Neural models for detection and classification of brain states and transitions

Check for updates

Arnau Marin-Llobet [1,2,5], Arnau Manasanch [1,3,5], Leonardo Dalla Porta [1], Melody Torao-Angosto [1] & Maria V. Sanchez-Vives [1,4] ✉

Exploring natural or pharmacologically induced brain dynamics, such as sleep, wakefulness, or anesthesia, provides rich functional models for studying brain states. These models allow detailed examination of unique spatiotemporal neural activity patterns that reveal brain function. However, assessing transitions between brain states remains computationally challenging. Here we introduce a pipeline to detect brain states and their transitions in the cerebral cortex using a dual-model Convolutional Neural Network (CNN) and a self-supervised autoencoder-based multimodal clustering algorithm. This approach distinguishes brain states such as slow oscillations, microarousals, and wakefulness with high confidence. Using chronic local field potential recordings from rats, our method achieved a global accuracy of 91%, with up to 96% accuracy for certain states. For the transitions, we report an average accuracy of 74%. Our models were trained using a leave-one-out methodology, allowing for broad applicability across subjects and pre-trained models for deployments. It also features a confidence parameter, ensuring that only highly certain cases are automatically classified, leaving ambiguous cases for the multimodal unsupervised classifier or further expert review. Our approach presents a reliable and efficient tool for brain state labeling and analysis, with applications in basic and clinical neuroscience.

Brain states are unique configurations of neural activity that are expressed in different behavioral and cognitive conditions, such as wakefulness, sleep, and anesthesia[1,2]. These states are characterized by distinct spatiotemporal patterns of neuronal firing and functional connectivity, which influence how the brain processes information and responds to external stimuli[3–7]. The significance of the properties and transitions between these brain states are not only relevant to the investigatation of the mechanisms at play in the healthy brain, but are also critical for the understanding of pathological brain states[8–10].

The classification and identification of sleep stages has been at the core of sleep studies, and has required human supervision for decades. Another relevant use of the identification of brain states is during anesthesia, since it is critical to know the level of anesthesia during medical procedures to ensure sedation and minimal side effects. The relevance of such identification is reflected in various recent investigations, using machine learning methods for the automatic detection of sleep stages[11,12] or anesthesia depth[13–16]. The

identification of consciousness levels, especially in disorders of consciousness[17], have also been a target of recent deep learning developments[18]. All these approaches primarily rely on features extracted from electroencephalography (EEG) to classify different stages. These cases illustrate that the identification and study of brain states and the transitions across them are of crucial interest for clinical applications.

The study of brain states is also critical for basic neuroscience, which investigates more detailed mechanisms involving the cellular and local network levels, leveraging techniques with higher spatial resolution than EEG[19,20]. Brain states can occur naturally or be artificially induced, with slow wave sleep and deep anesthesia being prime examples of each type, respectively. Both these states show similar brain activity patterns over time and space, making them ideal for studying how the brain can operate under conditions of lower excitability and synchronized activity patterns[21]. During wakefulness, the cortex is characterized by asynchronous activity that corresponds with heightened consciousness and cognitive capabilities.

[1]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Rosselló 149-153, 08036 Barcelona, Spain. [2]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Boston, MA, 02138, USA. [3]Faculty of Medicine and Health Sciences, University of Barcelona, 08036 Barcelona, Spain. [4]ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain. [5]These authors contributed equally: Arnau Marin-Llobet, Arnau Manasanch. ✉e-mail: msanche3@recerca.clinic.cat

However, in the context of deep sleep or deep anesthesia, the cortex predominantly exhibits synchronous slow oscillations (SO) that vary from 0.5 to 2 Hz for deep sleep, and 0.1 to 1 Hz for deep anesthesia[21–23]. These oscillations rhythmically shift between active "Up states" and quiescent "Down states"[24] (also known as Off-periods[8]), leading to a reduction in both information processing and spatiotemporal complexity[25–30]. Deep anesthesia provides a valuable model for exploring the brain's quieter operational modes and their transitions to more active states[19,20,31,32]. Unlike the unpredictability of spontaneous sleep transitions, anesthesia offers the advantage of precise experimental control over the state. Wakefulness and deep anesthesia are two extremes within a much broader spectrum of possible brain states[1,33], and the study of different levels of anesthesia on cortical dynamics provides an avenue for this research[34,35] generating diverse spatiotemporal patterns corresponding to brain states with physiological relevance[3,6].

One of the brain states that emerges when transitioning from deep to light anesthesia are microarousals. Microarousals are periods of desynchronized activity, arousal-like, that are well known in the field of physiological sleep, appearing in alternation with slow waves[36]. Microarousals are also relevant in different pathologies in which they present alterations, like in sleep disorders or different neurological and psychiatric conditions. Recent findings[37] suggest that transitioning from deep anesthesia to wakefulness is marked by both subtle and pronounced shifts in brain dynamics, as seen in local field potential (LFP) and multi-unit activity[38] (MUA). Microarousals can also be seen by injecting norepinephrine into the basal forebrain in subjects under desflurane anesthesia[38], and also by the spontaneously released norepinephrine by the locus coeruleus during slow wave sleep[39], which has an important role in memory. This intriguing behavior can be best understood as an evolving balance between two competing attractor states[37]. In particular, the transition does not merely involve a gradual evolution from one state to another. Instead, it showcases a period of alternating oscillation between the SO and more asynchronous states, like a back-and-forth salsa dance, with the latter progressively dominating as wakefulness approaches[37,40]. Research including studies by Tort-Colet et al.[37] and Camassa et al.[40], has effectively mapped brain states under anesthesia in rats, particularly focusing on microarousal (MA) states[41]. However, there are no computational tools capable of analyzing MA states, especially considering the challenges of inter-subject variability and the complexity of transitions between MA states[42,43].

Here we introduce a computational tool designed to (1) classify brain states that occur when going from deep anesthesia towards wakefulness and, (2) categorize the transitions between them, especially those occurring in the microarousals state. We use invasive LFP signals, providing more precise spatial resolution than non-invasive EEG. Our methodology employs deep learning techniques and dual-model Convolutional Neural Network (CNN) architecture along with an autoencoder-based clustering algorithm. This setup is designed and optimized to identify patterns with a high degree of certainty, setting a strict confidence threshold to ensure prediction accuracy. Any segments that do not meet this threshold are labeled as "unknown" and are further scrutinized through a self-supervised autoencoder-based clustering algorithm, inspired by recent advancements in computer vision research[44–47]. The latter model refines classifications by reconstructing the samples and clustering them in the frequency domain. With this layered approach, we can precisely categorize "unknown" samples as belonging to one state or transitioning between two. Overall, our methodology delivers accuracy rates that meet or exceed top-performing techniques in related fields such as sleep stage classification and comparable anesthesia studies using other techniques such as EEG[18,48,49]. In this paper, we focus on local neuronal activity and the differentiation of MA states and their transitions, areas that have recently gained experimental attention[37,39,50–52]. This highlights our approach's effectiveness and adaptability in specialized brain state categorization, including deep anesthesia scenarios and transitions. Furthermore, our work enhances current and common practices in visual inspection labeling and other manually curated methods.

Results

We analyzed in vivo cortical LFP recordings from chronically implanted rats ($n = 4$, including a total of over 60 h of recordings) to classify brain states and their transitions as the animal awakens from deep anesthesia. Our analysis focused on dynamic features of brain states, such as the amplitude, the frequency of the events, their shape, and their power spectral density (PSD). All these features are indicators used to classify the different brain states including SO, MA, and the awake state (AW). In line with recent findings, we further distinguish synchronous (slow MA) and asynchronous (asynch. MA) segments within the microarousal state[27] (Fig. 1a). We tested our dual-step machine learning system on data from four subjects across multiple sessions (up to four sessions per subject, each conducted on a separate day and lasting up to 6 h), ensuring a variety in data volume across subjects. Importantly, we employed a leave-one-out approach, meaning the data from one subject was never used in the training or validation for that same subject test.

Our method consisted of two main blocks, as shown in Fig. 1b. First, the general model, composed of two CNN models (see Supplementary Fig. 1 for an overview of the CNNs' architecture), predicts the label of the samples with a certain probability. If the probability of a sample belonging to one specific state is above the confidence level (CL), in this case 90%, they are classified as AW, SO and each type of MA. On the contrary, if the sample has a probability below the CL, it is marked as "unknown". These "unknown" samples are further processed via an autoencoder-based clustering, which provides a final prediction for every sample. The clustering method also allows the detection of the potential transitions between the MA states.

To evaluate the performance of our model (see Supplementary Table 1), we report various accuracies, including global and individual ones, for samples classified above the 90% CL, for those below the 90% CL, and for all samples, across grouped or individual states. Accuracies involving grouped states were calculated as the weighted average (in relation to the number of samples) of accuracies for each state. The global accuracy achieved, which combines high-confidence classifications and samples refined through the autoencoder-based clustering, was 91%. The global accuracy and the individual accuracies reported next highlight the robustness of the pipeline in processing both clear and ambiguous cases.

CNN classification

In the initial block of our system, all four subjects included in this study achieved high accuracies across the four main states, as shown in Fig. 2 and Supplementary Table 1. The first CNN model, which classified AW, SO, and MA for samples above the CL, achieved an accuracy of 97%, while the second CNN model, which classified slow MA and asynchronous MA, achieved an accuracy of 87%. Overall, for samples above the CL, combining both CNNs and across all subjects and states, we obtained an accuracy of 93%. For samples below the CL, we obtained an accuracy of 88% (see Supplementary Table 1). At the individual state level, the accuracy for the AW state was 81% ± 22%, while the SO and overall MA states demonstrated higher accuracies of 95% ± 4% and 98% ± 2%, respectively. Within the MA state, the asynchronous MA sub-state achieved an accuracy of 93% ± 5%, contrasting with the slow MA sub-state, which had a lower accuracy of 74% ± 38%. All the results presented, including those in Fig. 2a, use 90% as the CL. Data segments falling outside this threshold are noted as "unknown" and are processed in the subsequent autoencoder-based clustering phase of the system. Figure 2b illustrates examples of unknown samples under different confidence parameters. For a CL of 90%, the percentage of data that remains unlabeled and requires further classification is as follows: 14% ± 6% (AW), 11% ± 10% (SO), 6% ± 6% (MA), 38% ± 8% (slow MA), and 20% ± 11% (asynchronous MA) (see Supplementary Fig. 2 and Supplementary Table 1).

Autoencoder-based clustering performance

The second part of the system addresses the "unknown" samples, unclassifiable with high confidence by the CNN models. The samples are then reconstructed using autoencoders and its reconstruction is used to calculate

their PSD in the delta (0.1–1 Hz), theta (4–8 Hz) and gamma (100–500 Hz) frequency bands (Fig. 2c) (see Methods). To further distinguish between the asynchronous and synchronous phases of the MA state, we employed an unsupervised clustering algorithm that detects those samples that potentially come from transitions between the asynchronous and synchronous phases of the MA (see Methods). With the autoencoder-based clustering model we achieved individual state accuracies of 74% ± 26% (AW), 95% ± 5% (SO), 71% ± 28 (slow MA) and 60% ± 20% (asynchronous MA), see Supplementary Table 1.

In Fig. 2c we present in orange the transitions we can detect using our method. In the analysis of all sessions of one subject, most of the transitions

between the MA states we found using the method were well detected, having an average accuracy across subjects of 75% ± 6%. In Fig. 2d the traces identified as asynchronous MA, slow MA, or transitions across them are shown. Figure 2e illustrates the final accuracies of all states for the different prediction methods. The "Global" one includes all samples without adjusting for confidence levels or transitions. The "CNN" focuses only on unknown samples—those below the confidence threshold—and selects the predicted class as the one with the highest probability from the CNN model. The "centroid-with-transitions" method also considers only the unknown samples, but predicted classes are determined using a centroid-based classification method, assigning unknown samples to the closest centroid class,
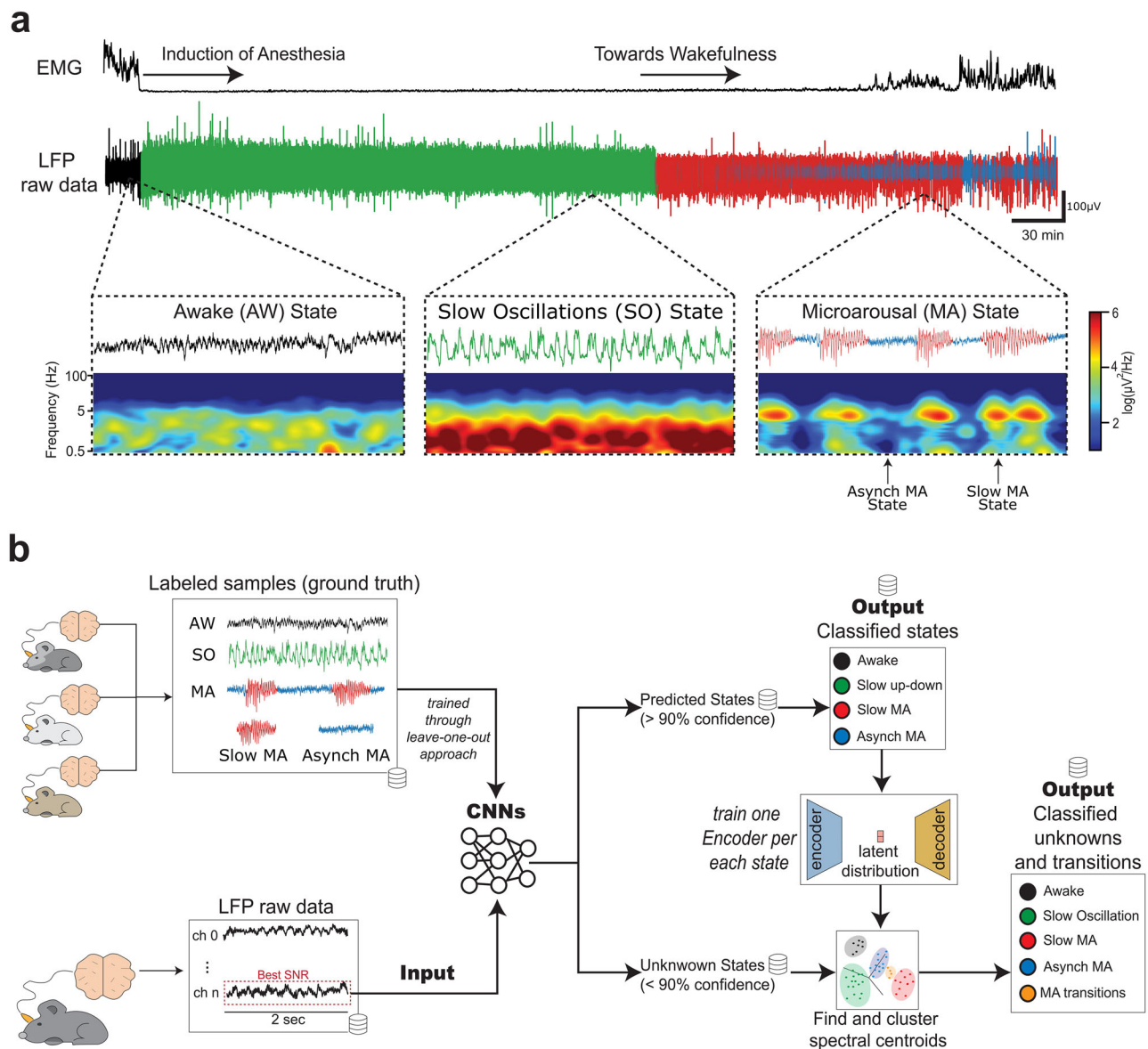

Fig. 1 | Brain state classification and network structure. a Recorded data and states. Schematics of the different Ground Truth brain states. The LFP of a full transition is shown. It starts at the Awake (AW) state and after the anesthesia induction it falls into the Slow Oscillation (SO) state. After a few hours it goes to the Microarousals (MA) state, approaching wakefulness. The samples used as input to the models are reshaped to a $M \times N$ matrix, where M is the number of samples and $N = 2000$ samples (2-s windows, fs = 1000 Hz). **b** Model overview, A model for the state classification is built using the leave-one-out strategy (i.e., the subject tested is not used for model training or validation). The model classifies the samples as the known states (AW, SO, MA, which can be divided into asynchronous MA and slow MA) if a confidence threshold (90%) is passed, otherwise, the samples are labeled as "unknown". From

the classified known samples, an autoencoder (AE) is trained on each state. Each known sample is then reconstructed with the corresponding state AE and the spectral power in the delta (0.1–4 Hz), theta (4–8 Hz), and gamma (100–500 Hz) bands is computed to find the centroids of each AE-reconstructed state. The unknown samples are also reconstructed with the AE that gives the minimum Mean Squared Error (MSE). The same spectral power is then calculated and projected over the known samples space, where these unknown samples are re-classified as the state whose centroid is the closest. We here add a new centroid, in between slow MA and asynchronous MA centroids to account for samples that may contain a transition between these two states.

including the computed "Transitions" centroid (see Methods). Lastly, the "centroid-no-transitions" method also applies a centroid-based classification but excludes samples belonging to the MA transitions centroid when determining the predicted classes.

The decline in accuracy from Global to CNN results from adding all samples in the Global approach, including those above the 90% confidence threshold. As evident in Fig. 2a, b and Supplementary Fig. 3, the accuracy of these samples is notably high. The decrease observed from CNN to centroid-with-transitions classified samples is attributed to the absence of transition labels (slow MA to async MA or vice versa) in the Ground Truth labels. The increase in accuracy of "centroid-no-transitions" highlights the contribution of the autoencoders to a better classification of those samples with a low confidence level, which in the case of AW and SO does not vary much, but for the two substates of the MA the accuracy increases with respect to the CNN-labeled "unknown" samples, giving autoencoders and the corresponding spectral clustering of the autoencoder-reconstructed samples a higher predictive power.

Transfer learning

We quantified the efficacy of transfer learning, with a focus on achieving high accuracy with minimal training sessions (Supplementary Fig. 4). This is exemplified in Supplementary Fig. 4a, b, where we observed a marked increase in mean accuracy as we conducted training over two to seven

sessions, using the predictors for sessions of the same subject. We attained a stable and high level of accuracy with as few as five training sessions across various states. Moreover, we found that the percentage of unknown samples decreased as we used more sessions, resulting in higher accuracy and confidence in the predictions. This inverse relationship between the number of sessions and the percentage of unknowns underscored the efficiency of our training process. The same results were evident when we employed a leave-one-out approach, similar to the rest of the study, as shown in Supplementary Fig. 4b, further highlighting the relationship between the increasing number of training sessions and the improvement in model performance.

Discussion

In recent decades, machine learning decoding algorithms have significantly advanced clinical applications for neurological disorders, particularly in the context of decoding for brain-computer interfaces[53]. These advancements have largely occurred within the field of brain-machine interfaces, which employ both invasive[54,55] and non-invasive methods[56–60] to decode movement commands. The use of these algorithms has recently allowed unprecedented advances in speech decoding[54,61], facilitating social communication and thus improving quality of life for these patients. Additionally, there is growing interest in applying similar machine learning approaches, including neural networks and autoencoders, to classify various brain states, not only in non-invasive studies of sleep and workload with EEG[18,48,62] but also in
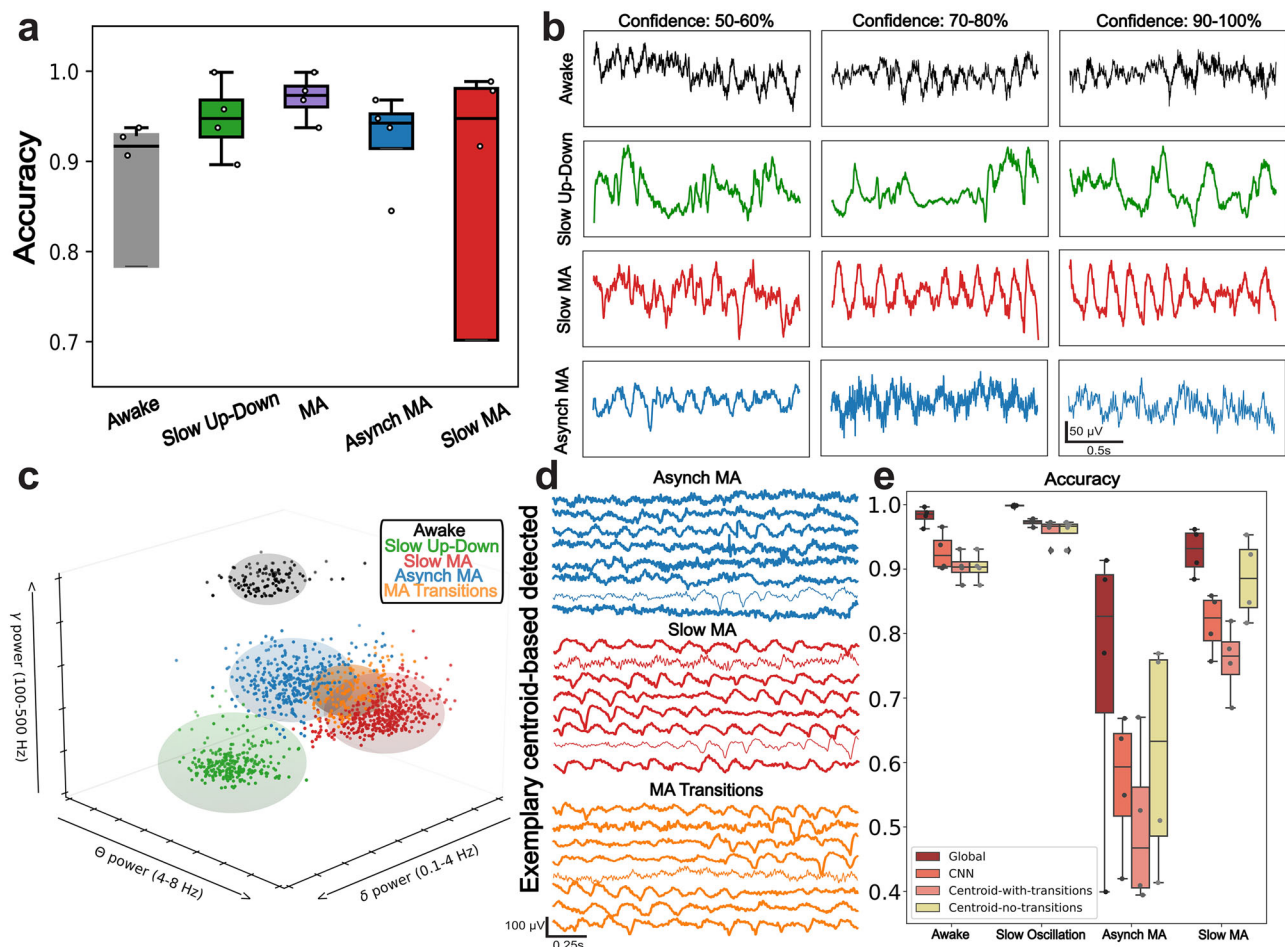

Fig. 2 | Model performance. a Averaged accuracies of the first CNN) model in classifying Awake (AW), Slow oscillations (SO) and Microarousals (MA) states in all subjects ($n = 4$) with a confidence threshold of 90% (unknowns are excluded from the confusion matrix). **b** Exemplary samples. Different state traces with the corresponding confidence level. **c** Power spectral density (PSD) clustered samples. The spectral power in the different bands was computed in the autoencoder-reconstructed unknown samples. **d** Exemplary unknowns classified with the PSD-cluster algorithm. **e** Accuracy across different models, Global, CNN, centroids-with-transitions and centroids-with-no-transitions. Boxplots show IQR, line is median, whiskers extend to 1.5× IQR or min/max non-outliers.

pathologies such as disorders of consciousness[18,63]. However, the application of machine learning algorithms to invasive LFPs for tracking spontaneous brain states and their transitions, through the analysis of invasive, high spatial resolution LFPs, remains relatively unexplored, as highlighted in this paper. A recent study investigating LFPs for different awake states[64] (thus different states to those studied here) used a different approach with unsupervised Gaussian mixture models and hidden semi-Markov model.

In this study, we presented a novel computational tool designed to classify and decode brain states across varying anesthesia levels, filling a noticeable gap in efficient labeling and anesthesia brain state categorization techniques. Our framework combines a one-dimensional-CNN approach to classify brain states in anesthesia (awake, AW; slow oscillations, SO; microarousals, MA). Moreover, with an additional classifier, we studied the differentiation within the MA states (namely asynchronous MA and slow MA), reflecting the latest advancements in the study of microarousal states[37,39,40] We achieved both high global and individual accuracies for each state. We achieved a 91% global accuracy for the full pipeline, which integrates high-confidence classifications from the dual-model CNNs with samples below the confidence level that were subsequently classified with the autoencoder-based clustering model. While we consider overall accuracy to be a crucial performance metric for the practical deployment of our model on arbitrary data, we also believe that evaluating the error rate for each brain state is essential. This approach provides a more detailed understanding of the model's strengths and weaknesses across different conditions. What sets our framework apart is its second phase—the autoencoder-based clustering strategy. The combination of CNNs and autoencoders in this dual-model training approach not only empowers neuroscientists to detect and classify samples that are challenging under visual inspection but also offers flexibility in dataset classification and the identification of transitions during state detection.

We wish to highlight that our machine-learning-based pipeline provides a specialized framework for detecting state transitions according to frequency bands following the reconstruction of carefully trained autoencoders. The use of these autoencoders is crucial in our pipeline, as they help refine the classification of samples that were either not predicted with high confidence in the initial models or were involved in a certain state transition. As part of our analysis, we found that the autoencoder's reconstruction provided better separation of the clusters, enabling clearer identification of the transitions. This improvement was particularly evident in the spectral power computed from the autoencoder-reconstructed samples, which offered a better cluster separation compared to the PSD obtained from the original time series (see Supplementary Fig. 5). In addition to improving cluster separation, we also explored the potential of autoencoders functioning as complex frequency filters. Our analysis reveals that the autoencoders, when trained on specific data subsets, can effectively alter the spectral composition of the signals. For example, autoencoders trained on 'Slow Oscillation' and 'Slow MA' samples were found to remove higher frequencies from synthetic signals, whereas those trained on 'Awake' and 'Asynchronous MA' samples did not exhibit this filtering effect (see Supplementary Fig. 6). This observation underscores the autoencoder's role in selectively processing the frequency content of the signal based on the underlying state. These findings highlight the autoencoder's dual function both as a classifier and as a frequency filter, which enhances the robustness of state transition detection.

We would also like to emphasize the strong performance of the CNN models in classifying instances of low-confidence data. We have focused on clustering within frequency bands to identify different states, which offers excellent discrimination power given that each band is representative of one state[46,65,66]. However, it is important to note that this is not the sole indicator for discerning state transitions. We also introduce a "confidence" parameter, which enables us to separate those samples that have been predicted with a high probability from those that have not. For future specific experiments where data quality is crucial, it will be extremely important how this confidence parameter is tweaked to ensure highly accurate predictions. On the

other hand, if the experiment requires a more lenient data quality, confidence levels can be adjusted accordingly. Since a lower confidence level suggests that a sample is less similar to its average spectral state characteristics, it would be logical to explore whether the confidence parameter could serve as a predictor of the samples' positions in frequency space, potentially aiding in the identification of state transitions. In our case there seems to be a strong correlation. We could eventually achieve higher accuracy in MA states at the expense of lower accuracy in AW or SO states, while incorporating other data modalities (such as electromyography (EMG) or video tracking) to fully label these states. This approach could ultimately lead to a more accurate and robust classification of the states in question, which is crucial for the objectives of our study.

Another key strength of this work is the specificity targeting intracortical LFP data in brain states occurring under anesthesia. Experimental neuroscientists using invasive approaches often struggle to access numerous valuable data channels; in response, our framework emphasizes simplicity. The implementation of a computationally efficient CNN with a single-channel input also facilitates the training phase. Recent studies demonstrate that, even with few layers, CNN models can offer significant insights into brain dynamics[50,66]. Simplicity and adaptability make it versatile for different subjects and sessions, optimizing the analysis of varied datasets. Another introduced feature is that, to cater to low-frequency classifications, we have implemented a sinusoid activation function, a method inspired by previous studies[67], as an alternative to the traditional ReLU (Rectified Linear Unit). However, when comparing both activation functions, no substantial differences in performance were observed, indicating that the choice of the sinusoidal activation function should not be viewed as critical to the success of the model.

Moreover, our decoding approach also has some limitations. The framework, as it stands, is optimized for a specific experiment dealing with standard anesthesia levels and then adds another classifier specifically to differentiate MA states. It is fine-tuned to tackle the need for reliable computational tools for automatic brain state classification during anesthesia stages. To use the same framework for other purposes (e.g., sleep), some refinement of the models, a reliable ground truth labels annotation, adapted preprocessing, and novel training of the models from scratch would be required. For instance, the current setup processes data in 2-s time windows. While this window suits our experiment, those delving into other brain state classifications might need to adjust this window size, with the trade-off of additional computational resources. Additionally, the dataset used in this study presents some degree of data imbalance given the duration of the awake periods, which were shorter than the other states. After randomly permuting the labels, our analysis confirms that the high accuracies obtained by the model are due to its effective performance rather than a consequence of data imbalance or random guessing. Still, whenever possible, larger data balance and larger numbers of subjects would be optimal, enhancing the robustness of the findings and avoiding additional model testing.

We acknowledge that there are areas where caution is necessary. For different applications, or when dealing with new datasets, pre-trained models might not perform well enough. In this paper, our goal was to focus on the shape and signature of the waves of the different states. In scenarios where specific state transitions (and their timings) are the focus, incorporating an RNN (Recurrent Neural Network)-based approach might be more apt. And while our framework aids in decision-making, determining hyperparameters of confidence still mandates expert visual inspection. Another potential limitation was whether selecting the autoencoder based on the previous CNN prediction introduced a potential bias in the overall accuracy, which we inspected and confirmed did not occur in our case.

In comparison with other models, our system can autonomously verify and highlight complex samples for precise classifications. In terms of accuracy, our approach yields comparable or even superior labeling outcomes to similar studies, which achieved 70–90% accuracy across subjects. Subjects with clearer signals experienced high accuracy across states. In

contrast, one subject's model occasionally confused the Awake state with the combined microarousal, attributed to the similarity in shape between asynchronous microarousal and awake segments. The fact that our method was designed to work as a leave-one-out approach, offers significant potential for generating *general* models that could be used for new subjects, meaning that every dataset we incorporate in our studies could make the model more generalizable and more robust. In this case, transfer learning nuances might have to be considered in the future[68]. Also, this gives us some pre-trained models that offer huge potential for deployment in micro-controllers for on-device inferences in the field of neuroscience and cognitive sciences, an area that has gained attention recently in preliminary studies[69,70].

Future perspectives of our approach in the medical field are worthy of exploration. Given its potential for analyzing single-channel activity, we could use signals from different areas both from multichannel recordings in experimental neuroscience to map across the brain of the behaviors detected, or from intracranial stereo EEG from humans in pathological conditions to identify anomalous local patterns.

Methods

End-to-end system overview

Our approach employs a two-step method that integrates two Convolutional Neural Networks (CNNs) and one autoencoder to classify brain states based on 2-s-long sequential data represented as a 1D array. This methodology aims to offer both a broad and refined classification and analysis of brain states.

Preliminary classification

The initial classification (step 1) employs a CNN with three convolutional layers. This model provides a general classification into three primary states: Awake (AW), Slow (SO) and Microarousals (MA). If this first CNN predicts a MA state, the data is further processed by a second CNN (step 2). This second CNN classifies the MA state into two substates: the asynchronous MA (asynch. MA) and the slow MA (slow MA) phases. Thus, after steps 1 and 2, we obtain four classified states, each with a given probability, namely: AW, SO, asynchronous MA and slow MA.

For both CNNs, we use a Confidence Level (CL) thresholding technique, inspired by semi-supervised self-labeling in image processing[71]. These models predict the class of each sample by estimating a probability. In our case, we set the CL at 90%. If the probability of a sample belonging to a specific state exceeds the CL, it is classified as AW, SO, or one of the MA phases. Conversely, if the probability falls below the CL, the sample is labeled as "unknown". These "unknown" samples are then further analyzed using an autoencoder-based clustering method, which ultimately provides a final classification for each sample (see next section).

Handling "unknown" and transition states with autoencoders

States classified as "unknown" resulting fromby the CNN are further processed using autoencoders. These unsupervised neural networks learn to recreate their input, and in our method, they play a pivotal role in classifying ambiguous data. We use the "unknown" samples from the CNN models (those that did not surpass the confidence level (CL) as an input for each autoencoder, which has been trained with the classified data from the CNN that did surpass the CL. Each unknown sample is reconstructed with the autoencoder of the brain state that obtained the highest probability from the CNN models.

This is followed by computing the PSD of the reconstructed samples, either above or below the CL, across three primary frequency bands: delta (0.1–1 Hz), theta (4–8 Hz) and gamma (100–500 Hz)[40], leading to a re-clustering of samples. The points belonging to the power in specific frequency bands of the reconstructed samples above the CL are projected onto the three-dimensional space (three frequency bands). A centroid for each state is computed as the center of mass corresponding to the specific state-labeled samples above the CL. A centroid to account for the transitions across microarousal states was manually added as the middle point that

separates the slow MA and asynchronous MA centroids. Each unknown sample is then projected onto the 3D space, where the centroids of the reconstructed samples above CL will determine to which state each unknown sample is assigned to. Any reconstructed sample closer to the "transitions" centroid than to any other will be automatically classified as a "transition".

Preliminary classification models - CNNs architecture

Both CNN models share a similar architecture with minimal nuances. The first CNN, responsible for general state classification, consists of three convolutional blocks with ReLU activation functions, followed by max-pooling layers (see Supplementary Fig. 1). The final layers include a flattening layer, a fully connected layer, and an output layer. The output provides the classification into the AW, SO, and MA states. Similar classification structures have been used in previous literature[50,72]. The second CNN, responsible for classifying the MA state into its two substates, consists of two convolutional layers with sinusoidal activation functions, inspired by relevant literature[67]. The sinusoidal activation function used in our experiments is defined as $f(x) = \sin(x)$, a simple function that does not include any additional hyperparameters or learnable parameters. The function outputs values in the range $[-1,1]$, introducing a periodic nature to the activation. This characteristic was initially chosen based on its theoretical relevance to our empirical data, where the slow MA substate is characterized by low-frequency components. However, when compared to the ReLU activation function, no substantial differences in performance were observed. Specifically, the performance for the asynchronous MA (slow MA) substate was 90% ± 7% (76% ± 38%) with ReLU and 93% ± 5% (74% ± 38%) with the sinusoidal activation. These results suggest that the specific periodic properties of the sinusoidal activation function may not be necessary to solve this task. The use of the sinusoidal activation function in this study should therefore be viewed as an exploratory alternative rather than a strict requirement. The second CNN also includes max-pooling and flattening operations and a dense layer with a sigmoid activation function for binary classification.

Autoencoders architecture

A single layer autoencoder was implemented for dimensionality reduction and feature extraction. The autoencoder architecture consists of an encoder and a decoder, designed to learn a compressed representation of the data by minimizing the reconstruction error. The encoder compresses the input data into a lower-dimensional latent space, represented by a hidden layer, which is often referred to as the 'bottleneck' layer. This hidden layer contains 75 neurons, which correspond to the encoding dimension. The size of this hidden layer defines the extent of data compression and captures the essential features of the input while discarding redundant information. The input layer feeds into this hidden layer using a ReLU activation function, facilitating non-linear transformations of the input features. Subsequently, the decoder reconstructs the original input data from this compressed latent space. The decoder consists of a single output layer with the same number of neurons as the input layer, ensuring that the reconstructed output has the same dimensionality as the original input. A linear activation function is employed in the output layer to allow the network to learn the identity mapping necessary for accurate reconstruction.

Models training

The system training process involves training two CNN models and a number of autoencoders equal to the total classes targeted for classification and refinement. Initially, we created concatenated matrices for each state using the corresponding data samples. This ensures that each state's training data exclusively consists of samples relevant to that specific state. This is accomplished by concatenating traces from the same state, based on the ground truth annotation, and transforming the final array into a matrix whose rows will be samples of 2000 points (2 s, fs (sampling rate) = 1 kHz). A separate array indicates the ground truth label for each row.

Training CNN models

Both CNN models are trained using the Adam optimizer and cross-entropy as the loss function. For the first CNN, training and validation data are loaded in mini batches with varying sizes. For the second CNN, training data is filtered to include only instances belonging to MA states. An 80–20 split is employed to create training and validation subsets, respectively. The model is trained for 10 epochs, using a batch size of 32. To facilitate model evaluation and potential resumption of training, a custom callback function is used to save the model every 5 epochs. Performance metrics (loss and accuracy) are monitored for both the training and validation sets.

Training autoencoders

To handle samples that the initial CNN models could not classify with high confidence (CL < 90%), we use specialized autoencoders tailored for each state we aim to study. These autoencoders are trained using samples that the preliminary models classified confidently (CL > 90). For example, to train the autoencoder for the "awake" state, we use samples identified as "awake" by the initial CNN model. The input dataset is split into training and testing sets, with the proportion of the split depending on the total size of the state-specific dataset: 10% for the awake class, given the small number of samples available and 20% for the other classes. These testing sets act as classification thresholds for each state. Model optimization uses the Adam optimizer to minimize the Mean Squared Error (MSE) over a specific number of epochs and a predetermined batch size.

Training set size and data efficiency analysis

We explored the optimal training set sizes required to attain satisfactory testing outcomes. Specifically, our aim was to determine the fewest sessions needed for training to accurately predict the three primary states: AW, SO, MA. Our approach involved varying the training set sizes: using two, five and seven sessions to forecast one random session of the testing subjects. For the pairing of train/test data, we employed a k-fold algorithm and further juxtaposed it with a modified k-fold strategy to evaluate its impact on the leave-one-out approach. All the parameters of these new CNN models are the same as in the Models architecture: CNN models, explained above.

Dataset

The dataset used in this study originates from the same set of experiments described in Torao-Angosto et al.[21]. Briefly, cortical LFPs from different brain areas and EMG were recorded from four freely moving Lister Hooded rats (6–10 months old) spanning multiple sessions (4, 2, 2, and 1 session, respectively), each session lasting an average of 6 ± 1 h. Although recordings were obtained from multiple areas, with up to five electrodes per animal, only the electrode with the highest signal-to-noise ratio (SNR) was used for the current analysis. Chronically implanted electrodes in each animal were located in the orbitofrontal cortex ($n = 2$) and the primary motor cortex ($n = 2$).

LFPs were acquired in different brain states (Fig. 1A). First, a baseline recording was obtained while the animal was awake. Then, the animal was sedated and anesthetized using a combination of ketamine (50 mg/ml) and medetomidine (1 mg/ml). Cortical activity was recorded from the induction of the anesthesia, through deep anesthesia and then until it showed the first signs of wakefulness. During this process, the animal went through the SO and the MA states before showing signs of awakening. Slow oscillations consist of an almost periodic pattern (<1 Hz) of Up (active) and Down (silent) states[73], and are referred in the figures as Slow Up-Down. Conversely, microarousals consists of periods of slow oscillatory activity (~5 Hz) interspersed by periods of awake-like activity[37]. In the manuscript, we refer to the former as slow MA (slow MA) and the latter as asynchronous MA (asynch. MA). The awake state is characterized by irregular and high-frequency fluctuations in the LFP and muscle activity (EMG)[21]. The LFP traces were normalized and downsampled to 1 kHz for computational reduction purposes, as in previous studies[50]. For a detailed description see Torao-Angosto et al.[21]. All experiments were carried out following Spanish regulatory laws (BOE-A-2013-1337), which comply with European Union

guidelines and were supervised and approved by the Animal Experimentation Ethics Committee of the Universitat de Barcelona (287/17 P3).

It is important to note the distribution of data within the dataset. Due to the brief duration of awake recordings, which were performed prior to anesthesia injection, these segments only comprised approximately 4% ± 3% of the total dataset. In contrast, SO and MA were more prevalent, constituting 48% ± 16% and 59% ± 16% of the segments, respectively. The number of segments used in this work is approximately 50k segments of 2 s each, which is about 28 h of recordings (after data pre-processing, e.g., removing artifacts) in total. To ensure that this distribution did not adversely impact our results, we computed the overall performance after randomly permuting the labels. Our analysis confirms that the high accuracies achieved by the model are due to its effective performance, rather than a consequence of data imbalance or random guessing.

Ground truth labeling

The ground truth annotations were carried out by an expert. This process involved defining three significant global states: Awake, Slow Oscillation and the asynchronous and synchronous parts of the Microarousals. Ground truth annotations were applied to the LFP channel exhibiting the highest SNR. The EMG channel, positioned around the animal's neck, was utilized to monitor animal movements and served as a supplementary source of information for this labeling process. The AW state pertained to time periods showing asynchronous cortical activity recorded before anesthesia administration. During this phase, the animal exhibited clear signs of wakefulness, such as responsiveness and movement, accompanied by a high EMG amplitude. SO labels were assigned to time segments characterized by a distinct Up/Down pattern with a frequency below 2.5 Hz. These segments immediately followed anesthesia induction and typically persisted for 2–3 h. The animal displayed no signs of wakefulness, and the EMG exhibited a flat profile. MA labels were designated for time samples that occurred once the stable SO state had concluded. These samples were expected to exhibit a shift toward wakefulness, as detailed in Tort-Colet et al.[37] They consisted of brief periods of asynchronous activity interspersed with intervals of slow oscillation characterized by a notably higher frequency (~6 Hz) compared to the typical SO observed after anesthesia induction. During the MA state, the animal remained unresponsive but began to display drowsiness-related symptoms, which were validated through EMG and video recordings.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The demonstration dataset is publicly available at https://doi.org/10.5281/zenodo.14990181. Other data are available from the corresponding author on reasonable request.

Code availability

The code used in this work is publicly available and will be publicly maintained and updated at https://github.com/arnaumarin/LFPDeepStates and is available at ref. 74.

References

1. McCormick, D. A., Nestvogel, D. B. & He, B. J. Neuromodulation of Brain State and Behavior. *Annu Rev. Neurosci.* **43**, 391–415 (2020).
2. Zagha, E. & McCormick, D. A. Neural control of brain state. *Curr. Opin. Neurobiol.* **29**, 178–186 (2014).
3. Zelmann, R. et al. Differential cortical network engagement during states of un/consciousness in humans. *Neuron* **111**, 3479–3495 e3476 (2023).

4. Barttfeld, P. et al. Signature of consciousness in the dynamics of resting-state brain activity. *Proc. Natl Acad. Sci. USA* **112**, 887–892 (2015).

5. Sanz Perl, Y., Escrichs, A., Tagliazucchi, E., Kringelbach, M. L. & Deco, G. Strength-dependent perturbation of whole-brain model working in different regimes reveals the role of fluctuations in brain dynamics. *PLoS Comput. Biol.* **18**, e1010662 (2022).

6. Tauber, J. M. et al. Propofol-mediated Unconsciousness Disrupts Progression of Sensory Signals through the Cortical Hierarchy. *J. Cogn. Neurosci.* **36**, 394–413 (2024).

7. Dasilva, M. et al. Modulation of cortical slow oscillations and complexity across anesthesia levels. *Neuroimage* **224**, 117415 (2021).

8. Rosanova, M. et al. Sleep-like cortical OFF-periods disrupt causality and complexity in the brain of unresponsive wakefulness syndrome patients. *Nat. Commun.* **9**, 4427 (2018).

9. Russo, S. et al. Focal lesions induce large-scale percolation of sleep-like intracerebral activity in awake humans. *Neuroimage* **234**, 117964 (2021).

10. Massimini, M. et al. Sleep-like cortical dynamics during wakefulness and their network effects following brain injury. *Nat. Commun.* **15**, 7207 (2024).

11. Sekkal, R. N., Bereksi-Reguig, F., Ruiz-Fernandez, D., Dib, N. & Sekkal, S. Automatic sleep stage classification: From classical machine learning methods to deep learning. *Biomed. Signal Process. Control* **77**, 103751 (2022).

12. Aboalayon, K. A., Almuhammadi, W. S. & Faezipour, M.Efficient obstructive sleep apnea classification based on EEG signals. In *2015 Long Island Systems, Applications and Technology*. 1-6 (IEEE, 2015).

13. Liu, Q. et al. Spectrum analysis of EEG signals using CNN to model patient's consciousness level based on anesthesiologists' experience. *IEEE Access* **7**, 53731–53742 (2019).

14. Patlatzoglou, K. et al. Deep Neural Networks for Automatic Classification of Anesthetic-Induced Unconsciousness. In *Brain Informatics: International Conference, BI 2018*. 216–225 (Springer, 2018).

15. Afshar, S., Boostani, R. & Sanei, S. A combinatorial deep learning structure for precise depth of anesthesia estimation from EEG signals. *IEEE J. Biomed. Health Inform.* **25**, 3408–3415 (2021).

16. Gu, Y., Liang, Z. & Hagihira, S. Use of multiple EEG features and artificial neural network to monitor the depth of anesthesia. *Sensors* **19**, 2499 (2019).

17. van der Lande, GJM et al. Brain state identification and neuromodulation to promote recovery of consciousness. *Brain Commun.* **6**, fcae362 (2024).

18. Lee, M. et al. Quantifying arousal and awareness in altered states of consciousness using interpretable deep learning. *Nat. Commun.* **13**, 1064 (2022).

19. Gervasoni, D. et al. Global forebrain dynamics predict rat behavioral states and their transitions. *J. Neurosci.* **24**, 11137–11147 (2004).

20. Greene, A. S., Horien, C., Barson, D., Scheinost, D. & Constable, R. T. Why is everyone talking about brain state? *Trends Neurosci.*, **46**, 508-524 (2023).

21. Torao-Angosto, M., Manasanch, A., Mattia, M. & Sanchez-Vives, M. V. Up and Down States During Slow Oscillations in Slow-Wave Sleep and Different Levels of Anesthesia. *Front. Syst. Neurosci.* **15**, 609645 (2021).

22. Neske, G. T. The Slow Oscillation in Cortical and Thalamic Networks: Mechanisms and Functions. *Front. Neural Circuits* **9**, 88 (2015).

23. Steriade, M., Nunez, A. & Amzica, F. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *J. Neurosci.* **13**, 3252–3265 (1993).

24. Camassa, A., Galluzzi, A., Mattia, M. & Sanchez-Vives, M. V. Deterministic and Stochastic Components of Cortical Down States: Dynamics and Modulation. *J. Neurosci.* **42**, 9387–9400 (2022).

25. Huber, R., Ghilardi, M. F., Massimini, M. & Tononi, G. Local sleep and learning. *Nature* **430**, 78–81 (2004).

26. Casali, A. G. et al. A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* **5**, 198ra105 (2013).

27. Sanchez-Vives, M. V., Massimini, M. & Mattia, M. Shaping the Default Activity Pattern of the Cortical Network. *Neuron* **94**, 993–1001 (2017).

28. Steriade, M., McCormick, D. A. & Sejnowski, T. J. Thalamocortical oscillations in the sleeping and aroused brain. *Science* **262**, 679–685 (1993).

29. Barbero-Castillo, A. et al. Impact of GABA(A) and GABA(B) Inhibition on Cortical Dynamics and Perturbational Complexity during Synchronous and Desynchronized States. *J. Neurosci.* **41**, 5029–5044 (2021).

30. D'Andola, M. et al. Bistability, Causality, and Complexity in Cortical Networks: An In Vitro Perturbational Study. *Cereb. Cortex* **28**, 2233–2242 (2018).

31. Kringelbach, M. L. & Deco, G. Brain States and Transitions: Insights from Computational Neuroscience. *Cell Rep.* **32**, 108128 (2020).

32. Mukamel, E. A. et al. A transition in brain state during propofol-induced unconsciousness. *J. Neurosci.* **34**, 839–845 (2014).

33. Harris, K. D. & Thiele, A. Cortical state and attention. *Nat. Rev. Neurosci.* **12**, 509–523 (2011).

34. Brown, E. N., Purdon, P. L. & Van Dort, C. J. General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annu Rev. Neurosci.* **34**, 601–628 (2011).

35. Bardon, A. G. et al. Convergent effects of different anesthetics are due to changes in phase alignment of cortical oscillations. *bioRxiv* **03**, 2024 (2024).

36. Smerieri, A., Parrino, L., Agosti, M., Ferri, R. & Terzano, M. G. Cyclic alternating pattern sequences and non-cyclic alternating pattern periods in human sleep. *Clin. Neurophysiol.* **118**, 2305–2313 (2007).

37. Tort-Colet, N., Capone, C., Sanchez-Vives, M. V. & Mattia, M. Attractor competition enriches cortical dynamics during awakening from anesthesia. *Cell Rep.* **35**, 109270 (2021).

38. Pillay, S., Vizuete, J. A., McCallum, J. B. & Hudetz, A. G. Norepinephrine infusion into nucleus basalis elicits microarousal in desflurane-anesthetized rats. *J. Am. Soc. Anesthesiologists* **115**, 733–742 (2011).

39. Kjaerby, C. et al. Memory-enhancing properties of sleep depend on the oscillatory amplitude of norepinephrine. *Nat. Neurosci.* **25**, 1059–1070 (2022).

40. Camassa, A. et al. The temporal asymmetry of cortical dynamics as a signature of brain states. *Sci. Rep.* **14**, 24271 (2024).

41. Cancino-Fuentes, N. et al. Recording physiological and pathological cortical activity and exogenous electric fields using graphene microtransistor arrays in vitro. *Nanoscale* **16**, 664–677 (2024).

42. Einevoll, G. T., Kayser, C., Logothetis, N. K. & Panzeri, S. Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nat. Rev. Neurosci.* **14**, 770–785 (2013).

43. Linden, H. et al. Modeling the spatial reach of the LFP. *Neuron* **72**, 859–872 (2011).

44. Roth, K., Vinyals, O. & Akata, Z. Non-Isotropy Regularization for Proxy-Based Deep Metric Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7420–7430 (IEEE, 2022).

45. Roth, K., Vinyals, O. & Akata, Z. Integrating Language Guidance Into Vision-Based Deep Metric Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 16177–16189 (IEEE, 2022).

46. Sehwag, V., Chiang, M. & Mittal, P. Ssd: A unified framework for self-supervised outlier detection. Preprint at http://arxiv.org/abs/2103.12051 (2021).

47. Desai, A., Freeman, C., Wang, Z. & Beaver, I. Timevae: A variational auto-encoder for multivariate time series generation. Preprint at http://arxiv.org/abs/2111.08095 (2021).

48. Mousavi, S., Afghah, F. & Acharya, U. R. SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach. *PloS One* **14**, e0216456 (2019).

49. Zhang, X. et al. Automated sleep state classification of wide-field calcium imaging data via multiplex visibility graphs and deep learning. *J. Neurosci. Methods* **366**, 109421 (2022).

50. Marin-Llobet, A., Manasanch, A. & Sanchez-Vives, M. V. Hopfield-Enhanced Deep Neural Networks for Artifact-Resilient Brain State Decoding. Preprint at http://arxiv.org/abs/2311.03421 (2023).

51. Liang, Y. et al. Complexity of cortical wave patterns of the wake mouse cortex. *Nat. Commun.* **14**, 1434 (2023).

52. Blackwood, E. B., Shortal, B. P. & Proekt, A. Weakly correlated local cortical state switches under anesthesia lead to strongly correlated global states. *J. Neurosci.* **42**, 8980–8996 (2022).

53. Myszczynska, M. A. et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat. Rev. Neurol.* **16**, 440–456 (2020).

54. Metzger, S. L. et al. A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* **620**, 1037–1046 (2023).

55. Gallego, J. A., Perich, M. G., Miller, L. E. & Solla, S. A. Neural Manifolds for the Control of Movement. *Neuron* **94**, 978–984 (2017).

56. Del, R. M. J. et al. A local neural classifier for the recognition of EEG patterns associated to mental tasks. *IEEE Trans. Neural Netw.* **13**, 678–686 (2002).

57. Ang, K. K. & Guan, C. EEG-based strategies to detect motor imagery for control and rehabilitation. *IEEE Trans. Neural Syst. Rehabilit. Eng.* **25**, 392–401 (2016).

58. Herman, P., Prasad, G., McGinnity, T. M. & Coyle, D. Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. *IEEE Trans. Neural Syst. Rehabilit. Eng.* **16**, 317–326 (2008).

59. Tabar, Y. R. & Halici, U. A novel deep learning approach for classification of EEG motor imagery signals. *J. Neural Eng.* **14**, 016003 (2017).

60. Orset, B., Lee, K., Chavarriaga, R. & Millan, J. D. R. User Adaptation to Closed-Loop Decoding of Motor Imagery Termination. *IEEE Trans. Biomed. Eng.* **68**, 3–10 (2021).

61. Willett, F. R. et al. A high-performance speech neuroprosthesis. *Nature* **620**, 1031–1036 (2023).

62. Yu, Y., Bezerianos, A., Cichocki, A. & Li, J. Latent Space Coding Capsule Network for Mental Workload Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **31**, 3417–3427 (2023).

63. Wang, F., Tian, Y. -C., Zhang, X. & Hu, F. Detecting disorders of consciousness in brain injuries from EEG connectivity through machine learning. *IEEE Trans. Emerg. Top. Computational Intell.* **6**, 113–123 (2020).

64. Weiss, D. A., Borsa, A. M., Pala, A., Sederberg, A. J. & Stanley, G. B. A machine learning approach for real-time cortical state estimation. *J Neural Eng* **21**. https://doi.org/10.1088/1741-2552/ad1f7b (2024).

65. Camassa, A. et al. Chronic full-band recordings with graphene microtransistors as neural interfaces for discrimination of brain states. *Nanoscale Horiz.* **9**, 589–597 (2024).

66. Maheswaranathan, N. et al. Interpreting the retinal neural code for natural scenes: From computations to neurons. *Neuron* **111**, 2742–2755 e2744 (2023).

67. Ramachandran, P., Zoph, B. & Le, Q. V. Searching for activation functions. Preprint at http://arxiv.org/abs/1710.05941 (2017).

68. Weiss, K., Khoshgoftaar, T. M. & Wang, D. A survey of transfer learning. *J. Big data* **3**, 1–40 (2016).

69. Wang, Y., Nahon, R., Tartaglione, E., Mozharovskyi, P. & Nguyen, V.-T. Optimized preprocessing and Tiny ML for Attention State Classification. In *2023 IEEE Statistical Signal Processing Workshop (SSP)*. 695–699 (IEEE).

70. Nguyen, V.-T., Tartaglione, E. & Dinh, T. AIoT-based Neural Decoding and Neurofeedback for Accelerated Cognitive Training: Vision, Directions and Preliminary Results. In *2023 IEEE Statistical Signal Processing Workshop (SSP)*. 705–709 (IEEE).

71. Chapelle, O., Scholkopf, B. & Zien. *Semi-Supervised Learning* (MIT Press, 2006).

72. Navas-Olive, A., Rubio, A., Abbaspoor, S., Hoffman, K. L. & de la Prida, L. M. A machine learning toolbox for the analysis of sharp-wave ripples reveals common waveform features across species. *Commun. Biol.* **7**, 211 (2024).

73. Ruiz-Mejias, M., Ciria-Suarez, L., Mattia, M. & Sanchez-Vives, M. V. Slow and fast rhythms generated in the cerebral cortex of the anesthetized mouse. *J. Neurophysiol.* **106**, 2910–2921 (2011).

74. Marin-Llobet, A. & Manasanch, A. (2025)/lfpdeepstates: Communications Biology. Version, Mar 2025. v.1.0 (Zenodo, 2025). https://doi.org/10.5281/zenodo.15066751.

Author contributions
A.M.L., A.M., L.D.P. and M.S.V. conceived the idea and designed the research. A.M.L. and A.M. designed and developed the neural models. M.T-A obtained the brain recordings. A.M. and M.T.-A. provided the ground truth analysis. A.M.L. and A.M. designed and developed the data pipeline. A.M.L., A.M., L.D.P. and M.S.V. provided critical discussions during the development. A.M.L. and A.M. conducted analyses on neural recording datasets. A.M.L., A.M., L.D.P., and M.S.V. prepared figures and wrote the manuscript. M.S.V. supervised the study.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-07991-3.

Correspondence and requests for materials should be addressed to Maria V. Sanchez-Vives.

Peer review information *Communications Biology* thanks Hernán Bocaccio and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Enzo Tagliazucchi and Joao Valente.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.