

# Epilepsy Seizure Prediction using Time Series Data

Fenil Suchak

Epilepsy afflicts nearly 1 percent of the world's population, and is characterized by the occurrence seizures. For many patients, medications can be given at sufficiently high doses to prevent seizures, but patients frequently suffer side effects. For 20-40 percent of patients with epilepsy, medications are not effective. Even after surgical removal of epilepsy, many patients continue to experience spontaneous seizures. Despite the fact that seizures occur infrequently, patients with epilepsy experience persistent anxiety due to the possibility of a seizure occurring. The aim of this research is to come up with Machine Learning based algorithms to warn patients of seizures some time  $t$  before the occurrence. Patients could avoid potentially dangerous activities like driving or swimming, and medications could be administered only when needed to prevent impending seizures, reducing overall side effects

## Literature Review

The ability to forecast seizures opens completely new possibilities to both study the seizure phenomenon and help patients from being harmed. Over the previous decade effort has been put towards identifying changes in EEG from minutes to hours before onset of seizure. (1) is based on pattern recognition method called N-gram technique which predicts the occurrence of a particular characteristic data sequence within data. The seizure horizon, ie.time before onset, was 20 min. The method achieved a sensitivity of 67 percent for temporal lobe cases and a FPR of 0.04 per hour. Frontal lobe cases showed lower specificity performance with 72 percent sensitivity and a false prediction rate of 0.61 per hour. By increasing the number of seizures in training, the method achieved 100 percent sensitivity and a false prediction rate of 0.06/h for the temporal cases. All results were statistically significant.

Two crowd-sourced prediction competition conducted by University of Melbourne(2), American epilepsy Society(3) hosted on Kaggle, had some really novel solutions and approach to solving epileptic seizure prediction data acquired from 3 subjects over a long period. Intra-Cranial EEG clips were taken and each segment was 10 minutes long, the data segment was annotated as either Preictal(Before Seizure) and Interictal(Normal). The sampling frequency was 400Hz. The challenge is to distinguish between ten minute long data clips covering an hour prior to a seizure, and ten minute iEEG clips of interictal activity. To avoid any potential contamination between interictal, preictal, and post-ictal EEG signals, interictal data segments were restricted to be at least four hours before or after any seizure. Interictal data segments were chosen at random within these restrictions. Average AUC across 3 patients was 0.8.

Some approaches of both competitions are listed below

1. From raw EEG time series data, features like Fourier Transform, Frequency Power in EEG bands, Summary features like mean, std, rms, kurtosis, skewness in the temporal domain, Correlation between channels in the frequency and temporal domains. Models used was an ensemble of Boosted Trees and Quadratic SVM classifier.
2. Features used Spectrum and Shannon's entropy at six frequency bands: delta (0.1-4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz), low-gamma (30-70Hz) and high gamma (70-180Hz), Spectral edge power of 50 percent power up to 40Hz, Shannon's entropy at dyadic frequency bands, Spectrum correlation across channels at dyadic frequency bands, Time-series correlation matrix and its eigenvalues, Fractal dimensions, Statistical moments: skewness and kurtosis.
3. Another novel method that was discovered was using Convolutional neural networks for automated feature generation. This requires feeding the raw time series data as a matrix to a CNN as used in (4). CNN acts as a feature extractor which is validated against a logloss and finally accuracy is measured.

Spike rate detection to differentiate between segments was also promising as done in (5). Using the open database, they divided the data for train and test. Morphological filter were used to detect spikes after removal of high frequency artifacts (using a 0.5–30 Hz band-pass filter). The spike rate was found to be significantly different for the interictal, preictal. The authors reported an average sensitivity of 75.8 percent and a false prediction rate of 0.09/h across 21 patients but they did not provide the seizure prediction horizon and intervention time values used to get these results. The mean prediction time was 49.7 min. The addition of spike rate feature to any of the existing state-of-the-art techniques can possibly lead to improvements and it will be explored in this project.

An algorithm based on univariate features and machine learning, named Advanced Seizure Prediction via Pre-Ictal Relabeling (ASPPR)(6), was trained and tested using the Freiburg database. The authors used 34 features. The authors introduced a new concept characterized by a "time-in-advance" parameter  $t$ . A relatively high specificity and accuracy (respectively, 99 percent and 97 percent) were achieved for all values of the intervention time equivalent parameter. The average sensitivity ranged between 89 percent and 93.5 percent.

# Bibliography

1. Amir Eftekhari, Walid Juffali, Jamil El-Imad, Timothy G. Constantinou, and Christofer Tournoux. Ngram-derived pattern recognition for the detection and prediction of epileptic seizures. *PLOS ONE*, 9(6):1–15, 06 2014. doi: 10.1371/journal.pone.0096235.
2. Dean R Freestone, Mark J Cook, David B Grayden, Philippa Karoly, Levin Kuhlmann, David T J Liley, Gareth Jones, Qingnan Tang, Irina Ivanenko, Oleg Panichev, Timothée Proix, Michal Náhlík, Daniel B Grunberg, Chip Reuben, Brian Litt, Benjamin H Brinkmann, Gregory Worrell, Andriy Temko, Alexandre Barachant, Feng Li, Jr. Titericz, Gilberto, Brian W Lang, Daniel Lavery, Derek Broadhead, Kelly Roman, and Scott Dobson. Epilepsyecosystem.org: crowdsourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain*, 141(9):2619–2630, 08 2018. ISSN 0006-8950. doi: 10.1093/brain/awy210.
3. "Crowdsourcing reproducible seizure forecasting in human B. H. Brinkmann, J. Wagenaar et al. and vol. 139 no. 6 pp. 1713–1722 2016. canine epilepsy, " *Brain*.
4. Nhan Duy Truong, Anh Duy Nguyen, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Samuel Ippolito, and Omid Kavehei. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 105:104 – 111, 2018. ISSN 0893-6080. doi: <https://doi.org/10.1016/j.neunet.2018.04.018>.
5. S. Li, W. Zhou, Q. Yuan, and Y. Liu. Seizure prediction using spike rate of intracranial eeg. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(6):880–886, Nov 2013. ISSN 1534-4320. doi: 10.1109/TNSRE.2013.2282153.
6. Negin Moghim and David W. Corne. Predicting epileptic seizures in advance. *PLOS ONE*, 9(6):1–17, 06 2014. doi: 10.1371/journal.pone.0099334.

## Supplementary Note 1: Drift correction

FluoCells slide #2 (Thermo Fisher Scientific) were imaged on a Ti2 microscope (Nikon) (Fig. ??) using a 100x TIRF objective (CFI Apochromat TIRF 100XC Oil, Nikon) onto a Prime 95B camera (Photometrics) in the GFP channel (tubulin), using an LED illumination source. 100 frames were acquired at 40 Hz and an artificially large drift was applied digitally.

## Supplementary Note 2: Channel alignment

Multicolor beads (TetraSpeck™ Microspheres, 0.1 µm, fluorescent in blue/green/orange/dark red, Thermo Fisher Scientific) were prepared on a #1.5 cover slip and embedded in water before being sealed with nail polish. The beads were then imaged on a Ti2 microscope (Nikon) (Fig. ??) using a 100x TIRF objective (CFI Apochromat TIRF 100XC Oil, Nikon) onto a Prime 95B camera (Photometrics), in both GFP and mCherry channels, using an LED illumination source.

## Supplementary Note 3: SRRF

**A. Cell lines.** Cos7 cells were cultured in phenol red-free Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific) supplemented with 10% (v/v) fetal bovine serum (FBS; Gibco), 1% (v/v) penicillin/streptomycin (Thermo Fisher Scientific) and 2 mM L-alanyl-L-glutamine (GlutaMAX™, Thermo Fisher Scientific) at 37 °C in a 5% CO<sub>2</sub> incubator.

**B. Sample preparation.** For live-cell imaging, Cos7 cells were seeded on 25mm-diameter #1.5 coverslips (Marienfeld) at a density of  $0.3 - 0.9 \times 10^5$  cells/cm<sup>2</sup>. One day after splitting, cells were transfected with a plasmid encoding the calponin homology domain of utrophin fused to GFP (GFP-UtrCH) (?) using Lipofectamin 2000 (Thermo Fisher Scientific) according to the manufacturer's recommendations. Cells were imaged 1–4 days post transfection in culture medium.

**C. Imaging.** LED-illumination widefield imaging of GFP-UtrCH in live Cos7 cells (Fig. ??) was performed in at 37 °C and 5% CO<sub>2</sub> on a N-STORM microscope (Nikon). A 100x TIRF objective (Apochromat 100x/1.49 Oil, Nikon) with additional 1.5x magnification was used to collect fluorescence onto an EMCCD camera (iXon Ultra 897, Andor), yielding a pixel size of 107 nm. Frames were acquired for 30 min with 30 ms exposure and 490 nm LED illumination at 5% of maximum output.

**D. SRRF reconstruction.** The dataset was reconstructed using NanoJ-SRRF with the following parameters: magnification: 5, radius: 0.5, number of axes: 6, number of frames per SRRF reconstruction: 100, temporal analysis: Temporal Radiality Average (TRA).

## Supplementary Note 4: SQUIRREL

**A. Cell lines.** Cos7 cells were cultured as described in Supplementary Note 3.

**B. Sample preparation.** Cos7 cells were fixed with glutaraldehyde and labelled with two monoclonal mouse anti-alpha tubulin antibodies (DM1A and B-5-1-2, both from Sigma) and a goat anti-mouse Alexa Fluor 647-conjugated secondary antibody (A21235, Thermo Fisher Scientific). Samples were mounted in Smart Buffer (Abbelight) for imaging.

**C. STORM acquisition and reconstruction.** STORM imaging of Cos7 cells (Fig. ??) was performed on an N-STORM (Nikon) microscope using a 100x TIRF objective (Apochromat 100x/1.49 Oil, Nikon). Prior to STORM imaging, a reference widefield image was acquired using low intensity 642 nm laser excitation. For subsequent STORM imaging, 60,000 frames were acquired with high intensity 642 nm laser excitation at 15 ms exposure with pixel size of 160 nm and 0.1248 photons per analog-to-digital unit. Localizations were detected using the N-STORM software (Nikon), and exported as a text file to be rendered using ThunderSTORM (? ).

**D. SQUIRREL analysis.** The initially acquired widefield image was used as the reference and the ThunderSTORM reconstruction as the super-resolution image. All parameters in the plugin were left at their default values. For local resolution measurements, two independent super-resolution images were created using ThunderSTORM; one comprising of the localisations from odd-numbered frames, and the other of even-numbered frames. These images were concatenated to form a two slice image stack and the 'Calculate FRC-map' method of NanoJ-SQUIRREL was run with 10 blocks per axis.

## Supplementary Note 5: VirusMapper

**A. Vaccinia virus sample preparation.** The L4-mCherry F17-EGFP virus was based on the WR strain and was described previously as WR EGFP-F17 VP8-mCherry (? ). A4-EGFP was described previously as A5-EGFP (? ). Virions were produced in BSC-40 cells, purified from cytoplasmic lysates and banded on a sucrose gradient. High performance coverslips (18x18

mm, #1.5H Zeiss) were washed with water, ethanol and acetone sequentially three times, then sonicated for 20 min in 1 M potassium hydroxide. The purified virus was diluted in 1  $\mu$ L of 1 mM Tris buffer (pH 9) and placed onto the coverslips. After 30 min, virus solution was removed and the virus was fixed with 4% formaldehyde in PBS for 20 min. Samples were washed 3 times with PBS and blocked for 30 min with 5% BSA in PBS. A17 was stained with rabbit polyclonal anti-A17 antibody, which was a kind gift from Jacomine Krijnse-Locker (Institut Pasteur, Paris, France) and Alexa Fluor 647-conjugated secondary antibody (A21245, Thermo Fisher Scientific). For STORM imaging, A4-EGFP virus was bound to coverslips as described above. Samples were permeabilised with 0.2% Triton-X 100 for 10 min, blocked 30 min with 5% BSA in PBS and labelled with anti-L1 antibody 7D11, which was purified from a mouse hybridoma cell line kindly provided by Bernard Moss (NIH, Bethesda, MD) with permission of Alan Schmaljohn (University of Maryland, Baltimore, MD), and Alexa Fluor 647-conjugated secondary antibody (A21235, Thermo Fisher Scientific).

**B. *Sulfolobus acidocaldarius* sample preparation.** *Sulfolobus acidocaldarius* strain DSM639 was grown in Brock's medium (pH 3.2), supplemented with 0.1% NZ-amine and 0.2% sucrose at 75 °C. Cells were fixed at OD600: 0.310 (i.e. within the limit of exponential growth) in 70% EtOH, rehydrated in PBST (0.2% v/v Tween) and stained overnight with polyclonal rabbit antibody for CdvB (?). Secondary staining was done with Alexa Fluor 647-conjugated Concanavalin A-antibody (C21421, Thermo Fischer Scientific) for labelling the S-layer and Alexa Fluor 546-conjugated secondary goat anti-rabbit antibody (A11035, Thermo Fischer Scientific) for visualising CdvB. The cells were spun down onto 2% polyethylenimine-coated coverslips before imaging.

**C. SIM imaging.** SIM imaging of vaccinia virus (Fig. ??a) was performed using a 63x objective (Plan-Apochromat 63x/1.4 oil DIC M27, Zeiss) and *Sulfolobus* imaging (Fig. ??b) was performed using a 100x TIRF objective (alpha Plan-Apochromat 100x/1.46 Oil DIC M27, Zeiss), both on an Elyra PS.1 microscope (Zeiss). Images were acquired using 5 phase shifts and 3 grid rotations, with the 647 nm (32  $\mu$ m grating period) and 561 nm (32  $\mu$ m grating period) and the 488 nm (32  $\mu$ m grating period) lasers, and filter set 3 (1850–553, Zeiss). 2D- (vaccinia) and 3D-images (*Sulfolobus*) were acquired on an sCMOS camera and processed using ZEN Black software (Zeiss). Channels were aligned with the NanoJ-Core Channel Alignment Tool.

**D. SIM/STORM imaging.** Correlative SIM/STORM imaging of vaccinia virus (Fig. S1) was performed on a Zeiss Elyra PS.1 inverted microscope with a 100x TIRF objective (alpha Plan-Apochromat 100x /1.46 NA oil DIC M27, Zeiss) objective with a 1.6x tube lens and an iXon 897 EMCCD camera (Andor). SIM imaging was performed as described above. STORM images were acquired at 25–30 ms exposure with 642 nm excitation at 100% laser power and a 655 nm LP filter. Activation was dynamically controlled with a 405 nm laser at 0–2% laser power. Images were processed using ThunderSTORM (?). Localisations were fitted with a maximum-likelihood estimator, lateral drift corrected by cross-correlation and localisations <20 nm apart within 2 frames merged.

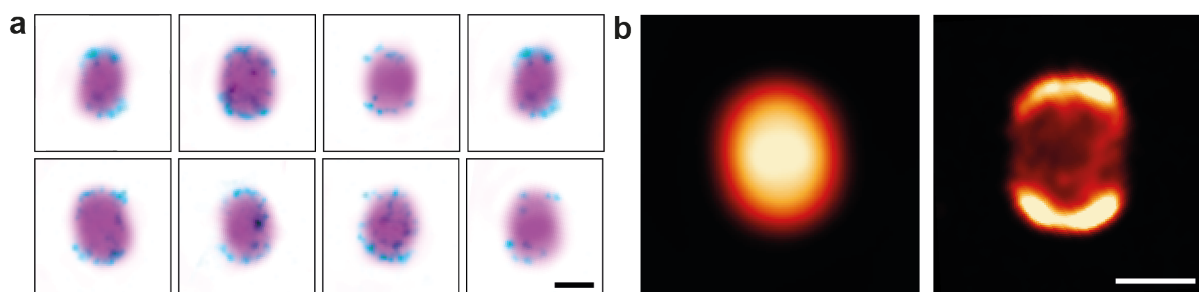
**E. VirusMapper processing.** Individual viral particles or archaea cells were extracted from the SIM or SIM/STORM images. Template images were generated with VirusMapper as described previously (?). 2- or 3-colour models of vaccinia virus proteins and 2-colour models of *Sulfolobus* S-layer and division ring in different orientations were then created by registration of the entire set of particles according to cross-correlation with the templates and calculation of a weighted average of a subset of particles.

## Supplementary Note 6: NanoJ-Fluidics

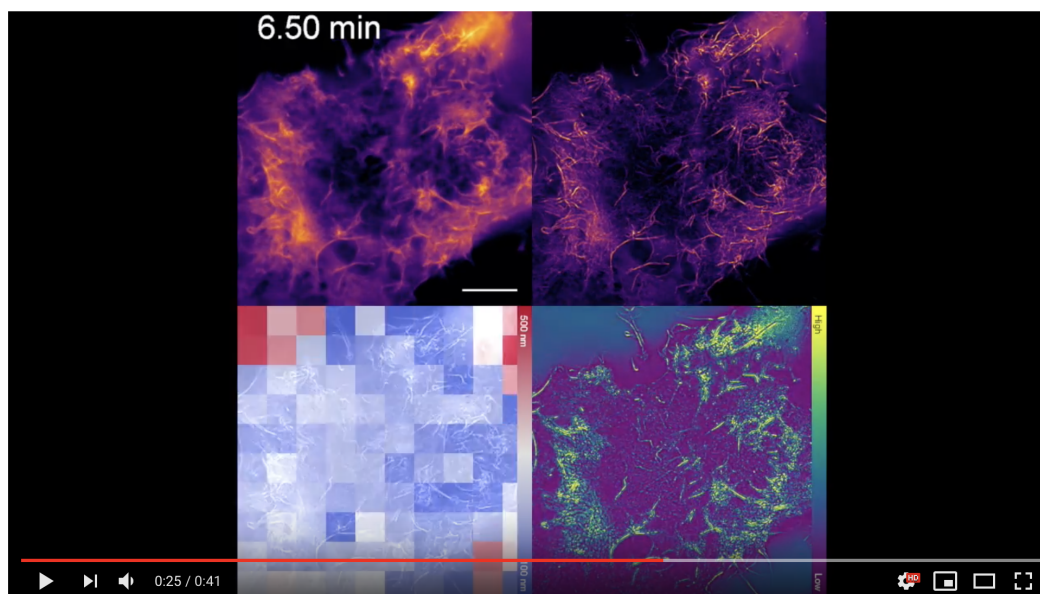
**A. Sample preparation.** Sample were prepared as described previously (?). Briefly, Cos7 cells were grown, seeded on glass coverslips and then fixed as described in Supplementary Note 4 and were labelled with primary antibodies overnight: mouse monoclonal anti-TOM20 (612278, BD Biosciences), rabbit polyclonal anti-clathrin heavy chain (ab21679, Abcam) and chicken poly-clonal anti-vimentin (919101, BioLegend). Cells were then incubated with Exchange-PAINT secondary antibodies coupled to DNA sequences (?): goat anti-mouse I1, goat anti-chicken I2 and goat anti-rabbit I3. After rinses, the cells were incubated with 12.5  $\mu$ M Atto488-conjugated phalloidin (Merck) for 90 min at room temperature and imaged within 3–4 days.

**B. NanoJ-Fluidics workflow and SMLM acquisition.** Imaging of Cos7 cells (Fig. ??) was performed as described in (?). The NanoJ-Fluidics pump array was installed on an N-STORM microscope (Nikon) using a 100x TIRF objective (CFI Apochromat 100x/1.49 Oil, Nikon). First, a STORM imaging of phalloidin-Atto488 was performed using 30,000 frames at 30 ms exposure. After injection of the I1-Atto655 (0.25 nM) and I2-CY3B (2 nM) imager strands, 60,000 frames were acquired in an alternating way (30,000 frames of each channel) to image TOM20 and vimentin, respectively. After rinses, I3-Cy3B (1 nM) was injected, and 30,000 frames were acquired to image clathrin.

**C. SMLM reconstruction.** Localisations were detected using the N-STORM software (Nikon), and exported as a text file before being filtering (number of photons between 700 and 50,000; number of detections (after linking across frames) < 50 frames) and rendered using ThunderSTORM.



**Fig. S1. NanoJ-VirusMapper analysis of SIM/SMLM data.** **a)** Example SIM/STORM images of A4-EGFP vaccinia virions labelled for the entry protein L1. SIM images were taken of the A4-EGFP core (magenta), followed by STORM images of the L1 membrane protein (cyan). The images were aligned and overlaid. **b)** VirusMapper SPA models of A4 and L1 using the SIM/STORM data. Particles were aligned in parallel and averaged to produce the model. Scale bars 200 nm.



NanoJ Sup. Movie 1 - Demonstration of SRRF and SQUIRREL analysis

**Fig. S2. Actin dynamics visualized at super-resolution by NanoJ-SRRF.** Cos7 cells expressing GFP-UtrCH imaged at 33.3 Hz, and reconstructed at 0.3Hz (100 frames) for 30 min. The image quality and resolution were assessed by NanoJ-SQUIRREL. Scale bars 10  $\mu$ m.

## Bibliography