Integrating Space, Time, and Orientation in Spiking Neural Networks: A Case Study on Multimodal Brain Data Modeling

Neelava Sengupta, Carolyn B. McNabb, Nikola Kasabov, *Fellow, IEEE*, and Bruce R. Russell

***Abstract*— Recent progress in a noninvasive brain data sampling technology has facilitated simultaneous sampling of multiple modalities of brain data, such as functional magnetic resonance imaging, electroencephalography, diffusion tensor imaging, and so on. In spite of the potential benefits from integrating predictive modeling of multiple modality brain data, this area of research remains mostly unexplored due to a lack of methodological advancements. The difficulty in fusing multiple modalities of brain data within a single model lies in the heterogeneous temporal and spatial characteristics of the data sources. Recent advances in spiking neural network systems, however, provide the flexibility to incorporate multidimensional information within the model. This paper proposes a novel, unsupervised learning algorithm for fusing temporal, spatial, and orientation information in a spiking neural network architecture that could potentially be used to understand and perform predictive modeling using multimodal data. The proposed algorithm is evaluated both qualitatively and quantitatively using synthetically generated data to characterize its behavior and its ability to utilize spatial, temporal, and orientation information within the model. This leads to improved pattern recognition capabilities and performance along with robust interpretability of the brain data. Furthermore, a case study is presented, which aims to build a computational model that discriminates between people with schizophrenia who respond or do not respond to monotherapy with the antipsychotic clozapine.**

***Index Terms*— Diffusion tensor imaging (DTI), functional magnetic resonance imaging (fMRI), multimodal brain data, NeuCube, response to treatment, schizophrenia, spike-time-dependent plasticity (STDP), spiking neural network.**

I. Introduction and Background Work

IN THE recent past, noninvasive brain data collection techniques, such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), diffusion tensor imaging (DTI), and others, have made significant contributions to understanding various structural and functional properties of the human brain. Considerable improvements in data sampling technology over the past few years have enabled simultaneous sampling of multiple modalities of brain data a reality. It is evident that each data modality provides a unique but limited perspective of the brain. For instance, fMRI measures neural activity indirectly by measuring changes in cerebral blood flow (the hemodynamic response) over time. Energy consumption increases in areas of the brain that are more active, leading to increases in blood flow to replace lost oxygen and glucose. This is a slow response, measured 6–10 s after the initial event of neuronal excitation. Despite poor temporal resolution, fMRI provides excellent spatial resolution, making it a useful tool for brain research. EEG, on the other hand, has outstanding temporal resolution at the expense of spatial resolution. EEG measures cortical electrical activity at the scalp surface, and though the scalp does not impede electrical signals temporally, it causes spreading of the signal from its origin to a wider area, making source localization much more complex. In the past, these data modalities were used independently for pattern recognition and overlooked the joint information present in the data [1]. Algorithms with the ability to extract and integrate relevant information from various data sources into a single model are crucial not only for the robustness and performance of predictive models but also in terms of understanding the spatiotemporal relationships within the data.

In a clinical sense, pattern recognition algorithms provide novel and practical means of understanding the differences between patients and healthy controls and predict individual patients' responses to treatment. Within psychiatric research, in particular, machine learning has gained considerable momentum as a useful tool for developing predictive models of treatment response. Incorporation of multiple imaging modalities into these algorithms could provide increased reliability, especially in disorders where clinical diagnosis does not necessarily guide treatment. Reference [2] recently applied machine learning algorithms to predict treatment response in late-life depression using a combination of clinical and imaging data. Comparing several algorithms, they determined that alternating decision trees could most accurately predict treatment outcome in this cohort using a combination of structural and functional connectivity data [2]. Reference [3] has used EEG data to predict response to selective serotonin reuptake inhibitors in major depressive disorder and to clozapine in people with treatment-resistant schizophrenia (TRS) [4].

Manuscript received October 11, 2016; revised August 5, 2017 and January 10, 2018; accepted January 16, 2018. *(Corresponding author: Neelava Sengupta.)*

N. Sengupta and N. Kasabov are with the Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Auckland 1010, New Zealand (e-mail: neelava.sengupta@aut.ac.nz).

C. B. McNabb is with the School of Psychology and Clinical Language Sciences, University of Reading, Reading RG6 6AH, U.K.

B. R. Russell is with the Pharmacy, Health Sciences Department, University of Otago, Dunedin 9054, New Zealand.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNNLS.2018.2796023

Reference [5] also employed machine learning algorithms to predict response to clozapine, instead using a combination of clinical and pharmacogenetic data as input. Reference [6] employed machine learning techniques to predict treatment outcome in social anxiety disorder. Using task-based fMRI, they accounted for 40% of the variance in treatment response [6]. The challenge now is to create an algorithm that can incorporate brain data from different modalities.

One modality with potential for inclusion in such a model is DTI. DTI measures the net movement of water within a voxel. Water trapped within axons or dendrites is restricted to movement along the direction of those axons or dendrites, respectively; this can be measured using DTI, providing a "map" of neuronal tracts (white matter) within the brain. Decreased anisotropy within a voxel is often interpreted as a reduction in white matter "integrity," implying that white matter has been damaged or is diminished in some ways. However, a reduction in anisotropy may also be attributed to an increase in the number of crossing fibers, or in fact a reduction in the uniformity of fiber orientations. In terms of incorporation into a multimodal prediction model, DTI information might be utilized for its orientational implications rather than any interpretation regarding white matter integrity.

Structural dysconnectivity, as measured by DTI, has been demonstrated in several psychiatric disorders and has been shown to reflect functional dysconnectivity in some cases [8], [9]. In accordance with these theories, it would be appealing to incorporate dysconnectivity information into any algorithm that is designed to classify or predict outcomes in people with psychiatric disorders. This paper discusses the steps that we are undertaking to develop a new algorithm that can incorporate orientation information from DTI along with the EEG/fMRI activity data as a data fusion approach. Most commonly used method of an integrated data analysis for this kind of problem is by "converging evidence" [10]. Typically, each data type is analyzed separately, and results from other analyses that support's ones finding are discussed. Reference [10] also discussed an alternative data fusion analysis called "computation neural modeling." This is done by creating biologically realistic neural network models, where each network simulates data of certain type and compared with observed data. One major setback for this paradigm of data analysis is that the hypothesis-driven neural network model is built under several assumptions for simulated data generation. Hence, it is difficult to know, if any lack of agreement between observed and simulated data is due to the assumptions in the model, or simply wrong. A comprehensive review of the research in the direction of a multimodal brain data analysis is summarized in [1]. Some of the prominent works include integration of fMRI/EEG [11], fMRI/MEG [12], fMRI/DTI [13], [14], and fMRI/gene expression [15]. There is a third alternative for multimodal data integration known as data fusion, which is a more data-driven approach. Direct data fusion encompasses technique of directly fusing multiple data sets using statistical and machine learning algorithms. The data-driven methods span across the domain of the combined blind source separation techniques, such as independent component analysis (ICA) [16], [17] and its variants, multimodal cross correlation analysis [18]–[20], partial least squares [21], and others.

This paper is further presented in four sections. Sections II and III describe the various components of the proposed architecture for recognizing patterns from fMRI and DTI data by using multimodal information in the form of space, time, and orientation. In Section IV, we have demonstrated the properties of the proposed architecture using synthetic data-driven experiments and further extended our proposed framework to a case study of discerning treatment response in people with schizophrenia using the fMRI and DTI data. The discussion and the future directions of this paper are presented in Section V.

II. NeuCube Personalized SNNcube Architecture

Reference [22] recently proposed NeuCube, a spiking neural network-based pattern recognition system for representing and learning spatiotemporal associations in the data. The NeuCube architecture is composed of the following.

1) *Temporal Encoding Module:* The data encoding layer transforms real continuous data $\mathbb{R}^{N\times T}$ (N is the number of features and T represents the time duration) into spike trains $\{0,1\}^{N\times T}$. Numerous temporal encoding algorithms, such as Bens Spiker Algorithm (BSA) [23], temporal contrast, and GAGamma [24], are proposed and used in an application specific manner. The data encoding module in NeuCube acts as a data compression system, which compresses data in the temporal domain by representing events by spike timings. In the temporal encoding scheme, the timings of the spike are considered to be useful rather than the quantity of the spike. Reference [25] discusses the temporal encoding by spike-time representation in the light of data compression and information theory and compares the capabilities of different temporal encoding algorithms.
2) *Personalised SNNcube (SNNc) Learning Module:* The SNNc learning module is an unsupervised learning module composed of a spatial grid of spiking and input neurons. Each neuron inside the grid has a spatial location and resides within a neighborhood of other neurons. This grid is known as the spiking neural network cube. The purpose of this layer is to transform the compressed spike representation of input data into a higher dimensional space through unsupervised learning ($g : \{0,1\}^{N\times T} \rightarrow \{0,1\}^{M\times T} | M \gg N$) inside an SNNc, using a modified Hebbian spike-time-dependent plasticity (STDP) learning [26]. The configuration of SNNc is loosely inspired from Kohonen's self-organizing maps (SOMs) in a way that addition of spatial organization in a network of neurons adds visual characteristics and interpretability. This paper proposes a personalized multi-instance-based SNNc architecture (Fig. 1). In the proposed architecture (see Fig. 1), we have used subject-specific SNNc's where each SNNc is responsible for transforming the data from a single subject. The outputs of the multiple SNNc are later merged in the supervised readout layer. The personalized SNNc architecture assumes the existence of temporal relationships within a

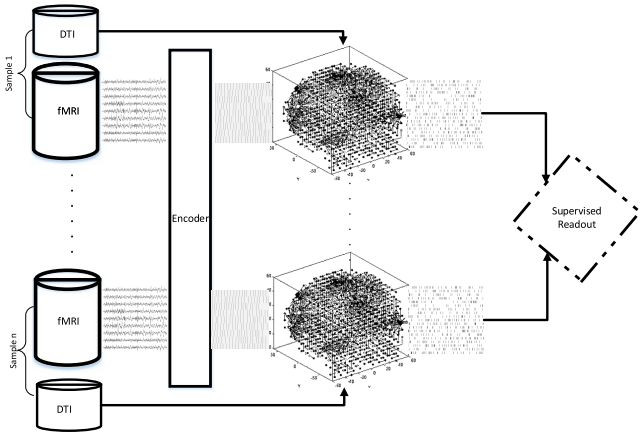


Fig. 1. Proposed NeuCube personalized SNNcube (SNNc) architecture. As a functional black box, this architecture acts as a machine learning model, which learns from static and dynamic data (during training phase) and can output predictions once trained. Every sample/observation in its raw form is assumed to be made up of dynamic activity data (fMRI/EEG) and static orientation data (DTI). The dynamic data are transformed to binary spike sequences. The unsupervised SNNc learning layer creates a single instance of SNNc for each subject data and learns from dynamic sample spike data using the static sample orientation data as structural constraints for learning. Each SNNc outputs subject-specific transformed higher dimensional spike data. These transformed higher order spike data are used by the supervised readout layer to build a classification model (e.g., a model that can discriminate a normal brain function from a disordered brain function) or regression model (e.g., a model that can predict IQ).

sample and not between samples, i.e., each sample maps its spatiotemporal relationship on its unique personal SNNc, which is used further for discrimination purposes.

3) *Supervised Readout Module:* This module uses the SNNc generated output spike sequences and/or the connection weights to learn a simple classifier or regressor [27]. KNN-based models [27] are the choice of supervised learning in almost all of the works done until now.

The NeuCube architecture is extended from the liquid state machines (LSMs)-based [28] reservoir computing paradigm. However, some of the considerable departures of NeuCube from LSM as discussed in [29] are as follows.

1) The first layer of the NeuCube framework encodes/compresses continuous streaming data into discrete spike timings using the state-of-the-art temporal encoding algorithms. The spike timings are considered to encode important events in the data stream.
2) The SNNc layer includes spatial mapping of input data into corresponding locations of the reservoir using 3-D spatial coordinates. The spatial mapping of neurons is inspired from the Kohonen's SOMs.
3) NeuCube has a learning reservoir or SNNc, which offers a certain degree of visual interpretability as a result of the spatial mapping and spatiotemporal learning in the SNNc.

We have, in this paper, analyzed and interpreted the SNNc (a special case of reservoir in LSM terminology) in Section IV. Within its vast range of applicability, this framework has been successfully used in classification and clustering of fMRI data [30].

III. Multimodal Information Integration in SNNc Using Orientation and Spike-Time Data

Section III-A formalizes the SNNc architecture as a directed graph and elaborates the connectivity scheme and initialization of the graph. The neuron model is discussed in Section III-B. Section III-C elaborates the SNNc learning scheme through the adaptation of synaptic efficacies by integrating orientation and spike-time information.

A. SNNc Architecture, Mapping, and Initialization Scheme

The SNNc architecture of NeuCube consists of a spatially arranged (in three dimensions) set of neurons. The neurons are partially connected by recurrent synapses forming a directed incomplete and acyclic graph. The neurons and synapses form the vertices and the edges of the graph. The SNNc architecture is thus formalized as a directed graph $G = \{M, C, W\}$, consisting of the set of neurons M, the set of directed synaptic connections C, and the corresponding synaptic weights W. The set M are categorized as follows.

1) *Input Neurons:* The input neurons, denoted by set $N \subset M$, feed the input spike data (generated by the encoding layer) to the SNNc. These neurons do not have any activations and do not perform any computations. The behavior of the input neurons is similar to the behavior of the input layer neurons in traditional neural networks. These neurons do not have presynaptic connections, i.e., a synapse can only originate from such a neuron.
2) *Spiking Neurons:* The spiking neurons, denoted by set $Q \subset M$, are the computational units of the network. The neurons, in this paper, are leaky integrate and fire (LIF) in nature. These neurons are responsible for performing computations on input data. The details of the neuron model of this category of neuron are described later. These neurons can act both as postsynaptic and presynaptic neurons, i.e., if we consider a pair of neurons connected by a directed synapse, the synapse can both originate and end at a spiking neuron.

The neurons in the SNNc are arranged spatially following the natural spatial arrangements in the data or by using automated mapping algorithms [31]. The spatial arrangement of the neurons is typically done in 2-D or 3-D space. The synaptic connectivity of the SNNc graph is typically heuristics driven. There are numerous heuristics to initialize the network, including data-driven and graph-theory-driven algorithms. Some preliminary works done on data-driven network initialization heuristics, however, showed no significant impact on robustness or performance of the model. In this paper, we have, therefore, used a graph-theory-based network initialization scheme known as the small-world connectivity (SWC) algorithm [32], [33]. The SWC algorithm connects a neuron to its spatial neighborhood (controlled by the hyperparameter radial distance r_{swc}) of neurons. For our purpose, contrary to the random initialization scheme of the synaptic weights, the synapses are initialized with a small constant weight of 0.05. We avoided random initialization of the network to avoid any predefined bias in the network.

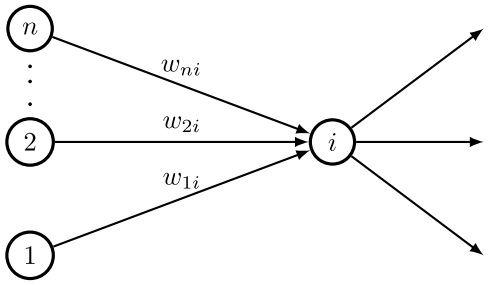


Fig. 2. Typical connectivity configuration of a spiking neuron i. A spiking neuron has a multi-input, multioutput configuration. A pair of neuron is connected by synapses. The synaptic strengths are represented by $\mathbf{w_i}$.

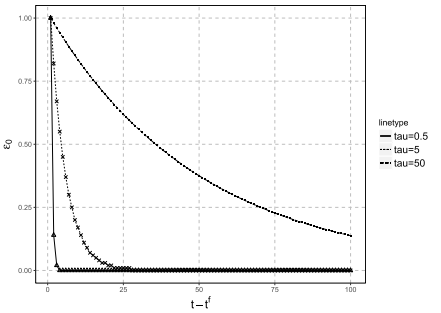


Fig. 3. Plot of the PSP trace ϵ_0 as a function of $\epsilon_0(t-t^f)$. This figure plots the PSP simulation for different values of τ_m.

B. Neuron Model

The activation of the spiking neurons present in the SNNc is modeled using the spike response model (SRM), which is a simplified realization of the LIF model. The SRM model generalizes the differential equation-based dynamics of the LIF model by replacing them with arbitrary kernels.

Fig. 2 shows a typical example of a spiking neuron's configuration. A spiking neuron receives spike data at different time instances from the presynaptic neurons and emits spike data when sufficiently stimulated. The activation state of a spiking neuron i is described by the membrane potential v_i. In a nonstimulated state, the membrane potential is said to be in a resting state $v_{\text{rest}} = 0$. The SRM model in our setup consists of multiple components.

1) Postsynaptic Potential Kernel: Firing of presynaptic neuron j at time t_j^f evokes a postsynaptic potential (PSP) in neuron i and is modeled by the kernel response function ϵ_0

$$\epsilon_0 = \exp\left(-\frac{t-t_j^f}{\tau_m}\right)\mathcal{H}\big(t-t_j^f\big) \tag{1}$$

where

$$\mathcal{H}(t-t_j^f) = \begin{cases} 1, & \text{if } t-t_j^f \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The PSP kernel is a function of $t-t_j^f$, representing the PSP trace over time generated by the firing of neuron j over time. Fig. 3 plots the PSP kernel as the function of $t-t^f$. τ_m [see (1)] is the membrane constant, which controls the decay rate of the PSP. In our experiments, we have used a constant $\tau_m = 0.5$. This means that the influence of a presynaptic spike diminishes from 1 to 0 within five discrete time intervals.

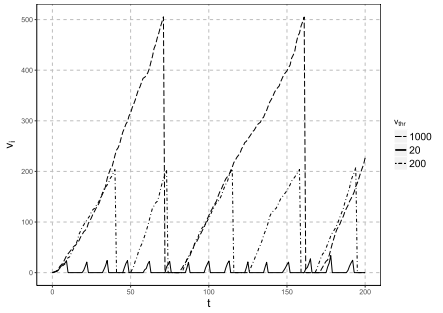


Fig. 4. Plot of the membrane potentials (v_i) of a neuron i simulated over $T = 200$ time points using the SRM model. For the simulations, we connected three predecessor neurons to a spiking neuron. The spike data from the predecessor neurons are sampled randomly from the uniform random distribution. η_{thr} for the spiking neuron was set to 10. Each of the three v_i traces corresponds to a preset v_{thr} mentioned in the label.

2) Temporal Integration of PSP Kernels and Condition for Spike Emission: Under the SRM model, the PSPs evoked by the presynaptic neurons are temporally integrated to activate the spiking neuron. The overall contribution of the presynaptic spikes elicited by the presynaptic neurons j at any time t is given as (3) describing the SRM model

$$v_i(t) = v_{\text{rest}} + \sum_{j\in T_i} w_{ji} \sum_{t_j^f \in F_j} \epsilon_0\big(t-t_j^f\big). \tag{3}$$

The temporal summation is a double summation operation. The inner sum adds up the PSP contributions due to the firings $t_j^f \in F_j$ of one presynaptic neuron. The outer sum adds up the PSP contributions of all the presynaptic neurons $j \in \mathcal{T}_i$ connected to neuron i. Equation (3) describes that the membrane potential (activation state) v_i of a spiking neuron i can be calculated by adding the resting potential term and the temporal PSP sum. Each incoming spike perturbs the value of v_i and if, after the summation of the inputs, the membrane potential reaches the threshold v_{thr}, then an output spike is generated. The firing time is given by the condition $v_i(t_i^f) >= v_{\text{thr}}$. After a neuron fires the neurons, membrane potential is reset to v_{rest}.

3) Refractory Period: After emitting the spike, a spiking neuron enters a refractory period, when the membrane potential is unaffected by incoming spikes. In the SRM model, the neuron behavior in the refractory period depends only on the last firing moment leading to a short-term memory in the neuron. In the literature, the refractory period is modeled by absolute and relative refractory period. During the absolute refractory period, the neurons do not accumulate membrane potential and hence cannot fire. During the relative refractory period, it is relatively difficult but not impossible to fire the neuron. In our implementation, we have used an absolute refractory period for the sake of simplicity. The absolute refractory period of a neuron is specified by the hyperparameter η_{thr}.

Fig. 4 shows a plot of three simulations of a spiking neuron for 200 discrete time with random spike inputs. Each simulation uses a preset v_{thr}. At the beginning of the simulation, the neuron is in a resting state $v_{t=0} = v_{\text{rest}}$. With arrival of spikes, the membrane potential increases in a linear fashion, and when sufficiently stimulated (sufficiency is

determined by v_{thr}), the neuron spikes and goes back to the resting state. At this point, the neuron is said to be in a refractory state. The neuron stays in this state for a predeterminer period η_{thr} and then goes back to the nonrefractory state.

C. Unsupervised Weight Adaptation of Synapses

The unsupervised learning algorithm in the SNNc is the most important aspect of our proposed architecture for integrating multimodal information and is the main contribution. In a neural network paradigm, learning is achieved through the modification of synaptic strength of the network. The learning behavior of the SNNc can be explained using the learning model of a single spiking neuron. Considering the single neuron architecture in Fig. 2, the unsupervised learning problem is a scheme of updating the w_{ji} values by $\Delta w_{ji}(t)$ over the simulation time T. In a recurrent SNNc layer, our aim is to learn both dynamic influences from dynamic data (fMRI or DTI) and static orientation influence from static data (DTI).

1) Dynamic Influence From fMRI/EEG and Spike-Time-Dependent Plasticity Learning: In majority of the machine learning applications, models are trained on static data, where a sample is represented by a vector of numbers $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, where each number represents the value of a feature. But in our case with fMRI or EEG data, a sample is represented by the form of a matrix $X_{\text{seq}} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$, where $\mathbf{x_i} = \{x_i(1), x_i(2), \ldots x_i(t)\}$. This form of sample representation is unique. It is not only 2-D sample representation but also is ordered in time (sequence). Learning from these types of data sequences in the machine learning is known as sequence learning and techniques, such as hidden Markov model, and flavors of recurrent neural network have shown promise in learning from such sequences. Here, we describe an unsupervised sequence learning framework in NeuCube SNNc that uses data sequences as part of its learning. This is referred to as dynamic influence and is denoted by ϕ. The NeuCube SNNc being a recurrent spiking neural network architecture receives the data in the form of binary spike sequence, i.e., $X = \{0, 1\}^{n \times t}$.

We modeled the dynamic influence from the spike-time data using the STDP learning method. STDP is a temporally asynchronous form of Hebbian learning ("neurons wire together, if they fire together") [34] induced by the temporal correlation of the spikes. This biological process in the brain adjusts the synaptic strength based on the relative timing of a neuron's input and the output spikes. With STDP, repeated presynaptic spike arrival just before postsynaptic spike leads to long-term potentiation (LTP) of the synapses establishing causal relationship, whereas repeated spike arrival after postsynaptic spikes leads to long-term depression (LTD) of the same synapse.

References [26] and [35] formalized the mathematical model of STDP learning as per (4) and (5). Symbols j and i are used to indicate presynaptic and postsynaptic neurons. In STDP learning, The dynamic influence ϕ_{ji} is estimated using a learning window function $W(\cdot)$. The learning window takes historical presynaptic firing times $\{t_j^1 \ldots t_j^f\}$ and postsynaptic firing times $\{t_i^1 \ldots t_i^g\}$ as an input and calculates the LTP and LTD traces. These historical firing times are nothing but indices of a historical spike sequence. For example, a historical spike sequence [01001011] can be rewritten as sequence of spike-time indices $t^f := \{1, 4, 6, 7\}$. Exponential decay functions are a popular choice for the learning window, and we use (5) as the learning window function for all our simulations. The κ_+ and κ_- parameters control the maximum LTP and LTD update, respectively, and we have chosen $\kappa_- = \kappa_+ = 1$ to keep the bounds of dynamic influence between $[-1, 1]$. From (5), it can be observed that the polarity of $(t_i^g - t_j^f)$ defines the polarity of ϕ_{ji}. This is a causal Hebbian relationship model, where synapses are rewarded positively (strengthened) for causal firing (i fires later than j, i.e., firing of i is caused by firing of j) and penalized (weakened) for noncausal firing. However, (4) and (5) describe a batch update scheme and require modification for online learning in the SNNc. Reference [36] proposed a modified online STDP update rule. In the online setup, ϕ is calculated every time neuron i fires a spike or receives a spike from neuron j. Equation (6) formalizes the dynamic influence update for the online mode. The first term in the right-hand side of (6) corresponds to the LTP update and is calculated when neuron i fires a spike at time t. The second term is the LTD update and is calculated when neuron i receives a spike from neuron j at time t. Both the batch and online formalizations of STDP learning are extended from [37], which discusses the properties of the STDP learning model extensively.

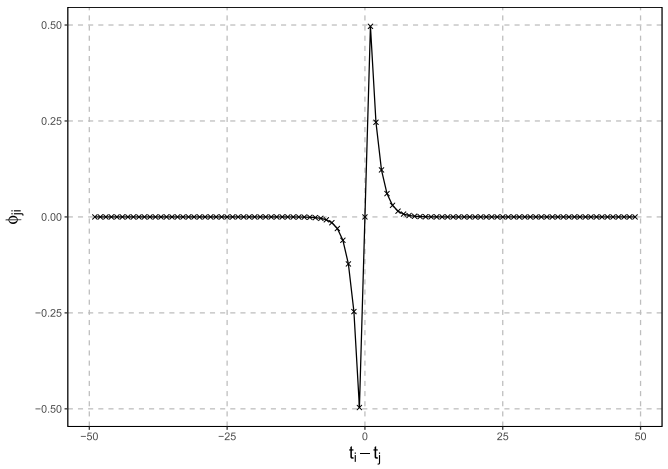


Fig. 5. STDP weight update as a function of the relative timing of the presynaptic and postsynaptic spikes.

Fig. 5 shows the plot of the STDP learning function where the dynamic influence in quadrants I and III corresponds to LTP and LTD, respectively

$$\phi_{ji} := \sum_f \sum_g W\left(t_i^g - t_j^f\right) \tag{4}$$

$$W(s) := \begin{cases} \kappa_+ \exp(-s), & \text{if } s > 0 \\ -\kappa_- \exp(-s), & \text{if } s < 0 \end{cases} \tag{5}$$

$$\phi_{ji}(t) := \sum_f \kappa_+ \exp\left(-\left(t - t_j^f\right)\right) - \sum_g \kappa_- \exp\left(-\left(t - t_i^g\right)\right). \tag{6}$$

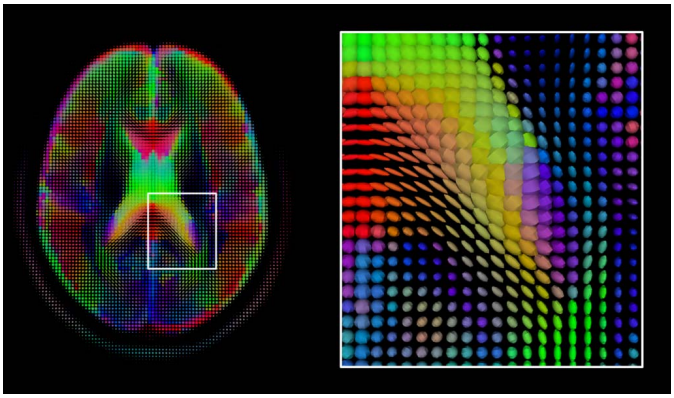


Fig. 6. Orientation information from DTI image. Left: axial slice of a single subject's DTI data, registered to structural and Montreal Neurological Institute (MNI) standard space. Right: close-up of the right posterior corpus callosum. Orientations corresponding to each color are as follows. Red: left to right/right to left. Green: anterior to posterior/posterior to anterior. Blue: superior to inferior/inferior to superior [7].

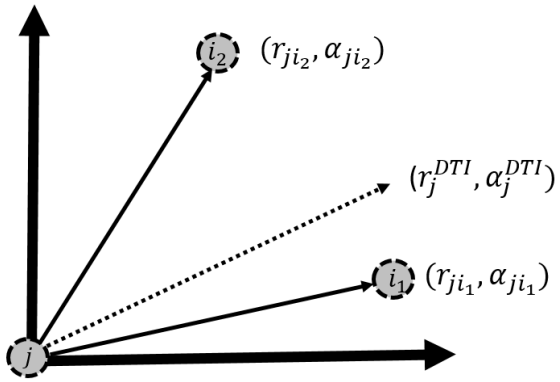


Fig. 7. Example of a presynaptic neuron j connected to two postsynaptic neurons i_1 and i_2. Each neurons spatial location is defined by the polar coordinates (r, α).

It is evident from the above discussion that the STDP learning rule enhances or depletes the synaptic strength of the connections based on the relative coincidence of the spikes. This behavior mimics the ability of the biological neurons to encode information by detecting the occurrence of temporally close but spatially distributed input signals and thus incorporating spatiotemporal information in the model.

2) DTI Fiber Tractography Data and Static Orientation Influence: Diffusion tensor weighed images are represented as a 3-D image made of a set of spatially arranged glyphs, as shown in Fig. 6. Each glyph/voxel (color and orientation) in the image is characterized by a rotation invariant ellipsoid representing the properties of the molecular diffusion of water in that region. Due to the tensorial nature of the ellipsoid, the raw DTI voxel information is stored as a second-order positive-definite tensor

$$D := \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix}.$$

The six unique elements of the tensor are the coefficients of the ellipsoid equation given by $D_{xx}x^2 + D_{yy}y^2 + D_{zz}z^2 + D_{yx}yx + D_{zx}zx + D_{xy}xy = 1$. The diffusion properties of the ellipsoid are characterized by the eigenvectors and eigenvalues of the tensor, which correspond to the magnitudes and directions of the anisotropy. For example, in the areas with isotropic diffusion, the shape of the ellipsoid will be nearly spherical with small anisotropy measure [38]. From this information, it is possible to determine fractional anisotropy but not the fiber orientation. Fiber tractography is a very elegant method for delineating individual fiber tracts from diffusion images. The primary fiber orientation (or primary diffusion direction) in each voxel can be estimated using the Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques (BEDPOSTX) [39] to give an orientation vector per voxel. In this paper, we have used the DTI data in the form of orientation vectors representing mean orientation of the fiber tract at different voxel locations. The orientation vector of a sample DTI image is represented by a matrix $\mathbf{X_{or}} \in \mathbb{R}^{n\times 3}$, where each feature is represented by a 3-D vector describing the orientation of the fiber in the Cartesian coordinate system.

Here, we are interested in establishing a learning rule that can not only accommodate both dynamic data influences as described in Section III-C1 but also static orientation influence from the DTI data. The orientation influence is denoted as ψ. The intuition behind the orientation influence can be explained by a small SNNc architecture consisting of three neurons, as shown in Fig. 7. Fig. 7 shows a single presynaptic neuron j connected to two postsynaptic neurons i_1 and i_2. The important thing to note here is that the neurons in this diagram have spatial allocations as opposed to the one in Fig. 2. The location of the neurons is defined by the radial and the angular coordinate in the polar coordinate system. Now, we are interested in calculating the orientation influence of neuron j on neurons $\{i_1, i_2\}$. We refer neuron j as the pivot neuron. Let the orientation vector of the pivot neuron (from DTI data) be represented by (r_j, α_j). The orientation influence of the pivot neuron on neurons $\{i_1, i_2\}$ are defined by their angular proximity pivot neuron's orientation vector. In that way, as per our hypothesis, the pivot neuron wields a stronger angular influence on the neurons as they lie closer in angular proximity to the orientation vector of the pivot neuron. Hence, the influence of neuron j can be arranged as $i_1 > i_2$ due to the angular proximity of i_1 and j being greater than i_2 and j.

Even though we have used a 2-D vector space for explaining the intuition of angular influence, our neurons in the SNNc reside in a 3-D space. The intuition is extendible to a 3-D vector space by adding another dimension in the coordinate representation. In 3-D, the spherical coordinates of a point are given by (r, α, β), where r is the scalar distance of the point from the center, and α and β are the elevation and azimuth angles from the center. A Gaussian radial basis function (GRBF) is used to realize the elevation and azimuth orientation influences given the elevation and azimuth data of the neurons. The orientation influence between the pivot presynaptic neuron j and postsynaptic neuron i is

$$\psi_{ji} = \frac{\psi_{ji}^{\alpha} + \psi_{ji}^{\beta}}{2}$$

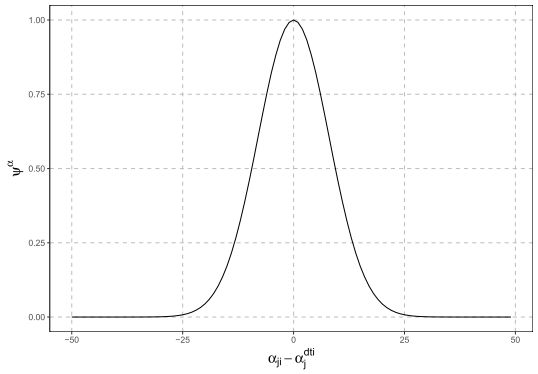


Fig. 8. Plot of the elevation influence ψ^α as a function of the radial distance $\alpha_{ji} - \alpha_j^{dti}$ and $\sigma = 8$.

$$\psi_{ji}^{\alpha} = \exp\left(\frac{\left\|\alpha_{ji} - \alpha_j^{dti}\right\|^2}{2\gamma^2}\right)$$
$$\psi_{ji}^{\beta} = \exp\left(\frac{\left\|\beta_{ji} - \beta_j^{dti}\right\|^2}{2\gamma^2}\right). \quad (7)$$

The GRBF exponentially decays the orientation influence as the Euclidean norm $||\alpha_{ji} - \alpha_j||$ and $||\beta_{ji} - \beta_j||$ increases. The variance hyperparameter σ^2 controls the speed with which the orientation influence decays with increasing radial distance (see Fig. 8). The overall orientation influence is calculated as the mean of the elevation and azimuth influence, as shown in (7). It is to be noted that apart from the orientational influence, a radial distance r between the neuron pairs is possible to be included using a similar formalization. The inclusion of a radial distance as a structural constraint will have a similar effect to using a radial distance for synaptic delay in the PSP calculation.

D. Orientation Influence-Driven STDP Learning in SNNc

Algorithm 1 describes the orientation influence-driven STDP (oiSTDP) learning algorithm step by step. The unsupervised learning is executed on a preinitialized SNNc G. The algorithm takes: 1) the dynamic data D_{seq} (fMRI/EEG) in the form of spikes; 2) the static orientation data D_{stat} as 3-D orientation vectors; and 3) the spatial coordinates loc of the SNNc neurons as an input. The output of the learning algorithm is the spike train O_{seq} in the higher dimensional space. Over simulation time T, the execution of the algorithm is divided into two blocks.

1) *Block 1 (Lines 6–17):* First, each spiking neuron in set Q is queried for the presynaptic neurons J_i^{spk}, which has fired a spike at time t. These connections are stored in C_{learn} for later updating the weights as per the LTD rule. Then, the spiking neuron is simulated to update the membrane potential. If the neuron spikes as a result of neuron simulation (line 10), a spike is added to O_{seq}, and the presynaptic connections are again added to the variable C_{learn} for later weight update as per the LTP rule.

Algorithm 1 oiSTDP-Based SNNc Learning Algorithm

1: input: $G = \{M, C, W\}$, $D_{seq} \in \{0,1\}^{|N|\times|T|}$, $D_{stat} \in \mathbb{R}^{|N|\times 3}$, $loc \in \mathbb{R}^{|M|\times 3}$,$\{hyperparameters = v_{thr}, \eta_{thr}, \kappa\}$
2: output: $O_{seq} \in \{0,1\}^{|M|\times|T|}$
3: $O_{seq}[N,T] \leftarrow D_{seq}$
4: **for** $t \in T$ **do**
5: initialise $C_{learn} \leftarrow \{\}$
6: **for** i in spiking neurons Q **do**
7: find firing(at time $t-1$) pre-synaptic neurons, $J_i^{spk(t)}$
8: set $C_i^{ltd} \leftarrow (J_i^{spk}, i)$
9: $C_{learn}+ \leftarrow C_i^{ltd}$
10: simulate i as per eq. 3
11: **if** i fires a spike **then**
12: $O_{seq}[i, t+1] \leftarrow 1$
13: find pre-synaptic neurons, J_i
14: set $C_i^{ltp} \leftarrow (J_i, i)$
15: set $C_{learn}+ \leftarrow C_i^{ltp}$
16: **end if**
17: **end for**
18: **for** c_{ji} in C_{learn} **do**
19: initialise $\psi_{ji} \leftarrow 1$
20: **if** j in N **then**
21: $(r_j^{dti}, \alpha_j^{dti}, \beta_j^{dti}) \leftarrow cart2sph(D_{stat}[j])$
22: $(r_{ji}, \alpha_{ji}, \beta_{ji}) \leftarrow cart2sph(loc[i] - loc[j])$
23: calculate ψ_{ji} as per eq. 7
24: **end if**
25: calculate ϕ_{ji} as per eq. 6
26: set $\Delta w_{ji} \leftarrow \psi_{ji} \cdot \phi_{ji}$
27: update $w_{ji} \leftarrow w_{ji} + \Delta w$
28: **end for**
29: **end for**

2) *Block 2 (Lines 18–28):* This block implements the weight update rule. For every time index t, synaptic strengths are updated for all the connections stored in C_{learn}. Lines 20–24 are the steps for calculating the orientation influence ψ. The function $cart2sph(.)$ takes a 3-D Cartesian coordinate (x, y, z) as an input and outputs the spherical coordinate (r, α, β). The following formula is used for this conversion:

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \alpha = \tan^{-1}\frac{\sqrt{x^2+y^2}}{z}$$
$$\beta = \tan^{-1}\frac{y}{z}.$$

The weight update Δw_{ji} for connection C_{ji} is the product of ψ_{ji} and ϕ_{ji} (line 27). In this way, we use the orientation influence as a modulation factor of the dynamic influence. ϕ and ψ are bound to $[-1, 1]$ and $[0, 1]$, respectively. Henceforth, Δw is bound between $[-1, 1]$. The most important characteristic of this learning rule is, of course, the inclusion of the orientation influence along with the dynamic influence in the weight update rule formulation. The rationale behind this formulation is to bind the coincidence (STDP) and tract information together in a way, such that the strongest

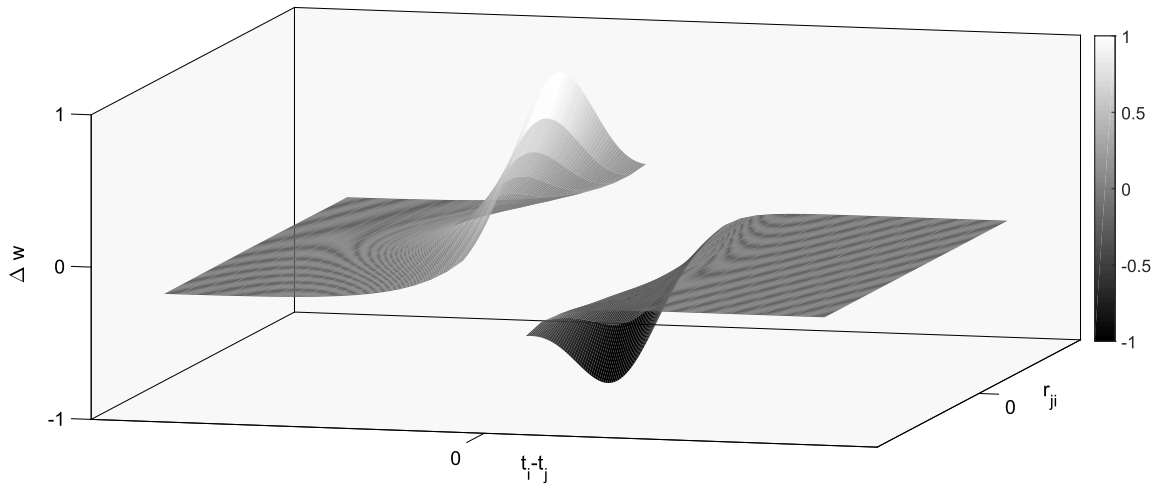


Fig. 9. Graph showing the relationship of oiSTDP weight update Δw with postsynaptic and presynaptic firing time difference $t_i - t_j$ and orientation distance r_{ji}. As the temporal difference between neuronal spikes decreases, the effect on weight updating increases, so that spikes timed closely together lead to greater increases in weight updating than spikes timed further apart. The order of spikes also affects weight updating. If neuron j fires before neuron i consistently, then the synaptic weight between them continues to increase; however, if the order switches, the weight is reduced.

weight update occurs between a neuron pair when: 1) there is maximal coincidence between the presynaptic and the postsynaptic firing and 2) the orientation of a neuron pair is in the matches the DTI orientation data of the presynaptic neuron. Hence, the relations observed in the synaptic strengths of the SNNc network are representative of both the spatiotemporal coincidence generated from the encoded spike sequence and the orientation information produced by the DTI data. In this way, we capture the spatial, temporal, and orientation information in the synaptic strengths. Fig. 9 shows the behavior of the oiSTDP update rule. The x- and y-axes are the radial orientation distance between the neuron pair and the temporal difference between the pre and postsynaptic firing time, respectively. Fig. 9 shows a mirrored inverted half Mexican hat behavior. It is visible that every slice across r_{ji}-axis mimics the STDP behavior shown in Fig. 5. The top half of the Mexican hat relates to the positive LTP weight update, which peaks at the minimum angular distance and decays with increasing angular distance. On the contrary, the bottom inverse half relates to negative LTD weight update, which achieves a negative peak at the minimum angular distance. In this way, the angular proximity of the neurons plays a role in modulating the spike synchronicity-driven dynamic influence.

IV. Experimental Results and Evaluation

This section is subdivided into two sections. In Section IV-A, we have analyzed the behavior and the effect of oiSTDP algorithm on synthetically generated data sets, and Section IV-B describes the application of the proposed algorithm on the clinical data.

A. Analysis of oiSTDP Algorithm Using Synthetic Data

To describe and evaluate the oiSTDP learning algorithm, we have used synthetically generated activity and orientation information. The input spike data D_{seq} are of size 128×14, and mimic a random sample of 1 s generated by a 14-channel EEG device with a sampling frequency of 128 Hz. We have used $M = 1485$ neurons with sparse recurrent connections. The neurons in the SNNc are spatially distributed to mimic the shape of the brain [22]. The location of the input spike train in the reservoir is resolved as per the natural spatial ordering of EEG channels-AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, and F8. The SWC algorithm is used to initialize the SNNc network. We have used $r_{\text{swc}} = 0.02$ (meaning connect neurons within 2% of the maximum distance) for connection generation and $W_{\text{init}} = 0.05$. The spiking neurons are simulated using the parameters ($\eta_{\text{thr}} = 4, \upsilon_{\text{thr}} = 0.1$, and $\kappa_- = \kappa_+ = 0.01$).

1) Effect of the Orientation Information on SNNc: Here, we show the systematic effect of orientation information in the SNNc map. As discussed in Section IV, the oiSTDP learning rule represents orientation information in conjunction with the spatiotemporal information in the connection strengths. As we are interested in the orientation information, we have sampled the spike train D_{seq} from Poisson's distribution to keep minimal spike synchronicity in the data. As per Algorithm 1, this would mean that the spike data will have a minimal influence in the SNNc map. Fig. 10 shows the systematic effect of a different input orientation information on the final 3-D SNNc map. In first of the three experiments, all the SNNc neurons were fed with input orientation data ($\alpha = 0°, \beta = 0°$), i.e., parallel to the x-axis and perpendicular to the z-axis. It is clearly visible from Fig. 10(a)–(c) that the strongest connections in the SNNc are representative of the orientation information provided. The second and the third experiment uses ($\alpha = 45°, \beta = 45°$) and no orientation information, respectively. It is evident from Fig. 10 that in absence of the temporal information (synchronicity), the orientation information is reflected in the strongest connections of the SNNc, and as such, in simple cases, they are visually discriminatory. Since each of these learned SNNc is represented by a directed

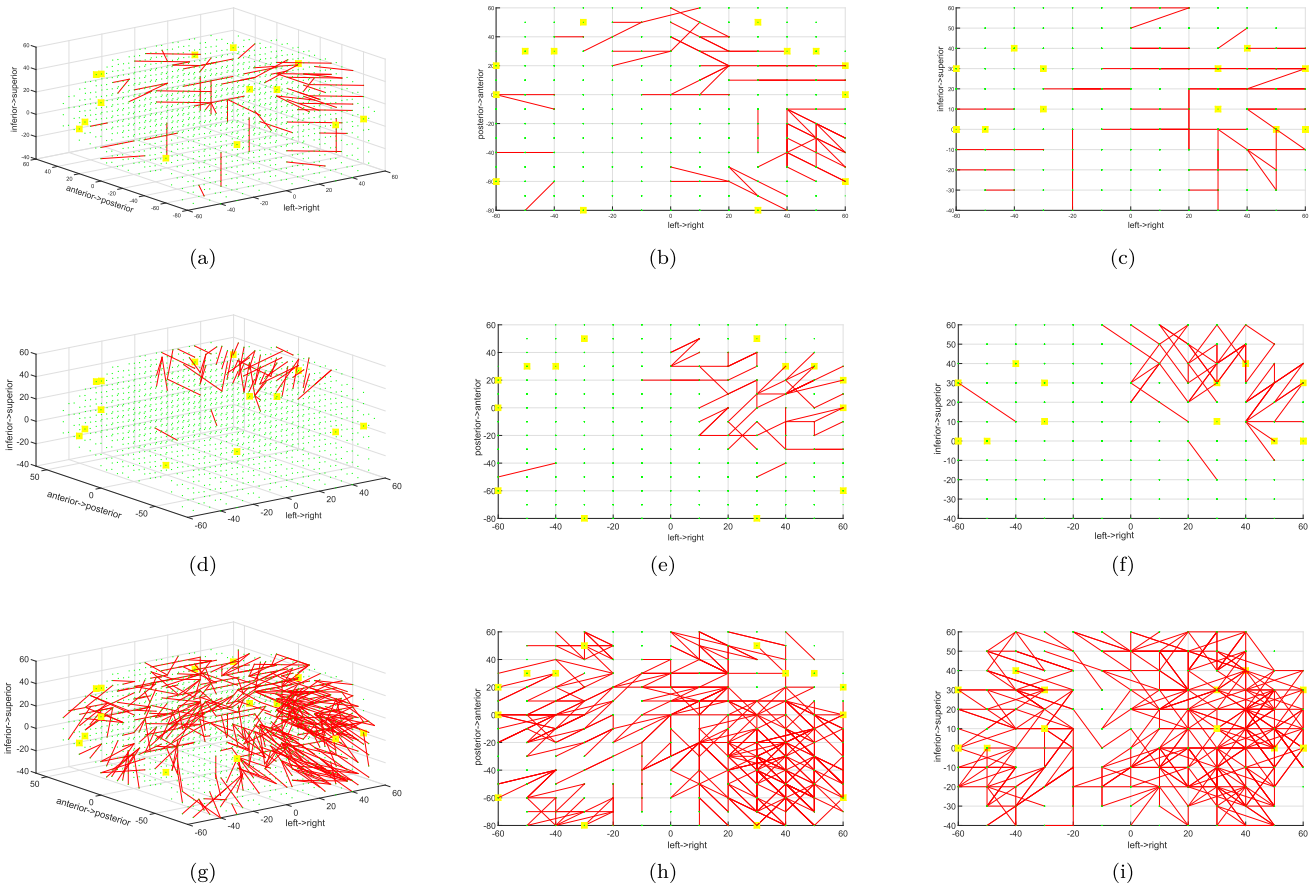


Fig. 10. Columns: 3-D, horizontal, and coronal views of the strongest connection weights in the 3-D SNNc at the end of the unsupervised learning. Each row corresponds to different input orientation data. Every neuron of the SNN in the first row was simulated with orientation data ($\alpha = 0°$, $\beta = 0°$). This resulted in the majority of the strong connections being oriented in the direction ($\alpha = 0°$, $\beta = 0°$). This shows the systematic bias toward orientation information in the absence of any dynamic bias. The second row shows a similar systematic bias toward the input orientation ($\alpha = 45°$, $\beta = 45°$). The third row is the baseline with no orientation influence where we cannot observe systematic orientation information bias.

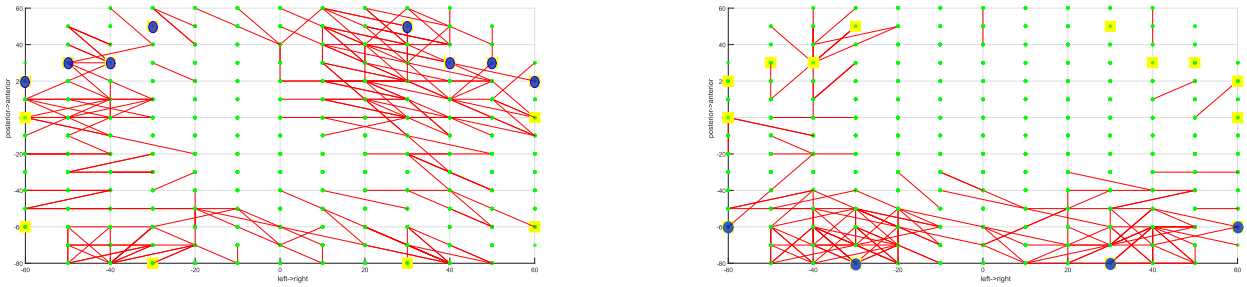
graph of weighed connections, connectomics analysis [40] can further be performed on the learned SNNc to extract new and useful knowledge in the spatiotemporal domain.

2) Effect of the Spike Synchronicity on SNNc: The aim of this experiment is to show the effect of spike synchronicity, i.e., the effect of STDP learning [see (6)] on the SNNc map for different spatiotemporal patterns. Per the STDP learning rule, greater synchronicity leads to stronger connections through LTP. To demonstrate the effect of the spatiotemporal synchronicity, we have created two samples of the input spike data. In the first sample, the spike sequences corresponding to the channels in the frontal lobe of the brain are kept the same (mimicking 100% synchronicity), and in the second sample, 100% spike synchronicity is kept at the occipital and parietal lobes. Fig. 11 shows the comparison between the two SNNc maps created by the oiSTDP learning algorithm. The "strongest connection" density is clearly more prominent in the frontal lobe in Fig. 11(a) due to the greater input spike synchronicity in that region. Similar clusters [Fig. 11(b)] at the parietal and occipital lobes can be seen with when the second sample is used. Through these analyses, we have demonstrated how different temporal patterns and the spatial arrangement of such patters affect the visual map of SNNc through the oiSTDP learning.

B. NeuCube With oiSTDP Learning for the Classification of Responders and Nonresponders to Clozapine Monotherapy

This paper was conducted as part of a large cross-sectional study investigating clozapine (CLZ) response in people with TRS using EEG, MRI, and genetic information (TRS study). The study was approved by the Health and Disability Ethics Committee and received locality approval from the Auckland and Counties Manukau District Health Boards of New Zealand. CLZ is uniquely effective for TRS. However, many people still suffer from residual symptoms or do not respond at all [ultratreatment-resistant schizophrenia (UTRS)] to CLZ.

In this paper, our aim was to build a model for discriminating CLZ monotherapy respondent and nonrespondent individuals from multimodal fMRI and DTI brain data. It must be noted that the data used to build this model have been collected post-CLZ treatment. For our investigation, we used a subset of data collected from the TRS study with the intention of classifying subjects into groups with either TRS or UTRS using the resting-state fMRI and DTI data. A total of 30 subjects with no recorded head injury and aged from 18 to 45 years were chosen for the study. Sixteen subjects belonged to the TRS group and fourteen to the UTRS group. The resting-state fMRI data were collected using a 3-T Siemens Skyra



(a) (b)

Fig. 11. Comparison the effect of synchronous input spikes on the SNNc map generated by algorithm 1. The blue dots show the synchronous input channels. (a) Synchronous input spike train at locations AF3, F7, F3, FC5, FC6, F4, F8, and AF4. (b) Synchronous input spike train at locations P7, O1, O2, and P8.

Magentom Scanner with TR = 3000 ms and TE = 30 ms for a duration of 8 min.

The fMRI and DTI data were preprocessed using standard preprocessing methods to control for head movement, registration, and normalization using FSL software (version 5.0.7) [41]. Both fMRI and DTI data for each subject were registered to a subject-specific structural image and normalized to the MNI-152 2 mm atlas [42], [43]. ICA-based automatic removal of motion artifacts was used to remove motion artifacts from the fMRI data utilizing FSL's FEAT output [44], [45]. Structural images were reoriented to a standard template, and brain tissue was extracted from raw image files using FSL's brain extraction tool (BET). If automatic brain extraction failed to eliminate all nonbrain tissues, the excess was removed manually. Magnitude images were subjected to the same process, after which brain-extracted images were eroded to ensure that no voxels containing nonbrain tissue remained. Fieldmaps were then created. The gradient-free image was used to create a binary mask with BET. Gradient distortions were corrected using FSL's fugue function and output registered to gradient-free images using the FMRIB Linear Image Registration Tool (FLIRT). Data were then corrected for head movement and eddy current distortions using FSL's eddy tool. Slices with average intensity at least four standard deviations lower than the expected intensity were interpolated with predictions made by the Gaussian process. DTIfit was used to independently fit diffusion tensors to each voxel, limited to brain space using the binary brain mask. Crossing fibers were modeled using BEDPOSTX. BEDPOSTX estimates of primary fiber orientations (dyads1) were then warped to a standard MNI template for use in the initial NeuCube construction. The second stage of data processing focused on selecting a set of voxels from the fMRI and DTI to be used to build the multimodal NeuCube model. As discussed in Section III, since a major component of our model captures temporal variations in data and the noise reduction capabilities of SNN architectures through encoding [46], we hypothesize that the discriminatory information is hidden in the voxels with significant variation in the activity over time. We selected a set of voxels with an absolute mean standard deviation of greater than 105 for our experiments. Fig. 12 shows the 3-D atlas locations of the selected voxels in the MNI coordinate system.

TABLE I

FREQUENCY TABLE OF ROIS OF THE SELECTED VOXELS

ROI	# voxel	%
Frontal Lobe	177	7.64
Insula	16	0.69
Temporal Lobe	138	5.95
Cerebellum	1557	67.17
Occipital Lobe	25	1.08
Parietal Lobe	134	5.78
Thalamus	187	8.07
Caudate	84	3.62

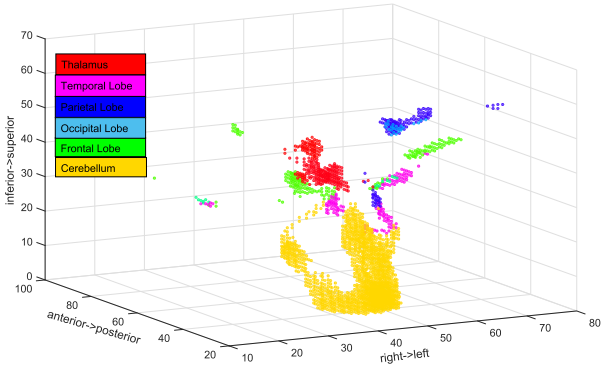


Fig. 12. Selected voxels for pattern recognition.

The selected voxels are predominantly (>67%) located in the cerebellum of the brain. The second and third column of the Region of Interest (ROI) frequency table (see Table I) also corresponds to the number and the percentage of voxels belonging to the different ROIs.

The final preprocessed data set consists of dynamic fMRI trials $\mathbf{fMRI_{seq}} \in \mathbb{R}^{30\times2318\times80}$ and static DTI orientation vector data $\mathbf{DTI_{stat}} \in \mathbb{R}^{30\times2318\times3}$ of 30 subjects and 2318 voxels. Each fMRI voxel is sampled over 80 time points within a trial. The voxels of the DTI orientation data are represented by the 3-D vector signifying the primary orientation of the fiber tract at the voxel location.

Our NeuCube personalized architecture consists of three modules, as described in Section III. The first step of the

process was to compress or encode the dynamic fMRI data from continuous real space to binary spike space. For the encoding step, we have chosen to use the BSA [23] algorithm due to its ability to represent brain data as important spike events and have shown promising results in [25] and [47]. For the second step, subject-specific SNNc's were initialized with $K = 1000$ spiking neurons, and $N = 2318$ input neurons and synapses were initialized using the SWC algorithm within the radial neighborhood of $r_{\text{swc}} = 0.05$ and $w_{ij}(\text{init}) = 0.05$. The spiking neurons were set with the hyperparameters $\{v_{\text{rest}} = 0, v_{\text{thr}} = 0.1, \eta_{\text{thr}} = 8\}$. These along with other hyperparameters were chosen by a grid-based hyperparameter search strategy. The encoded fMRI data and the DTI orientation vector data for each subject are then passed through the initialized SNNc, generating O_{seq} for each subject using algorithm 1. In the final step, we learn a lazy KNN binary classification model using 50% of the randomly chosen subjects. It is to be noted that we use a custom distance function as part of KNN algorithm for learning binary spike data. We have used a custom spike asynchronicity-based distance function as described in [25] elaborately. The KNN can, of course, be replaced by any supervised learning module that can learn from binary input data.

Due to the multimodular and rather flexible nature of the NeuCube architecture, selecting baselines for comparison is a challenging task. In this paper, we have used the NeuCube architecture as a combination of temporal feature compressor (TFC), spatial expander (SE), and classifier. The compressor and the SNNc module together are used as a spatiotemporal feature extraction module. The supervised readout layer then learns a model from the transformed feature transformed data. Hence, it is appropriate to compare our BSA + oiSTDP feature extraction module against other feature extraction methods in continuous data domain. We have compared the following feature extraction algorithms.

1) *Sparse Autoencoder [48]:* Autoencoders are shallow single-hidden layer neural networks that can perform identity mapping of the input. The hidden layer of the autoencoder learns nonlinear lower dimensional data representations. The sparse autoencoders are an extension of regular autoencoders that impose sparsity regularization constraint on the loss function. In our experiments, we have used the fMRI data to learn a sparse autoencoder (with 1000 ReLU units in the hidden layer and L1 regularization constraint of 10^{-5}) that encodes the data into 1000 D feature space using the python keras API [49].
2) *Principal Component Analysis:* Principal component analysis (PCA) is a standard orthogonal linear feature transformation technique that transforms features into principal components. We have used scikit-learn API [50] to fit and transform the fMRI data to 1000 principal components.
3) *Independent Component Analysis:* ICA is another statistical feature transformation technique, which is used to decompose feature space to statistically independent component space by maximizing statistical independence of the estimated components. We have used scikit-learn [50] API's FastICA algorithm to fit and transform the fMRI data to 1000 independent components.
4) *Restricted Boltzman's Machine [51]:* Restricted Boltzman's machine (RBM) is an unsupervised nonlinear feature learner based on a probabilistic model that has gained much popularity in the deep neural network domain. We have used the scikit-learn API [50] to learn a Bernoulli RBM network with 1000 components using stochastic maximum likelihood [52] learning.

Table II presents the experimental results. The rows of Table II compare the methods for the classification task. (C) and (E) in the method names correspond to the custom and Euclidean distance function used in the KNN. The framework column specifies the role of each component in the method names. For example, the proposed BSA + oiSTDP + KNN is a combination of TFC, SE, and classifier (C). The performance of the binary classification task is measured by overall accuracy and Cohen's κ statistic. The first three rows of Table II compare the different NeuCube architectures. The BSA + oiSTDP + KNN is the proposed architecture for fMRI and DTI integrated learning. The next two methods systematically removes: 1) orientation influence from SNNc learning (STDP) and 2) the SNNc module to show the effect of inclusion of these artifacts on the performance. The best performance across the different methods is achieved by the proposed BSA + oiSTDP + KNN architecture with the overall accuracy of $72.4 \pm 12.3\%$ and Cohen's kappa of 0.44 ± 0.25. The classification accuracy increases by $\approx 8\%$ and doubles the mean Cohen's κ statistic when oiSTDP-based SNNc learning is performed in the middle using the fMRI and DTI data. Rows 4–7 are the nontemporal feature extraction baselines described earlier. Due to the nontemporal nature of the baseline feature compressors, the fMRI data for each subject are input to these feature extractors as a single vector (created by concatenating the temporal dimension) leading to a massive feature vector space. We have used KNN ($K = 1$) as a classifier in the classification module to keep the comparisons as fair as possible. The disadvantage of the large feature space is quite imperative as it leads to potential overfitting of the data. We have avoided adding the DTI data to the already large feature space to avoid further overfitting. As the SNNc of NeuCube is a spiking recurrent neural network framework with temporal or sequential learning capabilities, we have also learned the binary classification task with other single-hidden layer recurrent neural network framework, such as long short-term memory (LSTM) [53] and gated recurrent units (GRUs) [54]. Both LSTM and GRU networks were designed as shallow single-hidden layered neural networks having 50 LSTM and GRU units. These networks were implemented in keras API [49] and learned by optimizing binary crossentropy loss function using an adaptive momentum optimizer. The results for the baselines show that the KNN

[1]TFC = Temporal feature compressor, NTFC = No-temporal feature compressor, SE=Spatial expander, and C = classifier.

[2]The performance metrics are computed as mean ± standard deviation of ten independent train/test runs of the classification module.

TABLE II
COMPARISON OF THE PATTERN RECOGNITION METHODS ON THE BINARY CLASSIFICATION TASK

Method	Framework[1]	Data	Relation learning capabilities		Performance[2]	
			Temporal	Multi-modal	Accuracy(%)	Cohen's κ
BSA+oiSTDP+KNN(C)	TFC+SE+C	fMRI+DTI	✓	✓	$\mathbf{72.3 \pm 12.3}$	$\mathbf{0.44 \pm 0.25}$
BSA+STDP+KNN(C)	TFC+SE+C	fMRI	✓	✗	69.4 ± 13.9	0.38 ± 0.28
BSA+KNN(C)	TFC+C	fMRI	✗	✗	64.2 ± 12.4	0.22 ± 0.26
Sparse Autoencoder+KNN(E)	NTFC+C	fMRI	✗	✗	56.1 ± 7.2	0.01 ± 0.11
PCA+KNN(E)	NTFC+C	fMRI	✗	✗	56.1 ± 11.3	0.13 ± 0.18
ICA+KNN(E)	NTFC+C	fMRI	✗	✗	62.8 ± 12.3	0.26 ± 0.23
RBM+KNN(E)	NTFC+C	fMRI	✗	✗	36.2 ± 4.9	-0.23 ± 0.11
LSTM	C	fMRI	✓	✗	45.7 ± 9.6	-0.15 ± 0.14
GRU	C	fMRI	✓	✗	45.2 ± 7.5	-0.018 ± 0.13

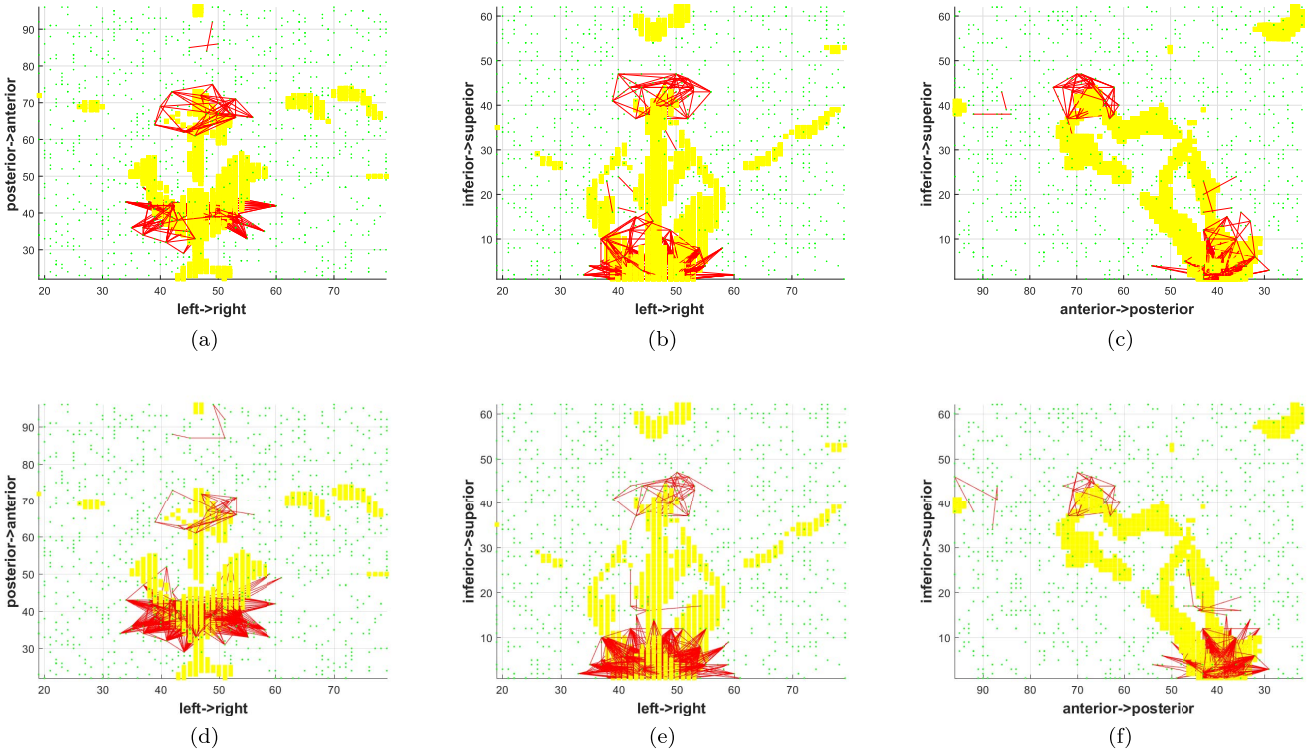


Fig. 13. Visual comparison of the strongest connections (mean weight across subjects within a group) formed in the SNNc of the TRS and the UTRS group. The top row shows the connections in the TRS group and the bottom row correspond to the UTRS group. The yellow colored cluster represents the input neurons and the green neurons are the computational spiking neurons. (a) and (d) Horizontal view. (b) and (e) Coronal view. (c) and (f) Sagittal view.

performs best on the ICA-based feature representation. PCA and sparse autoencoders are similar in accuracy, but PCA achieves stronger κ statistic. On the other hand, the baseline recurrent neural networks fail to learn the task leading to poor performance statistics.

Furthermore, we have individually scrutinized connection weights for the TRS and the UTRS group, generated by the oiSTDP learning algorithm. Fig. 13 shows a comparison of the strongest mean connection weights of the TRS and the UTRS group. Most of the strong connections are shown to be created in the lower cerebellum and thalamus across the UTRS and the TRS group. It has been shown that by connections via the thalamus, the cerebellum innervates with motor cortical, prefrontal, and parietal lobes [55]. Following cerebellar damage, neurocognitive symptoms and a cognitive affective syndrome, including blunted affect and inappropriate behavior, have been shown [56]. Recent fMRI and PET studies have also demonstrated the involvement of cerebellum and thalamus in sensory discrimination [57], attention [58], and complex problem solving [59]. All of which may be impaired in people with schizophrenia [60]. A recent study has also shown that the administration of CLZ in people with schizophrenia can restore cerebellar functions by altering the glutamatergic system and neuroplasticity [60]. This paper has additionally shown (Fig. 13) the presence of large density of strong connections in the cerebellum region of the model in the

UTRS group compared with the TRS group. Similarly, a large number of strong connections are present in the thalamic region of the TRS as opposed to UTRS. As the oiSTDP learning regulates the connection strength based on the joint information extracted from the spike activity and angular information, it can be stated that the distinctive connection density in the cerebellum and thalamus regions of the two groups in Fig. 13 is due to distinctive fMRI activity and DTI orientation information in the input data.

At this point, it is important to mention that the interpretation of connectivity in spatially mapped graph such as that shown here has significant conceptual resemblance to the area of brain network connectivity analysis, and especially dynamic causal modeling (DCM) in fMRI data [61], [62]. Numerous studies have been published over the past decade that has focused on such a connectivity analysis [63], [64]. DCM uses simple (deterministic) models of neural dynamics in a network of interacting nodes and models the change of neuronal state vectors in time, where brain regions are represented by hidden states, using bilinear differential equations. The objective of NeuCube SNNc is very much the same as DCM, i.e., to ask the question "what is the nature of the information that region A passes on to region B over time?" However, contrary to the causal models, the unit of information analyzed is not the raw data (fMRI or otherwise), but the data generated as a result of spike encoding and unsupervised learning. The spike encoding and learning transform the raw data to represent discriminatory patterns, which might be hidden in the noisy raw data.

V. Discussion and Future Directions

To the best of our knowledge, this is the first attempt to integrate multiple modalities of information in a spiking neural network architecture. The novelty of this approach lies in the proposed personalized SNNc-based architecture of NeuCube and, most importantly, the proposed oiSTDP learning algorithm, which can integrate multiple modalities of information including time, space, and orientation from data. Despite some assumptions being made on multimodal brain data, the proposed algorithm is not limited to brain data and can be extended to any data that has some form of spatial, temporal, and orientation information. Examples of such data are weather (change in temperature, wind movement, cloud movement, and so on) and traffic data.

The experiments shown here were conducted to demonstrate the ability of the algorithm to capture discriminative joint information present in the data and represent this information within its connection strengths. We have used the current design to incorporate DTI and fMRI from individuals initiating antipsychotic therapy to create a personalized classifier of treatment response in people with schizophrenia. Interrogation of the SNNc network revealed increased network connectivity in the cerebellar region of the model, potentially implicating activity in this area of the brain as a biomarker of treatment response in schizophrenia. Inclusion of more participants and studies using specific task-based designs may expose other markers not currently identified in the literature and provide novel hypotheses regarding why some individuals respond to clozapine monotherapy, while others do not. Additional applications of the algorithm may include other disorders where treatment or clinical outcome is poorly understood.

The ability to incorporate data from multiple imaging modalities simultaneously could increase reliability of the model to predict treatment outcomes in individuals. To date, studies have achieved high rates of accuracy in patient samples combining single imaging techniques alongside clinical and pharmacogenetic data [2], [3], though none of them have led to changes in clinical practice. This could potentially be a result of overfitting during training, which the algorithm proposed would be less likely to do. The proposed learning algorithm is formulated in a way that it favors the use of multimodal information so that additional data modalities can be included at a later stage, which is supported by the standard brain template used to design the SNNc.

In the future, apart from algorithmic refinement to further improve the performance, we also aim to include EEG data as part of the integrated brain data modeling using the NeuCube personalized SNNc approach. We expect further improvement of the classification performance through the addition of EEG data to the model. This strand of work will lead to new methods for the integration of multimodal data with heterogeneous spatial and temporal resolutions and scales.

References

[1] J. Sui, T. Adalı, Q. Yu, J. Chen, and V. D. Calhoun, "A review of multivariate methods for multimodal fusion of brain imaging data," *J. Neurosci. Methods*, vol. 204, no. 1, pp. 68–81, 2012.

[2] M. J. Patel, C. Andreescu, J. C. Price, K. L. Edelman, C. F. Reynolds, III, and H. J. Aizenstein, "Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction," *Int. J. Geriatric Psychiatry*, vol. 30, no. 10, pp. 1056–1067, 2015.

[3] A. Khodayari-Rostamabad, J. P. Reilly, G. M. Hasey, H. de Bruin, and D. J. MacCrimmon, "A machine learning approach using EEG data to predict response to ssri treatment for major depressive disorder," *Clin. Neurophysiol.*, vol. 124, no. 10, pp. 1975–1985, 2013.

[4] A. Khodayari-Rostamabad, G. M. Hasey, D. J. MacCrimmon, J. P. Reilly, and H. de Bruin, "A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy," *Clin. Neurophysiol.*, vol. 121, no. 12, pp. 1998–2006, 2010.

[5] C.-C. Lin *et al.*, "Artificial neural network prediction of clozapine response with combined pharmacogenetic and clinical data," *Comput. Methods Programs Biomed.*, vol. 91, no. 2, pp. 91–99, 2008.

[6] O. Doehrmann *et al.*, "Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging," *JAMA Psychiatry*, vol. 70, no. 1, pp. 87–97, 2013.

[7] *Medical Image Computing*. Accessed: Jan. 31, 2018. [Online]. Available: https://en.wikipedia.org/wiki/Medical_image_computing

[8] M. D. Greicius, K. Supekar, V. Menon, and R. F. Dougherty, "Resting-state functional connectivity reflects structural connectivity in the default mode network," *Cerebral Cortex*, vol. 19, no. 1, pp. 72–78, 2009.

[9] K. E. Stephan, K. J. Friston, and C. D. Frith, "Dysconnection in schizophrenia: From abnormal synaptic plasticity to failures of self-monitoring," *Schizophrenia Bull.*, vol. 35, no. 3, pp. 509–527, 2009.

[10] B. Horwitz and D. Poeppel, "How can EEG/MEG and fMRI/PET data be combined?" *Hum. Brain Mapping*, vol. 17, no. 1, pp. 1–3, 2002.

[11] P. A. Valdes-Sosa *et al.*, "Model driven EEG/fMRI fusion of brain oscillations," *Hum. Brain Mapping*, vol. 30, no. 9, pp. 2701–2721, 2009.

[12] S. M. Plis *et al.*, "Effective connectivity analysis of fMRI and MEG data collected under identical paradigms," *Comput. Biol. Med.*, vol. 41, no. 12, pp. 1156–1165, 2011.

[13] P. Stæmpfli, C. Reischauer, T. Jaermann, A. Valavanis, S. Kollias, and P. Boesiger, "Combining fMRI and DTI: A framework for exploring the limits of fMRI-guided DTI fiber tracking and for verifying DTI-based fiber tractography results," *NeuroImage*, vol. 39, no. 1, pp. 119–126, 2008.

[14] R. Kleiser, P. Staempfli, A. Valavanis, P. Boesiger, and S. Kollias, "Impact of fMRI-guided advanced DTI fiber tracking techniques on their clinical applications in patients with brain tumors," *Neuroradiology*, vol. 52, no. 1, p. 37, 2010.

[15] H. Yang, J. Liu, J. Sui, G. Pearlson, and V. D. Calhoun, "A hybrid machine learning method for fusing fMRI and genetic data: Combining both improves classification of schizophrenia," *Frontiers Hum. Neurosci.*, vol. 4, no. 192, p. 3389, 2010.

[16] V. D. Calhoun, J. Liu, and T. Adalı, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *NeuroImage*, vol. 45, no. 1, pp. S163–S172, 2009.

[17] S. J. Teipel *et al.*, "oller, and H. Hampel, "White matter microstructure underlying default mode network connectivity in the human brain," *NeuroImage*, vol. 49, no. 3, pp. 2021–2032, 2010.

[18] N. M. Correa, Y. O. Li, T. Adali, and V. D. Calhoun, "Canonical correlation analysis for feature-based fusion of biomedical imaging modalities and its application to detection of associative networks in schizophrenia," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 6, pp. 998–1007, Dec. 2008.

[19] N. M. Correa, T. Eichele, T. Adalı, Y.-O. Li, and V. D. Calhoun, "Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI," *NeuroImage*, vol. 50, no. 4, pp. 1438–1445, 2010.

[20] N. Correa, Y.-O. Li, T. Adalı, and V. D. Calhoun, "Examining associations between fMRI and EEG data using canonical correlation analysis," in *Proc. 5th IEEE Int. Symp. Biomed. Imag., Nano Macro (ISBI)*, May 2008, pp. 1251–1254.

[21] K. Chen *et al.*, "Linking functional and structural brain images with multivariate network analyses: A novel application of the partial least square method," *NeuroImage*, vol. 47, no. 2, pp. 602–610, 2009.

[22] N. K. Kasabov, "NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data," *Neural Netw.*, vol. 52, pp. 62–76, Apr. 2014.

[23] B. Schrauwen and J. Van Campenhout, "BSA, a fast and accurate spike train encoding scheme," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, vol. 4. Jul. 2003, pp. 2825–2830.

[24] N. Sengupta, N. Scott, and N. Kasabov, "Framework for knowledge driven optimisation based data encoding for brain data modelling using spiking neural network architecture," in *Proc. 5th Int. Conf. Fuzzy Neuro Comput. (FANCCO)*, 2015, pp. 109–118.

[25] N. Sengupta and N. Kasabov, "Spike-time encoding as a data compression technique for pattern recognition of temporal data," *Inf. Sci.*, vols. 406–407, pp. 133–145, Sep. 2017.

[26] S. Song, K. D. Miller, and L. F. Abbott, "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity," *Nature Neurosci.*, vol. 3, no. 9, pp. 919–926, 2000.

[27] N. Kasabov, K. Dhoble, N. Nuntalid, and G. Indiveri, "Dynamic evolving spiking neural networks for on-line spatio- and spectro-temporal pattern recognition," *Neural Netw.*, vol. 41, pp. 188–201, May 2013.

[28] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, 2009.

[29] N. K. Kasabov, M. G. Doborjeh, and Z. G. Doborjeh, "Mapping, learning, visualization, classification, and understanding of fMRI data in the NeuCube evolving spatiotemporal data machine of spiking neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 4, pp. 887–899, Apr. 2017.

[30] N. Kasabov, L. Zhou, M. G. Doborjeh, Z. G. Doborjeh, and J. Yang, "New algorithms for encoding, learning and classification of fMRI data in a spiking neural network architecture: A case on modeling and understanding of dynamic cognitive processes," *IEEE Trans. Cogn. Develop. Syst.*, vol. 9, no. 4, pp. 293–303, Dec. 2017.

[31] E. Tu, N. Kasabov, and J. Yang, "Mapping temporal variables into the NeuCube for improved pattern recognition, predictive modeling, and understanding of stream data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 6, pp. 1305–1317, Jun. 2017.

[32] N. Kasabov *et al.*, "Evolving spiking neural networks for personalised modelling, classification and prediction of spatio-temporal patterns with a case study on stroke," *Neurocomputing*, vol. 134, pp. 269–279, Jun. 2014.

[33] E. Tu *et al.*, "NeuCube$^{(ST)}$ for spatio-temporal data predictive modelling with a case study on ecological data," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2014, pp. 638–645.

[34] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Approach*. Hoboken, NJ, USA: Wiley, 1949.

[35] W. Gerstner, R. Kempter, J. L. van Hemmen, and H. Wagner, "A neuronal learning rule for sub-millisecond temporal coding," *Nature*, vol. 383, no. 6595, pp. 76–78, 1996.

[36] J. Sjöström and W. Gerstner, *Spike-Timing Dependent Plasticity*. 2010, p. 35, doi: 10.4249/scholarpedia.1362.

[37] J. Sjöström and W. Gerstner, "Spike-timing dependent plasticity," *Scholarpedia*, vol. 5, no. 2, p. 1362, 2010.

[38] E. van Aart, N. Sepasian, A. Jalba, and A. Vilanova, "CUDA-accelerated geodesic ray-tracing for fiber tracking," *J. Biomed. Imag.*, vol. 2011, Jan. 2011, Art. no. 6, doi: 10.1155/2011/698908.

[39] T. E. J. Behrens *et al.*, "Characterization and propagation of uncertainty in diffusion-weighted MR imaging," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 1077–1088, 2003.

[40] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[41] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *NeuroImage*, vol. 62, no. 2, pp. 782–790, 2012.

[42] V. Fonov *et al.*, "Unbiased average age-appropriate atlases for pediatric studies," *NeuroImage*, vol. 54, no. 1, pp. 313–327, 2011.

[43] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry, and D. L. Collins, "Unbiased average age-appropriate atlases for pediatric studies," *NeuroImage*, vol. 54, no. 1, pp. 313–327, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811910010062, doi: https://doi.org/10.1016/j.neuroimage.2010.07.033

[44] R. H. R. Pruim, M. Mennes, J. K. Buitelaar, and C. F. Beckmann, "Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI," *NeuroImage*, vol. 112, pp. 278–287, May 2015.

[45] R. H. R. Pruim, M. Mennes, D. van Rooij, A. Llera, J. K. Buitelaar, and C. F. Beckmann, "ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data," *NeuroImage*, vol. 112, pp. 267–277, May 2015.

[46] N. Kasabov *et al.*, "Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: Design methodology and selected applications," *Neural Netw.*, vol. 78, pp. 1–14, Jun. 2016.

[47] N. Nuntalid, K. Dhoble, and N. Kasabov, "EEG classification with BSA spike encoding algorithm and evolving probabilistic spiking neural network," in *Neural Information Processing*, B.-L. Lu, L. Zhang, and J. Kwok, Eds. Berlin, Germany: Springer, 2011, pp. 451–460.

[48] A. Ng, "Sparse autoencoder," *CS294A Lect. Notes*, vol. 72, pp. 1–19, 2011.

[49] F. Chollet *et al.* (2015). *Keras*. [Online]. Available: https://github.com/fchollet/keras

[50] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.

[51] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[52] T. Tieleman, "Training restricted boltzmann machines using approximations to the likelihood gradient," in *Proc. 25th Int. Conf. Mach. Learn.*, 2008, pp. 1064–1071.

[53] S. Hochreiter and J. Schmidhuber, "LSTM can solve hard long time lag problems," in *Proc. Adv. Neural Inf. Process. Syst.*, 1997, pp. 473–479.

[54] K. Cho *et al.* (2014). "Learning phrase representations using RNN encoder-decoder for statistical machine translation." [Online]. Available: https://arxiv.org/abs/1406.1078

[55] F. A. Middleton and P. L. Strick, "Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function," *Science*, vol. 266, no. 5184, pp. 458–461, 1994. [Online]. Available: http://science.sciencemag.org/content/266/5184/458, doi: 10.1126/science.7939688.

[56] H. Baillieux, W. Verslegers, P. Paquier, P. P. De Deyn, and P. Mariën, "Cerebellar cognitive affective syndrome associated with topiramate," *Clin. Neurol. Neurosurg.*, vol. 110, no. 5, pp. 496–499, 2008.

[57] J.-H. Gao, L. M. Parsons, J. M. Bower, J. Xiong, J. Li, and P. T. Fox, "Cerebellum implicated in sensory acquisition and discrimination rather than motor control," *Science*, vol. 272, no. 5261, pp. 545–547, 1996.

[58] E. Courchesne, N. A. Akshoomoff, J. Townsend, and O. Saitoh, "A model system for the study of attention and the cerebellum: Infantile autism," *Electroencephalogr. Clin. Neurophysiol. Suppl.*, vol. 44, pp. 315–325, Jan. 1995.
[59] S.-G. Kim, K. Uğurbil, and P. L. Strick, "Activation of a cerebellar output nucleus during cognitive processing," *Science*, vol. 265, no. 5174, pp. 949–951, 1994.
[60] P. Yeganeh-Doost, O. Gruber, P. Falkai, and A. Schmitt, "The role of the cerebellum in schizophrenia: From cognition to molecular pathways," *Clinics*, vol. 66, pp. 71–77, Mar. 2011.
[61] A. C. Marreiros, K. E. Stephan, and K. J. Friston, "Dynamic causal modeling," *Scholarpedia*, vol. 5, no. 7, p. 9568, 2010.
[62] K. J. Friston *et al.*, "Dynamic causal modelling revisited," *NeuroImage*, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811917301568, doi: https://doi.org/10.1016/j.neuroimage.2017.02.045
[63] K. J. Friston, J. Kahan, B. Biswal, and A. Razi, "A DCM for resting state fMRI," *NeuroImage*, vol. 94, pp. 396–407, Jul. 2014.
[64] J. Heinzle and K. E. Stephan, "Dynamic causal modeling and its application to psychiatric disorders," in *Computational Psychiatry*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 117–144.

Neelava Sengupta received the B.Tech. degree from the Maulana Abul Kalam Azad University of Technology, Kolkata, India, in 2009. He is currently pursuing the Ph.D. degree with the Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Auckland, New Zealand.

He was an Engineer with the Centre for Development of Advanced Computing, Pune, India, and a Research Assistant with the University of Hildesheim, Hildesheim, Germany. He is involved in the development of novel algorithms for integrated data modeling in spiking neural network architectures.

Carolyn B. McNabb received the Ph.D. degree from The University of Auckland, Auckland, New Zealand, with a focus on designing and validating a predictive model of clozapine response using structural and functional MRI data from individuals with treatment-resistant schizophrenia, under the supervision of Associate Professor B. Russell and Prof. R. Kydd.

She is currently a Post-Doctoral Research Fellow with the School of Psychology and Clinical Language Sciences, University of Reading, Reading, U.K.

Nikola Kasabov (F'10) received the M.Sc. degree in electrical engineering with specialization in computer science and the Ph.D. degree in mathematical sciences from Technical University, Sofia, Bulgaria, in 1971 and 1975, respectively.

He is currently the Director of the Knowledge Engineering and Discovery Research Institute and the Personal Chair of Knowledge Engineering, Auckland University of Technology, Auckland, New Zealand. He has published over 550 works in the areas of intelligent systems, neural networks, connectionist and hybrid connectionist systems, fuzzy systems, expert systems, bioinformatics, and neuroinformatics.

Dr. Kasabov is a fellow of the Royal Society of New Zealand and a Distinguished Visiting Fellow of the Royal Academy of Engineering UK, and a member of the International Neural Network Society (INNS) and Asia–Pacific Neural Network Assembly (APNNA) Governing Boards. He is a past President of the INNS and the APNNA.

Bruce R. Russell is currently an Associate Professor with the School of Pharmacy, University of Otago, Dunedin, New Zealand. His research is focused on the psychopharmacology surrounding severe mental illnesses, such as schizophrenia and drug addiction, where he is also seeking biomarkers to enable the prediction of a response to drug treatment.