A. SchulzeBonhage, "et al," *Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction*, vol. 2006, p. 16.

[24] L. M. Hively, V. A. Protopopescu, and G. Pc., "Timely detection of dynamical change in scalp eeg signals," *Chaos*, vol. 2000, p. 10.

[25] V. A. Protopopescu, L. M. Hively, and G. Pc., "Epileptic seizure forewarning from scalp eeg," *J Clin Neurophysiol*, vol. 2001, p. 18.

[26] L. M. Hively and P. Va., "Channel-consistent forewarning of epileptic events from scalp eeg," *IEEE Trans Biomed Eng*, vol. 2003, p. 50.

[27] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and S. Ds., "Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures," *J Comb Optim*, vol. 2001, p. 5.

[28] L. D. Iasemidis, D. S. Shiau, W. Chaovalitwongse, J. C. Sackellares, P. M. Pardalos, and J. C. Principe, "et al," *Adaptive epileptic seizure prediction system*, vol. 2003, p. 50.

[29] L. D. Iasemidis, D. S. Shiau, P. M. Pardalos, W. Chaovalitwongse, K. Narayanan, and A. Prasad, "et al," *Long-term prospective on-line real-time seizure prediction*, vol. 2005, p. 116.

[30] W. Chaovalitwongse, L. D. Iasemidis, P. M. Pardalos, P. R. Carney, D. S. Shiau, and S. Jc., *Performance of a seizure warning algorithm base*.

[31] B. Litt, R. Esteller, J. Echauz, M. D.Alessandro, R. Shor, and T. Henry, "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients," vol. 2001, p. 30.

[32] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. SchulzeBonhage, and J. Timmer, "Comparison of three non-linear seizure prediction methods by means of the seizure prediction characteristic," *Physica D*, vol. 2004, p. 194.

[33] S. Gigola, F. Ortiz, C. E. D-Attellis, W. Silva, and S. Kochen, "Prediction of epileptic seizures using accumulated energy in a multiresolution framework," *J Neurosci Methods*, vol. 2004, p. 138.

[34] R. Esteller, J. Echauz, M. D.Alessandro, G. Worrell,

S. Cranstoun, and G. Vachtsevanos, "et al," *Continuous energy variation during the seizure cycle: towards an on-line accumulated energy*, vol. 2005, p. 116.

[35] M. A. Harrison, M. G. Frei, and O. I. A. energy revisited *Clin Neurophysiol*, vol. 116, p. 527.

[36] S. Kalitzin, D. Velis, P. Suffczynski, and J. Parra, "Lopes da silva f," *Electrical brain-stimulation paradigm for estimating the seizure onset site and the time to ictal transition in temporal lobe epilepsy*, vol. 2005, p. 116.

[37] W. van Drongelen, S. Nayak, D. M. Frim, M. H. Kohrman, V. L. Towle, and H. C. Lee, "Seizure anticipation in paediatric epilepsy: use of kolmogorov entropy," vol. 2003, p. 29.

[38] R.-O. E. Cerf, "Spectral analysis of stereo-electroencephalograms: pre-ictal slowing in partial epilepsies," vol. 2000, p. 83.

[39] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: the long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2007.

[40] P. Mirowski, D. Madhavan, Y. LeCun, and R. Kuzniecky, "Classification of patterns of eeg synchronization for seizure prediction," *Clinical neurophysiology*, vol. 120, no. 11, pp. 1927–1940, 2009.

[41] L. Chisci, A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio, and F. Fuggetta, "Real-time epileptic seizure prediction using ar models and support vector machines," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1124–1132, 2010.

[42] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of eeg using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[43] M. J. Cook, T. J. O'Brien, S. F. Berkovic, M. Murphy, A. Morokoff, and G. Fabinyi, "Prediction of seizure likelihood with a long-term implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study," vol. 12, no. 6, pp. 563–571, 2013.

[44] S. Enshaeifar, L. Spyrou, S. Sanei, and C. C. Took, "A regularised eeg informed kalman filtering algorithm," *Biomedical Signal Processing and Control*, vol. 25, pp. 196–200, 2016.

[45] V. P. Oikonomou, A. T. Tzallas, S. Konitsiotis, D. G. Tsalikakis, and D. I. Fotiadis, *The use of Kalman filter in biomedical signal processing*. In Kalman Filter Recent Advances and Applications. InTech, 2009.

[46] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of non-linear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state, , vol. 64, no. 6, article id 061907, 8 pages," vol. 2001.

[47] E. A. Wan and R. V. D. Merwe, *The unscented Kalman filter for nonlinear estimation Sign In or Purchase*. The IEEE, 2000.

[48] E. A. Wan and R. V. D. Merwe, "The unscented kalman filter for nonlinear estimation," 2000.

[49] E. A. Wan and R. V. D. Merwe, *The unscented Kalman filter for nonlinear estimation*. The IEEE, 2000.

[50] D. M. Pierre, "Non linear filtering: Interacting particle solution," *Markov Processes and Related Fields*, vol. 2, pp. 555–580, 1996.

[51] N. G. D. S. A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEE Proceedings F - Radar and Signal Processing*, vol. 140, April 1993.