

Automated Detection of Seizures in Rodents: Use of Artificial Intelligence (AI) to Improve Racine Score Scoring

Summan Bahadur

University of the Punjab

Syed Faisal Bukhari

University of the Punjab

Aneeqa Saleem

`aneeqa.pharma@mul.edu.pk`

Minhaj University Lahore

Article

Keywords: Seizure, Epilepsy, Mice, Racine scale, YOLOv8, Deep SORT

Posted Date: August 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10034489/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Automated Detection of Seizures in Rodents: Use of Artificial Intelligence (AI) to Improve Racine Score Scoring

Summan Bahadur¹, Dr. Faisal Bukhari¹, Dr. Aneeqa Saleem^{2*}

¹ Department of Data Science, University of the Punjab, Lahore, Pakistan

² Department of Pharmacy, Minhaj University, Lahore, Pakistan

***Corresponding author: Dr. Aneeqa Saleem (aneeqa.pharma@mul.edu.pk)**

Abstract

Epilepsy is a chronic neurological condition that presents episodic and unprovoked seizures, resulting in significantly impacting the quality of life. The presented research aims to revolutionize conventional seizure detection methods by leveraging the power of AI and considerably improving the accuracy and efficiency of seizure detection. Specifically, the YOLOv8 model detects and classifies alongside Deep-Sort for tracking rodents. This research used a set of videos recorded under high-confidence, controlled conditions designed for developing novel plant-based therapy directed at epilepsy. The YOLOv8 model was trained and detected rodents with an average precision of 77.1%. Moreover, Deep-Sort provided successful and efficient tracking of rodents. The outcomes of the YOLOv8 classification model highlight its accuracy in predicting seizure classification. The model performed particularly well in classifying 'stage-4' and 'normal-stage' with high precision and accuracy, and 'Hind-Limb-Extension' with strong recall. These results imply that the model effectively identified hind limb extension and the two stages with a very low misclassification rate, demonstrating the model's utility for detecting both seizure and normal conditions.

Keywords: Seizure; Epilepsy; Mice; Racine scale; YOLOv8; Deep SORT

Introduction

Seizures are caused by dysrhythmia in the brain, which results in sudden electric impulses producing involuntary abnormal movements including changes in feelings, movements, behaviours, and consciousness level^{1,2}. Epilepsy is a disease that results from having an unprovoked seizure with a yielding threat for impending unprovoked seizures, or possibly having two or more unprovoked seizures^{3,4}. According to the World Health Organization (WHO), epilepsy is one of the most common neurological disorders worldwide, with an estimated 50 million people affected globally⁵. It is projected that nearly 70% of individuals living with epilepsy are likely to become seizure-free when appropriately identified and treated⁵, highlighting the critical need for improved diagnosis and monitoring.

Studies on epilepsy are performed using animal models to help understand how seizures work and to identify treatments that could potentially restore brain function⁶. Considerable importance has been given to establishing rodent models, especially albino mice, given their genetic equivalence with humans and the scope for genome manipulation capable of simulating human epilepsy⁷. In

such studies, one standard way of evaluating the effectiveness of therapeutic agents is to grade seizure severity according to the Racine scale, which provides a standardised method to estimate seizure severity and determine the effectiveness of antiepileptic drugs⁸. Nevertheless, the scale relies on manual observation and grading of seizures, which is laborious, time-consuming, and susceptible to mistakes from human error as well as observer bias⁹.

Artificial Intelligence (AI), especially with sophisticated algorithms, can improve the accuracy of seizure detection and grading, making it faster and more reliable¹⁰. Despite many promising applications of AI, the field has yet to fully flourish in terms of automatic detection and grading of seizures in rodent models. To address this gap, this study provides automatic seizure detection and grading using the You Only Look Once (YOLO) algorithm in combination with Deep-Sort tracking. The incorporation of these cutting-edge technologies is predicted to greatly improve the efficiency, accuracy, and consistency of seizure analysis in rodent models.

Results

Rodent Detection Using YOLOv8

The first model was trained on a dataset annotated with one class label—Rodent—for object detection. Training was performed using pre-trained weights (Yolov8n.pt) over 50 epochs with an image size of 640×640 pixels. The training and validation loss curves demonstrated convergence of the model, with performance metrics showing high accuracy and stability. Fluctuations observed in the validation metrics reflect performance variability arising from different validation subsets or from the model learning complex patterns across frames.

The confusion matrix confirmed high accuracy in predicting rodent presence. Validation batch outputs showed that the trained model achieved confident detections with a maximum confidence score of 0.8 and consistently accurate bounding boxes for all detected subjects (see Figure 9 in the Methods section).

Tracking with Deep-Sort

Deep SORT (Simple Online and Real-time Tracking) was employed to track detected rodents across all frames of each video. The algorithm assigned a unique tracking ID to each detected rodent and maintained this ID persistently across frames by combining motion-based prediction (Kalman filter) with deep learning-based appearance descriptors. The tracker inferred rodent positions in new frames by estimating location based on prior positions and updating with new detections, enabling robust tracking of individual rodents throughout the full video sequence.

Results demonstrated successful tracking with IDs correctly assigned to individual subjects. One limitation observed was that tracking IDs assigned by the algorithm were not sequential and did not correspond to the manually assigned IDs used by the researchers prior to recording. This can be addressed in future work by relying exclusively on the tracker for labelling rather than pre-labelling before recording.

Seizure Classification

The seizure classification model was trained over 100 epochs using an image size of 640×640. The model was tasked with classifying seizure stages based on the Racine scale. Due to the subtle nature of early-stage motor changes (stages 1–3), the model's focus was concentrated on stages 4 and 5 as well as tonic-clonic hind limb extension, which exhibit more pronounced and visually distinct behaviours.

The PR curves for the classification model showed values approaching 1, indicating high confidence in predictions. A 'Normal' class was introduced to represent non-seizure frames, which improved the balance of the dataset and reduced incorrect seizure classification.

Precision and Recall for Each Class

Average precision (AP) was calculated for each class using the standard detection metrics:

Class	TP	FP	FN	Precision
Hind-Limb-Extension	8	2	3	0.80 / 0.727
Stage-4	11	6	1	0.647 / 0.917
Normal	13	2	1	0.867 / 0.929

Table 1. Precision and recall metrics for each seizure classification class.

Precision and recall for the Hind-Limb-Extension class were 0.80 and 0.727 respectively (TP=8, FP=2, FN=3). For Stage-4, precision was 0.647 and recall was 0.917 (TP=11, FP=6, FN=1). For the Normal class, precision was 0.867 and recall was 0.929 (TP=13, FP=2, FN=1). The mean Average Precision (mAP) across all three classes was 0.771 (77.1%), calculated as the mean of per-class precision values. The model performed well in detecting 'Normal' and 'Hind-Limb-Extension' instances, with high precision and recall, while Stage-4 showed lower precision, indicating room for further optimisation.

Discussion

This study demonstrates that AI-based video analysis using YOLOv8 and Deep-Sort provides a viable automated framework for rodent detection, tracking, and seizure classification based on the Racine scale. The proposed system achieved an overall mAP of 77.1% and effectively identified hind limb extension and Stage-4 seizure behaviours, while also reliably distinguishing normal (non-seizure) frames. These results are consistent with prior work showing the promise of deep learning approaches for video-based seizure analysis in preclinical models^{19,20}.

The integration of Deep-Sort tracking with YOLOv8 detection addresses a key limitation of standalone object detectors, which do not track objects across frames. Deep-Sort's use of a deep learning appearance descriptor enabled persistent tracking of individual rodents through complex,

multi-rodent environments, maintaining tracking accuracy even in cases of brief occlusion^{36,37,38}.

A key challenge encountered was the limited availability of labelled examples across all Racine scale stages. Stages 1–3 exhibited only subtle motor changes that were difficult to distinguish visually, and Stage-5 events occurred only twice in the entire dataset. This data imbalance constrained the model's ability to generalise to underrepresented classes and highlights the need for larger, more balanced datasets in future studies. Augmentation techniques (rotation and flipping) were applied to partially address this imbalance for stages 4 and hind limb extension.

Additional data quality issues—including suboptimal camera angles, variable lighting, and background noise—further limited model generalisation. The heterogeneity of recording conditions across different experiments required that only the highest-quality videos be selected for analysis. Future studies should standardise recording protocols, including top-centred camera alignment, uniform lighting, and a neutral (black) background, to enhance dataset quality and model robustness.

From a computational perspective, the current pipeline required approximately 3,283 milliseconds per frame on a standard CPU, making processing of hour-long video recordings time-intensive. A multithreaded implementation that processes smaller video segments in parallel—skipping redundant frames—could substantially reduce processing time and improve practical applicability in research settings.

Broader implications of this work extend to improving consistency and reproducibility in epilepsy pharmacology research. Manual Racine scoring is inherently subjective and varies across observers; automated AI-based grading can provide standardised, quantitative assessments that reduce inter-rater variability and improve the reliability of drug efficacy evaluations. Future work should pursue larger datasets across diverse rodent epilepsy models, incorporation of additional Racine stages, and prospective validation of the system in active pharmaceutical research pipelines.

Methods

Ethical Statement

All experimental procedures involving animals were conducted by an experienced pharmaceutical researcher (D.A.S.) in compliance with applicable institutional and national guidelines for the care and use of laboratory animals. The experimental data used in this study were obtained from a published investigation on the antiepileptic activity of botanical extracts in a pentylenetetrazole (PTZ)-induced kindling mouse model⁴¹. All animal procedures in the source study were approved by the relevant institutional ethics committee. No new animal experiments were conducted as part of the present study.

Dataset and Source of Data

The dataset comprised video recordings of seizure activity in albino mice (*Mus musculus*), procured from a pharmaceutical researcher specialising in epilepsy pharmacology at the Department of Pharmacy, University of the Punjab, Lahore⁴¹. The recordings were generated in the context of evaluating the antiepileptic efficacy of a novel botanical drug formulation and captured mice during and between chemically induced seizure episodes. Video files were labelled with mouse ID, seizure induction method, and treatment details.

Experimental Subjects and Seizure Induction

Albino mice were used as experimental subjects due to their well-established genetic equivalence with human disease models and their utility in epilepsy research⁷. All subjects were housed in controlled environmental conditions to ensure experimental reliability and reproducibility. Seizures were induced using a standardised chemical protocol (pentylenetetrazole-induced kindling) that reproducibly generates pathophysiological changes analogous to human epilepsy, as described in the original pharmacological study⁴¹.

Video Recording

Mice were video-monitored from multiple angles using digital cameras placed in standardised enclosures to ensure precision and uniformity of capture. Recordings documented all seizure-associated movements and behaviours across the full duration of seizure induction experiments.

Data Pre-processing and Augmentation

Raw video data were reviewed and filtered to retain only recordings with adequate image quality—excluding clips with significant camera misalignment, poor lighting, or excessive background noise. Selected videos were segmented into individual frames, of which approximately 30% were allocated for model training and 70% for testing. Because the dataset was not originally generated for machine learning purposes, class distributions across Racine stages were uneven. To address underrepresentation of Stage-4 and hind limb extension classes, augmentation techniques including rotation (at multiple angles) and horizontal flipping were applied.

Dataset Annotation and Preparation

Video frames were imported into Roboflow³¹ for annotation. Two annotation tasks were defined: (1) rodent detection, in which bounding boxes were placed around each mouse in each frame; and (2) seizure classification, in which bounding boxes were associated with one of the following class labels: Stage-1, Stage-2, Stage-3, Stage-4, Stage-5, Hind-Limb-Extension, and Normal. A subset of frames was carefully selected for annotation to ensure high-quality training data, excluding blurred or uninformative frames. The annotated dataset was divided into training (70%), validation (20%), and test (10%) subsets for model evaluation.

Model Architecture: YOLOv8

The You Only Look Once version 8 (YOLOv8) architecture (Ultralytics, v8.0.196) was used for both rodent detection and seizure classification tasks^{27,28}. YOLOv8 is a single-stage object

detector that simultaneously predicts bounding boxes and class probabilities. Pre-trained weights (Yolov8n.pt) were used as initialisation for transfer learning on the custom dataset. The rodent detection model was trained for 50 epochs with an image size of 640×640. The seizure classification model was trained for 100 epochs with the same image size. Batch size was set to 16, with early stopping applied if mean average precision did not improve over 50–100 consecutive epochs. All experiments were conducted using Google Colaboratory Pro with a Google Compute Engine GPU backend, using Python 3.10.12 and PyTorch 2.3.0+cu121.

Multi-Object Tracking: Deep-Sort

Deep SORT (Simple Online and Real-time Tracking with a Deep Appearance Descriptor) was integrated with the YOLOv8 detection model for frame-to-frame tracking of individual rodents³⁸. Deep-Sort extends the original SORT framework by incorporating a convolutional neural network-based appearance descriptor that maintains object identity across frames, enabling robust tracking in complex and occluded environments^{36,37}. Hyperparameters used in the Deep-Sort implementation are listed in Table 2.

Parameter	Description	Value
max_age	Maximum duration of a track	5
nms_max_overlap	Maximum detection overlap	0.5
max_cosine_distance	Threshold for cosine distance	0.2

Table 2. Hyperparameters used in the Deep-Sort tracking algorithm.

Performance Metrics

Model performance was evaluated using standard object detection metrics: precision, recall, and mean Average Precision (mAP). Precision was defined as $TP / (TP + FP)$ and recall as $TP / (TP + FN)$, where TP, FP, and FN denote true positives, false positives, and false negatives respectively. Mean Average Precision was calculated as the arithmetic mean of per-class precision values across the three evaluated classes (Hind-Limb-Extension, Stage-4, Normal).

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability:

This study used as well as generated research data using You Only Look Once version 8 (YOLOv8) Ultralytics, v8.0.196 and others.

References

1. Jacobs, C. S. & Aisiku, I. P. Seizure activity. *Neurologic Emergencies: How to Do a Fast, Focused Evaluation of Any Neurologic Complaint* 59–84 (2018).
2. Leibetseder, A. et al. How to distinguish seizures from non-epileptic manifestations. *Epileptic Disord.* 22, 716–738 (2020).
3. Beghi, E. The epidemiology of epilepsy. *Neuroepidemiology* 54, 185–191 (2020).
4. Neligan, A. et al. Prognosis of adults and children following a first unprovoked seizure. *Cochrane Database Syst. Rev.* (1) (2023).
5. World Health Organization. Epilepsy. <https://www.who.int/news-room/fact-sheets/detail/epilepsy> (2024).
6. Golub, V. M. & Reddy, D. S. Post-traumatic epilepsy and comorbidities: advanced models, molecular mechanisms, biomarkers, and novel therapeutic interventions. *Pharmacol. Rev.* 74, 387–438 (2022).
7. Johan Arief, M. F. et al. A systematic review on non-mammalian models in epilepsy research. *Front. Pharmacol.* 9, 655 (2018).
8. Van Erum, J., Van Dam, D. & De Deyn, P. P. PTZ-induced seizures in mice require a revised Racine scale. *Epilepsy Behav.* 95, 51–55 (2019).
9. Sabnis, G. et al. Visual detection of seizures in mice using supervised machine learning. *bioRxiv* (2024).
10. Roy, S. et al. Evaluation of artificial intelligence systems for assisting neurologists with fast and accurate annotations of scalp electroencephalography data. *EBioMedicine* 66 (2021).
11. Yin, Y., Li, H. & Fu, W. Faster-YOLO: An accurate and faster object detection method. *Digit. Signal Process.* 102, 102756 (2020).
12. Gavrilescu, R. et al. Faster R-CNN: an approach to real-time object detection. In *Proc. Int. Conf. Expo. Electr. Power Eng. (EPE)* 165–168 (2018).
13. Diwan, T., Anirudh, G. & Tembhurne, J. V. Object detection using YOLO: Challenges, architectural successors, datasets and applications. *Multimed. Tools Appl.* 82, 9243–9275 (2023).
14. Golub, V. M. & Reddy, D. S. Post-traumatic epilepsy and comorbidities. *Pharmacol. Rev.* 74, 387–438 (2022).
15. Johan Arief, M. F. et al. A systematic review on non-mammalian models in epilepsy research. *Front. Pharmacol.* 9, 655 (2018).
16. Shimada, T. & Yamagata, K. Pentylentetrazole-induced kindling mouse model. *J. Vis. Exp.* (136) (2018).

17. Marshall, G. F., Gonzalez-Sulser, A. & Abbott, C. M. Modelling epilepsy in the mouse: challenges and solutions. *Dis. Model. Mech.* 14, dmm047449 (2021).
18. Xu, C. et al. New advances in pharmacoresistant epilepsy towards precise management. *Pharmacol. Ther.* 233, 108026 (2022).
19. Acharya, U. R., Hagiwara, Y. & Adeli, H. Automated seizure prediction. *Epilepsy Behav.* 88, 251–261 (2018).
20. Han, K., Liu, C. & Friedman, D. Artificial intelligence/machine learning for epilepsy and seizure diagnosis. *Epilepsy Behav.* 155, 109736 (2024).
21. Elger, C. E. & Hoppe, C. Diagnostic challenges in epilepsy: seizure underreporting and seizure detection. *Lancet Neurol.* 17, 279–288 (2018).
22. Tang, P. et al. Channel attention-based temporal convolutional network for satellite image time series classification. *IEEE Geosci. Remote Sens. Lett.* 19, 1–5 (2021).
23. Haleem, A., Javaid, M. & Khan, I. H. Current status and applications of artificial intelligence (AI) in medical field: An overview. *Curr. Med. Res. Pract.* 9, 231–237 (2019).
24. Dong, Y. et al. Research on how human intelligence, consciousness, and cognitive computing affect the development of artificial intelligence. *Complexity* 2020, 1680845 (2020).
25. Pinto-Coelho, L. How artificial intelligence is shaping medical imaging technology: A survey of innovations and applications. *Bioengineering* 10, 1435 (2023).
26. Khalifa, M. & Albadawy, M. AI in diagnostic imaging: Revolutionising accuracy and efficiency. *Comput. Methods Programs Biomed. Update* 100146 (2024).
27. Hussain, M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. *Machines* 11, 677 (2023).
28. Sirisha, U. et al. Statistical analysis of design aspects of various YOLO-based deep learning models for object detection. *Int. J. Comput. Intell. Syst.* 16, 126 (2023).
29. Maharana, K., Mondal, S. & Nemade, B. A review: Data pre-processing and data augmentation techniques. *Glob. Transit. Proc.* 3, 91–99 (2022).
30. Khosla, C. & Saini, B. S. Enhancing performance of deep learning models with different data augmentation techniques: A survey. In *Proc. Int. Conf. Intell. Eng. Manag. (ICIEM)* 79–85 (2020).
31. Alin, A. Y. et al. Data augmentation method on drone object detection with YOLOv5 algorithm. In *Proc. 8th Int. Conf. Informatics Comput. (ICIC)* 1–6 (2023).
32. Shandilya, S. K. et al. YOLO-based segmented dataset for drone vs. bird detection for deep and machine learning algorithms. *Data Brief* 50, 109355 (2023).
33. Lit, Z. et al. Multiple object tracking with GRU association and Kalman prediction. In *Proc. Int. Jt. Conf. Neural Netw. (IJCNN)* 1–8 (2021).
34. Rakai, L. et al. Data association in multiple object tracking: A survey of recent techniques. *Expert Syst. Appl.* 192, 116300 (2022).
35. Kuzdeba, S. et al. Kalman Filter Changepoint Detection and Trend Characterization. In *Proc. IEEE 29th Int. Workshop Mach. Learn. Signal Process. (MLSP)* 1–6 (2019).
36. Ahmedt-Aristizabal, D. et al. Deep learning approaches for seizure video analysis: A review. *Epilepsy Behav.* 154, 109735 (2024).

37. Sun, H., Li, C. & Zhang, H. Image Segmentation-P300 Selector: A Brain-Computer Interface System for Target Selection. *Comput. Mater. Continua* 79 (2024).
38. Abouelyazid, M. Comparative evaluation of SORT, DeepSORT, and ByteTrack for multiple object tracking in highway videos. *Int. J. Sustain. Infrastruct. Cities Soc.* 8, 42–52 (2023).
39. Inthiyaz, S. et al. YOLO (You Only Look Once) making object detection work in medical imaging on convolution detection system. *Int. J. Pharm. Res.* 12 (2020).
40. Wu, X., Sahoo, D. & Hoi, S. C. Recent advances in deep learning for object detection. *Neurocomputing* 396, 39–64 (2020).
41. Saleem, A. et al. Neuropharmacological-based in vivo and behavioral exploration of antiepileptic activity of leaves extract of *Terminalia bellirica* Roxb. in chronic pentylenetetrazole-induced kindling model in mice. *Chem. Pap.* 1–10 (2024).

Acknowledgements

The authors thank the Department of Data Science, University of the Punjab, for providing computational resources, and Dr. Aneeqa Saleem (Minhaj University) for providing the video dataset used in this research.

Author Contributions

Summan Bahadur designed the AI pipeline, performed model training and analysis, and drafted the manuscript.

Faisal Bukhari supervised the study, conceptualized the research framework, and critically revised the manuscript.

Zubair Nawaz contributed to methodology development and data analysis.

Aneeqa Saleem provided the experimental dataset and domain expertise in epilepsy pharmacology. All authors reviewed and approved the final manuscript.

Data Availability

The video dataset used in this study was obtained from an external collaborator and is not publicly available due to institutional restrictions. The model weights and code used for training and evaluation will be made available upon reasonable request to the corresponding author.

Competing Interests

The authors declare no competing interests.