Self-Supervised Learning for Improved Classification and Characterization of Sharp Wave Ripples

Saber Graf

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE : SCIENCES DE LA VIE ET DE LA SANTE

SPÉCIALITÉ : NEUROSCIENCES

Par **Saber GRAF**

Apprentissage auto-supervisé pour l'amélioration de la classification et de la caractérisation des ondes rapides hippocampiques

Sous la direction de : **Cyril Herry**

Soutenue le 13/12/2024

Membres du jury :

M. Georges François	Université de Bordeaux	Président du jury
Mme. Chien Sarina Hui-Lin	China Medical University	Rapporteur
M. Humeau Yann	Friedrich Miescher Institute for Biomedical Research	Rapporteur
M. Cheng Da-Chuan	China Medical University	Examinateur

Membres invités :

M. Tsai Feng-Sheng	China Medical University	Invité
M. Meyrand Pierre	Université de Bordeaux	Invité

Titre : Apprentissage auto-supervisé pour l'amélioration de la classification et de la caractérisation des ondes rapides hippocampiques

Résumé :

Les ondes rapides hippocampiques (ORH) sont des événements oscillatoires essentiels associés à la consolidation de la mémoire et au traitement cognitif. Cette recherche explore l'application de techniques d'apprentissage auto-supervisé (AAS) pour améliorer la classification des ORH, en se concentrant sur la distinction entre les événements avant et après l'apprentissage. Les méthodes traditionnelles échouent souvent à capturer les subtiles nuances des propriétés des ORH qui émergent à la suite de l'apprentissage. En intégrant l'AAS avec une architecture personnalisée de réseau de neurones convolutifs 1D (RNC 1D), cette étude introduit une approche novatrice pour relever ces défis, améliorant ainsi de manière significative la précision de la classification des ORH.

Le RNC 1D a démontré une performance solide dans la classification des ORH, avec une re-labellisation des données basée sur l'AAS qui a renforcé la robustesse du modèle en atténuant les effets du bruit des étiquettes. Cette approche exploite des caractéristiques critiques telles que les coefficients de la Transformée de Fourier Rapide (TFR) et l'entropie des ondelettes, identifiées comme des indicateurs clés des changements induits par l'apprentissage dans les propriétés spectrales et temporelles des ORH. Les résultats ont révélé des changements significatifs dans le contenu fréquentiel et la distribution de la puissance des ORH après l'apprentissage, soulignant la nature dynamique de l'activité hippocampique en réponse aux exigences cognitives. De plus, l'application de la méthode de visualisation par activation de classes pondérées par gradient (Grad-CAM) combinée à l'analyse des ondelettes de Morlet a validé la pertinence biologique des décisions de classification du modèle, apportant une meilleure compréhension des mécanismes neuronaux sous-jacents.

L'extension de l'analyse à un modèle murin transgénique de la maladie d'Alzheimer a révélé des différences distinctes dans les caractéristiques des ORH par rapport aux animaux de type sauvage, soulignant le potentiel de la classification des ORH en tant que biomarqueur des déficits cognitifs et des conditions neurodégénératives.

Cette étude démontre l'efficacité de l'apprentissage profond pour l'analyse des signaux neuronaux complexes, offrant une compréhension plus nuancée des changements induits par l'apprentissage dans les ORH.

Mots clés : Ondes rapides hippocampiques, Apprentissage auto-supervisé, Apprentissage profond, Hippocampe, Apprentissage spatial, Oscillations neuronales

Title: Self-Supervised Learning for Improved Classification and Characterization of Sharp Wave Ripples

Abstract :

Sharp wave ripples (SWRs) in the hippocampus are essential oscillatory events associated with memory consolidation and cognitive processing. This research explores the application of self-supervised learning (SSL) techniques to enhance the classification of SWRs, focusing on distinguishing between before and after learning events. Traditional methods often fail to capture the subtle nuances in SWR properties that emerge as a result of learning. By integrating SSL with a custom 1D convolutional neural network (CNN) architecture, this study introduces a novel approach to address these challenges, significantly improving the classification accuracy of SWRs.

The 1D CNN demonstrated strong performance in classifying SWRs, with SSL-based data re-labeling further enhancing the model's robustness by mitigating the effects of label noise. This approach leverages critical features such as Fast Fourier Transform (FFT) coefficients and wavelet entropy, which were identified as key indicators of learning-induced changes in the spectral and temporal properties of SWRs. The findings revealed significant shifts in frequency content and power distribution of SWRs following learning, highlighting the dynamic nature of hippocampal activity in response to cognitive demands. Additionally, the application of Gradient-weighted Class Activation Mapping (Grad-CAM) combined with Morlet wavelet analysis validated the biological relevance of the model's classification decisions, providing insights into the underlying neural mechanisms.

Extending the analysis to a transgenic mouse model of Alzheimer's disease uncovered distinct differences in SWR characteristics compared to wild-type animals, underscoring the potential of SWR classification as a biomarker for cognitive impairments and neurodegenerative conditions.

This study demonstrates the efficacy of deep learning for the analysis of complex neural signals, offering a more nuanced understanding of learning-induced changes in SWRs.

Keywords: Sharp Wave Ripples, Self-Supervised Learning, Deep Learning, Hippocampus, Spatial Learning, Neural Oscillations

Remerciements

Je tiens à exprimer ma profonde gratitude à tous ceux qui ont contribué à la réalisation de cette thèse et m'ont soutenu tout au long de ce parcours.

Je souhaite adresser mes sincères remerciements à Neuroservice et à ses directeurs, Bruno Buisson et Olivier Toury, pour leur soutien exceptionnel tout au long de ce parcours. La flexibilité accordée par l'entreprise, me permettant de travailler à temps partiel, a été déterminante dans la réalisation de cette thèse. Je suis profondément reconnaissant pour leur compréhension, leur confiance et leur engagement envers mon développement professionnel et académique.

Je remercie chaleureusement mon directeur de thèse, Cyril Herry, pour sa guidance exceptionnelle, son expertise et son soutien indéfectible. Son leadership, les précieuses discussions que nous avons eues, ses conseils avisés et ses révisions rigoureuses ont considérablement enrichi ce projet de recherche. Sa vision scientifique et son encouragement constant ont joué un rôle crucial dans l'aboutissement de cette thèse.

Je souhaite également exprimer ma sincère reconnaissance à mon co-directeur de thèse, Feng-Sheng Tsai, pour son expertise en apprentissage machine. Ses conseils techniques et son soutien dans l'application des techniques d'apprentissage profond ont été déterminants pour la réussite de mes analyses.

Un grand merci à Pierre Meyrand pour son encadrement rigoureux et ses conseils avisés qui ont été essentiels à la réussite de ce travail. Sa contribution a été inestimable tout au long de ce parcours.

Je tiens à remercier Tiaza Bem pour ses connaissances approfondies en neurosciences, qui ont été d'une aide précieuse. Sa collaboration et ses contributions ont significativement enrichi mes recherches et m'ont permis de mieux comprendre les subtilités de ce domaine complexe.

Enfin, je souhaite exprimer toute ma gratitude à ma femme et à mes enfants, dont l'amour, la patience et le soutien inconditionnel m'ont permis de traverser les moments difficiles de ce long parcours. Leur présence à mes côtés a été une source inestimable de force et de motivation.

Je remercie également ma famille pour leur soutien constant, leur encouragement et leur foi en moi, qui m'ont aidé à persévérer dans cette entreprise.

À vous tous, je vous adresse mes plus sincères remerciements.

Table of contents

Abbreviations List

AD	Alzheimer's Disease
AI	Artificial Intelligence
AL	After Learning
ANOVA	Analysis of Variance
BL	Before Learning
CA1	Cornu Ammonis 1 (region of the hippocampus)
CNN	Convolutional Neural Network
EEG	Electroencephalogram
FFT	Fast Fourier Transform
G1	Group 1 (re-labeled group from SSL)
G2	Group 2 (re-labeled group from SSL)
Grad-CAM	Gradient-weighted Class Activation Mapping
KDE	Kernel Density Estimation
LFP	Local Field Potential
LTP	Long-Term Potentiation
PCA	Principal Component Analysis
RCI	Ripple-Centered Interval
ReLU	Rectified Linear Unit
SEM	Standard Error of the Mean
SSL	Self-Supervised Learning
SWR	Sharp Wave Ripple
SWR_{art}	Synthetic Sharp Wave Ripple
TG	Transgenic
WT	Wild-Type

List of figures

I. Introduction

The field of neuroscience has undergone a significant transformation over the past decades, driven by technological advancements and an ever-growing understanding of the brain's complex functions. This evolution has been particularly evident in the domain of data analysis, where the progression from rudimentary recording techniques to sophisticated computational methods has improved our approach to understanding neural processes.

In recent years, the field has witnessed an increase in the application of artificial intelligence (AI) and machine learning techniques to data analysis. Deep learning algorithms, in particular, have shown considerable success in decoding neural signals, classifying brain states, and uncovering hidden patterns in neurophysiological data (Glaser et al. 2020). These advanced computational methods offer the potential to enhance our understanding of brain function and dysfunction, opening new avenues for research and clinical applications.

Among the myriad neural phenomena studied using these advanced techniques, local field potentials (LFPs) have emerged as a crucial source of information about neural activity. LFPs represent the aggregate electrical activity of neural populations and offer insights into various oscillatory patterns such as theta (4-8 Hz), gamma (30-100 Hz), and sharp wave ripples (SWRs). Of particular interest are SWRs, first described by Buzsáki et al. (G. Buzsáki, Leung, and Vanderwolf 1983). These brief, high-frequency oscillations (140-250 Hz) occur in the hippocampus during periods of quiet wakefulness and slow-wave sleep. SWRs have been implicated in memory consolidation, spatial navigation, and decision-making processes, making them a critical target for advanced analytical methods (Joo and Frank 2018).

This thesis aims to advance the field of SWR analysis by developing and implementing novel deep learning techniques, with a particular focus on self-supervised learning approaches. By leveraging the power of artificial intelligence and the wealth of available neural data, we seek to enhance our ability to classify, and characterize SWRs, contributing to a deeper understanding of their role in cognitive processes and their potential as biomarkers for neurological disorders.

To fully appreciate the significance of modern deep learning approaches in SWR analysis, it is essential to consider both the historical context of neurophysiological recording techniques and the evolution of data analysis methods. In the following sections, we will explore the progression of these techniques and methodologies.

A. Historical Perspective on Neurophysiological Recording Techniques

1.Early Challenges in Brain Activity Recording

The study of neural activity has been a cornerstone of neuroscience since its inception, with the quest to understand brain function driving continuous innovation in recording techniques. In the early 20th century, neuroscientists faced numerous challenges in capturing brain activity. The intricate nature of neural signals, characterized by their small amplitude and rapid temporal dynamics, posed significant technical hurdles. Early attempts to record neural activity were limited by the sensitivity and temporal resolution of available instruments.

The development of the electroencephalogram (EEG) by Hans Berger in the 1920s marked a significant breakthrough, allowing for the first non-invasive recordings of human brain activity (Swartz and Goldensohn 1998). However, EEG's spatial resolution was limited, providing only a broad view of cortical activity. This limitation spurred researchers to develop more precise recording techniques.

The advent of intracellular recording techniques in the 1940s, pioneered by Hodgkin and Huxley in their seminal work on action potentials, transformed the field by enabling the study of individual neurons (Hodgkin and Huxley 1952). This approach, using glass micropipettes to penetrate cell membranes, allowed for precise measurements of membrane potentials and ionic currents. The subsequent development of patch-clamp techniques by Neher and Sakmann in the 1970s further refined single-cell recordings, providing unprecedented insights into ion channel function and synaptic transmission (Neher and Sakmann 1976).

These early advancements in neurophysiological recording techniques laid the foundation for our understanding of neural function at the cellular level. However, as researchers began to appreciate the complexity of neural circuits and the importance of network-level interactions, it became clear that single-cell recordings, while invaluable, were insufficient to capture the full picture of brain function. This realization led to the development of techniques capable of recording from multiple neurons simultaneously, ushering in a new era of neurophysiological research.

2.From Single-Cell to Multi-Electrode Recordings

The progression from single electrodes to multi-electrode arrays (MEAs) in the late 20th century marked another pivotal advancement. MEAs enabled simultaneous recordings from multiple sites, providing a more comprehensive view of neural population dynamics. Early MEAs, developed in the 1970s, typically consisted of dozens of electrodes. Modern arrays can incorporate hundreds or even thousands of recording sites, dramatically increasing the spatial and temporal resolution of neural recordings (Jun et al. 2017).

The development of silicon-based probes in the 1990s and 2000s further expanded the capabilities of multi-electrode recordings. These probes offered higher density electrode arrangements and improved biocompatibility. The CMOS-based probes, introduced in the late 2010s, represent a significant leap forward, allowing for simultaneous recordings from hundreds of channels across multiple brain regions (Steinmetz et al. 2021).

Complementary to these invasive techniques, advances in non-invasive imaging methods have expanded our ability to study human brain function at various scales. Functional magnetic resonance imaging (fMRI), developed in the early 1990s, allowed researchers to observe brain activity indirectly through changes in blood oxygenation. Magnetoencephalography (MEG), which measures magnetic fields produced by electrical currents in the brain, offers high temporal resolution complementary to fMRI's spatial precision (Bandettini 2020).

Recent developments in optical imaging techniques, such as two-photon microscopy and light-sheet microscopy, have enabled researchers to visualize neural activity at cellular and subcellular resolution in living tissue. These methods, combined with genetically encoded calcium indicators, provide unprecedented insights into neural circuit function (Yang and Yuste 2017).

These advancements in multi-electrode recordings and imaging techniques have dramatically expanded our ability to observe and analyze neural activity across multiple spatial and temporal scales. From single neurons to large-scale brain networks, researchers now have access to an unprecedented wealth of data on brain function. However, this explosion of data has brought with it new challenges in data management, analysis, and interpretation. As we enter the era of big data in neuroscience, novel computational approaches are becoming increasingly crucial to extract meaningful insights from the vast amounts of neural data being generated.

3. The Big Data Era in Neuroscience

The advent of advanced recording technologies in neuroscience has ushered in an era of big data, transforming the landscape of brain research. This transition from single-cell recordings to large-scale, multi-electrode arrays has generated an unprecedented volume and complexity of neural data, presenting both opportunities and challenges for the field.

Modern recording techniques can generate terabytes of data in a single experiment. For instance, a typical Neuropixels probe recording session can produce over 1 TB of raw data per day. This deluge of information has overwhelmed traditional analysis methods, which were designed for smaller, more manageable datasets.

The challenges posed by big data in neuroscience are multifaceted. First, there is the sheer volume of data to contend with, which requires significant computational resources for storage, processing, and analysis. Cloud computing and distributed processing systems have become increasingly important in managing these massive datasets.

Second, the high dimensionality of neural data, encompassing spatial, temporal, and functional aspects of brain activity, necessitates sophisticated analytical techniques to uncover underlying patterns and relationships. Dimensionality reduction methods, such as principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE), have become crucial tools in visualizing and analyzing these complex datasets (Cunningham and Yu 2014).

Third, the heterogeneity of data types, ranging from electrophysiological recordings to imaging data and behavioral metrics, demands integrative approaches that can synthesize information across multiple modalities. This has led to the development of multimodal analysis techniques that can combine data from different recording methods to provide a more comprehensive view of brain function (Dipietro et al. 2023).

The big data era in neuroscience has fundamentally changed how we approach brain research, necessitating new tools and methodologies to extract meaningful insights from vast and complex datasets. This shift has not only pushed the boundaries of our understanding of brain function but has also highlighted the critical need for advanced computational methods in neuroscience. As we continue to generate increasingly large and complex neural datasets, the development of sophisticated data analysis techniques becomes paramount.

This evolution in data generation has been paralleled by an equally important evolution in data analysis methods. To fully appreciate the current state of neural data analysis and the potential of advanced techniques like deep learning, it is essential to understand the historical context of these analytical approaches.

B. Evolution of Data Analysis Methods

1. Traditional Approaches to Neural Data Analysis

Traditional approaches to neural data analysis were primarily rooted in statistical techniques and signal processing methods developed in the mid-20th century. These methods were designed to extract meaningful information from small and simple datasets, often focusing on single-unit recordings or local field potentials from a limited number of brain regions.

One of the foundational techniques in neural data analysis was spike sorting, which aimed to identify and classify action potentials from extracellular recordings. Early spike sorting methods relied on manual inspection of waveforms and simple clustering algorithms. As recording technologies advanced, more sophisticated automated spike sorting algorithms were developed (Buccino, Garcia, and Yger 2022).

Another key approach in traditional neural data analysis was spectral analysis, particularly the use of Fourier transforms to decompose neural signals into their frequency components. This technique proved invaluable in identifying oscillatory patterns in brain activity, such as theta and gamma rhythms, which have been implicated in various cognitive processes (Robinson et al. 2021).

As the complexity of neural recordings increased, researchers began to employ multivariate analysis techniques. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) became popular tools for dimensionality reduction and feature extraction in neural datasets (Greenacre et al. 2022). These methods allowed for the identification of underlying patterns in high-dimensional data, facilitating the interpretation of complex neural dynamics.

While these traditional approaches have been instrumental in advancing our understanding of brain function, they have increasingly shown limitations in the face of modern neuroscience data. The exponential growth in data volume, complexity, and dimensionality has pushed these methods to their limits, revealing the need for more advanced computational

techniques. As neuroscience enters the era of big data, new challenges have emerged that require novel analytical approaches.

2.Computational Challenges in Modern Neuroscience

The advent of high-density multi-electrode arrays, advanced imaging techniques, and large-scale recording technologies has brought about unprecedented challenges in neural data analysis. Modern neuroscience experiments can generate a large amount of data in a single session, far surpassing the capabilities of traditional analysis methods.

One of the primary challenges in modern neuroscience is the high dimensionality of neural data. Recordings from thousands of neurons across multiple brain regions result in datasets with numerous spatial and temporal dimensions. This " curse of dimensionality" complicates statistical analysis and increases the risk of spurious correlations (Cunningham and Yu 2014).

Furthermore, the non-stationary nature of neural activity presents challenges for traditional statistical approaches. Brain states can change rapidly in response to external stimuli or internal processes, necessitating analytical methods that can account for dynamic changes in neural representations (Linderman and Gershman 2017).

These challenges have pushed the boundaries of conventional data analysis techniques, revealing their limitations in handling the complexity and scale of modern neuroscience data. The need for more sophisticated computational approaches that can effectively process, analyze, and interpret these vast and intricate datasets has become increasingly apparent. As the field grapples with these computational hurdles, it has turned to cutting-edge technologies and methodologies from other domains, particularly artificial intelligence and machine learning.

3.Emergence of Artificial Intelligence in Neural Data Analysis

The confluence of big data in neuroscience and advancements in artificial intelligence (AI) has ushered in a new era of neural data analysis. Machine learning, and particularly deep learning, has emerged as a powerful tool for tackling the complexities of modern neuroscience datasets. In response to these challenges, the field has witnessed a rapid evolution of data analysis methods. Traditional manual approaches, which relied heavily on visual inspection and simple

statistical tests, have given way to automated algorithms and machine learning techniques. This shift has been driven by the need for more efficient, objective, and scalable methods of data analysis.

Convolutional Neural Networks (CNNs) have found wide application in analyzing spatial patterns in neural data, particularly in neuroimaging studies. These models have been used to decode visual stimuli from fMRI data, classify brain states, and identify structural abnormalities in clinical imaging (Celeghin et al. 2023).

Recurrent Neural Networks (RNNs), especially variants like Long Short-Term Memory (LSTM) networks, have proven effective in modeling temporal dynamics in neural data. These models have been applied to tasks such as predicting neural responses, decoding movement intentions from motor cortex activity, and analyzing sequential patterns in behavior (Glaser et al. 2020).

Recent advancements in artificial intelligence have further expanded the toolkit available to neuroscientists. Attention-based models, such as transformers, have shown promise in analyzing temporal sequences in neural data (Vaswani et al. 2017). These models have been applied to EEG analysis, neural decoding, and modeling of complex cognitive processes (Kostas et al. 2021). Graph neural networks have been applied to study brain connectivity and network dynamics (Bessadok, Mahjoub, and Rekik 2023). These approaches offer the advantage of automatically learning relevant features from data, potentially uncovering patterns that might be missed by traditional analysis methods.

Unsupervised learning techniques, such as autoencoders and generative adversarial networks (GANs), have been employed to discover latent structure in neural data. These methods can reveal low-dimensional representations of high-dimensional neural activity, potentially uncovering fundamental principles of neural computation (Pandarinath et al. 2018).

The rise of big data in neuroscience has also spurred the development of new statistical frameworks for hypothesis testing and inference. Methods like permutation tests and bootstrap resampling have become increasingly important for assessing the significance of findings in high-dimensional datasets where traditional parametric tests may not be appropriate (Lotte et al. 2018).

Despite the power of AI approaches, challenges remain in their application to neuroscience. Interpretability is a key concern, as the complex, non-linear transformations in deep learning models can be difficult to relate to underlying biological mechanisms. Recent work has focused

on developing more interpretable AI models and on methods for extracting biological insights from trained networks (Samek et al. 2021).

Another challenge is the need for large amounts of labeled data for supervised learning approaches. This has spurred interest in transfer learning and self-supervised learning techniques. Transfer learning aims to leverage knowledge gained from one dataset or task to improve performance on related problems with limited data. Self-supervised learning, on the other hand, focuses on training models using pretext tasks that do not require manual annotation, allowing them to learn valuable representations from large amounts of unlabeled data. These approaches offer promising solutions to the data scarcity problem in neuroscience (Zhao et al. 2024).

As we continue to push the boundaries of neural recording technologies, the challenges and opportunities presented by big data in neuroscience will only grow. The next frontier lies in developing integrative approaches that can synthesize information across multiple scales and modalities, from single-neuron activity to whole-brain dynamics. This comprehensive approach promises to provide a more complete understanding of brain function and may lead to new insights into cognition, behavior, and neurological disorders.

C. Research questions and objectives

Our study aims to leverage the power of deep learning and self-supervised learning techniques to enhance the analysis of sharp wave ripples (SWRs), with the overarching goal of advancing our understanding of these critical neural events and their role in memory processes.

The primary objective of this research is to develop and validate a novel, deep learning-based approach for the analysis of SWRs that surpasses the performance of traditional methods. Specifically, we aim to create a robust, automated system for SWR classification that can handle the variability and complexity inherent in neural recordings. This methodological advancement is crucial for improving the accuracy and efficiency of SWR analysis, particularly in the context of large-scale, long-duration recordings that are increasingly common in modern neuroscience research.

A key focus of our study is the implementation of self-supervised learning techniques in the context of SWR analysis. Self-supervised learning (SSL), a paradigm that leverages unlabeled data to create supervisory signals, offers a promising solution to the challenge of limited

labeled datasets in neuroscience (Banville et al. 2021). In our research, we specifically employ SSL to address the issue of label noise, which is a common problem in neurophysiological data. By applying this approach to SWR classification, we aim to develop models that can learn meaningful representations from large amounts of unlabeled neural data, potentially uncovering subtle patterns and features that may be overlooked by traditional analysis methods while simultaneously improving the quality of our labeled dataset.

Our approach to mitigating label noise through SSL involves a re-labeling process. We utilize SSL to generate new labels for our dataset, effectively redistributing the SWRs based on their learned features rather than their original temporal classification. This method aims to uncover potentially more nuanced groupings within the data that may not have been apparent in the original categorization, thereby reducing the impact of potential mislabeling in the initial dataset.

Furthermore, we seek to explore the interpretability of our deep learning models, addressing the common criticism of AI approaches as "black boxes". By employing techniques such as gradient-weighted class activation mapping (Grad-CAM), we aim to provide insights into the features and patterns that our models deem most relevant for SWR classification (Selvaraju et al. 2020). This interpretability is crucial not only for validating the biological relevance of our approach but also for potentially uncovering new insights into the characteristics and dynamics of SWRs.

In addition to these methodological objectives, our study aims to contribute to the broader understanding of SWRs and their functional significance. By developing more sensitive and precise analytical tools, we hope to enable more nuanced investigations into the relationship between SWR characteristics and various cognitive processes, particularly memory consolidation and retrieval. This could potentially lead to new insights into the neural mechanisms underlying memory formation and the role of SWRs in coordinating information transfer between the hippocampus and neocortex.

The study of SWRs inherently demands an understanding of their anatomical origins and the neural circuits that generate them. These complex oscillatory events arise from the intricate interplay of neuronal populations within specific hippocampal subregions. Therefore, to provide a comprehensive context for our research, the following section will offer an anatomical and functional overview of the hippocampus.

D. Anatomical and functional overview of the hippocampus

The hippocampus, a distinctive seahorse-shaped structure nestled within the medial temporal lobe of the mammalian brain, has been a subject of intense scientific scrutiny for decades. Its pivotal role in learning, memory formation, and spatial navigation has made it a cornerstone of neuroscientific research.

1. Structural organization of the hippocampus

The hippocampal formation is a complex structure comprising several interconnected regions, each with distinct cellular compositions and connectivity patterns. The primary components include the dentate gyrus (DG), the Ammon's Horn or *cornu Ammonis* (consisting of the CA1, CA2, and CA3 subfields), and the subiculum. These regions are organized into a largely unidirectional circuit known as the trisynaptic pathway, which forms the basis of information processing within the hippocampus (Hainmueller and Bartos 2020) (Figure 1).

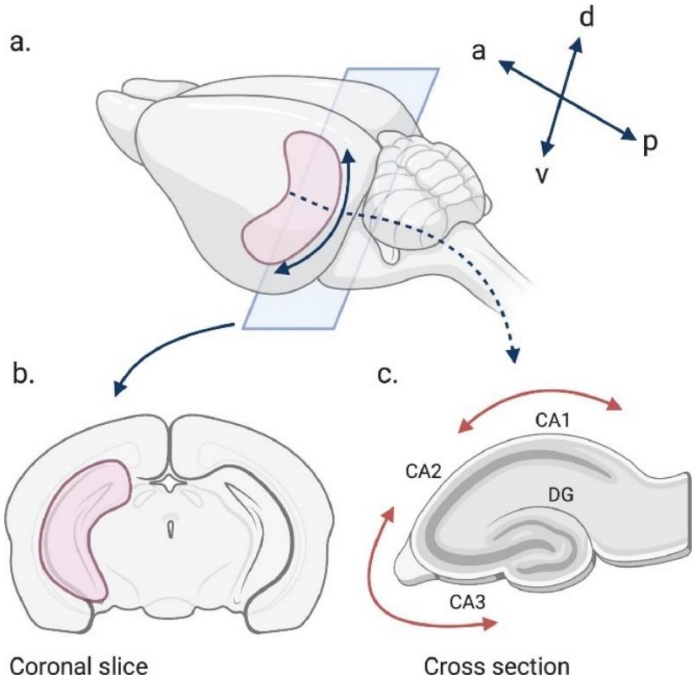

Figure 1 Anatomical representation of the rodent hippocampus.

(a) Lateral view of the rodent brain highlighting the hippocampus (pink). (b) Coronal slice showing the hippocampus. (c) Cross-sectional view of the hippocampus detailing major subfields: CA1, CA2, CA3, and dentate gyrus (DG).Adapted from (Gunnarsdóttir, Zerbi, and Kelly 2022), distributed under Creative Commons Attribution license.

The dentate gyrus, often considered the primary input region of the hippocampus, is characterized by its densely packed granule cell layer. These granule cells receive major afferent projections from layer II of the entorhinal cortex via the perforant path (Figure 2). The axons of the granule cells, known as mossy fibers, project to the CA3 region, forming distinctive synaptic connections called mossy fiber synapses.

The CA3 region is distinguished by its extensive recurrent collateral system, where pyramidal neurons form synaptic connections with other CA3 neurons (Figure 2). This auto-associative network architecture supports two complementary processes: pattern completion and pattern separation. Pattern completion allows incomplete or degraded sensory inputs to be used to retrieve entire memories (Nakazawa et al. 2002). Conversely, pattern separation, which is thought to primarily occur in the dentate gyrus but also involve CA3, enables the formation of distinct memory representations for similar experiences (Yassa and Stark 2011). The interplay between these processes in the hippocampal circuit is crucial for both the accurate recall of memories and the ability to distinguish between similar events. CA3 neurons project to the CA1 region via the Schaffer collaterals, forming the final step of the trisynaptic circuit.

The CA1 region, often viewed as the primary output region of the hippocampus, receives inputs from both CA3 and directly from layer III of the entorhinal cortex. The precise laminar organization of CA1, with distinct strata for different inputs, allows for complex integration of information from multiple sources (Spruston 2008). The stratum radiatum and stratum lacunosum-moleculare receive inputs from CA3 and the entorhinal cortex, respectively, while the cell bodies of CA1 pyramidal neurons are organized in the stratum pyramidale. This layered structure facilitates the comparison of current sensory input (via the direct pathway from the entorhinal cortex) with stored associations (via the CA3 pathway), potentially supporting match/mismatch detection and novelty recognition (Kumaran and Maguire 2007).

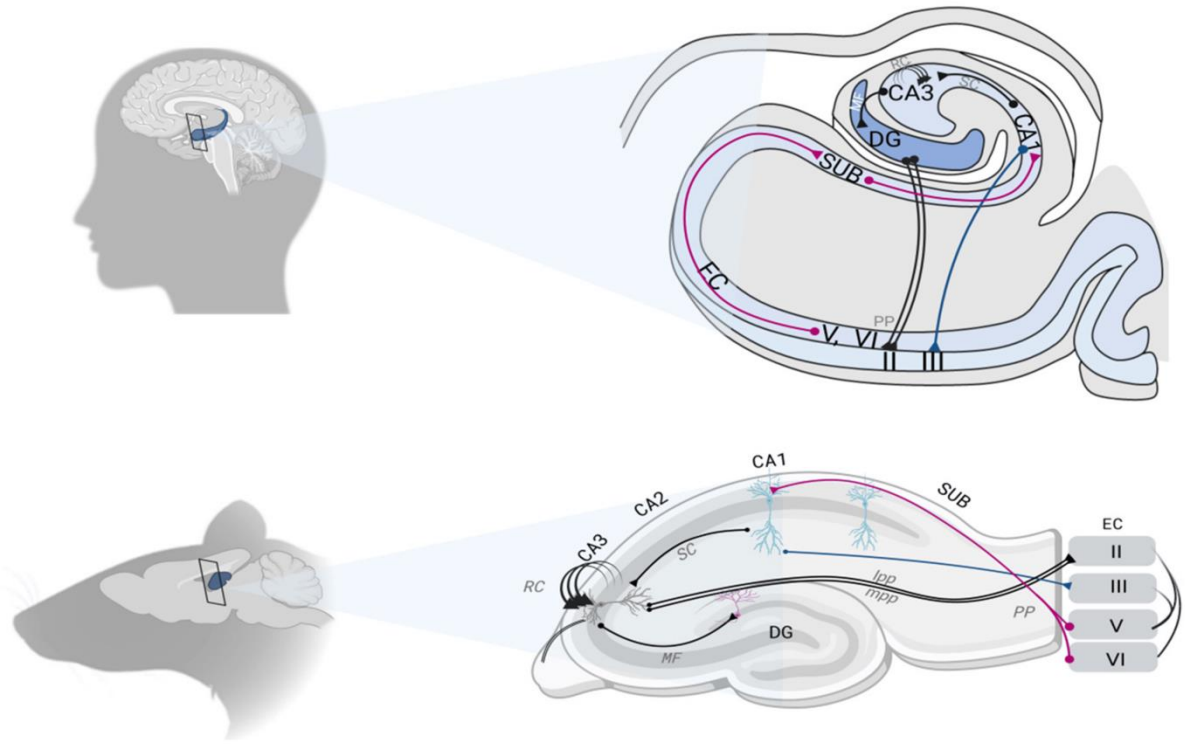

Figure 2 Anatomical structure and connectivity of the hippocampus.

(Top) The human brain highlighting the location of the hippocampus and its internal circuitry, including major pathways such as the perforant path (PP) and the connections between the dentate gyrus (DG), CA3, and CA1. (Bottom) The rodent brain illustrating similar hippocampal pathways and connections, emphasizing the comparative anatomy and function across species. Adapted from (Roux, Leger, and Freret 2021), distributed under Creative Commons Attribution license.

The CA2 region, although smaller and less studied than its counterparts, has gained attention for its unique properties and functional roles. CA2 pyramidal neurons receive inputs from both the entorhinal cortex and the supramammillary nucleus and project to CA1 and the deep layers of the entorhinal cortex. Recent research has implicated CA2 in social memory and the encoding of contextual information, highlighting its importance in hippocampal function despite its small size.

The subiculum, positioned between the hippocampus proper and the entorhinal cortex, serves as the main output structure of the hippocampal formation. It receives input primarily from CA1 and projects to various cortical and subcortical regions, including the prefrontal cortex, nucleus accumbens, and amygdala (Amaral and Lavenex 2007). The subiculum plays a crucial role in the distribution of processed information and has been implicated in spatial navigation and memory retrieval (O'Mara et al. 2001).

While the structural organization provides the physical substrate for hippocampal function, it is the dynamic interactions within and between these structures that give rise to the hippocampus's pivotal role in cognition. To fully appreciate the significance of Sharp Wave Ripples and their analysis, we must consider how this anatomical framework supports the diverse functional roles of the hippocampus.

2. Functional roles of the hippocampus

The hippocampus's involvement in memory formation and consolidation was first brought to light by the seminal case of patient H.M., who suffered severe anterograde amnesia following bilateral medial temporal lobe resection (Scoville and Milner 1957). This case underscored the critical role of the hippocampus in declarative memory formation, particularly in the initial encoding and consolidation of new experiences into long-term memory.

Subsequent research has elucidated the hippocampus's involvement in various aspects of memory processing. It plays a crucial role in the formation of episodic memories, which are autobiographical memories of specific events tied to particular times and places. The hippocampus is thought to bind together disparate elements of an experience – including sensory information, emotional content, and spatial context – into a coherent memory trace (Eichenbaum 2004).

Beyond its role in episodic memory, the hippocampus is integral to spatial navigation and the formation of cognitive maps. The discovery of place cells in the rat hippocampus by O'Keefe and Dostrovsky (O'Keefe and Dostrovsky 1971) reshaped our understanding of how the brain represents space. These neurons, which fire when an animal occupies specific locations in its environment, form the basis of a cognitive map that allows for flexible navigation and spatial memory. This work was later complemented by the discovery of grid cells in the entorhinal cortex (Hafting et al. 2005), suggesting a complex interplay between these structures in spatial representation and navigation.

The hippocampus has also been shown to be involved in imagining future scenarios, suggesting a role in prospective thinking and planning. This finding has led to the development of the constructive episodic simulation hypothesis, which posits that the same neural mechanisms involved in remembering past events are also engaged in imagining future possibilities (Schacter, Addis, and Buckner 2007). This perspective highlights the dynamic and constructive

nature of memory, moving beyond the notion of memories as static representations of past events.

The multifaceted functional roles of the hippocampus underscore its significance in cognition and behavior. From forming episodic memories to enabling spatial navigation and supporting flexible thinking, the hippocampus serves as a crucial hub for integrating and processing diverse types of information. This functional versatility is made possible by the complex anatomical organization discussed earlier, highlighting the intricate relationship between structure and function in the brain.

However, the hippocampus does not operate in isolation. Its diverse functions are supported and modulated by interactions with numerous other brain regions. Understanding these interactions is crucial for comprehending how the hippocampus contributes to larger-scale cognitive processes and how information is transferred between brain areas during memory formation and retrieval.

3. Hippocampal interactions with other brain regions

The hippocampus does not function in isolation but is part of a broader network involving various cortical and subcortical structures. Its interactions with the prefrontal cortex are particularly important for the integration of spatial and contextual information in goal-directed behavior. This hippocampal-prefrontal circuit is crucial for processes that require the flexible use of memory, such as in tasks involving delayed responses or rule learning (Preston and Eichenbaum 2013). These interactions support cognitive flexibility and the application of past experiences to guide future behavior. The role of this circuit in decision-making has been further elaborated by Yu and Frank (J. Y. Yu and Frank 2015), who discuss how hippocampal-prefrontal interactions support decision-making in spatial navigation and memory-guided choice.

The hippocampus's connections with the amygdala underlie its role in emotional memory and the contextual modulation of fear responses. This interaction is critical for the formation and expression of fear memories, as well as for their extinction (Maren and Fanselow 1995). The hippocampus-amygdala circuit allows for the association of emotional valence with specific contexts or stimuli, supporting adaptive behavior in potentially threatening situations (**Figure 3**).

Additionally, the hippocampus maintains reciprocal connections with various neocortical regions, including sensory and association areas. These connections allow for the integration of diverse sensory inputs into coherent memory representations and facilitate the reinstatement of these representations during memory retrieval (Ranganath and Ritchey 2012).

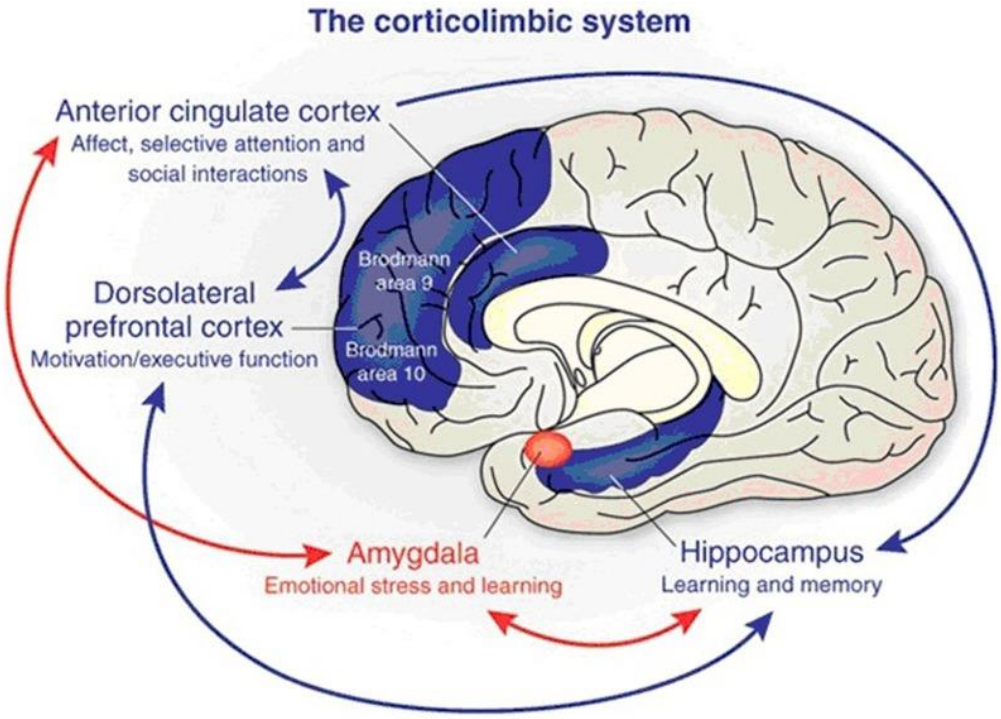

Figure 3 The corticolimbic system: hippocampus and its interactions.

Including the anterior cingulate cortex, dorsolateral prefrontal cortex, and amygdala. These interactions are crucial for learning, memory, affect, selective attention, motivation, executive function, and emotional stress. Adapted from (Leisman et al. 2012), distributed under Creative Commons Attribution license.

The intricate network of connections between the hippocampus and other brain regions highlights its central role in integrating diverse types of information and coordinating complex cognitive processes. These interactions enable the hippocampus to influence and be influenced by a wide range of neural systems, from those involved in emotional processing to those responsible for executive function and decision-making. Understanding these interactions is crucial for comprehending how the hippocampus contributes to various aspects of cognition and behavior, and how information is processed and transferred across different brain areas.

As we delve deeper into the hippocampus's role in memory processes, it becomes clear that these inter-regional interactions form the substrate upon which complex mnemonic functions are built. The hippocampus's ability to integrate information from multiple sources and coordinate activity across distributed brain networks is fundamental to its contributions to memory formation, consolidation, and retrieval.

E. Hippocampal contributions to memory

The hippocampus has long been recognized as a critical structure in memory formation and storage, with its role in various memory processes extensively studied over the past several decades.

1. Theoretical models of memory formation and storage

Several influential models have been proposed to explain the role of the hippocampus in memory formation and storage. One of the earliest and most influential is the Standard Model of Systems Consolidation (SMC) proposed by Squire and Alvarez (Squire and Alvarez 1995). This model postulate that the hippocampus is initially crucial for the encoding and retrieval of new memories, but its involvement gradually decreases over time as memories become consolidated in neocortical regions. According to this view, the hippocampus serves as a temporary storage site and index for distributed cortical representations, facilitating their integration and eventual independence from hippocampal support.

However, challenges to the SMC have led to the development of alternative models. The Multiple Trace Theory (MTT), proposed by Nadel and Moscovitch (Nadel and Moscovitch 1997), suggests that the hippocampus remains involved in the storage and retrieval of episodic memories throughout their lifetime. This theory posits that each reactivation of a memory creates a new hippocampal trace, leading to a distributed network of hippocampal-cortical connections that support the memory. The MTT accounts for findings that some very old memories, particularly those rich in episodic detail, continue to rely on hippocampal function.

A more recent model, the Complementary Learning Systems (CLS) theory, developed by Kumaran et al. (Kumaran, Hassabis, and McClelland 2016), attempts to reconcile these views. The CLS theory suggests that the hippocampus and neocortex work in complementary ways to support learning and memory. The hippocampus, with its ability to rapidly encode new

information, serves as a fast learning system that can acquire new memories without interfering with existing knowledge. In contrast, the neocortex acts as a slow learning system, gradually integrating new information into existing knowledge structures through repeated exposures and hippocampal-mediated replay during offline periods such as sleep.

While these theoretical models offer valuable perspectives on hippocampal function in memory, they are ultimately grounded in and informed by empirical evidence. A vast body of research has accumulated over the years, providing concrete support for the hippocampus's involvement in various types of memory processes.

2. Empirical evidence linking the hippocampus to various memory types

Extensive empirical research has linked the hippocampus to various types of memory, with particularly strong evidence for its role in episodic and spatial memory. Episodic memory, which involves the recollection of specific events tied to particular times and places, has been consistently associated with hippocampal function. Neuroimaging studies have shown increased hippocampal activation during both the encoding and retrieval of episodic memories (Dickerson and Eichenbaum 2010). Moreover, the degree of hippocampal activation during encoding has been found to predict subsequent memory performance, supporting its crucial role in the initial formation of episodic memories (Moscovitch et al. 2016).

While the hippocampus's role in episodic and spatial memory is well-established (Ekstrom and Ranganath 2018), its involvement in other forms of memory has been a subject of ongoing research and debate. Some studies have suggested a role for the hippocampus in certain types of semantic memory, particularly in the acquisition of new semantic information (Manns et al. 2003). However, the extent of hippocampal involvement in established semantic memories remains controversial.

The role of the hippocampus in working memory has also been investigated, with some studies suggesting hippocampal involvement in maintaining complex relational information over short delays (Ranganath and Blumenfeld 2005). However, the specific contributions of the hippocampus to working memory processes, as opposed to long-term memory encoding occurring during working memory tasks, remain a topic of ongoing research. Yonelinas et al. (Yonelinas et al. 2019) provide a contextual binding theory that reconsiders the relationship

between working memory, long-term memory, and hippocampal function, offering new perspectives on these intricate interactions.

The hippocampus has also been linked to imagination and future thinking. Studies have shown that hippocampal activity increases not only when individuals recall past events but also when they imagine future scenarios (Schacter et al. 2012). This finding has led to the development of the constructive episodic simulation hypothesis, which proposes that the same neural mechanisms involved in remembering the past are also engaged in imagining the future.

These diverse lines of research collectively highlight the hippocampus's multifaceted role in memory and cognition. From its well-established functions in episodic and spatial memory to its emerging roles in semantic memory, working memory, and future-oriented thinking, the hippocampus appears to be a crucial node in a complex network supporting various forms of memory and mental time travel. As research continues to evolve, our understanding of the hippocampus's contributions to different memory types and cognitive processes is likely to become even more nuanced and comprehensive.

One key mechanism that has emerged as fundamental to hippocampal function and memory processes is the presence of neural oscillations. These rhythmic fluctuations in neural activity provide a temporal framework for coordinating information processing within the hippocampus and between the hippocampus and other brain regions. Understanding these oscillations is crucial for comprehending how the hippocampus supports its various memory functions and how it interacts with other brain areas during memory formation, consolidation, and retrieval.

3. Neural oscillations in the hippocampus

Neural oscillations are a fundamental feature of brain activity, reflecting the rhythmic and synchronized firing of neuronal populations. In the hippocampus, these oscillations play a crucial role in organizing and coordinating neural activity, supporting various cognitive functions, particularly memory processes.

The brain's electrical activity exhibits rhythmic patterns across a wide range of frequencies, from slow oscillations below 1 Hz to ultra-fast oscillations above 100 Hz. These oscillations arise from the coordinated activity of large groups of neurons and are thought to provide a

temporal framework for information processing and communication between different brain regions (Buzsáki and Draguhn 2004).

In the hippocampus, three main types of oscillations have been extensively studied: theta (4-8 Hz), gamma (30-100 Hz), and sharp wave-ripples (SWRs) (Figure 4). Each of these oscillations is associated with specific behavioral states and cognitive processes , forming a complex symphony of neural activity that underlies hippocampal function.

Theta oscillations are perhaps the most prominent rhythm in the hippocampus. They are typically observed during active exploration and rapid eye movement (REM) sleep. Theta rhythms are generated through a complex interplay of intrinsic membrane properties, synaptic interactions, and inputs from other brain regions, particularly the medial septum (Colgin 2013).

The medial septum acts as a pacemaker, sending rhythmic inhibitory and excitatory inputs to hippocampal interneurons and pyramidal cells. This pacing signal interacts with the intrinsic resonance properties of hippocampal neurons to produce the theta rhythm. Hippocampal pyramidal cells, particularly in the CA1 region, exhibit intrinsic membrane oscillations in the theta frequency range, which are amplified and synchronized by the septal input (György Buzsáki 2002).

Theta oscillations play a crucial role in spatial navigation and memory encoding. O'Keefe and Recce (O'Keefe and Recce 1993) discovered that the firing of hippocampal place cells relative to the phase of theta oscillations carries information about an animal's position within its environment, a phenomenon known as phase precession. This temporal coding mechanism allows for the compression of spatial sequences within individual theta cycles, potentially facilitating the formation of associations between sequential events or locations.

The strength of theta oscillations during learning has been shown to predict subsequent memory performance (Sederberg et al. 2003). Moreover, disrupting theta rhythms through pharmacological or optogenetic manipulations impairs spatial memory and navigation (Wang et al. 2015), underscoring the critical role of these oscillations in hippocampal-dependent cognitive processes.

While theta oscillations dominate the hippocampal rhythm during active states, another important oscillatory pattern, gamma oscillations, plays a crucial role in local information processing and memory formation. Gamma oscillations (30-100 Hz) often co-occur with theta rhythms in the hippocampus and are thought to play a role in local information processing and

memory formation. Gamma oscillations can be further subdivided into slow gamma (30-50 Hz) and fast gamma (50-100 Hz), which may serve distinct functions (Colgin et al. 2009).

The generation of gamma oscillations primarily involves the interplay between excitatory pyramidal cells and inhibitory interneurons, particularly parvalbumin-positive basket cells. The rhythmic inhibition provided by these interneurons creates windows of opportunity for pyramidal cell firing, resulting in the gamma rhythm (Bartos, Vida, and Jonas 2007).

Gamma oscillations are thought to support memory encoding by promoting spike-timing-dependent plasticity. The precise timing of neuronal firing within gamma cycles may facilitate the strengthening of synaptic connections between co-active neurons, promoting the formation of cell assemblies representing specific memories or experiences (Fell and Axmacher 2011).

In addition to theta and gamma oscillations, which are prominent during active states, the hippocampus exhibits a third major oscillatory pattern during periods of rest and sleep: Sharp wave-ripples (SWRs). These brief (50-100 ms) high-frequency oscillations (140-250 Hz) superimposed on sharp waves, which are large amplitude deflections in the local field potential. SWRs typically occur during periods of quiet wakefulness and slow-wave sleep, when the hippocampus is relatively decoupled from external sensory inputs (György Buzsáki 2015).

The generation of SWRs involves a complex interplay of cellular and network mechanisms. Sharp waves originate from the synchronous firing of CA3 pyramidal cells, which then excite CA1 pyramidal cells and interneurons. The fast ripple oscillation is generated locally in CA1 through the interaction of pyramidal cells and fast-spiking interneurons (Ylinen et al. 1995).

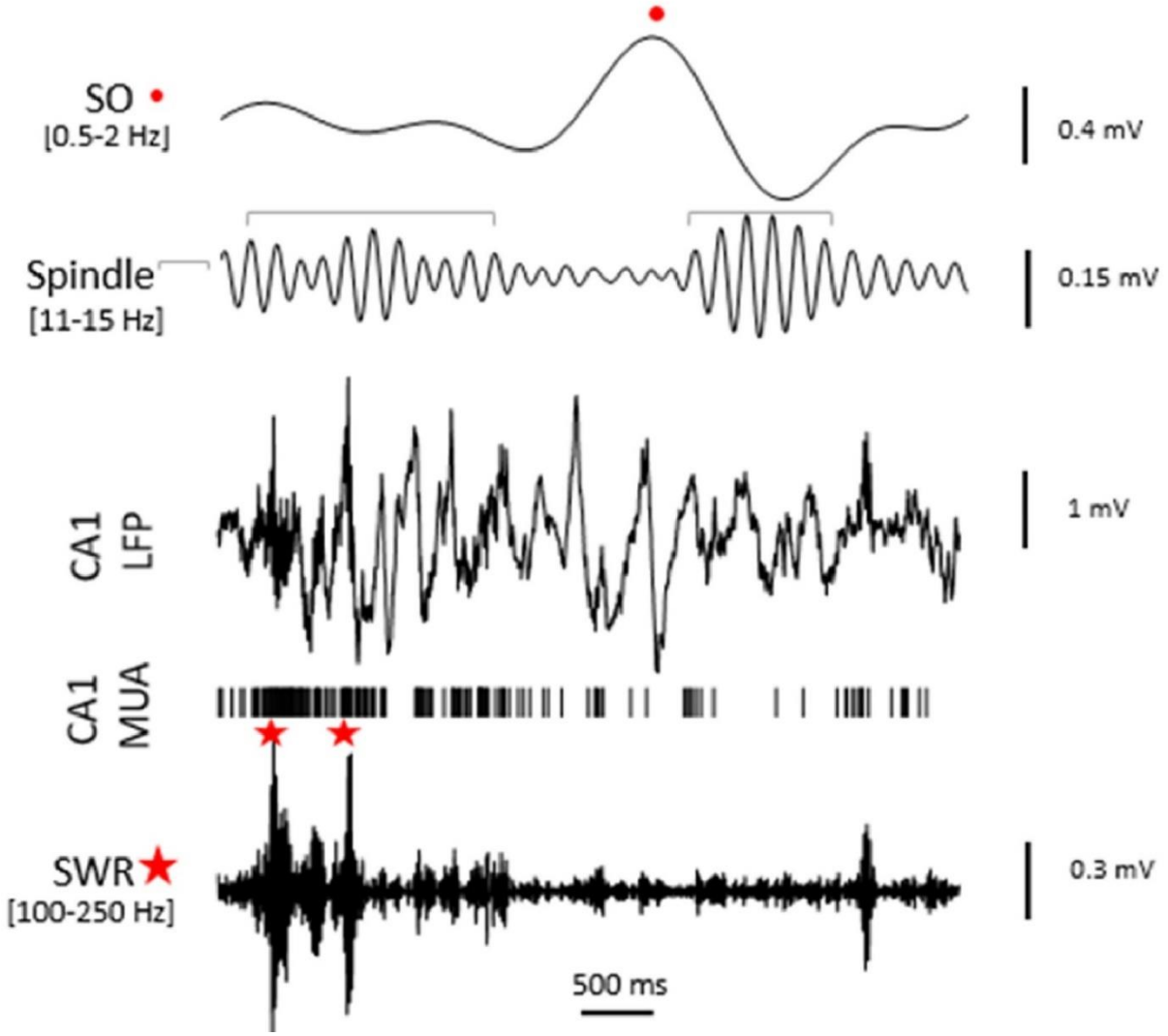

Figure 4 Oscillatory events involved in long-term memory consolidation.

Slow oscillations, sleep spindles, and sharp wave ripples (SWRs) are depicted, each contributing to memory replay and integration into long-term storage. Adapted from (Azimi, Alizadeh, and Ghorbani 2021), distributed under Creative Commons Attribution license.

The relationship between hippocampal oscillations and memory functions has been extensively studied, with each oscillatory pattern contributing uniquely to various aspects of memory processing. Gamma oscillations, particularly their coupling with theta rhythms, have been linked to successful memory encoding and retrieval. The strength of theta-gamma coupling has been shown to predict memory performance in both rodents and humans (Tort et al. 2009; Axmacher et al. 2010). This coupling may serve to organize discrete memory items within the theta cycle, allowing for the encoding and retrieval of ordered sequences.

SWRs have been most closely associated with memory consolidation processes (Figure 4). The selective disruption of SWRs during post-learning sleep has been shown to impair subsequent memory performance, highlighting their causal role in memory consolidation (Girardeau et al. 2009). Conversely, artificially inducing SWRs can enhance memory performance, further supporting their role in consolidation processes (Fernández-Ruiz et al. 2019).

The propagation of oscillatory activity across neural circuits is a key mechanism by which the hippocampus communicates with other brain regions during memory processes. Theta oscillations, for example, can synchronize activity between the hippocampus and prefrontal cortex during spatial working memory tasks (Benchenane et al. 2010). This long-range synchronization may facilitate the integration of spatial information with goal-directed behavior.

Similarly, SWRs in the hippocampus are temporally coordinated with oscillatory events in other brain regions, including the prefrontal cortex and the ventral striatum. This coordination is thought to support the transfer of reactivated memory traces from the hippocampus to neocortical areas for long-term storage (Peyrache et al. 2009).

Recent technological advances have allowed for more precise manipulation and recording of neural oscillations, providing new insights into their causal role in memory processes. Optogenetic techniques, for example, have enabled researchers to selectively manipulate specific neuronal populations at precise phases of ongoing oscillations, allowing for a more detailed understanding of how oscillatory dynamics contribute to memory functions (Iaccarino et al. 2016). Furthermore, closed-loop stimulation approaches have demonstrated that theta phase-specific manipulation of hippocampal activity can enhance encoding and retrieval functions, highlighting the critical role of oscillatory timing in memory processes (Siegle and Wilson 2014). These studies, among others, have Advanced our understanding of how neural oscillations orchestrate complex cognitive functions, revealing the intricate relationship between oscillatory phase, neuronal activity, and behavior (Benchenane et al. 2010; Fernández-Ruiz et al. 2019). Such phase-specific manipulations offer a powerful approach for dissecting the functional roles of neural oscillations in cognition and behavior, paving the way for more nuanced models of hippocampal function in memory and spatial navigation.

The study of neural oscillations in the hippocampus has provided crucial insights into the mechanisms underlying memory formation, consolidation, and retrieval. These rhythmic patterns of activity serve as a fundamental organizing principle for hippocampal function,

coordinating information processing both within the hippocampus and between the hippocampus and other brain regions. As our understanding of these oscillations continues to grow, so does our appreciation for their complexity and their critical role in cognitive processes.

Among the various oscillatory patterns observed in the hippocampus, Sharp Wave Ripples (SWRs) stand out as particularly intriguing and important for memory consolidation. These brief, high-frequency events have become a focal point of research due to their unique properties and their hypothesized role in the transfer of information from the hippocampus to neocortical areas. Given their significance in memory processes and their relevance to our research, a more detailed examination of SWRs is warranted.

F. Sharp Wave Ripples (SWRs) overview

1. Historical perspective on SWR research

Sharp wave ripples (SWRs) have captivated neuroscientists for decades, offering a window into the intricate workings of memory consolidation and information processing in the brain. The story of SWRs begins in the early 1970s with a groundbreaking discovery by O'Keefe and Dostrovsky (O'Keefe and Dostrovsky 1971). They identified place cells in the hippocampus, neurons that fire when an animal occupies specific locations in its environment. This finding sparked intense interest in hippocampal oscillations and laid the groundwork for future SWR research.

However, it was not until the early 1980s that SWRs themselves took center stage. György Buzsáki and his colleagues made a pivotal discovery in 1983, reporting high-frequency oscillations in the CA1 region of rat hippocampus during slow-wave sleep and periods of awake immobility (G. Buzsáki, Leung, and Vanderwolf 1983). These oscillations, which they initially called "ripples," were brief (lasting 50-100 milliseconds) and had a high-frequency component (140-250 Hz). This finding revealed a new form of synchronized neuronal activity occurring outside periods of active behavior.

The discovery of SWRs opened up a new avenue of research in hippocampal function. Throughout the 1980s and early 1990s, researchers focused on characterizing these events and understanding their relationship to other hippocampal oscillations, such as theta rhythms. The work of John O'Keefe and his colleagues was particularly influential during this period, as

they continued to explore the relationship between place cell activity and various hippocampal oscillatory patterns (O'Keefe and Recce 1993).

As our knowledge of SWRs has grown, so has our curiosity about the mechanisms underlying their generation. Understanding how these complex oscillatory events are produced in the hippocampal circuit is crucial for unraveling their role in memory processes and for developing targeted interventions to modulate memory function.

2. Generation mechanisms

SWRs are complex events consisting of two primary components: a sharp wave, which is a large amplitude deflection in the local field potential, and superimposed high-frequency ripple oscillations (György Buzsáki 2015; Colgin 2016) (Figure 6).

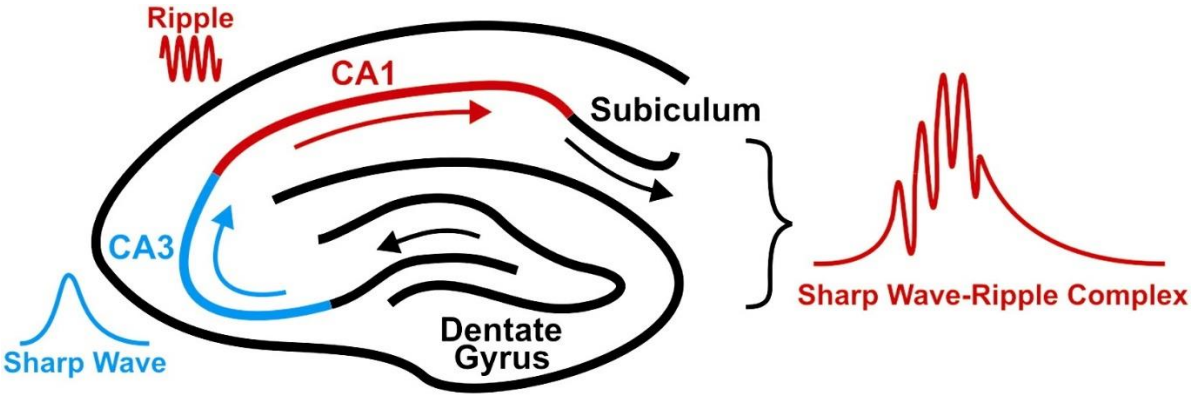

Figure 5 Schematic representation SWR generation in the hippocampus.

The sharp wave component (depicted in blue) originates from the CA3 region and travels to the CA1 region, where it induces a large amplitude deflection in the local field potential. The ripple component (depicted in red) superimposes on the sharp wave within the CA1 pyramidal cell layer, consisting of high-frequency oscillations in the 140-200 Hz range. The right side of the figure shows a typical sharp wave-ripple complex as recorded in the local field potential. Adapted from (Beenhakker and Huguenard 2009), distributed under Creative Commons Attribution license.

The sharp wave component is typically observed as a negative deflection in the CA1 stratum radiatum, reflecting strong depolarization of CA1 pyramidal cell dendrites. The ripple component, on the other hand, is most prominent in the CA1 pyramidal cell layer and consists of fast oscillations in the 140-200 Hz range (Figure 6).

The generation of SWRs involves an intricate interplay of cellular and network mechanisms. It all starts with a buildup of excitatory activity in the recurrent collateral system of CA3 pyramidal cells. This activity then propagates to CA1, triggering the fast oscillatory response known as the ripple. Various types of interneurons, particularly parvalbumin-positive basket cells, play a crucial role in orchestrating the precise timing and coordination of neuronal firing during SWRs (Schlingloff et al. 2014; Stark et al. 2014).

The exact mechanisms underlying ripple generation have been a subject of intense research. One prominent model suggests that the reciprocal interactions between pyramidal cells and fast-spiking interneurons create a feedback loop that generates and sustains the high-frequency oscillations (Ylinen et al. 1995). According to this model, the synchronized firing of pyramidal cells activates interneurons, which in turn provide phasic inhibition back onto the pyramidal cells. This rhythmic inhibition creates temporal windows for pyramidal cell firing, resulting in the observed high-frequency oscillations.

A recent computational study by Latimer et al. (2023) has provided further insights into the generation mechanisms of SWRs. Using a large-scale biophysically realistic model of the CA1 hippocampus, they demonstrated that the interplay between different neuron types and microcircuit motifs can reproduce key aspects of physiological SWRs. Their model highlights the role of chandelier interneurons in orchestrating the temporal structure and frequency bands of SWRs, offering new perspectives on the cellular and network mechanisms underlying these events.

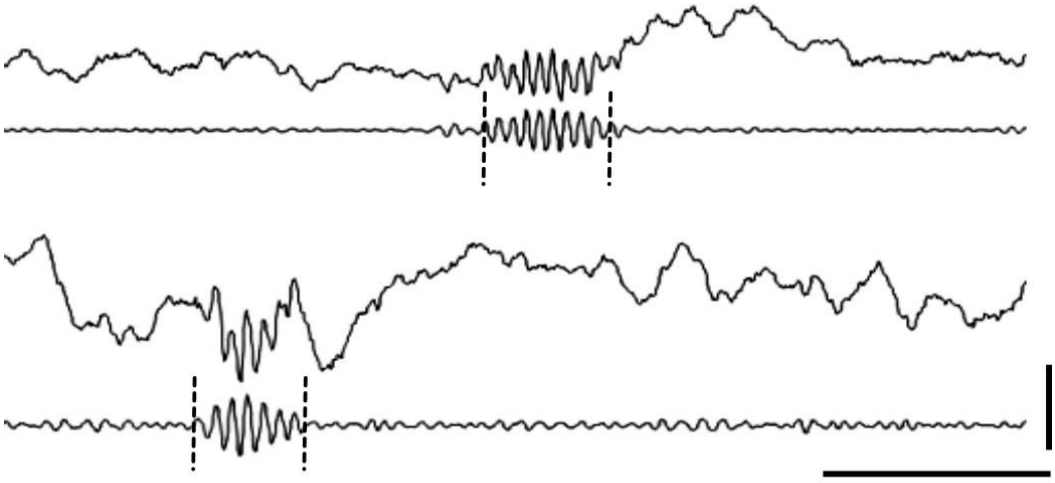

Figure 6 Example of sharp wave-ripple (SWR) events.

SWR events recorded in the hippocampus. Each example shows the raw signal (top) and the filtered signal (bottom), illustrating the sharp wave with superimposed high-frequency ripple oscillations. Horizontal scale 100 ms, vertical scale 1mV.

As we delve deeper into the functional role of SWRs, it becomes clear that these brief but powerful events play a pivotal role in various aspects of memory processing and consolidation. The unique properties of SWRs, stemming from their generation mechanisms, make them ideally suited for coordinating information transfer both within the hippocampus and between the hippocampus and other brain regions.

3. Functional role of SWRs in hippocampal operations

The functional role of SWRs in hippocampal operations and memory consolidation has been a subject of intense investigation. Numerous studies have shown that SWRs provide a temporal framework for the reactivation of neuronal ensembles representing past experiences. This reactivation, or replay, is thought to be critical for transferring information from the hippocampus to neocortical areas for long-term storage, a process central to systems consolidation theory (Buzsáki 2015; Joo and Frank 2018).

A landmark study by Wilson and McNaughton (Wilson and McNaughton 1994) demonstrated that hippocampal place cells that fired together during spatial exploration also tended to co-activate during subsequent sleep, specifically during SWR events. This observation provided

the first evidence for the replay of waking experiences during sleep, a phenomenon now recognized as crucial for memory consolidation.

Subsequent studies have further characterized the nature of this replay. Researchers have found that the reactivation of place cell sequences during SWRs often occurs in a temporally compressed manner, with events that took place over several seconds during behavior being replayed in just 50-100 milliseconds during a single SWR (Nádasdy et al. 1999). This temporal compression may facilitate synaptic plasticity mechanisms and allow for the rapid transfer of information to cortical areas.

As our understanding of SWRs' functional role continues to evolve, it becomes increasingly clear that these events are not confined to a single behavioral state. While initially discovered during slow-wave sleep, SWRs have been observed across various behavioral contexts, each potentially serving distinct cognitive functions. This realization has led to a more nuanced view of SWRs and their contributions to hippocampal operations.

4. SWRs in various behavioral states

While initially discovered during slow-wave sleep, SWRs have since been observed across various behavioral states, each potentially serving distinct functions. During slow-wave sleep, SWRs are thought to play a crucial role in memory consolidation, facilitating the transfer of information from the hippocampus to neocortical areas. The coordination of SWRs with neocortical slow oscillations and thalamocortical spindles during sleep has been proposed as a mechanism for this information transfer (Sirota et al. 2003) (Figure 4).

Interestingly, SWRs aren't just important during sleep. Recent research has revealed that they also occur during periods of awake immobility and may contribute to online cognitive processes such as decision-making and planning. Joo and Frank (Joo and Frank 2018) proposed that awake SWRs might support the retrieval of stored representations for immediate use in guiding behavior, suggesting a dual role for these events in both memory consolidation and retrieval. This idea is supported by studies showing that disrupting awake SWRs can impair performance on spatial memory tasks (Jadhav et al. 2012).

During active waking states, particularly during exploratory behavior, SWRs are less frequent but still occur during brief pauses in movement. These awake SWRs have been associated with

the replay of both past and potential future trajectories, suggesting a role in navigational planning and decision-making (Pfeiffer and Foster 2013).

Recent studies have also observed SWRs during rapid eye movement (REM) sleep, although they are less frequent and have different characteristics compared to those in non-REM sleep. The function of these REM sleep SWRs is still unclear, but they may play a role in emotional memory processing or in integrating new memories with existing knowledge (Boyce et al. 2016).

The presence of SWRs across various behavioral states underscores their versatility and importance in hippocampal function. From consolidating memories during sleep to guiding decision-making during wakefulness, SWRs appear to be a fundamental mechanism by which the hippocampus processes and utilizes information. This diversity of roles highlights the need for a nuanced understanding of SWRs that takes into account the behavioral context in which they occur.

The multifaceted nature of SWRs naturally leads us to a deeper examination of their role in memory processes. At the heart of this investigation lies the intricate relationship between SWRs, memory consolidation, and synaptic plasticity. These interconnected processes form the bedrock of our understanding of how memories are formed, stored, and retrieved in the brain.

G. SWRs in memory consolidation and synaptic plasticity

1. Hypotheses on SWRs and memory consolidation

The primary hypothesis regarding sharp wave ripples (SWRs) in memory consolidation centers on their ability to facilitate the transfer of information from the hippocampus to neocortical areas. This process, known as systems consolidation, is crucial for the formation of long-term memories. SWRs provide a mechanism for the compressed replay of recent experiences, allowing for the strengthening of synaptic connections that represent these memories in distributed cortical networks.

The "two-stage model" of memory consolidation, proposed by Buzsáki (G. Buzsáki 1989), postulates that initial encoding of experiences occurs during theta-dominated states (such as during active exploration), while subsequent consolidation occurs during SWR-dominated states (such as quiet wakefulness or slow-wave sleep). According to this model, SWRs serve as

a mechanism for transferring information from the hippocampus, which acts as a temporary storage buffer, to neocortical areas for long-term storage.

Another hypothesis suggests that SWRs play a role in the integration of new information with existing knowledge structures or schemas. This process may involve the coordinated reactivation of both recently acquired and older memories, facilitating the incorporation of new experiences into a broader cognitive framework. This idea is supported by studies showing that SWRs can reactivate neural patterns associated with remote as well as recent memories (Kudrimoti, Barnes, and McNaughton 1999).

Recent work has also proposed that SWRs might be involved in memory reconsolidation, a process by which previously consolidated memories become labile upon reactivation and require re-stabilization. According to this hypothesis, SWRs could provide a mechanism for updating existing memories with new information, allowing for the flexible adaptation of memory representations over time (Genzel et al. 2017).

These hypotheses collectively paint a picture of SWRs as a versatile mechanism for memory processing, capable of supporting not just the initial consolidation of memories, but also their integration, updating, and reconsolidation. This multifaceted role of SWRs in memory consolidation underscores their importance in maintaining a flexible and adaptive memory system.

While these hypotheses provide a theoretical framework for understanding the role of SWRs in memory consolidation, the underlying mechanisms by which SWRs influence memory at the cellular level involve modulation of synaptic plasticity. Understanding how SWRs affect synaptic strength and connectivity is crucial for bridging the gap between network-level phenomena and cellular-level changes that support memory formation and consolidation.

2. Role of SWRs in modulating synaptic plasticity

The ability of SWRs to influence synaptic plasticity lies at the heart of their role in memory consolidation. While we previously discussed systems consolidation, which involves the transfer of information from the hippocampus to neocortical areas, here we focus on cellular consolidation, the process by which memories are stabilized at the synaptic level. Synaptic plasticity refers to the capacity of synapses to strengthen or weaken over time in response to increases or decreases in their activity. This process is fundamental to learning and memory

formation. SWRs, with their unique temporal and spatial characteristics, create ideal conditions for inducing lasting changes in synaptic strength.

During SWRs, large populations of neurons fire in synchrony, creating a powerful burst of activity. This synchronized firing leads to substantial depolarization of postsynaptic neurons, which can trigger molecular cascades associated with synaptic strengthening. The high-frequency component of SWRs, the ripples, is particularly effective at inducing long-term potentiation (LTP), a form of synaptic strengthening thought to underlie memory formation (Buzsáki 2015).

Research has shown that the timing of neuronal activity relative to SWRs is critical for determining the direction and magnitude of synaptic changes. Neurons that fire consistently during SWRs are more likely to strengthen their connections, while those that remain silent may see their synapses weakened. This timing-dependent plasticity helps to reinforce neural circuits that represent specific memories or experiences (Sadowski, Jones, and Mellor 2016).

The role of SWRs in modulating synaptic plasticity provides a crucial link between network-level phenomena and the cellular-level changes that support memory formation and consolidation. By creating optimal conditions for synaptic strengthening and weakening, SWRs act as a powerful mechanism for selectively reinforcing and pruning neural connections, thereby shaping the neural networks that encode our memories and experiences.

To fully appreciate how SWRs influence memory at the cellular level, it's essential to delve deeper into the specific mechanisms of synaptic plasticity. Two key processes, long-term potentiation (LTP) and long-term depression (LTD), play crucial roles in shaping synaptic strength and, consequently, memory formation and storage.

3. Mechanisms of long-term potentiation and depression

Long-term potentiation (LTP) is a persistent strengthening of synapses based on recent patterns of activity. It is widely considered one of the major cellular mechanisms underlying learning and memory. SWRs create ideal conditions for LTP induction due to their ability to generate strong, synchronized depolarization of postsynaptic neurons.

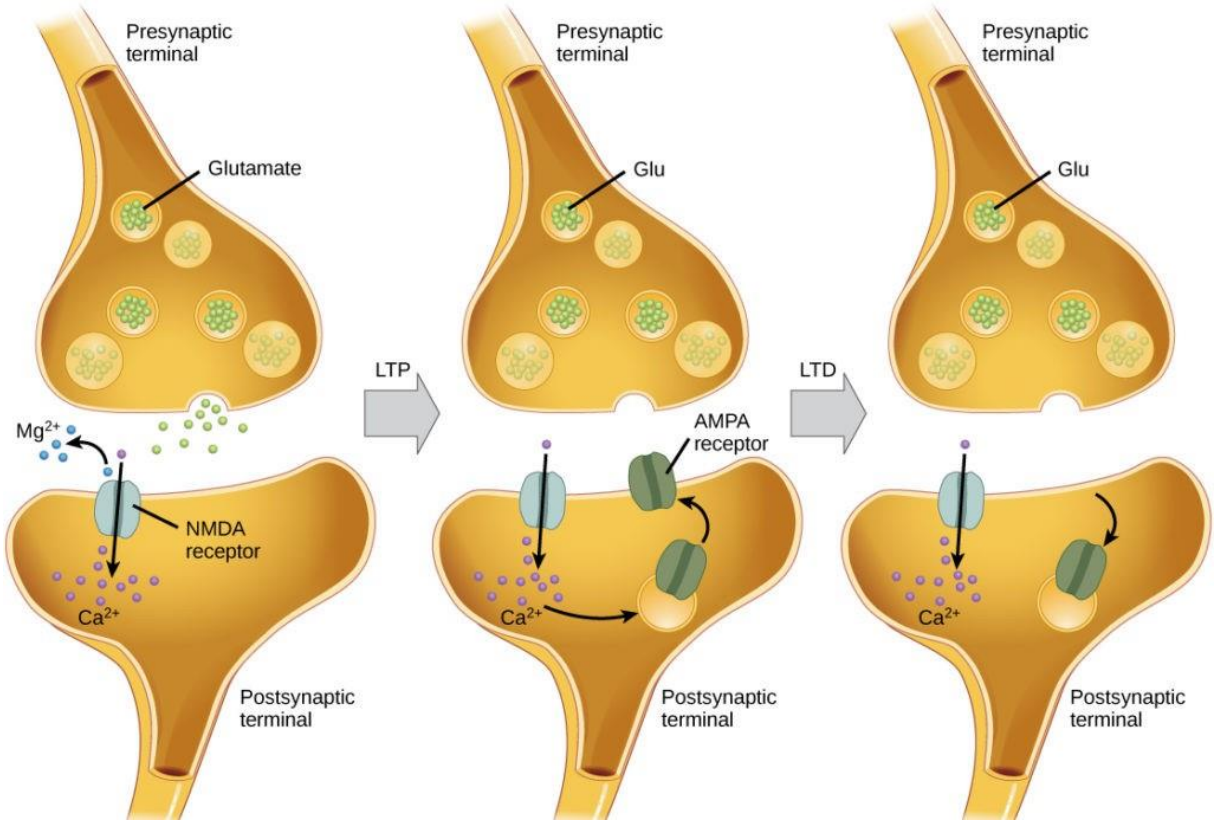

Figure 7 Synaptic mechanisms of Long-Term Potentiation (LTP) and Long-Term Depression (LTD).

During LTP, the repeated activation of synapses and high-frequency stimulation leads to the removal of the magnesium block from NMDA receptors, allowing calcium influx into the postsynaptic cell. This triggers a cascade of intracellular events that result in the insertion of AMPA receptors into the postsynaptic membrane, strengthening the synapse. In contrast, LTD is induced by lower frequency stimulation or specific patterns of synaptic activity, leading to a reduction in calcium influx and a decrease in AMPA receptor density, thereby weakening the synapse.

The high-frequency ripple component of SWRs is particularly effective at inducing LTP. These fast oscillations (140-200 Hz) fall within the frequency range that has been shown to be optimal for LTP induction in slice preparations of hippocampal tissue. When neurons fire repeatedly at this frequency, it leads to a substantial influx of calcium into postsynaptic neurons, triggering molecular cascades that result in the insertion of more AMPA receptors into the postsynaptic membrane, thereby strengthening the synapse (Magee and Johnston 1997).

Importantly, the coordination of neuronal firing during SWRs also promotes spike-timing-dependent plasticity (STDP). In STDP, the precise timing of pre- and postsynaptic spikes determines whether a synapse is strengthened or weakened. The compressed timescale of

neuronal sequence replay during SWRs provides an ideal framework for STDP, allowing for the strengthening of synapses between neurons that represent related aspects of an experience or memory (Bi and Poo 1998).

While much attention has been given to the role of SWRs in promoting LTP, these events can also induce long-term depression (LTD) under certain conditions. LTD, a long-lasting decrease in synaptic strength, is equally important for learning and memory as it allows for the weakening of irrelevant or competing neural pathways.

The induction of LTD during SWRs depends on the specific patterns of activity and the state of the synapse. Neurons that consistently fail to fire during SWRs, or that fire uncoordinated with the population, may undergo LTD. This process could help to refine memory representations by weakening connections that do not contribute to the reactivated memory trace (Collingridge et al. 2010).

Moreover, the relationship between SWRs and LTD may be important for memory updating and forgetting. By selectively weakening certain synapses, SWRs might facilitate the updating of existing memories with new information or the gradual forgetting of less relevant details (Hardt, Nader, and Nadel 2013).

To further our understanding of SWRs and their role in memory processes, it is essential to have robust methods for detecting and analyzing these events in neural recordings. The characterization of SWRs presents unique challenges due to their brief duration and the complex neural activity patterns they represent. In the following section, we will explore the various techniques used to detect, isolate, and analyze SWRs.

H. Characterization techniques for sharp wave ripples (SWRs)

1. Methodologies for detection and isolation of SWRs

The accurate characterization of SWRs is fundamental to understanding their function in hippocampal circuits and their broader impact on cognitive processes. The detection and isolation of SWRs from continuous neural recordings present significant challenges due to the complexity of neural signals and the variability of SWR events. Traditional methods for SWR detection often rely on threshold-based approaches applied to filtered local field potential (LFP) recordings. Typically, the LFP signal is band-pass filtered in the ripple frequency range (usually 80-250 Hz), and events exceeding a certain threshold (often set as a multiple of the

standard deviation of the filtered signal) are identified as potential SWRs. This approach, while straightforward, can be sensitive to noise and may miss subtle SWR events.

More sophisticated detection methods have been developed to improve the accuracy and reliability of SWR identification. For instance, Stark et al. (Stark et al. 2014) proposed a method that combines multi-channel recordings with independent component analysis (ICA) to isolate SWR-associated components from background activity. This approach leverages the spatial coherence of SWRs across multiple recording sites to enhance detection accuracy.

Another advanced approach involves the use of wavelet transform techniques, which provide better time-frequency resolution compared to traditional Fourier-based methods. Sullivan et al. (Sullivan et al. 2011) demonstrated the effectiveness of continuous wavelet transforms in detecting and characterizing SWRs, allowing for a more nuanced analysis of their spectro-temporal properties.

Recently, machine learning approaches, particularly deep learning models, have shown promise in SWR detection. Navas-Olive et al. (Navas-Olive et al. 2022) introduced a convolutional neural network-based method for SWR detection that outperformed traditional threshold-based approaches in terms of both accuracy and computational efficiency. These data-driven approaches have the potential to capture more subtle and variable SWR events that might be missed by conventional methods.

While accurate detection of SWRs is crucial, it is only the first step in unraveling their complex roles in hippocampal function and memory processes. To truly understand the significance of SWRs, researchers have developed a range of innovative approaches that go beyond mere detection. These state-of-the-art techniques aim to decipher the information content of SWRs, their relationship to behavior, and their causal role in memory formation and consolidation.

2. State of the art approaches to decipher the role of SWRs

The investigation of sharp wave ripples (SWRs) in the context of learning tasks has been a cornerstone of hippocampal research, providing crucial insights into the neural mechanisms underlying memory formation and consolidation. One of the seminal studies in this field was conducted by Wilson and McNaughton (Wilson and McNaughton 1994), who demonstrated the reactivation of place cell ensembles during SWRs following spatial exploration. Using multi-electrode recordings in the hippocampus of rats, they showed that pairs of neurons with

overlapping place fields during maze running exhibited increased co-activation during subsequent sleep, particularly during SWR events. This study established a crucial link between waking experience and SWR-associated replay, laying the groundwork for subsequent research on the role of SWRs in memory consolidation.

Building on this foundation, Lee and Wilson (Lee and Wilson 2002) employed similar multi-electrode recording techniques to investigate the temporal structure of hippocampal replay during SWRs. They found that sequences of place cell activations observed during running were replayed in the same order during subsequent sleep, but at a compressed timescale. This temporal compression of neural sequences during SWRs suggested a mechanism for rapid memory consolidation and information transfer to neocortical regions.

A significant advancement in SWR research came with the development of closed-loop stimulation techniques, which allowed for the selective manipulation of SWRs during specific behavioral states. Girardeau et al. (Girardeau et al. 2009) used this approach to investigate the causal role of SWRs in memory consolidation. By delivering electrical stimulation to the ventral hippocampal commissure to suppress SWRs during post-learning sleep, they demonstrated that SWR disruption impaired spatial memory performance in rats. This study provided compelling evidence for the necessity of SWRs in memory consolidation processes.

Complementing these disruption studies, Fernández-Ruiz et al. (Fernández-Ruiz et al. 2019) employed optogenetic techniques to selectively enhance SWRs during slow-wave sleep. Using a transgenic mouse line expressing channelrhodopsin in pyramidal cells, they were able to induce artificial SWRs with light stimulation. They found that augmenting SWRs during post-learning sleep improved memory performance on a spatial task, further supporting the causal role of SWRs in memory consolidation.

The investigation of awake SWRs and their role in ongoing cognitive processes has been another important area of research. Pfeiffer and Foster (Pfeiffer and Foster 2013) used high-density electrophysiological recordings in rats performing a spatial memory task to examine the content of awake SWRs. They discovered that these events often contained representations of future trajectories to goal locations, suggesting a role for SWRs in planning and decision-making. This study highlighted the potential importance of awake SWRs in online cognitive processes, expanding our understanding of their function beyond offline consolidation.

Jadhav et al. (Jadhav et al. 2012) provided causal evidence for the role of awake SWRs in spatial working memory. Using a closed-loop disruption paradigm, they selectively suppressed awake SWRs as rats performed a spatial alternation task. The disruption of awake SWRs led to impaired performance on the working memory component of the task, demonstrating their importance in ongoing memory processes.

Research on SWRs has also extended to non-spatial learning paradigms. For instance, Ólafsdóttir et al. (Ólafsdóttir, Bush, and Barry 2018) investigated SWR replay in an object-location association task. Using a combination of electrophysiological recordings and optogenetic manipulations in mice, they demonstrated that SWRs support the reactivation of object-location pairs, suggesting a role for these events in binding different elements of episodic memories.

The development of advanced recording techniques has allowed for increasingly detailed analyses of SWR-associated neural activity. Jun et al. (Jun et al. 2017) introduced the Neuropixels probe, a high-density silicon probe capable of recording from hundreds of neurons simultaneously across multiple brain regions. This technology has enabled researchers to study the coordination of SWRs across different hippocampal subregions and between the hippocampus and other brain areas involved in memory processing.

As our understanding of SWRs deepens and our methodologies become more sophisticated, new questions and challenges emerge. One area that continues to present difficulties is the accurate classification of SWRs, particularly in the context of their diverse functional roles and the variability in their characteristics across different behavioral states and cognitive processes.

3. Limitations of current classification methods

Despite these significant advances, current methods for classifying and analyzing SWRs have several limitations. One major challenge is the variability in SWR detection methods across studies. While most approaches rely on thresholding of filtered local field potentials, the specific frequency bands and threshold criteria can vary widely. This lack of standardization can make it difficult to compare results across different studies and laboratories.

Another limitation is the potential for false positives and false negatives in SWR detection. High-frequency oscillations similar to ripples can occur during epileptiform activity or as a result of movement artifacts, leading to misclassification. Conversely, genuine SWRs may be

missed if they do not meet the predefined detection criteria, particularly if they are of low amplitude or short duration.

The analysis of SWR content presents additional challenges. While many studies have focused on the reactivation of place cell sequences, characterizing the content of SWRs in non-spatial tasks or for more complex memory representations remains difficult. Current methods may not fully capture the richness and complexity of information encoded in these events.

Furthermore, the relationship between SWRs and behavior is not always straightforward. While some studies have found clear correlations between SWR properties and subsequent memory performance, others have reported more complex or task-specific relationships. This variability highlights the need for more sophisticated analytical approaches that can account for the multifaceted nature of memory processes.

The limitations of current classification methods have spurred the development of new approaches, particularly in the realm of machine learning and artificial intelligence. Hsu et al. (Hsu et al. 2021) introduced a deep learning-based method for SWR classification, which demonstrated improved performance compared to traditional threshold-based approaches. Such machine learning techniques offer the potential to more accurately detect and characterize SWRs, potentially revealing subtle features that may be missed by conventional methods.

The potential of machine learning and artificial intelligence to address these limitations points to a broader trend in neuroscience: the increasing integration of advanced computational methods into neurological research. This intersection of neuroscience and artificial intelligence has a rich history and has led to significant advancements in both fields.

I. Application of deep learning in neuroscience

1. Historical overview of AI in neuroscience

The intersection of artificial intelligence (AI) and neuroscience has a rich and complex history, dating back to the mid-20th century. This convergence of disciplines has not only advanced our understanding of the brain but has also inspired new approaches in AI development. The historical overview of AI in neuroscience reflects a bidirectional relationship, where each field has significantly influenced the other's progress and conceptual frameworks.

The roots of this interdisciplinary connection can be traced to the 1940s and 1950s, with the emergence of cybernetics and the development of the first artificial neural networks. Warren McCulloch and Walter Pitts (McCulloch and Pitts 1943) proposed a mathematical model of a neuron, which laid the foundation for the field of neural networks. This model, though simplistic by today's standards, represented a crucial step in bridging the gap between biological neurons and computational units.

In the 1950s, Frank Rosenblatt introduced the perceptron, an early machine learning algorithm inspired by the structure of the visual system (Rosenblatt 1958). The perceptron was designed to perform binary classification and was hailed as a significant breakthrough in AI. However, its limitations, particularly in solving non-linearly separable problems, were later highlighted by Minsky and Papert (Minsky and Papert 1969), leading to a temporary decline in neural network research.

The 1960s and 1970s saw the rise of symbolic AI, which focused on rule-based systems and logical reasoning. While this approach diverged from the neural network paradigm, it still drew inspiration from cognitive science and theories of human problem-solving. For instance, Newell and Simon's (Newell and Simon 1972) work on the General Problem Solver was influenced by theories of human cognition and aimed to mimic human problem-solving strategies.

The 1980s marked a resurgence of interest in neural networks, due to the development of new learning algorithms and architectural innovations. The introduction of backpropagation by Rumelhart, Hinton, and Williams (Rumelhart, Hinton, and Williams 1986) provided an efficient method for training multi-layer neural networks, overcoming many of the limitations of earlier perceptrons. This period also saw the emergence of connectionist models in cognitive science, which sought to explain cognitive processes in terms of distributed representations and parallel processing (Rumelhart, McClelland, and AU 1986).

The 1990s and early 2000s witnessed a growing synergy between neuroscience and machine learning. Advances in neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), generated vast amounts of brain data, creating new opportunities for applying machine learning algorithms to neuroscientific problems. Concurrently, neuroscientific insights into brain structure and function continued to inspire new AI architectures and learning algorithms.

A significant milestone in this period was the development of convolutional neural networks (CNNs) by LeCun et al. (Lecun et al. 1998), which were inspired by the hierarchical organization of the visual cortex. CNNs proved highly effective in image recognition tasks and have since become a cornerstone of modern deep learning applications in computer vision.

The turn of the millennium saw an explosion of interest in deep learning, characterized by neural networks with many layers capable of learning hierarchical representations from data. This resurgence was driven by advances in computing power, the availability of large datasets, and algorithmic innovations. Notably, the work of Hinton, et al. (Hinton, Osindero, and Teh 2006) on deep belief networks demonstrated the potential of unsupervised pre-training for deep neural networks, sparking renewed interest in neural network research.

In recent years, the interplay between AI and neuroscience has intensified, with each field contributing to the other's advancement. Neuroscience-inspired AI models, such as recurrent neural networks with long short-term memory (LSTM) units (Chambers et al. 2024), have found wide applications in sequence learning tasks. Conversely, AI techniques have become invaluable tools in neuroscience for analyzing complex brain data, decoding neural signals, and modeling brain function.

2. Overview of deep learning applications in neuroscience

Deep learning, a subset of machine learning characterized by artificial neural networks with multiple layers, has found numerous applications in neuroscience, transforming the way researchers analyze and interpret complex brain data. These applications span a wide range of areas, from neuroimaging analysis to neural decoding and brain-computer interfaces.

One of the most prominent applications of deep learning in neuroscience is in the analysis of neuroimaging data. Functional Magnetic Resonance Imaging (fMRI) studies, which generate vast amounts of high-dimensional data, have particularly benefited from deep learning approaches. Convolutional Neural Networks (CNNs), originally developed for computer vision tasks, have been adapted to analyze fMRI data, enabling more sophisticated feature extraction and pattern recognition. For instance, Wen et al. (Wen et al. 2020) demonstrated the use of CNNs for decoding cognitive states from fMRI data, achieving superior performance compared to traditional machine learning methods.

Deep learning has also made significant contributions to the field of connectomics, which aims to map the neural connections in the brain. The process of reconstructing neural circuits from electron microscopy images is labor-intensive and time-consuming. Deep learning models, particularly U-Net architectures, have been employed to automate the segmentation of neuronal structures in these images. Januszewski et al. (Januszewski et al. 2018) developed a flood-filling network that significantly improved the accuracy and efficiency of neuron reconstruction from electron microscopy data.

In the domain of electrophysiology, deep learning has been employed to analyze and interpret complex neural signals, with applications ranging from decoding movement intentions to detecting specific neural events. Recurrent Neural Networks (RNNs), particularly those incorporating Long Short-Term Memory (LSTM) units, have shown great promise in processing temporal sequences of neural activity. Glaser et al. (Glaser et al. 2020) utilized LSTM networks to decode movement intentions from motor cortex recordings in non-human primates, demonstrating the potential of these models for brain-computer interface applications. Extending the application of deep learning to electrophysiological signal processing, Navas-Olive et al. (Navas-Olive et al. 2022) introduced a deep learning-based approach for the detection of sharp wave ripples (SWRs) in hippocampal recordings. Their method demonstrated improved performance over traditional threshold-based approaches, highlighting the potential of deep learning in analyzing complex neural oscillations. These studies collectively underscore the versatility and effectiveness of deep learning techniques in addressing various challenges in electrophysiological research.

Deep learning has also found applications in modeling neural responses and predicting brain activity. Yamins and DiCarlo (Yamins and DiCarlo 2016) showed that deep CNNs trained on object recognition tasks could predict neural responses in the primate visual cortex, providing insights into the hierarchical processing of visual information in the brain. This work exemplifies how deep learning can serve as a bridge between artificial and biological neural networks, offering new perspectives on brain function.

The application of deep learning to neuroscience has not been without challenges. Issues such as the interpretability of complex models, the need for large datasets, and the risk of overfitting are ongoing concerns. However, recent advancements in explainable AI and transfer learning are addressing some of these challenges, making deep learning models more transparent and applicable to smaller datasets typical in neuroscience research.

As deep learning continues to evolve, its applications in neuroscience are likely to expand further. The development of more sophisticated models, coupled with advances in neuroimaging and electrophysiological recording techniques, promises to provide even deeper insights into brain function and dysfunction in the coming years.

3. Specific techniques for neural data analysis

The analysis of neural data presents unique challenges due to its high dimensionality, temporal complexity, and inherent variability. To address these challenges, researchers have developed a variety of specific techniques tailored to the intricacies of neural data analysis. These techniques span a wide range of approaches, from traditional statistical methods to advanced machine learning algorithms, each offering distinct advantages in extracting meaningful information from complex neural signals.

One fundamental technique in neural data analysis is spike sorting, which aims to identify and separate the activity of individual neurons from extracellular recordings. This process is crucial for understanding how information is encoded in neural populations. Traditional spike sorting methods often rely on clustering algorithms applied to features extracted from spike waveforms. However, recent advancements have leveraged deep learning approaches to improve the accuracy and efficiency of spike sorting. For instance, Yger et al. (Yger et al. 2018) introduced a spike sorting algorithm that combines density-based clustering with convolutional neural networks, demonstrating superior performance in handling overlapping spikes and complex noise conditions.

Time-frequency analysis is another essential technique for neural data analysis, particularly for studying oscillatory activity in local field potentials (LFPs). Wavelet transforms have become a popular tool in this domain due to their ability to provide good time and frequency resolution simultaneously. The continuous wavelet transform, using wavelets such as the Morlet wavelet, has been widely employed to analyze the spectral content of neural signals across different time scales. For example, Tallon-Baudry et al. (Tallon-Baudry et al. 1997) used wavelet analysis to investigate induced gamma-band activity in humans, revealing important insights into the temporal dynamics of cognitive processes.

Dimensionality reduction techniques play a crucial role in neural data analysis, given the high-dimensional nature of many neural datasets. Principal Component Analysis (PCA) remains a

widely used method for reducing the dimensionality of neural data while preserving its main features. However, more advanced nonlinear dimensionality reduction methods have gained popularity in recent years. t-Distributed Stochastic Neighbor Embedding (t-SNE), introduced by Maaten and Hinton (Maaten and Hinton 2008), has been particularly useful for visualizing high-dimensional neural data in a lower-dimensional space while preserving local structure. This technique has been applied to various types of neural data, from single-cell gene expression profiles to population-level neural activity patterns.

Decoding algorithms form another important class of techniques in neural data analysis, aimed at inferring external variables (such as stimuli or behaviors) from observed neural activity. Bayesian decoding approaches have been widely used in this context, leveraging probabilistic models to estimate the most likely stimulus or behavior given the observed neural data. Zhang et al. (K. Zhang et al. 1998) pioneered the use of Bayesian decoding for reconstructing animal position from hippocampal place cell activity, a method that has since been refined and applied to various neural decoding problems.

In recent years, deep learning techniques have made significant inroads into neural data analysis. Convolutional Neural Networks (CNNs), originally developed for computer vision tasks, have been adapted for analyzing spatiotemporal neural data. Bashivan et al. (Bashivan et al. 2016) demonstrated the effectiveness of CNNs in classifying cognitive states from EEG data, showcasing the potential of these models in extracting relevant features from complex neural signals. Recurrent Neural Networks (RNNs), particularly those incorporating Long Short-Term Memory (LSTM) units, have shown promise in modeling the temporal dynamics of neural activity. Sussillo et al. (Sussillo et al. 2016) used LSTM networks to model and predict neural population dynamics in motor cortex, providing insights into the underlying computational principles.

Another powerful technique in neural data analysis is Granger causality, which aims to infer causal relationships between different neural signals or brain regions. This method, based on the concept of predictive causality, has been widely used to study functional connectivity in the brain. Seth et al. (Seth, Barrett, and Barnett 2015) provided a comprehensive review of Granger causality applications in neuroscience, highlighting its utility in uncovering directed interactions in neural systems.

Independent Component Analysis (ICA) has emerged as a valuable tool for separating mixed signals in neural data, particularly in the context of EEG and fMRI analysis. ICA can effectively

isolate independent sources of neural activity, helping to remove artifacts and identify functionally distinct neural processes. Makeig et al. (Makeig et al. 2004) demonstrated the power of ICA in decomposing EEG data into independent components, revealing insights into the spatiotemporal dynamics of cognitive processes.

State-space models provide a framework for analyzing neural dynamics as a sequence of latent states evolving over time. These models have been particularly useful for understanding decision-making processes and motor control. Smith and Brown (Smith and Brown 2003) introduced a state-space approach for analyzing neural spiking activity, allowing for the simultaneous estimation of firing rates and underlying state dynamics.

Cross-frequency coupling analysis has gained prominence as a technique for studying interactions between different frequency bands in neural oscillations. This approach has revealed important insights into how information is integrated across different temporal scales in the brain. Tort et al. (Tort et al. 2010) developed methods for quantifying phase-amplitude coupling, demonstrating its relevance in understanding hippocampal-cortical interactions during learning and memory processes.

As the field of neural data analysis continues to evolve, new techniques are constantly being developed and refined. The integration of multiple analysis approaches, often combining traditional statistical methods with advanced machine learning algorithms, is becoming increasingly common. This multi-faceted approach allows researchers to leverage the strengths of different techniques, providing a more comprehensive understanding of complex neural phenomena.

The ongoing development of these analytical techniques is closely tied to advancements in neural recording technologies. As methods for collecting neural data become more sophisticated, capable of recording from larger populations of neurons with higher temporal and spatial resolution, the need for advanced analytical tools grows. This symbiotic relationship between data collection and analysis techniques continues to drive innovation in the field of neuroscience, pushing the boundaries of our understanding of brain function and dysfunction.

4. Advantages and challenges of deep learning for SWR classification

The application of deep learning techniques to the classification of sharp wave ripples (SWRs) represents a significant advancement in the field of neuroscience, offering both substantial advantages and notable challenges. This approach has the potential to greatly enhance our understanding of these critical neural events and their role in memory consolidation and cognitive processes. However, the implementation of deep learning for SWR classification also presents unique obstacles that researchers must carefully navigate.

One of the primary advantages of deep learning for SWR classification is its ability to automatically learn relevant features from raw data. Traditional methods of SWR detection and classification often rely on hand-crafted features and threshold-based approaches, which may not capture the full complexity of these neural events. Deep learning models, particularly convolutional neural networks (CNNs), can learn hierarchical representations directly from the raw electrophysiological signals. This capability allows for the detection of subtle patterns and characteristics that might be overlooked by conventional analysis methods. For instance, Hsu et al. (Hsu et al. 2021) applied convolutional neural networks to classify SWRs before and after learning in wild-type and Alzheimer's disease model mice. Their study not only showcased the ability of deep learning models to distinguish pre- and post-learning SWRs in healthy animals but also revealed impairments in this classification for Alzheimer's disease models. This approach demonstrated the potential of deep learning to uncover subtle learning-related changes in SWR characteristics and their alterations in pathological conditions, highlighting the power of these techniques in advancing our understanding of SWR dynamics in both normal and disease states.

Another significant advantage of deep learning is its potential for improved generalization across different datasets and experimental conditions. SWRs can exhibit considerable variability across subjects, brain states, and recording conditions. Deep learning models, when trained on diverse datasets, have the potential to capture this variability and generalize well to new, unseen data. This generalization capability is particularly important in the context of translational research, where findings from animal models need to be applied to human studies or clinical settings. The ability of deep learning models to transfer knowledge across different domains could facilitate more robust and widely applicable SWR classification methods.

Deep learning approaches also offer the advantage of scalability in handling large volumes of neural data. As recording technologies advance, allowing for simultaneous measurements from an increasing number of neurons and brain regions, the amount of data generated in neuroscience experiments has grown exponentially. Deep learning models are well-suited to process and analyze these large-scale datasets, potentially uncovering complex patterns and relationships that would be difficult or impossible to detect with traditional analysis methods. This scalability is particularly relevant for studying SWRs in the context of broader neural circuits and their interactions with other brain regions.

Furthermore, deep learning models have the potential to integrate multiple data modalities, which is particularly advantageous for comprehensive SWR analysis. For example, a model could simultaneously consider local field potentials, single-unit activity, and even behavioral data to provide a more holistic classification of SWRs. This multi-modal approach could offer insights into the relationship between SWRs and various physiological and behavioral states, enhancing our understanding of their functional significance.

Despite these advantages, the application of deep learning to SWR classification also presents several challenges that researchers must address. One of the primary challenges is the need for large, well-annotated datasets for training deep learning models. SWR detection and classification traditionally rely on expert annotation, which is time-consuming and potentially subject to inter-rater variability. Obtaining sufficiently large datasets with reliable ground truth labels can be a significant hurdle. To address this challenge, researchers have explored various approaches, including the use of semi-supervised learning techniques and data augmentation methods to maximize the utility of limited labeled data.

Another significant challenge is the interpretability of deep learning models. While these models can achieve high performance in SWR classification tasks, understanding the basis of their decisions can be difficult. This "black box" nature of deep learning models can be particularly problematic in neuroscience research, where understanding the underlying mechanisms is often as important as the classification accuracy itself. Recent advances in explainable AI techniques, such as gradient-weighted class activation mapping (Grad-CAM), offer promising avenues for gaining insights into the decision-making processes of deep learning models applied to SWR classification (Selvaraju et al. 2020). However, bridging the gap between model performance and neurophysiological interpretation remains an ongoing challenge.

The risk of overfitting is another concern in applying deep learning to SWR classification. Given the complexity of neural data and the potentially limited size of datasets, deep learning models may learn to fit noise or dataset-specific idiosyncrasies rather than generalizable features of SWRs. This risk is particularly pronounced when working with high-dimensional data from a relatively small number of subjects or recording sessions. Techniques such as regularization, dropout, and careful cross-validation are essential to mitigate overfitting and ensure the robustness of deep learning models in this context.

Additionally, the computational resources required for training and deploying deep learning models can be substantial. This requirement can be a limiting factor, especially for real-time applications or for researchers with limited access to high-performance computing facilities. Balancing model complexity with computational efficiency is an ongoing challenge in the application of deep learning to SWR classification.

Another challenge lies in the integration of domain-specific knowledge into deep learning models. While the ability of these models to learn features automatically is a significant advantage, incorporating prior neurophysiological knowledge about SWRs and hippocampal function into the model architecture or training process can be complex. Striking the right balance between data-driven learning and domain-specific constraints is crucial for developing models that are both powerful and physiologically plausible.

Lastly, the validation of deep learning models for SWR classification in diverse experimental contexts and across species presents a significant challenge. Ensuring that models trained on data from one experimental paradigm or animal model generalize well to others is critical for the broader applicability of these approaches. This challenge is particularly relevant when considering the translation of findings from animal studies to human applications, where differences in neural architecture and recording techniques must be carefully considered.

While deep learning offers powerful advantages for SWR classification, including automatic feature learning, improved generalization, and the ability to handle large-scale, multi-modal data, it also presents significant challenges. These challenges include the need for large, annotated datasets, issues of model interpretability, risks of overfitting, computational demands, integration of domain knowledge, and cross-species validation. Addressing these challenges will be crucial for realizing the full potential of deep learning in advancing our understanding of SWRs and their role in neural information processing and memory consolidation.

5. Self-supervised learning techniques and their potential in neuroscience

Self-supervised learning (SSL) has emerged as a powerful paradigm in machine learning, bridging the gap between unsupervised and supervised learning approaches. This technique has shown remarkable potential in various domains, including computer vision and natural language processing (Kexin Zhang et al. 2024). In recent years, the application of SSL to neuroscience has gained traction, offering promising avenues for analyzing complex neural data and uncovering underlying patterns in brain activity. This section explores the principles of SSL, its applications in neuroscience, and its potential for advancing our understanding of neural processes, particularly in the context of sharp wave ripple (SWR) analysis.

At its core, SSL leverages the inherent structure of unlabeled data to create supervisory signals, allowing models to learn meaningful representations without the need for explicit labels. This approach is particularly valuable in neuroscience, where obtaining large-scale labeled datasets can be challenging and time-consuming. In SSL, the model is trained on a pretext task, which is designed to capture relevant features of the data without requiring manual annotations. Once trained on this pretext task, the model can be fine-tuned for specific downstream tasks with a smaller amount of labeled data.

One of the pioneering applications of SSL in neuroscience was demonstrated by Banville et al. (Banville et al. 2021), who applied contrastive learning, a popular SSL technique, to electroencephalography (EEG) data. Their approach, termed SSL-EEG, involved training a model to distinguish between temporally close EEG segments (positive pairs) and distant segments (negative pairs). This pretext task encouraged the model to learn representations that captured the temporal structure of EEG signals. The authors showed that the learned representations could be effectively used for various downstream tasks, such as sleep stage classification and pathology detection, often outperforming fully supervised approaches trained on smaller labeled datasets.

The potential of SSL in neuroscience extends beyond EEG analysis. In the context of SWR research, SSL techniques could be particularly valuable for addressing the challenges associated with SWR detection and classification. Traditional methods for SWR analysis often rely on expert-defined criteria and manual annotation, which can be subjective and time-consuming. SSL offers the possibility of learning robust representations of SWRs from large

amounts of unlabeled hippocampal recordings, potentially capturing subtle features and variations that might be missed by conventional approaches.

One possible SSL approach for SWR analysis could involve a pretext task based on temporal consistency. For instance, a model could be trained to predict the temporal order of short segments of hippocampal local field potentials (LFPs). This task would encourage the model to learn representations that capture the temporal dynamics of neural activity, including the characteristic patterns associated with SWRs. These learned representations could then be fine-tuned for specific tasks such as SWR detection, classification, or content analysis with a relatively small amount of labeled data.

Another promising SSL technique for neuroscience applications is the masked autoencoder approach, which has shown remarkable success in computer vision (He et al. 2022). Adapted to neural time series data, this method could involve masking portions of the input signal and training a model to reconstruct the masked regions. Such an approach could help the model learn robust representations of the underlying neural dynamics, potentially capturing both the low-frequency components of sharp waves and the high-frequency ripple oscillations.

The application of SSL to multi-modal neuroscience data presents another exciting avenue for research. For instance, Eldele et al. (Eldele et al. 2021) proposed an SSL framework for time series data that combines temporal and contextual contrasting. While their work was not specifically focused on neuroscience, the principles could be readily adapted to neural time series. In the context of SWR research, this approach could be extended to incorporate multiple data streams, such as LFPs from different hippocampal subregions or simultaneous recordings from hippocampus and cortical areas. By learning to contrast and align these different data modalities, the model could potentially uncover complex relationships between SWRs and broader neural circuit dynamics.

One of the key advantages of SSL in neuroscience is its potential to leverage large amounts of unlabeled data, which are often more readily available than expertly annotated datasets. This is particularly relevant for SWR research, where continuous long-term recordings can generate vast amounts of data, but only a small portion may be manually labeled. SSL techniques could allow researchers to make use of these extensive datasets to learn general representations of neural activity, which could then be fine-tuned for specific analyses with smaller amounts of labeled data.

Furthermore, SSL has the potential to address some of the challenges associated with inter-subject and inter-session variability in neural recordings. By learning representations from diverse datasets encompassing multiple subjects and recording sessions, SSL models may develop more robust and generalizable features. This could be particularly valuable for translating SWR analysis methods across different experimental conditions or from animal models to human studies.

Despite its promise, the application of SSL to neuroscience, and specifically to SWR analysis, is not without challenges. One key consideration is the design of appropriate pretext tasks that capture relevant aspects of neural dynamics. Unlike in computer vision or natural language processing, where the structure of the data often suggests intuitive pretext tasks, the complexity of neural signals may require careful consideration to design tasks that lead to meaningful representations.

Another challenge lies in the interpretation of the representations learned through SSL. While these models may achieve high performance on downstream tasks, understanding what specific features or patterns they have captured can be difficult. This challenge is particularly relevant in neuroscience, where relating model behavior to underlying neural mechanisms is often crucial. Developing methods for visualizing and interpreting SSL-derived representations in the context of neural data will be an important area for future research.

The computational demands of training SSL models on large-scale neural datasets also present a practical challenge. However, as high-performance computing resources become more accessible and efficient SSL algorithms are developed, this limitation is likely to become less significant.

Self-supervised learning techniques offer exciting possibilities for advancing neuroscience research, particularly in the domain of SWR analysis. By leveraging large amounts of unlabeled data to learn robust and meaningful representations of neural activity, SSL has the potential to enhance our ability to detect, classify, and interpret complex neural events like SWRs. As these techniques continue to evolve and are increasingly adapted to the specific challenges of neuroscience data, they promise to provide new insights into the functioning of the brain and to drive forward our understanding of fundamental neural processes underlying memory and cognition.

II. Material and methods

A. Experimental design

1. Subjects

The present study employed a cohort of mice comprising two distinct genotypes, six wild-type (WT) and six transgenic (TG) animals. The TG mice were bred as a model for Alzheimer's disease (AD) through crossing APPswe, Tg2576, and PS1dE9 lines, resulting in an accelerated AD model on a mixed B6J/B6SJL background (Lagadec et al. 2012). All mice were aged 8-9 months at the time of experimentation. Genotyping was confirmed via polymerase chain reaction (PCR) of tail biopsy samples. This model was selected for its well-characterized phenotype, exhibiting amyloid pathology and cognitive deficits that closely mimic early-stage AD in humans (Jankowsky et al. 2004).

Animals were individually housed in a controlled environment (temperature: 22 ± 1°C, humidity: 50 ± 10%) under a 12-hour light/dark cycle. Prior to experimentation, mice had ad libitum access to food and water. During the study, food was restricted to maintain body weight at 85% of ad libitum weight, ensuring motivation for the behavioral tasks, while water access remained unrestricted. All experimental procedures adhered to European Guidelines for laboratory animal care (directive 2010/63/UE) and were approved by the University of Bordeaux's ethical committee (protocols A50120159 and A16323) (Jura et al. 2019).

2. Behavioral paradigm

The study employed an elevated eight-arm radial maze (IMETRONIC, Pessac, France) to assess spatial learning and memory (Olton, Collison, and Werz 1977). The maze consisted of a central circular platform (30 cm diameter) with eight radiating arms (50 cm long, 11 cm wide), each equipped with automated sliding doors and food wells. To control for olfactory cues, automatic feeders beneath each arm created a consistent background odor.

The maze was situated in a dimly lit room with various distal visual cues positioned between arms, 1-2 meters from the arm ends, ensuring the task remained hippocampal-dependent. This configuration required mice to integrate multiple spatial cues for accurate navigation, rather than relying on simple cue-following behavior.

The experimental procedure consisted of habituation and learning phases. During the two-day habituation phase, all eight arms were baited. Each day began with a 90-minute recording session in the home cage, followed by maze exposure until all arms were visited and at least one reward consumed, concluding with another 90-minute recording session.

The learning phase spanned six consecutive days. Only three specific arms were baited: two adjacent and one separated by a non-baited arm. This configuration remained constant for each animal throughout the learning period. Daily, animals underwent a 90-minute pre-learning recording in their home cage, followed by six maze trials, and concluded with a 90-minute post-learning recording session. Each trial ended when all rewards were consumed.

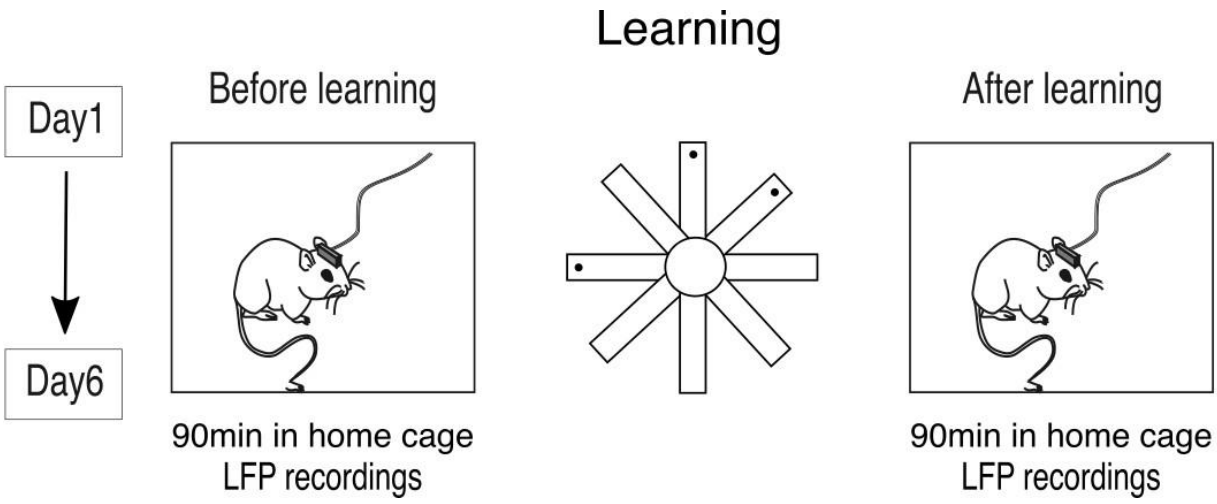

Figure 8 Experimental design.

Depicts the setup for maze learning and electrophysiological recording. Animals were connected to a recording device in their home cage both pre- and post-maze trials. The maze task involved finding three baited arms over six trials, repeated for six days. Electrophysiological data were collected for 90 minutes before and after each learning session.

This comprehensive paradigm, combining a well-established spatial memory task with controlled environmental variables, provided a robust framework for investigating the relationship between sharp wave ripples (SWRs) and memory processes. The design allowed for examination of SWR characteristics changes with learning, and how these changes might differ between WT and AD model mice, offering insights into neural mechanisms underlying spatial memory formation in both normal and pathological conditions.

Figure 8 illustrates the daily experimental protocol during the learning phase, highlighting the pre-learning recordings, maze task, and post-learning recordings. This structured approach ensured consistent data collection across all subjects and experimental days, crucial for subsequent analysis of SWRs in relation to spatial learning and memory consolidation.

B. Electrophysiological recordings

1. Setup

Electrophysiological recordings were obtained from the CA1 area of the hippocampus in both hemispheres. Insulated tungsten wire electrodes (35 µm diameter, California Fine Wires) were implanted using stereotaxic coordinates (AP: +2.0 mm, L: ±1.5 mm, VD: -1.05 mm) under deep isoflurane anesthesia. Reference and ground electrodes were placed in the cerebellum, while an electromyogram (EMG) electrode was inserted into the neck muscles. This setup allowed for precise recording of hippocampal activity while monitoring overall brain states and muscle activity.

The choice of tungsten wire electrodes was based on their excellent signal-to-noise ratio and stability over extended recording periods (Buzsáki, Anastassiou, and Koch 2012). Post-surgery, animals were allowed a 3-4 week recovery period to ensure stable recordings and minimize any potential confounds from the surgical procedure.

2. Signal acquisition

During recording sessions, the mouse head connector was linked to amplifiers via a soft cable, permitting free movement. Neurophysiological and EMG signals were acquired at a sampling rate of 40 kHz using a 128-channel Plexon system. This high sampling rate was chosen to capture the full spectral range of neural activity, including high-frequency oscillations like SWRs. To facilitate further analysis, the data were down-sampled to 2000 Hz using Matlab's 'decimate' function, which applies an anti-aliasing filter before resampling to prevent aliasing artifacts.

Behavioral activity was simultaneously tracked using a video camera, providing context for the recorded neural signals. This multi-modal approach to data acquisition allows for a

comprehensive analysis of neural activity in relation to behavioral states and task performance.

C. Signal processing and SWR detection

1. Filtering

The electrophysiological signals underwent filtering using a 4th-order Chebyshev Type II filter. This filter type was chosen for its steep roll-off characteristics and minimal ripple in the passband, ensuring effective isolation of the frequency bands of interest. For the detection of sharp wave ripples (SWRs), signals were specifically filtered in the 100-250 Hz frequency band. This frequency range has been established in previous studies as optimal for isolating SWR events (Buzsáki 2015). As shown in Figure 9, the raw local field potential (LFP) signal from the CA1 region was filtered to highlight the SWR events. Following filtration, the Hilbert transform was applied to compute the signal envelope, representing the instantaneous amplitude of the filtered signal.

2. SWR identification

SWR events were identified based on specific criteria applied to the z-scored envelope of the filtered signal. The envelopes were z-scored using standard deviation values calculated from slow-wave sleep (SWS) bouts during baseline recording sessions before learning. This approach normalizes the signal relative to the background activity, allowing for consistent SWR detection across different recording sessions and subjects.

Ripple bouts were defined as epochs where the envelope exceeded 2 standard deviations (SDs) of the signal and reached a peak of 5 SDs (Figure 9). The onset and offset of each SWR were determined by the time points at which the envelope crossed the 2 SD threshold.

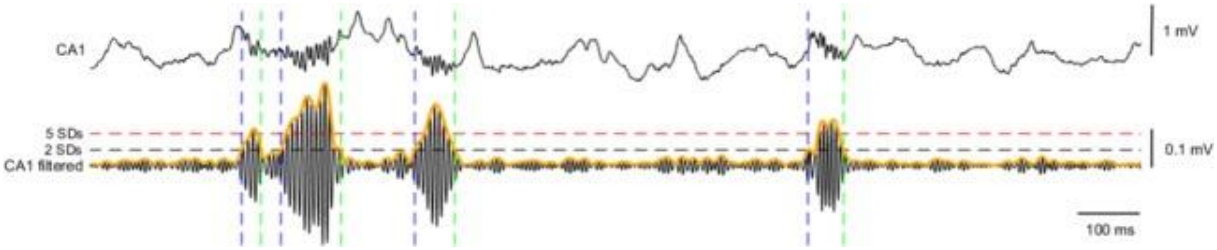

Figure 9 Signal processing and SWR detection.

Upper trace: Raw local field potential (LFP) recorded from the CA1 region of the hippocampus. Lower trace: Filtered LFP signal (100-250 Hz) highlighting sharp wave ripple (SWR) events. The yellow shaded areas represent the signal envelope derived using the Hilbert transform. Dashed lines indicate the thresholds used for SWR detection: 2 standard deviations (SDs) for onset/offset (gray) and 5 SDs for the peak (red). Vertical lines (blue and green) mark the onset and offset of detected SWR events. Adapted from (Jura et al. 2019), distributed under Creative Commons Attribution license.

To account for closely spaced events and exclude unusually long oscillations, episodes spaced less than 20 ms apart were merged, while those longer than 100 ms were discarded. This approach ensured the capture of genuine SWR events while minimizing false positives and the inclusion of other high-frequency oscillations that may not be true SWRs.

Additionally, the intrinsic frequency of each SWR was calculated using the Hilbert transform of the signal, specifically as the time derivative of the instantaneous phase of the transformed signal.

3. Ripple-Centered Intervals (RCIs)

Following the methodology outlined by Hsu et al. (2021), the local field potential (LFP) signals were segmented into non-overlapping Ripple-Centered Intervals (RCIs) of 256 milliseconds. This segmentation process was crucial for preparing the data for subsequent deep learning analysis. Each RCI was centered on a detected SWR event, ensuring that the full temporal extent of the SWR and its immediate context were captured (Figure 10).

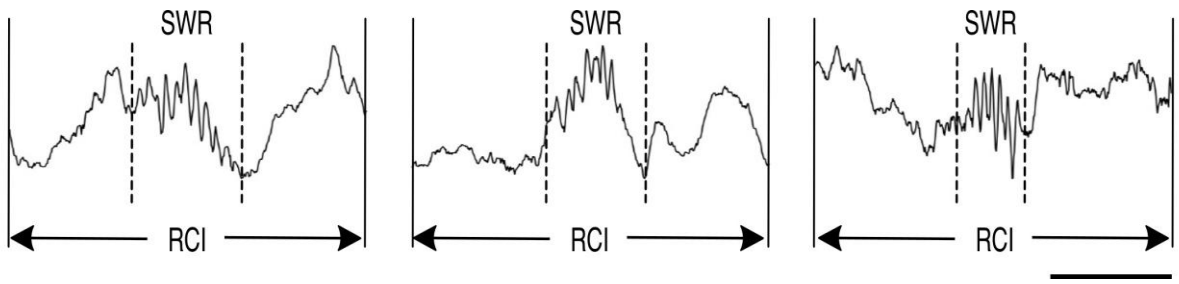

Figure 10 Examples of Ripple-Centered Intervals (RCIs).

Each panel shows a segment of the local field potential (LFP) signal centered on a detected sharp wave ripple (SWR) event. The black dashed lines mark the onset and offset of the SWR, and the arrows indicate the 256 ms duration of the RCI. This segmentation method ensures that the entire SWR event and its immediate neural context are captured for subsequent analysis.

The use of non-overlapping intervals prevents artificial correlations that could arise from data overlap, ensuring the independence of each sample. This is particularly important for the subsequent machine learning analyses, as it helps to prevent overfitting and ensures that the model learns genuine patterns in the data.

The choice of a 256 ms interval length was carefully considered to balance multiple factors. This duration is sufficiently long to encompass the entire SWR event, which typically lasts 50-100 ms, while also capturing pre- and post-ripple neural activity. This additional context is crucial for understanding the broader neural dynamics associated with SWR generation and propagation, including potential preparatory activity before the SWR and immediate consequences in local network activity.

Furthermore, the 256 ms length, being a power of 2, is computationally efficient for subsequent frequency domain analyses. It is also compatible with many deep learning architectures, allowing for efficient processing and analysis (Hsu et al., 2021).

D. Synthetic dataset creation (SWR_{art})

To rigorously evaluate our classification algorithms and assess the impact of label noise, we developed a synthetic dataset of artificial sharp wave-ripples (SWR_{art}). This approach allowed for precise control over signal characteristics and provided a well-defined ground truth for algorithm validation. Each SWRart was modeled as a Gaussian-modulated sinusoidal wave with frequencies ranging from 120 to 250 Hz, consistent with the spectral properties of

biological SWRs observed in vivo. The duration of each SWR_{art} was randomized following a normal distribution between 100 and 200 data points, introducing variability that mimics the natural heterogeneity of these events (Figure 11).

To enhance biological realism, we incorporated additional oscillatory components characteristic of hippocampal activity during slow-wave sleep. Specifically, delta (1-3 Hz) and low gamma (20-40 Hz) oscillations were added, with frequencies randomly selected from their respective ranges using a normal distribution (Oliva et al. 2018). Gaussian white noise was also introduced to both the SWR_{art} and the entire ripple-centered interval (RCI) to simulate background neural activity. This multi-component approach ensured that our synthetic data closely approximated the complexity of real neurophysiological recordings.

We generated a total of 20,000 SWR_{art} instances, equally divided into two classes: Class 1 representing SWRs not affected by learning, and Class 2 representing SWRs transformed by learning. The key distinction between these classes lay in the variability of their parameters, reflecting hypothesized changes in signal coherence following learning (Jura et al. 2019).

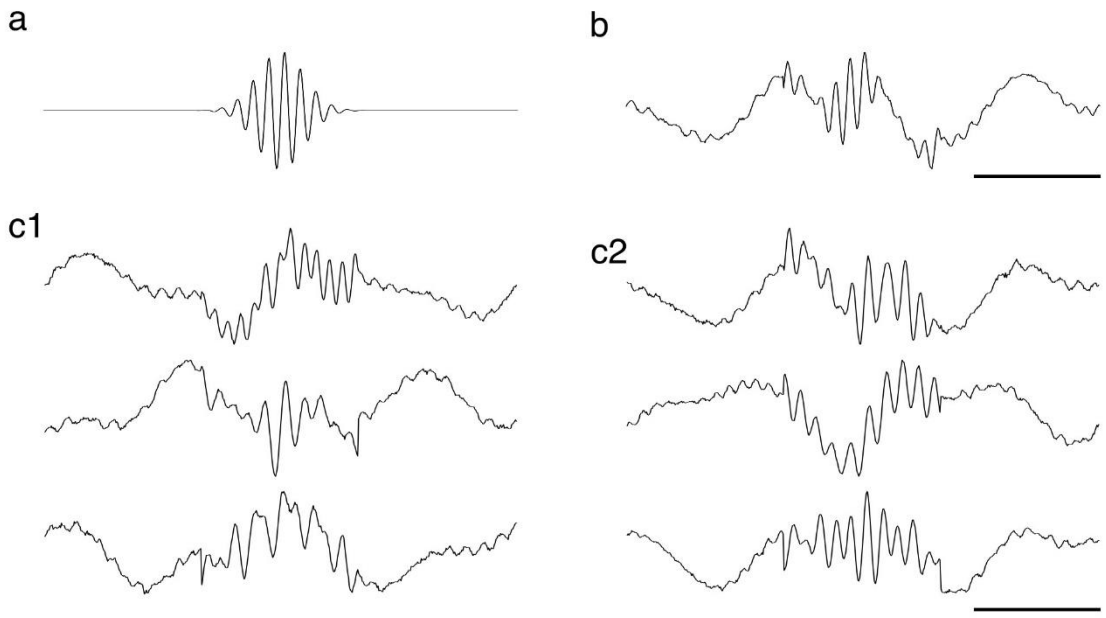

Figure 11 Synthetic SWR_{art} dataset.

(a) Example of a synthetic sharp wave ripple (SWR_{art}) modeled as a Gaussian-modulated sinusoidal wave. (b) SWR_{art} with added delta (1-3 Hz) and low gamma (20-40 Hz) oscillations to enhance biological realism. (c1) Examples of Class 1 SWR_{art} instances, characterized by greater variability in frequency and amplitude, and added Gaussian white noise to simulate background neural activity. (c2) Examples of Class 2 SWR_{art} instances, reflecting a post-learning state with narrower parameter ranges and lower variability in frequency and amplitude.

Class 1 SWR_{art} instances were designed with greater variability in frequency and amplitude. The frequency was modulated to span 0.8 to 1.2 times the base frequency of 160 Hz, following a normal distribution. The amplitude similarly varied between 0.8 to 1.2 times a standardized base amplitude. The noise level incorporated Gaussian white noise with a mean of zero and a standard deviation of 0.1, derived empirically from biological SWR recordings (Table 1).

Multiplication factors	Frequency	Amplitude	Noise
Class 1	0.8–1.2	0.8–1.2	1–1.5
Class 2	1.1–1.3	0.95–1.05	0.5–1

Table 1: *Multiplication factors of the frequency, amplitude, and noise level used for construction of SWR_{art}. Base frequency: 160 Hz; Base amplitude: a range between -1 and 1; Base noise level (standard deviation): 0.1.*

In contrast, Class 2 SWR_{art} instances were constructed with narrower parameter ranges, reflecting the hypothesized post-learning state. The frequency ranged from 1.1 to 1.3 times the base frequency, while the amplitude varied from 0.95 to 1.05 times the base amplitude. The noise level standard deviation ranged from 0.5 to 1 times the base noise level (Table 1).
This carefully constructed synthetic dataset provided a controlled environment for testing our classification algorithms, allowing us to systematically vary the level of label noise and assess its impact on classification performance.

E. Deep learning models selection

In our comprehensive approach to classifying sharp wave ripples (SWRs), we carefully selected a diverse array of deep learning models for benchmarking. This selection process was guided by the unique characteristics of our time-series data and the specific requirements of our classification task. Our aim was to evaluate both one-dimensional (1D) and two-dimensional (2D) models to thoroughly assess different approaches to SWR classification.
For processing the raw temporal sequences of SWRs directly, we employed 1D models, specifically Long Short-Term Memory (LSTM) networks and 1D Convolutional Neural Networks

(CNNs). LSTMs, as described by Hochreiter (Hochreiter and Schmidhuber 1997), are particularly well-suited for capturing long-term dependencies in sequential data, which is crucial for understanding the temporal dynamics of SWRs. These networks have demonstrated remarkable success in various time-series analysis tasks, making them a natural choice for our study. The ability of LSTMs to selectively remember or forget information over long sequences aligns well with the complex temporal patterns present in SWR data.

1D CNNs, on the other hand, have shown exceptional performance in various time-series classification tasks, as demonstrated by Kiranyaz et al. (Kiranyaz et al. 2021) in their comprehensive review. These models excel at extracting local patterns and features from sequential data, which is particularly relevant for identifying the distinctive spectral and temporal characteristics of SWRs. The hierarchical feature learning capability of CNNs allows them to capture both low-level and high-level features of the SWR signals, potentially revealing nuanced differences between pre- and post-learning states.

To explore alternative representations of our data, we also investigated 2D models by converting the SWR temporal sequences into image-like inputs. This approach allowed us to leverage powerful architectures from the field of computer vision, which have demonstrated remarkable success in image classification tasks. We selected MobileNet (Howard et al. 2017) for its efficiency. The lightweight nature of MobileNet, achieved through the use of depthwise separable convolutions, makes it an attractive option for potential real-time SWR classification. EfficientNet (Tan and Le 2020) was chosen for its balanced trade-off between accuracy and computational cost. The compound scaling method employed by EfficientNet, which uniformly scales network width, depth, and resolution, offers the potential for highly efficient SWR classification. This model's ability to achieve state-of-the-art accuracy with significantly fewer parameters than traditional CNNs made it an intriguing candidate for our benchmarking process.

We included VGG (Simonyan and Zisserman 2015) in our selection for its simplicity and effectiveness. Despite being an older architecture, VGG's straightforward design of stacked convolutional layers followed by max-pooling has proven to be remarkably robust across various domains. Its inclusion in our benchmarking allows us to compare more recent architectures against this well-established baseline.

Lastly, we incorporated ConvNeXt (Liu et al. 2022) as a state-of-the-art model that bridges the gap between CNNs and Transformers. ConvNeXt's design, which modernizes the standard

ResNet architecture with techniques borrowed from vision Transformers, offers a fresh perspective on convolutional architectures. Its potential for capturing both local and global dependencies in the data made it an exciting addition to our benchmarking suite.

In addition to these models, we incorporated autoencoders into our approach to address the challenge of label noise in our dataset. Autoencoders, as described by Vincent et al. (Vincent et al. 2010), are unsupervised learning models that can learn compact representations of input data. We utilized them to detect and potentially correct mislabeled instances by comparing reconstruction errors against a defined threshold.

To further enhance our ability to handle label noise, we implemented genetic algorithms, following the approach outlined by Srinivas and Patnaik (Srinivas and Patnaik 1994). These evolutionary algorithms were employed to optimize the labeling configurations, evolving solutions based on fitness criteria designed to maximize classification accuracy.

Recognizing the potential of self-supervised learning in scenarios with limited labeled data, we also incorporated this approach into our methodology. Inspired by the work of Eldele et al. (Eldele et al. 2021) on time-series representation learning, we developed a self-supervised pretraining stage. This stage allowed our models to learn meaningful features from unlabeled SWR data before fine-tuning on the labeled pre- and post-learning dataset, potentially improving the robustness and generalizability of our classifications.

This comprehensive selection of models and techniques enabled us to not only evaluate different architectural approaches for SWR classification but also to address the critical issue of label noise in our dataset. By incorporating autoencoders, genetic algorithms, and self-supervised learning alongside traditional supervised models, we aimed to develop a robust and accurate classification framework capable of handling the complexities and potential inconsistencies in neurophysiological data.

F. 1D CNN Model implementation

1. Model architecture

Our primary model for SWR classification was a custom-designed 1D CNN, optimized for processing the temporal sequences of ripple-centered intervals (RCIs). Throughout this text, the terms SWRs and their corresponding segmented intervals, RCIs, are used interchangeably. This is because our analytical focus is on RCIs, segments centered around SWRs. The

architecture of this model was carefully crafted to balance the trade-off between model complexity and performance, taking inspiration from successful applications of 1D CNNs in time-series analysis (Kiranyaz et al. 2021).

The mathematical formulation for our CNN model (Figure 12) can be described as follows: Let x be the input RCI and y be the output prediction vector indicating the class membership ("before learning" or "after learning"). The model aims to minimize the loss function $L(y, \tilde{y})$, where $\tilde{y}$ is the ground truth, by adjusting the weight vector w using Stochastic Gradient Descent (SGD) (Ruder 2017):

$$w(t+1) = w(t) - \alpha \nabla_w L(y(t), \tilde{y}(t))$$

where α is the learning rate and ∇_w is the gradient of the loss function with respect to the weights.

The input layer of our 1D CNN accepts RCIs of 512 data points, corresponding to 256 ms of recorded neural activity. This temporal window was chosen to encompass the full duration of typical SWR events, which usually last between 50-100 ms, while also capturing the immediate pre- and post-ripple neural context. The inclusion of this surrounding activity is crucial for understanding the broader neural dynamics associated with SWR generation and propagation.

The first convolutional layer consists of 32 filters with a kernel size of 3 and uses Rectified Linear Unit (ReLU) activation. This initial layer is designed to capture low-level features in the SWR signals, such as local fluctuations and short-term patterns. The relatively small kernel size allows the model to detect fine-grained temporal features that may be characteristic of SWRs. The convolution operation in the k^{th} layer is given by:

$$y_k = \sum_m x_{k+m} w_m$$

Here, x_{k+m} are the input values at positions $k+m$, where k is the current position in the input sequence and m is the relative position within the filter window, while w_m are the filter weights learned during training.

Following the principles of hierarchical feature learning in CNNs, the subsequent convolutional layers progressively increase the number of filters. The second layer uses 64 filters, while the third layer employs 128 filters. This expansion in the number of filters allows the network to learn increasingly complex and abstract representations of the data as it progresses through

the layers. The deeper layers can potentially capture more sophisticated patterns, such as the specific frequency components of SWRs or their temporal evolution over the course of the RCI. After each convolutional layer, we incorporate a batch normalization layer. This technique, introduced by Ioffe and Szegedy (Ioffe and Szegedy 2015), normalizes the inputs to each layer, which has several beneficial effects on the training process. Batch normalization helps to mitigate the internal covariate shift problem, where the distribution of each layer's inputs changes during training as the parameters of the previous layers are updated. By stabilizing these distributions, batch normalization allows for higher learning rates, potentially leading to faster convergence. Additionally, it provides a regularizing effect that can help prevent overfitting.

Following batch normalization, we apply the ReLU activation function. The choice of ReLU is motivated by its effectiveness in deep neural networks, as demonstrated by Nair and Hinton (Nair and Hinton, 2010). ReLU introduces non-linearity into the model, allowing it to learn complex, non-linear mappings from the input space to the output space. Moreover, ReLU activations help mitigate the vanishing gradient problem that can occur in deep networks with other activation functions, facilitating the training of deeper architectures.

After each set of convolution, batch normalization, and ReLU activation (Figure 12), we implement a max pooling layer with a pool size of 2. These max pooling operations serve multiple purposes in our architecture. They reduce the spatial dimensions of the feature maps, decreasing the computational load and the number of parameters in the model, which helps to prevent overfitting. Max pooling also introduces a degree of translational invariance to the features learned by the network, allowing the model to detect key features regardless of their exact temporal position within the RCI. Furthermore, by selecting the maximum value in each pooling window, max pooling helps the network focus on the most salient features, promoting a hierarchical representation of the data. The max-pooling operation is defined as:

$$y_k = max\{x_{2k-1}, x_{2k}\}$$

In the first convolutional block, we include a dropout layer following the max pooling operation. Dropout, as introduced by Srivastava et al. (Srivastava et al. 2014), is a powerful regularization technique that helps prevent overfitting by randomly setting a fraction of input units to 0 during training. In our model, we employ a dropout rate of 0.2, meaning that 20%

of the units are randomly dropped during each training iteration. The inclusion of dropout in the early stages of the network serves to create an ensemble effect, where different subsets of the network are trained on each batch. This encourages the network to learn more robust and generalizable features, reducing its reliance on any single set of neurons.

After the series of convolutional, batch normalization, ReLU activation, and max pooling operations, we incorporate a flatten layer. This layer transforms the 3D output of the convolutional layers (channels × time steps × features) into a 1D vector. This flattening operation is necessary to transition from the convolutional part of the network to the fully connected layers that follow.

The flattened output is then fed into a linear (fully connected) layer. This layer allows the model to learn complex combinations of the high-level features extracted by the convolutional layers. The linear layer performs a weighted sum of all inputs, followed by a bias term, enabling it to capture intricate relationships between features that may be crucial for distinguishing pre- and post-learning SWRs. The number of neurons in this linear layer was determined through our hyperparameter optimization process, balancing the model's capacity to learn complex representations with the risk of overfitting.

The final component of our architecture is the output layer, which consists of a linear layer with two neurons, corresponding to our binary classification task (pre-learning vs. post-learning SWRs). The activations of these neurons are passed through a softmax function, which converts the raw scores into a probability distribution over the two classes. The softmax function ensures that the output probabilities sum to 1, providing a natural interpretation of the model's confidence in its classification decision. The softmax function is defined as:

$$\text{softmax}\,(y_k) = \frac{e^{y_k}}{\sum_i e^{y_i}}$$

where y_i are the raw scores (logits) produced by the final layer.

This carefully designed architecture, combining convolutional layers for feature extraction, regularization techniques to prevent overfitting, and fully connected layers for final classification, provides a powerful framework for distinguishing between pre- and post-learning SWRs.

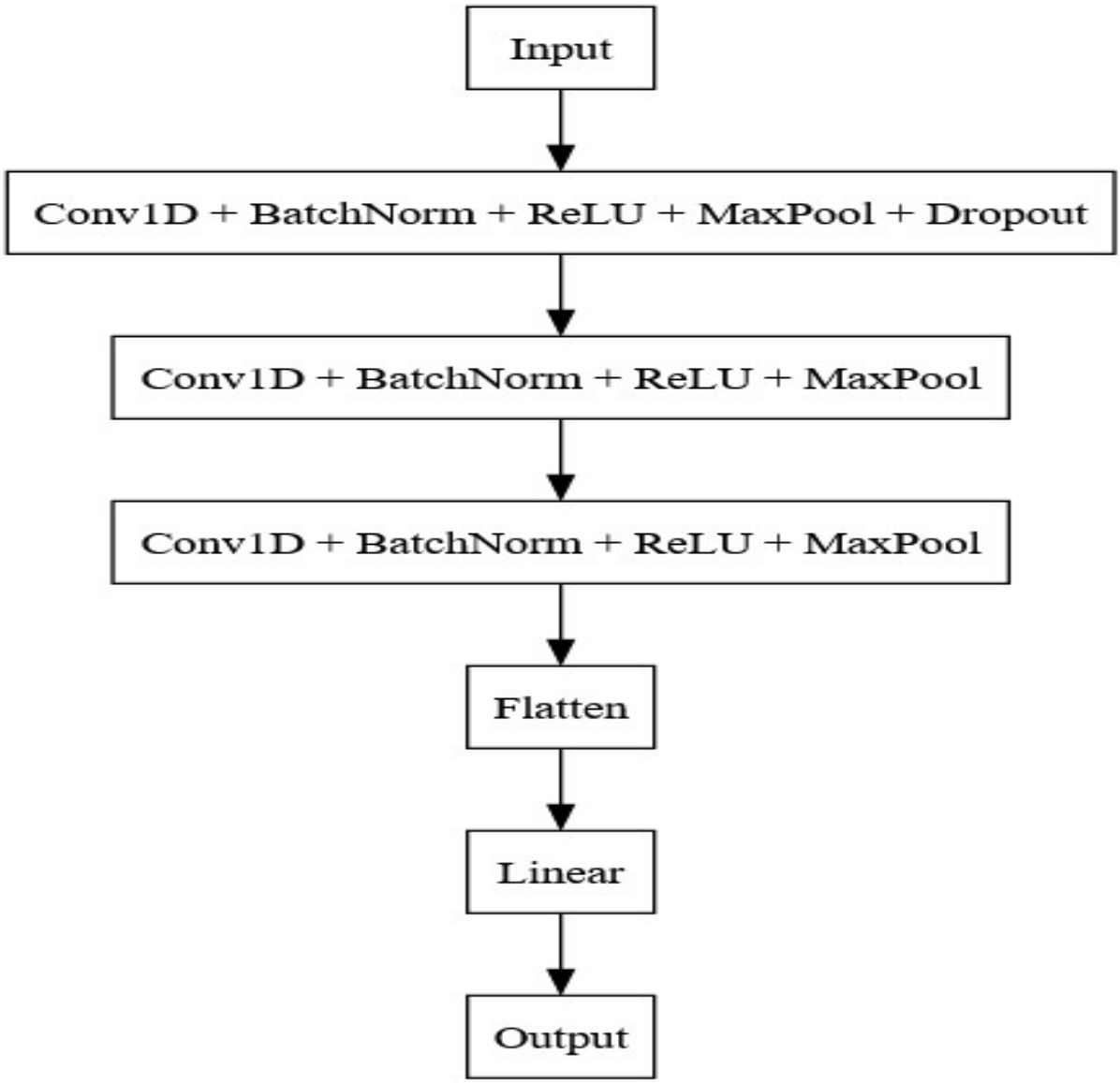

Figure 12 Architecture of the 1D Convolutional Neural Network (CNN) model.

Architecture of the 1D CNN used in our study for classifying sharp wave-ripples (SWRs). It consists of multiple layers, including convolutional layers with batch normalization and ReLU activation, max pooling layers, dropout, a flatten layer, a fully connected layer, and an output layer with softmax activation for binary classification.

2. Hyperparameter optimization

To maximize the performance of our 1D CNN model, we conducted a thorough hyperparameter optimization process. We employed a combination of grid search and random search techniques (T. Yu and Zhu 2020), to efficiently explore the hyperparameter space.

The key hyperparameters we optimized included the number of convolutional layers (ranging from 2 to 5), the number of filters in each convolutional layer (ranging from 32 to 512), the

kernel size for convolutional layers (ranging from 3 to 5), the number of dense layers (ranging from 1 to 3), the number of neurons in dense layers (ranging from 64 to 512), the dropout rate (ranging from 0.2 to 0.5), the learning rate (ranging from 1e-4 to 1e-2), and the batch size (ranging from 16 to 128).

We used 5-fold cross-validation to evaluate the performance of each hyperparameter configuration, ensuring robustness of our results. The primary metric for optimization was the F1 score, which provides a balanced measure of precision and recall.

The grid search allowed us to systematically explore the hyperparameter space for key parameters such as the number of convolutional layers and the number of filters. This approach ensured that we thoroughly investigated the impact of these critical architectural choices on model performance. Complementing this, we used random search for continuous parameters like learning rate and dropout rate, as well as for exploring less critical discrete parameters. Random search has been shown to be more efficient than grid search for high-dimensional spaces, allowing us to cover a wider range of possible configurations with fewer trials.

During the optimization process, we also monitored secondary metrics such as accuracy, precision, recall, and area under the ROC curve (AUC-ROC) to ensure a comprehensive evaluation of model performance. We paid particular attention to the model's generalization ability, carefully monitoring the gap between training and validation performance to avoid overfitting.

The final hyperparameter configuration was selected based on the best average F1 score across the cross-validation folds. However, we also considered the stability of performance across folds and the model's performance on secondary metrics in making our final selection. The optimal hyperparameters were as follows: Number of convolutional layers: 4, Number of filters in convolutional layers: [32, 64, 128, 256], Kernel size for convolutional layers: 3, Number of dense layers: 2, Number of neurons in dense layers: [256, 128], Dropout rate: 0.3, Learning rate: 0.001, Batch size: 128. This comprehensive approach to hyperparameter optimization resulted in a model architecture that demonstrated superior and robust performance in classifying SWRs into before and after learning categories.

G. Self-Supervised Learning (SSL) implementation

Our study leveraged the Time-Series Representation Learning via Temporal and Contextual Contrasting (TS-TCC) framework proposed by Eldele et al. (Eldele et al. 2021) to enhance the performance of our sharp wave ripple (SWR) classification model. This framework is particularly well-suited for capturing the temporal relationships inherent in the sequential nature of our neural recording data from the maze experiment. The TS-TCC framework addresses the challenges associated with labeled time-series data, where obtaining accurate labels can be time-consuming and costly.

The framework comprises two primary components: a temporal contrasting module and a contextual contrasting module. These modules work synergistically to extract robust and discriminative features from the unlabeled SWR data. By combining these components, the framework aims to learn representations that are both temporally coherent and contextually meaningful, which is crucial for accurately classifying SWRs before and after learning events.

1. Contrastive learning mechanism

At the core of our SSL model is the contrastive learning paradigm, which aims to learn representations by contrasting positive and negative examples. This approach is particularly effective for self-supervised learning (Kexin Zhang et al. 2024), where explicit labels are not used. Instead, the model leverages the inherent structure of the data to learn meaningful representations.

The contrastive learning mechanism operates on three main types of examples: the anchor, which serves as the reference point in the feature space; positive examples, which are similar to the anchor and should be close in the latent space; and negative examples, which are dissimilar to the anchor and should be farther away in the latent space.

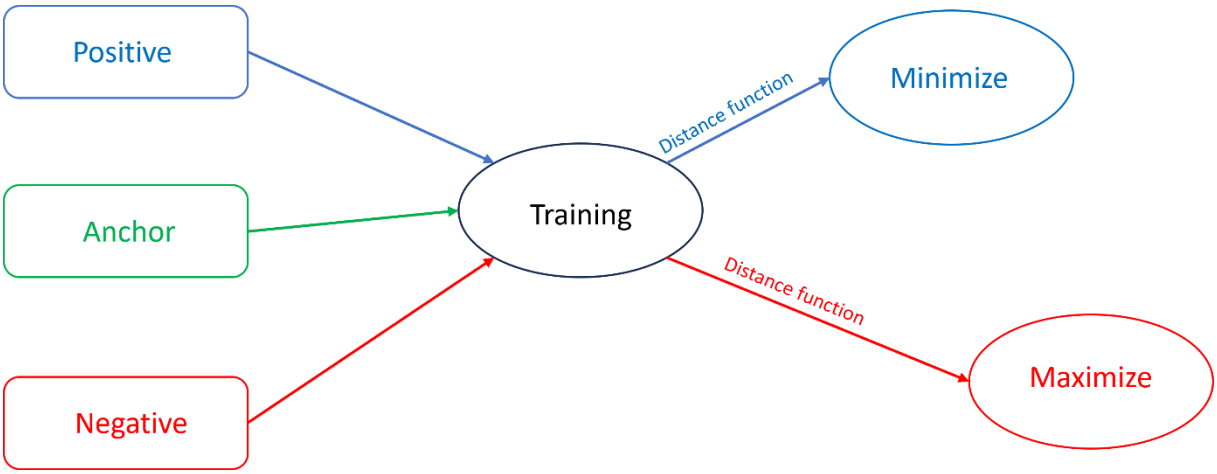

Figure 13 Contrastive learning mechanism.

The anchor represents a base data sample from which similarities or dissimilarities are calculated. The positive example is a data point similar to the anchor, belonging to the same class, while the negative example belongs to a different class. The objective of the training is to minimize the distance between the anchor and the positive example while maximizing the distance between the anchor and the negative example. This is achieved through a loss function specifically designed for contrastive learning.

Figure **13** illustrates this mechanism, depicting how the training objective aims to minimize the distance between the anchor and positive examples while maximizing the distance to negative examples. This approach is implemented in both the Temporal and Contextual Contrasting modules of our architecture.

2. Data augmentation

The first stage of our SSL model involves data augmentation to produce two different views of each SWR sample. Data augmentation is a critical component of our framework, as it allows the model to learn invariant representations by exposing it to various transformed versions of the same input. We employ two distinct augmentation techniques: weak augmentation and strong augmentation.

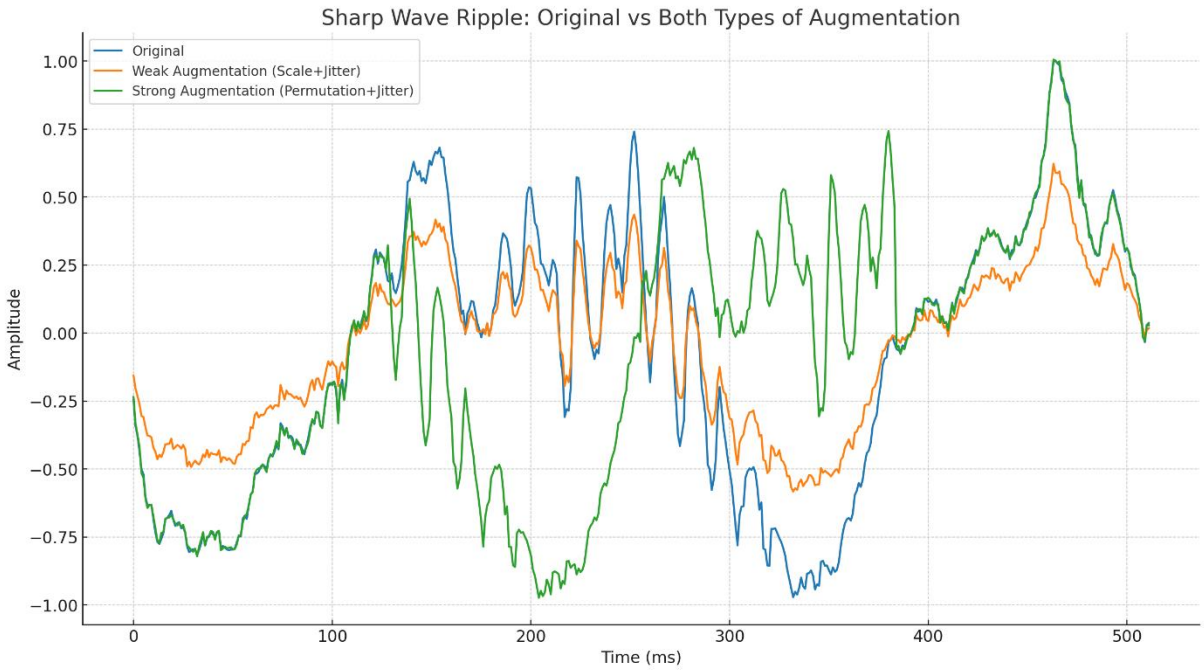

*Figure **14 Data augmentation.***

Examples of data augmentation applied to a single sharp wave-ripple (SWR). Two types of augmentations are shown: weak and strong. Weak augmentation involves a scale and jitter modification to the original SWR, subtly altering its temporal properties. Strong augmentation is more extensive, applying both permutation and jittering to transform the SWR significantly.

Weak augmentation, denoted as x_w, applies a jitter-and-scale strategy, subtly altering the temporal properties of the SWR. This involves adding random variations to the signal and scaling up its magnitude. The jittering helps the model learn robustness to small temporal shifts, while scaling aids in learning amplitude-invariant features.

Strong augmentation, denoted as x_s, implements a more extensive permutation-and-jitter strategy, significantly transforming the SWR. This involves splitting the signal into a random number of segments (with a maximum of M segments) and randomly shuffling them, followed by the addition of random jitter. The permutation encourages the model to learn features that are invariant to the specific ordering of sub-sequences within an SWR, while still maintaining the overall temporal structure.

Figure 14 provides examples of these augmentation techniques applied to a single SWR. This augmentation strategy is crucial for creating diverse yet correlated views of the SWR data, enabling the model to learn invariant representations.

The choice of augmentation techniques and their parameters is critical and depends on the nature of the time-series data.

3. Encoder architecture

Following augmentation, the data is passed through an encoder composed of a 3-block convolutional architecture. This encoder maps each input x into a high-dimensional latent space $z = [z_1, z_2, \dots z_T]$, where T is the total timesteps and d is the feature length. The same encoder is used for both strongly and weakly augmented views, producing corresponding latent features z_s and z_w.

The encoder plays a crucial role in transforming the raw time-series data into a high-dimensional feature space where meaningful patterns can be more easily discerned. The use of convolutional layers allows the encoder to capture local patterns and hierarchical features in the time-series data, which is particularly important for SWRs given their complex temporal structure.

4. Temporal contrasting module

The Temporal Contrasting module is a key component of our SSL framework, designed to capture the intricate temporal dynamics inherent in SWR signals. This module employs an autoregressive model, specifically a Transformer architecture (Vaswani et al. 2017), to summarize past latent features into a context vector c_t.

A key innovation in this module is the implementation of a cross-view prediction task. In this task, the context derived from one augmented view of the SWR data is used to predict future timesteps of the other augmented view. Specifically, the context from the strong augmentation $c_{s,t}$ is used to predict the future latent features of the weak augmentation $z_{w,t+k}$ and vice versa. This approach forces the model to learn representations that are robust against perturbations introduced by different augmentations and timesteps.

Two loss functions, $L_{s,TC}$ and $L_{w,TC}$, are calculated to serve this cross-view prediction objective. These loss functions aim to minimize the dot product between the predicted representation and the true one of the same sample, while maximizing the dot product with the other samples within the mini-batch.

The use of a Transformer as the autoregressive model is motivated by its efficiency and ability to capture long-range dependencies in sequential data. The Transformer architecture consists of multi-headed attention mechanisms followed by feed-forward networks, allowing it to effectively model complex temporal relationships in the SWR data.

5. Contextual contrasting module

Building upon the Temporal Contrasting module, the Contextual Contrasting module aims to learn more discriminative representations. This module applies a non-linear transformation to the contexts produced by the temporal module, mapping them into a space where contrastive learning is applied.

The objective of this module is to maximize the similarity between contexts derived from different augmented views of the same sample (positive pairs) while minimizing similarity with contexts from different samples (negative pairs). This ensures that the model learns not only temporal features but also the specific context in which these features are most informative.

The loss function for this module encourages the model to learn representations that are discriminative across different samples while being consistent across different augmented views of the same sample. It achieves this by using a cosine similarity measure and a temperature parameter to control the concentration of the distribution of similarities.

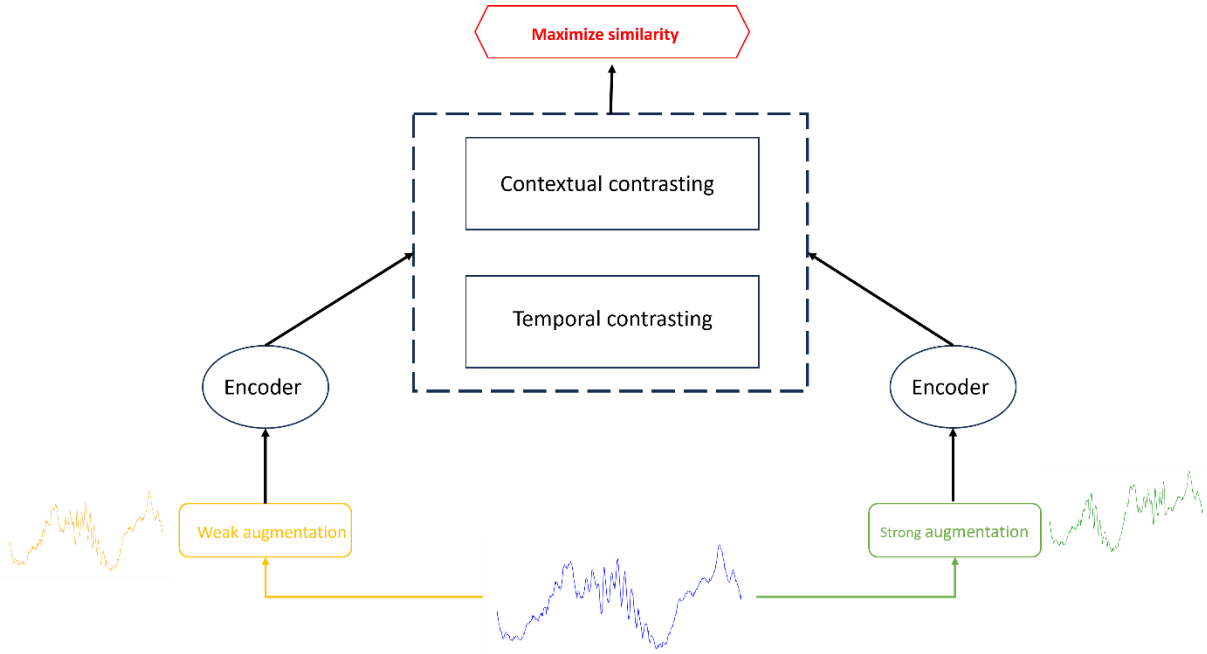

Figure 15 Self-supervised model architecture.

The self-supervised learning model architecture consists of four core components: Data Augmentation, Encoder, Temporal Contrasting, and Contextual Contrasting. Initially, Data Augmentation techniques are applied to raw input. Next, the augmented data passes through an Encoder to generate feature vectors. These vectors are then subjected to Temporal Contrasting, which aligns temporally close features. Finally, Contextual Contrasting is used to discriminate the feature vectors from varying contexts. Together, these elements work in harmony to enable effective feature learning.

6. Final architecture

Figure 15 provides a detailed schematic of our complete SSL model architecture, illustrating the seamless integration of these components: Data Augmentation, Encoder, Temporal Contrasting, and Contextual Contrasting. This architecture is designed to capture the most informative features of SWRs both before and after learning in the maze experiment, enhancing the robustness and generalizability of the model. The combination of temporal and contextual contrasting allows the model to learn representations that are both temporally coherent and contextually meaningful, which is crucial for accurately classifying SWRs in different learning states.

7. Training process

We trained the SSL model for 500 epochs on our unlabeled SWR or SWR_{art} datasets, which included instances from both before and after learning phases. The use of a large number of epochs allows the model to fully exploit the unlabeled data and learn rich, generalizable representations.

The primary objective during training was to minimize the combined temporal and contextual contrastive loss. This combined loss function balances the importance of temporal features and contextual information, allowing the model to learn comprehensive representations of the SWR data.

We used the Adam optimizer with carefully tuned learning rate and weight decay parameters to prevent overfitting. The batch size was adjusted based on the amount of available data.

The Transformer used in the temporal contrasting module was configured with multiple layers and attention heads, with the specific architecture determined through extensive experimentation and validation. We also implemented dropout in the Transformer to further prevent overfitting and improve generalization.

By incorporating these advanced self-supervised techniques into our SWR classification framework, we aimed to address the unique challenges posed by complex neural data and label noise.

8. Re-labeling process

Our SSL model was employed to re-label the SWR or SWR_{art} datasets in an unsupervised manner. Following the training period, the SSL model utilized its learned representations to generate new labels for the dataset. This re-labeling process resulted in the creation of two new groups: Group 1 (G1) and Group 2 (G2) or $G1_{art}$ and $G2_{art}$. The assignment of SWRs to these new groups was based on the proximity of features in the learned representation space, as evaluated by the SSL model (Figure 16).

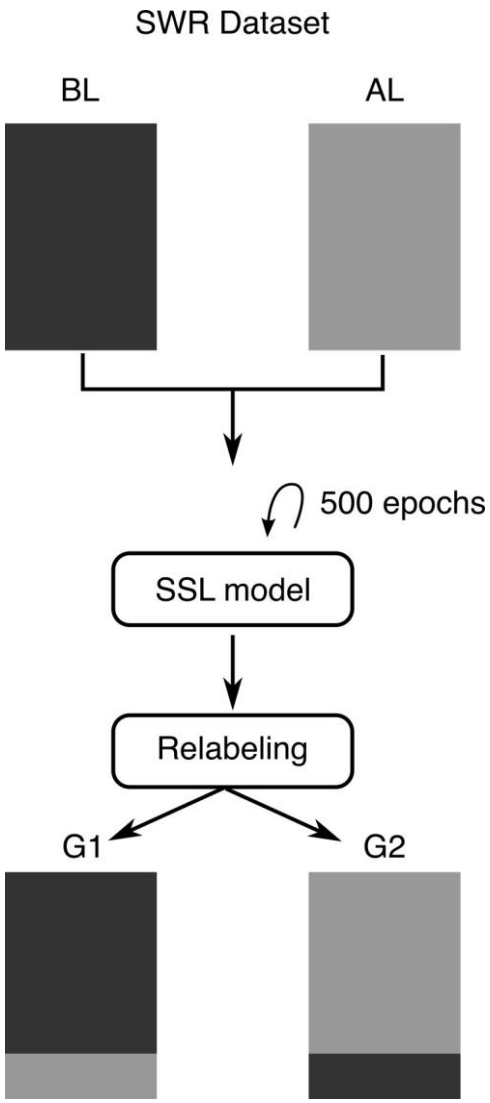

Figure 16 SSL re-labeling.

The SSL model re-labels the SWR dataset, categorizing before-learning (BL) and after-learning (AL) instances into new groupings, Group 1 (G1) and Group 2 (G2), based on similarities in features.

This approach aimed to uncover potentially more nuanced groupings within the data that may not have been apparent in the original categorization. The re-labeling process effectively redistributed the SWRs into the new groups, potentially capturing subtle differences in SWR characteristics that could be associated with learning-induced changes.

By leveraging the SSL model's ability to learn from unlabeled data, this re-labeling strategy sought to provide a more refined classification of SWRs that could lead to improved insights into the neural processes underlying learning and memory consolidation.

H. Model evaluation and performance metrics

1. Cross-validation strategy

To ensure the robustness and generalizability of our deep learning model for classifying sharp wave ripples (SWRs), we implemented a rigorous 5-fold cross-validation strategy. This approach, widely recognized in machine learning literature for its ability to provide a comprehensive assessment of model performance (Kohavi 1995), was particularly crucial given the complex nature of neurophysiological data and the potential for overfitting.

Our dataset, comprising Ripple-Centered Intervals (RCIs) from both pre- and post-learning sessions, was first pooled and then randomly divided into five equal subsets or folds. This process ensured that each fold contained a representative distribution of SWRs from both learning conditions. In each iteration of the cross-validation process, four folds were used for training the model, while the remaining fold served as the validation set. This procedure was repeated five times, with each fold serving as the validation set exactly once, as illustrated in Figure 17.

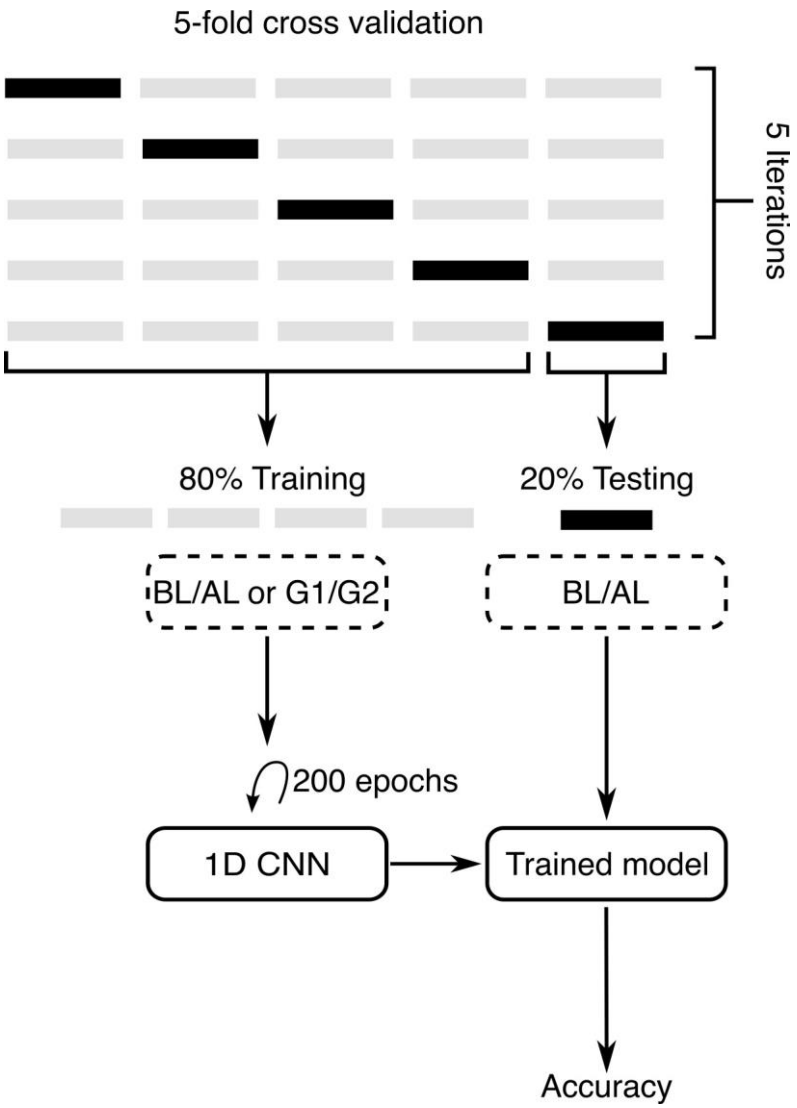

Figure 17 1D CNN training paths.

In each iteration of a 5-fold cross-validation, 80% of the data is allocated to two distinct training sets: with original labels (BL/AL) or re-labeled (G1/G2), and 20% with original BL/AL labels is used for testing. The 1D CNN is then trained over 200 epochs, which is followed by an evaluation phase where the model's accuracy is assessed using the test data.

The 5-fold cross-validation strategy offers several advantages in the context of our study. Firstly, it allows for a more efficient use of our dataset, which is particularly important given the challenges in acquiring large volumes of high-quality neurophysiological data. Secondly, it provides a more reliable estimate of the model's performance on unseen data, thereby addressing concerns about overfitting that are common in complex neural network models (Cawley and Talbot 2010).

2. Performance metrics

Accuracy is the most straightforward metric, and it is defined as the ratio of correctly classified samples to the total number of samples in the dataset. Mathematically, it can be represented as:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Samples}}{\text{Total Number of Samples}}$$

Accuracy is particularly useful when the classes are balanced, meaning that the number of samples in each class is roughly equal.

I. Model interpretability

Model interpretability is of paramount importance in machine learning. While achieving high classification accuracy is crucial, understanding the underlying decision-making process of the model can provide invaluable insights into the features that distinguish sharp wave ripples (SWRs) before and after learning. To this end, we employed Gradient-weighted Class Activation Mapping (Grad-CAM), a state-of-the-art visualization technique that elucidates the regions of input data most influential in the model's classification decisions.

1. Gradient-weighted Class Activation Mapping (Grad-CAM)

Grad-CAM, introduced by (Selvaraju et al. 2020), is an extension of Class Activation Mapping (CAM) that is applicable to a wide variety of CNN model architectures. This technique uses the gradients of any target concept (in our case, pre- or post-learning SWRs) flowing into the final convolutional layer to produce a coarse localization map highlighting the important regions in the input for predicting the concept.

The Grad-CAM algorithm begins with a forward propagation of an input SWR through the CNN to obtain raw class scores. Subsequently, it computes the gradient of the score for the target class with respect to feature map activations of the last convolutional layer. These gradients undergo global average pooling to obtain weights for each feature map. A weighted combination of forward activation maps is then generated, followed by the application of a ReLU function to focus on features that have a positive influence on the class of interest (Figure 18). Mathematically, the Grad-CAM technique can be represented as follows:

$$L^c = \mathrm{Re}LU\left(\sum_k \alpha_k^c A^k\right)$$

Where L^c is the Grad-CAM localization map for class c, $ReLU$ is the rectified linear unit activation function, A^k are the activation maps, and α_k^c are the weights corresponding to the gradients for class c. The weights α are computed as:

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}$$

Here y^c is the score for class c, Z is a normalization constant, and A_{ij}^k is the activation at spatial location i, j in feature map k.

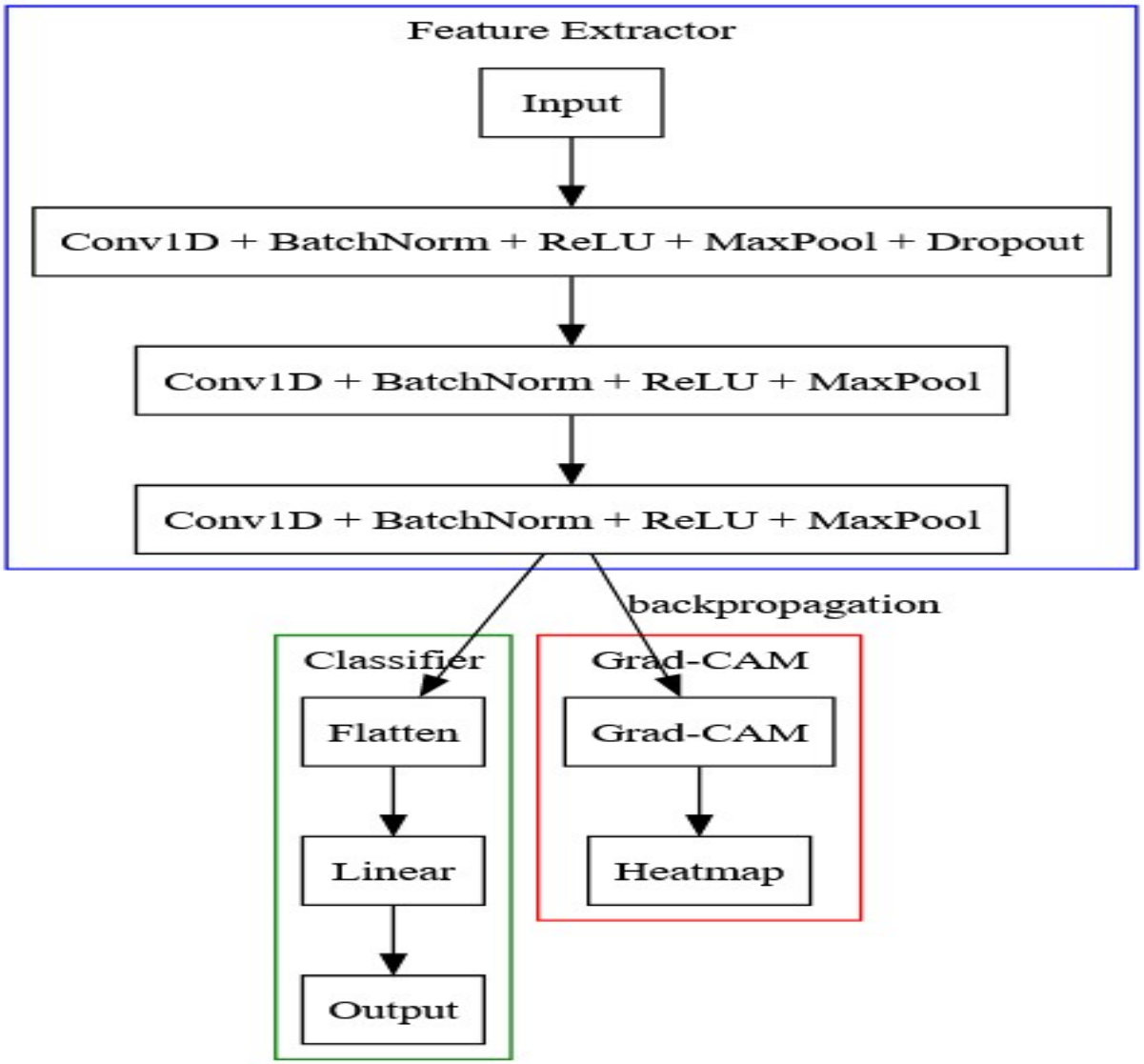

Figure 18 Grad-CAM on 1D CNN model architecture.

Integration of Grad-CAM with our 1D CNN architecture. Grad-CAM utilizes the Conv1D layer to generate a heatmap that highlights the regions in the input sequence most relevant for classification. The Conv1D layer's activations are weighted by the gradients flowing back from the output layer, effectively spotlighting the critical features for model decision-making. This provides invaluable insights into the model's internal workings and aids in interpretability, allowing us to understand which parts of the input data are essential.

In the context of our SWR classification task, Grad-CAM allowed us to visualize which temporal or spectral components of the SWR were most crucial for the model's decision. This is particularly important given the complex and multifaceted nature of SWRs, where subtle changes in waveform characteristics might signify important physiological differences between pre- and post-learning states.

Figure 18 illustrates the integration of Grad-CAM with our 1D CNN architecture. The heatmap generated by Grad-CAM highlights the regions in the input SWR that are most relevant for classification, providing a visual representation of the model's focus during decision-making. This visualization enables us to identify the temporal localization of important features within the SWR.

J. Feature analysis

1. Feature extraction and feature importance

To extract meaningful features from our sharp wave ripple (SWR) data, we employed two robust libraries: Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh) and Time Series Feature Extraction Library (TSFEL). The tsfresh library, developed by Christ et al. (Christ et al. 2018), offers an automated approach to feature engineering in time-series data, providing 794 features derived from 63 time series characterization methods. Its capability to conduct feature selection based on statistically significant hypothesis tests ensured the inclusion of only genuinely relevant features. Complementing this, TSFEL (Barandas et al. 2020) computed over 60 features across temporal, statistical, and spectral domains, offering valuable insights into the computational complexity of the extracted features.

Following feature extraction, we utilized PyCaret (Pycaret 3.0.4) to discern the relative importance of these features. PyCaret's automated machine learning capabilities were particularly useful in quantifying feature importance for our binary classification task of categorizing SWRs before and after learning. This analysis highlighted two standout features: Fast Fourier Transform (FFT) coefficients and wavelet entropy. FFT coefficients provided crucial information about the frequency domain characteristics of the SWR signals, capturing key aspects such as amplitude and phase. Wavelet entropy, on the other hand, offered insights into the complexity and unpredictability of the time-series data.

2. Morlet wavelet analysis and wavelet entropy

Wavelet analysis, particularly using the Morlet wavelet, proved to be a powerful tool for time-frequency analysis of our non-stationary SWR signals. The Morlet wavelet, defined mathematically as:

$$\psi(t) = \pi^{-1/4} e^{i\omega t} e^{-t^2/2}$$

where ω is the angular frequency of the complex sinusoid, offers excellent localization in both time and frequency domains. This property makes it especially useful for analyzing complex signals like SWRs.

The Morlet wavelet transform of a function $f(t)$ is computed as:

$$W_f(a,b) = \int_{-\infty}^{\infty} f(t)\psi^*_{a,b}(t)dt$$

where $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$ and $\psi^*_{a,b}(t)$ is the complex conjugate of $\psi_{a,b}(t)$.

Complementing the Morlet wavelet analysis, we utilized wavelet entropy to measure the disorder or complexity of the SWR signals in the wavelet domain. Wavelet entropy is defined as:

$$H = -\sum_{j=1}^{J} p_j \log_2\left(p_j\right)$$

where p_j is the normalized energy at the j-th scale and J is the number of decomposition levels. This metric provided a robust way to encapsulate the complexity and energy distribution of the SWR signals across different frequency bands into a single scalar value.

3. Feature ablation study

To corroborate the importance of the selected features, we conducted a feature ablation study. This involved running our model multiple times, systematically omitting one or more features in each iteration, and observing the impact on the model's performance metrics, particularly accuracy.

K. Statistical methods

Statistical methods provide the backbone for validating the findings of this research, ensuring that the observed differences and trends are not mere coincidences but are statistically significant. The specific statistical tests employed in this study include the T-Test, Wilcoxon Signed-Rank Test, and Repeated Measures Analysis of Variance (ANOVA). Each of these tests serves a unique purpose and offers distinct advantages depending on the distribution and nature of the data.

1. T-Test

The Student's t-test, a fundamental statistical technique in neuroscience research, was employed in this study. This parametric test assumes that the data follows a normal distribution and that the variances of the two groups are approximately equal (Zimmerman 1997). In our analysis, we utilized the independent samples t-test, as our SWR data from pre- and post-learning conditions were considered independent.

The t-test evaluates the null hypothesis that there is no significant difference between the means of the two groups. The test statistic t is calculated as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means, s_1^2 and s_2^2 are the sample variances, and n_1 and n_2 are the sample sizes. A p-value less than the chosen alpha level (typically 0.05) indicates a statistically significant difference between the groups, allowing us to reject the null hypothesis.

2. Wilcoxon signed-rank test

To account for potential non-normality in our data distribution, we also employed the Wilcoxon signed-rank test, a non-parametric alternative to the paired t-test. This test is particularly useful when the assumption of normality is violated or when dealing with ordinal data (Wilcoxon 1945). The Wilcoxon signed-rank test evaluates the null hypothesis that the median difference between paired observations is zero.

The test procedure involves calculating the differences between paired observations, ranking these differences, and then comparing the sum of positive ranks to the sum of negative ranks. The test statistic W is the smaller of these two sums. For large samples, the distribution of W

approximates a normal distribution, allowing for the calculation of a z-score and corresponding p-value.

In our study, the Wilcoxon signed-rank test provided a robust method for comparing SWR characteristics before and after learning, particularly when the data did not meet the assumptions required for parametric tests.

3. Repeated measures ANOVA

To analyze the temporal dynamics of SWR characteristics, we utilized repeated measures Analysis of Variance (ANOVA). This statistical method is particularly suited for analyzing data collected from the same subjects under different conditions or time points, allowing us to control for individual variability (Girden 1992).

The repeated measures ANOVA model can be expressed as:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where Y_{ijk} is the dependent variable, μ is the grand mean, α_i is the effect of the i^{th} level of the independent variable, β_j is the effect of the j^{th} subject, $(\alpha\beta)_{ij}$ is the interaction term, and ϵ_{ijk} is the residual error term.

In our analysis, the repeated measures ANOVA allowed us to assess how SWR characteristics changed over the course of the maze learning experiment, accounting for both within-subject and between-subject variability. This approach provided insights into the temporal dynamics of SWR modulation during the learning process.

L. Software and tools

The implementation of our research methodology required a robust and versatile development environment capable of handling complex data processing, statistical analysis, and deep learning model development. To this end, we utilized a comprehensive suite of software tools and libraries, each chosen for its specific strengths and compatibility with our research objectives.

1. Python programming language

At the core of our development environment was Python (version 3.10), a high-level, interpreted programming language widely recognized for its simplicity, readability, and extensive ecosystem of scientific computing libraries (Van Rossum and Drake 2009). Python's versatility and extensive support for numerical and scientific computing made it an ideal choice for our research, allowing for efficient data manipulation, analysis, and visualization.

2. Deep learning frameworks

For the development and training of our deep learning models, we employed two leading frameworks: PyTorch (version 2.1.1) and TensorFlow (version 2.13). PyTorch, developed by Facebook's AI Research lab, offers dynamic computational graphs and imperative programming, which proved particularly useful for our experimental iterations (Paszke et al. 2019). TensorFlow, created by the Google Brain team, provided a comprehensive ecosystem for machine learning and deep neural network modeling (Abadi et al. 2016). The use of both frameworks allowed us to leverage their respective strengths and ensure the robustness of our results across different implementations.

3. Scientific computing and statistical analysis

For statistical analysis and scientific computing tasks, we relied heavily on SciPy (version 1.11.4). This open-source library builds on NumPy arrays and provides a collection of tools for optimization, linear algebra, integration, and statistics (Virtanen et al. 2020). SciPy's statistical functions were particularly crucial for implementing our t-tests, Wilcoxon signed-rank tests, and repeated measures ANOVA.

4. Data visualization

To create clear and informative visualizations of our data and results, we utilized Matplotlib (version 3.8.2). This comprehensive library for creating static, animated, and interactive visualizations in Python allowed us to generate publication-quality figures and plots (Hunter 2007). Matplotlib's flexibility and extensive feature set enabled us to create custom

visualizations tailored to the specific needs of our research, including time series plots, spectrograms, and statistical graphics.

5. Custom algorithm development

All basic and advanced algorithms used in this work were developed and implemented using custom-written Python scripts. This approach allowed us to tailor our methods precisely to the unique challenges of analyzing sharp wave ripples in the context of learning and memory. By developing custom algorithms, we ensured that our analysis pipeline was optimized for our specific research questions and data characteristics.

III. Results

This section delineates the findings from our multifaceted approach, which combines deep learning algorithms, statistical analyses, and advanced signal processing techniques to interrogate the complex nature of SWRs before and after a maze learning task.

Our investigation builds upon the growing body of evidence suggesting that SWRs play a crucial role in memory consolidation and retrieval (Buzsáki 2015; Joo and Frank 2018). By leveraging state-of-the-art machine learning methodologies, we sought to discern subtle yet significant changes in SWR characteristics that may reflect the neural processes underlying spatial learning and memory formation.

The results presented herein are structured to address several key aspects of our research. First, we report on the efficacy of our deep learning model in classifying SWRs, providing insights into the model's performance and the challenges encountered in this classification task. Subsequently, we delve into the impact of our novel self-supervised learning approach on classification accuracy, exploring how this method mitigates the issue of label noise inherent in complex neurophysiological data (Frénay and Kabán 2014).

Our analysis extends beyond mere classification to encompass a detailed examination of the features that distinguish pre- and post-learning SWRs. Through rigorous feature importance analysis and ablation studies, we shed light on the specific spectral and temporal characteristics that undergo modification as a result of the learning process. This approach aligns with recent efforts to uncover the mechanistic underpinnings of experience-dependent changes in SWR properties (Jura et al. 2019).

Furthermore, we employ advanced visualization techniques, such as Gradient-weighted Class Activation Mapping (Grad-CAM), to elucidate the decision-making process of our deep learning model (Selvaraju et al. 2020). This analysis is complemented by a comprehensive Morlet wavelet decomposition, providing a time-frequency perspective on SWR dynamics that corroborates and extends our machine learning findings.

To ensure the robustness and reliability of our results, we subject our data to a battery of statistical tests, including parametric and non-parametric analyses. These tests not only validate our primary findings but also reveal nuanced patterns in SWR modulation that may have implications for our understanding of hippocampal function in learning and memory.

Finally, we extend our investigation to a transgenic mouse model of Alzheimer's disease, exploring the potential of our analytical framework to identify SWR alterations in pathological conditions. This translational aspect of our study addresses the growing interest in SWRs as potential biomarkers for neurodegenerative disorders (Ognjanovski et al. 2018).

Through this comprehensive and multifaceted analysis, we aim to contribute novel insights into the nature of SWRs and their role in spatial learning, while also demonstrating the power of advanced computational techniques in unraveling the complexities of neural information processing.

A. SWR dataset exploratory data analysis (EDA)

1. Dataset overview

The sharp wave ripple (SWR) dataset used in this study was derived from electrophysiological recordings obtained from the hippocampal CA1 region of mice during a spatial learning task. The experimental paradigm involved an elevated eight-arm radial maze, which was used to assess spatial learning and memory (Olton, Collison, and Werz 1977).

Each day of the six-day experiment followed a structured protocol. It began with a 90-minute recording session in the home cage, followed by maze exposure. During the maze task, only three specific arms were baited: two adjacent and one separated by a non-baited arm. This configuration remained constant for each animal throughout the learning period. Animals underwent six maze trials per day, with each trial ending when all rewards were consumed. Following the maze trials, another 90-minute recording session was conducted in the home cage (Figure 8).

The study employed a cohort of mice comprising two distinct genotypes: six wild-type (WT) and six transgenic (TG) animals. The TG mice were bred as a model for Alzheimer's disease (AD). All mice were aged 8-9 months at the time of experimentation.

Electrophysiological recordings were obtained from the CA1 area of the hippocampus in both hemispheres. During recording sessions, the mouse head connector was linked to amplifiers via a soft cable, permitting free movement. Neurophysiological and EMG signals were acquired at a sampling rate of 40 kHz using a 128-channel Plexon system. This high sampling rate was chosen to capture the full spectral range of neural activity, including high-frequency oscillations like SWRs. To facilitate further analysis, the data were down-sampled to 2000 Hz.

SWR detection involved several preprocessing steps. Initially, the local field potential (LFP) signals were filtered using a 4th-order Chebyshev Type II filter in the 100-250 Hz range to isolate SWR events. The Hilbert transform was then applied to compute the signal envelope, representing the instantaneous amplitude of the filtered signal. SWR events were identified based on the z-scored envelope of the filtered signal, normalized using standard deviation values from slow-wave sleep (SWS) bouts during baseline recording sessions. Ripple bouts were defined as epochs where the envelope exceeded 2 standard deviations (SDs) and peaked at 5 SDs. Events spaced less than 20 ms apart were merged, and those longer than 100 ms were discarded (Figure 9).

Following the methodology outlined by Hsu et al. (2021), the LFP signals were segmented into non-overlapping Ripple-Centered Intervals (RCIs) of 256 milliseconds (Figure 10). This segmentation process was crucial for preparing the data for subsequent deep learning analysis. Each RCI was centered on a detected SWR event, ensuring that the full temporal extent of the SWR and its immediate context were captured. The resulting dataset comprises 38,822 RCIs.

To gain insights into the characteristics of our SWR dataset, we conducted a comprehensive exploratory data analysis. This analysis involved various statistical and visualization techniques to uncover patterns, distributions, and potential relationships within the data.

2. Distribution of SWRs before and after Learning

Figure 19 presents a bar plot showing the distribution of SWRs before and after the learning task. The relatively balanced nature of the dataset is evident, with 21,244 pre-learning SWRs and 17,578 post-learning SWRs.

While there is a slight predominance of pre-learning SWRs, it's important to note that we implemented measures to ensure perfect class balance in our deep learning analyses. Specifically, for all our training procedures, we used exactly the same number of SWRs from each class. This approach of using balanced subsets of the data eliminates any potential bias that could arise from class imbalance during model training.

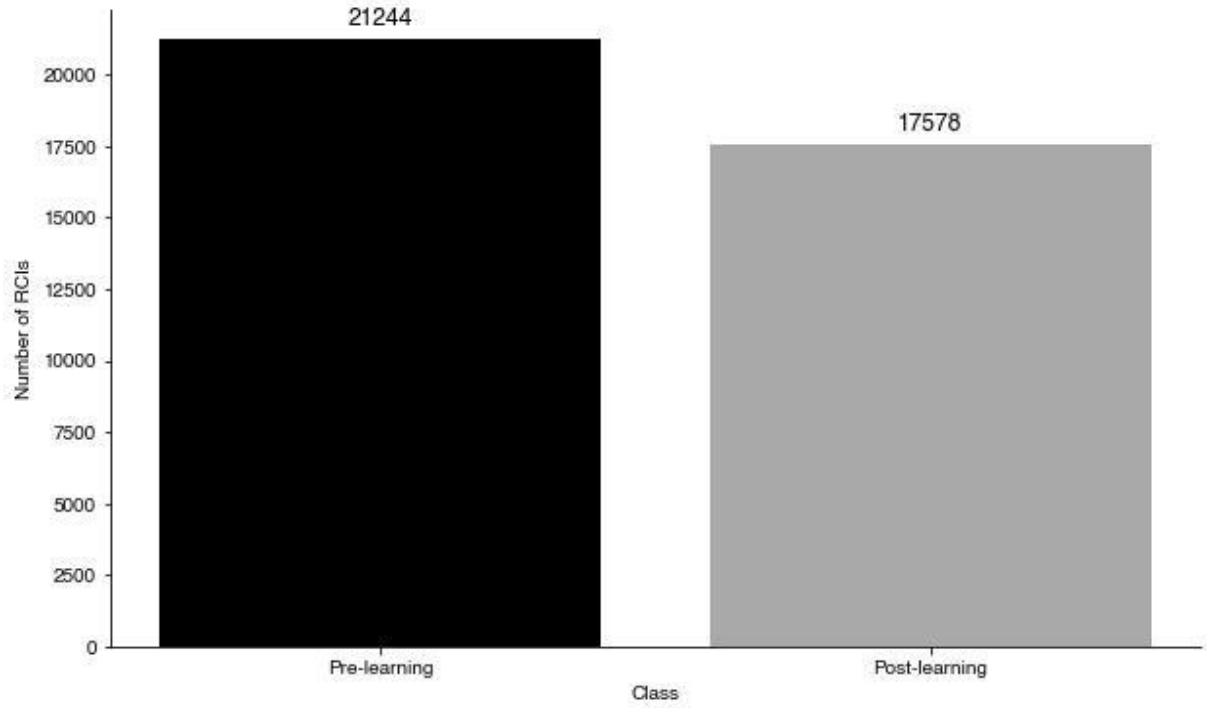

Figure 19 Distribution of SWRs before and after learning.

Bar plot illustrating the distribution of SWRs before and after the learning task.

By enforcing this strict balance, we ensure that our models learn equally from both pre- and post-learning SWRs, thereby providing a fair and unbiased basis for comparing the characteristics of these two classes. This methodological choice strengthens the reliability of our subsequent analyses and the interpretability of our results, as any observed differences can be attributed to genuine distinctions between pre- and post-learning SWRs rather than artifacts of class imbalance.

3. SWR duration analysis

Figure 20 shows the distribution of SWR durations for both before and after learning events, using a combination of histogram and kernel density estimation (KDE) plot. SWR duration, defined as the time between the onset and offset of the ripple oscillation, is a critical parameter that may reflect underlying network dynamics (Fernández-Ruiz et al. 2019). Our analysis reveals subtle differences in the duration distributions between the two classes, which could indicate learning-induced changes in the temporal structure of SWRs.

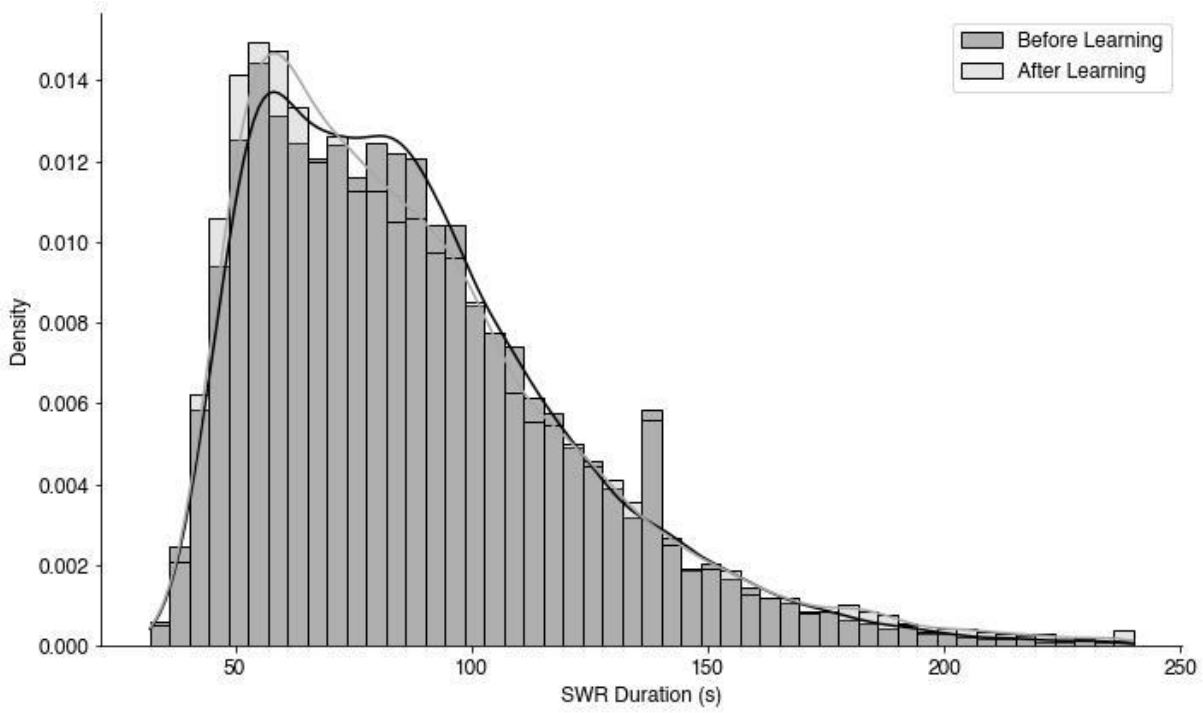

Figure 20 SWR duration analysis.

Distribution of SWR durations for both pre- and post-learning events

The observed variations in SWR duration align with previous findings suggesting that learning experiences can modulate the temporal characteristics of ripple events. For instance, Girardeau et al. (Girardeau, Cei, and Zugaro 2014) reported changes in SWR duration following spatial learning tasks, potentially reflecting the encoding of new spatial information into hippocampal circuits.

4. Intrinsic frequency analysis

Figure 21 presents a combined histogram and kernel density estimation (KDE) plot of the intrinsic frequencies of SWRs in our dataset. The intrinsic frequency, calculated as the peak frequency within the ripple band (typically 140-250 Hz in rodents), provides insights into the oscillatory nature of these events (Buzsáki 2015). Our analysis reveals a distribution centered around the expected ripple frequency range, with notable variations between pre- and post-learning SWRs.

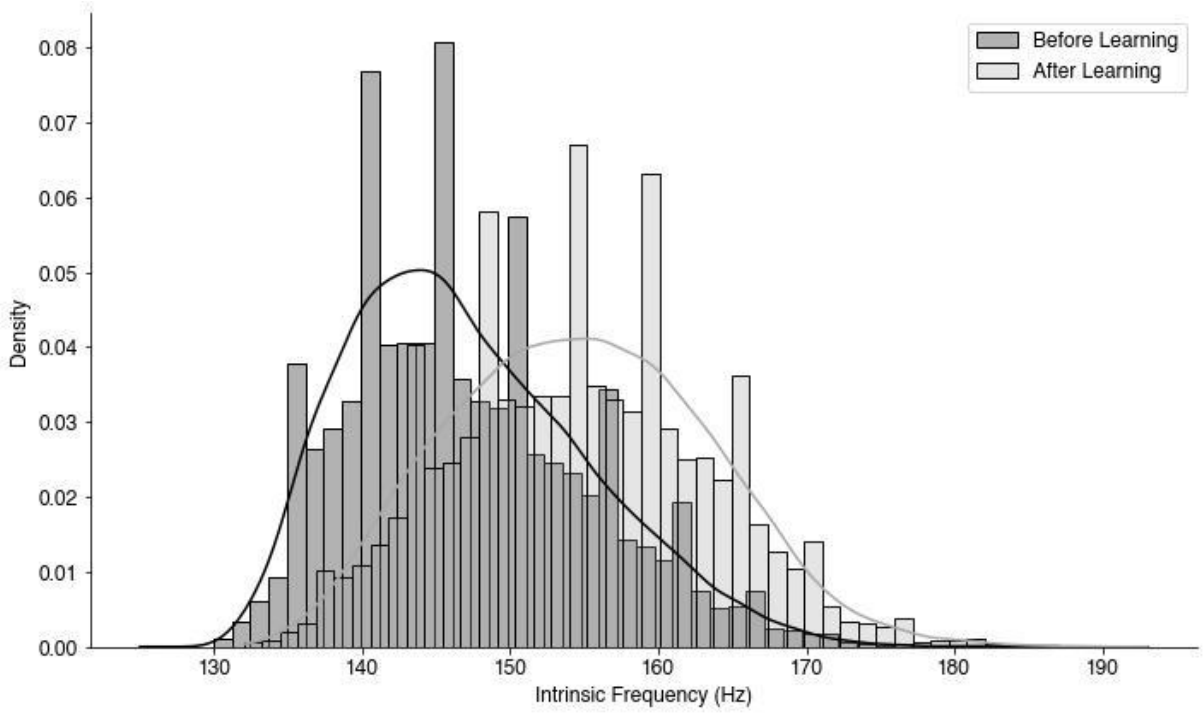

Figure 21 Distribution of SWR intrinsic frequencies before and after learning.

The plot combines histogram bars with overlaid KDE curves, showing a clear rightward shift in frequency distribution after learning.

As evident in Figure 21, the observed frequency distribution demonstrates a significant shift towards higher frequencies in post-learning SWRs. The histogram bars provide a detailed view of the frequency counts, while the smooth KDE curves offer a clear visualization of the overall distribution shift. This aligns with previous studies that have reported learning-induced changes in SWR spectral properties. For example, Jura et al. (Jura et al. 2019) found that the frequency content of SWRs can be modulated by recent learning experiences.

The variations in intrinsic frequency between before and after learning SWRs are visually apparent and statistically significant ($p < 0.001$), indicating substantial changes in the underlying network dynamics following the learning task. The before learning distribution shows a peak around 140-150 Hz, while the after learning distribution exhibits a broader peak shifted towards 150-160 Hz. Interestingly, our data shows a broader distribution of frequencies after learning, with a noticeable increase in the occurrence of higher frequency SWRs.

B. Classification of SWRs using deep learning

1. Benchmarking various architectures

In our pursuit of enhancing sharp wave ripple (SWR) classification accuracy, we conducted a comprehensive evaluation of six distinct deep learning algorithms. This comparative analysis aimed to surpass the performance established by Hsu et al. (Hsu et al. 2021) and identify the most suitable model for our specific SWR classification task. Our approach encompassed both one-dimensional (1D) and two-dimensional (2D) processing methodologies, allowing for a thorough exploration of diverse model architectures and their applicability to SWR data.

For 2D processing, we transformed the SWR temporal sequences into image-like representations, enabling the application of state-of-the-art convolutional neural network (CNN) architectures renowned for their image classification capabilities. The models employed in this category included MobileNet, EfficientNet, VGG16, and ConvNeXt. MobileNet, utilizes depthwise separable convolutions to achieve computational efficiency while maintaining high accuracy (Howard et al. 2017). EfficientNet, known for its compound scaling method, optimizes network depth, width, and resolution to achieve state-of-the-art performance on image classification tasks (Tan and Le 2020). Despite its earlier inception, VGG16 continues to serve as a robust baseline due to its simplicity and effectiveness (Simonyan and Zisserman 2015). ConvNeXt, a more recent development, modernizes traditional CNN architectures by incorporating design elements from vision transformers (Liu et al. 2022).

In parallel, we explored 1D processing methods that operate directly on the raw temporal sequences of SWRs. This approach utilized Long Short-Term Memory (LSTM) networks and 1D Convolutional Neural Networks (1D CNNs). LSTMs, with their ability to capture long-term dependencies in sequential data, have been widely applied in various neuroscience applications, including EEG signal analysis (Hochreiter and Schmidhuber 1997; Craik, He, and Contreras-Vidal 2019). 1D CNNs have gained prominence in recent years for their efficiency in extracting local patterns from time-series data while maintaining sensitivity to scale and translation invariance (Kiranyaz et al. 2021).

Our results, as illustrated in Figure 22, reveal a consistent performance across all tested models, with classification accuracies ranging from 69% to 73%. This narrow range of performance suggests that the models are capturing similar underlying patterns in the SWR data, regardless of their specific architectural differences. Among the tested architectures, the

1D CNN, following careful hyperparameter optimization, achieved the highest accuracy of 73%.

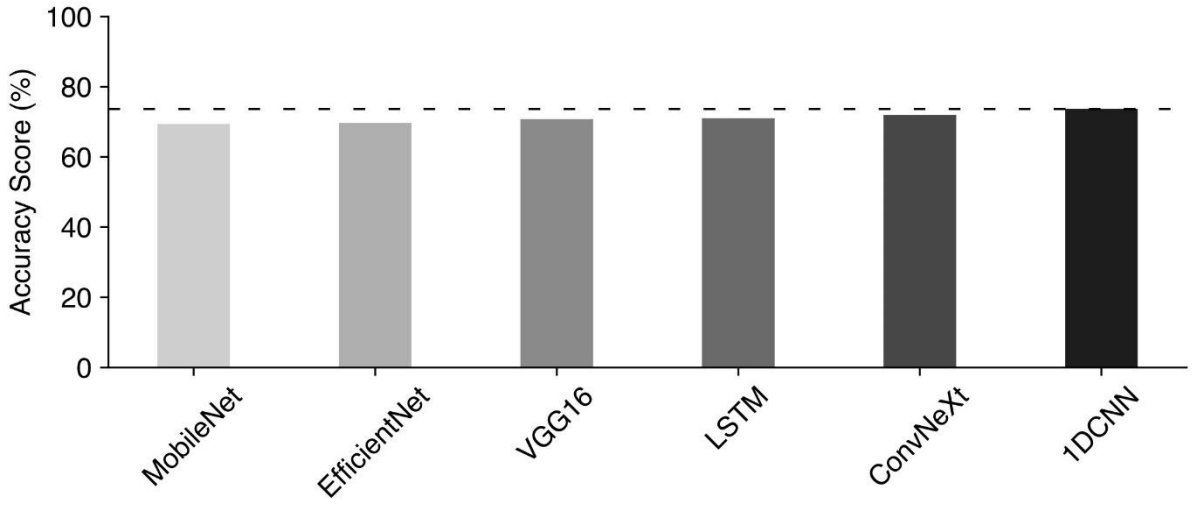

Figure 22 Performance comparison of deep learning models.

MobileNet achieved the lowest accuracy score, while 1D CNN reached the highest accuracy. However, all models exhibited relatively similar performance levels.

The 2D models (MobileNet, EfficientNet, VGG16, and ConvNeXt) exhibited slightly lower performance compared to the 1D models, with accuracies ranging from 69% to 71%. This marginal difference may be attributed to the inherently sequential nature of SWR data, which appears to be more effectively captured by models specifically designed for time-series analysis. Notably, the LSTM model, despite its theoretical advantages in capturing long-term dependencies, did not outperform the 1D CNN. This observation suggests that the most discriminative features for SWR classification may reside in local temporal patterns rather than long-range dependencies.

2. Performance limitation analysis

The consistent accuracy ceiling observed across different models, with performance plateauing around 73%, suggests the presence of underlying factors limiting classification accuracy beyond model architecture. Among these factors, label noise emerges as a critical issue deserving particular attention in the context of our sharp wave ripple (SWR) classification task.

Initially, the labeling of SWRs as 'before' or 'after' learning, based on their temporal occurrence relative to the maze experiment, seemed straightforward. However, the complexity of this labeling scheme became apparent as our study progressed through six days of repeated experimental conditions. This extended timeline introduced nuances that challenged the simplicity of our binary classification approach.

A primary concern arose from the possibility that SWRs generated before a learning task on later experimental days might express features acquired during previous days' learning sessions. This phenomenon aligns with our understanding of memory consolidation and retrieval processes, where previously formed memories can be reactivated in subsequent sessions (Carr, Jadhav, and Frank 2011). Such reactivation could produce SWRs before a learning session that bear close resemblance to those typically observed after learning, thus blurring the distinction between our predefined categories.

Furthermore, we recognized that not all SWRs are equally impacted by learning experiences. Some SWRs generated after a learning session may retain characteristics similar to those produced before learning, possibly due to their involvement in processes unrelated to the specific learning task or individual variability in neural plasticity (Buzsáki 2015). This inherent variability in SWR responses to learning adds another layer of complexity to the classification task.

These observations led us to hypothesize that our initial labeling scheme, while logical from an experimental design perspective, might not accurately reflect the intrinsic structure of the data. The potential mismatch between assigned labels and the true neurophysiological state of the network introduces significant label noise, a well-recognized challenge in machine learning applications (Frénay and Kabán 2014).

Label noise can severely impact the performance of classification models by providing inconsistent or incorrect information during the training process. In our case, it could lead to models learning to distinguish between 'before' and 'after' categories based on features that are not truly representative of learning-induced changes, but rather artifacts of our labeling scheme. This misalignment between labels and underlying data structure likely contributes to the performance ceiling observed across various model architectures in our initial benchmarking efforts.

C. Addressing label noise in SWR classification

The challenge of label noise in our sharp wave ripple (SWR) classification task necessitated the exploration of advanced techniques to mitigate its impact on model performance. We implemented three distinct approaches: genetic algorithms, autoencoders, and self-supervised learning (SSL). Each method offered unique advantages in addressing the complexities of label noise in neurophysiological data.

1. Genetic algorithm approach

Our initial attempt to address label noise employed a genetic algorithm, a method inspired by the principles of natural selection and evolution. This approach has shown promise in optimizing complex, multi-dimensional problems in various domains, including bioinformatics and neural network optimization (Srinivas and Patnaik 1994). In our implementation, we encoded the labels of SWRs as binary strings within chromosomes, with each gene representing the label of a single SWR instance. The fitness function was designed to maximize the classification accuracy of our 1D CNN model when trained on the evolved labels.

The genetic algorithm proceeded through multiple generations, with each generation involving selection, crossover, and mutation operations. Selection favored chromosomes (label configurations) that resulted in higher classification accuracies. Crossover allowed for the exchange of label information between high-performing chromosomes, while mutation introduced random alterations to maintain genetic diversity and explore the solution space.

Despite the theoretical advantages of this approach in navigating complex fitness landscapes, the genetic algorithm yielded only modest improvements in classification accuracy. After 100 generations, with a population size of 50 chromosomes per generation, the best-performing label configuration achieved an accuracy of 70.5%, representing a marginal improvement over our initial baseline.

2. Autoencoders implementation

Our second approach utilized autoencoders, a type of artificial neural network that learns to encode data into a compressed representation and then decode it back to its original form. Autoencoders have been successfully applied in various domains for anomaly detection and

data denoising (Vincent et al. 2010). We hypothesized that mislabeled SWRs might manifest as anomalies in the learned representations, allowing for their identification and potential correction.

The autoencoder architecture consisted of an encoder network that compressed the input SWR data into a lower-dimensional latent space, followed by a decoder network that reconstructed the original input from this latent representation. We employed a convolutional autoencoder design, leveraging the temporal structure of SWR data. The network was trained to minimize reconstruction error on the entire dataset, irrespective of labels.

Post-training, we analyzed the reconstruction error for each SWR instance. Instances with reconstruction errors exceeding a statistically determined threshold were flagged as potential mislabels. These flagged instances underwent a secondary classification process using the trained 1D CNN model, and labels were adjusted based on the model's high-confidence predictions.

This autoencoder-based approach resulted in a modest improvement in classification accuracy, reaching 71.2%. While this represented a step forward, the improvement was not substantial enough to fully address the challenges posed by label noise in our dataset.

3. Self-supervised learning

The most significant breakthrough in addressing label noise came with the implementation of a self-supervised learning (SSL) framework. SSL has emerged as a powerful paradigm in machine learning, allowing models to learn meaningful representations from unlabeled data by solving pretext tasks (Gui et al. 2024). We adapted the Time-Series Representation Learning via Temporal and Contextual Contrasting (TS-TCC) framework proposed by Eldele et al. (Eldele et al. 2021) to our SWR classification task.

Our SSL model architecture comprised two primary components: a temporal contrasting module and a contextual contrasting module. These modules worked together to maximize the similarity between augmented views of the same SWR sample while minimizing similarity with views from different samples. Data augmentation played a crucial role, with weak and strong augmentation techniques applied to enhance the model's learning.

The SSL framework allowed the model to learn rich, generalizable representations of the SWR data without relying on potentially noisy labels. Following the SSL training, we used the learned representations to generate new labels for our dataset, effectively redistributing the SWRs based on their learned features rather than their original temporal classification.

D. Improved classification through re-labeling by SSL model

1. Study of labels reassigned by self-supervised learning model

The SSL model's re-labeling process (Figure 16) resulted in a substantial reorganization of our SWR dataset, offering new perspectives on the underlying structure of the data. This redistribution of labels suggests that the SSL model identified features within the SWRs that transcend the simple temporal demarcation of before learning (BL) and after learning (AL) used in our initial classification scheme. The nuanced redistribution reveals the complex nature of SWRs and their relationship to learning and memory processes, challenging our initial assumptions about the clear-cut distinction between pre- and post-learning neural activity.

Upon applying the SSL model to our dataset, we observed a noteworthy redistribution of SWRs into two new groups (Figure 23), designated as Group 1 (G1) and Group 2 (G2). This redistribution revealed interesting patterns in the relationship between the original temporal labels and the new feature-based groupings. Group 1 (G1) comprised 87% of SWRs originally labeled as before learning (BL) and 11.2% of SWRs originally labeled as after learning (AL). Conversely, Group 2 (G2) contained 88.8% of SWRs originally labeled as after learning (AL) and 13% of SWRs originally labeled as before learning (BL).

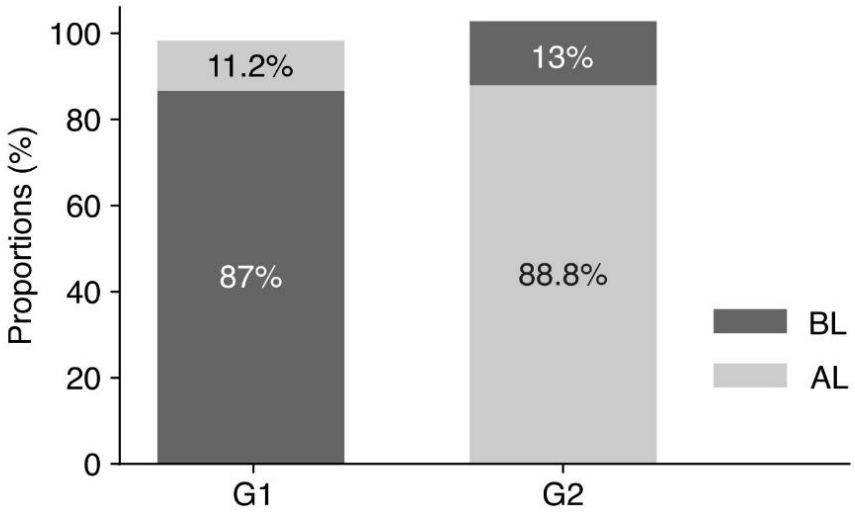

Figure 23 Distribution of SWRs post-SSL re-labeling.

The bar graph illustrates the new distribution of SWRs after applying the self-supervised learning (SSL) re-labeling process. It shows the proportion of SWRs originally recorded before learning (BL) and after learning (AL) within the newly established groups, G1 and G2.

This redistribution indicates a strong, but not perfect, correlation between the original temporal labels and the new feature-based groupings. The presence of a substantial proportion of AL SWRs in G1 (11.2%) and BL SWRs in G2 (13%) suggests that the SSL model identified subtle features that differentiate SWRs beyond their temporal occurrence relative to the learning task.

2. Effect of self-supervised learning on classification accuracy

To evaluate the impact of SSL re-labeling on classification accuracy, we conducted a series of experiments using our 1D CNN model (Figure 17). Initially, we established a baseline performance by training the 1D CNN model on the original BL/AL labels, which resulted in an accuracy of 73.28% (Figure 24, left bar). This baseline represents the performance achievable when relying solely on the temporal occurrence of SWRs relative to the learning task as the basis for classification.

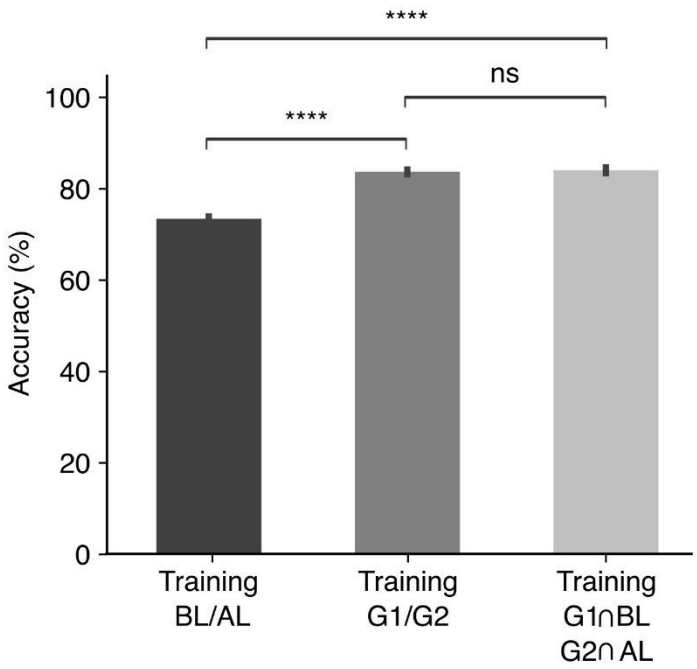

Figure 24 Enhanced accuracy of 1D CNN model following SSL re-labeling.

*The bar graph illustrates the accuracy of the 1D CNN model trained with the original dataset (BL/AL) (left bar), the re-labeled data (G1/G2) (middle bar), and the intersection of G1 with BL and G2 with AL (right bar). All evaluations were performed on the BL/AL dataset. Statistical significance, determined via t-tests, is marked directly above the bars: ****p < 0.0001, and 'ns' indicates non-significance with p > 0.05. Vertical lines denote the standard error of the mean (SEM).*

Following the application of SSL re-labeling, we observed a substantial improvement in classification accuracy. Training the 1D CNN model on the SSL re-labeled data (G1/G2) resulted in a significant increase in classification accuracy to 83.66% (Figure 24, middle bar). This improvement, confirmed through statistical analysis (N=5, t-test, $p<0.0001$), represents a gain of over 10 percentage points compared to the original labeling scheme. Such a substantial increase in accuracy suggests that the SSL approach successfully identified and leveraged more informative features within the SWR data, enabling the 1D CNN model to make more accurate distinctions between different classes of SWRs.

To further explore the relationship between the original temporal labels and the SSL-derived groupings, we examined the intersection of SSL re-labeled groups with the original labels. Specifically, we created a dataset comprising the intersection of Group 1 (G1) with BL instances and Group 2 (G2) with AL instances. Training the 1D CNN model on this intersected dataset yielded an even higher accuracy of 84.11% (Figure 24, right bar). This result was also

statistically significant (N=5, t-test, $p<0.0001$) when compared to the performance on the original dataset.

Interestingly, the performance of the model trained on the intersected groups (G1∩BL, G2∩AL) was not significantly different from that trained on the full G1 and G2 groups (N=5, t-test, $p>0.05$). This similarity in performance suggests that the SSL model effectively identified a core set of features that distinguish between the two classes of SWRs, regardless of their original temporal labels.

E. Evaluating SSL relabeling using SWR_{art} dataset

To rigorously assess the efficacy of our self-supervised learning (SSL) relabeling approach and its impact on sharp wave ripple (SWR) classification, we developed and utilized a synthetic dataset of artificial sharp wave ripples (SWR_{art}). This controlled environment allowed us to systematically investigate the performance of our SSL method under various conditions of label noise, providing insights into its robustness and generalizability.

1. SWR_{art} dataset construction

The SWR_{art} dataset was designed to closely mimic the key characteristics of biological SWRs while allowing for precise control over signal properties and noise levels. Each SWR_{art} was modeled as a Gaussian-modulated sinusoidal wave, with frequencies ranging from 120 to 200 Hz, aligning with the spectral properties of SWRs observed in vivo (Buzsáki 2015). To enhance biological realism, we incorporated delta (1-3 Hz) and low gamma (20-40 Hz) oscillations, as well as Gaussian white noise (Figure 11).

We created two distinct classes of SWR_{art}: Class 1 representing SWRs not affected by learning, and Class 2 representing SWRs transformed by learning. These classes differed in their frequency and amplitude variability, as well as noise levels, to simulate the hypothesized changes in neural activity following learning. For a detailed description of the SWR_{art} construction process, including specific parameter ranges and noise levels, please refer to the Methods section.

2. Evaluation of SSL re-labeling efficacy

To assess the robustness of our SSL re-labeling approach, we conducted a series of experiments using the SWR_{art} dataset. These experiments were designed to test the SSL model's ability to accurately adjust labels under controlled conditions with varying levels of label noise.

We created artificial before learning (BL_{art}) and after learning (AL_{art}) groups by combining SWR_{arts} from Class 1 and Class 2 in different proportions. The proportion of mixing, which we termed "label noise" ranged from 0% to 50%. For instance, at 20% label noise, the BL_{art} group comprised 80% of Class 1 and 20% of Class 2 SWR_{arts}, while the AL_{art} group contained 20% of Class 1 and 80% of Class 2 SWR_{arts}. This mixing strategy allowed us to simulate the complexities and ambiguities often present in real neurophysiological data, where the distinction between pre- and post-learning neural states may not always be clear-cut.

The SSL model was then applied to this mixed dataset to generate new labels, creating $G1_{art}$ and $G2_{art}$ groups. This process mirrored our approach with the biological SWR data, allowing us to evaluate the SSL model's performance in a controlled setting.

We employed a rigorous 5-fold cross-validation strategy to assess the robustness of our approach. The dataset was systematically divided into 80% for training and 20% for testing, ensuring that the model was exposed to diverse segments of the data during both training and validation. This cross-validation approach was crucial for obtaining reliable estimates of the model's performance and generalizability.

3. Impact on 1D CNN model accuracy

To quantify the impact of SSL re-labeling on classification accuracy, we compared the performance of our 1D CNN model under different labeling conditions. The model was first trained and tested using the original BL_{art} and AL_{art} labels, establishing a baseline performance level. Subsequently, we trained the model using the SSL re-labeled data ($G1_{art}$ and $G2_{art}$) and evaluated its performance on the original BL_{art} and AL_{art} test set.

Our results demonstrated a significant improvement in classification accuracy when using the SSL re-labeled data. The baseline model, trained on BL_{art} and AL_{art} without SSL re-labeling, achieved an accuracy of 73.23% at 20% label noise (Figure 25, left bar). In contrast, when trained on the SSL re-labeled data ($G1_{art}$ and $G2_{art}$), the model's accuracy increased to 82.16%

(N=5, t-test, p<0.001) (Figure 25, middle bar). This substantial improvement of nearly 9 percentage points underscores the effectiveness of our SSL approach in mitigating the adverse effects of label noise.

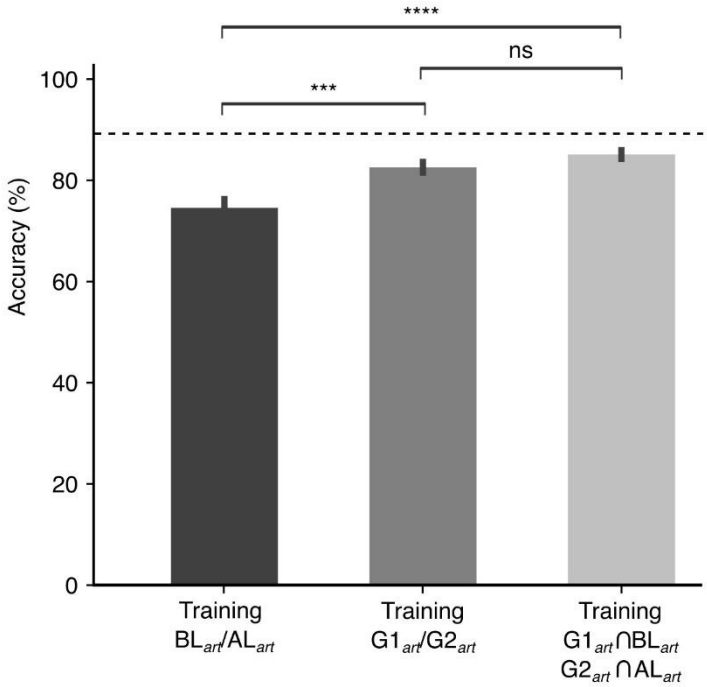

Figure 25 Accuracy comparison of 1D CNN model using SWR$_{art}$ dataset.

*The bar graph compares the accuracy of the 1D CNN model trained under three different labeling conditions: BLart/ALart labels created by mixing 20% of Class 1 and Class 2 (left bar), SSL re-labeled data (G1art/G2art) (middle bar), and the intersection of G1art with BLart and G2art with ALart (right bar). The dashed line represents the ideal accuracy of the 1D CNN model with no label noise. Statistical significance, determined via t-tests, is marked directly above the bars: ****p < 0.0001, ***p < 0.001, and 'ns' for non-significance with p > 0.05. Vertical lines indicate the standard error of the mean (SEM).*

To further validate our findings, we also examined the performance of the model when trained on the intersection of G1$_{art}$ with BL$_{art}$ and G2$_{art}$ with AL$_{art}$. This intersected dataset yielded an even higher accuracy of 83.91% (Figure 25, right bar), which was significantly better than the accuracy obtained with BL$_{art}$/AL$_{art}$ (N=5, t-test, p<0.0001). Interestingly, the performance on this intersected dataset was not significantly different from that obtained using the full G1$_{art}$ and G2$_{art}$ groups, suggesting that the SSL model effectively identified core features that distinguish between the two classes of SWR$_{arts}$, regardless of their original labels.

F. Robustness of SSL re-labeling method against label noise

1. Impact assessment at varying noise levels

Our analysis focused on examining the performance of the 1D CNN model across a spectrum of label noise levels, ranging from 0% to 50%. This comprehensive evaluation allowed us to observe how the model's classification accuracy responded to increasing levels of label uncertainty and to assess the efficacy of SSL re-labeling in counteracting these effects.

The results, as visualized in Figure 26, reveal several key insights into the behavior of our model under different labeling conditions. The black dashed line represents an ideal scenario where the model is trained on perfectly labeled data (BL_{art} equal to Class 1, AL_{art} equal to Class 2). This curve serves as an upper bound, illustrating the best possible performance achievable under noise-free conditions. The gradual decline in this curve as noise levels increase is attributed to the model being tested on increasingly noisy data, even though it was trained on clean data.

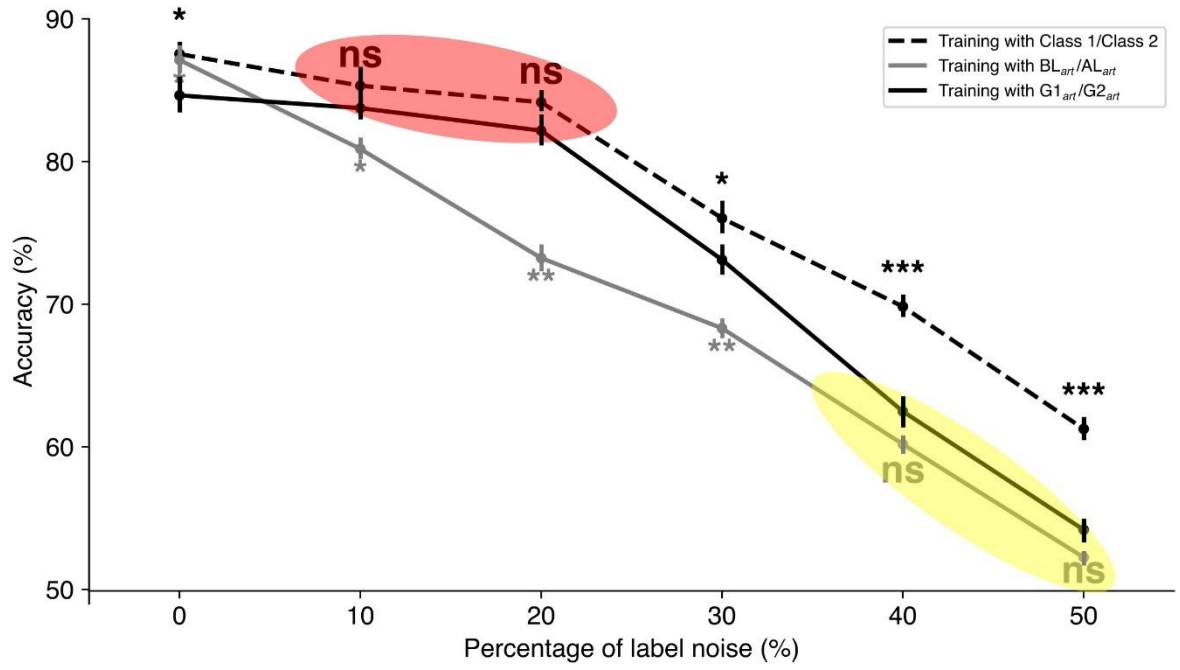

Figure 26 Effects of SSL re-labeling on 1D CNN model accuracy across various levels of label noise.

The graph shows the 1D CNN model's classification accuracy as the percentage of label noise increases (gray solid line), alongside the effects of SSL re-labeling (black solid line) and a noise-free dataset (black dashed line) applied during 1D CNN training. SSL re-labeling is most effective at label noise levels between 10% and 20%, where it nearly matches the ideal scenario's performance (red ellipse). The efficacy of SSL re-labeling decreases as label noise

*approaches 40% and 50% (yellow ellipse). Statistical significance, determined via t-tests, ***p < 0.001, **p < 0.01, *p < 0.05, and 'ns' indicates non-significance with p > 0.05. Vertical lines indicate the standard error of the mean (SEM).*

In contrast, the gray solid line depicts the performance of the 1D CNN model when both trained and tested on data containing label noise. As expected, this curve shows a more pronounced decline in accuracy as noise levels increase, reflecting the challenges posed by inconsistent and unreliable labels during both training and testing phases.

The most intriguing aspect of our results is represented by the black solid line, which illustrates the performance of the model trained on SSL re-labeled data. At 0% noise, we observed that the model trained on the original, noise-free labels slightly outperformed the SSL approach. This finding aligns with observations made by Eldele et al. (Eldele et al. 2021), who noted that supervised learning can outperform SSL methods when working with completely clean, correctly labeled datasets.

However, the true value of our SSL re-labeling method becomes apparent as label noise is introduced. In the range of 10% to 20% label noise, marked by the red ellipse in Figure 26, the SSL re-labeling approach demonstrates remarkable effectiveness. In this range, the performance of the SSL-enhanced model nearly matches that of the ideal scenario, significantly outperforming the model trained on noisy labels. This result suggests that our SSL method is particularly adept at correcting moderate levels of label noise, effectively restoring the dataset to a state that closely approximates noise-free conditions.

As the noise level increases to around 30%, the SSL re-labeling method continues to show significant benefits. While its performance begins to drop below the ideal scenario, it still maintains a substantial advantage over the model trained directly on noisy labels. This persistent advantage underscores the robustness of our SSL approach in handling considerable levels of label uncertainty.

It's noteworthy that at extreme noise levels of 40-50%, the improvements offered by SSL re-labeling, while still present, become less pronounced. This observation, highlighted by the yellow ellipse in Figure 26, indicates the limitations of our method in scenarios of severe label corruption. However, it's important to contextualize this finding, such extreme levels of label noise are rare in carefully collected neurophysiological datasets, and the ability of our method to provide meaningful improvements even under these conditions is noteworthy.

2. Comparison with other noise mitigation techniques

To further validate the efficacy of our SSL re-labeling approach, we conducted a comparative analysis with two recent techniques designed to address label noise in time-series data: Self-Re-Labeling with Embedding Analysis (SREA) and Confident Time-Warping (CTW).

SREA, which employs a multi-task deep learning approach combining an autoencoder, classifier, and constrained clustering module to gradually correct mislabeled samples, achieved an accuracy of 76.23% when applied to our SWR dataset (Castellani, Schmitt, and Hammer 2021). However, we observed that SREA's performance peaked early and subsequently declined, suggesting a tendency towards over-correction of labels. This behavior highlights a common challenge in label correction methods – the risk of introducing new errors while attempting to fix existing ones.

CTW uses a small-loss criterion to identify confident samples and applies time-warping augmentation, showed high volatility in test accuracy (Ma et al. 2023). The results ranged from 69.81% to 83.33%, with indications of significant overfitting as evidenced by the disparity between training and test accuracies. This volatility underscores the challenges of maintaining consistent performance across different subsets of the data, a critical consideration in neurophysiological studies where robust, generalizable results are essential.

In contrast, our SSL method demonstrated remarkable stability and consistency, maintaining an accuracy of around 84% across various experimental conditions. This performance not only surpasses that of SREA and CTW but also addresses key limitations observed in these methods. Unlike SREA, our approach appears to avoid the pitfall of over-correction, suggesting a more robust mechanism for identifying truly mislabeled instances. The stability of our results compared to CTW indicates better generalization capability and resistance to overfitting, crucial attributes when dealing with complex, variable neurophysiological data.

G. Model interpretability through Gradient-weighted Class Activation Mapping

After demonstrating the efficiency of our SSL relabeling method and the significant improvement in 1D CNN performance, we turned our attention to a crucial aspect of our research: understanding the underlying characteristics of Sharp Wave Ripples (SWRs) that

distinguish pre- and post-learning states. While achieving high classification accuracy is valuable, our ultimate goal is to gain deeper insights into the neurophysiological processes associated with learning and memory consolidation.

Interpreting complex models like Convolutional Neural Networks (CNNs) presents a significant challenge. This is particularly true when dealing with intricate neurophysiological data such as SWRs. Our 1D CNN model, while highly accurate, operates as a "black box," making decisions based on learned features that are not immediately apparent to human observers. To truly advance our understanding of SWRs and their role in learning, we need to peek inside this black box and decipher which aspects of the SWR signals are most influential in distinguishing between pre- and post-learning states.

To address this challenge and bridge the gap between machine learning performance and neuroscientific insight, we implemented Gradient-weighted Class Activation Mapping (Grad-CAM). This advanced visualization technique allows us to identify and localize the parts of the input signal that most strongly influence our model's classifications. By providing heatmap visualizations of these influential regions, Grad-CAM not only enhances our understanding of the model's decision-making process but also offers a unique opportunity to validate and potentially expand our existing knowledge of SWRs.

1. Grad-CAM implementation for 1D CNN

Adapting Grad-CAM to our 1D CNN model presented unique challenges due to the temporal nature of our data. Unlike its application in image classification tasks, our focus was on identifying temporally relevant features within the Ripple-Centered Intervals (RCIs). Our implementation computed the importance of different time points in the input signal for the model's classification decision.

Figure 27 illustrates the Grad-CAM visualization process for a single SWR sample. The top panel shows the original input signal, representing a typical SWR. The bottom panel displays the same signal overlaid with a color-coded heatmap generated by Grad-CAM. This heatmap provides a visual representation of the regions that most influenced the model's decision, with warmer colors (red and yellow) indicating areas of higher importance and cooler colors (blue) signifying less influential regions.

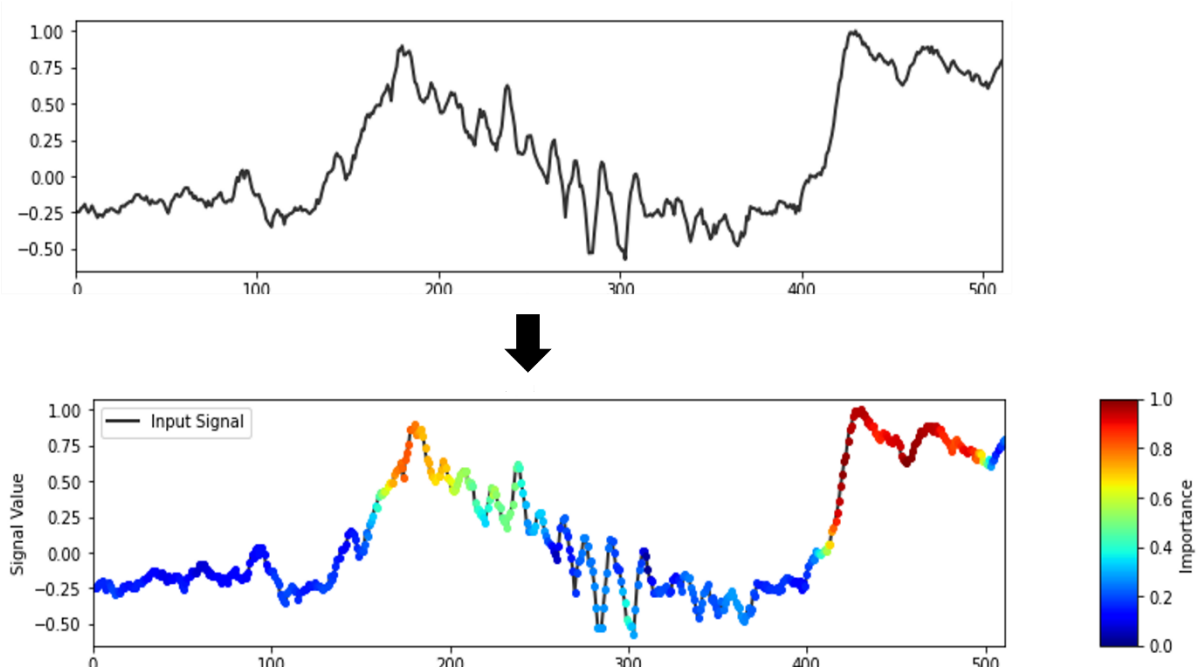

Figure 27 Grad-CAM visualization on a single SWR.

Example of the application of Grad-CAM on a single SWR. The visualization aims to highlight the regions within the 1D CNN layer that were most influential in classifying the SWR. The heatmap generated by Grad-CAM overlays the original SWR signal, providing insights into the specific features of the SWR that the deep learning model focuses on for classification. This serves as an interpretative tool for understanding the model's decision-making process.

The heatmap in Figure 27 reveals a clear pattern of activation across the SWR signal. We observe that the central portion of the signal, corresponding to the actual ripple event, shows the highest activation (red and orange colors). This indicates that the model is placing the most emphasis on the high-frequency oscillations characteristic of SWRs when making its classification decision. The regions immediately following the central ripple also show significant activation (yellow and green colors), suggesting that post-ripple activity also plays a role in the model's decision-making process. In contrast, the early portions of the signal, preceding the ripple, show minimal activation (blue colors), indicating that these regions have less influence on the model's classification.

2. Temporal localization of discriminative features

To facilitate a more nuanced understanding of the model's focus, we divided each RCI into three distinct regions: the pre-ripple, peri-ripple, and post-ripple regions. Figure 28 presents

this segmentation, with vertical dashed lines delineating the boundaries between these regions.

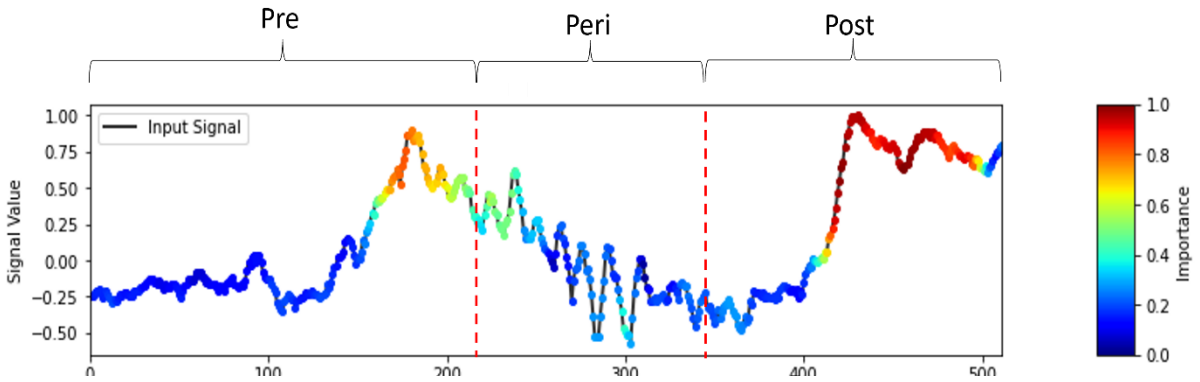

Figure 28 Segmentation of RCIs into Pre, Peri, and Post-Ripple Regions.

The division of Ripple-Centered Intervals (RCIs) into three distinct temporal segments: the 'Pre' region occurring before the ripple event, the 'Peri' region aligning with the ripple, and the 'Post' region immediately following the ripple. Each segment is delineated with markers to clearly indicate the boundaries. This segmentation aims to facilitate a nuanced analysis of the RCI components, thereby aiding in the understanding of different temporal characteristics associated with sharp wave-ripples.

The pre-ripple region encompasses the signal preceding the SWR event. In Figure 28, we observe that this region generally displays cooler colors (blues and greens), indicating lower activation in the Grad-CAM heatmap. This suggests that the model places less emphasis on the neural activity leading up to the SWR when making its classification decisions.

The peri-ripple region, which coincides with the SWR event itself, consistently shows the warmest colors (reds and oranges) in the heatmap. This observation aligns well with our neurophysiological understanding of SWRs, where the high-frequency oscillations during the ripple are thought to play a crucial role in memory consolidation and information transfer. The model's strong focus on this region validates its ability to identify the most salient features of SWRs.

The post-ripple region, following the SWR event, displays intermediate levels of activation (yellows and greens). This suggests that the neural activity immediately following the SWR contains important information for distinguishing between pre- and post-learning states.

To provide a more comprehensive view of these patterns across multiple samples, Figure 29 presents box plots of Grad-CAM activations in RCIs, separated by pre, peri, and post-ripple

regions. This figure offers a quantitative perspective on the observations made from individual heatmaps.

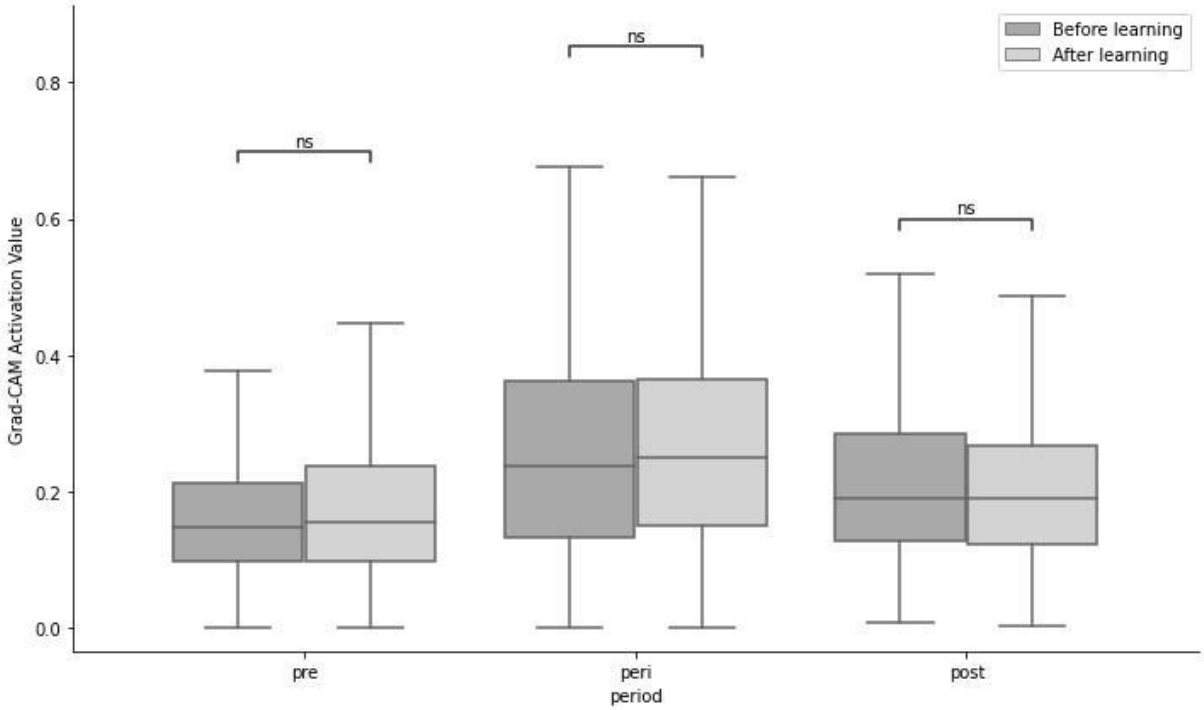

Figure 29 Grad-CAM activation values for SWRs before and after learning.

The box plot illustrates the Grad-CAM activation values for SWRs during the pre, peri, and post periods, comparing the "Before learning" (black) and "After learning" (dark gray) conditions.

The box plots in Figure 29 consistently show that the peri-ripple region has the highest median activation values and the largest interquartile ranges. This confirms our earlier observation that the model places the most emphasis on the actual ripple event when making its classification decisions. The post-ripple region generally shows the next highest activation levels, though with notable variability. This variability suggests that the importance of post-ripple activity may differ depending on specific learning experiences. The pre-ripple region consistently shows the lowest activation levels. This reinforces our earlier observation that the model considers the neural activity preceding the SWR less informative for distinguishing between pre- and post-learning states.

Statistical analysis showed no significant differences in Grad-CAM activation values between the "Before learning" and "After learning" conditions in any of the three regions. This lack of statistical difference could be because the 1D CNN model uses the same regions to make decisions, but the features within these regions differ. This observation prompted us to conduct feature extraction and feature importance analysis, which will be discussed in the next section.

H. Feature importance and impact of feature ablation study

Building on the insights gained from the Grad-CAM analysis, we turned our attention to advanced feature extraction and analysis techniques to further unravel the complex mechanisms underlying sharp wave ripple (SWR) classification before and after learning. This approach was crucial not only for enhancing the interpretability of our deep learning model but also for identifying the neurophysiological characteristics of SWRs that are most relevant for distinguishing between pre- and post-learning states. The lack of statistical difference in Grad-CAM activation values suggested that the 1D CNN model utilized the same regions for decision-making, though the features within these regions varied. This prompted a deeper investigation into the specific features contributing to the classification, which will be detailed in the following section.

1. Time-series feature extraction methodology

To extract meaningful features from our SWR data, we utilized two robust libraries: The tsfresh library (Christ et al. 2018), offers an automated approach to feature engineering in time-series data, providing 794 features derived from 63 time series characterization methods. Its capability to conduct feature selection based on statistically significant hypothesis tests ensured the inclusion of only genuinely relevant features. Complementing this, TSFEL (Barandas et al. 2020) computed over 60 features across temporal, statistical, and spectral domains, offering valuable insights into the computational complexity of the extracted features.

The integration of these libraries allowed us to capture a comprehensive set of features that encompass various aspects of the SWR signals. This approach is particularly valuable in the

context of neurophysiological data, where the underlying patterns may not be immediately apparent through visual inspection or simple statistical measures.

2. Ranking of feature importance

Following feature extraction, we utilized PyCaret (Pycaret 3.0.4) to discern the relative importance of these features. PyCaret's automated machine learning capabilities were particularly useful in quantifying feature importance for our binary classification task of categorizing SWRs before and after learning. This analysis highlighted two standout features: Fast Fourier Transform (FFT) coefficients and wavelet entropy (Figure 30).

FFT coefficients provided crucial information about the frequency domain characteristics of the SWR signals, capturing key aspects such as amplitude and phase. Specifically, FFT Coefficient 3 and FFT Coefficient 10 emerged as the most significant features, with importance scores of approximately 5.8 and 5.0 respectively.

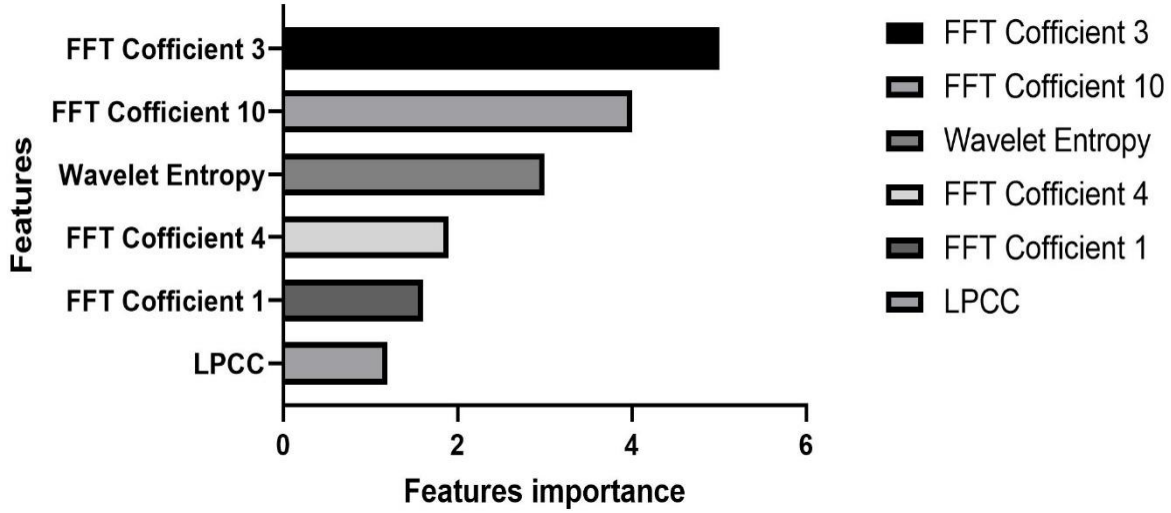

Figure 30 Feature extraction and importance analysis.

The bar graph displays the relative importance of various features. Notably, the Fourier Transform Coefficients (specifically at indices 3 and 10) and wavelet entropy emerged as the most critical features in the classification process.

Wavelet entropy, the third-ranked feature with an importance score of about 3.8, offered insights into the complexity and unpredictability of the time-series data. This measure, derived

from wavelet analysis, quantifies the degree of order/disorder in a signal, providing a unique perspective on the temporal dynamics of SWRs. The prominence of wavelet entropy in our feature importance analysis suggests that the complexity of SWR signals may be a key differentiator between before and after learning states. This observation lays the groundwork for a more detailed examination of the temporal dynamics of SWRs through Morlet wavelet analysis, which will be explored in a later section.

3. Feature ablation study

To corroborate the importance of the selected features, we conducted a feature ablation study. This involved running our model multiple times, systematically omitting one or more features in each iteration, and observing the impact on the model's performance metrics, particularly accuracy.

Our initial focus was on the two most prominent features identified in our importance analysis: FFT Coefficient 3 and FFT Coefficient 10. We created modified versions of our dataset where these specific coefficients were removed, effectively ablating these features from the input signals. The model was then retrained and evaluated on these modified datasets.

Figure 31 illustrates the effect of removing these key FFT coefficients from the SWR signals. The left panel shows an original SWR signal, while the right panel displays the same signal after the removal of FFT Coefficients 3 and 10. The visual difference between these signals is subtle, highlighting the challenge of distinguishing pre- and post-learning SWRs through simple visual inspection and underscoring the value of our deep learning approach.

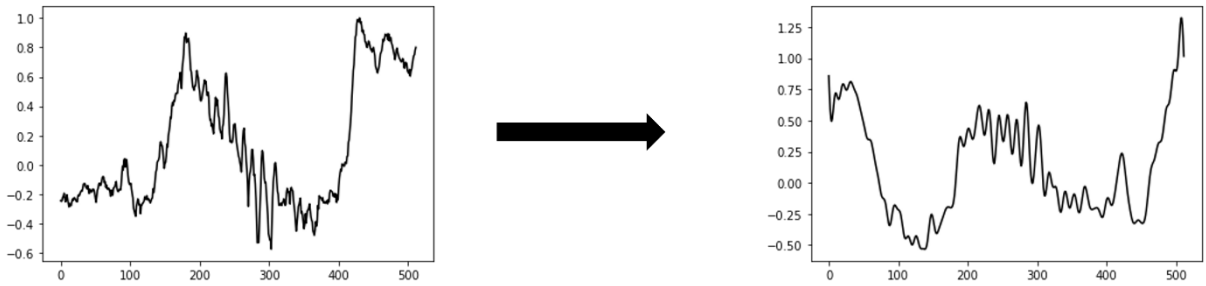

Figure 31 Example of features ablation of FFT coefficients.

Displayed are the SWR alongside version where FFT coefficients 3 and 10 have been removed.

As shown in Figure 32, removal of FFT Coefficient 3 and FFT Coefficient 10 led to a significant drop in classification accuracy, from the original 84% to 75%. This 9 percentage point decrease underscores the critical role these specific frequency components play in distinguishing between before and after learning SWRs. The magnitude of this impact suggests that these FFT coefficients capture essential information about the spectral changes that occur in SWRs as a result of learning experiences.

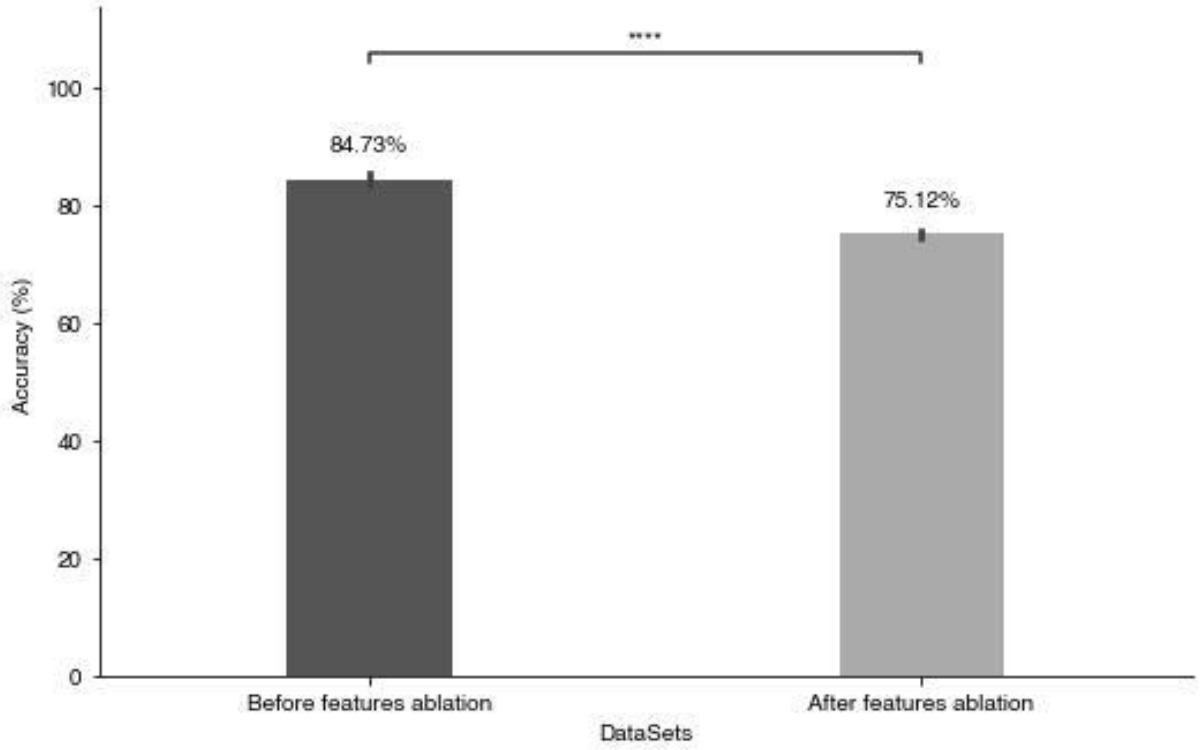

Figure 32 Drop in 1D CNN accuracy post-FFT coefficients ablation.

Impact of feature ablation on the classification accuracy of the 1D Convolutional Neural Network (CNN) model. Specifically, the graph shows a drop in accuracy from 84% to 75% following the removal of Fourier Transform Coefficients.

The feature ablation study not only validated our feature importance analysis but also provided quantitative evidence for the significance of specific spectral features in SWR classification. This approach offers a bridge between the often opaque nature of deep learning models and the need for interpretable results in neuroscientific research.

4. Impact of learning on SWR frequency bands

Building upon our previous analyses of sharp wave ripples (SWRs) and their modulation by learning, we conducted a series of statistical tests to further elucidate the changes in frequency bands associated with the learning process. These analyses provide a more nuanced understanding of the spectral characteristics of SWRs and how they are altered by learning experiences.

To investigate the changes in spectral content of SWRs before and after learning, we performed a mixed-design analysis of variance (ANOVA) on Fast Fourier Transform (FFT) coefficients across different frequency bands. This analysis allowed us to examine both the main effects of learning and frequency bands, as well as their interaction.

As illustrated in Figure 33, we categorized the FFT coefficients into six distinct frequency bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), low gamma (30-50 Hz), and middle gamma (50-100 Hz). These bands were chosen based on their established relevance in hippocampal function and memory processes (Buzsáki and Draguhn 2004).

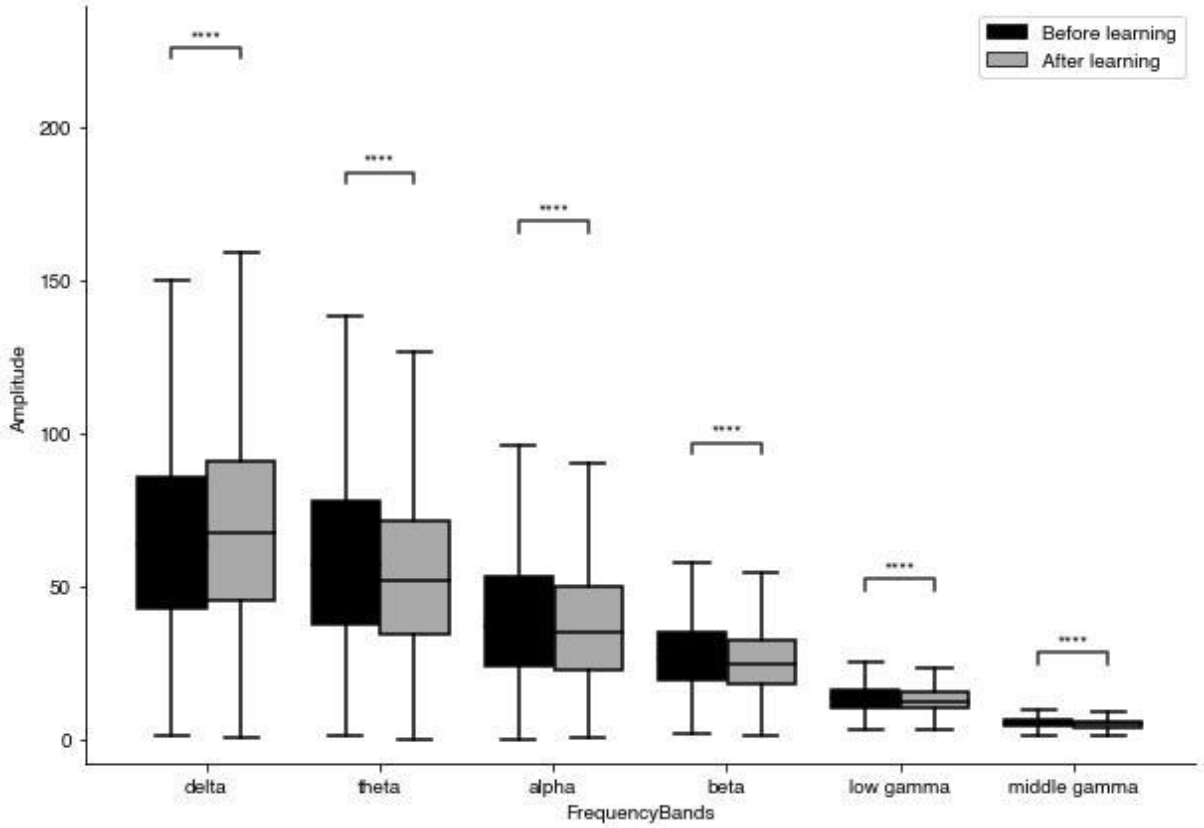

Figure 33 Comparison of power amplitude across different frequency bands.

*The box plot categorizes the FFT coefficients into six distinct frequency bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), low gamma (30-50 Hz), and middle gamma (50-100 Hz). The power amplitudes for these frequency bands are compared between 'Before learning' (black) and 'After learning' (dark gray) conditions. Statistical significance between the groups is indicated: ***p < 0.001. The results show significant differences in power amplitude across most frequency bands, reflecting changes in hippocampal activity associated with the learning process.*

The mixed-design ANOVA revealed several significant findings. First, we observed a main effect of frequency band ($F(5, 590) = 142.37$, $p < 0.001$, $\eta^2 = 0.547$), indicating substantial differences in power across the various frequency ranges.

More intriguingly, we found a significant interaction between learning stage and frequency band ($F(5, 590) = 3.82$, $p < 0.05$, $\eta^2 = 0.031$). This interaction suggests that learning alters the spectral composition of SWRs in a frequency-dependent manner. Specifically, post-hoc t-tests revealed significant decrease in power for theta, alpha, beta, and gamma bands following learning, while the delta band showed a significant increase in power.

I. Morlet wavelet decomposition analysis

Building upon our feature importance analysis, which highlighted the significance of wavelet entropy in distinguishing SWRs before and after learning, we delved deeper into the time-frequency characteristics of these neural events using Morlet wavelet decomposition. This advanced analytical technique allowed us to explore the intricate spectral dynamics of SWRs and their potential modulation by learning processes.

The transition from broad feature analysis to detailed wavelet decomposition represents a natural progression in our quest to unravel the subtle differences between before and after learning SWRs. While the feature importance analysis provided a high-level view of relevant characteristics, Morlet wavelet decomposition offers a fine-grained perspective on the time-frequency structure of these complex neural signals.

1. Wavelet transform implementation

To implement the Morlet wavelet decomposition, we utilized the continuous wavelet transform (CWT) with a Morlet wavelet as the mother wavelet. The Morlet wavelet is particularly well-suited for analyzing neural oscillations due to its optimal time-frequency resolution trade-off. Our implementation focused on a frequency range of 1-250 Hz to capture both the slow components of sharp waves and the high-frequency ripple oscillations.

Figure 34 illustrates the outcome of this analysis for two SWRs, one from before and one after the learning task. The raw SWR traces are shown alongside their corresponding Morlet wavelet transformations, revealing the distinct temporal and spectral dynamics of these neural events.

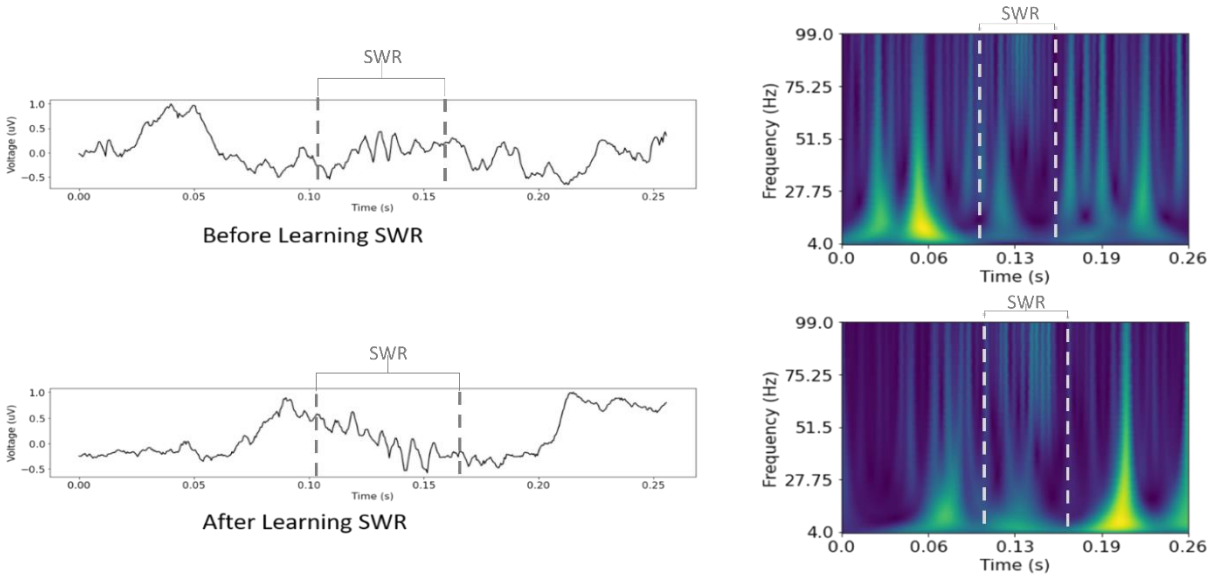

Figure 34 Morlet wavelet analysis of SWRs.

Two representative examples of Sharp Wave-Ripples (SWRs), one from before and the other after the learning classes. Morlet wavelet analysis is applied to these SWRs to explore their frequency-time characteristics. Each subplot illustrates the raw SWR trace along with their corresponding Morlet wavelet transformation, providing insights into the distinct temporal and spectral dynamics of SWRs.

This visualization provides valuable insights into the complex spectro-temporal structure of SWRs. In both before and after learning examples, subtle differences in the distribution and intensity of this power become apparent, hinting at potential learning-induced changes in SWR dynamics.

2. Spectral power distribution analysis

Our analysis of the spectral power distribution across different temporal regions of the RCIs provided insights into the potential impact of learning on SWR characteristics. We divided each RCI into three distinct regions: pre-ripple, peri-ripple, and post-ripple, aligning with our previous analyses.

To quantify these observations rigorously, we computed the mean Morlet wavelet power across the three RCI regions for both learning states, as illustrated in Figure 35. This visualization provides a clear representation of the temporal dynamics of spectral power within SWRs and how they might be modulated by learning.

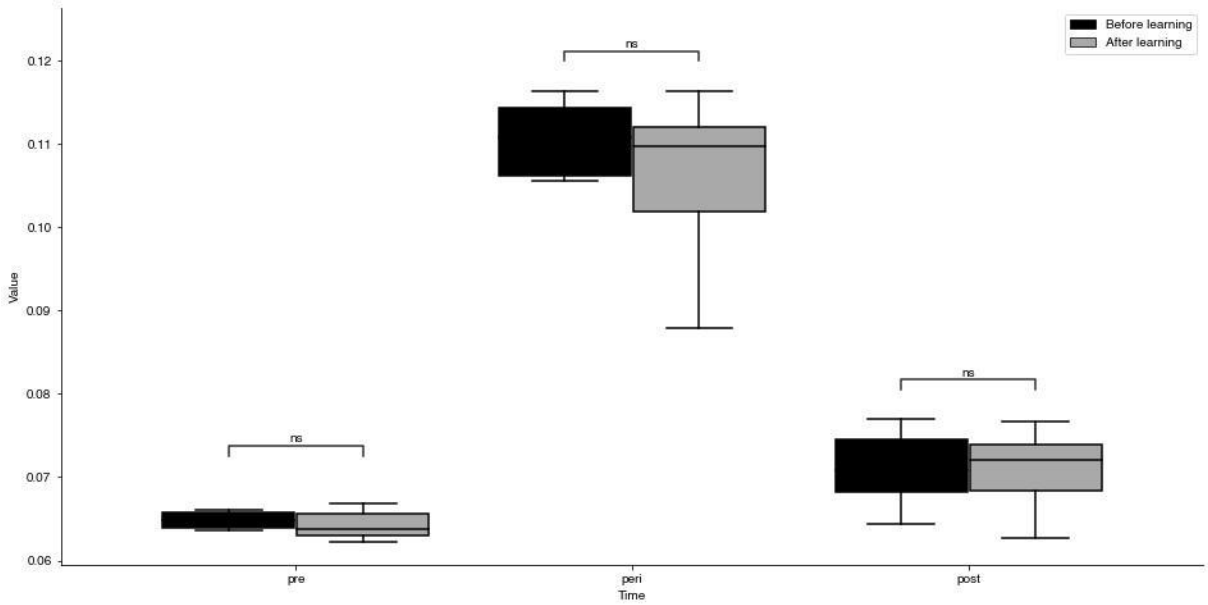

Figure 35 Morlet wavelet power distribution across periods for each class.

Mean Morlet wavelet power across the three distinct regions ('pre', 'peri', 'post') of the Ripple-Centered Intervals (RCIs) for both 'before' and 'after' learning classes.

To statistically evaluate the observed differences in spectral power distribution, we employed a mixed Analysis of Variance (ANOVA). This statistical approach allowed us to simultaneously assess the effects of time (within-subject factor: pre-, peri-, and post-ripple regions) and learning state (between-subject factor: before and after learning) on wavelet power.

The mixed ANOVA revealed a significant main effect of time on wavelet power, indicating substantial changes in spectral power across the three RCI regions.

Interestingly, we found no significant main effect of learning state and no significant interaction between time and learning state. These results suggest that while there are clear temporal dynamics within SWRs, these changes do not appear to be significantly modulated by the learning experience, at least in terms of overall wavelet power.

The absence of a significant learning effect on wavelet power distribution contrasts with our earlier findings from the feature importance analysis, which highlighted wavelet entropy as a key discriminative feature. This apparent discrepancy underscores the complex nature of learning-induced changes in SWRs and suggests that the effects of learning may manifest in more subtle ways than overall power changes.

J. Correlation analysis between Grad-CAM and Wavelet decomposition

In our pursuit of a comprehensive understanding of sharp wave ripple (SWR) classification and its neurophysiological underpinnings, we conducted a correlation analysis between the Gradient-weighted Class Activation Mapping (Grad-CAM) outputs and the results of our Morlet wavelet decomposition. This analysis aimed to bridge the gap between our deep learning model's decision-making process and the established time-frequency characteristics of SWRs. Among our findings, the most captivating result is the significant correlation between the Grad-CAM heatmaps and the regions of highest Morlet wavelet power. This alignment underscores the relevance of the model's focus on specific SWR features and validates our approach in linking machine learning insights with neurophysiological phenomena.

1. Methodology for correlation assessment

To assess the correlation between Grad-CAM activations and high-power regions identified through wavelet analysis, we developed a novel methodological approach. First, we generated Grad-CAM heatmaps for each SWR in our dataset, using the technique described by Selvaraju et al. (Selvaraju et al. 2020). These heatmaps provided a visual representation of the regions within each SWR that were most influential in the model's classification decision.

Concurrently, we identified the regions of highest Morlet wavelet power for each SWR. We defined these high-power regions as those exceeding a threshold of 50% of the maximum wavelet power observed within each SWR. This threshold was chosen to capture the most significant spectral components while excluding background noise.

To quantify the correlation, we developed an overlap metric. This metric calculated the percentage of Grad-CAM activation above a certain threshold (set at 0.7 in our analysis) that coincided with the high-power regions identified through wavelet analysis. We computed this metric separately for SWRs from the 'before learning' and 'after learning' classes to investigate potential learning-induced changes in the correlation patterns.

Figure 36 provides a visual representation of this correlation analysis for two SWRs, one from the 'before learning' class and one from the 'after learning' class. The figure juxtaposes the Grad-CAM heatmaps with the corresponding Morlet wavelet power plots, allowing for a qualitative assessment of the overlap between regions of high Grad-CAM activation and high wavelet power.

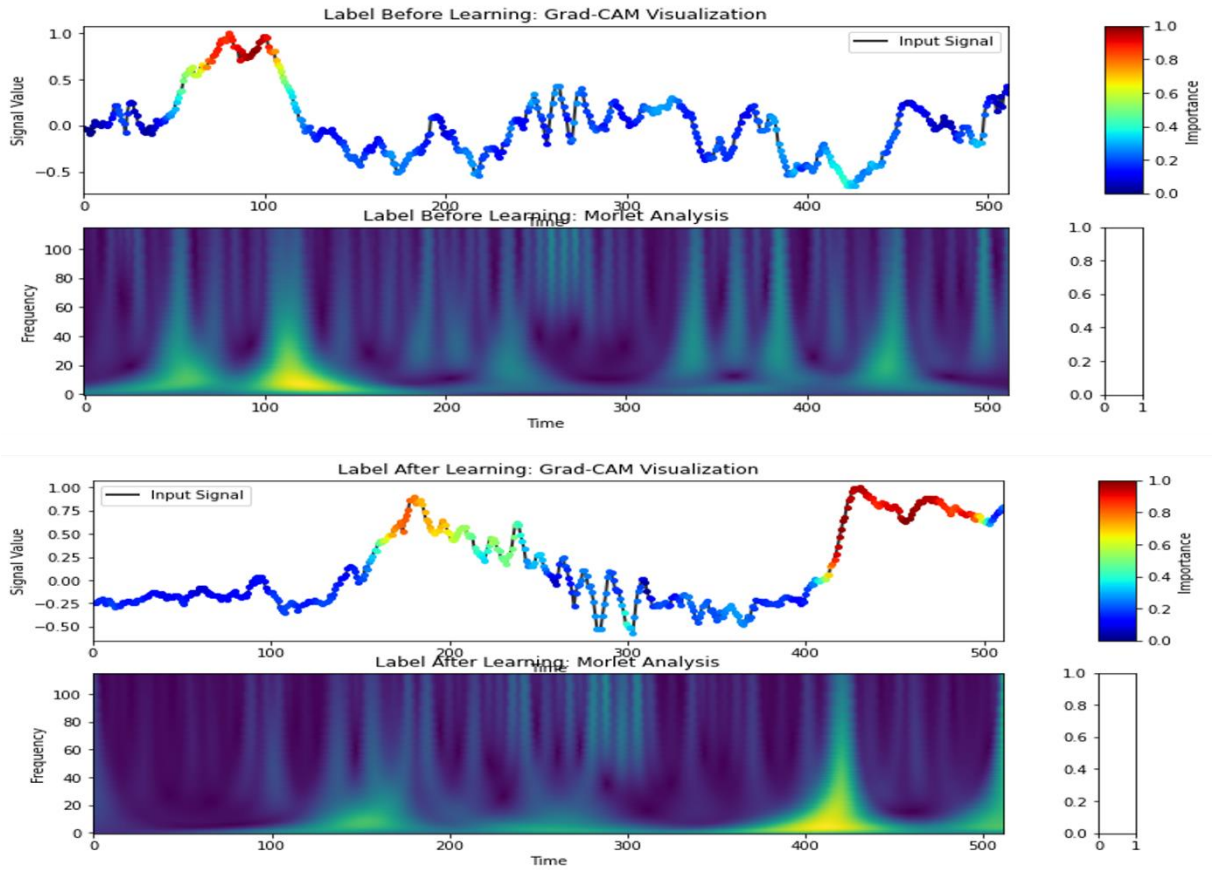

Figure 36 Examples of correlation between Grad-CAM and Morlet analysis

Two representative examples showing the correlation between Grad-CAM heatmaps and regions of highest Morlet wavelet power, one from 'before learning' and another from 'after learning' sessions. The Grad-CAM heatmaps are aligned alongside the Morlet wavelet power plots for each of the two cases.

2. Quantification of correlation strength

The results of our correlation analysis revealed a striking correspondence between the regions of high Grad-CAM activation and those of high Morlet wavelet power. As illustrated in Figure 37, we observed a robust correlation in both the 'before learning' and 'after learning' classes.

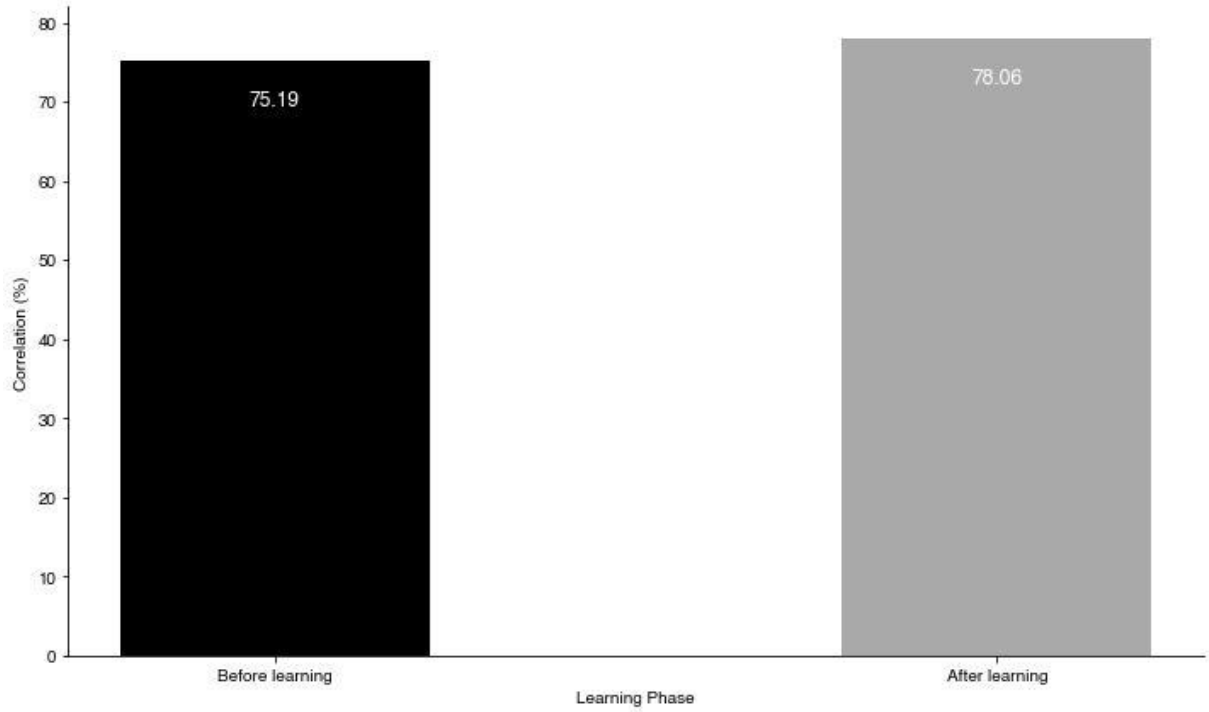

Figure 37 Correlation between Grad-CAM and Morlet wavelet analysis.

The bar graph shows the correlation between the regions of high Grad-CAM activation and high Morlet wavelet power for SWRs in the 'before learning' and 'after learning' classes.

Specifically, for SWRs in the 'before learning' class, we found that 75.19% of the high Grad-CAM activation regions (above the 0.7 threshold) overlapped with the high-power regions identified through wavelet analysis. This strong correlation suggests that our model was predominantly focusing on the spectral components that are characteristic of SWRs when making its classification decisions.

Interestingly, the correlation was even stronger for SWRs in the 'after learning' class, with 78.06% overlap between high Grad-CAM activation and high wavelet power regions. This slight increase in correlation post-learning could indicate that learning experiences lead to more pronounced or consistent spectral features in SWRs, which our model was able to capture.

The strong correlation between Grad-CAM activations and regions of high Morlet wavelet power has profound implications for the validity of our deep learning model and its neurophysiological relevance. This correlation provides strong evidence that our model is indeed focusing on physiologically relevant aspects of the SWRs when making its classification decisions.

K. Comparative analysis of wild-type and transgenic Alzheimer's model animals

In this section, we extend our investigation to include transgenic (TG) Alzheimer's disease model animals, providing a comparative analysis with wild-type (WT) animals. This comparison offers valuable insights into how Alzheimer's pathology affects the neural dynamics underlying learning and memory processes, particularly in the context of sharp wave ripples (SWRs).

1. Classification performance in TG vs WT animals

Our analysis of SWR classification performance revealed striking differences between WT and TG animals. As illustrated in Figure 38, the 1D CNN classifier achieved an initial accuracy of 58.47% in TG animals, which increased marginally to 63.77% after self-supervised learning re-labeling. This performance stands in stark contrast to the high accuracy observed in WT animals, where we achieved 84% accuracy using the same methodological approach.

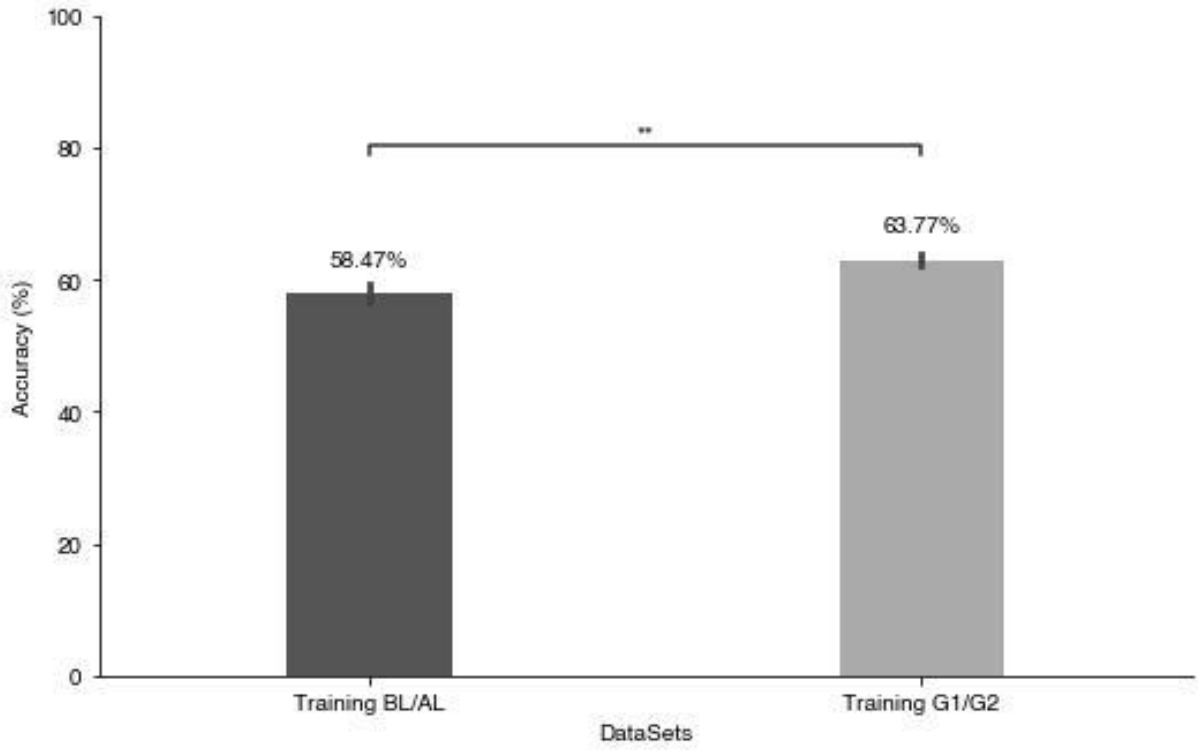

Figure 38 Comparison of 1D CNN model accuracy on SWRs from transgenic animals.

Bars plot that quantitatively compares the accuracy of two deep learning models—1D CNN and 1D CNN combined with SSL relabeling when applied to SWRs recorded from transgenic animals.

The reduced classification accuracy in TG animals suggests a fundamental alteration in the distinguishability of pre- and post-learning SWRs in the Alzheimer's disease model. This finding aligns with previous studies demonstrating disrupted hippocampal function in Alzheimer's models (Palop et al. 2007). The inability of our model to effectively differentiate between pre- and post-learning SWRs in TG animals may reflect a broader impairment in the formation or consolidation of new memories, a hallmark of Alzheimer's disease.

Moreover, the limited improvement in accuracy following self-supervised learning in TG animals (from 58% to 63%) contrasts sharply with the substantial gains observed in WT animals. This discrepancy suggests that the inherent structure and variability of SWRs in TG animals may be fundamentally altered, limiting the effectiveness of our machine learning approach in extracting meaningful patterns associated with learning.

2. Differential analysis of SWR characteristics

Despite the marked differences in classification performance, our analysis of FFT coefficients revealed intriguing similarities between WT and TG animals. As shown in Figure 39, significant differences in spectral power across frequency bands before and after learning were observed in TG animals, mirroring the pattern seen in WT animals. This preservation of frequency-dependent modulation suggests that some aspects of neural plasticity associated with learning remain intact in the Alzheimer's model.

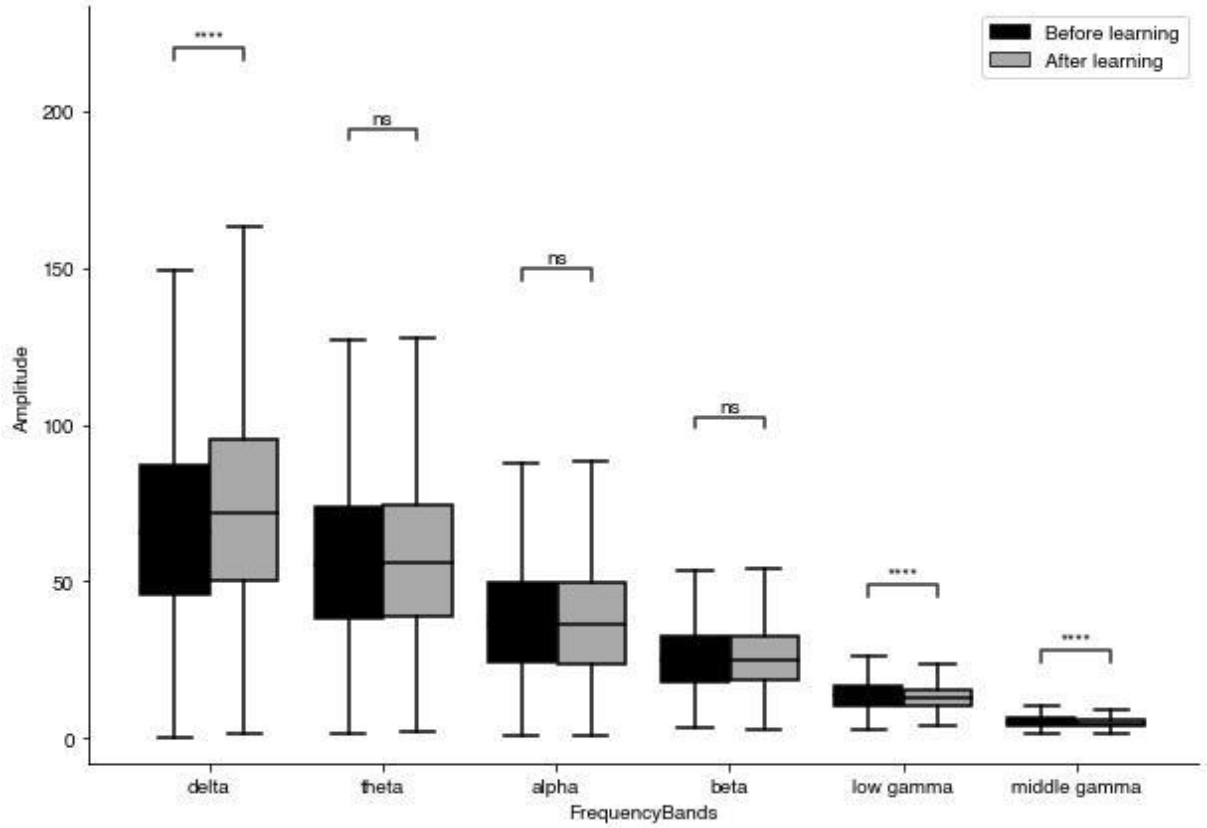

Figure 39 FFT coefficients across frequency bands before and after learning in TG animals.

Statistical analysis of Fast Fourier Transform (FFT) coefficients, categorized by different frequency bands (delta, theta, alpha, beta, low gamma, and middle gamma), before and after learning sessions in TG animals.

However, a crucial distinction emerged in our wavelet entropy analysis. While WT animals exhibited significant changes in wavelet entropy following learning, TG animals showed no such modulation. This absence of learning-induced changes in signal complexity, as measured by wavelet entropy, points to a specific disruption in the information processing capabilities of the hippocampus in Alzheimer's disease.

The divergence in wavelet entropy results between WT and TG animals is particularly noteworthy. It suggests that while basic frequency-dependent neural mechanisms may be relatively preserved in Alzheimer's disease, more complex aspects of neural information processing are disrupted.

IV. Discussion

A. Overview of key findings

The application of advanced machine learning techniques, particularly self-supervised learning (SSL), has yielded significant advancements in the classification and characterization of sharp wave ripples (SWRs) in the context of spatial learning. Our research has demonstrated a marked improvement in SWR classification accuracy, elevating it from 73.28% to 84.11% through the implementation of a 1D Convolutional Neural Network (CNN) model coupled with SSL-based relabeling. This substantial enhancement in accuracy underscores the potential of SSL in addressing the persistent challenge of label noise in neurophysiological data (Frénay and Kabán 2014).

The success of our approach can be attributed to the SSL model's ability to leverage the inherent structure of unlabeled data, thereby refining the labeling process and mitigating the impact of noise. This methodology aligns with recent advancements in self-supervised learning that emphasize the importance of learning temporally consistent features in time-series data (Zhang et al. 2024). By employing a combination of weak and strong augmentation techniques, our SSL framework encouraged the model to learn robust and invariant representations of SWRs, capturing both local temporal shifts and global, order-invariant features.

A critical outcome of our research was the identification of key features that distinguish pre- and post-learning SWRs. Through rigorous feature extraction and importance analysis, we discovered that Fast Fourier Transform (FFT) coefficients and wavelet entropy emerged as the most significant features for SWR classification. The prominence of FFT coefficients, particularly FFT Coefficient 3 and FFT Coefficient 10, underscores the importance of specific frequency components in characterizing SWRs. This finding aligns with previous studies that have reported learning-induced alterations in the spectral content of hippocampal oscillations (Tort et al. 2009).

The identification of wavelet entropy as a key feature provides a complementary perspective on the neural changes associated with learning. As a measure of signal complexity, wavelet entropy captures subtle alterations in the temporal organization of SWRs that may not be immediately apparent from spectral analysis alone. The significance of this feature suggests that learning experiences may lead to changes in the overall complexity or regularity of SWR

patterns, potentially reflecting alterations in the coordination of neuronal firing patterns within the hippocampal circuit.

These findings demonstrate the power of advanced machine learning techniques in uncovering subtle patterns in complex neural data. The improvement in classification accuracy and the identification of key features open new avenues for investigating the role of SWRs in memory processes and their potential as biomarkers for cognitive function and dysfunction.

B. Methodological advancements in SWR analysis

The field of sharp wave ripple (SWR) analysis has undergone a significant transformation in recent years, marked by a shift from traditional threshold-based methods to more sophisticated deep learning approaches. This transition represents a paradigm shift in how we approach the detection and classification of these complex neural events. Traditional methods, which often relied on simple amplitude thresholds or basic spectral criteria, were limited in their ability to capture the full complexity and variability of SWRs across different experimental conditions and subjects (Buzsáki 2015). In contrast, deep learning models, particularly convolutional neural networks (CNNs), have demonstrated a remarkable ability to learn and leverage intricate spatiotemporal features directly from raw neurophysiological data (Navas-Olive et al. 2023).

Our implementation of a 1D CNN model for SWR classification builds upon recent advancements in the application of deep learning to time-series data analysis. The success of CNNs in this domain can be attributed to their ability to automatically learn hierarchical representations of input data, capturing both local and global patterns that may be challenging to define manually (Hsu et al. 2021). This approach aligns with the complex nature of SWRs, which exhibit variations in frequency, amplitude, and temporal structure that can be subtle yet highly informative.

The introduction of self-supervised learning (SSL) techniques in our methodology marks a significant advancement in addressing one of the most persistent challenges in neurophysiological data analysis: label noise. Label noise, which refers to inaccuracies or inconsistencies in the assigned labels of data points, can severely impact the performance of machine learning models, particularly in the context of complex neural signals (Frénay and Kabán 2014). Our SSL approach, inspired by the Time-Series Representation Learning via

Temporal and Contextual Contrasting (TS-TCC) framework (Eldele et al. 2021), offers a novel solution to this problem by leveraging the intrinsic structure of the data itself to generate more reliable labels.

The key innovation in our SSL methodology lies in its ability to learn meaningful representations from unlabeled data through carefully designed pretext tasks. By employing both temporal and contextual contrasting modules, our model learns to capture both the fine-scale temporal dynamics of SWRs and their broader contextual relationships. The temporal contrasting module, which implements a cross-view prediction task, encourages the model to learn representations that are robust against perturbations introduced by different augmentations and timesteps. This approach is particularly well-suited to the analysis of SWRs, which exhibit complex temporal patterns that may vary subtly between pre- and post-learning states.

Complementing this, the contextual contrasting module enhances the discriminative power of the learned representations by maximizing similarity between contexts derived from different augmented views of the same sample while minimizing similarity with contexts from different samples. This dual-pronged approach to feature learning – focusing on both temporal coherence and contextual distinctiveness – was instrumental in capturing the nuanced differences between before and after learning SWRs, ultimately leading to the significant improvement in classification accuracy from 73.28% to 84.11%.

The effectiveness of our SSL approach in mitigating label noise is further evidenced by its performance on synthetic datasets with controlled levels of noise. Our experiments demonstrated that the SSL-enhanced model maintained high classification accuracies (between 82% and 85%) even with moderate levels of label noise (10% to 20%), nearly matching the performance observed under ideal, noise-free conditions. This robustness is particularly significant in the context of neurophysiological data, where moderate levels of label noise are common due to factors such as inter-rater variability and the inherent complexity of neural signals (Atkinson and Metsis 2020).

To contextualize the performance of our SSL approach, we conducted a comparative analysis against other recent techniques designed to address label noise in time-series data, specifically Self-Re-Labeling with Embedding Analysis (SREA) and Confident Time-Warping (CTW). SREA, which employs a multi-task deep learning approach combining an autoencoder, classifier, and constrained clustering module, achieved an accuracy of 76.23% when applied to our SWR

dataset (Castellani, et al 2021). However, we observed that SREA's performance peaked early and subsequently declined, suggesting a tendency towards over-correction of labels.

CTW, which uses a small-loss criterion to identify confident samples and applies time-warping augmentation, showed high volatility in test accuracy, ranging from 69.81% to 83.33% (Ma et al. 2023). This volatility underscores the challenges of maintaining consistent performance across different subsets of the data, a critical consideration in neurophysiological studies where robust, generalizable results are essential.

In contrast, our SSL method demonstrated remarkable stability and consistency, maintaining an accuracy of around 84% across various experimental conditions. This performance not only surpasses that of SREA and CTW but also addresses key limitations observed in these methods. Unlike SREA, our approach appears to avoid the pitfall of over-correction, suggesting a more robust mechanism for identifying truly mislabeled instances. The stability of our results compared to CTW indicates better generalization capability and resistance to overfitting, crucial attributes when dealing with complex, variable neurophysiological data.

The success of our SSL approach in improving SWR classification accuracy has broader implications for the field of neuroscience. It demonstrates the potential of unsupervised learning techniques to extract meaningful features from complex neural data, even in the presence of significant label noise. This is particularly relevant in the context of electrophysiological recordings, where the ground truth for neural states is often ambiguous or difficult to establish (Navas-Olive et al. 2022). By providing a more reliable method for labeling and classifying neural events, our approach opens new avenues for investigating the subtle changes in neural activity associated with learning and memory processes.

C. The transformative potential of AI and deep learning in neuroscience

The integration of artificial intelligence (AI) and deep learning techniques into neuroscience research represents a paradigm shift in our ability to analyze and interpret complex neural data. The transformative potential of these approaches lies in their capacity to uncover intricate patterns and relationships within neurophysiological signals that often elude traditional analysis methods. Our work on sharp wave ripple (SWR) classification exemplifies this potential, demonstrating how deep learning can reveal subtle yet significant changes in neural activity associated with learning and memory processes.

Deep learning models, particularly convolutional neural networks (CNNs), have shown remarkable success in capturing the complex spatiotemporal dynamics of neural signals. Unlike traditional analysis methods that often rely on predefined features or simplified statistical models, CNNs can automatically learn hierarchical representations directly from raw data (Glaser et al. 2020). This ability to extract relevant features without explicit programming is particularly valuable in neuroscience, where the underlying patterns in neural activity may be too complex or subtle for human experts to fully characterize.

Our implementation of a 1D CNN for SWR classification, combined with self-supervised learning (SSL), showcases the power of this approach. By achieving an improvement in classification accuracy, our model demonstrates its ability to capture nuanced differences in SWR characteristics. This improvement is not merely a statistical achievement but represents a significant advancement in our ability to differentiate between neural states associated with different stages of learning.

The capacity of AI to handle the immense complexity and volume of modern neuroscience data is another key aspect of its transformative potential. As neuroscience technologies continue to advance, enabling simultaneous recordings from thousands of neurons across multiple brain regions, the volume and dimensionality of neural data have grown exponentially. Traditional analysis methods often struggle to cope with this data deluge, leading to potential loss of valuable information. Deep learning models, in contrast, thrive on large datasets, with their performance often improving as the amount of available data increases (Wen et al. 2020).

Our SSL approach exemplifies how AI can leverage large amounts of unlabeled data to improve model performance. By learning meaningful representations from the entire dataset, regardless of labels, our model was able to mitigate the effects of label noise and extract more robust features. This ability to learn from unlabeled data is particularly valuable in neuroscience, where obtaining accurate labels for neural states can be challenging and time-consuming.

Furthermore, the generalizability of our SSL approach suggests its potential applicability to a wide range of neurophysiological data analysis tasks. This versatility could prove invaluable in neuroscience applications, where the variability in neural signals across patients and conditions often presents significant challenges for traditional analysis methods. By providing a more robust and flexible framework for analyzing complex neural data, AI and deep learning

approaches have the potential to accelerate the translation of neuroscientific discoveries into clinical practice.

D. Challenges and considerations in applying deep learning to neuroscience

While the application of deep learning to neuroscience has yielded remarkable insights, it also presents significant challenges and considerations that must be carefully addressed. One of the primary concerns in utilizing deep learning models for neuroscientific research is their often opaque decision-making processes, commonly referred to as the "black box" problem. This lack of transparency can be particularly problematic in a field where understanding the underlying mechanisms is as crucial as achieving high performance.

To address this challenge, we implemented Gradient-weighted Class Activation Mapping (Grad-CAM) in our study, a technique that provides visual explanations for the decisions made by convolutional neural networks (Selvaraju et al. 2020). Our application of Grad-CAM to the 1D CNN model used for sharp wave ripple (SWR) classification revealed crucial insights into the specific features and temporal regions that the model deemed most influential in distinguishing between before and after learning states. The Grad-CAM analysis consistently highlighted the central portion of the SWR signal, corresponding to the actual ripple event, as the area of highest activation.

Interestingly, our Grad-CAM analysis also revealed significant activation in the regions immediately following the central ripple, suggesting that post-ripple activity plays a crucial role in the model's decision-making process. This observation provides new insights into the temporal dynamics of SWRs and their potential role in memory consolidation, highlighting the ability of deep learning models to capture subtle yet significant aspects of neural activity that might be overlooked by traditional analysis methods.

To further validate the biological relevance of our model's focus, we conducted a correlation analysis between the Grad-CAM activations and the results of Morlet wavelet decomposition. This analysis revealed a striking correspondence between regions of high Grad-CAM activation and those of high Morlet wavelet power, with overlap percentages of 75.19% for before learning SWRs and 78.06% for after learning SWRs. This strong correlation not only validates the physiological relevance of our model's attention but also demonstrates how deep learning models can align with established analytical techniques in neuroscience.

Despite these advancements in model interpretability, challenges remain in striking the right balance between model complexity and biological relevance. While more complex models may achieve higher classification accuracy, they risk overfitting to noise in the data or learning features that are not biologically meaningful. Conversely, overly simplistic models may fail to capture the intricate dynamics of neural activity. Our approach of combining a relatively simple 1D CNN architecture with sophisticated SSL techniques represents an attempt to navigate this trade-off, achieving high performance while maintaining interpretability and biological relevance.

The ethical considerations surrounding the use of AI in neuroscience research and potential clinical applications are numerous and complex. As deep learning models become increasingly capable of analyzing and interpreting neural data, questions arise about privacy, consent, and the potential for misuse of this information. For instance, the ability to detect subtle changes in SWR characteristics associated with learning or cognitive states raises concerns about the protection of cognitive privacy and the potential for unintended inferences about an individual's mental state or cognitive abilities.

Moreover, as AI-based diagnostic tools for neurological disorders become more prevalent, there is a need to ensure their reliability, fairness, and transparency. The potential for bias in AI models, whether due to imbalances in training data or inherent algorithmic biases, could lead to disparities in diagnosis and treatment if not carefully addressed. Our work on mitigating label noise through SSL techniques represents one approach to improving the reliability of AI models in neuroscience, but ongoing vigilance and research into fairness and bias mitigation strategies are crucial.

Another ethical consideration is the responsible communication of AI-derived insights in neuroscience. Given the complex and often probabilistic nature of deep learning models, there is a risk of overinterpretation or misrepresentation of results, particularly when translating research findings to clinical applications or public discourse. Researchers and clinicians must strive to communicate the capabilities and limitations of AI models accurately, ensuring that stakeholders, including patients and policymakers, have a realistic understanding of what these technologies can and cannot do.

E. Application to Alzheimer's disease model

Our investigation into the classification of sharp wave ripples (SWRs) extends beyond normal cognitive function to include a transgenic (TG) mouse model of Alzheimer's disease (AD). This comparative analysis between wild-type (WT) and TG animals provides valuable insights into how AD pathology affects the neural dynamics underlying learning and memory processes, particularly in the context of SWRs.

The most striking finding from our analysis is the marked difference in classification performance between WT and TG animals. While our 1D CNN classifier achieved an impressive accuracy of 84% in WT animals after self-supervised learning (SSL) re-labeling, the performance in TG animals was significantly lower. Initially, the classifier achieved an accuracy of only 58.47% in TG animals, which increased marginally to 63.77% after SSL re-labeling. This substantial disparity in classification accuracy suggests a fundamental alteration in the distinguishability of pre- and post-learning SWRs in the AD model.

This reduced classification accuracy in TG animals aligns with previous studies demonstrating disrupted hippocampal function in AD models. For instance, Palop et al. (Palop et al. 2007) reported aberrant excitatory neuronal activity and compensatory inhibitory mechanisms in the hippocampal circuits of AD mouse models. Our findings extend these observations by suggesting that AD pathology not only affects general hippocampal function but also specifically impairs the learning-induced modulation of SWRs.

The limited improvement in accuracy following SSL in TG animals (from 58% to 63%) contrasts sharply with the substantial gains observed in WT animals. This discrepancy suggests that the inherent structure and variability of SWRs in TG animals may be fundamentally altered, limiting the effectiveness of our machine learning approach in extracting meaningful patterns associated with learning. This observation is consistent with the hypothesis that AD pathology disrupts the fine-scale temporal dynamics of hippocampal oscillations, potentially interfering with normal learning and memory processes (Goutagny and Krantic 2013).

Despite the marked differences in classification performance, our analysis of FFT coefficients revealed intriguing similarities between WT and TG animals. Significant differences in spectral power across frequency bands before and after learning were observed in TG animals, mirroring the pattern seen in WT animals. This preservation of frequency-dependent

modulation suggests that some aspects of neural plasticity associated with learning remain intact in the AD model, at least at the stage of pathology represented in our TG animals.

However, a crucial distinction emerged in our wavelet entropy analysis. While WT animals exhibited significant changes in wavelet entropy following learning, TG animals showed no such modulation. Wavelet entropy, as a measure of signal complexity, provides insights into the overall organization and variability of neural oscillations. The absence of learning-induced changes in wavelet entropy in TG animals points to a specific disruption in the information processing capabilities of the hippocampus in AD.

This divergence in wavelet entropy results between WT and TG animals is particularly noteworthy. It suggests that while basic frequency-dependent neural mechanisms may be relatively preserved in AD, more complex aspects of neural information processing are disrupted. This finding aligns with the hypothesis that AD pathology preferentially affects higher-order cognitive functions while sparing some fundamental neural processes (Selkoe 2002).

The distinct patterns of SWR characteristics observed in our TG AD model have significant implications for the development of early diagnostic tools and novel treatment strategies for Alzheimer's disease. The ability of our deep learning model to detect subtle differences in SWR properties between WT and TG animals suggests that SWR analysis could serve as a sensitive biomarker for AD-related neural dysfunction.

F. Limitations of the current study

Despite the significant advancements our study presents in SWR classification, it is important to acknowledge several limitations. Firstly, the generalizability of our results across species and different learning paradigms remains a key concern. Our focus on rodent SWRs during a spatial learning task may not fully capture the variability in SWR characteristics across species and cognitive domains. Human SWRs, for instance, often recorded from intracranial electrodes in epilepsy patients, may exhibit different spectral or temporal properties compared to rodent SWRs (Jiang, et al. 2019). Future studies should validate our approach across a broader range of species and learning paradigms to establish its wider applicability.

The specificity of our SSL model's learned features to our particular experimental setup and recording conditions presents another limitation. Different electrode configurations, brain

states, or behavioral contexts could potentially alter the most salient features for SWR classification. This necessitates caution when applying our trained model to data collected under significantly different experimental conditions, emphasizing the need for appropriate validation and potential retraining.

From a computational standpoint, while our SSL approach significantly improved classification accuracy, it also increased computational complexity compared to traditional methods. The training of deep learning models, especially those involving SSL techniques, demands substantial computational resources and time. This could pose challenges for real-time applications or for researchers with limited access to high-performance computing facilities.

The interpretability of deep learning models remains an ongoing challenge in the field, despite the valuable insights provided by our Grad-CAM analysis. The complex, non-linear transformations within these models can make it difficult to directly relate their learned features to underlying neurophysiological mechanisms. This highlights the need for further development of interpretability techniques specifically tailored to time-series data and neurophysiological signals.

Lastly, while our improved classification method enhances our ability to distinguish and characterize SWRs, fundamental questions about SWR function and generation mechanisms persist. The precise contribution of different hippocampal subregions and cell types to SWR generation remains a subject of active research. Similarly, the mechanisms by which SWRs support memory consolidation and retrieval at the network level are not fully understood. Our improved classification method provides a powerful tool for investigating these questions, but it does not directly address the underlying physiological processes.

G. Future directions

The success of our self-supervised learning (SSL) approach in improving the classification of sharp wave ripples (SWRs) opens up exciting avenues for future research in neuroscience. One promising direction is the application of our methodology to other oscillatory phenomena in the brain. For instance, theta oscillations (4-8 Hz), crucial in spatial navigation and memory formation (Colgin 2013), could benefit from our SSL-enhanced classification technique. This could potentially reveal subtle changes in these oscillations associated with different cognitive states or stages of memory processing. Similarly, our approach could uncover new insights into

how gamma oscillations (30-100 Hz), implicated in attention and information integration (Fries, 2009), contribute to cognitive functions.

Extending our method to these other oscillatory phenomena could provide a more comprehensive understanding of how different brain rhythms interact and coordinate to support complex cognitive processes. Our SSL approach, with its ability to capture subtle features in time series data, could be adapted to analyze complex interactions such as cross-frequency coupling between theta and gamma oscillations in memory formation (Tort et al. 2009).

Integration of our SSL-based classification approach with other neuroimaging techniques presents another exciting future direction. Combining our method with functional magnetic resonance imaging (fMRI) could provide a powerful tool for linking oscillatory activity to broader patterns of brain activation. Simultaneous EEG-fMRI recordings have shown promise in elucidating the relationship between SWRs and large-scale network dynamics (Logothetis et al. 2012). Applying our improved classification technique to such multimodal datasets could potentially identify more precise correlations between specific SWR characteristics and patterns of activation.

The development of real-time SWR detection and modulation techniques based on our improved classification method is particularly promising. By implementing our SSL-enhanced classification algorithm in a real-time processing pipeline, we could potentially develop more accurate and responsive closed-loop systems compared to current systems that often rely on simple threshold-based detection methods (Fernández-Ruiz et al. 2019). Such systems could be used to study the causal role of SWRs in memory consolidation by selectively enhancing or suppressing specific subclasses of SWRs based on their learned features.

Future work should also focus on optimizing the computational efficiency of our approach, potentially through model compression techniques or the development of more streamlined SSL architectures. This could address the challenges posed by the increased computational complexity of our method, making it more accessible for real-time applications and researchers with limited computational resources.

V. Conclusion

This thesis has presented a comprehensive investigation into the classification and characterization of sharp wave ripples (SWRs) in the context of spatial learning, leveraging advanced machine learning techniques, particularly self-supervised learning (SSL). Our research has made significant advancements in enhancing the accuracy of SWR classification and providing novel insights into the neurophysiological changes associated with learning. The application of advanced machine learning techniques to neuronal data has opened new avenues for understanding the intricate workings of the brain.

The success of our self-supervised learning approach in improving SWR classification accuracy demonstrates the immense potential of AI in neuroscience research. It highlights how machine learning can uncover patterns and relationships in neural data that may not be immediately apparent through traditional analysis methods. This synergy between AI and neuroscience is not just a methodological advancement; it represents a paradigm shift in how we approach the study of brain function.

However, as we push the boundaries of what is possible with these new tools, we must remain cognizant of the challenges and limitations they present. The interpretability of complex models and the biological relevance of machine-learned features remain ongoing concerns. We must strive to bridge the gap between computational insights and neurophysiological understanding, ensuring that our technological advancements translate into meaningful progress in brain science.

The extension of our work to a model of Alzheimer's disease underscores the potential clinical implications of this research. It reminds us that beyond the pursuit of fundamental knowledge, our efforts have the potential to impact real-world medical challenges. The development of novel biomarkers and diagnostic tools based on SWR analysis could reshape how we approach neurological disorders, potentially enabling earlier intervention and more effective treatments.

Looking forward, the field of computational neuroscience stands on the brink of exciting possibilities. The integration of multiple data modalities, the application of even more sophisticated AI algorithms, and the extension of these techniques to human studies promise to yield ever deeper insights into brain function. As we continue this journey, it is crucial to

foster interdisciplinary collaboration, bringing together expertise from neuroscience, computer science, mathematics, and clinical medicine.

In reflecting on this work, we are reminded of the profound complexity of the brain and the enormity of the task that lies ahead in fully understanding its workings. Yet, it is precisely this complexity that makes the pursuit so compelling. Each new discovery, each refined method, brings us a step closer to unraveling the mysteries of memory, cognition, and consciousness itself.

As we close this chapter of research, we open the door to countless new questions and possibilities. The study of SWRs, enhanced by machine learning, serves as a microcosm of the broader endeavor to understand the brain. It exemplifies how the convergence of diverse scientific disciplines can lead to breakthroughs in our understanding of the most complex organ in the known universe.

VI. References

Abadi, Martín, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. 2016. "TensorFlow: A System for Large-Scale Machine Learning." arXiv. https://doi.org/10.48550/arXiv.1605.08695.

Amaral, David, and Pierre Lavenex. 2007. "Hippocampal Neuroanatomy." In *The Hippocampus Book*, 37–114. New York, NY, US: Oxford University Press.

Atkinson, Gentry, and Vangelis Metsis. 2020. "Identifying Label Noise in Time-Series Datasets." In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, 238–43. UbiComp/ISWC '20 Adjunct. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3410530.3414366.

Axmacher, Nikolai, Michael X. Cohen, Juergen Fell, Sven Haupt, Matthias Dümpelmann, Christian E. Elger, Thomas E. Schlaepfer, Doris Lenartz, Volker Sturm, and Charan Ranganath. 2010. "Intracranial EEG Correlates of Expectancy and Memory Formation in the Human Hippocampus and Nucleus Accumbens." *Neuron* 65 (4): 541–49. https://doi.org/10.1016/j.neuron.2010.02.006.

Azimi, Amin, Zahra Alizadeh, and Maryam Ghorbani. 2021. "The Essential Role of Hippocampo-Cortical Connections in Temporal Coordination of Spindles and Ripples." *NeuroImage* 243 (November):118485. https://doi.org/10.1016/j.neuroimage.2021.118485.

Bandettini, Peter A. 2020. *Fmri*. MIT Press. https://books.google.com/books?hl=en&lr=&id=waXODwAAQBAJ&oi=fnd&pg=PR7&dq=info:Z2THkABte4gJ:scholar.google.com&ots=pXtOoDRjeZ&sig=QtjsICt5sCv87rYlBbetAWmpqjs.

Banville, Hubert, Omar Chehab, Aapo Hyvärinen, Denis-Alexander Engemann, and Alexandre Gramfort. 2021. "Uncovering the Structure of Clinical EEG Signals with Self-Supervised Learning." *Journal of Neural Engineering* 18 (4). https://doi.org/10.1088/1741-2552/abca18.

Barandas, Marília, Duarte Folgado, Letícia Fernandes, Sara Santos, Mariana Abreu, Patrícia Bota, Hui Liu, Tanja Schultz, and Hugo Gamboa. 2020. "TSFEL: Time Series Feature Extraction Library." *SoftwareX* 11 (January):100456. https://doi.org/10.1016/j.softx.2020.100456.

Bartos, Marlene, Imre Vida, and Peter Jonas. 2007. "Synaptic Mechanisms of Synchronized Gamma Oscillations in Inhibitory Interneuron Networks." *Nature Reviews Neuroscience* 8 (1): 45–56. https://doi.org/10.1038/nrn2044.

Bashivan, Pouya, Irina Rish, Mohammed Yeasin, and Noel Codella. 2016. "Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks." arXiv. https://doi.org/10.48550/arXiv.1511.06448.

Beenhakker, Mark P., and John R. Huguenard. 2009. "Neurons That Fire Together Also Conspire Together: Is Normal Sleep Circuitry Hijacked to Generate Epilepsy?" *Neuron* 62 (5): 612–32. https://doi.org/10.1016/j.neuron.2009.05.015.

Benchenane, Karim, Adrien Peyrache, Mehdi Khamassi, Patrick L. Tierney, Yves Gioanni, Francesco P. Battaglia, and Sidney I. Wiener. 2010. "Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal- Prefrontal Network upon Learning." *Neuron* 66 (6): 921–36. https://doi.org/10.1016/j.neuron.2010.05.013.

Bessadok, Alaa, Mohamed Ali Mahjoub, and Islem Rekik. 2023. “Graph Neural Networks in Network Neuroscience.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45 (5): 5833–48. https://doi.org/10.1109/TPAMI.2022.3209686.

Bi, G. Q., and M. M. Poo. 1998. “Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 18 (24): 10464–72. https://doi.org/10.1523/JNEUROSCI.18-24-10464.1998.

Boyce, Richard, Stephen D. Glasgow, Sylvain Williams, and Antoine Adamantidis. 2016. “Causal Evidence for the Role of REM Sleep Theta Rhythm in Contextual Memory Consolidation.” *Science (New York, N.Y.)* 352 (6287): 812–16. https://doi.org/10.1126/science.aad5252.

Buccino, Alessio P., Samuel Garcia, and Pierre Yger. 2022. “Spike Sorting: New Trends and Challenges of the Era of High-Density Probes.” *Progress in Biomedical Engineering* 4 (2): 022005. https://doi.org/10.1088/2516-1091/ac6b96.

Buzsáki, G. 1989. “Two-Stage Model of Memory Trace Formation: A Role for ‘Noisy’ Brain States.” *Neuroscience* 31 (3): 551–70. https://doi.org/10.1016/0306-4522(89)90423-5.

Buzsáki, G., L. W. Leung, and C. H. Vanderwolf. 1983. “Cellular Bases of Hippocampal EEG in the Behaving Rat.” *Brain Research* 287 (2): 139–71. https://doi.org/10.1016/0165-0173(83)90037-1.

Buzsáki, György. 2002. “Theta Oscillations in the Hippocampus.” *Neuron* 33 (3): 325–40. https://doi.org/10.1016/s0896-6273(02)00586-x.

———. 2015. “Hippocampal Sharp Wave-ripple: A Cognitive Biomarker for Episodic Memory and Planning.” *Hippocampus* 25 (10): 1073–1188. https://doi.org/10.1002/hipo.22488.

Buzsáki, György, Costas A. Anastassiou, and Christof Koch. 2012. “The Origin of Extracellular Fields and Currents--EEG, ECoG, LFP and Spikes.” *Nature Reviews. Neuroscience* 13 (6): 407–20. https://doi.org/10.1038/nrn3241.

Buzsáki, György, and Andreas Draguhn. 2004. “Neuronal Oscillations in Cortical Networks.” *Science (New York, N.Y.)* 304 (5679): 1926–29. https://doi.org/10.1126/science.1099745.

Carr, Margaret F., Shantanu P. Jadhav, and Loren M. Frank. 2011. “Hippocampal Replay in the Awake State: A Potential Substrate for Memory Consolidation and Retrieval.” *Nature Neuroscience* 14 (2): 147–53. https://doi.org/10.1038/nn.2732.

Castellani, Andrea, Sebastian Schmitt, and Barbara Hammer. 2021. “Estimating the Electrical Power Output of Industrial Devices with End-to-End Time-Series Classification in the Presence of Label Noise.” In , 12975:469–84. https://doi.org/10.1007/978-3-030-86486-6_29.

Cawley, Gavin C., and Nicola L. C. Talbot. 2010. “On Over-Fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation.” *Journal of Machine Learning Research* 11 (70): 2079–2107.

Celeghin, Alessia, Alessio Borriero, Davide Orsenigo, Matteo Diano, Carlos Andrés Méndez Guerrero, Alan Perotti, Giovanni Petri, and Marco Tamietto. 2023. “Convolutional Neural Networks for Vision Neuroscience: Significance, Developments, and Outstanding Issues.” *Frontiers in Computational Neuroscience* 17 (July). https://doi.org/10.3389/fncom.2023.1153572.

Chambers, Jordan D., Mark J. Cook, Anthony N. Burkitt, and David B. Grayden. 2024. “Using Long Short-Term Memory (LSTM) Recurrent Neural Networks to Classify Unprocessed

EEG for Seizure Prediction." *Frontiers in Neuroscience* 18 (October). https://doi.org/10.3389/fnins.2024.1472747.

Christ, Maximilian, Nils Braun, Julius Neuffer, and Andreas W. Kempa-Liehr. 2018. "Time Series FeatuRe Extraction on Basis of Scalable Hypothesis Tests (Tsfresh – A Python Package)." *Neurocomputing* 307 (September):72–77. https://doi.org/10.1016/j.neucom.2018.03.067.

Colgin, Laura Lee. 2013. "Mechanisms and Functions of Theta Rhythms." *Annual Review of Neuroscience* 36 (July):295–312. https://doi.org/10.1146/annurev-neuro-062012-170330.

———. 2016. "Rhythms of the Hippocampal Network." *Nature Reviews Neuroscience* 17 (4): 239–49. https://doi.org/10.1038/nrn.2016.21.

Colgin, Laura Lee, Tobias Denninger, Marianne Fyhn, Torkel Hafting, Tora Bonnevie, Ole Jensen, May-Britt Moser, and Edvard I. Moser. 2009. "Frequency of Gamma Oscillations Routes Flow of Information in the Hippocampus." *Nature* 462 (7271): 353–57. https://doi.org/10.1038/nature08573.

Collingridge, Graham L., Stephane Peineau, John G. Howland, and Yu Tian Wang. 2010. "Long-Term Depression in the CNS." *Nature Reviews. Neuroscience* 11 (7): 459–73. https://doi.org/10.1038/nrn2867.

Craik, Alexander, Yongtian He, and Jose L. Contreras-Vidal. 2019. "Deep Learning for Electroencephalogram (EEG) Classification Tasks: A Review." *Journal of Neural Engineering* 16 (3): 031001. https://doi.org/10.1088/1741-2552/ab0ab5.

Cunningham, John P., and Byron M. Yu. 2014. "Dimensionality Reduction for Large-Scale Neural Recordings." *Nature Neuroscience* 17 (11): 1500–1509. https://doi.org/10.1038/nn.3776.

Dickerson, Bradford C., and Howard Eichenbaum. 2010. "The Episodic Memory System: Neurocircuitry and Disorders." *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 35 (1): 86–104. https://doi.org/10.1038/npp.2009.126.

Dipietro, Laura, Paola Gonzalez-Mego, Ciro Ramos-Estebanez, Lauren Hana Zukowski, Rahul Mikkilineni, Richard Jarrett Rushmore, and Timothy Wagner. 2023. "The Evolution of Big Data in Neuroscience and Neurology." *Journal of Big Data* 10 (1): 116. https://doi.org/10.1186/s40537-023-00751-2.

Eichenbaum, Howard. 2004. "Hippocampus: Cognitive Processes and Neural Representations That Underlie Declarative Memory." *Neuron* 44 (1): 109–20. https://doi.org/10.1016/j.neuron.2004.08.028.

Ekstrom, Arne D., and Charan Ranganath. 2018. "Space, Time, and Episodic Memory: The Hippocampus Is All over the Cognitive Map." *Hippocampus* 28 (9): 680–87. https://doi.org/10.1002/hipo.22750.

Eldele, Emadeldeen, Mohamed Ragab, Zhenghua Chen, Min Wu, Chee Keong Kwoh, Xiaoli Li, and Cuntai Guan. 2021. "Time-Series Representation Learning via Temporal and Contextual Contrasting." In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 2352–59. Montreal, Canada: International Joint Conferences on Artificial Intelligence Organization. https://doi.org/10.24963/ijcai.2021/324.

Fell, Juergen, and Nikolai Axmacher. 2011. "The Role of Phase Synchronization in Memory Processes." *Nature Reviews. Neuroscience* 12 (2): 105–18. https://doi.org/10.1038/nrn2979.

Fernández-Ruiz, Antonio, Azahara Oliva, Eliezyer Fermino de Oliveira, Florbela Rocha-Almeida, David Tingley, and György Buzsáki. 2019. “Long-Duration Hippocampal Sharp Wave Ripples Improve Memory.” *Science* 364 (6445): 1082–86. https://doi.org/10.1126/science.aax0758.

Frénay, Benoît, and Ata Kabán. 2014. “A Comprehensive Introduction to Label Noise.” *Computational Intelligence*.

Genzel, Lisa, Janine I. Rossato, Justin Jacobse, Roddy M. Grieves, Patrick A. Spooner, Francesco P. Battaglia, Guillen Fernández, and Richard G. M. Morris. 2017. “The Yin and Yang of Memory Consolidation: Hippocampal and Neocortical.” *PLOS Biology* 15 (1): e2000531. https://doi.org/10.1371/journal.pbio.2000531.

Girardeau, Gabrielle, Karim Benchenane, Sidney I. Wiener, György Buzsáki, and Michaël B. Zugaro. 2009. “Selective Suppression of Hippocampal Ripples Impairs Spatial Memory.” *Nature Neuroscience* 12 (10): 1222–23. https://doi.org/10.1038/nn.2384.

Girardeau, Gabrielle, Anne Cei, and Michaël Zugaro. 2014. “Learning-Induced Plasticity Regulates Hippocampal Sharp Wave-Ripple Drive.” *Journal of Neuroscience* 34 (15): 5176–83. https://doi.org/10.1523/JNEUROSCI.4288-13.2014.

Girden, Ellen. 1992. *ANOVA*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc. https://doi.org/10.4135/9781412983419.

Glaser, Joshua I., Ari S. Benjamin, Raeed H. Chowdhury, Matthew G. Perich, Lee E. Miller, and Konrad P. Kording. 2020. “Machine Learning for Neural Decoding.” *eNeuro* 7 (4). https://doi.org/10.1523/ENEURO.0506-19.2020.

Goutagny, Romain, and Slavica Krantic. 2013. “Hippocampal Oscillatory Activity in Alzheimer’s Disease: Toward the Identification of Early Biomarkers?” *Aging and Disease* 4 (3): 134–40.

Greenacre, Michael, Patrick J. F. Groenen, Trevor Hastie, Alfonso Iodice D’Enza, Angelos Markos, and Elena Tuzhilina. 2022. “Principal Component Analysis.” *Nature Reviews Methods Primers* 2 (1): 1–21. https://doi.org/10.1038/s43586-022-00184-w.

Gui, Jie, Tuo Chen, Jing Zhang, Qiong Cao, Zhenan Sun, Hao Luo, and Dacheng Tao. 2024. “A Survey on Self-Supervised Learning: Algorithms, Applications, and Future Trends.” arXiv. https://doi.org/10.48550/arXiv.2301.05712.

Gunnarsdóttir, Brynja, Valerio Zerbi, and Clare Kelly. 2022. “Multimodal Gradient Mapping of Rodent Hippocampus.” *NeuroImage* 253 (June):119082. https://doi.org/10.1016/j.neuroimage.2022.119082.

Hafting, Torkel, Marianne Fyhn, Sturla Molden, May-Britt Moser, and Edvard I. Moser. 2005. “Microstructure of a Spatial Map in the Entorhinal Cortex.” *Nature* 436 (7052): 801–6. https://doi.org/10.1038/nature03721.

Hainmueller, Thomas, and Marlene Bartos. 2020. “Dentate Gyrus Circuits for Encoding, Retrieval and Discrimination of Episodic Memories.” *Nature Reviews. Neuroscience* 21 (3): 153–68. https://doi.org/10.1038/s41583-019-0260-z.

Hardt, Oliver, Karim Nader, and Lynn Nadel. 2013. “Decay Happens: The Role of Active Forgetting in Memory.” *Trends in Cognitive Sciences* 17 (3): 111–20. https://doi.org/10.1016/j.tics.2013.01.001.

He, Kaiming, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. 2022. “Masked Autoencoders Are Scalable Vision Learners.” In , 16000–9. https://openaccess.thecvf.com/content/CVPR2022/html/He_Masked_Autoencoders_Are_Scalable_Vision_Learners_CVPR_2022_paper.html.

Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. 2006. “A Fast Learning Algorithm for Deep Belief Nets.” *Neural Computation* 18 (7): 1527–54. https://doi.org/10.1162/neco.2006.18.7.1527.

Hochreiter, Sepp, and Jürgen Schmidhuber. 1997. “Long Short-Term Memory.” *Neural Computation* 9 (8): 1735–80. https://doi.org/10.1162/neco.1997.9.8.1735.

Hodgkin, A. L., and A. F. Huxley. 1952. “A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve.” *The Journal of Physiology* 117 (4): 500–544.

Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.” arXiv. https://doi.org/10.48550/arXiv.1704.04861.

Hsu, Sheng-Yi, Bartosz Jura, Mau-Hsiang Shih, Pierre Meyrand, Feng-Sheng Tsai, and Tiaza Bem. 2021. “Recognition of Post-Learning Alteration of Hippocampal Ripples by Convolutional Neural Network Differs in the Wild-Type and AD Mice.” *Scientific Reports* 11 (October):21241. https://doi.org/10.1038/s41598-021-00598-8.

Hunter, John D. 2007. “Matplotlib: A 2D Graphics Environment.” *Computing in Science & Engineering* 9 (3): 90–95. https://doi.org/10.1109/MCSE.2007.55.

Iaccarino, Hannah F., Annabelle C. Singer, Anthony J. Martorell, Andrii Rudenko, Fan Gao, Tyler Z. Gillingham, Hansruedi Mathys, et al. 2016. “Gamma Frequency Entrainment Attenuates Amyloid Load and Modifies Microglia.” *Nature* 540 (7632): 230–35. https://doi.org/10.1038/nature20587.

Ioffe, Sergey, and Christian Szegedy. 2015. “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.” arXiv. https://doi.org/10.48550/arXiv.1502.03167.

Jadhav, Shantanu P., Caleb Kemere, P. Walter German, and Loren M. Frank. 2012. “Awake Hippocampal Sharp-Wave Ripples Support Spatial Memory.” *Science* 336 (6087): 1454–58. https://doi.org/10.1126/science.1217230.

Jankowsky, Joanna L., Hilda H. Slunt, Victoria Gonzales, Nancy A. Jenkins, Neal G. Copeland, and David R. Borchelt. 2004. “APP Processing and Amyloid Deposition in Mice Haplo-Insufficient for Presenilin 1.” *Neurobiology of Aging* 25 (7): 885–92. https://doi.org/10.1016/j.neurobiolaging.2003.09.008.

Januszewski, Michał, Jörgen Kornfeld, Peter H. Li, Art Pope, Tim Blakely, Larry Lindsey, Jeremy Maitin-Shepard, Mike Tyka, Winfried Denk, and Viren Jain. 2018. “High-Precision Automated Reconstruction of Neurons with Flood-Filling Networks.” *Nature Methods* 15 (8): 605–10. https://doi.org/10.1038/s41592-018-0049-4.

Jiang, Xi, Jorge Gonzalez-Martinez, and Eric Halgren. 2019. “Coordination of Human Hippocampal Sharpwave Ripples during NREM Sleep with Cortical Theta Bursts, Spindles, Downstates, and Upstates.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 39 (44): 8744–61. https://doi.org/10.1523/JNEUROSCI.2857-18.2019.

Joo, Hannah R., and Loren M. Frank. 2018. “The Hippocampal Sharp Wave–Ripple in Memory Retrieval for Immediate Use and Consolidation.” *Nature Reviews Neuroscience* 19 (12): 744–57. https://doi.org/10.1038/s41583-018-0077-1.

Jun, James J., Nicholas A. Steinmetz, Joshua H. Siegle, Daniel J. Denman, Marius Bauza, Brian Barbarits, Albert K. Lee, et al. 2017. “Fully Integrated Silicon Probes for High-Density

Recording of Neural Activity." *Nature* 551 (7679): 232–36. https://doi.org/10.1038/nature24636.

Jura, Bartosz, Nathalie Macrez, Pierre Meyrand, and Tiaza Bem. 2019. "Deficit in Hippocampal Ripples Does Not Preclude Spatial Memory Formation in APP/PS1 Mice." *Scientific Reports* 9 (December):20129. https://doi.org/10.1038/s41598-019-56582-w.

Kiranyaz, Serkan, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J. Inman. 2021. "1D Convolutional Neural Networks and Applications: A Survey." *Mechanical Systems and Signal Processing* 151 (April):107398. https://doi.org/10.1016/j.ymssp.2020.107398.

Kohavi, Ron. 1995. "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection." In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, 1137–43. IJCAI'95. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Kostas, Demetres, Stéphane Aroca-Ouellette, and Frank Rudzicz. 2021. "BENDR: Using Transformers and a Contrastive Self-Supervised Learning Task to Learn From Massive Amounts of EEG Data." *Frontiers in Human Neuroscience* 15 (June). https://doi.org/10.3389/fnhum.2021.653659.

Kudrimoti, H. S., C. A. Barnes, and B. L. McNaughton. 1999. "Reactivation of Hippocampal Cell Assemblies: Effects of Behavioral State, Experience, and EEG Dynamics." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 19 (10): 4090–4101. https://doi.org/10.1523/JNEUROSCI.19-10-04090.1999.

Kumaran, Dharshan, Demis Hassabis, and James L. McClelland. 2016. "What Learning Systems Do Intelligent Agents Need? Complementary Learning Systems Theory Updated." *Trends in Cognitive Sciences* 20 (7): 512–34. https://doi.org/10.1016/j.tics.2016.05.004.

Kumaran, Dharshan, and Eleanor A. Maguire. 2007. "Match-Mismatch Processes Underlie Human Hippocampal Responses to Associative Novelty." *The Journal of Neuroscience* 27 (32): 8517–24. https://doi.org/10.1523/JNEUROSCI.1677-07.2007.

Lagadec, Chann, Erina Vlashi, Lorenza Della Donna, Carmen Dekmezian, and Frank Pajonk. 2012. "Radiation-Induced Reprogramming of Breast Cancer Cells." *Stem Cells (Dayton, Ohio)* 30 (5): 833–44. https://doi.org/10.1002/stem.1058.

Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86 (11): 2278–2324. https://doi.org/10.1109/5.726791.

Lee, Albert K., and Matthew A. Wilson. 2002. "Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep." *Neuron* 36 (6): 1183–94. https://doi.org/10.1016/s0896-6273(02)01096-6.

Leisman, Gerry, Calixto Macahdo, Robert Melillo, and Raed Mualem. 2012. "Intentionality and 'Free-Will' from a Neurodevelopmental Perspective." *Frontiers in Integrative Neuroscience* 6 (June). https://doi.org/10.3389/fnint.2012.00036.

Linderman, Scott W., and Samuel J. Gershman. 2017. "Using Computational Theory to Constrain Statistical Models of Neural Data." *Current Opinion in Neurobiology* 46 (October):14–24. https://doi.org/10.1016/j.conb.2017.06.004.

Liu, Zhuang, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. 2022. "A ConvNet for the 2020s." arXiv. https://doi.org/10.48550/arXiv.2201.03545.

Logothetis, N. K., O. Eschenko, Y. Murayama, M. Augath, T. Steudel, H. C. Evrard, M. Besserve, and A. Oeltermann. 2012. “Hippocampal-Cortical Interaction during Periods of Subcortical Silence.” *Nature* 491 (7425): 547–53. https://doi.org/10.1038/nature11618.

Lotte, F., L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. 2018. “A Review of Classification Algorithms for EEG-Based Brain–Computer Interfaces: A 10 Year Update.” *Journal of Neural Engineering* 15 (3): 031005. https://doi.org/10.1088/1741-2552/aab2f2.

Ma, Peitian, Zhen Liu, Junhao Zheng, Linghao Wang, and Qianli Ma. 2023. “CTW: Confident Time-Warping for Time-Series Label-Noise Learning.” In , 4:4046–54. https://doi.org/10.24963/ijcai.2023/450.

Maaten, Laurens van der, and Geoffrey Hinton. 2008. “Visualizing Data Using T-SNE.” *Journal of Machine Learning Research* 9 (86): 2579–2605.

Magee, J. C., and D. Johnston. 1997. “A Synaptically Controlled, Associative Signal for Hebbian Plasticity in Hippocampal Neurons.” *Science (New York, N.Y.)* 275 (5297): 209–13. https://doi.org/10.1126/science.275.5297.209.

Makeig, Scott, Arnaud Delorme, Marissa Westerfield, Tzyy-Ping Jung, Jeanne Townsend, Eric Courchesne, and Terrence J. Sejnowski. 2004. “Electroencephalographic Brain Dynamics Following Manually Responded Visual Targets.” *PLOS Biology* 2 (6): e176. https://doi.org/10.1371/journal.pbio.0020176.

Manns, Joseph R., Ramona O. Hopkins, Jonathan M. Reed, Erin G. Kitchener, and Larry R. Squire. 2003. “Recognition Memory and the Human Hippocampus.” *Neuron* 37 (1): 171–80. https://doi.org/10.1016/s0896-6273(02)01147-9.

Maren, S., and M. S. Fanselow. 1995. “Synaptic Plasticity in the Basolateral Amygdala Induced by Hippocampal Formation Stimulation in Vivo.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 15 (11): 7548–64. https://doi.org/10.1523/JNEUROSCI.15-11-07548.1995.

McCulloch, Warren S., and Walter Pitts. 1943. “A Logical Calculus of the Ideas Immanent in Nervous Activity.” *The Bulletin of Mathematical Biophysics* 5 (4): 115–33. https://doi.org/10.1007/BF02478259.

Minsky, Marvin, and Seymour Papert. 1969. *Perceptrons*. Perceptrons. Oxford, England: M.I.T. Press.

Moscovitch, Morris, Roberto Cabeza, Gordon Winocur, and Lynn Nadel. 2016. “Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation.” *Annual Review of Psychology* 67:105–34. https://doi.org/10.1146/annurev-psych-113011-143733.

Nádasdy, Zoltán, Hajime Hirase, András Czurkó, Jozsef Csicsvari, and György Buzsáki. 1999. “Replay and Time Compression of Recurring Spike Sequences in the Hippocampus.” *Journal of Neuroscience* 19 (21): 9497–9507. https://doi.org/10.1523/JNEUROSCI.19-21-09497.1999.

Nadel, L., and M. Moscovitch. 1997. “Memory Consolidation, Retrograde Amnesia and the Hippocampal Complex.” *Current Opinion in Neurobiology* 7 (2): 217–27. https://doi.org/10.1016/s0959-4388(97)80010-4.

Nair, Vinod, and Geoffrey E Hinton. n.d. “Rectified Linear Units Improve Restricted Boltzmann Machines.”

Nakazawa, Kazu, Michael C. Quirk, Raymond A. Chitwood, Masahiko Watanabe, Mark F. Yeckel, Linus D. Sun, Akira Kato, et al. 2002. “Requirement for Hippocampal CA3 NMDA

Receptors in Associative Memory Recall." *Science (New York, N.Y.)* 297 (5579): 211–18. https://doi.org/10.1126/science.1071795.

Navas-Olive, Andrea, Rodrigo Amaducci, Maria-Teresa Jurado-Parras, Enrique R Sebastian, and Liset M de la Prida. 2022. "Deep Learning-Based Feature Extraction for Prediction and Interpretation of Sharp-Wave Ripples in the Rodent Hippocampus." Edited by Adrien Peyrache and John R Huguenard. *eLife* 11 (September):e77772. https://doi.org/10.7554/eLife.77772.

Navas-Olive, Andrea, Adrian Rubio, Saman Abbaspoor, Kari L. Hoffman, and Liset M. de la Prida. 2023. "A Machine Learning Toolbox for the Analysis of Sharp-Wave Ripples Reveal Common Features across Species." *bioRxiv: The Preprint Server for Biology*, July, 2023.07.02.547382. https://doi.org/10.1101/2023.07.02.547382.

Neher, E., and B. Sakmann. 1976. "Single-Channel Currents Recorded from Membrane of Denervated Frog Muscle Fibres." *Nature* 260 (5554): 799–802. https://doi.org/10.1038/260799a0.

Newell, Allen, and Herbert A. Simon. 1972. *Human Problem Solving*. Human Problem Solving. Oxford, England: Prentice-Hall.

Ognjanovski, Nicolette, Christopher Broussard, Michal Zochowski, and Sara J. Aton. 2018. "Hippocampal Network Oscillations Rescue Memory Consolidation Deficits Caused by Sleep Loss." *Cerebral Cortex (New York, N.Y.: 1991)* 28 (10): 3711–23. https://doi.org/10.1093/cercor/bhy174.

O'Keefe, J., and J. Dostrovsky. 1971. "The Hippocampus as a Spatial Map. Preliminary Evidence from Unit Activity in the Freely-Moving Rat." *Brain Research* 34 (1): 171–75. https://doi.org/10.1016/0006-8993(71)90358-1.

O'Keefe, J., and M. L. Recce. 1993. "Phase Relationship between Hippocampal Place Units and the EEG Theta Rhythm." *Hippocampus* 3 (3): 317–30. https://doi.org/10.1002/hipo.450030307.

Ólafsdóttir, H. Freyja, Daniel Bush, and Caswell Barry. 2018. "The Role of Hippocampal Replay in Memory and Planning." *Current Biology* 28 (1): R37–50. https://doi.org/10.1016/j.cub.2017.10.073.

Oliva, Azahara, Antonio Fernández-Ruiz, Eliezyer Fermino de Oliveira, and György Buzsáki. 2018. "Origin of Gamma Frequency Power during Hippocampal Sharp-Wave Ripples." *Cell Reports* 25 (7): 1693-1700.e4. https://doi.org/10.1016/j.celrep.2018.10.066.

Olton, David S., Christine Collison, and Mary Ann Werz. 1977. "Spatial Memory and Radial Arm Maze Performance of Rats." *Learning and Motivation* 8 (3): 289–314. https://doi.org/10.1016/0023-9690(77)90054-6.

O'Mara, S. M., S. Commins, M. Anderson, and J. Gigg. 2001. "The Subiculum: A Review of Form, Physiology and Function." *Progress in Neurobiology* 64 (2): 129–55. https://doi.org/10.1016/s0301-0082(00)00054-x.

Palop, Jorge J., Jeannie Chin, Erik D. Roberson, Jun Wang, Myo T. Thwin, Nga Bien-Ly, Jong Yoo, et al. 2007. "Aberrant Excitatory Neuronal Activity and Compensatory Remodeling of Inhibitory Hippocampal Circuits in Mouse Models of Alzheimer's Disease." *Neuron* 55 (5): 697–711. https://doi.org/10.1016/j.neuron.2007.07.025.

Pandarinath, Chethan, Daniel J. O'Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D. Stavisky, Jonathan C. Kao, Eric M. Trautmann, et al. 2018. "Inferring Single-Trial Neural Population Dynamics Using Sequential Auto-Encoders." *Nature Methods* 15 (10): 805–15. https://doi.org/10.1038/s41592-018-0109-9.

Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, et al. 2019. “PyTorch: An Imperative Style, High-Performance Deep Learning Library.” arXiv. https://doi.org/10.48550/arXiv.1912.01703.

Peyrache, Adrien, Mehdi Khamassi, Karim Benchenane, Sidney I. Wiener, and Francesco P. Battaglia. 2009. “Replay of Rule-Learning Related Neural Patterns in the Prefrontal Cortex during Sleep.” *Nature Neuroscience* 12 (7): 919–26. https://doi.org/10.1038/nn.2337.

Pfeiffer, Brad E., and David J. Foster. 2013. “Hippocampal Place-Cell Sequences Depict Future Paths to Remembered Goals.” *Nature* 497 (7447): 74–79. https://doi.org/10.1038/nature12112.

Preston, Alison R., and Howard Eichenbaum. 2013. “Interplay of Hippocampus and Prefrontal Cortex in Memory.” *Current Biology: CB* 23 (17): R764-773. https://doi.org/10.1016/j.cub.2013.05.041.

“PyCaret — Pycaret 3.0.4 Documentation.” n.d. Accessed July 28, 2024. https://pycaret.readthedocs.io/en/latest/.

Ranganath, Charan, and Robert S. Blumenfeld. 2005. “Doubts about Double Dissociations between Short- and Long-Term Memory.” *Trends in Cognitive Sciences* 9 (8): 374–80. https://doi.org/10.1016/j.tics.2005.06.009.

Ranganath, Charan, and Maureen Ritchey. 2012. “Two Cortical Systems for Memory-Guided Behaviour.” *Nature Reviews Neuroscience* 13 (10): 713–26. https://doi.org/10.1038/nrn3338.

Robinson, Peter A., James A. Henderson, Natasha C. Gabay, Kevin M. Aquino, Tara Babaie-Janvier, and Xiao Gao. 2021. “Determination of Dynamic Brain Connectivity via Spectral Analysis.” *Frontiers in Human Neuroscience* 15 (July). https://doi.org/10.3389/fnhum.2021.655576.

Rosenblatt, F. 1958. “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.” *Psychological Review* 65 (6): 386–408. https://doi.org/10.1037/h0042519.

Roux, Candice M., Marianne Leger, and Thomas Freret. 2021. “Memory Disorders Related to Hippocampal Function: The Interest of 5-HT4Rs Targeting.” *International Journal of Molecular Sciences* 22 (21): 12082. https://doi.org/10.3390/ijms222112082.

Ruder, Sebastian. 2017. “An Overview of Gradient Descent Optimization Algorithms.” arXiv. https://doi.org/10.48550/arXiv.1609.04747.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. 1986. “Learning Representations by Back-Propagating Errors.” *Nature* 323 (6088): 533–36. https://doi.org/10.1038/323533a0.

Rumelhart, David E., James L. McClelland, and AU. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. The MIT Press. https://doi.org/10.7551/mitpress/5236.001.0001.

Sadowski, Josef H. L. P., Matthew W. Jones, and Jack R. Mellor. 2016. “Sharp-Wave Ripples Orchestrate the Induction of Synaptic Plasticity during Reactivation of Place Cell Firing Patterns in the Hippocampus.” *Cell Reports* 14 (8): 1916–29. https://doi.org/10.1016/j.celrep.2016.01.061.

Samek, Wojciech, Grégoire Montavon, Sebastian Lapuschkin, Christopher J. Anders, and Klaus-Robert Müller. 2021. “Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications.” *Proceedings of the IEEE* 109 (3): 247–78. https://doi.org/10.1109/JPROC.2021.3060483.

Schacter, Daniel L., Donna Rose Addis, and Randy L. Buckner. 2007. “Remembering the Past to Imagine the Future: The Prospective Brain.” *Nature Reviews. Neuroscience* 8 (9): 657–61. https://doi.org/10.1038/nrn2213.

Schacter, Daniel L., Donna Rose Addis, Demis Hassabis, Victoria C. Martin, R. Nathan Spreng, and Karl K. Szpunar. 2012. “The Future of Memory: Remembering, Imagining, and the Brain.” *Neuron* 76 (4): 10.1016/j.neuron.2012.11.001. https://doi.org/10.1016/j.neuron.2012.11.001.

Schlingloff, Dániel, Szabolcs Káli, Tamás F. Freund, Norbert Hájos, and Attila I. Gulyás. 2014. “Mechanisms of Sharp Wave Initiation and Ripple Generation.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34 (34): 11385–98. https://doi.org/10.1523/JNEUROSCI.0867-14.2014.

Scoville, William Beecher, and Brenda Milner. 1957. “Loss of Recent Memory after Bilateral Hippocampal Lesions.” *Journal of Neurology, Neurosurgery & Psychiatry* 20:11–21. https://doi.org/10.1136/jnnp.20.1.11.

Sederberg, Per B., Michael J. Kahana, Marc W. Howard, Elizabeth J. Donner, and Joseph R. Madsen. 2003. “Theta and Gamma Oscillations during Encoding Predict Subsequent Recall.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 23 (34): 10809–14. https://doi.org/10.1523/JNEUROSCI.23-34-10809.2003.

Selkoe, Dennis J. 2002. “Alzheimer’s Disease Is a Synaptic Failure.” *Science (New York, N.Y.)* 298 (5594): 789–91. https://doi.org/10.1126/science.1074069.

Selvaraju, Ramprasaath R., Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2020. “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization.” *International Journal of Computer Vision* 128 (2): 336–59. https://doi.org/10.1007/s11263-019-01228-7.

Seth, Anil K., Adam B. Barrett, and Lionel Barnett. 2015. “Granger Causality Analysis in Neuroscience and Neuroimaging.” *Journal of Neuroscience* 35 (8): 3293–97. https://doi.org/10.1523/JNEUROSCI.4399-14.2015.

Siegle, Joshua H., and Matthew A. Wilson. 2014. “Enhancement of Encoding and Retrieval Functions through Theta Phase-Specific Manipulation of Hippocampus.” *eLife* 3 (July):e03061. https://doi.org/10.7554/eLife.03061.

Simonyan, Karen, and Andrew Zisserman. 2015. “Very Deep Convolutional Networks for Large-Scale Image Recognition.” arXiv. https://doi.org/10.48550/arXiv.1409.1556.

Sirota, Anton, Jozsef Csicsvari, Derek Buhl, and György Buzsáki. 2003. “Communication between Neocortex and Hippocampus during Sleep in Rodents.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (4): 2065–69. https://doi.org/10.1073/pnas.0437938100.

Smith, Anne C., and Emery N. Brown. 2003. “Estimating a State-Space Model from Point Process Observations.” *Neural Computation* 15 (5): 965–91. https://doi.org/10.1162/089976603765202622.

Spruston, Nelson. 2008. “Pyramidal Neurons: Dendritic Structure and Synaptic Integration.” *Nature Reviews. Neuroscience* 9 (3): 206–21. https://doi.org/10.1038/nrn2286.

Squire, L. R., and P. Alvarez. 1995. “Retrograde Amnesia and Memory Consolidation: A Neurobiological Perspective.” *Current Opinion in Neurobiology* 5 (2): 169–77. https://doi.org/10.1016/0959-4388(95)80023-9.

Srinivas, M., and L.M. Patnaik. 1994. “Genetic Algorithms: A Survey.” *Computer* 27 (6): 17–26. https://doi.org/10.1109/2.294849.

Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting.” *Journal of Machine Learning Research* 15 (56): 1929–58.

Stark, Eran, Lisa Roux, Ronny Eichler, Yuta Senzai, Sebastien Royer, and György Buzsáki. 2014. “Pyramidal Cell-Interneuron Interactions Underlie Hippocampal Ripple Oscillations.” *Neuron* 83 (2): 467–80. https://doi.org/10.1016/j.neuron.2014.06.023.

Steinmetz, Nicholas A., Cagatay Aydin, Anna Lebedeva, Michael Okun, Marius Pachitariu, Marius Bauza, Maxime Beau, et al. 2021. “Neuropixels 2.0: A Miniaturized High-Density Probe for Stable, Long-Term Brain Recordings.” *Science* 372 (6539): eabf4588. https://doi.org/10.1126/science.abf4588.

Sullivan, David, Jozsef Csicsvari, Kenji Mizuseki, Sean Montgomery, Kamran Diba, and György Buzsáki. 2011. “Relationships between Hippocampal Sharp Waves, Ripples, and Fast Gamma Oscillation: Influence of Dentate and Entorhinal Cortical Activity.” *The Journal of Neuroscience* 31 (23): 8605–16. https://doi.org/10.1523/JNEUROSCI.0294-11.2011.

Sussillo, David, Sergey D. Stavisky, Jonathan C. Kao, Stephen I. Ryu, and Krishna V. Shenoy. 2016. “Making Brain–Machine Interfaces Robust to Future Neural Variability.” *Nature Communications* 7 (1): 13749. https://doi.org/10.1038/ncomms13749.

Swartz, B. E., and E. S. Goldensohn. 1998. “Timeline of the History of EEG and Associated Fields.” *Electroencephalography and Clinical Neurophysiology* 106 (2): 173–76.

Tallon-Baudry, Catherine, Olivier Bertrand, Claude Delpuech, and Jacques Pernier. 1997. “Oscillatory γ-Band (30–70 Hz) Activity Induced by a Visual Search Task in Humans.” *Journal of Neuroscience* 17 (2): 722–34. https://doi.org/10.1523/JNEUROSCI.17-02-00722.1997.

Tan, Mingxing, and Quoc V. Le. 2020. “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.” arXiv. https://doi.org/10.48550/arXiv.1905.11946.

Tort, Adriano B. L., Robert Komorowski, Howard Eichenbaum, and Nancy Kopell. 2010. “Measuring Phase-Amplitude Coupling between Neuronal Oscillations of Different Frequencies.” *Journal of Neurophysiology* 104 (2): 1195–1210. https://doi.org/10.1152/jn.00106.2010.

Tort, Adriano B. L., Robert W. Komorowski, Joseph R. Manns, Nancy J. Kopell, and Howard Eichenbaum. 2009. “Theta-Gamma Coupling Increases during the Learning of Item-Context Associations.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (49): 20942–47. https://doi.org/10.1073/pnas.0911331106.

Van Rossum, Guido, and Fred L. Drake. 2009. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2023. “Attention Is All You Need.” arXiv. https://doi.org/10.48550/arXiv.1706.03762.

Vincent, Pascal, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. 2010. “Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion.” *Journal of Machine Learning Research* 11 (110): 3371–3408.

Virtanen, Pauli, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, et al. 2020. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.” *Nature Methods* 17 (3): 261–72. https://doi.org/10.1038/s41592-019-0686-2.

Wang, Yingxue, Sandro Romani, Brian Lustig, Anthony Leonardo, and Eva Pastalkova. 2015. "Theta Sequences Are Essential for Internally Generated Hippocampal Firing Fields." *Nature Neuroscience* 18 (2): 282–88. https://doi.org/10.1038/nn.3904.

Wen, Junhao, Elina Thibeau-Sutre, Mauricio Diaz-Melo, Jorge Samper-González, Alexandre Routier, Simona Bottani, Didier Dormont, Stanley Durrleman, Ninon Burgos, and Olivier Colliot. 2020. "Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation." *Medical Image Analysis* 63 (July):101694. https://doi.org/10.1016/j.media.2020.101694.

Wilcoxon, Frank. 1945. "Individual Comparisons by Ranking Methods." *Biometrics Bulletin* 1 (6): 80–83. https://doi.org/10.2307/3001968.

Wilson, Matthew A., and Bruce L. McNaughton. 1994. "Reactivation of Hippocampal Ensemble Memories During Sleep." *Science* 265 (5172): 676–79. https://doi.org/10.1126/science.8036517.

Yamins, Daniel L. K., and James J. DiCarlo. 2016. "Using Goal-Driven Deep Learning Models to Understand Sensory Cortex." *Nature Neuroscience* 19 (3): 356–65. https://doi.org/10.1038/nn.4244.

Yang, Weijian, and Rafael Yuste. 2017. "In Vivo Imaging of Neural Activity." *Nature Methods* 14 (4): 349–59. https://doi.org/10.1038/nmeth.4230.

Yassa, Michael A., and Craig E. L. Stark. 2011. "Pattern Separation in the Hippocampus." *Trends in Neurosciences* 34 (10): 515–25. https://doi.org/10.1016/j.tins.2011.06.006.

Yger, Pierre, Giulia LB Spampinato, Elric Esposito, Baptiste Lefebvre, Stéphane Deny, Christophe Gardella, Marcel Stimberg, et al. 2018. "A Spike Sorting Toolbox for up to Thousands of Electrodes Validated with Ground Truth Recordings in Vitro and in Vivo." Edited by David Kleinfeld. *eLife* 7 (March):e34518. https://doi.org/10.7554/eLife.34518.

Ylinen, A., A. Bragin, Z. Nádasdy, G. Jandó, I. Szabó, A. Sik, and G. Buzsáki. 1995. "Sharp Wave-Associated High-Frequency Oscillation (200 Hz) in the Intact Hippocampus: Network and Intracellular Mechanisms." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 15 (1 Pt 1): 30–46. https://doi.org/10.1523/JNEUROSCI.15-01-00030.1995.

Yonelinas, Andrew P., Charan Ranganath, Arne D. Ekstrom, and Brian J. Wiltgen. 2019. "A Contextual Binding Theory of Episodic Memory: Systems Consolidation Reconsidered." *Nature Reviews. Neuroscience* 20 (6): 364–75. https://doi.org/10.1038/s41583-019-0150-4.

Yu, Jai Y., and Loren M. Frank. 2015. "Hippocampal-Cortical Interaction in Decision Making." *Neurobiology of Learning and Memory* 117 (January):34–41. https://doi.org/10.1016/j.nlm.2014.02.002.

Yu, Tong, and Hong Zhu. 2020. "Hyper-Parameter Optimization: A Review of Algorithms and Applications." arXiv. https://doi.org/10.48550/arXiv.2003.05689.

Zhang, K., I. Ginzburg, B. L. McNaughton, and T. J. Sejnowski. 1998. "Interpreting Neuronal Population Activity by Reconstruction: Unified Framework with Application to Hippocampal Place Cells." *Journal of Neurophysiology* 79 (2): 1017–44. https://doi.org/10.1152/jn.1998.79.2.1017.

Zhang, Kexin, Qingsong Wen, Chaoli Zhang, Rongyao Cai, Ming Jin, Yong Liu, James Zhang, et al. 2024. "Self-Supervised Learning for Time Series Analysis: Taxonomy, Progress, and Prospects." arXiv. https://doi.org/10.48550/arXiv.2306.10125.

Zhao, Zehui, Laith Alzubaidi, Jinglan Zhang, Ye Duan, and Yuantong Gu. 2024. "A Comparison Review of Transfer Learning and Self-Supervised Learning: Definitions, Applications,

Advantages and Limitations." *Expert Systems with Applications* 242 (May):122807. https://doi.org/10.1016/j.eswa.2023.122807.

Zimmerman, Donald W. 1997. "Teacher's Corner: A Note on Interpretation of the Paired-Samples t Test." *Journal of Educational and Behavioral Statistics* 22 (3): 349–60. https://doi.org/10.3102/10769986022003349.