Seizure: European Journal of Epilepsy 71 (2019) 124–131

Contents lists available at ScienceDirect

Seizure: European Journal of Epilepsy

journal homepage: www.elsevier.com/locate/seizure

Automated spectrographic seizure detection using convolutional neural networks

Peter Z. Yan[a,b,*], Fei Wang[b], Nathaniel Kwok[c], Baxter B. Allen[e], Sotirios Keros[f], Zachary Grinspan[a,d]

[a] Department of Neurology, Weill Cornell Medicine, 525 E. 68th St F-610, New York, NY 10065, United States
[b] Department of Health Policy & Research, Weill Cornell Medicine, 402 E. 67th St, New York, NY 10065, United States
[c] Weill Cornell Medical College, Weill Cornell Medicine, 1300 York Ave, New York, NY 10065, United States
[d] Department of Pediatric Neurology, Weill Cornell Medicine, 505 E. 70th St, New York, NY 10021, United States
[e] Department of Neurology, David Geffen School of Medicine at UCLA, 10833 Le Conte Ave, Los Angeles, CA 90095, United States
[f] Department of Pediatric Neurology, University of South Dakota Sanford School of Medicine, 1400 W 22nd St, Sioux Falls, SD 57105, United States

ARTICLE INFO

Keywords:
Video-EEG monitoring
Electroencephalogram quantitative analysis
Critical care
Computer application in medicine

ABSTRACT

Purpose: Non-convulsive seizures are common in critically ill patients, and delays in diagnosis contribute to increased morbidity and mortality. Many intensive care units employ continuous EEG (cEEG) for seizure monitoring. Although cEEG is continuously recorded, it is often reviewed intermittently, which may delay seizure diagnosis and treatment. This may be mitigated with automated seizure detection. In this study, we develop and evaluate convolutional neural networks (CNN) to automate seizure detection on EEG spectrograms.
Methods: Adult EEGs (12 patients, 12 EEGs, 33 seizures) from New-York Presbyterian Hospital (NYP) and pediatric EEGs (22 patients, 130 EEGs, 177 seizures) from Children's Hospital Boston (CHB) were converted into spectrograms. To simulate a telemetry display, seizure and non-seizure events on spectrograms were sequentially sampled as images across a detection window (26,380 total images). Four CNN models of increasing complexity (number of layers) were trained, cross-validated, and tested on CHB and NYP spectrographic images. All CNNs were based on the VGG-net architecture, with adjustments to alleviate overfitting.
Results: For spectrographically visible seizures, two CNN models (containing 4 and 7 convolution layers) achieved > 90% seizure detection sensitivity and specificity on the CHB test set and > 90% sensitivity and 75–80% specificity on the NYP test set. The one CNN model (10 convolution layers) did not converge during training; while another CNN (2 convolution layers) performed poorly (60% sensitivity and 32% specificity) on the NYP test set.
Conclusions: Seizure detection on EEG spectrograms with CNN models is feasible with sensitivity and specificity potentially suitable for clinical use.

1. Background

Non-convulsive seizures (NCS) are common in critically ill patients (8–50%) [1–6], and if untreated, are associated with high mortality (17–51%) [6–8]. Furthermore, effective treatment must be timely, as delayed treatment can lead to long term neurologic disability [9–11]. Many intensive care units (ICU) now perform continuous EEGs (cEEG) to monitor for NCS. However, although continuously recorded, cEEGs are rarely continuously monitored. Even in large academic medical centers, cEEGs are typically reviewed 2–3 times daily by a neurophysiologist. In some centers, this may be supplemented by an EEG technician or nurse who provide more frequent, but nonetheless intermittent, monitoring. This intermittent review model can lead to delays in diagnosis and treatment, because a patient may have a non-convulsive or subtle seizure that can remain undetected for several hours until the next review [12].

One potential solution is to transform the EEG into a more readily interpretable form (e.g., a spectrogram) through quantitative EEGs (qEEG) techniques. This enables the bedside clinician to rapidly review the EEG and decide appropriate clinical management. A well-studied qEEG modality is the compressed spectral array (CSA). CSAs transform the clinical EEG (18 waveforms, around 15 s epochs per screen) into colored spectrograms (2–8 spectrograms, 1–2 hour epochs per screen) [13,14]. Despite this considerable reduction in visual complexity,

* Corresponding author at: 525 E68th St F610, New York, NY 10065, United States.
E-mail address: PZY9001@med.cornell.edu (P.Z. Yan).

https://doi.org/10.1016/j.seizure.2019.07.009
Received 21 May 2019; Received in revised form 5 July 2019; Accepted 6 July 2019

interpreting multiple spectrograms simultaneously still remain a complex task. Furthermore, seizure detection by bedside physicians (non-neurophysiologists) using CSAs vary widely (65–93% sensitivity and 68–95% specificity) depending on seizure duration and intensity [14,15], and may require lengthy training up to 6 h [16].

An alternative to the traditional CSA is the median power spectrogram (MPS) [17]. The MPS condenses the EEG into 1–3 spectrograms, and more clearly delineates the sloped harmonic banding pattern associated with the evolving rhythmic electrical activity that occur during electrographic seizures [18]. Seizures are more visually distinct on MPS, and clinicians need only brief (5 min) training to recognize them [17]. The brief training does not sacrifice performance, as physicians detect seizures on the MPS with comparable performance to studies using more complex qEEGs with longer training (65–86% sensitivity, 62–80% specificity) [13,14,19–21]. The visual distinctiveness of seizures on the MPS suggest that it is well suited as a training substrate for automated seizure detection methods.

A problem common to all qEEG interpretation is user inconsistency. Most spectrograms, including the MPS, demonstrate only moderate agreement among users (Cohen's Kappa 0.41–0.60) [14,17,20]. Furthermore, in typical clinical scenarios, the neurophysiologist and the bedside clinician cannot continuously monitor the qEEG. Thus, an automated mechanism that monitors and alerts bedside staff of potential seizures in a consistent and rapidly manner would be valuable.

In this paper, our objective is to present an approach to automate seizure detection on spectrograms using Convolutional Neural Networks (CNN) [22]. We hypothesis that since the spectrogram is essentially an image, seizure detection on spectrograms can be conceptualized as an image classification task, which can be achieved with sufficiently trained CNNs. Image classification using CNNs has demonstrated success in medical imaging applications [23,24], but not been applied to EEG spectrograms. Because seizures appear distinctive on the MPS, we use snapshots of these spectrographic seizures as training images for the CNNs. The trained CNNs can then recognize spectrographic patterns that resemble seizures, and thus detect them automatically. After training the CNN models, we test their seizure detection performance on EEG spectrograms in both adult and pediatric patients from two large academic medical centers.

2. Materials and methods

2.1. Study overview

We developed and evaluated CNN models to detect seizures using retrospectively collected EEGs from Children's Hospital Boston – Massachusetts Institute of Technology (CHB-MIT) and New York Presbyterian – Weill Cornell Medical Center (NYP-WC). This study was approved by the Weill Cornell Institutional Review Board.

From CHB-MIT EEGs, 177 seizures and 177 randomly selected non-seizure epochs were converted to spectrograms. The EEGs were converted into spectrograms via the MPS method [17]. 90% of these seizure and non-seizure spectrograms were randomly selected to train and cross validate four different CNN models. The other 10% was set aside to test CNN performance after training. Images were then sampled from spectrograms at seizure and non-seizure locations in a manner that simulated telemetry monitoring (see below, Spectrogram Sampling). These sampled images (n = 21,900) were labeled as seizure or non-seizure based on which spectrogram they were sampled from. They were then directly supplied to the CNN models for training and testing (Fig. 1a).

From NYP-WC EEGs, 33 seizures and 24 randomly selected non-seizure epochs were converted to spectrograms, and sampled into images (n = 3160), in the same manner as the CHB-MIT EEGs. However, NYP-WC EEGs were used for testing the CNN models only (Fig. 1b). Furthermore, because some seizures were EEG visible, but not spectrographically visible, a subset of NYP-WC EEGs containing only spectrographically visible seizures were separately tested (Fig. 1c).

2.2. EEG data

The CHB-MIT EEGs were acquired from PhysioNet.org [25,26]. The NYP-WC EEGs were acquired from the NYP-WC clinical EEG database. The mean seizure duration for both data sets was 60 s (6 s–12.5 min).

CHB-MIT EEGs were collected from 22 patients (ages 1.5–19). The waveforms were sampled at 256 Hz and included annotations of the seizure's start and end identified by neurophysiologists at CHB. Annotations were further verified by a neurophysiologist (ZG) at WC. There were 130 EEGs with 177 seizures, and 549 EEGs without seizures in this data set.

NYP-WC EEGs were collected from a convenience sample of 12 patients (ages 18–99). The waveforms were sampled at 256 Hz, and included seizure start and end times identified by the neurophysiologist at the time of care. These annotations were further independently verified by two neurophysiologists (BA and SK). There were 12 EEGs containing 33 seizures. All EEGs contained at least one seizure.

2.3. Spectrogram sampling

EEGs were converted to MPS, with a single spectrogram representing EEG signal from all four scalp quadrants [17]. Samples of gray-scale MPS images containing seizures are shown in Fig. 2. For each image, height of a single pixel corresponded a frequency bin of 0.125 Hz (i.e. 0.125 Hz/pixel; F_s [sampling rate] = 265, n [length of Fast Fourier transform] = 2048, F_s/n = 0.125 Hz). Each image included spectral power in the 0–20 Hz range and were thus 160px (20 Hz/[0.125 Hz/px] = 160px) in height. The MPS was sampled with a 120 s window, with the seizure/non-seizure event advancing across at 1 s increments. Thus, width of a single pixel represented 1 s of elapsed time. All sampled images were 160 × 120 pixels, i.e. spectral between 0 and 20 Hz over 120 s.

Spectrographic images containing seizures were sampled starting with the seizure's leading edge in the middle of the 120 s window. As the seizure's leading edge traveled across the sampling window at 1 s increments, a snapshot of the window was taken, resulting in 60 images per seizure (Fig. 2a). We used this sampling method to simulate telemetry monitoring where a spectrogram would travel across a 120 s seizure detection window. Furthermore, we selected a 120 s window to coincide with the maximum clinically recommended duration to initiate treatment on a patient with continuous or near continuous seizures [27]. Note that with our method, for seizures > 60 s in duration, only the first 60 s were sampled.

Spectrogram images without seizures were sampled as above, but with two different methods in selecting the starting locations between the CHB-MIT and NYP-WC spectrograms. For CHB-MIT spectrograms, a random start location was set on each of 177 randomly selected spectrograms without seizures, and 60 images sampled from each location (Fig. 2b). For NYP-WC spectrograms, because all spectrograms contained at least one seizure, seizure-free images were sampled from spectrogram epochs before the first seizure and after the last seizure. The sampling locations were randomly selected from epochs such that the detection window would not overlap with any seizures (Fig. 2c).

2.4. Spectrogram review

While all electrographic seizures are by definition EEG visible, there is no guarantee that all seizures will be spectrographically visible (very subtle seizures on the EEG may blend into the spectrogram background). Therefore, all spectrograms in this study were also reviewed for seizure visibility. From our previous work, all 177 CHB-MIT seizures were visible on MPS [17]. For the NYP-WC EEGs, all spectrograms were initially unlabeled and underwent blinded review by a neurophysiologist (PY). 17/33 seizures were not visible on the spectrogram, with 8/

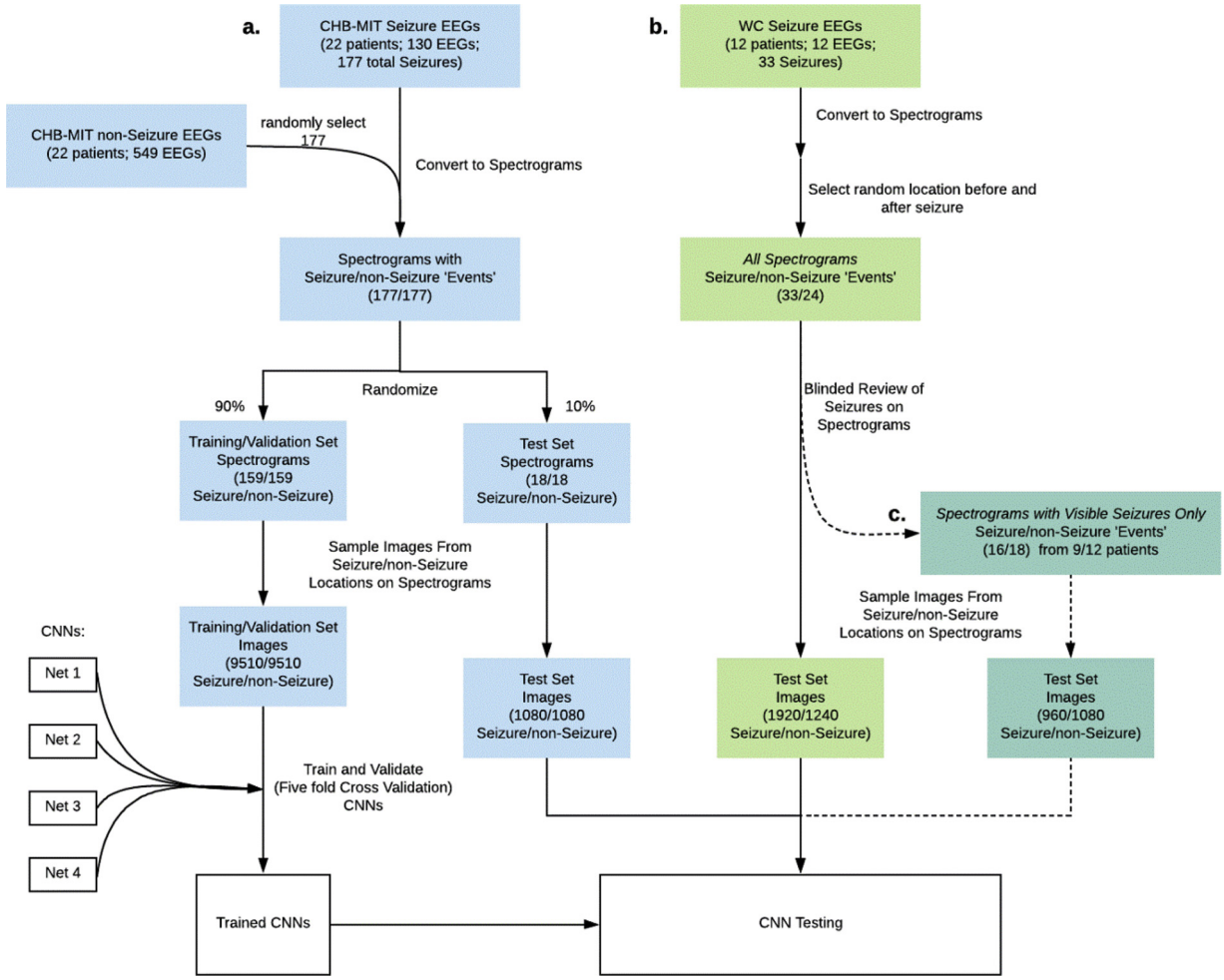

Fig. 1. Study design. (a) CNNs were trained and cross-validated using data from CHB-MIT. 90% of CHB-MIT was allocated for CNN training and cross-validation, with the remainder for testing CNN performance. (b) All data from WC were used for testing only. (c) Seizures were only visible in a subset of WC spectrograms (dashed arrows), and CNNs were tested separately on this subset.

17 seizures from one patient.

2.5. Spectrogram image partitioning

Spectrogram images were partitioned so that CHB-MIT images were used for CNN training/cross-validation (90%) and testing (10%), whereas all NYP-WC images were used for testing only. Images were partitioned based on seizures (i.e. images belonging to an individual seizure were partitioned together into the same group). Because not all seizures were visible on the NYP-WC images, a subset containing only spectrogram-visible seizures was also created (Fig. 1).

2.6. CNN architecture

The CNN is a specific type of deep learning neural network model. It is composed of sequential layers of convolution and sub-sampling operations [28]. There are many CNN architectures that differ in the composition and connections among different layers. In this study, we use VGG-net, a well-known CNN architecture. We selected VGG-net due to its modular design, consisting of convolution and sub-sampling operations grouped in repetitive blocks. The VGG-net is a high performing CNN model in the ImageNet classification task [29]. There were four CNN models used in this study. All are based on VGG-net, and each contained a different numbers of convolution blocks (2–3 convolutional layers per block) connected in series to a classification block (Fig. 3). The output of each convolution block is a set of 'higher-level' feature representations that describe the block's input (feature representations from the previous convolution block), which backtracking through the CNN, is derived from the input image. This new representation of the input image can then be further processed by additional convolutional blocks, until the final 'optimal' feature representation of the input image is learned in the classification block (a fully connected artificial neural net that functions as a flexible non-linear classifier) [30]. This process is remarkably similar to the image processing in the human visual cortex [31].

2.7. CNN training/validation

The CNN model's convolution blocks can be conceptualized as a series of weighted functions that extract the 'optimal' features describing a spectrographic seizure. The training process is to determine the optimal weights for these functions. All CNNs were trained using stochastic gradient descent with momentum (batch size = 256, $\gamma = 0.9$, learning rate = 0.01). To alleviate overfitting, L2 regularization (with regularization parameter 0.0005) was used and 0.5 dropout rate was used in the fully connected (FC) layers during classification. All training proceeded for 50 epochs in all models, at each cross-validation step. After cross-validation, the CNN was trained on all samples from the training/validation set. All CNNs were trained with MATLAB 2017a (Mathworks, Natick, MA) using a Titan Xp GPU (NVIDIA, Santa Clara, CA).

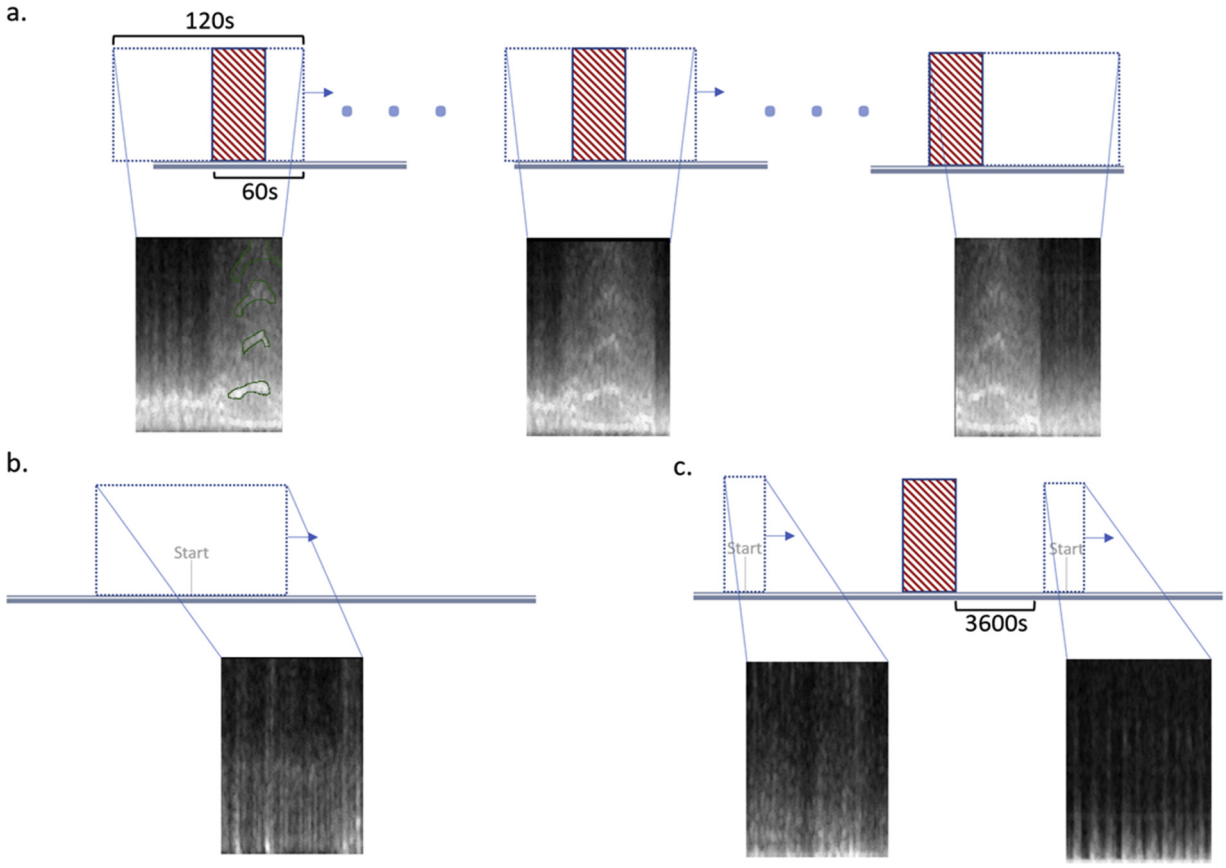

Fig. 2. Sampling of seizure and non-seizure spectrogram images from the MPS. Images were mapped to gray scale, with brighter pixels indicating more power. For visual analysis, a color map would be applied for better visualization; however, for computer analysis, the choice of color map is irrelevant. Note that the sloped banding pattern specific to seizures is outlined in green. (a) For MPS containing a seizure (filled box), snap shots began with the seizure's leading edge at the middle of the 120 s wide window (dashed box). The seizure then advanced across the window at one second increments until the seizure's leading edge reached the end of the window. A snapshot was taken at each 1 s stride, resulting in 60 images per seizure. (b) For MPS from CHB-MIT EEGs not containing seizures, a random location was selected, and spectrogram images were sampled in the same manner as in (a). (c) Because all NYP-WC EEGs contained seizures, to sample the non-seizure spectrogram images, random non-seizure locations before and after the first and last seizures of each EEG were selected and sampled in the same manner as (a). Note that the second non-seizure location was selected one-hour post-seizure. This was to avoid sampling any residual post-ictal EEG change. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

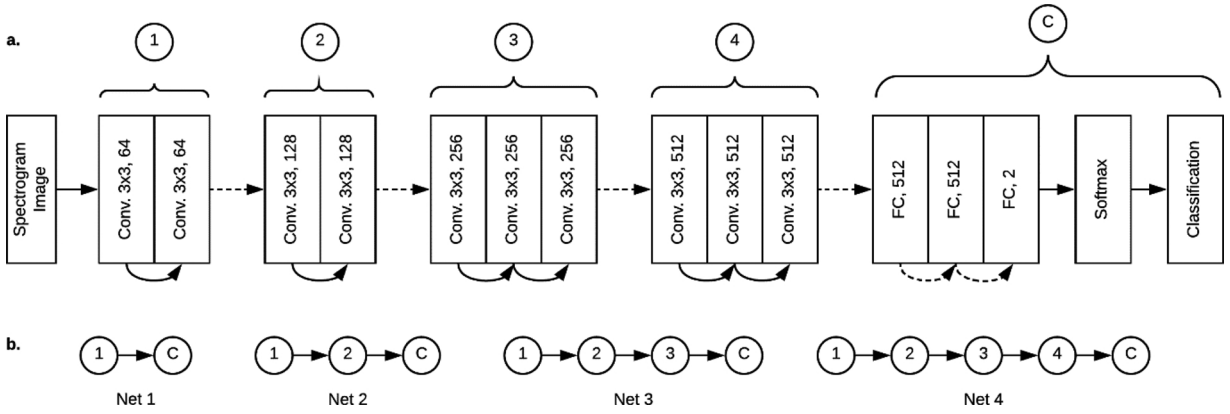

Fig. 3. CNN schema. (a) The CNNs in this study is based on the VGG net architecture and consists of convolution blocks (1–4) connected to a classification block [C]. Within each convolution block, individual convolution layers have a 3 × 3-pixel convolution filter applied at 1-pixel steps, and multiple filters are applied at each layer in each block (64, 128, 256, 512). The result from each layer is then passed through rectified linear units to the next convolution layer (solid curved arrows). Results from each convolution block then passes to the next block via 2 × 2 max pooling (i.e. subsampling; dashed straight arrows). Finally, features from the convolution blocks are passed to the classification block, which consists of full connected (FC) units of artificial neuron layers with dropout between layers (dashed curved arrows). The output is then passed to a softmax layer, which finally leads to classification of the initial spectral image as containing a seizure or not. (b) The four different CNNs used in this study, each net contains more convolution blocks and extracts increasingly 'higher level' descriptors before passing them to the classification block.

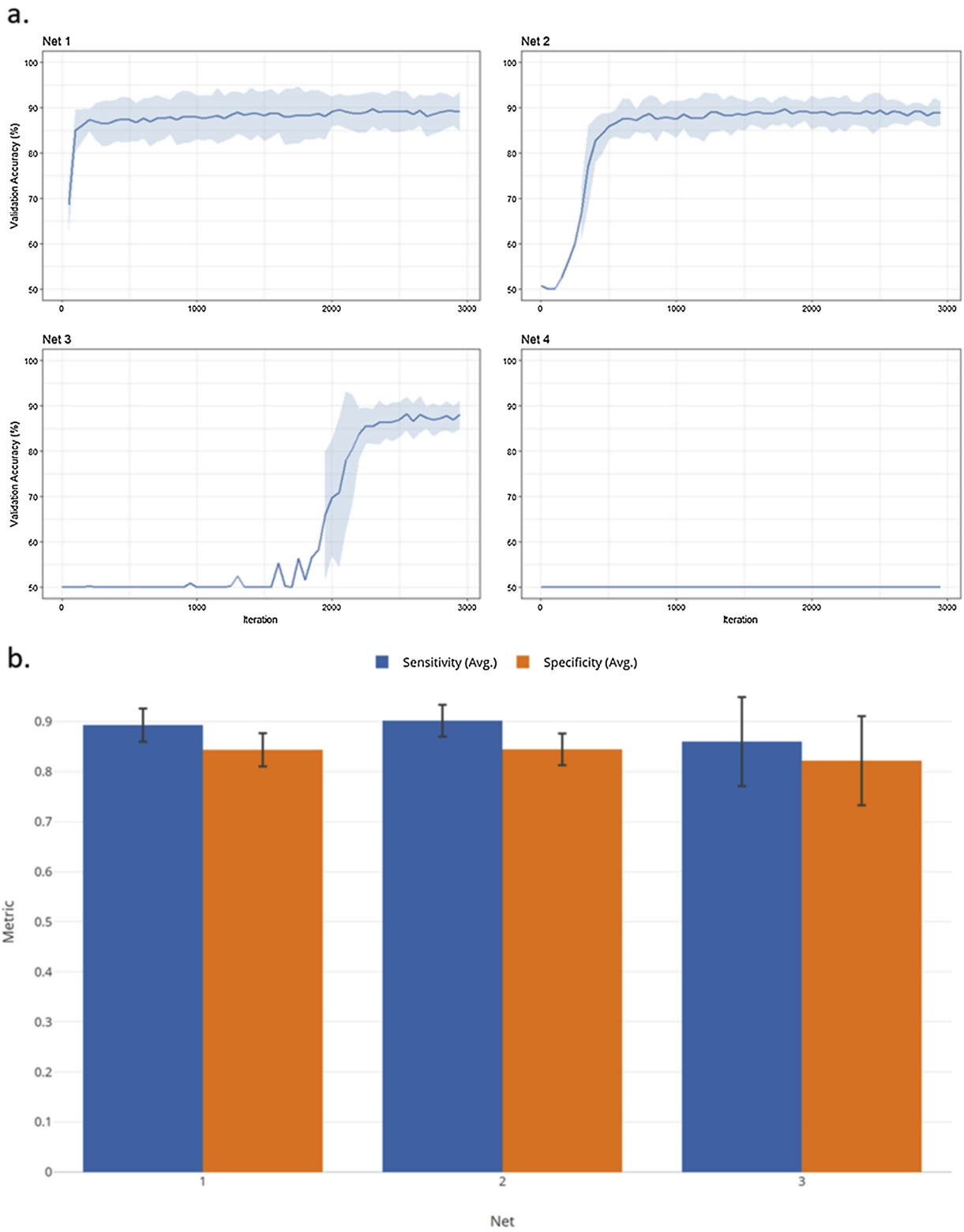

Fig. 4. CNN training and 5-fold cross validation results. (a) Average validation accuracy across the five folds during cross validation (blue ribbon indicates 95% confidence interval [CI]). Each CNN was trained for 50 epochs (just under 3000 iterations). Net 4 never converged to a solution. (b) Average cross-validation sensitivity and specificity across the 5 folds for nets 1–3 (error bars denote 95% CI). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.8. CNN testing

The trained CNN was used to detect seizures from three test sets: the CHB-MIT test set, the NYP-WC test set with all spectrograms, and the NYP-WC test set with only spectrographically visible seizures. Detection performance was calculated at the level of individual seizures, not individual images. The presence of a seizure was determined by the number of *N*-consecutive images classified as containing a seizure (e.g. for $N = 10$, 10 images must be consecutively classified as containing a seizure before flagging the event as a seizure). *N* was varied between 1

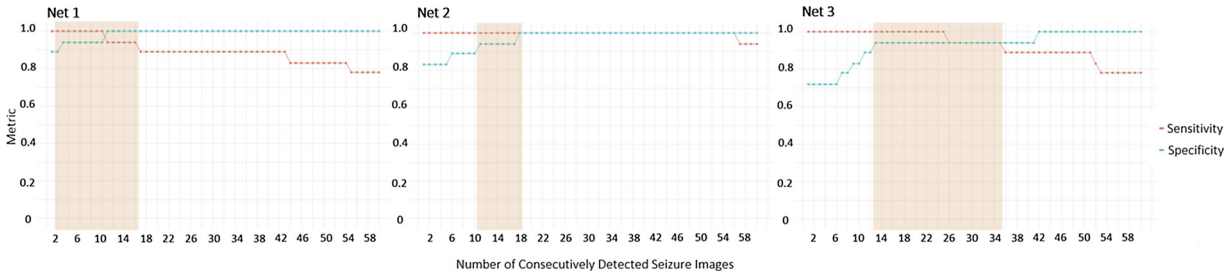

Fig. 5. CNN performance on the CHB-MIT test set. A positive seizure detection by the CNN is defined as N-consecutive spectrogram images having been classified as containing a seizure (e.g. for N = 10, 10 images must be consecutively detected as containing a seizure before the event is flagged as a seizure). Varying N yields a range of seizure detection sensitivity (red) and specificity (green). For all three CNNs, there were ranges of N (highlighted regions) that may be suitable for seizure screening (> 90 sensitivity, ≥75 specificity). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and 60, with sensitivity and specificity calculated at each *N*.

3. Results

Three of four CNNs converged during training. Net 4, the neural net with the most convolutions (4 convolution blocks/10 layers), did not converge during training and five-fold cross-validation. (Fig. 4a) The cross-validation performance for the remaining three CNNs are summarized in Fig. 4b. All three CNNs (nets 1–3) had average sensitivity and specificity between 80–90%, though the 95% confidence interval was wider for net 3.

For the CHB-MIT test set, seizure detection performance is shown in Fig. 5. For neural nets 1–3, seizure detection performance varied depending on the number of *N*-consecutive images identified as containing a seizure (e.g. for $N = 10$, the CNN must classify 10 consecutive spectrogram images as containing a seizure before flagging the event as a seizure). For each CNN, there was a range of *N* where seizure detection sensitivity and specificity were > 90%. This range was wider for net 2 ($N = 11–56$) and net 3 ($N = 13–35$) compared with net 1 ($N = 3–10$).

For the NYP-WC test set, in the subset of spectrographically visible seizures, net 2 and 3 had a range of *N* that yielded > 90% sensitivity and ≥75% specificity (Fig. 6a, highlighted regions). This range of *N* was much narrower compared to their CHB-MIT performance (net 2: $N = 8–10$ vs. 11–56 and net 3: $N = 5–8$ vs.13–35). Furthermore, for net 1, there was no range of *N* that yield similar performance as nets 2 and 3. Notably in this NYP-WC test subset, seizure detection sensitivity for nets 2 and 3 remained > 90%, but there was a specificity decrease to 75–80%, when compared to the models' prior CHB-MIT test set performance. When all NYP-WC spectrograms (including those without spectrographically visible seizures) were used as the test set, there was no range of *N* that yielded ≥75% sensitivity and specificity for any of the CNNs (Fig. 6b).

4. Discussion

4.1. CNN performance

The CNN models in this study achieved > 90% sensitivity and specificity in seizure detection when tested on pediatric CHB-MIT spectrograms, and > 90% sensitivity and 75–80% specificity in seizure detection when tested on adult NYP-WC spectrograms with spectrographically visible seizures. This demonstrates performance potentially suitable for clinical applications. While encouraging, there remains several limitations in this novel approach to seizure detection.

While nets 1–3 converged during training, net 4 did not. This is likely related to image complexity of the spectrographic seizures and the number of convolution layers in the neural net. The sloped banding pattern (Fig. 2a) characteristic of spectrographic seizures is composed of medium level image features (combinations of simple edges, corners,

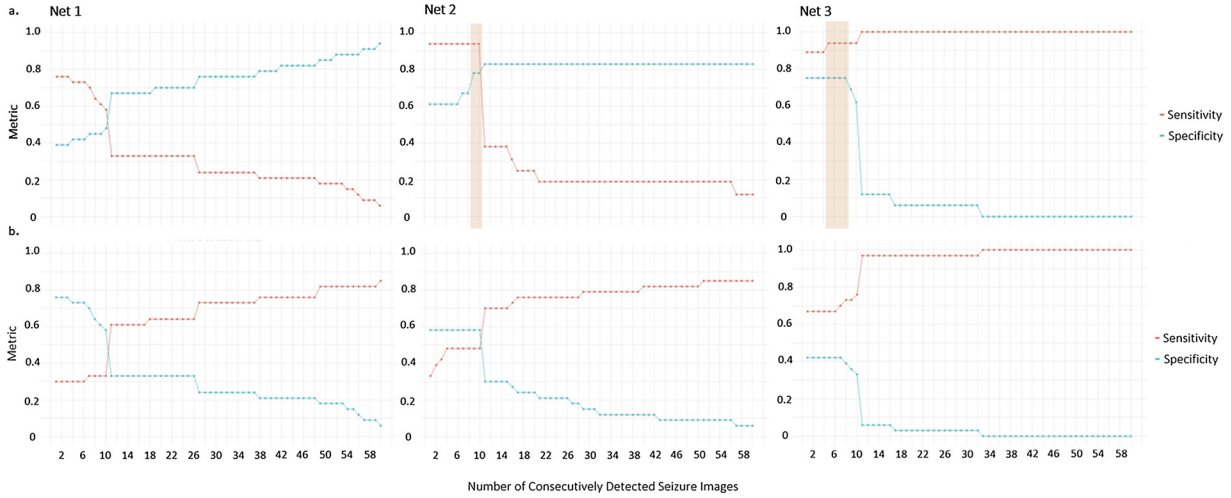

Fig. 6. CNN performance on the NYP-WC test set. Varying N yields a range of seizure detection sensitivity (red) and specificity (green). (a) Performance in the subset containing only spectrographic visible seizures. Potentially suitable performance for seizure screening (> 90 sensitivity, ≥75 specificity) of N was only present in net 2 & 3 (highlighted regions) and were narrower compared to their corresponding performance on the CHB-MIT test set in Fig. 5. There was also 10–15% decrease in specificity. (b) Performance on all NYP-WC spectrograms, regardless of seizure visibility. The performance diminished markedly across all CNNs, which is expected, as the CNN will not be able to detect a seizure if the seizure itself is not visible on the spectrogram. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and shading). In CNNs, each subsequent convolution block effectively extracts a 'high level' feature summary of the features from the previous block. This 'summarization' melds feature details from the previous convolution block in an attempt to produce new features that are more broadly descriptive. However, after too many convolution blocks, the produced features are too broadly descriptive and insufficiently specific for the CNN to achieve its classification task [32,33]. The CNN then does not converge on a solution (a set of weights that when applied to the final features allows the neural net to distinguish between seizure and non-seizure images). This is likely the reason for non-convergence in net 4, the CNN containing the most convolution blocks.

While performance on the CHB-MIT test set was comparable to the cross-validation results, performance on the NYP-WC test set was less specific, and with a narrower suitable seizure screening margin (the number of *N*-consecutive images detected as containing a seizure before the event is flagged as a seizure). This suggests overfitting because all CNN models were trained on the CHB-MIT data, and when tested on data from a very similar source (i.e. the same institution) the CNN performs well, but when tested on data from a different source (NYP-WC), the model suffers performance loss. Overfitting is a common problem in neural nets, and it persisted in our CNNs despite using dropout and L2 regularization to mitigate its effects. There are three likely reasons for overfitting in our study. First, the seizures in our training set originate from a small patient cohort and the number of seizures is unevenly distributed across patients, with some patients contributing disproportionally more seizures. While the sloped banding pattern is prototypical of spectrographic seizures, most patients have specific versions of these patterns. Thus, CNNs trained on CHB-MIT spectrograms, while generally able to recognize the sloped bands, are predisposed to detect banding patterns similar to those found in the CHB-MIT spectrograms. Second, the CHB-MIT and NYP-WC patients have different demographics (children aged 1.5–19 vs. adults aged 18–99). Children and adults have different prevailing seizure etiologies and background rhythms, which can systematically affect the seizure's EEG appearance. This then leads to more variation in spectrographic morphology between CHB-MIT and NYP-WC seizures, and decreases performance when a pediatric trained CNN model is tested on adult data. Third, the CHB-MIT training spectrograms and NYP-WC test spectrograms are from two different institutions. Minor differences in EEG acquisition (e.g. variations in electrodes, electrode gels, and other acquisition techniques) between institutions may introduce small but systematic differences between CHB-MIT and NYP-WC spectrograms.

While spectrograms have demonstrated effectiveness in seizure screening [34], some subtle seizures may be only minimally visible or not spectrographically visible at all. This is a notable limitation of seizure detection on spectrograms, regardless of human or computer screening. Unsurprisingly, neural net performance suffers considerably when spectrogram without visible seizures were retained in the test set. The practical implication is that a neurophysiologist, if available, should still intermittently review the underlying EEG waveforms while the automated detection provides continuous screening.

4.2. Clinical relevance

cEEG monitoring is resource intensive and in demand. In a large (n = 97) multicenter survey of tertiary care centers, 18% of ICU physicians would have increased cEEG duration had more resources been available, and at 17% of these institutions, there was no attending neurophysiologist EEG interpretation overnight. Furthermore, demand for cEEG is increasing, with 43% of institutions reporting an increase in cEEGs per month compared to the prior year [35]. Automated seizure detection with CNNs can help address these issues by either supplementing seizure screening for the neurophysiologist between reviews, or providing standalone telemetry seizure monitoring for the bedside physician.

For the bedside physician, the MPS offers a concise EEG visualization, and the CNN provides automated telemetry monitoring for seizures. In our CNN training and testing, images were sampled as the spectrogram moved across a 120 s window, which simulated telemetry data acquisition. Acquisition of these seizure 'snapshots' for CNN training only began when the seizure's leading edge had reached the middle of the window. This was done to provide training images with ample seizure content. In practice, this means that the CNN may not detect a seizure until 60 s after it had initially occurred. This is clinically acceptable as most seizures self-remit between 30–60 s [36], and it is recognizing seizure occurrence and initiating treatment within a reasonable time window (on the order of minutes to tens of minutes) that lead to improved outcomes [37]. Furthermore, CNN performance (> 90% sensitivity and 75–80% specificity) is comparable to clinician (non-neurophysiologist) performance in detection (65–85% sensitivity and 62–80% specificity) on the MPS [17].

For the neurophysiologist, our CNN approach can assist in seizure screening during cEEG monitoring. In our study, the CNN was trained on MPS based spectrographic images, but the same approach can be applied to other types of spectrograms. Since most qEEG implementations include at least one spectrogram, and many neurophysiologists already use qEEGs during their review [34,35]. Our CNN based approach can augment existing workflow by detecting potential seizures and highlighting them as areas of interest for the neurophysiologist. This may increase review speed, which would alleviate the increasing demand for more cEEG monitoring. Additionally, *N* (the threshold of consecutive images that must contain an event prior to declaring the event a seizure) is a parameter that can be set on an individual basis. This allows for a neurophysiologist to tune the detection threshold in applicable cases, thereby mitigating potential over-alerting or alert-fatigue, which would be particularly helpful in ICU patients with periodic/rhythmic EEG patterns equivocal for seizures or patients prone to frequent seizures.

The CNN's real-world performance will be influenced by the patient population's underlying seizure prevalence. In critically ill patients, the prevalence ranges 8–50% [1–6], and assuming the CNN's lower end performance (90% sensitivity, 75% specificity), this translates to a positive predictive value (PPV) of 25–78% and negative predictive value (NPV) of 88–98%. The high NPV indicates that the CNN is better suited for seizure screening. Additionally, the wide PPV range underscores the clinician's role in judiciously selecting patients for cEEG, as those patients with higher seizure likelihood will derive more benefit from cEEG monitoring, which applies to all cEEG scenarios, regardless of human or computer-based seizure screening.

4.3. Future work

While the CNN models in this study was trained and tested on 26,380 total spectrographic images, our study is limited by its small number of patients from which the images were derived. A larger patient sample with stratification of different patient groups in a future study is warranted. In this study, we hypothesize that the CNN is looking for the distinct sloped banding pattern as the differentiator between seizure and non-seizure events. However, as with all neural nets, the CNN's actual features of interest are not explicit. Explicitly revealing these features will involve application of a deconvolution neural net in a future study [38]. Finally, we selected the VGG-net architecture due to its modular design, but it may not be the optimal architecture for our application. Both the neural net architecture as well as its hyperparameters can undergo evolution and optimization [29], and the optimal neural net architecture for this application is yet to be determined.

5. Conclusion

Automated seizure detection on EEG spectrograms using CNN models is feasible and may be suitable for seizure screening. Further

study with a more diverse patient sample and additional neural net optimization is warranted.

Declarations of interest

None.

Acknowledgements

Support for this study was provided by Weill Cornell Medicine, Department of Neurology (seed funding to PY), and special thanks to NVIDIA Corporation with the donation of the Titan Xp GPU used for this study.

References

[1] Laccheo I, Sonmezturk H, Bhatt AB, Tomycz L, Shi Y, Ringel M, et al. Non-convulsive status epilepticus and non-convulsive seizures in neurological ICU patients. Neurocrit Care 2015;22:202–11. https://doi.org/10.1007/s12028-014-0070-0.

[2] Towne AR, Waterhouse EJ, Boggs JG, Garnett LK, Brown AJ, Smith JR, et al. Prevalence of nonconvulsive status epilepticus in comatose patients. Neurology 2000;54:340–5. https://doi.org/10.1212/WNL.54.2.340.

[3] Claassen J, Mayer Sa, Kowalski RG, Emerson RG, Hirsch LJ. Detection of electrographic seizures with continuous EEG monitoring in critically ill patients. Neurology 2004;62(10):1743–8. https://doi.org/10.1212/01.WNL.0000125184.88797.62.

[4] Vespa PM, Nuwer MR, Nenov V, Ronne-Engstrom E, Hovda DA, Bergsneider M, et al. Increased incidence and impact of nonconvulsive and convulsive seizures after traumatic brain injury as detected by continuous electroencephalographic monitoring. J Neurosurg 1999;91:750–60. https://doi.org/10.3171/jns.1999.91.5.0750.

[5] Oddo M, Carrera E, Claassen J, Mayer SA, Hirsch LJ. Continuous electroencephalography in the medical intensive care unit. Crit Care Med 2009;37(6):2051–6. https://doi.org/10.1097/CCM.0b013e3181a00604.

[6] Young GB, Jordan KG, Doig GS. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. Neurology 1996;47(1):83–9. https://doi.org/10.1212/WNL.47.1.83.

[7] Litt B, Wityk RJ, Hertz SH, Mullen PD, Weiss H, Ryan DD, et al. Nonconvulsive status epilepticus in the critically ill elderly. Epilepsia 1998;39(11):1194–202.

[8] Shneker BF, Fountain NB. Assessment of acute morbidity and mortality in non-convulsive status epilepticus. Neurology 2003;61(8):1066–73.

[9] Egawa S, Hifumi T, Kawakita K, Manabe A, Nakashima R, Matsumura H, et al. Clinical characteristics of non-convulsive status epilepticus diagnosed by simplified continuous electroencephalogram monitoring at an emergency intensive care unit. Acute Med Surg 2017;4:31–7. https://doi.org/10.1002/ams2.221.

[10] Schramm P, Luczak J, Engelhard K, El Shazly J, Juenemann M, Tschernatsch M. Continuous electroencephalography in a mixed non-neurological intensive care population, an observational study. J Crit Care 2017;39:62–5. https://doi.org/10.1016/j.jcrc.2017.01.015.

[11] Van de Vel A, Cuppens K, Bonroy B, Milosevic M, Jansen K, Van Huffel S, et al. Non-EEG seizure detection systems and potential SUDEP prevention: state of the art. Seizure 2016;41:141–53. https://doi.org/10.1016/j.seizure.2016.07.012.

[12] Trevathan E. Rapid EEG analysis for intensive care decisions in status epilepticus. Am J Electroneurodiagn Technol 2006;46:4–17.

[13] Stewart CP, Otsubo H, Ochi A, Sharma R, Hutchison JS, Hahn CD. Seizure identification in the ICU using quantitative EEG displays. Neurology 2010;75:1501–8. https://doi.org/10.1212/WNL.0b013e3181f9619e.

[14] Pensirikul AD, Beslow LA, Kessler SK, Sanchez SM, Topjian AA, Dlugos DJ, et al. Density spectral array for seizure identification in critically ill children. J Clin Neurophysiol 2013;30(4):371–5. https://doi.org/10.1097/WNP.0b013e31829de01c.

[15] Topjian AA, Fry M, Jawad AF, Herman ST, Nadkarni VM, Ichord R, et al. Detection of electrographic seizures by critical care providers using color density spectral array after cardiac arrest is feasible. Pediatr Crit Care Med 2015;16:461–7. https://doi.org/10.1097/PCC.0000000000000352.

[16] Dericioglu N, Yetim E, Bas DF, Bilgen N, Caglar G, Arsava EM, et al. Non-expert use of quantitative EEG displays for seizure identification in the adult neuro-intensive care unit. Epilepsy Res 2015;109(1):48–56. https://doi.org/10.1016/j.eplepsyres.2014.10.013.

[17] Yan P, Melman T, Yan S, Otgonsuren M, Grinspan Z. Evaluation of a novel median power spectrogram for seizure detection by non-neurophysiologists. Seizure 2017;50:109–17. https://doi.org/10.1016/j.seizure.2017.06.016.

[18] Britton JW, Frey LC, Hopp JL, Korb P, Koubeissi M, Lievens W, et al. Electroencephalography (EEG): an introductory text and atlas of normal and abnormal findings in adults, children, and infants [Internet]. Am Epilepsy Soc 2016. https://doi.org/10.5698/978-0-9979756-0-4.

[19] Haider HA, Esteller R, Hahn CD, Westover MB, Halford JJ, Lee JW, et al. Sensitivity of quantitative EEG for seizure identification in the intensive care unit. Neurology 2016;87:935–44. https://doi.org/10.1212/WNL.0000000000003034.

[20] Swisher CB, White CR, Mace BE, Dombrowski KE, Husain A, Kolls BJ, et al. Diagnostic accuracy of electrographic seizure detection by neurophysiologists and non-neurophysiologists in the adult ICU using a panel of quantitative EEG trends. J Clin Neurophysiol 2015;32:324–30. https://doi.org/10.1097/WNP.0000000000000144.

[21] Williamson CA, Wahlster S, Shafi MM, Westover MB. Sensitivity of compressed spectral arrays for detecting seizures in acutely ill adults. Neurocrit Care 2014;20(1):32–9. https://doi.org/10.1007/s12028-013-9912-4.

[22] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–324. https://doi.org/10.1109/5.726791.

[23] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402. https://doi.org/10.1001/jama.2016.17216.

[24] Hoseini F, Shahbahrami A, Bayat P. An efficient implementation of deep convolutional neural networks for MRI segmentation. J Digit Imaging 2018;31(5):738–47. https://doi.org/10.1007/s10278-018-0062-2.

[25] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000;101:E215–20. https://doi.org/10.1161/01.CIR.101.23.e215.

[26] Shoeb A. Application of machine learning to epileptic seizure onset detection and treatment. 2009.

[27] Claassen J, Hirsch LJ. Status epilepticus. In: Frontera J, editor. Decision making in neurocritical care. 1st ed.New York: Thieme; 2009. p. 63–75.

[28] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–44. https://doi.org/10.1038/nature14539.

[29] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Computer vision and pattern recognition. 2014.

[30] Aloysius N, Geetha M. A review on deep convolutional neural networks. 2017 international conference on communication and signal processing (ICCSP). IEEE 2017:0588–92. https://doi.org/10.1109/ICCSP.2017.8286426.

[31] Grill-Spector K, Weiner KS. The functional architecture of the ventral temporal cortex and its role in categorization. Nat Rev Neurosci 2014;15(8):536–48. https://doi.org/10.1038/nrn3747.

[32] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. Lect Notes Comput Sci 2013;11212 LNCS:3–20. https://doi.org/10.1007/978-3-030-01237-3_1.

[33] Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H. Understanding neural networks through deep visualization. 2015.

[34] Moura LMVR, Shafi MM, Ng M, Pati S, Cash SS, Cole AJ, et al. Spectrogram screening of adult EEGs is sensitive and efficient. Neurology 2014;83:56–64. https://doi.org/10.1212/WNL.0000000000000537.

[35] Gavvala J, Abend N, LaRoche S, Hahn C, Herman ST, Claassen J, et al. Continuous EEG monitoring: A survey of neurophysiologists and neurointensivists. Epilepsia 2014;55:1864–71. https://doi.org/10.1111/epi.12809.

[36] Jenssen S, Gracely EJ, Sperling MR. How long do most seizures last? A systematic comparison of seizures recorded in the epilepsy monitoring unit. Epilepsia 2006;47(9):1499–503. https://doi.org/10.1111/j.1528-1167.2006.00622.x.

[37] Naylor DE. Treating acute seizures with benzodiazepines: does seizure duration matter? Epileptic Disord 2014;16 Spec No(16):S69–83. https://doi.org/10.1684/epd.2014.0691.

[38] Stutz D. Understanding convolutional neural networks. Nips 2016;2014(3):1–23. https://doi.org/10.1016/j.jvcir.2016.11.003.