Seizure Susceptibility Prediction in Uncontrolled Epilepsy

Nhan Duy Truong [1,2], Yikai Yang [1], Christina Maher [1], Levin Kuhlmann [3,4], Alistair McEwan [1], Armin Nikpour [5,6] and Omid Kavehei [1,2]*

[1] Australian Research Council Training Centre for Innovative BioEngineering, School of Biomedical Engineering, Faculty of Engineering, The University of Sydney, Sydney, NSW, Australia, [2] The University of Sydney Nano Institute, Sydney, NSW, Australia, [3] Faculty of Information Technology, Monash University, Melbourne, VIC, Australia, [4] Department of Medicine - St. Vincent's Hospital Melbourne, The University of Melbourne, Fitzroy, VIC, Australia, [5] Comprehensive Epilepsy Service and Department of Neurology at the Royal Prince Alfred Hospital, Sydney, NSW, Australia, [6] Faculty of Medicine and Health, Central Clinical School, The University of Sydney, Sydney, NSW, Australia

Epileptic seizure forecasting, combined with the delivery of preventative therapies, holds the potential to greatly improve the quality of life for epilepsy patients and their caregivers. Forecasting seizures could prevent some potentially catastrophic consequences such as injury and death in addition to several potential clinical benefits it may provide for patient care in hospitals. The challenge of seizure forecasting lies within the seemingly unpredictable transitions of brain dynamics into the ictal state. The main body of computational research on determining seizure risk has been focused solely on prediction algorithms, which involves a challenging issue of balancing sensitivity and false alarms. There have been some studies on identifying potential biomarkers for seizure forecasting; however, the questions of "What are the true biomarkers for seizure prediction" or even "Is there a valid biomarker for seizure prediction?" are yet to be fully answered. In this paper, we introduce a tool to facilitate the exploration of the potential biomarkers. We confirm using our tool that interictal slowing activities are a promising biomarker for epileptic seizure susceptibility prediction.

Keywords: epileptic seizure forecasting, probabilistic programming, Bayesian, variational inference, uncertainty level

1. INTRODUCTION

There has been great interest recently in identifying biomarkers for seizure susceptibility by looking into critical transitions in brain dynamics in order to enhance the precision of seizure forecasting in a cohort of patients with focal epilepsy (1–3). These studies often require a very long recording that is not available and, in fact, are critically lacking. Chronic and often intracranial electroencephalogram (EEG) recordings demonstrated some limited evidence of circadian, multidien, and circannual cycles in epileptic brain dynamics (4–6). In determining seizure-risk, we believe that understanding what features or biomarkers in the EEG signals lead to such seizure-risk level.

The availability of a seizure forecasting system that can notify patients or their carers about forthcoming seizure-risk can drastically improve patients' quality of life and the chance to develop innovative interventions and preventative therapies. Many studies have been on forecasting seizures; most of them used the signal-modal approach based on electroencephalogram (EEG) signals. These studies can be grouped into two categories: (1) finding discriminative features with various signal processing and transformation techniques and (2) leveraging deep learning's
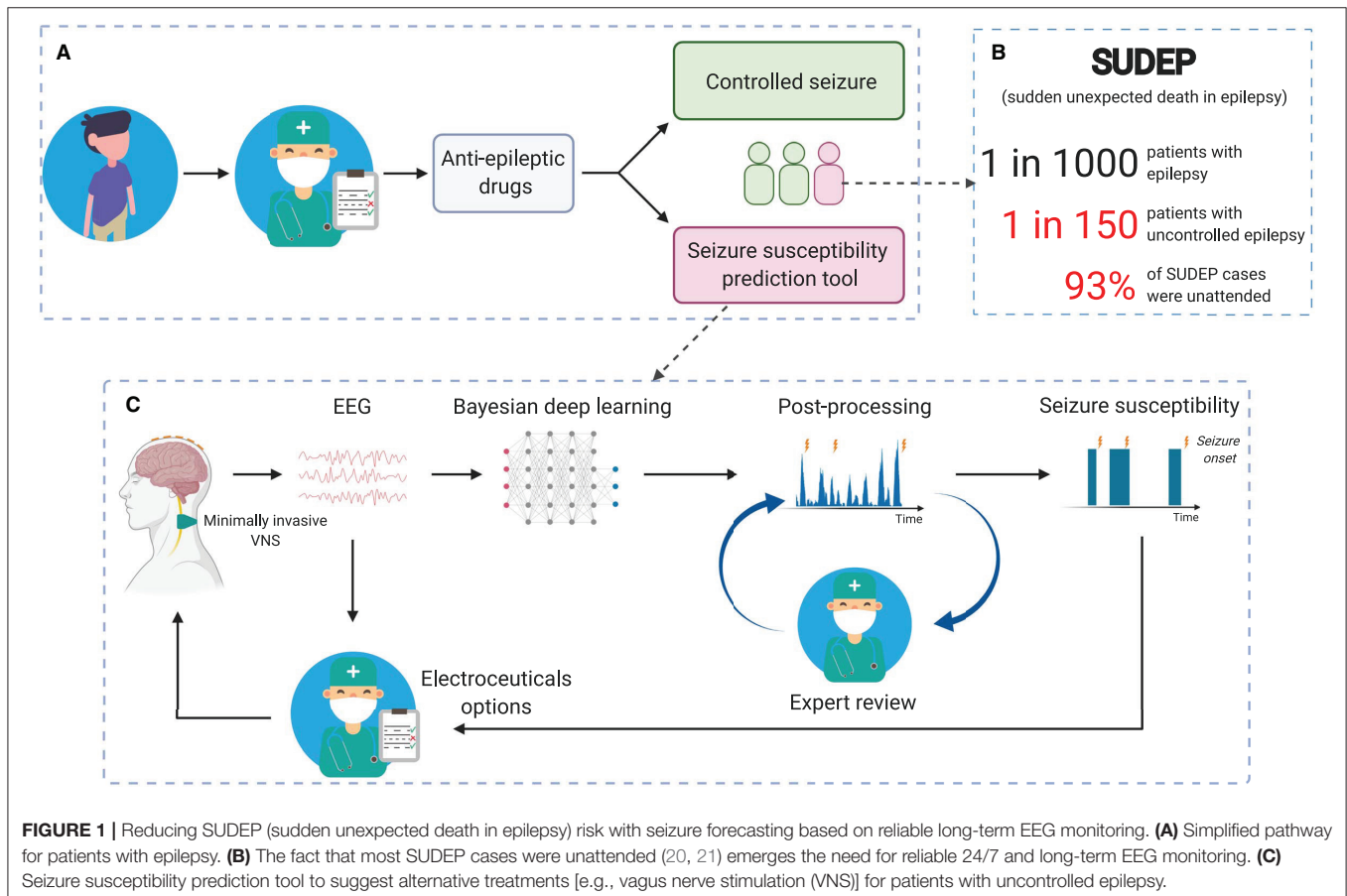
capability of extracting high-level features. In the first group, the most common approach is to use spatio-temporal correlation features, auto-regressive modeling predictive error, Hjorth parameters, spectral power, energy wavelet coefficients, and other statistics (7, 8). Other discriminative features include phase and amplitude lock values (9), common spatial pattern (10), permutation entropy (11), bispectrum features (12). In the second group, the convolutional neural network (CNN) and recurrent neural network (RNN) have shown their capability to extract high-level features that can be used for forecasting seizures. Particularly, CNN was used on the EEG signal spectrogram (13), raw EEG, and fast Fourier transform (FFT) of raw EEG (14), local mean decomposition of raw EEG (15), and the common spatial pattern of multi-channel EEG signals (16). CNN was also used in unsupervised learning as effective feature extraction for seizure prediction (17). To further extract the temporal characteristics over time-series data, Wei et al. (18) applied CNN with long short-term memory recurrent network on the spectrogram of EEG signals. The combination of convolutional and recurrent neural networks is also effective when using multi-timescale of raw time-series EEG signals (19).

In this work, we propose a framework (see **Figure 1**) to minimize the risk of sudden unexpected death in epilepsy (SUDEP), especially for patients with uncontrolled epilepsy. We also introduce a tool to facilitate the exploration of biomarkers for epileptic seizure forecasting. Specifically, we use probabilistic programming and propose a framework to incorporate other relevant information into an EEG-based seizure forecasting system. As an advantage of using probabilistic programming, our system not only can forecast impending seizures but also quantify the uncertainty level of its decision-making.

2. DATASET

EPILEPSIAE is the largest epilepsy database that contains EEG data from 275 patients (22). However, up to the time of this writing, only 30 surface EEG and 30 invasive EEG datasets are made available (23). We believe the use of surface EEG is more beneficial because it is non-invasive so it can be applied to a broader group of patients. In this study, we analyze scalp-EEG of 30 patients with 261 leading seizures and 2881.4 interictal hours in total in this work. The time-series EEG signals were recorded at a sampling rate of 256 Hz and from 19 electrodes. Seizure onset information obtained by two methods, namely EEG based and video analysis, is provided. In our study, we use seizure onset information using an EEG based technique, where the onsets were determined by visual inspection of EEG signals performed

FIGURE 1 | Reducing SUDEP (sudden unexpected death in epilepsy) risk with seizure forecasting based on reliable long-term EEG monitoring. **(A)** Simplified pathway for patients with epilepsy. **(B)** The fact that most SUDEP cases were unattended (20, 21) emerges the need for reliable 24/7 and long-term EEG monitoring. **(C)** Seizure susceptibility prediction tool to suggest alternative treatments [e.g., vagus nerve stimulation (VNS)] for patients with uncontrolled epilepsy.

by an experienced clinician (22). **Table 1** provides a summary of the dataset being studied in this work.

3. METHOD

3.1. Pre-processing

We split EEG signals into 30-s segments with 50% overlap. We perform a short-time Fourier transform (STFT) with a cosine window of 1-s length and 50% overlap on each 30-s segment and get data with a dimension of ($n \times 59 \times 129$), where n is the number of EEG channels. We remove the first and last two elements along the second axis, which corresponds to time, to eliminate any potential disruption of signal near the signal window's edges. We also remove the dc component of the STFT, which is the first element along the last axis. The final dimension of pre-processed data is ($n \times 56 \times 128$).

TABLE 1 | The EPILEPSIAE scalp-EEG dataset.

Patient	Gender	Age	No. of seizures	No. of leading seizures[*]	Interictal hours
Pat1	Male	36	11	11	68.9
Pat2	Female	46	8	8	114.9
Pat3	Male	41	8	8	96.3
Pat4	Female	67	5	5	126
Pat5	Female	52	8	8	204.1
Pat6	Male	65	8	7	92.2
Pat7	Male	36	5	5	75.7
Pat8	Male	26	22	11	65.6
Pat9	Male	47	6	6	51.1
Pat10	Male	44	11	11	60.7
Pat11	Male	48	14	14	57.8
Pat12	Male	28	9	9	94.1
Pat13	Male	46	8	8	101.3
Pat14	Female	62	6	6	115.7
Pat15	Female	41	5	5	82.8
Pat16	Female	15	6	6	51.1
Pat17	Female	17	9	9	82.4
Pat18	Male	47	7	6	133
Pat19	Male	32	22	21	75.4
Pat20	Male	47	7	7	115.3
Pat21	Female	31	8	8	106.6
Pat22	Male	38	7	7	88.2
Pat23	Male	50	9	9	179.6
Pat24	Female	54	10	10	36.2
Pat25	Male	42	8	8	109.8
Pat26	Male	13	9	9	97.1
Pat27	Male	58	9	8	99.9
Pat28	Female	35	9	9	95.2
Pat29	Male	50	10	10	111.9
Pat30	Female	16	12	12	92.5

We are considering leading seizures only. Seizures that are <30 min away from the previous one are considered as one seizure only, and the onset of a leading seizure is used as the onset of the combined seizure.

3.2. Bayesian Convolutional Neural Network

In this paper, we will use variational inference to approximate posterior densities for Bayesian models (24). Consider $x = x_{1:n}$ as a set of observed variables and $z = z_{1:m}$ as a set of hidden variables, with joint density $p(z, x)$. The inference problem calculates the conditional density of the hidden variables given the observed variables, $p(z|x)$.

$$p(z|x) = \frac{p(z, x)}{p(x)}, \tag{1}$$

where $p(x)$ is intractable in many models (24).

Variational inference overcomes this by specifying a variational family $\mathcal{Q}$ over the hidden variables (24). The inference problem becomes finding the best candidate $q(z) \in \mathcal{Q}$ that is closest in Kullback-Leibler (KL) divergence to $p(z|x)$. The optimization subsequently can be achieved by maximizing a function called the evidence lower bound (ELBO) which is equivalent to minimizing the KL divergence between $q(z)$ and $p(z|x)$. ELBO is expressed as follows (24):

$$\begin{aligned} ELBO(q) &= \mathbb{E}\big[\log p(z, x)\big] - \mathbb{E}\big[\log q(z)\big] \\ &= \mathbb{E}\big[\log p(x|z)\big] + \mathbb{E}\big[\log p(z)\big] - \mathbb{E}\big[\log q(z)\big] \\ &= \mathbb{E}\big[\log p(x|z)\big] - \mathrm{KL}\big(q(z)\|p(z)\big) \end{aligned} \tag{2}$$

The stochastic variational inference was proposed by Hoffman et al. (25) to help Bayesian neural networks scale efficiently to large datasets. Particularly, this method generates noisy estimates of the natural gradient of the ELBO by repeatedly sub-sampling (mini-batch) the dataset. The loss function can be defined as the negative of ELBO, i.e., minimizing the loss is equivalent to maximizing the ELBO.

$$\mathrm{loss} = -ELBO(q) = -\mathbb{E}\big[\log p(x|z)\big] + \mathrm{KL}\big(q(z)\|p(z)\big) \tag{3}$$

In an EEG-based seizure prediction system, x is the EEG signals, and z is a variable indicating a seizure to occur in the time window $\mathcal{T} = [SPH : SPH + SOP]$. SPH stands for seizure prediction horizon that is defined as the period where seizure should not occur after an alarm rises. SOP stands for seizure occurrence period that is defined as the interval where seizure onset is expected to occur (26).

3.3. Probabilistic Convolutional Neural Network With Data Fusion

In this section, we will incorporate signals other than EEG signals into the Bayesian CNN. We want to estimate the probability of having a seizure given EEG signals, $p(z|x)$, which is the Bayesian CNN's output. Besides EEG signals, we have other relevant data and want to combine all the seizure forecasting information. Circadian information or time of the day has been used to improve the performance of a seizure prediction system (27). For another instance, electrocardiogram that could change around and even before seizure onsets has been shown helpful in predicting epileptic seizures (28, 29). Other physiological signals

that have been observed to change prior to seizure onset, such as blood oxygenation, metabolism, can be used as auxiliary data for seizure prediction (30, 31).

Let us start with EEG signals and one extra signal called d. Using Bayes theorem, the posterior probability of having a seizure in the time-window $\mathcal{T}$ can be expressed as:

$$p(z|x, d) = \frac{p(d|z, x)p(z|x)}{p(d|x)} \qquad (4)$$

Assume x and d are independent, (e.g., EEG signals are independent with the time of the day and can be considered independent with blood oxygenation), we can rewrite (4) as follows.

$$p(z|x, d) = \frac{p(d|z)p(z|x)}{p(d)} \qquad (5)$$

Similarly, for two extra signals, d_1 and d_2 with an assumption that x, d_1, and d_2 are independent of each other (e.g., time of the day and blood oxygenation), the posterior probability of having seizure in the time window $\mathcal{T}$ can be expressed as:

$$
\begin{aligned}
p(z|x, d_1, d_2) &= \frac{p(d_1|z, d_2)p(z|x, d_2)}{p(d_1|d_2)} \\
&= \frac{p(d_1|z)p(z|x, d_2)}{p(d_1)}
\end{aligned}
\qquad (6)
$$

By substituting Equation (5) (with d replaced by d_2) to Equation (6), we have:

$$p(z|x, d_1, d_2) = \frac{p(d_1|z)p(d_2|z)p(z|x)}{p(d_1)p(d_2)} \qquad (7)$$

To estimate $p(d_1|z)$ and $p(d_2|z)$, we applied a kernel density estimation using Gaussian kernels on a histogram containing time of the day (ToD) of seizure occurrences (see **Figure 2**) (32). Regarding the kernel density estimation parameters, we used Scott's rule for bandwidth selection and assumed all data points are equally weighted. Note that here we approximate $p(d_1|z) \approx$

FIGURE 2 | Distribution of time-of-day of seizure occurrences in the EPILEPSIAE scalp-EEG dataset.

$p(d_1|z')$ and $p(d_2|z) \approx p(d_2|z')$, where z' is the variable indicating an occurrence of seizure. The approximation is reasonable because we choose the time window $\mathcal{T} = [5 : 35\ \text{min}]$ which is <1 h.

To incorporate Equation (7) into the training of the Bayesian CNN, we modify the output of the last fully-connected layer (see Fig. 3), before softmax activation (33) as follows.

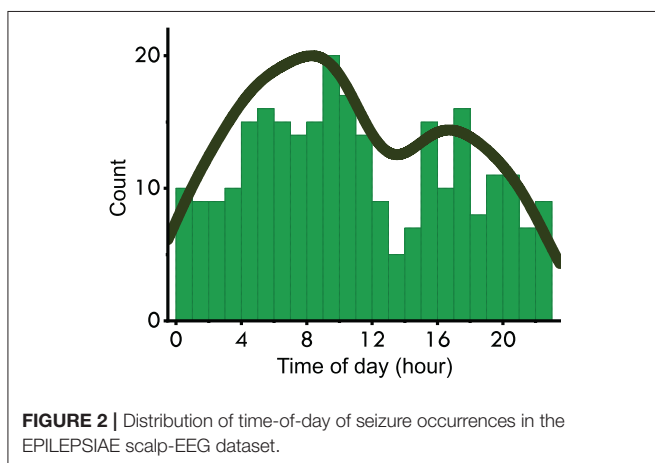$$\text{new-output}_l = \frac{p(d_1|z')p(d_2|z') \times \text{output}_l}{p(d_1)p(d_2)}, \qquad (8)$$
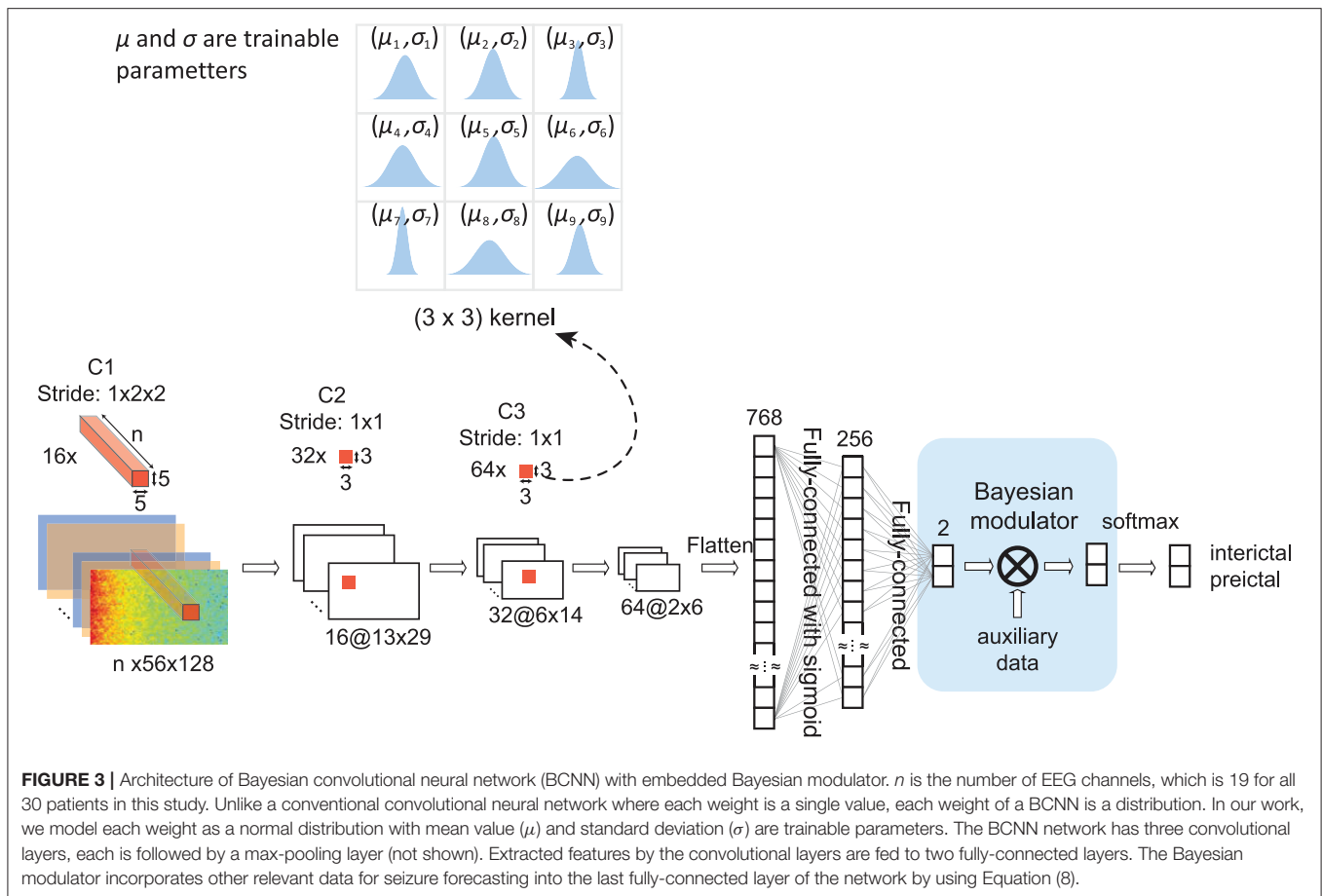
where $p(d_1|z')$ and $p(d_2|z')$ can be derived from the kernel density estimation. For example of time of the day, $p(d_1) = 1/24$ because the probability of having the auxiliary signal at a given hour is 1/24; $p(d_1|z')$ can be inferred from **Figure 2**. Note that Equations (7) and (8) can be extended with more extra signals d given that they are independent on each other.

The Bayesian convolutional neural network (BCNN) with Bayesian modulator as data fusion is depicted in **Figure 3**. Unlike a conventional CNN, where each weight is a single value, each weight of a BCNN is a distribution estimated during the training phase. In this work, we model each weight as a Gaussian distribution with mean and standard deviation values are trainable parameters. Input to the BCNN is the STFT of 30-second windows with size of ($n \times 56 \times 128$) (see Session 3.1 for details). The network starts with a convolutional layer consisted of 16 3-dimensional kernels of size ($n \times 5 \times 5$), valid padding, and a stride of ($1 \times 2 \times 2$). A max-pooling layer follows the first convolutional layer with a pooling size of ($1 \times 2 \times 2$). The network continues with two blocks of convolutional-pooling combinations, each consists of one convolutional layer with a kernel size of (3×3), valid padding and stride of (1×1), and one max-pooling layer with a pooling size of (2×2). The number of convolutional kernels in the two blocks is 64 and 128. The next two layers are fully-connected layers with output sizes of 256 and 2, respectively. The output of the last fully-connected layer is fed to the Bayesian modulator where we apply Equation 8 for data fusion, then is applied softmax activation to get the final output of the network.

4. RESULTS

This section tests the Bayesian convolutional neural network (BCNN) with the EPILEPSIAE scalp EEG dataset with and without auxiliary signal: time-of-day (ToD). Following Truong et al. (13), we use SPH of 5 min and SOP of 30 for calculating the performance. We also compare a seizure prediction system using a convolutional neural network (CNN) proposed by Truong et al. (13) as a baseline. **Figure 4A** shows the overall performance of the BCNN with and without auxiliary signal and the baseline CNN. Compared to the CNN that has an average AUC of 71.65%, BCNN achieves an AUC of 68.69% that is around 3% lower than that of CNN. By using the time-of-day information, the overall performance of BCNN-ToD is slightly improved by 0.3–69.03%. There is strong agreement between the methods that

FIGURE 3 | Architecture of Bayesian convolutional neural network (BCNN) with embedded Bayesian modulator. n is the number of EEG channels, which is 19 for all 30 patients in this study. Unlike a conventional convolutional neural network where each weight is a single value, each weight of a BCNN is a distribution. In our work, we model each weight as a normal distribution with mean value (μ) and standard deviation (σ) are trainable parameters. The BCNN network has three convolutional layers, each is followed by a max-pooling layer (not shown). Extracted features by the convolutional layers are fed to two fully-connected layers. The Bayesian modulator incorporates other relevant data for seizure forecasting into the last fully-connected layer of the network by using Equation (8).

is reflected via scatter plots of AUC between each pair of them (see **Figure 4B**). We did a one-tailed Wilcoxon signed-rank test and a one-tailed t-test between BCNN-ToD and CNN and found that the two methods' performance is not significantly different at the confidence level of 0.05 with p-values of 0.063 and 0.243, respectively. However, the BCNN-ToD provides more insights into how the prediction works. Particularly, because each weight of the BCNN or BCNN-ToD is a distribution, we sample those distributions to calculate the output for each forward pass of an input. By running multiple forward passes of the same input, we can estimate the distribution of the corresponding output. The output's distribution can then be used to quantify the uncertainty of the model's decision-making that will be explored in section 5.
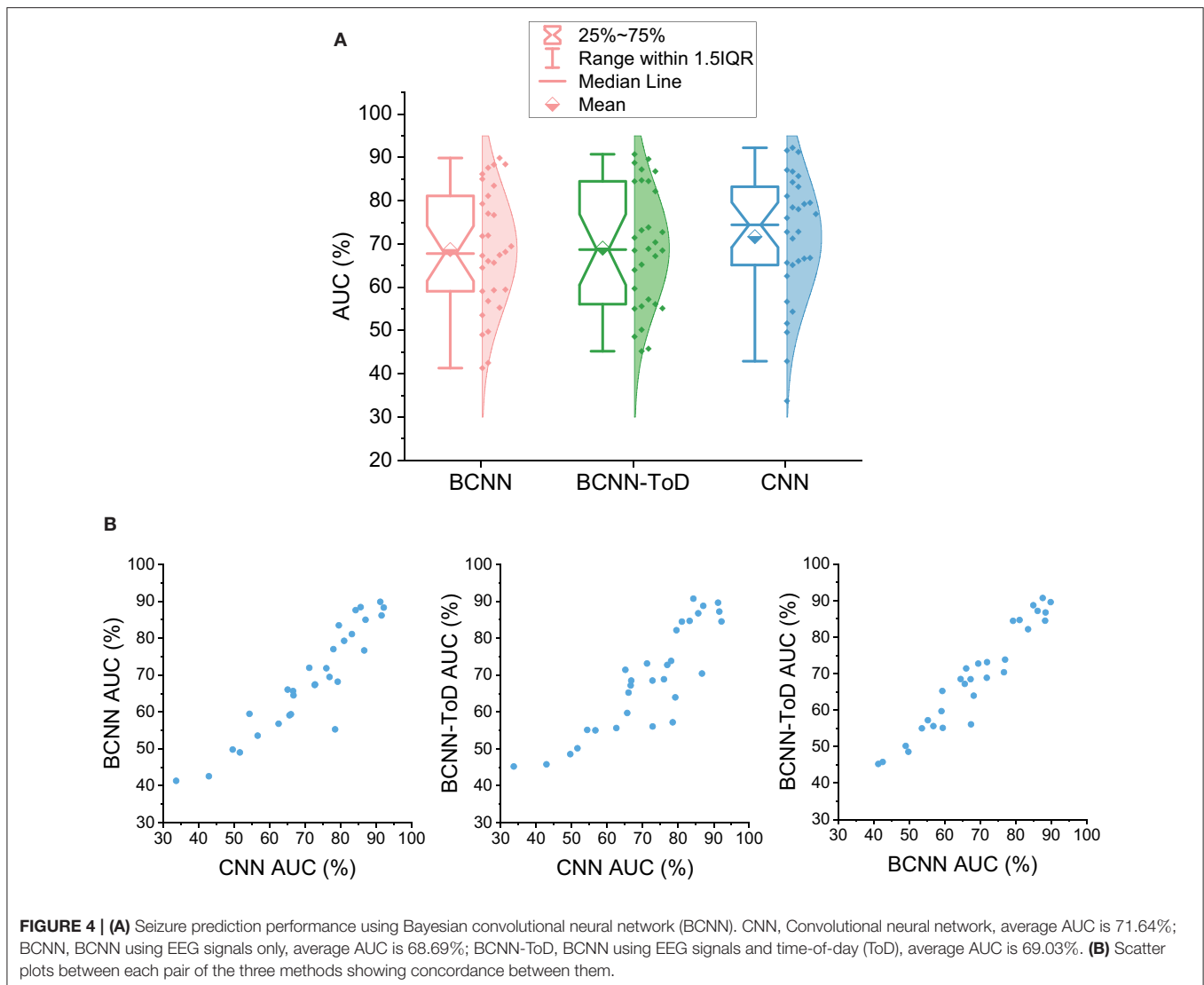
5. DISCUSSION

Bayes convolutional neural network (BCNN) can generate the distribution of its output for each input. We sampled the output of the BCNN by feeding forward the same input through the BCNN 500 times. We quantify the uncertainty level of the BCNN's decision making with Equation (9) below. The numerator takes into account the variability of the output with the standard deviation (std). The denominator considers the case where the output has a uniform-like distribution. Uncertainty

levels of different types of prediction distributions are illustrated in **Figure 5**.

$$\text{Uncertainty level} = \frac{\text{std}_{\text{inference values}}}{\left| \text{mean}_{\text{inference values}} - 0.5 \right|} \quad (9)$$

We trained the BCNN with two types of EEG signals: preictal— 35 to 5 min before seizure onset, and interictal—at least 4 h away from any seizures, we are interested in how the BCNN performs with continuous EEG recording. We ran inference over 13 h of continuous EEG recording for one of the best performers, Patient 4, consisting of two seizures. In **Figure 6**, we plot both the prediction scores (from 0 to 1, where higher values indicate a higher probability of having a seizure) and the corresponding uncertainty levels of the BCNN. In general, the prediction scores get higher values when it is closer to the first seizure onset. Interestingly, at around time 40 and 80 min (around 200 and 160 min before the first seizure onset), there are two predictions with high scores. However, the uncertainty levels were also high, which means that the BCNN "thinks" that there might be a seizure incoming, but it has very low confidence about its decision. From about 1 h before seizure to seizure onset, we can see prediction scores were mostly high, but there were also many low prediction scores with high uncertainty levels. We suggest that the "patterns" or bio-markers related to seizure prediction only

FIGURE 4 | (A) Seizure prediction performance using Bayesian convolutional neural network (BCNN). CNN, Convolutional neural network, average AUC is 71.64%; BCNN, BCNN using EEG signals only, average AUC is 68.69%; BCNN-ToD, BCNN using EEG signals and time-of-day (ToD), average AUC is 69.03%. **(B)** Scatter plots between each pair of the three methods showing concordance between them.

occur at certain particular points in time rather than consistently throughout the whole preictal duration.

Furthermore, we were able to use the trained BCNN as a tool to extract potential bio-markers from EEG signals. We fed 30-s EEG segments to the trained BCNN to sample the output, i.e., run multiple inferences with the same input; in this work, we ran 100 times. Uncertainty level and mean prediction score were extracted from the output samples. For every 30-s segment that has an uncertainty level below 0.1 and means prediction score above 0.9, we extract the attention map over time by accumulating over time the positive values of the feature map of the first convolutional layer. An example is illustrated in **Figure 7**. We observed that, from patients with high performance and with focal seizures, the BCNN focuses on slow EEG activity when performing seizure forecasting. Slow EEG activity has been shown as an important biomarker for studying epilepsy (34).

To verify the possibility of seizure forecasting, we ran the inference over the three patients' EEG recordings with the best seizure prediction performance in the EPILEPSIAE dataset, namely Pat-3, Pat-4, and Pat-12. We used different trained BCNN models at different periods separated by ictal segments to ensure that the trained BCNN being used did not see the current period's preictal segment during training. By doing that, we can have a retrospective risk of having a seizure over time, as shown in **Figure 8**. Generally, the risks are higher when it is closer to the seizure onset, indicating successful seizure forecasting. However, there are cases that we consider as false positive alarms if the risk is assessed as high, but it is too far from the seizure onset, e.g., the day before. For instance, Pat-3 receives a high-risk alarm almost 16 h before the first seizure onset.

Finally, we introduced the Bayesian modulator as a data fusion technique to incorporate relevant auxiliary signals for improving seizure prediction performance. In this work, we
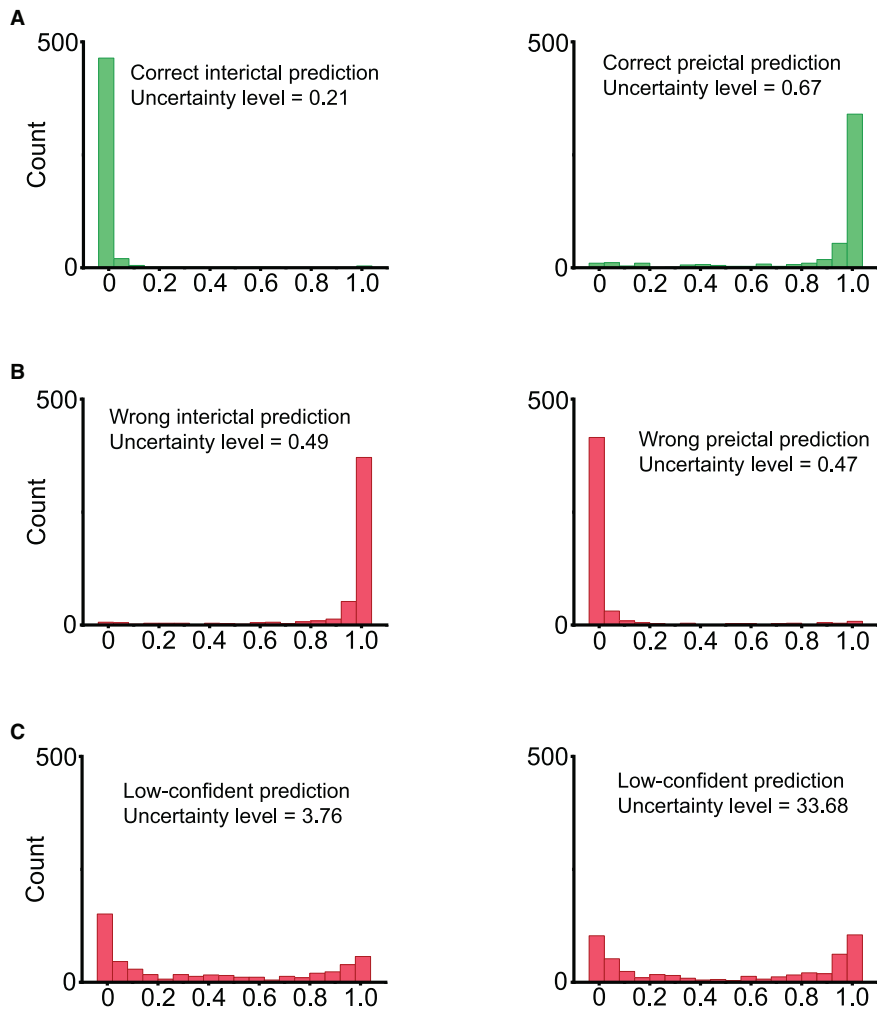
FIGURE 5 | Inference values by sampling the output of Bayesian convolutional neural network 500 times. **(A)** Correct predictions. Left: ground truth is 0 (interictal), most of the model's output samples are close to 0, indicating a correct prediction with high confidence. Right: ground truth is 1 (preictal), most of the model's output samples are close to 1, indicating a correct prediction with high confidence. **(B)** Wrong predictions. Most of the model's output samples are close to the wrong value with high confidence. This is an undesirable case. **(C)** Low-confident predictions with high uncertainty levels. The model's output samples spread randomly between 0 and 1, indicating the high uncertainty of the model.
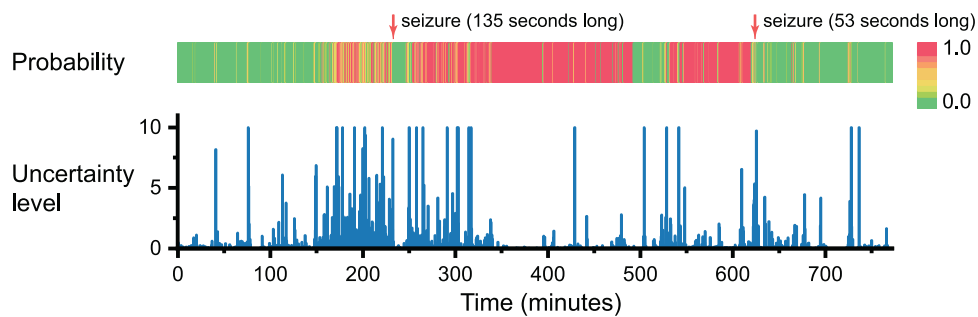
FIGURE 6 | Prediction score and uncertainty level produced by Bayesian convolution neural network in seizure forecasting task over 13 h of continuous EEG recording. For the sake of visualization, for all uncertainty levels higher than 10, we set them to 10.

have only used one extra signal, which is the time of the day, and the performance was not increased significantly. However, we argue that with the Bayesian modulator's capability to embed multiple auxiliary signals, we can achieve a boost in performance which is our aim in future works. Other relevant signals that can be used for data fusion include heart rate variability, blood oxygenation, metabolism. Some signals that have been shown related to seizure onsets, such as electrodermal activity, near-infrared spectroscopy, skin temperature, and respiratory monitor (35), can be used for data fusion as experimental exploration. For example, to use heart rate variability (HRV) for data fusion, one can plot a distribution of HRV during preictal periods (i.e., 35 to 5 min prior to seizure onset), and then apply the kernel density estimation. Lastly, we relied on the assumption that the auxiliary signals are independent of each other and independent of the main signal, i.e., EEG, to derive Equation 7. We are aware that this assumption may not always be entirely met. However, we argue that machine learning models may still work even if the assumptions are weakly met or violated; e.g., in the field of reinforcement learning, Markov property usually is not satisfied, but many models have shown working effectively in practice (36).

6. CONCLUSION

Epileptic seizure forecasting is still a substantially challenging task, but it has a consequential impact on patients' quality of life and their caregivers. While some patient-specific demonstrated excellent performance in a subset of patients, generalized predictions on non-invasive EEG recordings can work well on most patients, which has been a great challenge. This work presented an innovative approach to incorporate uncertainty and auxiliary signals information in seizure-risk forecasting. These informative warning signals will be invaluable for decision-making in employing any risk-mitigation intervention or therapies. We built our method based on the Bayesian convolutional neural network to provide an insight into the uncertainty level of seizure-risk prediction.
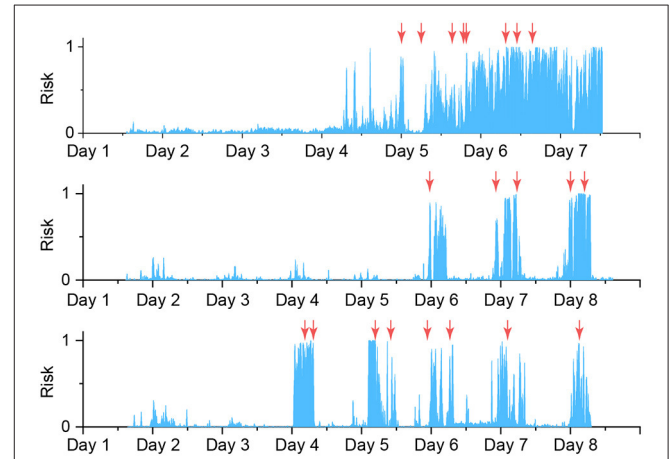

FIGURE 8 | Risk level of having seizures for top three good performance patients, from top to bottom: Pat-3, Pat-4, Pat-12. All decisions with an uncertainty level higher than 1 are discarded and not shown. Moving average of 50 steps backwards is applied to the probability score (risk level). Orange arrows indicate the seizure onsets.

Green to Red indicates increasing interest of the Bayesian neural network on EEG signal
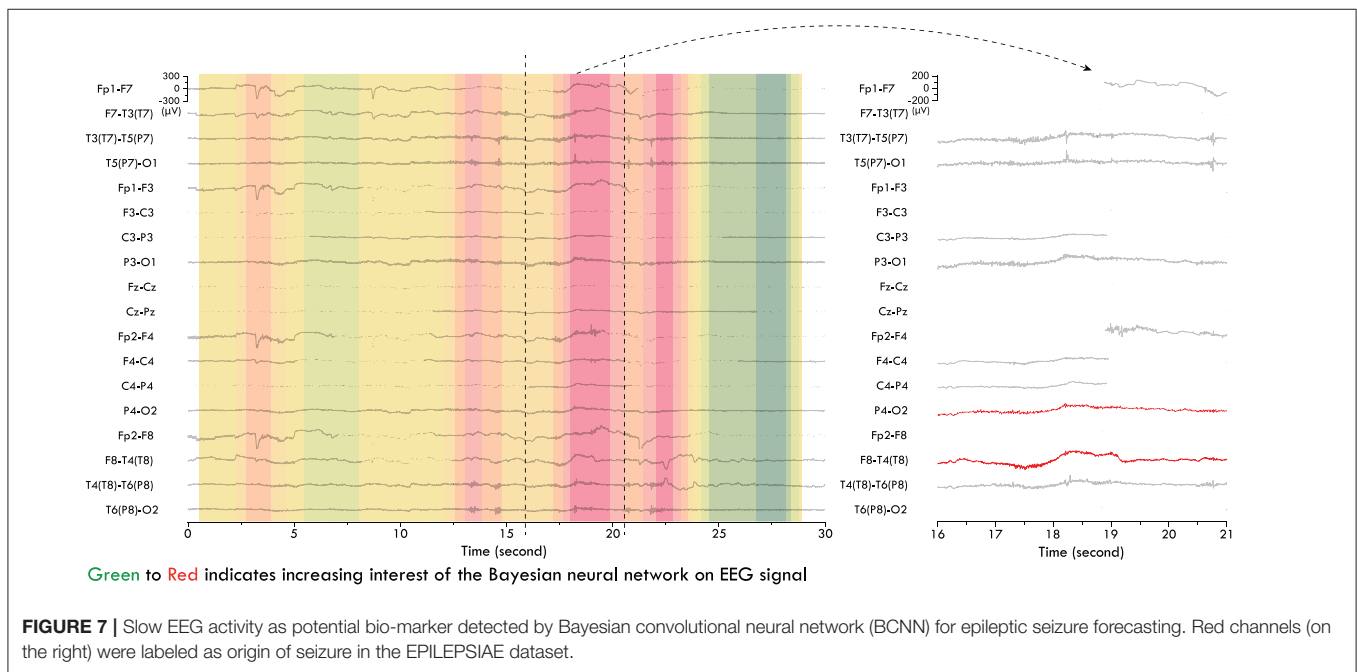
FIGURE 7 | Slow EEG activity as potential bio-marker detected by Bayesian convolutional neural network (BCNN) for epileptic seizure forecasting. Red channels (on the right) were labeled as origin of seizure in the EPILEPSIAE dataset.

CODE AVAILABILITY

The code used to generate all results in this manuscript can be made available upon request.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: European database on epilepsy (http://www.epilepsy-database. eu/).

AUTHOR CONTRIBUTIONS

NT carried out data analysis and wrote the manuscript. OK suggested the project and supervised the work. AN provided clinical and diagnostic information. YY, CM, LK, AM, and OK revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

1. Litt B, Esteller R, Echauz J, D'Alessandro M, Shor R, Henry T, et al. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. *Neuron*. (2001) 30:51–64. doi: 10.1016/S0896-6273(01)00262-8

2. Maturana MI, Meisel C, Dell K, Karoly PJ, D'Souza W, Grayden DB, et al. Critical slowing down as a biomarker for seizure susceptibility. *Nat Commun*. (2020) 11:2172. doi: 10.1038/s41467-020-15908-3

3. Bosl WJ, Leviton A, Loddenkemper T. Prediction of seizure recurrence. A note of caution. *Front Neurol*. (2021) 12:773. doi: 10.3389/fneur.2021.675728

4. Re CJ, Batterman AI, Gerstner JR, Buono RJ, Ferraro TN. The molecular genetic interaction between circadian rhythms and susceptibility to seizures and epilepsy. *Front Neurol*. (2020) 11:520. doi: 10.3389/fneur.2020.00520

5. Zhong D, Luo S, Zheng L, Zhang Y, Jin R. Epilepsy occurrence and circadian rhythm: a bibliometrics study and visualization analysis via CiteSpace. *Front Neurol*. (2020) 11:984. doi: 10.3389/fneur.2020.00984

6. Baud MO, Proix T, Rao VR, Schindler K. Chance and risk in epilepsy. *Curr Opin Neurol*. (2020) 33:163–72. doi: 10.1097/WCO.0000000000000798

7. Williamson JR, Bliss DW, Browne DW, Narayanan JT. Seizure prediction using EEG spatiotemporal correlation structure. *Epilepsy Behav*. (2012) 25:230–8. doi: 10.1016/j.yebeh.2012.07.007

8. Direito B, Teixeira CA, Sales F, Castelo-Branco M, Dourado A. A realistic seizure prediction study based on multiclass SVM. *Int J Neural Syst*. (2017) 27:1750006. doi: 10.1142/S012906571750006X

9. Myers MH, Padmanabha A, Hossain G, de Jongh Curry AL, Blaha CD. Seizure prediction and detection via phase and amplitude lock values. *Front Hum Neurosci*. (2016) 10:80. doi: 10.3389/fnhum.2016.00080

10. Alotaiby TN, Alshebeili SA, Alotaibi FM, Alrshoud SR. Epileptic seizure prediction using CSP and LDA for scalp EEG signals. *Comput Intell Neurosci*. (2017) 2017:1240323. doi: 10.1155/2017/1240323

11. Yang Y, Zhou M, Niu Y, Li C, Cao R, Wang B, et al. Epileptic seizure prediction based on permutation entropy. *Front Comput Neurosci*. (2018) 12:55. doi: 10.3389/fncom.2018.00055

12. Bou Assi E, Gagliano L, Rihana S, Nguyen DK, Sawan M. Bispectrum features and multilayer perceptron classifier to enhance seizure prediction. *Sci Rep*. (2018) 8:15491. doi: 10.1038/s41598-018-33969-9

13. Truong ND, Nguyen AD, Kuhlmann L, Bonyadi MR, Yang J, Ippolito S, et al. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Netw*. (2018) 105:104–11. doi: 10.1016/j.neunet.2018.04.018

14. Liu C, Xiao B, Hsiao W, Tseng VS. Epileptic seizure prediction with multi-view convolutional neural networks. *IEEE Access*. (2019) 7:170352–61. doi: 10.1109/ACCESS.2019.2955285

15. Yu Z, Nie W, Zhou W, Xu F, Yuan S, Leng Y, et al. Epileptic seizure prediction based on local mean decomposition and deep convolutional neural network. *J Supercomput*. (2020) 76:3462–76. doi: 10.1007/s11227-018-2600-6

16. Zhang Y, Guo Y, Yang P, Chen W, Lo B. Epilepsy seizure prediction on EEG using common spatial pattern and convolutional neural network. *IEEE J Biomed Health Inform*. (2020) 24:465–74. doi: 10.1109/JBHI.2019.2933046

17. Truong ND, Kuhlmann L, Bonyadi MR, Querlioz D, Zhou L, Kavehei O. Epileptic seizure forecasting with generative adversarial networks. *IEEE Access*. (2019) 7:143999–4009. doi: 10.1109/ACCESS.2019.2944691

18. Wei X, Zhou L, Zhang Z, Chen Z, Zhou Y. Early prediction of epileptic seizures using a long-term recurrent convolutional network. *J Neurosci Methods*. (2019) 327:108395. doi: 10.1016/j.jneumeth.2019.108395

19. Duan L, Hou J, Qiao Y, Miao J. Epileptic seizure prediction based on convolutional recurrent neural network with multi-timescale. In: *Proceedings of Intelligence Science and Big Data Engineering* (2019) p. 139–50. doi: 10.1007/978-3-030-36204-1_11

20. Shorvon S, Tomson T. Sudden unexpected death in epilepsy. *Lancet*. (2011) 378:2028–38. doi: 10.1016/S0140-6736(11)60176-1

21. Verducci C, Hussain F, Donner E, Moseley BD, Buchhalter J, Hesdorffer D, et al. SUDEP in the North American SUDEP registry: the full spectrum of epilepsies. *Neurology*. (2019) 93:e227–36. doi: 10.1212/WNL.0000000000007778

22. Klatt J, Feldwisch-Drentrup H, Ihle M, Navarro V, Neufang M, Teixeira C, et al. The EPILEPSIAE database: an extensive electroencephalography database of epilepsy patients. *Epilepsia*. (2012) 53:1669–76. doi: 10.1111/j.1528-1167.2012.03564.x

23. EPILEPSIAE. *The European Epilepsy Database* (2012). Available online at: http://www.epilepsy-database.eu/ (accessed July 13, 2021).

24. Blei DM, Kucukelbir A, McAuliffe JD. Variational inference: a review for statisticians. *J Am Stat Assoc*. (2017) 112:859–77. doi: 10.1080/01621459.2017.1285773

25. Hoffman MD, Blei DM, Wang C, Paisley J. Stochastic variational inference. *J Mach Learn Res*. (2013) 14:1303–47.

26. Maiwald T, Winterhalder M, Aschenbrenner-Scheibe R, Voss HU, Schulze-Bonhage A, Timmer J. Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic. *Phys D Nonlinear Phenomena*. (2004) 194:357–68. doi: 10.1016/j.physd.2004.02.013

27. Karoly PJ, Ung H, Grayden DB, Kuhlmann L, Leyde K, Cook MJ, et al. The circadian profile of epilepsy improves seizure forecasting. *Brain*. (2017) 140:2169–82. doi: 10.1093/brain/awx173

28. Zijlmans M, Flanagan D, Gotman J. Heart rate changes and ECG abnormalities during epileptic seizures: prevalence and definition of an objective clinical sign. *Epilepsia*. (2002) 43:847–54. doi: 10.1046/j.1528-1157.2002.37801.x

29. Pavei J, Heinzen RG, Novakova B, Walz R, Serra AJ, Reuber M, et al. Early seizure detection based on cardiac autonomic regulation dynamics. *Front Physiol*. (2017) 8:765. doi: 10.3389/fphys.2017.00765

30. Schwartz TH. Neurovascular coupling and epilepsy: hemodynamic markers for localizing and predicting seizure onset. *Epilepsy Curr*. (2007) 7:91–4. doi: 10.1111/j.1535-7511.2007.00183.x

31. Zhao M, Nguyen J, Ma H, Nishimura N, Schaffer CB, Schwartz TH. Preictal and ictal neurovascular and metabolic coupling surrounding a seizure focus. *J Neurosci.* (2011) 31:13292–300. doi: 10.1523/JNEUROSCI.2597-11.2011

32. Scipy. *Kernel Density Estimation* (2019). Available online at: https://docs.scipy.org.

33. Dunne RA, Campbell NA. On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation function. *Proc Conf Neural Netw.* (1997) 181:185.

34. Tao JX, Chen XJ, Baldwin M, Yung I, Rose S, Frim D, et al. Interictal regional delta slowing is an EEG marker of epileptic network in temporal lobe epilepsy. *Epilepsia.* (2011) 52:467–76. doi: 10.1111/j.1528-1167.2010. 02918.x

35. Ulate-Campos A, Coughlin F, Gaínza-Lein M, Fernández IS, Pearl PL, Loddenkemper T. Automated seizure detection systems and their effectiveness for each type of seizure. *Seizure.* (2016) 40:88–101. doi: 10.1016/j.seizure.2016.06.008

36. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction.* 2nd ed. The MIT Press (2014).