

Research Articles | Neurobiology of Disease

Machine Learning Framework for Integration of Video and EEG Reveals Distinct Seizure Phenotypes in Mouse Models of Epilepsy

<https://doi.org/10.1523/JNEUROSCI.0280-26.2026>

Received: 12 February 2026

Revised: 27 August 2026

Accepted: 23 September 2026

Copyright © 2026 Liu et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

This Early Release article has been peer reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.jneurosci.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

Machine Learning Framework for Integration of Video and EEG Reveals Distinct Seizure Phenotypes in Mouse Models of Epilepsy

Abbreviated title: Machine Learning analysis of rodent video-EEG data

Author names and affiliations: Christine J. Liu,^{1,2,3} Caryn Martin¹, Allison C. Lee¹, Tristan T. Sands¹, Wayne N. Frankel⁴, and Christopher D. Makinson^{1,3*}

¹Department of Neurology, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, 10032, USA

²Kavli Center for Theoretical Neuroscience, Mortimer B. Zuckerman Mind Brain Behavior Institute, New York, NY, 10027, USA

³Department of Neuroscience, Mortimer B. Zuckerman Mind Brain Behavior Institute, New York, NY, 10027, USA

⁴Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY 10032, USA

*Correspondence: cm3966@cumc.columbia.edu

Pages: 15

5 Figures, 5 Supplemental Figures, 2 Movies

Abstract Word Count: 217

Introduction Word Count: 560

Discussion Word Count: 1457

Conflict of Interest: The authors declare no conflict of interests.

Acknowledgements: We thank Drs Larry Abbott and Mu Yang for their support and insightful feedback on the manuscript.

Machine Learning Framework for Integration of Video and EEG Reveals Distinct Seizure Phenotypes in Mouse Models of Epilepsy

Christine J. Liu,^{1,2} Caryn Martin¹, Allison C. Lee¹, Tristan T. Sands¹, Wayne N. Frankel⁴, and Christopher D. Makinson^{1,3*}

¹Department of Neurology, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, 10032, USA

²Kavli Center for Theoretical Neuroscience, Mortimer B. Zuckerman Mind Brain Behavior Institute, New York, NY, 10027, USA

³Department of Neuroscience, Mortimer B. Zuckerman Mind Brain Behavior Institute, New York, NY, 10027, USA

⁴Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY 10032, USA

*Correspondence: cm3966@cumc.columbia.edu

Abstract

Video-electrocorticogram (VEEG) recordings are the gold-standard diagnostic tool for detecting and characterizing seizures in rodent models of seizure and epilepsy. Manual/visual scoring remains the prevailing strategy despite significant limitations. Alternatively, automated and semi-automated analyses have been applied to either EEG or video features alone but integrating these disparate data forms presents a challenge that has not been sufficiently addressed. Here, we introduce an integrated machine learning pipeline that jointly leverages EEG and behavioral video data to enable scalable and objective event classification. Integrated video-EEG analysis was initially validated using a well-established pharmacologically induced seizure paradigm in male mice, where overt generalized motor seizures provide a clear benchmark. We then tested the pipeline on a genetic mouse model of developmental and epileptic encephalopathy (DEE) that displays a complex spike-wave discharge (SWD) seizure phenotype that poses significant challenges for traditional scoring methods. Machine learning-powered integrated analysis of video-EEG recordings outperformed unimodal EEG- or video-based approaches and performed favorably to experienced rodent EEG electrophysiologists while also uncovering distinct phenotypic relationships, including classes of atypical behavioral and typical non-behavioral SWD events. This framework provides a flexible tool for high-throughput, precise analysis of VEEG data, with potential to accelerate discovery in epilepsy and seizure research. Furthermore, this approach represents a generalizable strategy for the integration of diverse neurobehavioral time series data types.

Keywords: Epilepsy, Seizure detection, Machine learning, Multimodal analysis, Video-EEG, Kainic acid, Spike-wave discharge

Significance

Seizure detection in animal models of epilepsy has relied heavily on manual analysis of video-EEG recordings, a process that is slow and subjective. We present a machine learning pipeline that integrates EEG and video data from mice to classify seizures with high accuracy. This work was validated across two models, first a pharmacologically induced generalized seizure model, and second, a spontaneous spike-wave discharge (SWD) model in genetic epilepsy. Our approach detected events missed by human reviewers and revealed previously uncharacterized SWD subtypes distinguished by their behavior. By combining open-source tools in a flexible framework, this work lowers the barrier to high-throughput, unbiased seizure phenotyping, and also offers a broadly applicable strategy for linking neural activity to behavior in neuroscience research.

Introduction

Video-electrocorticogram (VEEG) recordings provide a powerful means for linking neural activity to behavior in animal models of seizure and epilepsy, but their analysis remains a persistent challenge (Svantesson, 2024). The detection and classification of seizures has been inextricably linked to the behavioral manifestations of these events since the earliest known descriptions of seizures in antiquity (Wilson JV, 1990) (Kaculini CM, 2021), extending to early direct measurements of synchronous brain activity in humans and the establishment of EEG signatures as a core diagnostic criteria for seizure in the early 20th century (Berger, 1929). The integration of EEG and behavioral features is the basis of the detailed and evolving system of seizure classification used in animals (Wang Y, 2022) and humans (Wirrell EC, 2022) to the present day.

Seizures result from transient, hypersynchronous neuronal activity within brain networks, whereas epilepsy is defined as a chronic neurological disorder characterized by a persistent predisposition to generate recurrent unprovoked seizures (Fisher RS, 2014). These events can emerge from diverse neurophysiological mechanisms, including imbalance in excitatory and inhibitory signaling, inappropriate network connectivity, and alterations in synaptic plasticity (McCormick DA, 2001). This distinction underscores the importance of accurately detecting and classifying both ictal and interictal events, as well as their associated behavioral manifestations.

Recently, automated methods have emerged for the detection of seizures, but to-date these approaches rely on the separate analysis of either EEG or video features alone (EEG alone: (Jing J, 2020) (Tveit J, 2023) (Yan, 2025) (Kotloski, 2023) (Manzouri F, 2018) (Lucas, 2024) Video only: (Garção VM, 2023) (Li, 2014) (Y. Yang, 2021) (Liu C, 2020)). EEG-based approaches are prone to misclassification of artifacts, such as those associated with movement, while video-based analyses are limited in temporal resolution and are not well suited to the detection of non-motor seizures, such as some classes of partial or spike-and-wave (SWD) types, and interictal events. As a result, automated seizure detection is mostly used as an aid to visual inspection, intended to increase the speed and accuracy of seizure detection. Because EEG and behavior provide complementary information that is relevant to the detection and classification of seizures, an ideal analysis structure would integrate both modalities to improve accuracy, generalizability, and scalability.

Here, we present an integrated machine learning pipeline that combines EEG and behavioral data for automated event detection in rodent VEEG. Our framework synchronizes spike detection in EEG with behavior classification using DeepLabCut and SimBA, enabling joint analysis across modalities (Mathis, 2018) (Goodwin, 2024). We first validated the method using an established kainic acid-induced seizure paradigm that produces progressive large amplitude high frequency seizure events in the EEG and stereotyped seizure-related behaviors (Sabilallah M, 2016). We then applied the approach to detect spike-and-wave discharges (SWDs), hallmarks of absence epilepsy that are often difficult to distinguish from background or non-seizure events. In this setting, multimodal analysis not only improved detection accuracy compared to unimodal methods but also uncovered additional phenotypic relationships, including a distinction between SWD types characterized by typical spindle-form non-behavioral SWDs and atypical behavioral SWDs.

Together, these findings demonstrate that multimodal video-EEG analysis can enhance precision and reliability in epilepsy research. While we focused here on kainic acid-induced seizures and spontaneous SWDs in a genetic model of epilepsy, this pipeline provides a generalizable framework that could be extended to other seizure types or behavioral-neural coupling problems, offering a scalable alternative to manual annotation.

Materials and Methods

EEG Surgery

Male mice aged 3–5 weeks were prepared for electrode implantation under aseptic conditions as previously described. (Wang W, 2024) (Purcell RH, 2013) Animals received pre-operative analgesia, and their heads were shaved prior to surgery. Anesthesia was induced with 4% isoflurane, after which mice

were secured in a stereotaxic frame. The scalp was sterilized, and a midline incision was made to expose the skull. Four burr holes were drilled relative to bregma (AP -0.5 mm, ML +1.0 mm; AP -1.0 mm, ML -1.7 mm). Additional burr holes were made over the olfactory bulb and cerebellum to serve as ground and reference, respectively. PFA-coated silver wires (AM-Systems, 786000) attached to a Mill-Max headmount (Digikey, 853-93-100-10-001000) were inserted under the skull and secured with stainless steel screws (McMaster-Carr, 91794A658). To record electromyogram (EMG) activity, the incision was extended to expose dorsal neck muscles, and the remaining electrode wires were directly inserted into the musculature. Dental cement (Parkell, S380) was applied to anchor electrodes and form a stable headcap. Postoperative analgesia was administered 24 hours later, and animals were monitored for three days.

EEG Data Collection and Processing

A telemetric transmitter (Pinnacle Technology, 8274-SE3), powered by a zinc air battery (Pinnacle Technology, 9033-AZ-PR48; discontinued and replaced by 9033-HCAZ-PR48), was used to capture EEG and EMG activity. The device was first wirelessly paired to a USB dongle (Pinnacle Technology, 8274-D) to activate data streaming. The transmitter was then securely attached to the Mill-Max headmount, establishing electrical contact with the implanted EEG and EMG leads. Each recording was made using a Pinnacle Technology Sirenia Acquisition Software system at 1028 Hz and consisted of three channels of continuous data.

EEG signals were acquired from five heterozygous *Grin2d*-V664I mice, three wild-type littermates, and nine C57BL/6J wild-type mice that received kainic acid injections, yielding a total of 17 animals.

EEG signals were band-pass filtered between 1–100 Hz to remove slow drifts and high-frequency noise, and notch-filtered at 60 Hz to eliminate mains interference. Artifacts exceeding 5 standard deviations from the mean amplitude were automatically rejected. Spectral features were calculated using a short-time Fourier transform (STFT) with 1-second Hanning windows, extracting power in frequency bands relevant to rodent SWDs (7–9 Hz).

Spike detection employed a custom algorithm built on SciPy's `find_peaks` function. Candidate spikes were identified as local maxima exceeding an amplitude threshold. To account for variability across animals and sessions, an adaptive thresholding procedure was implemented: the EEG signal was segmented into overlapping 10,000-sample windows, and spikes were retained if their amplitudes exceeded the local mean by a multiple of the standard deviation. Non-seizure epochs were defined as periods prior to the first electrographic event and interictal windows between discrete seizures or SWDs for both the kainic acid model and the *Grin2d* model.

Application to kainic acid seizures.

Kainic acid-induced seizures were characterized by large-amplitude, rhythmic bursts. These overt electrographic events served as a benchmark for validating the classification pipeline. Performance was assessed on a validation set of five entirely held-out animals.

Application to spike-and-wave discharges.

For absence seizures, spike-and-wave discharges (SWDs) were detected using a sliding-window frequency analysis. A putative SWD was defined as a period in which at least six spikes occurred within a 6-second window; the window was extended until spike frequency dropped below this threshold. Events were further classified as short (<5 minutes) or long (≥5 minutes). This threshold was optimized through an unsupervised learning approach. The threshold of six spikes per six seconds was determined using a Gaussian mixture model to cluster spike frequencies, selecting the boundary between high-frequency (SWD) and low-frequency (background) clusters. This data-driven approach ensured robust, unbiased separation of SWDs from background activity.

To further distinguish seizure-related versus non-seizure EEG events, a support vector machine (SVM) classifier was trained on manually annotated SWDs. The SVM optimization problem is defined as:

Minimize

175 $\frac{1}{2} ||\mathbf{w}'||^2 + C \sum_{i=1}^n \xi_i$

176 subject to:

177 $y_i (\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall i = 1, \dots, n$

178 where $\mathbf{x}_i$ are EEG-derived feature vectors, $y_i \in \{-1, +1\}$ are class labels (seizure vs. non-seizure), $\mathbf{w}$ is the
179 weight vector, b the bias, and ξ_i slack variables allowing soft-margin classification with penalty parameter
180 C . The classifier's decision function is given by:

181 $f(\mathbf{x}) = \text{sign}(\mathbf{w} \cdot \mathbf{x} + b)$.

182 For both EEG- and behavior-based classifiers, input features were z-scored to normalize across animals.
183 An RBF kernel was employed, and hyperparameters (C and γ) were optimized using 5-fold cross-
184 validation on a training subset. Classifier performance was evaluated on a held-out test set of 1 animal's
185 entire data plus additional randomized held-out samples from the remaining animals to comprise an 80-
186 20 train/test split, and sensitivity, specificity, and Cohen's κ were computed per recording session to
187 account for inter-animal variability.

188 Features such as spike count, inter-spike interval, and spectral power were used as inputs to the
189 classifier. The trained SVM enabled automated discrimination between true SWDs and false positives
190 arising from background noise or artifacts.

191 Metrics including spike counts, SWD counts and durations, and percent recording time in SWDs were
192 extracted and averaged across channels. Group-level comparisons were performed using the Mann-
193 Whitney test. To facilitate validation, detected spikes and SWDs were visualized on raw traces, with
194 events color-coded by duration.

195 **Behavioral Data Collection and Processing**

196 Behavioral recordings were acquired concurrently with EEG using two Microsoft Azure high-speed
197 cameras positioned for both top-down and lateral views at 30 FPS, enabling comprehensive monitoring of
198 motor activity. Animal movement was tracked with DeepLabCut, a deep learning-based pose estimation
199 framework, using a custom-trained convolutional neural network to detect multiple landmarks (head,
200 snout, body, tail) and generate fine-grained kinematic trajectories.

201 A total of 975 frames were extracted from 12 videos spanning more than 144 hours of continuous
202 recording. Frames were sampled across animals and behavioral states to capture variability in posture,
203 movement, and lighting conditions. Training performance plateaued at approximately 200,000 iterations,
204 as indicated by convergence of the training loss. Based on empirical evaluation across multiple training
205 conditions, a model trained with 200,000 iterations, shuffle = 1, and batch size = 8 was selected for final
206 analysis. This model achieved a training error of 3.46 pixels and a test error of 6.29 pixels after applying a
207 likelihood threshold (p-cutoff) of 0.6 to exclude low-confidence detections.

208 Raw tracking outputs were processed in SimBA to extract quantitative behavioral metrics, including
209 velocity, angular displacement, and positional stability. A training set of manually annotated recordings
210 was used to label seizure-related behaviors such as behavioral arrest, rearing, darting, and other subtle
211 movements temporally associated with electrographic events. Feature selection was applied to reduce
212 redundancy and enhance classifier interpretability.

213 To validate behavioral classification, a subset of frames ($n = 25$ per behavior) was manually inspected
214 and annotated by an experienced experimenter. SimBA predictions were compared to manual labels to
215 assess agreement. Across behaviors, SimBA demonstrated strong agreement with manual annotations,
216 with performance varying slightly by class: arrest (accuracy = 0.80, F1 = 0.83), head-movement (accuracy
217 = 0.78, F1 = 0.74), locomotion (accuracy = 0.81, F1 = 0.80), and rearing (accuracy = 0.76, F1 = 0.75),
218 indicating generally robust classification performance with some class-specific differences in precision
219 and recall.

Behavioral metrics were normalized per animal to account for differences in overall activity. Feature selection was performed using recursive feature elimination (RFE) with the classification model to iteratively rank features according to their contribution to model performance and remove the least informative features. This process was repeated until the optimal reduced feature set was obtained, minimizing feature dimensionality while retaining the features that contributed most strongly to seizure classification. Feature selection was performed using the training data only, and the selected features were subsequently applied unchanged to the held-out test animals. RFE also reduced redundancy among correlated behavioral and kinematic measures, ensuring that classification was driven primarily by features with the greatest predictive value for seizure-associated behavior.

Seizure-related versus non-seizure behaviors were then classified using a support vector machine (SVM), implemented with the same mathematical framework described above for EEG-based SWD detection. In this case, the feature vectors x_i consisted of behavioral metrics, and class labels $y_i \in \{-1, +1\}$ indicated the presence or absence of seizure-associated behavior. Hyperparameters, including the penalty parameter C and kernel parameters, were optimized via grid search to maximize classification accuracy.

Behavioral Classification and Feature Definitions

Raw tracking outputs from DeepLabCut were processed in SimBA (Simple Behavioral Analysis) to automatically extract quantitative behavioral metrics and classify seizure-related behaviors. SimBA's machine learning algorithms identified the following behavioral states based on tracked body keypoint positions and velocities:

Behavioral arrest was automatically detected as periods of complete motor immobility, characterized by minimal displacement of tracked body regions across consecutive frames (velocity threshold <2 cm/s for ≥ 0.5 seconds).

Head-related movements were classified into three categories based on head region kinematics: (1) head nodding: rhythmic vertical oscillations of the head detected through periodic velocity peaks in the dorsal-ventral axis; (2) head rearing: sustained elevation of the head/snout region relative to the body centroid; and (3) head movement: non-rhythmic directional changes in head position. SimBA grouped these as a composite feature due to their frequent co-occurrence during seizures.

Locomotor activity included (1) running: rapid, erratic whole-body movement detected through high velocity (>20 cm/s) combined with frequent directional changes throughout the entire cage; and (2) darting: rapid short-distance movements with sudden trajectory reversals.

Rearing was detected as sustained elevation of the front limbs with the body in a vertical posture, automatically distinguished from head rearing by the spatial position of tracked forelimb keypoints relative to the ground plane.

SimBA extracted these behaviors as probabilistic outputs for each frame, which were converted to binary state variables (behavior present/absent) based on a confidence threshold of >0.5 and used as input features for the seizure detection classifier.

Multimodal Integration

EEG-based and behavior-based classifications were temporally aligned to produce synchronized multimodal seizure detection. EEG and behavioral data were temporally synchronized using shared TDT system timestamps, correcting for the known 33 ms delay of the video acquisition system. Events were aligned to the nearest frame to ensure precise correspondence between electrographic and behavioral seizure markers. Performance was benchmarked against unimodal approaches using Cohen's κ , sensitivity, and specificity. The combined method consistently outperformed EEG-only and behavior-only analyses, particularly in detecting subtle events. Notably, multimodal analysis uncovered distinct phenotypic subtypes, including behavioral versus non-behavioral SWDs, that were not reliably captured by unimodal methods.

Regarding temporal locking: EEG and behavioral features were temporally synchronized using shared timestamps, correcting for a 33 ms delay of the Azure camera acquisition system. Features were aligned to the nearest frame. Regarding temporal resolution: each classification instance was defined as a 1 to 6-second window depending on continuity of the EEG signal centered on the detected spike event within which all EEG spectral and spike features were computed. Sequence or transitions between behavioral states were not explicitly modeled as features in the current implementation- behavioral states were treated as instantaneous binary variables within the classification window.

Initial Classifier Training on Kainic Acid Data

The multimodal seizure detection pipeline was initially developed and validated using kainic acid-induced seizures as a benchmark. Support vector machine (SVM) classifiers were trained on manually annotated seizure events from four 2-month-old wild-type mice on the C57BL/6J background (Jackson Laboratory) that received intraperitoneal kainic acid injections (20 mg/kg). Five separate models were trained: EEG only, behavior only, top camera only, side camera only, and a combined multimodal model integrating all features.

For both EEG- and behavior-based classifiers, input features were z-scored to normalize across animals. All models were trained using an 80/20 initial train/test split. An RBF kernel was employed, and hyperparameters (C and γ) were optimized using 5-fold cross-validation on a training subset. Classifier performance was evaluated on a held-out test set using ROC curves, precision-recall curves, and metrics including accuracy, area under the curve (AUC), F1 score, precision, and recall. This initial training phase established baseline classifier performance on overt seizure phenotypes (Fig. 2).

The **behavior classifier** used SimBA-derived behavioral state features (arrest, rearing, head movement, etc.) extracted from the combined top and side camera views as inputs to the SVM. The **top-mounted camera classifier** and **side-mounted classifier** used movement features extracted by DeepLabCut from the top-mounted and side-mounted cameras respectively, processed through SimBA independently for each view. These are therefore camera-perspective-specific behavioral feature sets, not raw frames. The **combined multimodal classifier** concatenated EEG features, SimBA behavioral state features, and movement features from both camera views into a single feature matrix ($X = [X_Behavior \mid X_EEG]$).

Application to Spontaneous Seizures in *Grin2d*-V664I Animals

The trained classifiers were then applied to spontaneous seizures in *Grin2d*-V664I mice to test generalizability to more subtle seizure phenotypes. Three male heterozygous *Grin2d*-V664I mice and three male wild-type littermates were monitored using the same telemetric EEG and dual-camera video setup. The kainic acid-trained classifiers were applied without retraining to detect spontaneous spike-wave discharges (SWDs) and assess performance on this distinct seizure type.

Retraining on Combined Dataset for Genotype Classification

To enable classification of genotype-specific seizure patterns, classifiers were retrained on a combined dataset incorporating both kainic acid-induced seizures and spontaneous SWDs from *Grin2d* mutant and wild-type mice. This expanded training set included data from all twelve animals (four kainic acid-treated C57BL/6J mice, five *Grin2d*-V664I mice, and three wild-type littermates). The retrained multimodal classifier was then evaluated on its ability to distinguish between four classes: wild-type SWDs, wild-type non-seizure activity, *Grin2d* mutant SWDs, and *Grin2d* mutant non-seizure activity. Multi-class classification performance was assessed using confusion matrices and density distributions of principal component scores (Fig. 5F-G).

This two-stage training strategy allowed us to: (1) validate the pipeline on overt seizures with clear ground truth, (2) test generalizability to spontaneous seizures without retraining, and (3) demonstrate that retraining on diverse seizure phenotypes enables genotype-specific classification.

Statistical Analysis

All analyses were conducted in Python. Distribution normality was tested with the Shapiro–Wilk test. Nonparametric comparisons between groups were performed using the Mann–Whitney test, and repeated-measures comparisons were conducted with the Friedman test followed by Bonferroni-corrected

post hoc analyses. Statistical significance was set at $p < 0.05$. Statistical comparisons of classifier performance metrics (sensitivity, specificity, κ) across methods were corrected for multiple comparisons using the Holm–Bonferroni procedure. Group-level analyses accounted for repeated measures by averaging metrics across recordings per animal.

This integrated EEG–behavior classification pipeline represents, to our knowledge, the first multimodal approach validated across both overt (kainic acid–induced GTCSs) and subtle (SWDs) seizure phenotypes. By leveraging machine learning–based behavioral tracking alongside EEG classification, the method provides an efficient and scalable framework for automated seizure detection, reducing dependence on manual annotation and enabling discovery of additional phenotypic relationships.

Manual Annotation

To validate automated detection, EEG recordings were also reviewed by multiple human experts with a combined >30 years of experience scoring rodent EEGs, as well as one epileptologist (specializes in clinical epilepsy) who validates the scores independently. The reviewers performed an initial scan of each recording based on the predefined SWD parameters, then adjusted detection thresholds such that false negatives did not exceed approximately 5% of events using the software Assyst. (Casillas-Espinosa, 2019) This approach provided semi-statistical normalization across animals.

However, we note that manual scoring inherently involves an element of subjectivity. Detection rates can vary depending on the reviewer’s chosen balance between sensitivity and effort, as thresholds may be set more or less stringently depending on the desired level of exhaustiveness. Consequently, while the comparison between manual and automated detections is informative, it should be interpreted with awareness of this flexibility in traditional human scoring approaches. While no ground truth in neurophysiology is absolute, this approach reflects the current gold standard in the field.

For generalized tonic–clonic seizures, the same reviewers manually identified events based on electrographic signatures and corresponding video confirmation. These annotations served as an independent reference standard for evaluating automated classifications.

Results

Introduction to Multimodal Pipeline

To develop an automated approach for seizure detection, we constructed an integrated pipeline combining telemetric EEG recording with dual-perspective video analysis. DeepLabCut was used to extract kinematic features from video data, while SimBA classified behavioral states. (Mathis, 2018) (Goodwin, 2024) These behavioral features were then combined with EEG-derived metrics and used to train support vector machine (SVM) classifiers for seizure detection (Fig. 1).

Validation in Kainic Acid Model

Wild-type C57BL/6J mice ($n = 9$) were injected intraperitoneally with kainic acid (20 mg/kg) as previously described and monitored using telemetric EEG (Pinnacle Technology) and dual video recordings from top- and side-mounted Microsoft Azure cameras (Fig. 2A) (Makinson, Tanaka, Lamar, Goldin, & Escayg, 2014). After data collection, seizure state classifications were generated based on EEG traces, behavioral observations, and combined EEG/behavioral annotations. Five separate classification models were then trained: EEG only, behavior only, top-camera only, side-camera only, and a combined multimodal model integrating all modalities (Fig. 2B).

The kainic acid seizure model exhibits generalized seizures which provide a clear benchmark for classification. To ensure generalizability, classifiers were initially trained on events from four mice and tested on five held-out mice. Even under this strict evaluation, the combined EEG–video model outperformed unimodal approaches, achieving the highest AUC and F1 scores. Notably, the performance advantage of multimodal integration was consistent for both within-animal and across-animal validation, indicating that the model captures shared seizure features while avoiding overfitting to individual animals.

Across models, seizure detection performance varied substantially. Receiver operating characteristic (ROC) and precision–recall (PR) curves indicated that the combined multimodal classifier consistently

outperformed unimodal approaches, achieving the highest area under the curve (AUC) values and precision–recall balance (Fig. 2C). Classifier confidence distributions further highlighted that the combined model more effectively separated seizure from non-seizure states, with reduced overlap compared to EEG- or video-only models. The near-ceiling performance observed in the KA dataset likely reflects the high separability of overt seizure events from baseline activity rather than overfitting.

To further assess the robustness of this performance to the composition of the training cohort, we performed a second validation analysis in which the classifier was retrained using data from seven mice and evaluated using an 80/20 train–test split, with two additional mice held out as an independent validation set. Thus, the animals used for independent validation were not included in either the training or test sets. Training and validation results are provided in Supplemental Figure 1. Performance following retraining was comparable to that observed in the original four-animal training paradigm, demonstrating that classifier performance was robust to changes in the training cohort and generalized across independently recorded animals (Supp. Fig. 1). Given the overt electrographic and behavioral manifestations of kainic-acid-induced seizures, these results further support the ability of the classifier to identify consistent features that distinguish seizure from non-seizure states.

Feature importance analyses revealed that theta-band EEG power (5.5–7.5 Hz) and spike rate were among the most predictive electrophysiological features, while behavioral arrest, head nodding, and darting or running were the strongest behavioral indicators (Fig. 2E). Video-derived movement features from both top and side cameras contributed to classification but carried less weight relative to canonical seizure behaviors.

Quantitative performance metrics confirmed the superiority of multimodal integration. The combined model achieved the highest accuracy, F1 score, precision, and recall compared to individual modality classifiers (Fig. 2D). Among unimodal classifiers, EEG-based detection generally outperformed single-camera models, with behavior-based classification showing intermediate performance. Together, these results demonstrate that integrating electrophysiological and behavioral data yields more reliable and accurate seizure detection following kainic acid administration.

Classifier Captures Seizure Progression and Behavioral Complexity

To examine whether the classifier could distinguish seizure severity and behavioral features, we analyzed the temporal evolution of kainic acid-induced seizures in the original cohort of four mice used for the initial characterization of the KA model. Representative EEG traces demonstrated the characteristic evolution that generally follows from baseline activity through epileptiform discharges to lower frequency and amplitude “moderate seizures” that could include tonic-clonic behavior without a corresponding high amplitude electrographic seizure pattern, and ultimately to severe generalized tonic-clonic seizures (GTCS), characterized by both tonic-clonic behavior of all four limbs, loss of posture, and a prominent high amplitude electrographic discharge (Fig. 3A).

Analysis of seizure progression across individual mice revealed the expected stereotyped temporal sequence, with latency to different seizure behaviors increasing with behavioral complexity (Fig. 3B). Classifier confidence also showed a monotonic increase with number of behavior events (Supp. Fig. 2). Simple behaviors such as head nodding and rearing emerged earliest (10–30 min post-injection), followed by tonic-clonic movements and terminal GTCS at later timepoints (40–60 min). This progression was accompanied by increasing behavioral complexity, quantified as the number of distinct behaviors expressed during each seizure event (Fig. 3C).

To assess the reliability of automated detection, we directly compared machine learning classifications with human expert scoring as previously described (Wong JC, 2018) (Makinson, Tanaka, Lamar, Goldin, & Escayg, 2014). Of the total events analyzed, 29 were identified by both human and machine, demonstrating strong concordance (Fig. 3D). The machine learning approach detected 18 additional events (~37.5%) not initially scored by human observers, while a singular additional tonic-clonic seizure (~2.08%) was identified exclusively by a human expert (Fig. 3D). Visual inspection of representative traces confirmed that events detected by both methods displayed unambiguous seizure characteristics, while discordant events often represented brief or ambiguous discharges at the threshold of manual detection (Fig. 3E). These results validate that the automated pipeline reliably captures the severity spectrum and temporal dynamics of kainic acid-induced seizures.

Application to Spontaneous Seizures in a *Grin2d* –DEE Mouse Model

Having validated the multimodal pipeline on overt kainic acid-induced seizures, we next tested its ability to detect spontaneous seizures in a *Grin2d* model of epilepsy- a more challenging problem given the subtle and variable nature of seizure events in this model. We used a conditional knock-in mouse model carrying the V664I missense mutation in the *Grin2d* gene, which encodes the GluN2D NMDA receptor subunit (Fig. 4A,B) (Teoh J, 2025). This mutation, located within the agonist-binding domain of the GluN2D subunit, has been identified in patients with developmental and epileptic encephalopathy and is associated with gain-of-function effects on NMDA receptor activity (Teoh J, 2025) (Camp CR, 2020). To introduce the V664I mutation broadly throughout the animal we crossed Sox2-Cre animals (The Jackson Laboratory) to heterozygous floxed *Grin2d* animals to generate Sox2-Cre/*Grin2d*^{V664I/+} mice and wild-type littermate controls (Sox2-Cre/*Grin2d*^{+/+}). Consistent with previous reports examining this *Grin2d* variant (Teoh J, 2025), (Yam M, 2025), Sox2-Cre/*Grin2d*^{V664I/+} mice exhibited spontaneous seizures including frequent SWDs, providing a rigorous test of classifier generalizability.

We next applied the classifier, trained on kainic acid-induced seizures, to spontaneous SWDs in *Grin2d*-V664I mice without retraining. Performance remained high (AUC > 0.8 across animals), demonstrating robust generalization to subtle, spontaneous SWD events (Supp. Fig. 3). Retraining the classifier on a combined dataset (KA + *Grin2d*-V664I) further improved performance, particularly in distinguishing *Grin2d*-specific seizure types (Figure 4). Together, these results establish that multimodal EEG–video classification generalizes across both seizure types and genetic backgrounds.

We retrained five separate SVM classifiers using EEG features, behavioral features, top-camera video, side-camera video, or combined multimodal data (Fig. 4C). ROC and precision–recall analyses again demonstrated that the combined model achieved superior performance compared to unimodal approaches (Fig. 4D).

Examination of classifier confidence distributions revealed that the combined model produced more decisive classifications, with prediction scores clustering near 0 (non-seizure) and 1 (seizure) (Fig. 4D). Distance from the decision boundary further confirmed greater separation between seizure and non-seizure states for the combined model relative to individual modalities. Notably, the combined model maintained high performance values in each category, including more subtle spontaneous SWDs, indicating robust generalization of the model for diverse seizure types (Fig. 4E).

Analysis of feature importance identified the most discriminative variables for detecting spontaneous seizures in the *Grin2d* mutants (Fig. 4F). Consistent with the kainic acid experiments, EEG theta power (5.5–7.5 Hz) and spike rate were highly weighted in models incorporating electrophysiological data. Behavioral arrest and head rearing again emerged as critical behavioral indicators, while movement features from both camera angles contributed to video-based classification. The combined model effectively integrated complementary information across modalities, validating the generalizability of the multimodal approach to spontaneous seizure detection.

Identification of Spike-Wave Discharge Subtypes

Detailed analysis of the *Grin2d*-V664I seizure phenotype revealed heterogeneous spike-wave discharge (SWD) patterns that could be categorized into distinct electrographic subtypes. Representative EEG traces illustrated typical SWDs characterized by regular, spindleform, high-amplitude, spike-wave complexes, as well as atypical SWDs exhibiting irregular morphology, variable amplitude, and inconsistent frequency (Fig. 5A). Principal component analysis of SWD features demonstrated clear separation between typical SWDs, atypical SWDs, and non-seizure activity in multidimensional feature space (Fig. 5B, Supp. Fig. 4B), but not in interpretable feature space (Supp. Fig. 4A), indicating that these represent distinguishable electrographic entities.

Quantitative comparison between wild-type and *Grin2d*-V664I mice revealed significant differences in seizure burden. Mutant animals exhibited a significantly higher rate of SWDs per minute compared to wild-type littermates (Fig. 5C, left; $p < 0.01$). SWD duration showed a trend toward longer events in V664I mice, though this did not reach statistical significance (Fig. 5C, middle; $p = 0.097$). The fundamental frequency of SWDs did not differ between genotypes (Fig. 5C, right; n.s.).

Next, we sought to determine whether SWDs in *Grin2D*-V664I animals themselves represented a homogeneous seizure type or comprised distinct subtypes. Further examination of SWD heterogeneity through PCA revealed three clusters corresponding to non-seizure activity that we refer to as typical SWDs and atypical SWDs (Fig. 5D, left). These clusters occupied separable regions of feature space, confirming their distinct electrographic profiles. Representative EEG traces from each cluster illustrated the behavioral correlates of these subtypes (Fig. 5D, right). Typical SWDs were predominantly associated with behavioral arrest, consistent with absence-like seizures (Movie 1). By contrast, atypical SWDs occurred during active behavioral states including rearing, head movement, and running or darting, and notably did not interrupt ongoing behavior or trigger behavioral arrest (Movie 2). Importantly, the detected atypical SWDs were not explained by differences in head movement: WT and *Grin2D*-V664I mice exhibited comparable levels of head movement despite the markedly higher proportion of SWD-like events in the mutants (Supp. Fig. 5). A similar qualitative distinction between SWD-like events with and without behavioral arrest was previously noted by the authors in Teoh et al. (Teoh J, 2025). The present approach builds on these observations by providing a systematic and data-driven characterization of SWD heterogeneity, enabling a more precise delineation of seizure subtypes.

Quantification of these behavioral associations confirmed that typical SWDs were significantly more frequent than atypical discharges and showed a significantly stronger association with behavioral arrest compared to periods of behavioral activity (Fig. 5E; $p < 0.0001$). This distinction between behavioral and non-behavioral SWDs is an example of phenotypic classification enabled by the multimodal approach.

To assess the sensitivity and specificity of automated detection, we directly compared machine learning classifications with human expert scoring of SWD events. Comparison revealed substantial concordance, with 2,189 events identified by both methods (Fig. 5F). Notably, the machine learning approach detected an additional 687 events (~23.23% of total events) not initially scored by human observers suggesting enhanced sensitivity for brief or subtle discharges. Conversely, 81 events were identified exclusively by the human expert (~2.74% of total events), likely representing edge cases where contextual judgment informed manual scoring. Indeed, agreement between human and machine detection decreases as events approach the non-seizure cluster in PCA space (Fig. 5G).

PCA visualization of the classified events demonstrated that human-only, machine-only, and concordantly identified events occupied overlapping regions of feature space, with concordant detections forming a distinct, tight cluster (Fig. 5G). This spatial organization suggests that while human and machine approaches show strong overall agreement, they exhibit differences in detection thresholds and feature weighting. Machine-detected events that were missed by human scoring occupied similar feature space to concordant events, supporting their validity as true SWDs rather than false positives. Together, these results demonstrate that the multimodal pipeline achieves high sensitivity for detecting spontaneous SWDs while maintaining strong agreement with expert manual scoring, and importantly, uncovers phenotypic diversity (behavioral vs. non-behavioral SWDs) that would be difficult to identify reliably and quantify through traditional manual scoring approaches alone.

Discussion

We present an integrated machine learning pipeline that combines EEG and video data for automated seizure detection in rodent models. By validating this approach across two distinct seizure paradigms (kainic acid-induced seizures and complex spontaneous spike-wave seizures in a *Grin2d*-V664I mouse model) we demonstrate that multimodal integration consistently outperforms unimodal approaches and reveals phenotypic diversity that would be difficult to capture and quantify through traditional manual scoring methods.

The superior performance of the combined EEG-video classifier reflects the complementary nature of these input modalities. EEG provides high temporal resolution and direct measurement of neural synchrony, enabling detection of seizure events that do not have obvious behavioral correlates. Conversely, video-derived behavioral features offer critical contextual information that can disambiguate movement artifacts from true seizure activity and identify motor manifestations that may not be detectable using EEG or EMG. This complementarity was evident in both seizure models: in the kainic acid paradigm, the combined model achieved more decisive classification boundaries and better separated ambiguous events near the detection threshold, while in the *Grin2d* model, multimodal integration

enabled detection of 687 additional SWD events beyond human expert scoring while maintaining high concordance with manual annotations.

Importantly, our validation strategy, which consisted of training on overt kainic acid seizures and then testing generalization to spontaneous SWDs, demonstrated robustness across seizure types that differ substantially in their electrographic signatures, behavioral manifestations, and temporal dynamics. The consistency of feature importance patterns across models (theta power, spike rate, and specific behavior) suggests that the classifier captures biologically relevant seizure features rather than model-specific artifacts. Generalizability is critical for practical application, as it indicates the pipeline can be adapted to diverse seizure models without extensive retraining.

Defining SWDs subclasses

SWDs are seizure events that typically appear in mice as brief 5-8 Hz high amplitude events in the EEG. SWDs have been identified in numerous rodent models carrying single gene defects, as well as some inbred rodent lines (Letts VA, 2014). However, SWDs have also been observed in out-bred populations of rodents (Polack, 2006), and their resemblance to natural brain states such as sleep spindles has called into question the pathogenic nature of SWDs in rodents (Kozák G, 2020), (Beenhakker MP, 2009). Adding to this controversy, there is a long-standing appreciation that there exists developmentally-regulated changes in SWD occurrence and features and highly variable impact of these events on cognition and motor behaviors (Pearce PS, 2014).

The neurophysiological basis of SWDs is a topic of debate as elegant studies have demonstrated a central role of the thalamocortical loop in SWD generation (Crunelli V, 2011) (Paz JT, 2011) (Makinson CD, 2017). While SWD events have been detected within cortical circuits that are disconnected from the thalamocortical loop and following lesioning of the somatosensory thalamus (Steriade M, 1993). The most parsimonious explanation for these findings is that SWD events may be generated and regulated by distinct cortical and subcortical oscillators. Particularly in the case of genetic absence epilepsy involving widely expressed genes, such as *Grin2d*, impairments in multiple SWD circuits may generate multiple SWD types.

A key finding of this study is the detection of distinct SWD subtypes that can be defined, in part, by behavioral phenotypes. Traditional approaches to SWD analysis focus primarily on interpretable electrographic features such as frequency, amplitude, and duration, treating behavioral manifestations as secondary. Our multimodal analysis revealed that SWDs reliably cluster into at least two distinct categories: typical SWDs strongly associated with behavioral arrest, resembling typical absence seizures, and atypical SWDs that occur during active behaviors without interrupting ongoing motor activity.

We speculate that behavioral SWDs, which arrest ongoing activity, represent more severe disruptions of cortical rhythm-generating circuits and may correlate with the cognitive impairments documented in patients with absence epilepsy and some rodent models (Buzsáki, 2006), (Blumenfeld, 2005). Previous reports of SWDs generated from isolated cortex appear to resemble more closely the behavioral seizures we report in *Grin2d*-V664I animals (McCafferty C, 2018) (Kandel A, 1997) (Meeren HK, 2002). Notably, behavioral SWDs tend to be treatment resistant and are a hallmark of Lennox–Gastaut syndrome in humans, suggesting that these events reflect a particularly pathological form of network dysfunction (Arzimanoglou, 2009) (Polavarapu, 2024).

Multimodal analysis commentary

The automated pipeline offers several advantages over traditional manual scoring. First, it eliminates inter-rater variability, which is particularly problematic for ambiguous events such as brief SWDs or low-amplitude epileptiform discharges. Second, it enables high-throughput analysis of long-duration recordings, facilitating longitudinal studies and large-scale phenotyping efforts that would be prohibitively time-consuming with manual approaches. Third, automated analysis can detect subtle events that human observers may miss, potentially increasing sensitivity for rare or brief seizures. Indeed, prior work using hybrid wearable/EEG systems has shown strong sensitivity (~0.91) versus the video-EEG gold standard, albeit with non-negligible false positive rate (Naganur V, 2022).

However, several limitations of the study should be acknowledged. One limitation is that the current implementation focuses on binary seizure/non-seizure classification within predefined categories (e.g.,

SWDs). While we successfully identified subtypes within the SWD category post hoc using unsupervised clustering, the pipeline does not automatically discover novel seizure types without human guidance. Incorporating unsupervised or weakly supervised learning methods could enable true discovery-driven phenotyping, where the algorithm identifies electrographic-behavioral patterns without predefined labels.

The relatively small sample size ($n=9$ KA, $n=5$ *Grin2d*-V664I, $n=3$ WT) we recorded from represent a practical limitation imposed by the labor-intensive nature of VEEG experiments and the need to limit the use of animals to the smallest number required to train each model. Our cross-validation strategy, including held-out animal testing, and the strong concordance with human expert scoring suggest that the findings are robust with the number of animals included in the study. Moreover, the sample sizes are comparable to or larger than many published rodent EEG studies, and the discovery of distinct SWD subtypes with clear statistical separation argues against these being spurious findings. (Sanchez-Pernaute, 2022)

Reproducibility/accessibility/simplicity

A key design principle of this pipeline was accessibility and reproducibility. We deliberately chose established, open-source tools (DeepLabCut and SimBA) rather than *de novo* developed algorithms, ensuring that other laboratories can implement and adapt the approach without specialized expertise or resources (Nath T, 2019). The use of support vector machines, while potentially less sophisticated than deep learning approaches, offers several important advantages: SVMs are interpretable (enabling feature importance analysis), require relatively modest training data, perform well in high-dimensional feature spaces, and are computationally efficient (Saeidi M, 2021). Our results demonstrate that carefully engineered features with a classical machine learning approach can match or exceed the performance of end-to-end deep learning when training data are limited (Saeidi M, 2021).

The dual-camera setup provides complementary perspectives that improve behavioral classification, though our results indicate that a single top-mounted camera captures most of the relevant information for seizure detection. This suggests that laboratories with simple recording setups can still implement multimodal analysis.

Temporal synchronization between EEG and video is critical for accurate multimodal integration. We addressed the 33 ms delay inherent to the Azure camera system through timestamp alignment, but researchers using different hardware should carefully characterize and correct for system-specific latencies. Poor synchronization could lead to misalignment of electrographic and behavioral features, degrading classifier performance.

Broader Implications and Future Directions

While we focused on seizure detection, the framework is broadly applicable to many problems requiring integration of neural and behavioral data. Examples include studies of movement disorders, sleep architecture, social behaviors, or decision-making tasks where linking neural activity to behavioral output is scientifically central. We propose that a productive future application of this approach may be for the tracking of state transition dynamics. The pipeline could also be extended to include the simultaneous analysis of multiple interacting animals, additional data modalities such as local field potentials, single-unit recordings, or even non-neurological physiological measures (heart rate, respiration), potentially revealing even richer structure in body/brain-behavior relationships.

Conclusion

Integrating video behavioral analysis with electrophysiology through machine learning provides a useful framework for automated seizure detection and phenotyping. This approach not only matches or exceeds human expert performance but also can effectively reveal hidden phenotypic diversity. By demonstrating robust performance across fundamentally different seizure models and providing a reproducible, accessible implementation, this work establishes multimodal integration as a practical tool for epilepsy research. More broadly, it illustrates how thoughtful combination of complementary data modalities with interpretable machine learning can enhance both the efficiency and scientific depth of neurobehavioral analysis.

The discovery of SWD subtypes with distinct behavioral correlates underscores the value of moving beyond unimodal analyses. Future progress in understanding seizure mechanisms, developing

therapeutic interventions, and identifying biomarkers will likely require similar integrative approaches that capture the full complexity of brain-behavior relationships. The framework presented here provides a foundation for such efforts, offering a scalable, reproducible, and scientifically generative approach to analyzing the rich multimodal data generated by modern neuroscience experiments.

Acknowledgements:

We thank Drs Larry Abbott and Mu Yang for their support and insightful feedback on the manuscript. This work was supported by the National Institutes of Health (NIH) R21 NS121980 to C.D.M. and W.N.F., a United States - Israel Binational Science Foundation (BSF) grant to W.N.F. and C.D.M., a Cure Lennox-Gastaut Syndrome 365 Research Grant to C.D.M., a generous donation from an anonymous donor, a National Science Foundation Graduate Research Fellowships Program (GRFP) DGE-2036197 award to C.J.L. and the NIH/NINDS Training Fellowship T32NS064929 to C.J.L.

References

- Arzimanoglou, A. J.-P. (2009). Lennox-Gastaut syndrome: a consensus approach on diagnosis, assessment, management, and trial methodology. *The Lancet Neurology*, 8(1), 82-93.
- Beenhakker MP, H. J. (2009). Neurons that fire together also conspire together: is normal sleep circuitry hijacked to generate epilepsy? *Neuron*, 62(5), 612-632.
- Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1), 527-570.
- Blumenfeld, H. (2005). Cellular and network mechanisms of spike-wave seizures. *Epilepsia*, 46, 21-33.
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press.
- Camp CR, Y. H. (2020). GRIN2D/GluN2D NMDA receptor: Unique features and its contribution to pediatric developmental and epileptic encephalopathy. *Eur J Paediatr Neurol*, 24, 89-99.
- Casillas-Espinosa, P. &. (2019). A universal automated tool for reliable detection of seizures in rodent models of acquired and genetic epilepsy. *Epilepsia*.
- Colombo S, R. H. (2023). Epilepsy in a mouse model of GNB1 encephalopathy arises from altered potassium (GIRK) channel signaling and is alleviated by a GIRK inhibitor. *Front Cell Neuroscience*.
- Crunelli V, C. D. (2011). Transition to absence seizures and the role of GABA(A) receptors. *Epilepsy Res*, 97(3), 283-9.
- Fisher RS, A. C. (2014). ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*, 475-82.
- Garção VM, A. M. (2023). A novel approach to automatic seizure detection using computer vision and independent component analysis. *Epilepsia*, 64, 2472-2483.

657 Goodwin, N. C. (2024). Simple Behavioral Analysis (SimBA) as a platform for explainable
658 machine learning in behavioral neuroscience. *Nature Neuroscience*, 27, 1411-1424.

659 Jing J, S. H. (2020). Development of Expert-Level Automated Detection of Epileptiform
660 Discharges During Electroencephalogram Interpretation. *JAMA Neurol*, 77(1), 103-108.

661 Kaculini CM, T.-L. A. (2021). The History of Epilepsy: From Ancient Mystery to Modern
662 Misconception. *Cureus.*, 13(3).

663 Kandel A, B. G. (1997). Cellular-synaptic generation of sleep spindles, spike-and-wave
664 discharges, and evoked thalamocortical responses in the neocortex of the rat. *J*
665 *Neurosci*, 17(17), 6783-97.

666 Kotloski, R. (2023). A machine learning approach to seizure detection in a rat model of post-
667 traumatic epilepsy. *Science Rep*, 13.

668 Kozák G, F. T. (2020). Spike-and-Wave Discharges Are Not Pathological Sleep Spindles,
669 Network-Level Aspects of Age-Dependent Absence Seizure Development in Rats.
670 *eNeuro*, 7(1).

671 Letts VA, B. B. (2014). Hidden in plain sight: spike-wave discharges in mouse inbred strains.
672 *Genes Brain Behav*, 13(6), 519-526.

673 Li, J. Z. (2014). Classifying Normal and Abnormal Status Based on Video Recordings of
674 Epileptic Patients. *The Scientific World Journal*.

675 Liu C, S. J. (2020). An Evolving-Dynamic Network Activity Approach to Epileptic Seizure
676 Prediction using Machine Learning. *Arxiv*.

677 Lucas, A. R. (2024). Artificial intelligence in epilepsy — applications and pathways to the clinic.
678 *Nature Rev Neurology*, 20, 319-336.

679 Makinson CD, T. B. (2017). Regulation of Thalamic and Cortical Network Synchrony by Scn8a.
680 *Neuron*, 93(5), 1165-1179.

681 Makinson, C., Tanaka, B., Lamar, T., Goldin, A., & Escayg, A. (2014). Role of the hippocampus
682 in Nav1.6 (Scn8a) mediated seizure resistance. *Neurobiology of Disease*.

683 Manzouri F, H. S.-B. (2018). A Comparison of Machine Learning Classifiers for Energy-Efficient
684 Implementation of Seizure Detection. *Front Syst Neurosci*, 12.

685 Mathis, A. M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with
686 deep learning. *Nature Neuroscience*, 21, 1281-1289.

687 McCafferty C, D. F. (2018). Cortical drive and thalamic feed-forward inhibition control thalamic
688 output synchrony during absence seizures. *Nature Neuroscience*, 21(5), 744-756.

689 McCormick DA, C. D. (2001). On the cellular and network bases of epileptic seizures. *Annu Rev*
690 *Physiol*, 815-46.

691 Meeren HK, P. J. (2002). Cortical focus drives widespread corticothalamic networks during
692 spontaneous absence seizures in rats. *J Neuroscience*, 22(4), 1480-95.

693 Naganur V, S. S.-B. (2022). Automated seizure detection with noninvasive wearable devices: A
694 systematic review and meta-analysis. *Epilepsia*, 63(8), 1920-1941.

695 Nath T, M. A. (2019). Using DeepLabCut for 3D markerless pose estimation across species and
696 behaviors. *Nat Protoc*, 14(7), 2152-2176.

697 Paz JT, B. A. (2011). new mode of corticothalamic transmission revealed in the Gria4(-/-) model
698 of absence epilepsy. *Nature Neuroscience*, 14(9), 1167-73.

699 Pearce PS, F. D. (2014). Spike-wave discharges in adult Sprague-Dawley rats and their
700 implications for animal models of temporal lobe epilepsy. *Epilepsy Behav*, 32, 121-131.

701 Polack, P.-O. a. (2006). Intracellular activity of cortical and thalamic neurones during high-
702 voltage rhythmic spike discharge in Long-Evans rats in vivo. *The Journal of Physiology*,
703 571, 461-476.

704 Polavarapu, A. (2024). Epilepsy with myoclonic-atonic seizures. *MedLink Neurology*.

705 Purcell RH, P. L. (2013). Effects of an epilepsy-causing mutation in the SCN1A sodium channel
706 gene on cocaine-induced seizure susceptibility in mice. *Psychopharmacology (Berl)*.

707 Sabilallah M, F. P. (2016). Evidence for Status Epilepticus and Pro-Inflammatory Changes after
708 Intranasal Kainic Acid Administration in Mice. *PLOS ONE*, 11(3).

709 Saeidi M, K. W.-J. (2021). Neural Decoding of EEG Signals with Machine Learning: A
710 Systematic Review. *Brain Sci*, 11(11).

711 Sanchez-Pernaute, G. A. (2022). Mice in translational neuroscience: What R we doing?
712 *Progress in Neurobiology*, 217.

713 Steriade M, N. A. (1993). Intracellular analysis of relations between the slow (< 1 Hz)
714 neocortical oscillation and other sleep rhythms of the electroencephalogram. *J Neurosci*,
715 13(8), 3266-83.

716 Svantesson, M. &. (2024). Interrater agreement of annotations of epileptiform discharges and its
717 impact on deep learning: A pilot study. *medRxiv*.

718 Teoh J, S. J. (2025). Synaptic dysregulation in a mouse model of GRIN2D developmental and
719 epileptic encephalopathy. *Brain*.

720 Tveit J, A. H. (2023). Automated Interpretation of Clinical Electroencephalograms Using Artificial
721 Intelligence. *Jama Neurol*, 80(8), 805-812.

722 Wang W, W. D. (2024). Impaired axon initial segment structure and function in a model of
723 ARHGEF9 developmental and epileptic encephalopathy. *Proc Natl Acad Sci*.

724 Wang Y, W. P. (2022). Animal Models of Epilepsy: A Phenotype-oriented Review. *Aging Dis*,
725 13(1), 215-231.

726 Wilson JV, R. E. (1990). Texts and documents. Translation and analysis of a cuneiform text
727 forming part of a Babylonian treatise on epilepsy. *Medical History*., 34(2), 185-198.

- Wirrell EC, N. R. (2022). Methodology for classification and definition of epilepsy syndromes with list of syndromes: Report of the ILAE Task Force on Nosology and Definitions. . *Epilepsia.*, 63(6), 1333-1348.
- Wong JC, M. C. (2018). Selective targeting of Scn8a prevents seizure development in a mouse model of mesial temporal lobe epilepsy. *Science Reports*.
- Y. Yang, R. A. (2021). Video-Based Detection of Generalized Tonic-Clonic Seizures Using Deep Learning. *IEEE Journal of Biomedical and Health Informatics*, 25(8), 2997-3008.
- Yam M, N. J.-C.-V. (2025). A mouse model of GRIN2D developmental and epileptic encephalopathy recapitulates the human disease. *Brain*.
- Yan, K. L. (2025). Automated seizure detection in epilepsy using a novel dynamic temporal-spatial graph attention network. *Science Rep*, 15.

Figure Legends:

Figure 1. Integrated pipeline for multimodal seizure detection combining EEG and video analysis.

Schematic overview of the automated seizure detection workflow. Mice are monitored simultaneously using telemetric EEG (left) and dual-perspective video recording from top and side views (center). DeepLabCut is used to track seven body keypoints (snout, left ear, right ear, mid body, tail base, tail mid, tail tip) from video data. Tracked keypoints are input into SimBA for automated detection of behavioral states (Behavior A, Behavior B, etc.). In parallel, EEG traces undergo automated feature extraction to identify electrographic characteristics. Behavioral features (X_Behavior) and EEG features (X_EEG) are concatenated into a combined feature matrix ($X = [X_Behavior \mid X_EEG]$). Support vector machine (SVM) classifiers are trained on these features to generate seizure state classifications. The pipeline involves iterative refinement: initial classifier outputs inform refined EEG feature definitions and behavior/EEG state classifications, which are used to retrain classifiers for improved performance. Final outputs include EEG-based features, behavior-based features, and integrated EEG/behavior seizure state classifications.

Figure 2. Multimodal seizure detection outperforms unimodal approaches in kainic acid-induced seizures.

(A) Experimental design. Wild-type mice received intraperitoneal kainic acid injection (20 mg/kg) to induce acute seizures and were monitored using telemetric EEG and dual video cameras. Seizure state classifications were independently generated from EEG traces, behavioral observations, and combined EEG/behavioral annotations for classifier training. (B) Five separate SVM classifiers were trained and compared: EEG only, behavior only, top camera only, side camera only, and a combined multimodal model integrating all features. (C) Classifier performance comparisons. Left: Receiver operating characteristic (ROC) curves plotting true positive rate versus false positive rate, with the combined model achieving the highest area under the curve. Right: Precision-recall (PR) curves demonstrating maintained precision across recall thresholds, with superior performance of the multimodal classifier. Classifier confidence distributions showing prediction score frequencies for seizure versus non-seizure states. The combined model produces more decisive classifications with reduced overlap between classes. (D) Quantitative performance metrics for each classifier. Confusion matrix shows accuracy, F1 score, precision, and recall. (E) Feature importance analysis across SVM classifiers. Y axis represents the relative weight (W) of each feature in classification decisions. Theta-band EEG power (5.5–7.5 Hz) and spike rate are the most discriminative electrophysiological features. Behavioral arrest,

head nodding, and wild running are the strongest behavioral predictors. Movement features from top and side cameras contribute to video-based models.

Figure 3. Automated detection captures seizure progression and behavioral complexity. (A) Representative EEG traces of WT mice's baseline activity, epileptiform discharges, moderate seizures, and severe generalized tonic-clonic seizures (GTCS) following kainic acid injection. Scale bars: 500 μ V, 1 s (top three traces), 10 s (bottom trace). (B) Latency to onset of different seizure behaviors across four representative mice. Behaviors progress from simple (head nodding, rearing) to complex (tonic-clonic, terminal GTCS) manifestations. (C) Principal component analysis of seizure events demonstrating clustering by severity. PC1-PC3 effectively separate baseline, epileptiform activity, moderate seizures, and GTCS. Black dots denote lethal seizure. Relationship between seizure severity and behavioral complexity is quantified as number of distinct behaviors observed per event and shown via size. (E) Comparison of seizure event identification between human expert scoring and machine learning classification. Venn diagram shows 29 events detected by both methods, 18 by machine learning approach only, and 1 by human only. (E) Principal component analysis of seizure events detected by machine learning approach only, human only, and both, as well as some baseline examples. Right: Representative EEG traces illustrate examples of events classified by machine only. Scale bars: 500 μ V, 250 ms or 2 s as indicated.

Figure 4. Multimodal seizure detection generalizes to spontaneous seizures in Grin2d mutant mice. (A) GluN2D receptor structure showing the V664I mutation location in the ligand-binding domain (ABD: agonist-binding domain; ATD: amino-terminal domain; CTD: C-terminal domain; M1-M4: transmembrane domains) (B): Top: Schematic of the Grin2d "knockout first" conditional knock-in model. Bottom: Experimental setup showing simultaneous telemetric EEG recording with dual-perspective video monitoring (top and side views) applied to Grin2d mutant mice. (C) Classification framework comparing five SVM models: EEG only, behavior only, top camera only, side camera only, and combined multimodal approach. (D) Classifier performance metrics. Left: ROC curves showing true positive rate versus false positive rate. Middle: Precision-recall curves. Right: Distribution of prediction scores (classifier confidence) for seizure versus non-seizure states. The combined model shows superior performance across all metrics. Classifier confidence distributions measured as absolute distance from decision boundary. The combined model achieves greater separation between seizure and non-seizure predictions. (E) Quantitative performance metrics for each classifier. Confusion matrix shows accuracy, F1 score, precision, and recall. The combined multimodal model consistently achieves the highest performance across all metrics, demonstrating the advantage of integrating electrophysiological and behavioral data for SWD detection. (F) Feature importance weights (W) across SVM classifiers. Y axis indicates relative contribution of each feature to classification decisions. EEG features (theta power, spike rate), behavioral features (arrest, rearing, wild running), and movement metrics from both camera perspectives contribute differentially across models, with the combined model integrating the most informative features from each modality.

Figure 5. Automated detection and classification of spike-wave discharge subtypes in Grin2d mutant mice. (A) Representative EEG traces showing typical SWDs (regular, high-amplitude spike-wave complexes) and atypical SWDs (irregular morphology and frequency) in Grin2d-V664I mice. Scale bars: 500 μ V, 1 s. (B) Principal component analysis demonstrating separation between typical SWDs, atypical SWDs, and non-seizure activity in multidimensional feature space. (C) Quantitative comparison of SWD characteristics between wild-type (WT) and Grin2d-V664I (V664) mice. Box-and-whisker plots show SWD frequency (left; events per minute, $^{**}p < 0.01$), SWD duration (middle; $p = 0.097$, n.s.), and fundamental frequency (right; n.s.). (D) Left: PCA visualization of events classified as not-seizure (blue), typical SWDs (red), and atypical seizures that didn't fit the classic SWD (purple). Right: Representative EEG traces from each cluster showing SWD events with different behavioral states (rearing, head movement, darting, behavioral arrest). Scale bars: 500 μ V, 1 s. (E) Quantification of typical versus atypical SWD frequency

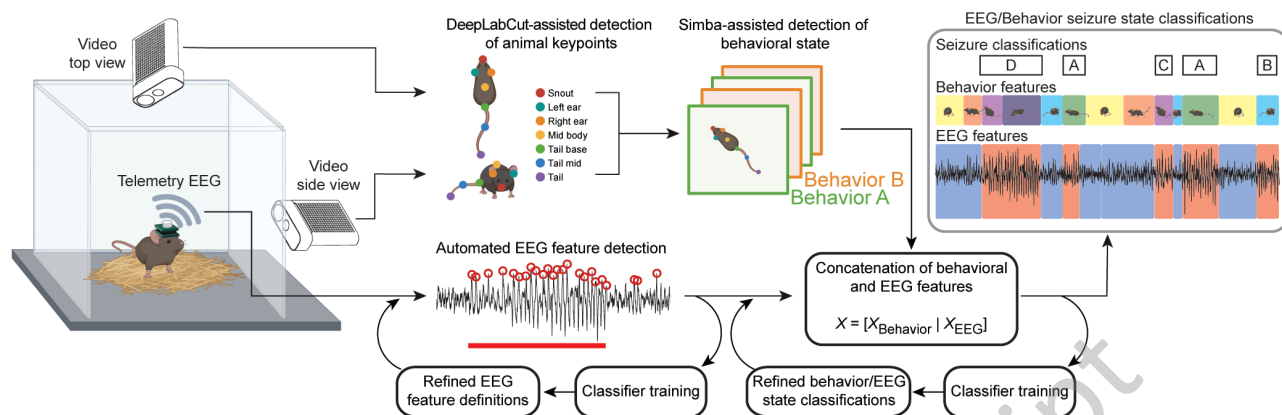
and association with behavioral arrest. Typical SWDs are significantly more frequent and more strongly associated with behavioral arrest (**** $p < 0.0001$). (F) Comparison of SWD identification between human expert scoring and machine learning classification. Venn diagram shows 2,189 events identified by both methods and 687 additional events detected by machine learning alone and 81 events detected by only the expert human. (G) PCA visualization of events classified by human only (blue), machine only (red), both methods (green), or not seizure (gray), demonstrating a distinct cluster of concordant events.

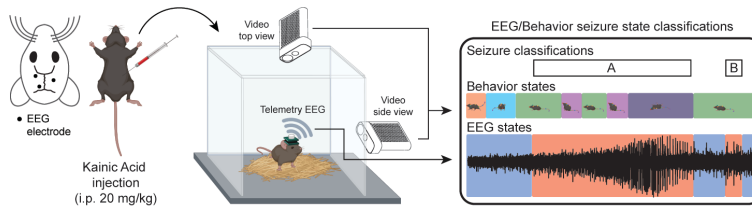
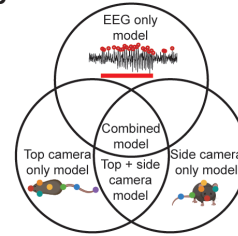
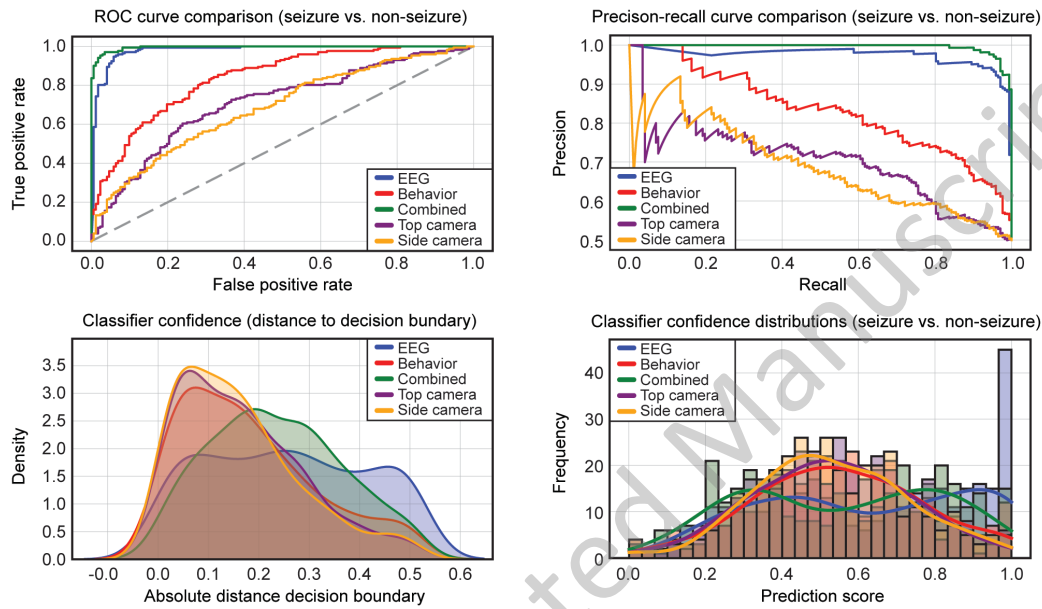
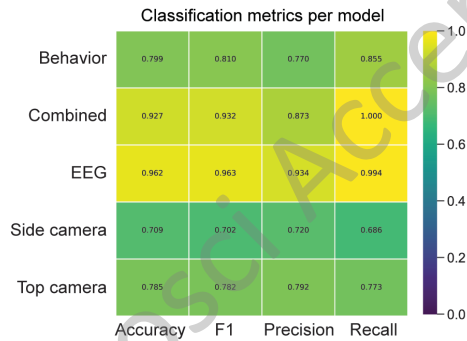
Movie 1. Video-EEG recording of a typical SWD event in a *Grin2d* mutant animal

Representative movie showing concurrent video and telemetric EEG recording of a typical SWD event in a *Grin2d* mutant animal. The EEG record is shown on the left, with simultaneous video recordings of mouse behavior from the side-mounted camera (bottom) and top-mounted camera (right). A SWD that occurs during a period of behavioral arrest is highlighted in red.

Movie 2. Video-EEG recording of an atypical behavioral SWD event in a *Grin2d* mutant animal

Representative movie showing concurrent video and telemetric EEG recording of an atypical SWD event in a *Grin2d* mutant animal. The EEG record is shown on the left, with simultaneous video recordings of mouse behavior from the side-mounted camera (bottom) and top-mounted camera (right). A SWD that is accompanied by a behavioral manifestation is highlighted in red.



a**b****c****d****e**