eXCube1: Explainable Neuromorphic Framework for Modelling Conscious Perception of Stimuli from fMRI Data

Nik Kirilov Kasabov *, Alexander Yang , Iman Abouhassan , Calum Chace , Assia Nikolova Kassabova , Teodoros Lappas

Posted Date: 8 April 2026

doi: 10.20944/preprints202604.0605.v1

Keywords: spiking neural networks; NeuCube; evolving spatio-temporal associative memory; music; video data

 doi:10.20944/preprints202604.0605.v1

Article

eXCube1: Explainable Neuromorphic Framework for Modelling Conscious Perception of Stimuli from fMRI Data

Nik Kirilov Kasabov [1,2,3,*], Alexander Yang [4], Iman Abouhassan [3,5], Calum Chace [6], Assia Nikolova Kassabova [3] and Teodoros Lappas [7]

[1] School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology (AUT), WZ Building, St. Paul Street, Auckland 1010, New Zealand
[2] Institute for Information and Communication Technologies (IICT), Bulgarian Academy of Sciences, Acad. G. Bonchev St., Block 2, 1113, Sofia, Bulgaria
[3] Knowledge Engineering Consulting Ltd
[4] Mana Bridge Ltd, New Zealand
[5] Technical University of Sofia, 8 St. Kliment Ohridski Blvd, 1000 Sofia, Bulgaria
[6] Consicum Ltd, UK
[7] Athens University of Economics and Business, 47A Evelpidon Str. & 33 Lefkados Str., Athens 11362, Greece
* Correspondence: nkasabov@aut.ac.nz

Abstract

Can artificial systems form internal representations that resemble conscious perception? Here we introduce **eXCube1**, a brain-inspired spiking neural network (BI-SNN) framework that learns *evolving spatio-temporal associative memories* (ESTAMs) from fMRI data. ESTAMs provide an interpretable, causal account of how neural activity propagates across space and time, bridging statistical neuroimaging analysis and mechanistic modelling. We show that eXCube1 learns discriminative ESTAMs that separate meaningful from meaningless visual and auditory stimuli without access to semantic content. Across two fMRI case studies, the learned models achieve high classification accuracy while revealing structured, modality-specific spatio-temporal dynamics consistent with known neurobiological pathways. Notably, ESTAMs can be robustly recalled from partial temporal input, demonstrating emergent associative memory properties. By explicitly modelling directed spike-based dynamics, eXCube1 moves beyond correlation-based analysis toward causal, interpretable representations of neural computation. These results position ESTAMs as a candidate computational substrate for aspects of machine consciousness grounded in spatio-temporal neural dynamics.

Keywords: spiking neural networks; NeuCube; evolving spatio-temporal associative memory; music; video data

1. Introduction

The rapid advancement of artificial intelligence (AI) has revolutionised technology; however, the question of whether AI can develop machine consciousness remains open and highly debated [1,2]. Neuroscientific research suggests that conscious perception emerges from dynamic neural activity and associative memory structures in biological brains. [3–5]. While some studies emphasise the inherently biological nature of consciousness, computational functionalism argues that if the functional properties of consciousness can be realised computationally, then machine consciousness may emerge in artificial systems.

Current AI models, including large language and vision models, demonstrate remarkable reasoning capabilities but are not brain-inspired and lack intrinsic conscious properties. Their

 doi:10.20944/preprints202604.0605.v1

reasoning is largely based on static associations learned from data, without adaptive perception or internally grounded awareness.

Tononi's Integrated Information Theory (IIT) [6] defines five postulates for consciousness: intrinsicality, information, integration, exclusion, and composition. Causality plays a central integrative role within this framework.

Our central hypothesis is that conscious perception and decision-making can be represented within a brain-inspired computational architecture through evolving spatio-temporal associative memories (ESTAMs) [7–10]. The ESTAM framework complies with the IIT postulates while introducing time as a fundamental dimension, enabling machine consciousness to be defined as an emergent property of past, present, and anticipated neural dynamics.

Distinguishing meaningful from meaningless stimuli at an abstract level underlies intelligent behaviour. In humans, even brief meaningful signals can trigger complex responses, while irrelevant data can be ignored [11–13] . Replicating this capability in AI systems is essential for operation in real-world environments.

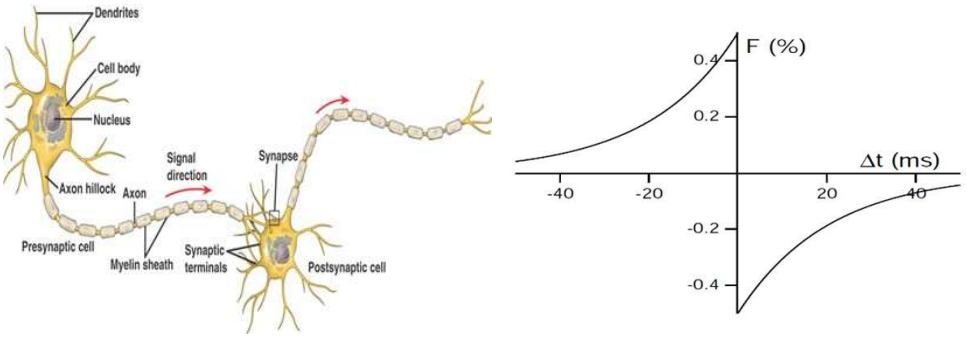
2. The Proposed eXCube1 SNN Framework for Learning Explainable, Evolving Spatio-Temporal Associative Memories (ESTAM) from fMRI Data

2.1. Spiking Neural Networks

Brain processes unfold through cascades of neural activity propagating across distributed networks. Spiking neural networks (SNNs) provide a biologically grounded computational framework by explicitly modelling discrete spike events and their timing [14–16].

Spike-timing-dependent plasticity (STDP) enables synaptic adaptation based on temporal relationships between neurons. The NeuCube framework extends this paradigm by organising neurons according to neuroanatomical templates and learning spatio-temporal patterns from data.

Spiking neural networks (SNNs) offer a biologically grounded approach to modelling such dynamics [15–18]. Unlike conventional artificial neural networks that use continuous activation values, SNNs explicitly represent discrete spike events and their precise timing. This enables spatio-temporal dynamics to be captured, which is essential for modelling C. A popular learning method is spike timing-dependent plasticity (STDP), a learning mechanism where synaptic connections strengthen or weaken based on the relative timing of pre- and post-synaptic spikes (Figure 1).



(**a**) (**b**)

Figure 1. a,b. Connection weights between 2 spiking neurons (Figure 1a) [18] are defined by using the STDP learning rule (Figure 1b) [19], which shows the change in the connection weights based on the difference between the pre-synaptic neuronal spike and the post-synaptic neuronal spike.

The NeuCube framework [17,18,20] pioneered the application of brain-inspired SNNs (BI-SNN) to spatiotemporal brain data by organising reservoir neurons according to neuroanatomical templates and using STDP for unsupervised learning.

When such BI-SNNs are trained on data from different stimulus conditions, the patterns that distinguish those conditions become encoded in the network's evolved connectivity and activity

 doi:10.20944/preprints202604.0605.v1

dynamics. Rather than requiring external statistical tests to compare conditions, the learned differences are embedded within the machine itself. This represents a fundamentally different approach: a BI-SNN model could potentially discover and internalise what differentiates different categories of stimuli. So, can BI-SNNs be used to mimic the way the human brain forms consciously abstract concepts as evolving spatio-temporal associative memory (ESTAM)? The paper addresses this challenge.

2.2. Evolving Spatio-Temporal Associative Memories (ESTAMs)

The concept of ESTAMs was first presented and illustrated by Kasabov in [7] and the utilised in [8–10]. ESTAMs represent how information propagates through neural populations via directed spiking dynamics. They integrate structural connectivity and temporal activity into interpretable representations of causal information flow.

An ESTAM captures:
coordinated activity across neural clusters,
directed propagation of information,
discriminative patterns between stimulus classes,
shared and subject-specific representations.

Unlike static connectivity analyses, ESTAMs explicitly model temporal evolution and causality.

An ESTAM reflects the causal pathways that have been strengthened through STDP as the network learned to process specific types of inputs. Unlike static connectivity matrices or correlation maps, ESTAMs explicitly represent temporal evolution and directedness, showing not just which regions interact, but how activity flows from sender to receiver populations over time. Importantly, ESTAMs reveal how the brain or a machine has organised information processing in response to training data, providing interpretable insight into what has been learned. This distinction is important (see Figure 2 for examples).

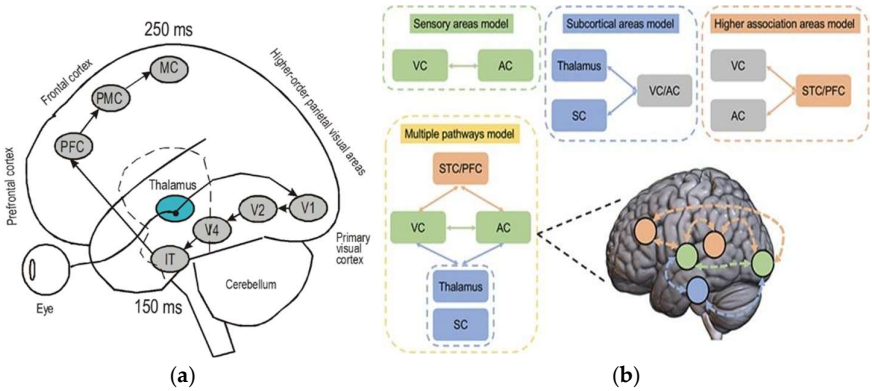


Figure 2. Hypothetical ESTAMs revealing schematically how the brain processes consciously: (a) visual stimuli (from [18]) and (b) audiovisual stimuli (from [21]).

2.3. The NeuCube Brain-Inspired Computational Architecture

We talk about BI-SNN architectures if they use neuroanatomical coordinates for organising neurons, so that evolved patterns of ESTAM in a BI-SNN can reflect such patterns in the human brain.

NeuCube is a brain-inspired computational architecture designed to learn and model spatiotemporal data using spiking neural networks (SNNs) [17,20]. Input data are first encoded into spike trains and projected onto a three-dimensional reservoir structured SNNcube, according to a general or personalised brain template (e.g., MNI space). Within this reservoir, neurons are connected, and their synaptic strengths evolve through spike-timing-dependent plasticity (STDP) [19], allowing the network to capture both spatial structure and temporal dynamics.

The NeuCube architecture integrates four main functional modules (Figure 3):

 doi:10.20944/preprints202604.0605.v1

1. Encoding spatiotemporal input data into spike sequences.
2. Unsupervised reservoir training, where neural connections self-organise in response to input dynamics.
3. Pattern recognition, using classifiers such as dynamic evolving SNNs (deSNNs) [22] to identify learned activity patterns.
4. Visualisation and interpretation of neural dynamics and information flow.

NeuCube has been successfully applied in domains ranging from neuroimaging and brain–computer interfaces to complex time-series modelling and multimodal data integration, providing a biologically grounded platform for both predictive modelling and exploratory neuroscience research [18]. Optionally, NeuCube can include gene regulatory networks as gene parameters. Different types of spiking neuronal models can be used in addition to the integrate-and-fire neurons, and the framework has been implemented on neuromorphic hardware such as SpiNNaker [23,24].

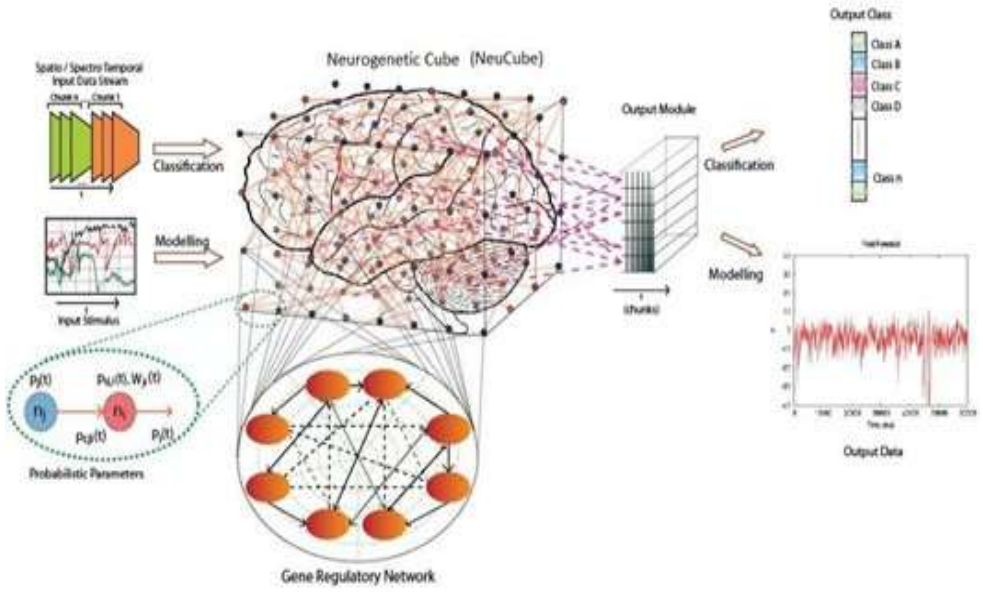


Figure 3. A NeuCube model encodes input data into spike sequences, trains a 3D SNNcube in an unsupervised mode using STDP and learns in a supervised mode the learned patterns in the SNNcube for a classification or regression (from [17]).

2.4. The eXCube1 Framework and Algorithms for Learning ESTAMs from fMRI Data

The proposed eXCube1 framework learns ESTAMs from spatio-temporal data. First, a feature interaction network (FIN) is constructed, which represents the exchange of spikes (information) between clusters of neurons centred by the input neurons (input variables) during learning from data or recall of new data. To construct a FIN, we first match each input to its nearest reservoir neuron from the SNNcube via a k-d tree, defining input seeds. Given spike activity and reservoir weights W, we compute an effective weight matrix W_{eff} by scaling W with per-neuron spike counts. We form a symmetrised affinity matrix and propagate labels from the input seeds using a normalised similarity matrix S to assign each neuron to an input-anchored cluster. The FIN is then the cluster-level adjacency matrix whose entries are the sum of the symmetrised affinity matrix across all neuron pairs spanning two clusters. This aggregates reservoir-level connectivity into interpretable ROI-to-ROI functional interactions.

The process of mapping 24 voxel variables from fMRI data into SNNcube, its training with STDP and the extraction of a FIN is shown in Figure 4 (a,b,c). The thicker the lines that connect two input neuronal clusters, the more these clusters are causally associated in time by exchanging information (spikes).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 April 2026 doi:10.20944/preprints202604.0605.v1

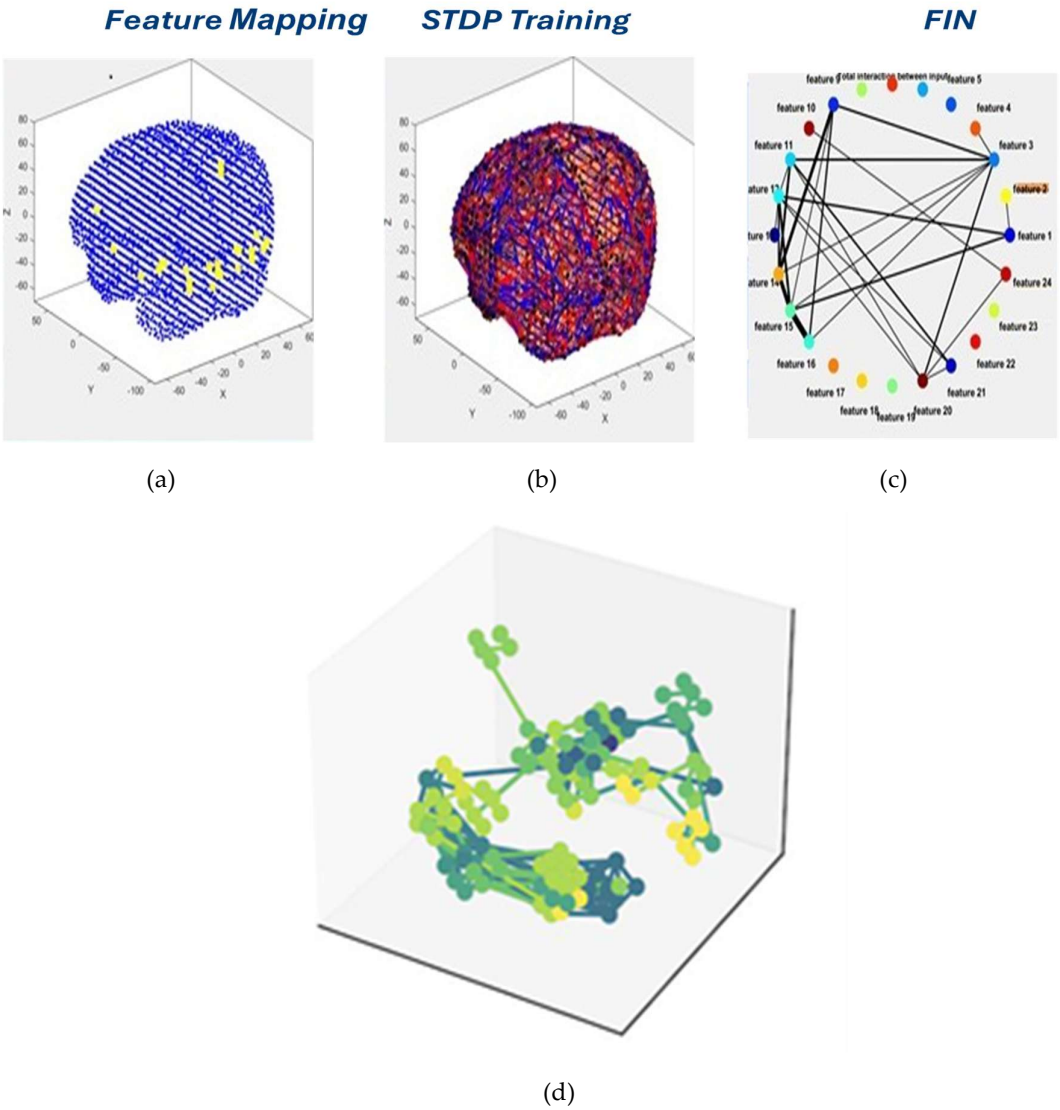


Figure 4. (a,b,c,d). The process of ESTAM extraction after training an eXCube1 model on 24 voxel variables of fMRI data for case study 1 in the paper, described below; the thicker a line is, the more causal associations are; (d) a 3D directed acyclic graph (DAG); darker colours mean spikes earlier in time, lighter colours represent spikes later in time.

While a FIN captures the interaction evolved within the neuron reservoir, 3D spatio-temporal patterns can be extracted from spike pathways formed by chains of reservoir neuron firings across time. Interactions between neuronal clusters within the SNN can serve as one way to distinguish different stimuli in a machine brain over time, which is manifested in an ESTAM (Figure 4d). In addition to FINs, synthetic 3D trajectories of ESTAMs can be extracted, e.g. of meaningful vs meaningless information perception, and compared.

Similar to how the effective weight matrix W_{eff} was made for FIN analysis, for each sample n, spike counts are computed as:

$$c_n(v) = \sum_{t=1}^{T} s_n(t, v), \quad (1)$$

where, $s_n(t, v) \in \{0,1\}$ indicates whether neuron v spiked at time t. These counts were used to scale the latent weight matrix $W \in \mathbb{R}^{v \times v}$, yielding an effective connectivity matrix weighted by presynaptic activity:

 doi:10.20944/preprints202604.0605.v1

$$W_n^{eff}(u,v) = c_n(u) \,.\, W(u,v), \tag{2}$$

where $c_n(u)$ is the total spike count of presynaptic neuron u during sample n.

This weighting captures the principle that synaptic influence scales with presynaptic firing frequency: a connection from a highly active neuron has greater effective strength than the same connection from a silent neuron.

Time-resolved directed acyclic graphs (DAGs) are then constructed by introducing edges. For each individual sample n, we construct a DAG that captures the causal spiking pathways through the reservoir. Nodes in the DAG represent events, denoted as (u, t), where $\mathbf{u}$ is the neuron index and $\mathbf{t}$ is the discrete time step when the neuron is activated.

A directed edge is created from spike event (u, t_u) to spike event (v, t_v) if two conditions are satisfied:

1. Temporal precedence: $t_u \geq t_v$ (the postsynaptic spike occurs at or after the presynaptic spike)
2. Structural connectivity: The learned synaptic weight $W_n^{\text{eff}}[u,v] > 0$, indicating that neuron $\mathbf{u}$ has a non-zero connection to neuron $\boldsymbol{v}$ in the trained reservoir. Formally, the edge set for sample n is defined as:

$$E_n \left\{ (u \longrightarrow v, t) \mid s_n(t,u) = 1, \left| W_n^{\text{eff}}(u,v) \right| \geq 0 \right\} \tag{3}$$

where there is a threshold (typically set to 0 to include all non-zero connections). Each edge carries attributes including connection weight $W_n^{\text{eff}}[u,v]$ *and* spike time t.

This construction reflects the principle that *causal influence* can only flow forward in time through structurally connected pathways.

The resulting per-sample ESTAM, G_n is a sparse directed graph where edge density reflects the interaction between network connectivity structure and the specific temporal pattern of spiking activity elicited by that sample (Figure 4d).

This DAG spatio-temporal dynamic graph-based representation (Figure 4d) departs significantly from the traditional time series clustering techniques [25] and also from the topological clustering as in SOM [26]. An ESTAM can also be represented as a set of spatio-temporal rules. The evolution of these rules can be traced and reported during incremental transfer learning of stimuli, representing new classes [27–29]. The proposed here method is a new direction in the field of fuzzy clustering [30,31], as it introduces for the first time evolving spatio-temporal fuzzy clustering of brain neuroimaging data.

To compare ESTAMS of two classes means to compare their spatio-temporal DAG graphs. Here we define a time-shifted distance metric:

$$D_\tau(G_i, G_j) = \frac{\left| E_i \cap \text{shift}(E_j, \tau) \right|}{\left| E_i \cup \text{shift}(E_j, \tau) \right|}, \tag{4}$$

where τ accounts for temporal misalignment.

The discriminative ESTAM is selected from class 1 candidates by maximising the separation margin with class 2 data.

$$M(G) = \frac{1}{|C_2|} \sum_{j \in C_2} \min_\tau D_\tau(G, G_j) - \frac{1}{|C_1|} \sum_{j \in C_1} \min_\tau D_\tau(G, G_i), \tag{5}$$

The optimal discriminative ESTAM, G^*, is thus the graph that minimises within-class variability while maximising between-class separability. This principle extends the fuzzy C-means clustering method [30].

Box 1 presents a meta algorithm for training an eXCube1 system and for extraction of an ESTAM.

Box 1. The meta-algorithm for the creation of an ESTAM in an eXCube1 model

 doi:10.20944/preprints202604.0605.v1

1. Encode input temporal or spatio-temporal data into spike sequences: continuous value input data is encoded into trains of spikes.
2. Create a 3D SNN based on a brain template (fMRI; MNI; Talairch, etc.[18])
3. Map the input variables/features into a 3D SNN.
4. Train in an unsupervised mode using STDP the 3D SNN, to learn spatio-temporal associations between activated neurons from the spike encoded input data.
5. Extract a FIN from a trained SNN representing spatio-temporal association between defined input clusters
6. Extract a 3D spatio-temporal pattern to represent the internal ESTAM model.
7. Continue training the model on new data and use it to recall previously learned spatio-temporal patterns.

An ESTAM has the property for learned spatio-temporal patterns to be recalled successfully if only part of the used for training spatial or temporal or spatio-temporal input data is presented [7–10], illustrated in section 3 of the paper. This is a result of the synfire process in the SNNcube, so that a learned ESTAM pattern can be fired by partial input data.

3. Two Case Studies for Modelling Conscious Perception of Meaningful vs Meaningless Visual and Auditory Stimuli

3.1. Rationale for the Case Studies

We demonstrate how ESTAMs are created with the use of the eXCube1 BI-SNN on two fMRI datasets examining meaningful versus meaningless perception from visual [11] and from auditory [12] stimuli.

In both experiments, an eXCube1 model has learned from fMRI data to distinguish meaningful from meaningless stimuli regardless of the concrete content of the video or music data. Analysis of the extracted ESTAMs revealed discriminative features that separate the two classes through evolved network dynamics. The ESTAM patterns are justified through neuroscience analysis and classification results.

These results demonstrate that brain-inspired architectures can autonomously discover interpretable spatiotemporal patterns that differentiate stimulus categories, encoding the distinction within learned connectivity rather than requiring external statistical comparison.

Some experimental results are described in the two sub-sections below.

3.2. An eXCube1 Model That Learns from fMRI Data Human Perception of Meaningful Versus Meaningless Video Stimuli

An eXCube1 model trained on fMRI data successfully distinguished meaningful video stimuli from meaningless noise. The model achieved high classification accuracy and revealed interpretable spatio-temporal connectivity patterns consistent with neuroscience findings. Importantly, the model maintained strong performance even when recalled using partial temporal data, demonstrating robust associative memory properties.

We have created an eXCube1 model from fMRI data of participants viewing meaningful (video) versus meaningless (static; noise) stimuli to evolve condition-specific ESTAMs. The data used here are from Boly et al. [11]. Preprocessed fMRI data included 24 (first experiment) and 47 (second experiment) voxel locations representing cortical areas (precentral gyrus, insula, postcentral gyrus, fusiform gyrus, middle temporal gyrus, cuneus, middle occipital gyrus, lingual gyrus) and subcortical areas (caudate, putamen, globus pallidus, nucleus accumbens, hippocampus, amygdala, thalamus) [11] (see Appendix A). The dataset comprised 160 samples (80 per class) from 6 subjects, with 16 time points per sample. More information about data preparation is in Appendix B. A BI-SNN model with 5,029 spiking neurons was created in a 3D structure with neurons spatially positioned according to the down sampled MNI brain template [11].

 doi:10.20944/preprints202604.0605.v1

The goal of this experiment is using fMRI data and the eXCube1 BI-SNN to identify spatio-temporal brain signatures in time and space, i.e. ESTAMS, that distinguish between:

- Class 1: Watching a meaningful video
- Class 2: Watching meaningless TV static noise/blank screen

Figures 5a,b show 24 fMRI normalised voxel values over 20 seconds when a person is watching meaningful or meaningless videos.

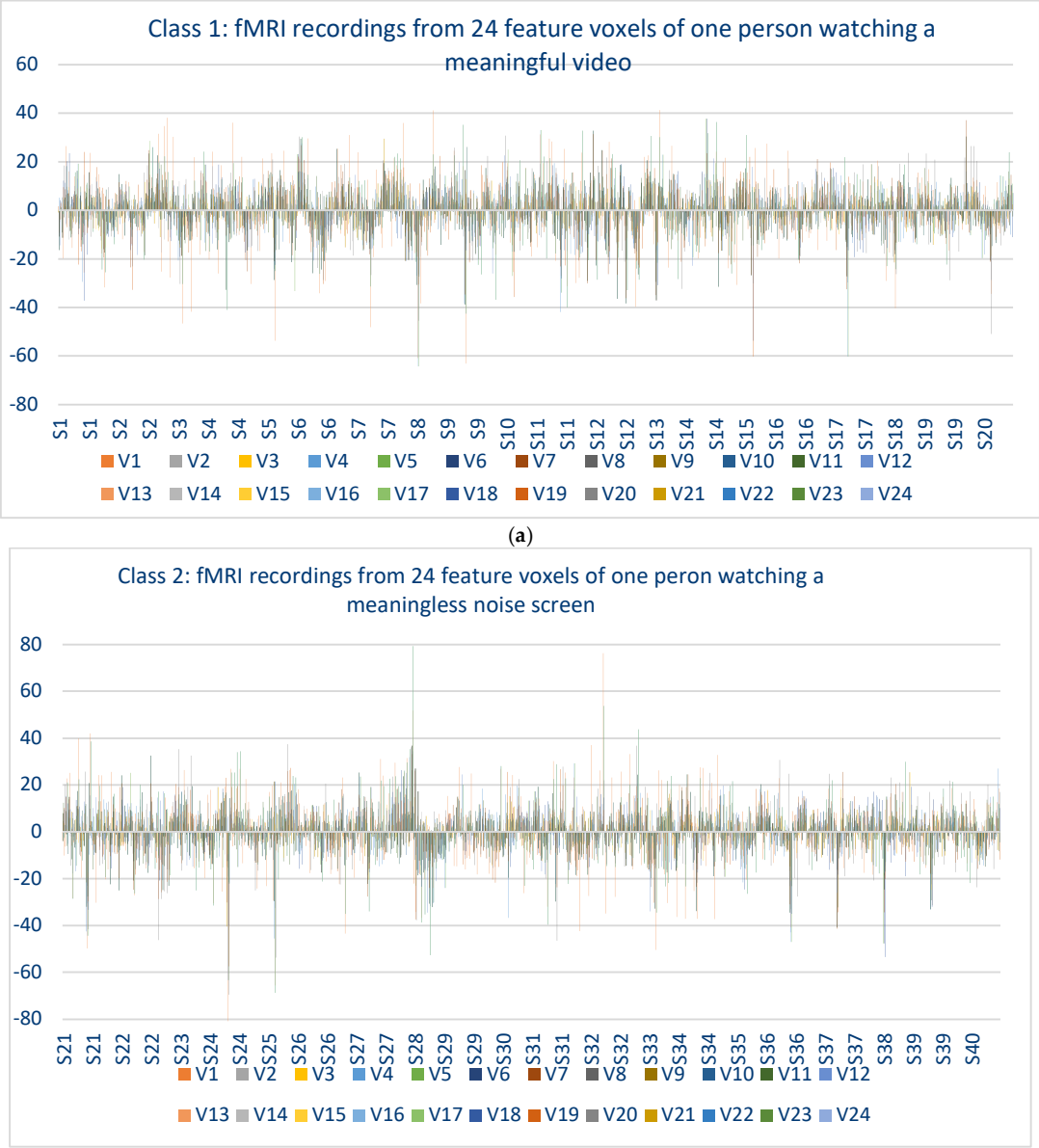


Figure 5. 24 fMRI normalised voxel values over 20 seconds when a person is watching meaningful (a) or meaningless videos (b).

Here the algorithm from Box 1 is applied to create an eXCube1 model. Figure 6a shows the FIN and the DAG representation of ESTAMs of one person perceiving meaningful vs meaningless video stimuli extracted after training of the model.

 doi:10.20944/preprints202604.0605.v1

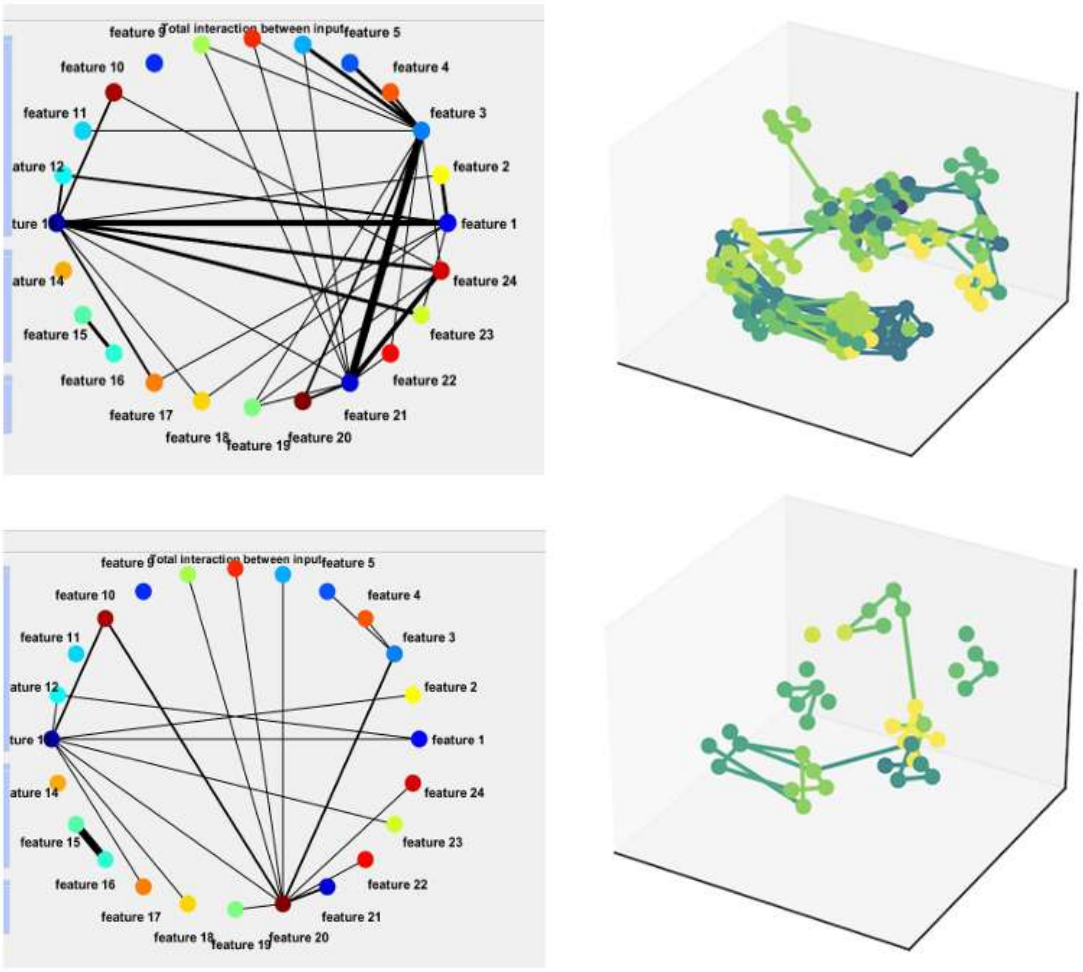


Figure 6. a. FIN and DAG representations of ESTAMs of one person perceiving meaningful (top raw) vs meaningless (bottom raw) video stimuli, along with the used 24 voxel features. The first ESTAM represents strong spatio-temporal interactions between 1, 24 and 23 with 13 and also 3 with to 21, while the second one does not show prominent interactions (for explanation, see the text below). The features from 1 to 24 represent the brain areas as follows: 1: right middle occipital gyrus;**7**: left cuneus; **17**: right cuneus; **16**: right postcentral gyrus; **13**: right fusiform gyrus; **14**: left insula; **15**: right postcentral gyrus; **16**: right postcentral gyrus; **12**: right fusiform gyrus; **9:** left precentral gyrus; **3**: left middle temporal gyrus; **15**: right postcentral gyrus; **14**: left insula; **12**: right fusiform gyrus; **9:** left precentral gyrus; **3**: left middle temporal gyrus; **23**: right cuneus; **21:** left middle occipital gyrus; **24:** left lingual gyrus.

After training the model on fMRI data, the model is recalled on the same data to extract one feature vector of 24 features for each training sample. Ranking the features (using SNR (signal-noise ratio), or T-test method) shows the importance of these features (voxels) to discriminate the two classes – Figure 6b.

 doi:10.20944/preprints202604.0605.v1

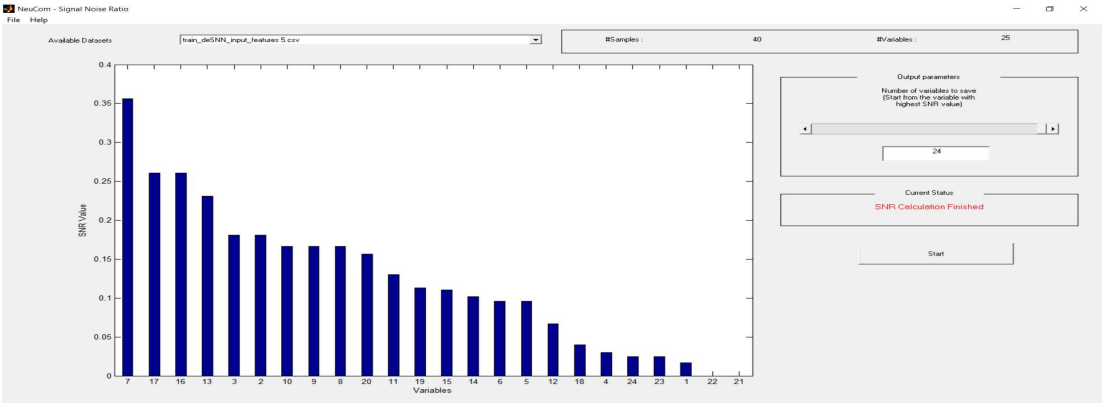


Figure 6. b Using SNR method to rank the features/voxels in terms of discriminating the two classes. Dominant features are: 7 (left cuneus), 17 (right cuneus), 16 (right postcentral gyrus), 13 (right fusiform gyrus), 3 (left middle temporal gyrus), 9 (left precentral gyrus).

These vectors are then classified in a classifier and tested in a cross-validation mode to confirm that, once a model is trained on two class fMRI data for one person, the model can recognise if the person is seeing meaningful or a meaningless new videos. Using an ECF classier from the NeuCom software (https://theneucom.com) the accuracy was 97% - Figure 6c.

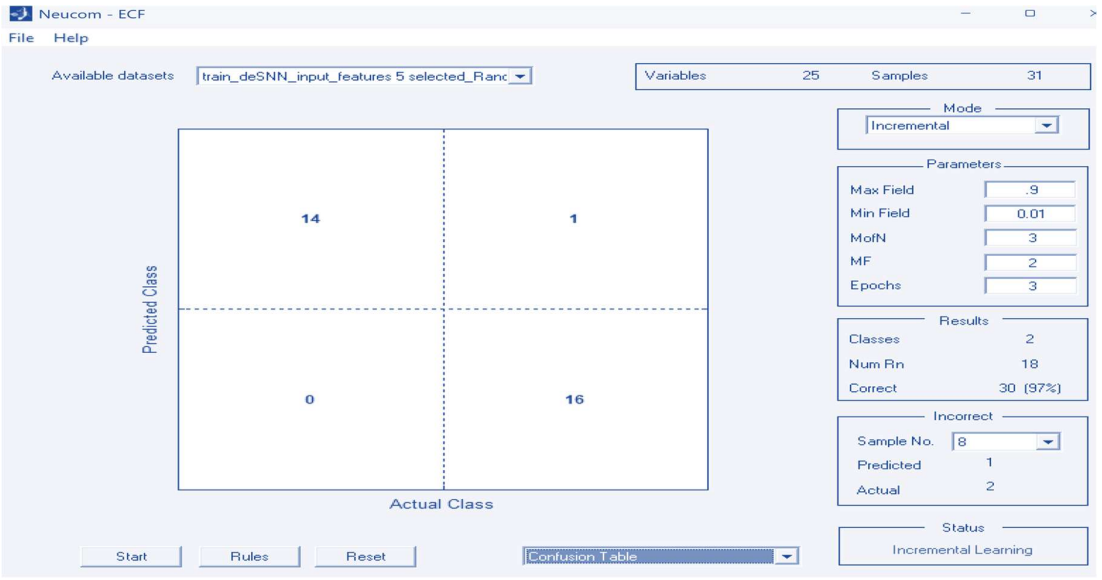


Figure 6. c. Classification accuracy, when an ECF classification method, in its incremental learning mode, is used to classify the 2-class feature vectors. The achieved 97% accuracy shows that the model has learned well to discriminate the person's perception of meaningful vs meaningless videos on new data. A 2D plot of the top 2 features as the 2 dimensions showing the centres of 18 cluster centres of the data from Figure 6c is shown in Figure 6d.

 doi:10.20944/preprints202604.0605.v1

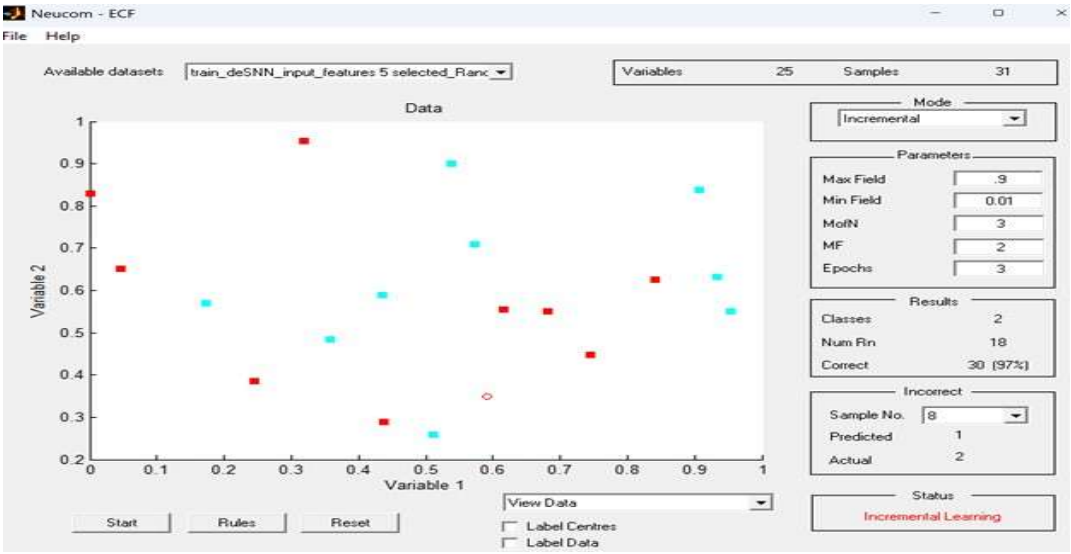


Figure 6. d. A 2D plot of the top 2 features as the 2 dimensions showing the centres of 18 cluster centres of the data from Figure 6c.

An ESTAM has the property for learned spatio-temporal patterns to be recalled successfully if only part of the used for training spatial or temporal or spatio-temporal input data is presented [7–10]. This is illustrated here on an eXCube1 model trained on two class data of one person (20 samples for each class, 16 time points each) and recalled on only 8 time points of the used data, which is 50% of the training data. This is a results of the synfire process in the SNNcube, so that a learned ESTAM pattern can be fired by partial input data.

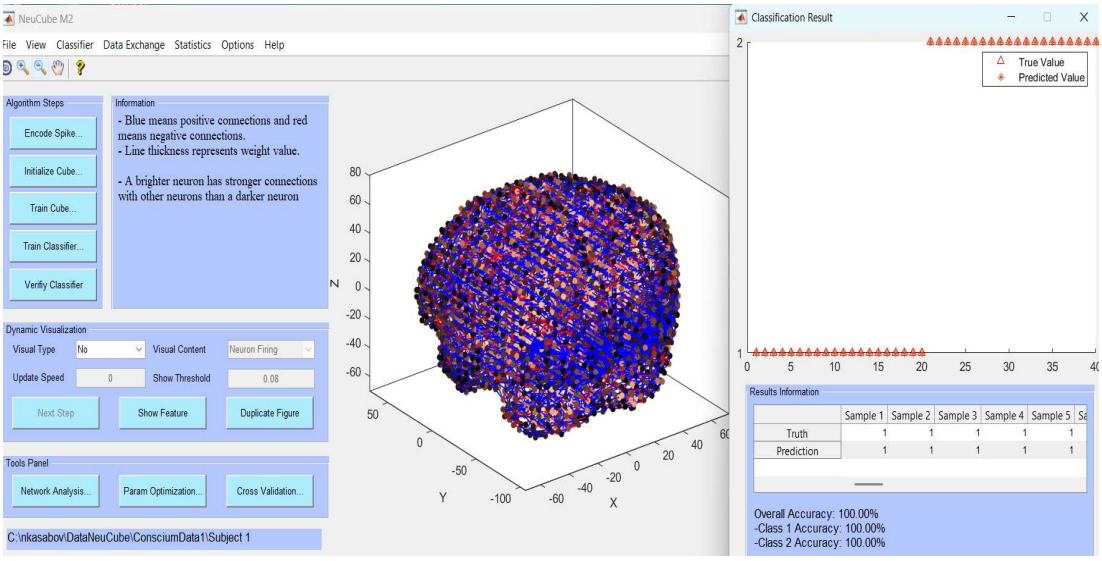


Figure 6. e. Testing results when an eXCube1 model is trained on two class data of one person and recalled on only 8 time points of the used data (50%) of the full time of 16 time points data used for training the model. When only 4 time points of data (20%) instead of 16, were used for a recall of the trained model on 16 time points data, the classification accuracy dropped to 87.5% (80% and 95% for class 1 and class 2 correspondingly). The results when recalling the model on a shorter time than 4 time points were not acceptable. .

The experiment above on one person data demonstrated that an eXCube1 model can learn ESTAMs of meaningful and meaningless stimuli perceptions of one person. Would an eXCUbe1

 doi:10.20944/preprints202604.0605.v1

model be able to learn ESTAMs of multiple subject as individual signatures? Figure 7a compares one person FIN, learned in a single person eXCube1 model, with the FIN of multiple subjects learned in a multiple subject eXCube1 model.

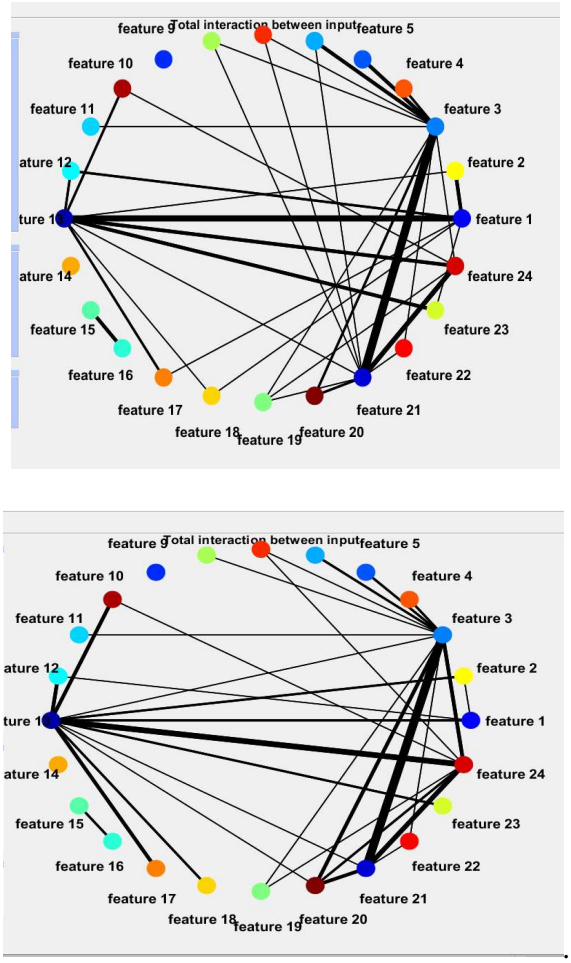


Figure 7. a. Common patterns of perception of meaningful visual stimuli can be identified across subjects: (top) one subject; (bottom) 4 subjects. Still, subjects have their subjective perception also captured in the FIN diagrams. The features from 1 to 24 represent the brain areas as follows: 1: right middle occipital gyrus;**7**: left cuneus; **17**: right cuneus; **16**: right postcentral gyrus; **13**: right fusiform gyrus; **14**: left insula; **15**: right postcentral gyrus; **16**: right postcentral gyrus; **12**: right fusiform gyrus; **9:** left precentral gyrus; **3**: left middle temporal gyrus; **15**: right postcentral gyrus; **14**: left insula; **12**: right fusiform gyrus; **9:** left precentral gyrus; **3**: left middle temporal gyrus; **23**: right cuneus; **21:** left middle occipital gyrus; **24:** left lingual gyrus.

When a classifier is applied to classify the ESTAMS extracted from the fMRI data in one eXCube1 model trained on multiple subjects data, to test if the model can distinguish persons from each other based on their extracted ESTAMS, a good classification was achieved even if only part of the temporal samples, as per the experiment in Figure 6e, were used for recall. Using only 50% of the temporal data for recall, an overall accuracy of recognising each of the 4 persons was 100% (Figure 7b). This means that an eXCube1 model can learn specific/subjective individual patterns of multiple subjects perceiving meaningful and meaningless visual stimuli as unique, subjective experience for each of the subjects.

 doi:10.20944/preprints202604.0605.v1

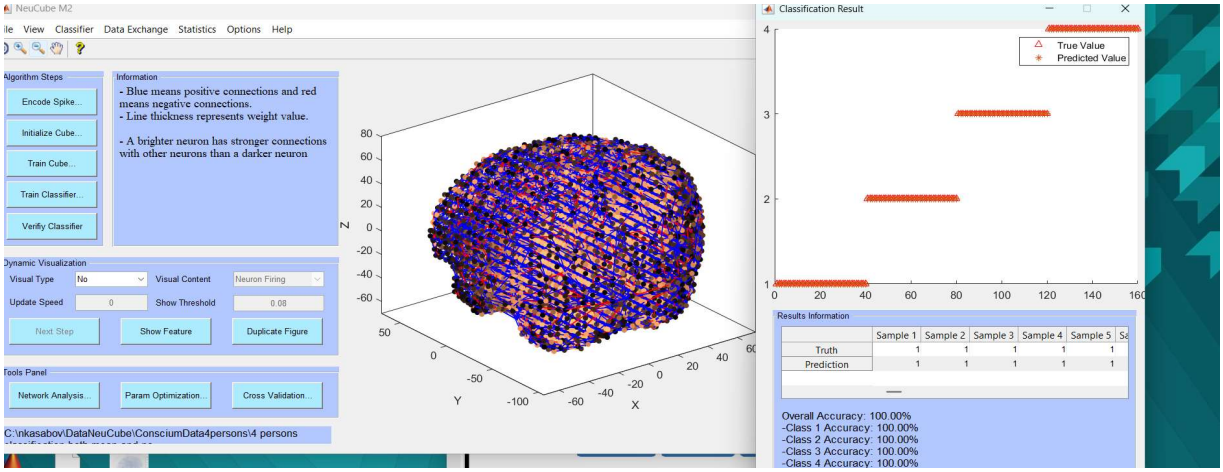


Figure 6. d. Using 50% and even 20% of the temporal data for recall, an overall accuracy of recognising each of the 4 persons in the model from Figure 7a was 100%. Further reducing the time steps to 10% (using only 2 initial time steps from all 16 time steps in the data) the recognition rate dropped dramatically to 23%, which was expected. This confirmed the strong performance of the model as spatio-temporal associative memory when learning to discriminate individual ESTAMS. Each subject was presented by 20 fMRI samples of each class. .

The experiments above were done on 24 feature/voxel variables. More experimental results on 47 features/voxels are presented in Appendix C.

The analysis of the above experiments revealed distinct connectivity signatures distinguishing meaningful from scrambled movies. Meaningful movies showed stronger connectivity patterns between motion-sensitive extrastriate regions (middle occipital gyrus, MOG) and biological motion perception regions (middle temporal gyrus, MTG), anatomical pathways known from prior work to support biological motion perception [32,33], ventral stream circuits previously associated with scene processing, object recognition, and memory encoding [34,35], and limbic-striatal networks involving nucleus accumbens (NAc), amygdala (AMY), and globus pallidus (GP) known to mediate reward and emotional salience in biological systems [36,37]. The model's learned connectivity patterns preferentially engaged these anatomical circuits for coherent narrative content containing recognisable actions and emotionally salient events. Conversely, TV noise (meaningless) showed markedly stronger bilateral early visual connectivity, a pattern consistent with compensatory low-level processing observed when top-down semantic predictions fail in biological systems [38,39].

Temporal analysis of input activation from the ESTAMs revealed hierarchical progression in which meaningful movies elicited early motion area and limbic engagement, followed by sustained higher-order processing in insula and recurrent thalamic feedback loops. This temporal sequence of rapid salience detection followed by prolonged emotional and semantic integration aligns with hierarchical predictive coding models [39–41].

3.3. An eXCube1 Model That Learns from fMRI Data Human Perception of Meaningful Versus Meaningless Audio Stimuli

A second study demonstrated that the framework can distinguish meaningful music from resting-state or noise conditions. The resulting ESTAMs revealed modality-specific processing differences, with auditory perception exhibiting more spatially localised causal dynamics.

Data from [12] is used in this experiment. The goal of this experiment is to create ESTAMS from fMRI data of brain patterns when a person is listening to music as a meaningful audio stimuli versus when they are at rest/exposed to noise. The fMRI data samples include 10 genres, each repeated twice: blues, classical music; country music; disco; hip-hop; jazz; metal; pop; reggae; rock. Eighty records of fMRI data, one each of 80 subjects, are recorded when people are at rest or exposed to noise. Similar

 doi:10.20944/preprints202604.0605.v1

to the previous experiments, a BI-SNN model was created to discover the ESTAMS of each of the two classes for a single person and for multiple persons as shown in Figure 8 (a,b,c).

The subtraction of FIN matrices (Figure 8c) revealed even more substantial differences in inter-cluster interaction strength between meaningful and meaningless conditions. Interestingly, the meaningless resting state condition showed more extensive inter-cluster connectivity across many regions compared to the meaningful music condition. However, the discriminative ESTAM shows that causal spike propagation for meaningful auditory perception remains spatially constrained to localised regions, contrasting sharply with the distributed visual ESTAM. This pattern suggests that auditory meaningful information is processed in the model through local sequential computation within a confined spatial area, while the widespread inter-cluster interactions observed in the FIN provide modulatory context or reflect broader network state rather than forming the core causal processing pathway.

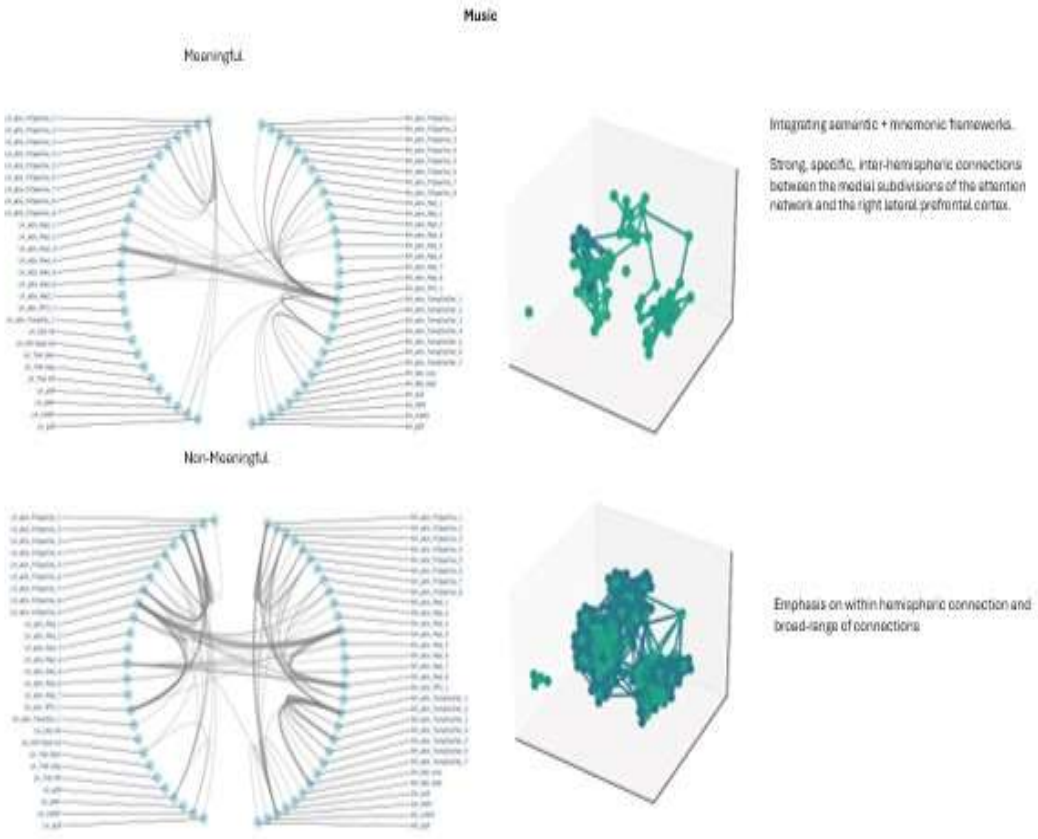


Figure 8. a. ESTAMs of one person listening to 10 genres of music (meaningful stimuli) and one person at rest or listening to noise (not meaningful stimuli). The brain areas are named in Appendix A.

 doi:10.20944/preprints202604.0605.v1

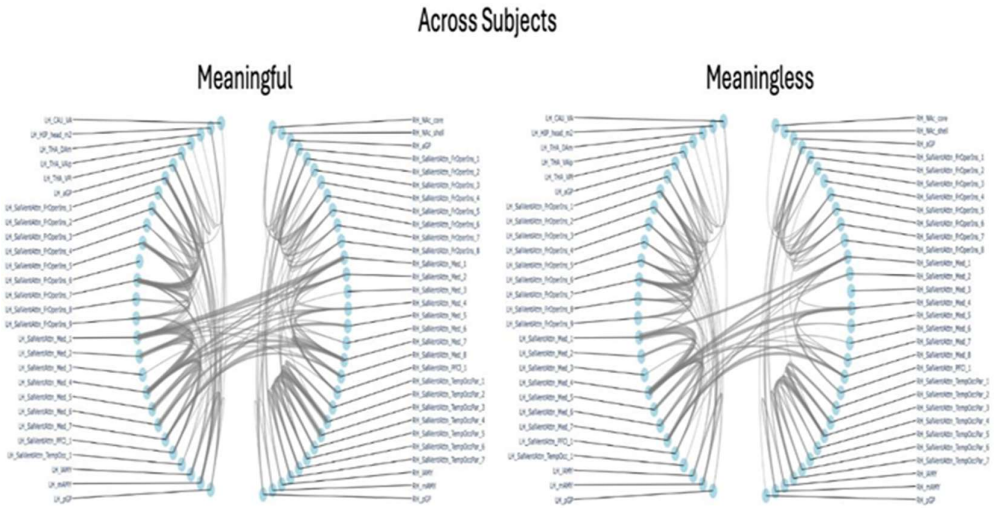


Figure 8. b. ESTAMs of meaningful and meaningless auditory stimuli perception for multiple subjects. The names of the nodes as brain areas are explained in Appendix A.

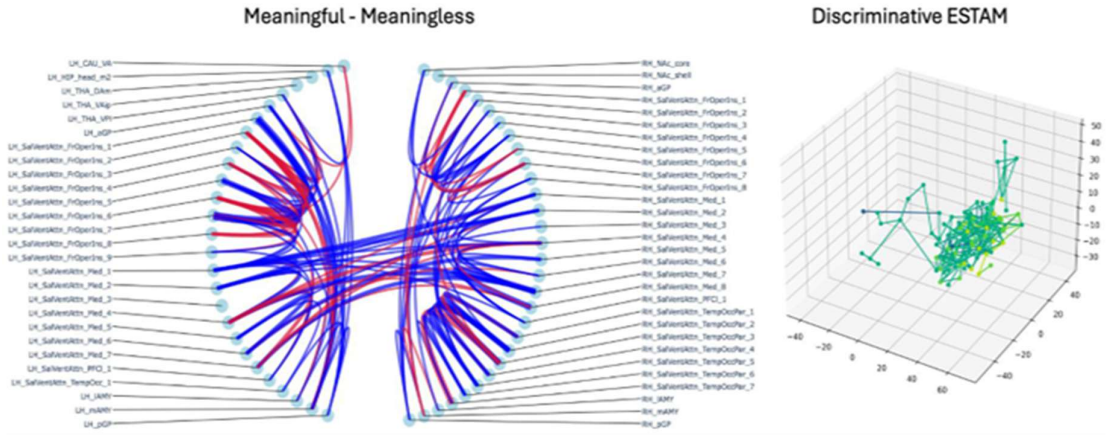


Figure 8. c. The difference between meaningful and meaningless auditory perception across subjects; blue means increase in interaction, red means decrease in interaction. The thicker the line, the more the magnitude of interaction shifts (left). The names of the nodes are explained in Appendix A. (right) ESTAM that discriminates between the two classes of perception. .

FIN analysis revealed connectivity patterns distinguishing music listening from resting state. Music listening showed stronger connectivity in fronto-opercular-insular circuits (inferior frontal gyrus, IFG, and anterior insula), regions previously associated with auditory syntax and working memory in biological systems [42,43], medial salience regions including anterior cingulate cortex (ACC) and medial prefrontal cortex (mPFC) linked to expectation monitoring [44], and limbic-striatal pathways connecting amygdala (AMY) with nucleus accumbens (NAc), circuits known to mediate emotional and reward responses [45,46]. The model's learned patterns preferentially engaged these anatomical circuits for structured musical content compared to passive resting state. Conversely, resting state showed stronger connectivity in temporo-parietal junction (TPJ) regions, a pattern potentially consistent with spontaneous cognition and mind-wandering processes [47].

Temporal analysis of input activation from the ESTAM revealed meaningful music triggered early salience network and limbic activation, followed by sustained fronto-opercular engagement. This temporal pattern of rapid relevance detection preceding prolonged syntactic processing matches auditory hierarchical timescales and predictive coding frameworks where salience signals prime subsequent detailed analysis [48,49].

 doi:10.20944/preprints202604.0605.v1

4. Conclusions, Discussions and Future Work

The eXCube1 framework demonstrates that brain-inspired SNNs can autonomously learn interpretable spatio-temporal representations that distinguish meaningful from meaningless stimuli.

However, limitations remain, including simplified voxel representations and lack of biologically detailed connectivity constraints. While ESTAM provides a novel generic theory for modelling causal spatiotemporal dynamics, several important limitations of the eXCube1 framework remain to be addressed before machine consciousness is achieved, at least to a certain degree. The models receive single-voxel input from each region of interest (ROI), with fMRI voxels representing volumes where spatial smoothing introduces considerable signal blur across neighbouring voxels [50,51]. This single-voxel representation fails to capture the distributed activation patterns characteristic of cortical ROIs as defined by standard parcellation atlases [52–54]. Implementing anatomically faithful propagation dynamics would require incorporating white matter tractography to define connectivity pathways [55,56], axonal conduction velocity parameters, and distance-scaled transmission delays [57,58].

Despite the above limitations, the eXCube1 framework is a contribution to the area of brain-inspired systems for modelling and understanding of brain data and the brain processes that generate them. It delivers specific contributions at the intersection of brain sciences, computational neuroscience and machine learning. First, the spiking dynamics and spatial organisation of our modelling enables transparent analysis of how input features interact across time within the model [16,59]. Unlike black-box deep learning approaches, we can trace how signals from different ROIs temporally coordinate, compete, or integrate through observable spike trains and synaptic weight evolution. This interpretability extends beyond traditional fMRI statistical analyses such as GLM-based activation detection or seed-based connectivity by explicitly modelling emergent spatiotemporal pattern dynamics [60,61]. Second, our work reveals patterns that distinguish meaningful from meaningless information across participants and personalised patterns reflecting subject-specific signatures for each condition. These patterns exist at the level of sensor input dynamics, treating each ROI as an information source entering a processing network. While we use fMRI data from neuroanatomical coordinates, the framework generalises to any spatially distributed, temporally evolving sensor data where understanding cross-sensor interactions and temporal pattern evolution is valuable [62]. Our primary aim is not to reverse-engineer brain function, but to identify reproducible spatiotemporal patterns through ESTAMs. Our SNN architecture should be understood as brain-inspired rather than brain-simulating. It leverages biological principles including spike-based processing, spatial organisation, and temporal plasticity to extract interpretable spatiotemporal patterns from neural data, without claiming to replicate physiological mechanisms. Future work validating these patterns against empirically measured spatiotemporal BOLD propagation and investigating causal relationships through neuromodulation experiments would be necessary to establish neurobiological correspondence [63].

Finally, a distinctive property of the ESTAMs learned by our models is that they naturally embed three temporal dimensions. The past is encoded in synaptic weights and memory traces accumulated from prior activity. The present is reflected in ongoing spike responses to current stimuli. The future can be inferred by extrapolating learned trajectories, allowing prediction of what patterns are likely to emerge next. This temporal triad highlights how ESTAMs align with theories of predictive coding, providing a bridge between memory, emotion, perception, and anticipation. Future work should formalise these temporal roles and test their predictive capacity across domains.

Our results demonstrate distinct spatiotemporal integration patterns between sensory modalities. Visual stimuli elicited reduced inter-cluster connectivity in specific regions alongside distributed causal spike propagation for meaningful compared to meaningless conditions. Auditory processing showed extensive inter-cluster connectivity during the meaningless resting state condition compared to meaningful music listening, while causal spike propagation remained spatially constrained during meaningful perception. These patterns suggest that spike activity within each model contains information about the underlying computational principles distinguishing meaningful from meaningless processing.

 doi:10.20944/preprints202604.0605.v1

Future work should explore more selective criteria for causality, such as prioritising spikes with higher firing frequency, weighting contributions by temporal proximity, or incorporating probabilistic causality models. Timing sequence and feedback interactions carry important information, yet their full representation can be computationally expensive and difficult to visualise. Developing scalable methods that capture the richness of temporal dynamics without overwhelming interpretability remains an open research problem. Hybrid approaches combining graph-theoretic simplifications with circuit-level analytics may provide a path forward.

One of the most promising aspects of ESTAMs is their ability to recall patterns from incomplete inputs, enabling predictive recognition of meaningful stimuli [7–10]. Future studies should systematically quantify this predictive capacity, identifying the minimum subset of data required for accurate recall and exploring its generalisation across tasks and modalities. Such work could link ESTAMs more directly to theories of predictive coding and anticipatory perception [64].

ESTAMs can be extended to model layered analytics of circuit interactions, feedback, timing of pattern emergence, and hierarchical cascades, paralleling techniques in systems neuroscience. This would allow the framework to approximate not only perceptual meaning-making but also higher-order processes such as contextual reasoning, affective modulation, and decision-making. A challenge that can also be addressed is dealing with multimodal neuroimaging data [65].

Future work will focus on developing novel BI-SNN for ESTAM. By leveraging retinotopic [66] and tonotopic [18] encoding methods, such models could identify generic patterns of stimuli perception in machines. This opens the way for brain-inspired machines that not only compute but also interpret the world in ways that may align with principles of adaptive, context-aware, conscious intelligence [5,67–69].

The extracted ESTAMS, being dynamic patterns of interaction between areas of the brain over time, captured major features of consciousness manifested in the brain, such as dynamic spatio-temporal interactions, memory, emotions, goals, abstract information processing, causal learning and reasoning, subjective experience, awareness, and predictive coding to mention only some of them. They also comply with the postulates of the IIT [6] and other theories of consciousness, which makes ESTAMs promising for future research.

Overall, the proposed eXCube1 framework represents a step toward biologically grounded, interpretable AI systems capable of modelling aspects of conscious perception through evolving spatio-temporal associative memories, ESTAMs.

5. Patents

The authors declare that there are no patents associated with this work.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. The NeuCube Python implementation (NeuCubePy) is available online at https://github.com/KEDRI-AUT/NeuCube-Py. The NeuCom software is available from: https://theneucom.com.

Author Contributions: N.K. designed the eXCube1 framework and the NeuCube architecture and wrote a major part of the paper. A.Y. designed the models, implemented them in Python, conducted the experiments, and contributed to the paper preparation. I.A. contributed to model experiments and paper preparation. A.K. contributed to the project discussions and paper preparation. C.C. and T.L. contributed to the problem specification, interpretation of the results and paper preparation. All authors have read and agreed to the published version of the manuscript.

Funding: There was no funding received for the preparation and submission of the paper.

Data Availability Statement: The Python implementation of the eXCube1 framework and the case study models presented in this paper are available upon request, subject to copyright restrictions (contacts: T.L. and A.Y.). NeuCube is available from https://kedri.aut.ac.nz/neucube.

 doi:10.20944/preprints202604.0605.v1

Acknowledgments: The authors would like to thank all collaborators who contributed to discussions and development related to this work. The paper experiments were conducted with the use of the eXCube1 software, developed in KECL and ManaBridge and funded by Conscium Ltd. This software utilised some functions from the NeuCubePy software, developed at KEDRI/AUT NZ by B.Singh in 2023 and funded by the MBIE (Ministry of Business, Innovation and Enterprise) of New Zealand. The authors would like to thank the MDPI Publisher for the strong encouragement to submit this paper.

Conflicts of Interest: The authors declare no conflicts of interests. The part-time association of N.K., I.A. and A.K. to KEC Ltd., and the association of A.Y. with ManaBridge and C.C. and T.L. with Conscium Ltd. did not involve funding for this paper preparation and publication. The authors declare that there are no patents associated with this work.

Appendix A: 47 Anatomical Regions Used in the Experimental Case Study Models

ACC - Anterior Cingulate Cortex, aGP - Anterior Globus Pallidus, AMY - Amygdala,

CAU - Caudate Nucleus, CAU_DA - Caudate Dorsal Anterior, CAU_VA - Caudate Ventral Anterior, FrOperIns - Fronto-Opercular-Insular, GP - Globus Pallidus, HIP - Hippocampus, HIP_head_m1 - Hippocampus Head Medial 1, HIP_head_m2 - Hippocampus Head Medial 2, IFG - Inferior Frontal Gyrus, lAMY - Lateral Amygdala, LH - Left Hemisphere,

mAMY - Medial Amygdala, Med - Medial, MOG - Middle Occipital Gyrus, mPFC - Medial Prefrontal Cortex, MTG - Middle Temporal Gyrus, NAc - Nucleus Accumbens, NAc_core - Nucleus Accumbens Core, NAc_shell - Nucleus Accumbens Shell, pGP - Posterior Globus Pallidus, PFCl - Lateral Prefrontal Cortex, PUT - Putamen, PUT_VA - Putamen Ventral Anterior, RH - Right Hemisphere, SalVentAttn - Salience/Ventral Attention Network,

TempOcc - Temporo-Occipital, TempOccPar - Temporo-Occipital-Parietal, THA - Thalamus, THA_VPl - Thalamus Ventral Posterolateral, TPJ - Temporo-Parietal Junction

Appendix B: Data Description and Data Preparation for the Experimental Studies

B.1. fMRI Preprocessing

All fMRI data were preprocessed in Python using NiBabel, ANTsPy, NumPy, SciPy, and Matplotlib. The pipeline comprised: (i) initial volume removal, (ii) rigid-body motion correction, (iii) spatial normalization to MNI space, (iv) spatial smoothing, (v) temporal high-pass filtering, (vi) voxelwise standardization, and (vii) brain masking. No slice-timing correction was applied.

B.2. Initial Volume Removal

We followed similar preprocessing by discarding the first 35 volumes at the start of each run for visual datasets [8] and music noise datasets [47]. For music datasets we did not remove volumes.

B.3. Motion Correction

Each 4D run was realigned to its first volume using ANTs rigid registration (type_of_transform = “Rigid”). Transforms were estimated per timepoint and applied to produce a motion-corrected time series.

 doi:10.20944/preprints202604.0605.v1

B.4. Spatial Normalization

The mean of the first 10 volumes served as the EPI reference for the visual and music noise dataset and was nonlinearly registered to the MNI152NLin2009cAsym T1 template (ANTs SyN affine was used for fallback on failure). For the music dataset the first volume served as the EPI reference. The resulting transforms were applied to every timepoint, yielding data in template space.

B.5. Spatial Smoothing

We applied Gaussian smoothing with an 8 mm full width at half maximum (FWHM) kernel to each timepoint.

B.6. Temporal Filtering

To reduce low-frequency drifts, we performed a Gaussian-weighted high-pass filter with a 60 s cutoff. The filter used the dataset-specific TR.

B.7. Voxelwise Standardization

For each voxel, we removed the temporal mean and scaled by the temporal standard deviation via z-scoring, producing zero-mean, unit-variance time series.

B.8. Brain Masking

After normalization, we applied an MNI template gray-matter probabilistic mask (tpl-MNI152NLin2009cAsym_res-01_label-GM_probseg.nii.gz); nonzero mask voxels were retained and all others set to zero. Visual QC was performed by comparing axial slices.

Appendix C: Using 47 Input Feature Voxels to Create eXCube1 Models and to Extract ESTAMs from Them

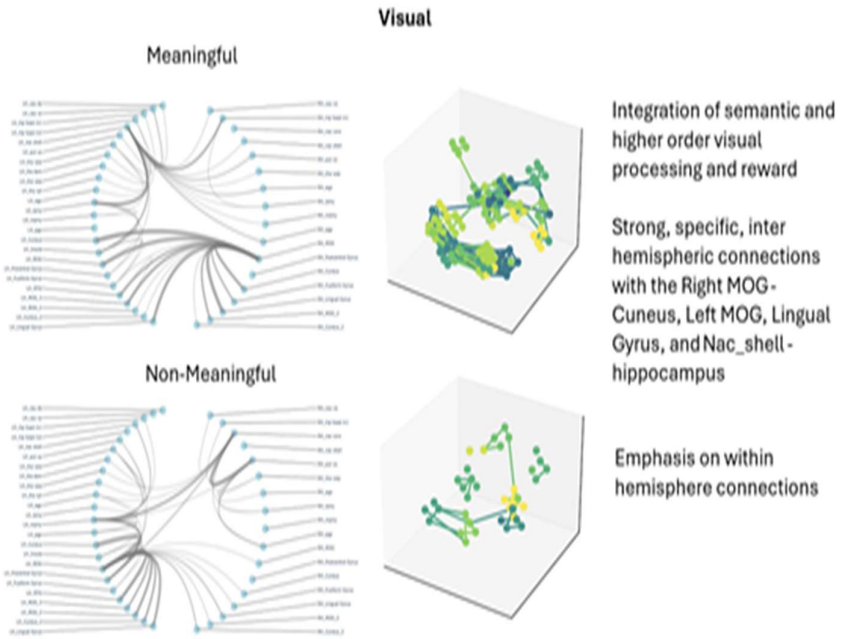


Figure C1. FINs of meaningful and meaningless visual perception for a single subject using 47 features. The names of the nodes are brain areas as explained in Appendix A.

 doi:10.20944/preprints202604.0605.v1

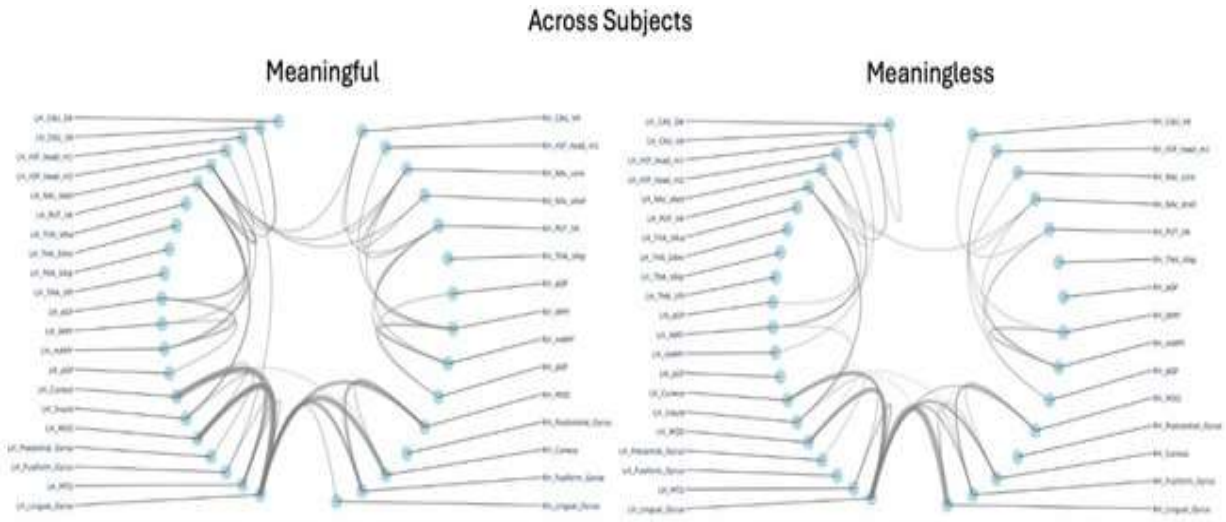


Figure C2. FIN of meaningful and meaningless visual perception, for multiple subjects using 47 feature voxel variables. The names of the nodes are explained in Appendix A.

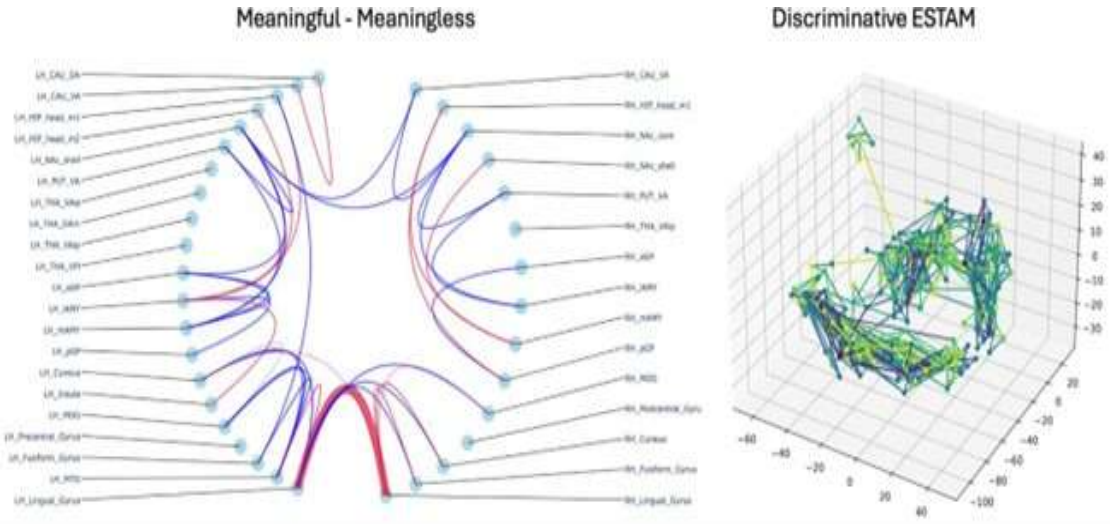


Figure C3. Differential ESTAMs for meaningful vs. meaningless visual perception across subjects. Blue edges indicate stronger interactions during meaningful conditions; red edges indicate stronger interactions during meaningless conditions. Edge width represents difference magnitude (threshold: ≥5% of maximum difference) (left). The names of the nodes are explained in Appendix A. ESTAM that discriminates between the two conditions (right). Darker colours mean spikes earlier in time, lighter colours represent spikes later in time.

The ESTAMs from Figures C1,2,3 show spatially distributed causal spiking patterns that unfold over time across several cortical and subcortical areas. The temporal evolution is visualised through colour, with darker shades indicating earlier spikes and lighter shades indicating later spikes in the sequence. The pattern suggests that visual meaningful information, as processed by the model, requires coordinated sequential processing spanning distributed regions, even when direct inter-cluster connections between specific areas are reduced. The complexity emerges not from strong bilateral connections, but from how information propagates through polysynaptic pathways across the network over time.

References

1. Butlin M, Lappas T 2024 Principles for Responsible AI Consciousness Research. J AI Research.

 doi:10.20944/preprints202604.0605.v1

2. Chalmers D 1996 The Conscious Mind. Oxford University Press.
3. Koch C 2004 The Quest for Consciousness: A Neurobiological Approach. Roberts & Company Publishers.
4. Dehaene S 2014 Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts. Penguin.
5. G Findlay, W Marshall, L Albantakis, I David, WGP Mayner, C Koch, G Tononi, Dissociating Artificial Intelligence from Artificial Consciousness, arXiv preprint arXiv:2412.04571, 2024•arxiv.org.
6. Hasson U, et al. 2008 A hierarchy of temporal receptive windows in human cortex. J Neurosci 28(10):2539-2550.
7. Kasabov, N K 2023, STAM-SNN: Spatio-Temporal Associative Memories in Brain-inspired Spiking Neural Networks: Concepts and Perspectives. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.23723208.v1
8. Kasabov NK, Bahrami H, Doborjeh M, Wang A 2023 Brain inspired spatio-temporal associative memories for neuroimaging data: EEG and fMRI. MDPI Bioengineering Preprints.
9. Kasabov NK 2024 STAM-SNN: Spatio-Temporal Associative Memory in Brain-Inspired Spiking Neural Networks. In: Kovács L, et al. (eds) Recent Advances in Intelligent Engineering. Topics in Intelligent Engineering and Informatics, vol 18. Springer.
10. Kasabov NK 2024 Life-long learning and evolving associative memories in brain-inspired spiking neural networks. MOJ App Bio Biomech 8(1):56–57.
11. Boly M, Sasai S, Gosseries O, et al. 2015 Stimulus Set Meaningfulness and Neurophysiological Differentiation: A Functional Magnetic Resonance Imaging Study. PLoS ONE 10(5): e0125337.
12. Nakai T, Koide-Majima N, Nishimoto S 2021 Music genre neuroimaging dataset. Data Brief 40:107675.
13. Orłowski P, Bola M 2023 Sensory modality defines the relation between EEG Lempel–Ziv diversity and meaningfulness of a stimulus. Sci Rep 13:3453.
14. Christoff K, et al. 2009 Experience sampling during fMRI reveals demfault network and executive system contributions to mind wandering. Proc Natl Acad Sci 106(21):8719-8724.
15. Izhikevich EM 2006 Polychronization: computation with spikes. Neural Comput 18(2):245-282.
16. Maass W 1997 Networks of spiking neurons: the third generation of neural network models. Neural Networks 10(9):1659-1671.
17. Kasabov NK 2014 NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. Neural Networks 52:62–76.
18. Kasabov NK 2018 Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence. Springer.
19. G.Tononi, An information integration theory of consciousness. BMC Neurosci. 5, 42 2004.
20. Kasabov, N. 2012. NeuCube EvoSpike Architecture for Spatio-temporal Modelling and Pattern Recognition of Brain Signals. In: Mana, N., Schwenker, F., Trentin, E. (eds) Artificial Neural Networks in Pattern Recognition. ANNPR 2012. Lecture Notes in Computer Science(), vol 7477. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33212-8_21.
21. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87 1999.
22. Kasabov NK, Dhoble K, Nuntalid N, Indiveri G 2013 Dynamic evolving spiking neural networks for on-line spatio- and spectro-temporal pattern recognition. Neural Networks 41:188-201.
23. Furber SB, Galluppi F, Temple S, Plana LA 2014 The spinnaker project. Proc IEEE 102(5):652–665.
24. Behrenbeck J, et al. 2019 Classification and regression of spatio-temporal signals using NeuCube and its realization on SpiNNaker neuromorphic hardware. J Neural Eng 16(2).
25. M. E. Futschik and N. K. Kasabov, "Fuzzy clustering of gene expression data," 2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No.02CH37291), Honolulu, HI, USA, 2002, pp. 414-419 vol.1, doi: 10.1109/FUZZ.2002.1005026.
26. Kohonen, Teuvo 1982. Self-Organized Formation of Topologically Correct Feature Maps". Biological Cybernetics. 43 (1): 59–69. doi:10.1007/bf00337288.
27. K.Kumarasinghe, N.Kasabov, D.Taylor, Deep Learning and Deep Knowledge Representation in Spiking Neural Networks for Brain-Computer Interfaces, Neural Networks, vol.121, Jan 2020, 169-185, doi: https://doi.org/10.1016/j.neunet.2019.08.029.

 doi:10.20944/preprints202604.0605.v1

28. Tan, C., Šarlija, M. & Kasabov, N. Spiking Neural Networks: Background, Recent Development and the NeuCube Architecture. Neural Process Lett 52, 1675–1701 2020. https://doi.org/10.1007/s11063-020-10322-8.
29. Kasabov, Nikola; Tan, Yongyao Tan; Doborjeh, Maryam; Tu, Enmei; Yang, Jie; Goh, Wilson; Lee, Jimmy 2023: Transfer Learning of Fuzzy Spatio-Temporal Rules in the NeuCube Brain-Inspired Spiking Neural Network: A Case Study on EEG Spatio-temporal Data. IEEE Transactions on Fuzzy Systems, vol.31, issue 12, Dec.2023, 4542-4552, Print ISSN: 1063-6706, Online ISSN: 1941-0034, DOI: https://doi.org/10.1109/TFUZZ.2023.3292802.
30. James C. Bezdek, Robert Ehrlich, William Full, FCM: The fuzzy c-means clustering algorithm, Computers & Geosciences, Volume 10, Issues 2–3, 1984, Pages 191-203, ISSN 0098-3004, https://doi.org/10.1016/0098-.
31. 3004(84)90020-7.
32. Paulun L, Wendt A, Kasabov NK 2018 A retinotopic spiking neural network system for accurate recognition of moving objects using neucube and dynamic vision sensors. Front Comput Neurosci , 12.
33. Grossman, E. D. & Blake, R. Brain areas active during visual perception of biological motion. Neuron 35, 1167–1175 2002.
34. Beauchamp, M. S. et al. Human MST but not MT responds to tactile stimulation. J. Neurosci. 22, 6196–6206 2002.
35. Grill-Spector, K. & Weiner, K. S. The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. 15, 536–548 2014.
36. Ranganath, C. et al. Working memory maintenance contributes to long-term memory formation: neural and behavioral evidence. J. Cogn. Neurosci. 17, 994–1010 2005.
37. Berridge, K. C. & Kringelbach, M. L. Pleasure systems in the brain. Neuron 86, 646–664 2015.
38. Adolphs, R. Fear, faces, and the human amygdala. Curr. Opin. Neurobiol. 18, 166–172 2008.
39. Rainer, G. et al. The effect of image scrambling on visual cortical activity. Exp. Brain Res. 138, 455–461 2001.
40. Lamme, V. A. & Roelfsema, P. R. The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. 23, 571–579 2000.
41. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47 1991.
42. Bar, M. The proactive brain: using analogies and associations to generate predictions. Trends Cogn. Sci. 11, 280–289 2007.
43. Koelsch, S. et al. Bach speaks: a cortical “language-network” serves the processing of music. Neuroimage 17, 956– 966 2002.
44. Zatorre, R. J., Belin, P. & Penhune, V. B. Structure and function of auditory cortex: music and speech. Trends Cogn. Sci. 6, 37–46 2002.
45. Seeley, W. W. et al. Dissociable intrinsic connectivity networks for salience processing and executive control. J. Neurosci. 27, 2349–2356 2007.
46. Salimpoor, V. N. et al. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat. Neurosci. 14, 257–262 2011.
47. Blood, A. J. & Zatorre, R. J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc. Natl Acad. Sci. USA 9811818–11823 2001.
48. Pearce, M. T. & Wiggins, G. A. Auditory expectation: the information dynamics of music perception and cognition. Top. Cogn. Sci. 4, 625–642 2012.
49. Rauschecker, J. P. & Scott, S. K. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12, 718–724 2009.
50. Song S, Miller KD, Abbott LF 2000 Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat Neurosci 3(9):919-926.
51. Hopfinger JB, et al. 2000 A study of analysis parameters that influence the sensitivity of event-related fMRI analyses. NeuroImage 11(4):326-333.
52. Poldrack RA, Mumford JA, Nichols TE 2011 Handbook of Functional MRI Data Analysis. Cambridge University Press.

 doi:10.20944/preprints202604.0605.v1

53. Tzourio-Mazoyer N, et al. 2002 Automated anatomical labeling of activations in SPM. NeuroImage 15(1):273-289.
54. Desikan RS, et al. 2006 An automated labeling system for subdividing the human cerebral cortex. NeuroImage 31(3):968-980.
55. Honey CJ, et al. 2009 Predicting human resting-state functional connectivity from structural connectivity. Proc Natl Acad Sci 106(6):2035-2040.
56. Sporns O, Tononi, G., Koetter, R. 2005 The human connectome: a strucxtural description of the human brain , PloS Comput Biol, 1(4), e42.
57. Hagmann P, et al. 2008 Mapping the structural core of human cerebral cortex. PLoS Biol 6(7), e159 28.
58. Deco G, Jirsa VK, McIntosh AR 2009 Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci 12(1):43-56.
59. Breakspear M 2017 Dynamic models of large-scale brain activity. Nat Neurosci 20(3):340-352.
60. Paugam-Moisy H, Bohte SM 2012 Computing with spiking neuron networks. In: Handbook of Natural Computing. Springer, pp 335-376.
61. Sporns O 2013 Structure and function of complex brain networks. Dialogues Clin Neurosci 15(3):247-262.
62. Friston KJ 2011 Functional and effective connectivity: a review. Brain Connect 1(1):13-36.
63. Pfeiffer M, Pfeil T 2018 Deep learning with spiking neurons: opportunities and challenges. Front Neurosci 12:774.
64. Vuust, P. & Witek, M. A. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. Front. Psychol. 5, 1111 2014.
65. Bestmann S, de Berker AO, Bonaiuto J 2015 Understanding the behavioural consequences of noninvasive brain stimulation. Trends Cogn Sci 19(1):13-20.
66. Jiawei Li, Jinyuan Liu, Shihua Zhou , Qiang Zhang and Nikola K. Kasabov, GeSeNet: A General Semantic-Guided Network With Couple Mask Ensemble for Medical Image Fusion, IEEE Transactions on neural networks and learning systems, vol.235, 11, pp.16248- 16261, Nov., 2024, https://doi.org/10.1109/TNNLS.2023.3293274.
67. Chuanji Gao, Jessica J Green, Xuan Yang, Sewon Oh, Jongwan Kim, Svetlana V Shinkareva, Audiovisual integration in the human brain: a coordinate-based meta-analysis, Cerebral Cortex, Volume 33, Issue 9, 1 May 2023, Pages 5574 – 584, https://doi.org/10.1093/cercor/bhac443.
68. Baars, B. J. 2005. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. Progress in brain research, 150, 45-53.
69. Dehaene S. Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts. New York: Penguin, 2014.
70. Song Q. and N. Kasabov, NFI: A Neuro-Fuzzy Inference Method for Transductive Reasoning, IEEE Transactions on Fuzzy Systems, Volume 13, Issue 6, pp 799-808, 2005, DOI:10.1109/TFUZZ.2005.859311.
71. Tian Y, et al. 2020 Topographic organization of the human subcortex unveiled with functional connectivity gradients. Nat Neurosci 23(11):1421-1432.
72. Salami M, et al. 2003 Change of conduction velocity by regional myelination yields constant latency. Proc Natl Acad Sci 100(10):6174-6179.
73. Lawrence M. Ward, The thalamic dynamic core theory of conscious experience, Consciousness and Cognition, 20, 2, 2011, Pages 464-486, https://doi.org/10.1016/j.concog.2011.01.007.