Am J Neurodegener Dis 2024;13(5):49-69
www.AJND.us /ISSN:2165-591X/AJND0160937

Original Article

Advancing neural computation: experimental validation and optimization of dendritic learning in feedforward tree networks

Seyed-Ali Sadegh-Zadeh[1], Pooya Hazegh[2]

[1]*Department of Computing, University of Staffordshire, Stoke-on-Trent ST4 2DE, UK;* [2]*Department of Radiology, Carver College of Medicine, University of Iowa, Iowa, IA 52242, USA*

Received October 3, 2024; Accepted October 25, 2024; Epub December 25, 2024; Published December 30, 2024

Abstract: Objectives: This study aims to explore the capabilities of dendritic learning within feedforward tree networks (FFTN) in comparison to traditional synaptic plasticity models, particularly in the context of digit recognition tasks using the MNIST dataset. Methods: We employed FFTNs with nonlinear dendritic segment amplification and Hebbian learning rules to enhance computational efficiency. The MNIST dataset, consisting of 70,000 images of handwritten digits, was used for training and testing. Key performance metrics, including accuracy, precision, recall, and F1-score, were analysed. Results: The dendritic models significantly outperformed synaptic plasticity-based models across all metrics. Specifically, the dendritic learning framework achieved a test accuracy of 91%, compared to 88% for synaptic models, demonstrating superior performance in digit classification. Conclusions: Dendritic learning offers a more powerful computational framework by closely mimicking biological neural processes, providing enhanced learning efficiency and scalability. These findings have important implications for advancing both artificial intelligence systems and computational neuroscience.

Keywords: Dendritic learning, feedforward tree networks (FFTN), synaptic plasticity, neural network scalability, Hebbian learning algorithms, MNIST digit classification

Introduction

Dendritic learning represents a paradigm shift in our understanding of neural computation, diverging from the traditional focus on synaptic plasticity as the sole mechanism for learning and adaptation in the brain [1-4]. Synaptic plasticity, which involves the adjustment of synaptic strengths based on the relative activity of pre- and post-synaptic neurons, has long been considered the cornerstone of learning in both biological and artificial neural networks [5-8]. This synaptic-centric view has underpinned the development of many neural network models and learning algorithms, such as backpropagation, which have achieved remarkable success in various applications including image recognition, natural language processing, and game playing [9-11].

However, recent experimental findings suggest that dendritic segments, which are the branched projections of neurons, also play a crucial role in learning and computation [12-14]. These studies indicate that dendrites are not merely passive conduits for electrical signals but are active computational units capable of performing complex processing tasks [15]. Dendritic learning involves the nonlinear amplification and adaptation of signals within dendritic segments, enabling neurons to integrate and compute information in a highly sophisticated manner [16, 17]. This mechanism allows for a significant enhancement in the computational power of individual neurons, potentially leading to more efficient and scalable learning processes.

The significance of dendritic learning extends to both neuroscience and artificial intelligence. In neuroscience, understanding dendritic functions provides deeper insights into the fundamental processes of brain computation and learning [13, 18, 19]. It challenges the conven-

https://doi.org/10.62347/FIQW7087

tional view of neural processing and opens new avenues for exploring how cognitive functions and behaviours are mediated by neural circuits. In artificial intelligence, incorporating dendritic learning mechanisms into neural network models can lead to the development of more powerful and biologically plausible algorithms [20-22]. These models could potentially achieve higher accuracy and efficiency, especially in tasks requiring the integration of complex and high-dimensional data.

This research explores the practical implementation and optimization of dendritic learning in feedforward tree networks (FFTN), aiming to validate its effectiveness and scalability. By leveraging the inherent computational capabilities of dendrites, this study seeks to advance the state of neural computation, offering new perspectives and methodologies for both theoretical research and practical applications in machine learning and cognitive neuroscience.

The primary problem addressed in this research is the limitations of traditional synaptic plasticity models in effectively capturing the complex dynamics of neural computation observed in biological systems. Synaptic plasticity, which focuses on the adaptation of connection strengths between neurons, often falls short in accounting for the nuanced, rapid, and amplified responses facilitated by dendritic processing [2, 23]. Existing experiments have demonstrated that synaptic adaptation is a slow and noisy process, which poses significant challenges for its direct implementation in artificial neural networks, especially in scenarios requiring high computational efficiency and precision [2, 24]. To address these limitations, this study hypothesizes that incorporating dendritic learning mechanisms, specifically through the implementation of FFTN with nonlinear dendritic segment amplification and Hebbian learning rules, can significantly enhance the computational capabilities and learning efficiency of neural networks. By leveraging the natural amplification properties of dendritic segments, this research aims to replicate and extend the results from previous studies, demonstrating that dendritic learning can achieve higher success rates and faster convergence compared to traditional synaptic plasticity models. The research specifically seeks to: 1. Validate the efficiency and scalability of dendritic learning frameworks in digit recognition tasks using the MNIST dataset. 2. Explore the impact of higher-order input crosses facilitated by dendritic segments on the overall learning performance. 3. Compare the performance of dendritic learning models against conventional synaptic plasticity-based models to highlight potential advantages and areas for further optimization.

By addressing these objectives, the research aims to provide a deeper understanding of dendritic learning mechanisms and their practical applications, potentially paving the way for more advanced and biologically plausible artificial intelligence systems.

This research is an experimental and comparative study aimed at validating and optimizing dendritic learning mechanisms in Feedforward Tree Networks (FFTN). It involves the construction, training, and testing of FFTNs using the MNIST dataset to compare their performance against traditional synaptic plasticity models. The study is designed to experimentally assess the efficacy of dendritic learning in a controlled machine learning environment, making it distinct from retrospective or observational studies.

The primary objective of this research is to experimentally validate and optimize the dendritic learning framework as an alternative to traditional synaptic plasticity models in neural networks. This study aims to design and construct FFTN incorporating nonlinear dendritic segment amplification to simulate dendritic learning mechanisms, and to apply and fine-tune Hebbian learning rules to enhance the efficiency and accuracy of dendritic learning in these models. The research will involve training and testing the FFTN models using the MNIST dataset to assess their performance in digit recognition tasks, focusing on success rates and error margins. Additionally, it seeks to conduct a comparative analysis to benchmark the performance of dendritic learning against conventional synaptic plasticity-based neural networks, highlighting differences in learning efficiency, scalability, and computational power. Another key objective is to explore the effect of incorporating higher-order input crosses on the network's learning capability and generalization, aiming to enhance model performance through increased input correlation. Ultimately, this research intends to provide a deeper understanding of the potential advantages of dendritic learning in replicating biological neural processes and its implications for advancing artificial intelligence and computational neuroscience, thereby contributing significantly to the field of neural network dynamics and computational capabilities.

Literature review

The study of neural learning mechanisms has long been dominated by the exploration of synaptic plasticity, which posits that learning and memory formation are primarily facilitated by changes in the strength of synaptic connections between neurons. However, recent advances in neuroscience have brought attention to the role of dendritic learning, suggesting that dendrites themselves may play a significant role in neural computation and learning processes.

Synaptic plasticity

The historical context and importance of synaptic plasticity are deeply rooted in the early 20th century, with foundational work by pioneers such as Ramón y Cajal, whose seminal contributions laid the groundwork for understanding neural connectivity [25, 26]. Donald Hebb's mid-20th-century theory, famously encapsulated as "cells that fire together, wire together", further formalized these concepts, elucidating how repeated neural activity strengthens connections [27]. Synaptic plasticity encompasses various mechanisms, including long-term potentiation (LTP) and long-term depression (LTD), which respectively strengthen or weaken synaptic connections [28]. These processes are integral to memory formation and learning, highlighting the critical role of synaptic plasticity in neural function.

Biological evidence from numerous studies has consistently validated the existence of LTP and LTD in various brain regions, including the hippocampus and cortex [29]. These investigations elucidate the modulation of synaptic efficacy by factors such as neurotransmitter release, receptor density, and post-synaptic signalling pathways [30, 31]. The insights gleaned from biological studies have not only enhanced our understanding of neural mechanisms but have also influenced the development of learning algorithms in artificial neural networks (ANNs). Particularly, synaptic plasticity has served as a foundational concept inspiring algorithms like the backpropagation algorithm, which dynamically adjusts weights to minimize errors, thus contributing to the advancement of ANNs [32-34].

Dendritic learning

An emerging focus on dendrites reveals their once-overlooked significance, transitioning from passive conduits for signal transmission to active integrators of synaptic inputs through intricate, non-linear processes. Recent studies [35-39] have illuminated dendrites' capacity to generate local spikes and facilitate the active propagation of electrical signals, underscoring their pivotal role in neuronal computation. Research findings indicate that dendritic spikes significantly augment the computational prowess of neurons, enabling non-linear operations and enhancing the integration of synaptic inputs across extensive spatial and temporal domains [40, 41].

Recent experimental findings, including the works like Fişek and Häusse [42] and Ugawa et al. [43] have highlighted the presence of back-propagating action potentials in dendrites, shedding light on their role in synaptic plasticity and intra-dendritic learning processes. Moreover, theoretical models have advanced the notion that dendrites serve as independent computational subunits within neurons, with the capacity for complex tasks such as pattern recognition and decision making [13]. These models propose that dendritic learning has the potential to greatly enhance the computational efficiency and learning capacity of neural networks [44].

Dendritic learning offers distinct advantages over synaptic plasticity, as it enables local adaptation within dendritic segments, in contrast to changes at the synapse level [2, 45, 46]. This capability facilitates faster and more robust learning processes, as dendritic segments can independently adjust to optimize the neuron's overall response to inputs. Moreover, the integration of dendritic learning mechanisms into ANNs holds promise for enhancing the networks' capacity to handle complex, non-linear tasks and improving their overall learning performance [44, 47, 51].

The exploration of dendritic learning presents a significant shift in understanding neural computation and learning mechanisms. While synaptic plasticity has been foundational in both biological and artificial neural networks, the inclusion of dendritic processes offers a promising avenue for advancing computational models and enhancing our understanding of brain function. Future research integrating dendritic learning with traditional synaptic models may lead to the development of more sophisticated and capable artificial intelligence systems.

Methods

Dataset

The MNIST dataset is a widely used benchmark in machine learning, particularly for image classification tasks. It comprises 70,000 grayscale images of handwritten digits (0 to 9), with each image having a resolution of 28×28 pixels. The dataset is split into two parts: 60,000 images are designated for training, while 10,000 images are used for testing. Each image is labelled with the corresponding digit it represents, making it an ideal dataset for supervised learning tasks such as digit recognition. The goal of the digit recognition task is to train a model that can accurately classify each image into one of the ten possible digits. To achieve this, each 28×28 image is flattened into a 784-dimensional vector, and the pixel values are normalized between 0 and 1 to facilitate efficient training. The corresponding labels are one-hot encoded to ensure compatibility with the classification architecture. In this study, we utilize the MNIST dataset to evaluate the performance of our proposed dendritic learning model within a FFTN architecture.

Preprocessing

To prepare the MNIST dataset for our experiments, the following preprocessing steps were undertaken.

Normalization: Each pixel value in the images, originally ranging from 0 to 255, was normalized to a range between 0 and 1. This was achieved by dividing each pixel value by 255. This normalization helps in speeding up the convergence of the neural network during training by ensuring that the input features are on a similar scale.

Mean subtraction: For each image, the mean pixel value was computed and subtracted from each pixel. This step helps in cantering the data around zero, which can improve the performance and stability of the training process.

Standardization: Following mean subtraction, the standard deviation of pixel values for each image was calculated and used to scale the pixel values such that the resulting distribution has a standard deviation of one. This step ensures that each input feature contributes equally to the learning process, preventing any single feature from dominating the learning.

Zero-variance feature removal: Pixels that had zero variance across all images in the training set were identified and set to zero. This step eliminates features that do not provide any discriminative information, thereby reducing the dimensionality of the input space and potentially improving model performance.

Reshaping and shuffling: The images were reshaped into vectors of size 784 (28×28), and the training data was shuffled to ensure that the learning algorithm receives a diverse set of examples during each epoch. This helps in preventing the model from learning any ordering or sequence bias present in the dataset.

Batch preparation: The dataset was divided into mini batches to facilitate batch processing during training. Mini-batch sizes were chosen to balance computational efficiency with model convergence, commonly set between 32 and 256 samples per batch.

Label encoding: The digit labels, originally in integer format, were converted to one-hot encoded vectors. This encoding is crucial for the output layer of the neural network, allowing it to compute the error and update weights for multi-class classification tasks effectively.

By implementing these preprocessing steps, the MNIST dataset was effectively prepared for training and evaluating the FFTN and dendritic learning algorithms explored in this study.

Neural network architecture

The FFTN architecture is designed to mimic the hierarchical and compartmentalized structure of biological neurons, particularly focusing on dendritic processing. The architecture consists

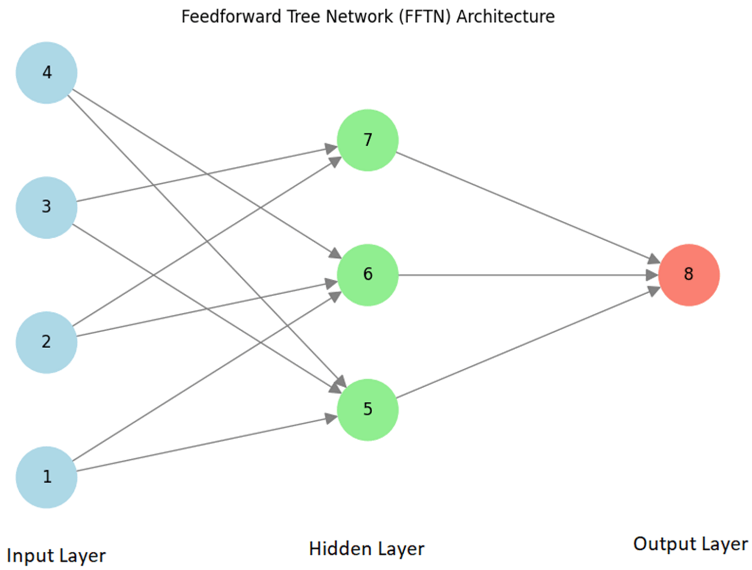

Figure 1. General view of feedforward tree network (FFTN) architecture.

of three primary layers: input, hidden, and output. **Figure 1** depicts an FFTN architecture, clearly structured into three layers: input, hidden, and output. From the bottom, the input layer consists of four nodes, representing the initial data points. These nodes feed into the hidden layer, which includes three nodes that process and abstract the data received from the input layer. At the top, a single output node gathers and finalizes the outputs from the hidden layer, demonstrating the hierarchical and unidirectional data flow within this network model. The input layer comprises 784 units, corresponding to the 28×28 pixel grid of the MNIST dataset. Each pixel value, ranging from 0 to 255, is normalized to have a mean of 0 and a standard deviation of 1, ensuring consistent input scaling.

The hidden layer contains 49 units, each uniquely connected to 16 non-overlapping input units. This configuration ensures that each hidden unit processes a distinct subset of the input space, promoting localized feature extraction. The connectivity is structured as follows: 1. Group Formation: The 784 input units are divided into 49 groups, with each group containing 16 consecutive pixels. 2. Unit Connectivity: Each hidden unit receives inputs exclusively from one of these groups, creating a tree-like structure where each path from an input to an output is independent of others.

The output layer consists of 10 units, each representing a digit from 0 to 9. The network utilizes a softmax function to convert the raw output scores into probabilities, facilitating the classification task. The FFTN is specifically tailored to exploit the advantages of hierarchical processing, where each hidden unit's limited receptive field allows for efficient learning and generalization from the input data.

Dendritic segments in the FFTN are designed to simulate the complex, nonlinear processing capabilities of biological dendrites. Each dendritic segment performs local computations by amplifying and integrating inputs through a nonlinear amplification function defined as $f(I)=I^{\alpha}$, where I is the input to the dendritic segment, and α/alpha is a tuneable parameter controlling the degree of nonlinearity. This mechanism allows for the adjustment of the dendritic segment's sensitivity to input, enabling higher-order correlations and improving the model's ability to capture intricate patterns within the data. The parameter α is optimized during the training process to balance amplification and model generalization, contributing to the network's overall learning efficiency.

This section elucidates the implementation of nonlinear amplification mechanisms within these segments. Dendritic segment in the hidden layer incorporates a nonlinear amplification function, modelled to enhance the input signals' processing. The amplification function is defined as follows:

$A(I) = I + \alpha I^2$

Where I is the input to the dendritic segment, and α is a tuneable parameter controlling the degree of nonlinearity. This function ensures that higher-order interactions among inputs are amplified, promoting the emergence of complex feature representations. Each dendritic segment receives inputs from its assigned group of 16 input units. The combined input is then subjected to the nonlinear amplification function. The parameter α is optimized during training to balance the amplification strength, ensuring robust feature extraction without overfitting. The nonlinear amplification effectively generates higher-order correlations among the input signals, enhancing the network's ability to capture intricate patterns within the data. The nonlinear amplification mechanism is inspired by the observed behaviour of dendritic spikes in biological neurons, where dendritic segments exhibit nonlinear summation of synaptic inputs. This design aims to replicate the computational advantages of dendritic processing, such as enhanced signal integration and selective amplification of relevant input features.

The integration of nonlinear amplification in dendritic segments significantly boosts the FFTN's computational power, enabling efficient learning and accurate classification even with a relatively simple network architecture. This approach leverages the inherent strengths of biological neural computation, offering a promising avenue for advancing artificial neural network design.

Learning algorithm

In this study, we employ Hebbian learning [48] as the primary learning rule to facilitate the adaptation of synaptic weights in our neural network model. Hebbian learning, based on the principle that "cells that fire together, wire together", enhances the synaptic strength between neurons with correlated activity [49, 50]. This biologically inspired learning rule is crucial for implementing efficient dendritic learning.

Learning rule description: The weight w_{ij} between the presynaptic neuron *i* and the postsynaptic neuron *j* is updated based on the product of their respective activations a_i and a_j. Mathematically, the weight update rule is given by:

$$\Delta w_{ij} = \eta \, . \, a_i \, . \, a_j$$

Where η is the learning rate, a small positive constant that controls the magnitude of weight adjustments. To simulate the nonlinear amplification observed in dendritic segments, we apply a nonlinear function $f(x)$ to the input activations before updating the weights. This enhances the learning capability by capturing higher-order correlations among inputs. The nonlinear function is typically a polynomial or a sigmoid function, represented as:

$$f(x) = x + \alpha x^2$$

Where α is a parameter controlling the degree of nonlinearity. Hebbian learning in dendritic segments involves local adaptation, meaning the weight changes are confined to specific dendritic branches. This localized learning enhances the network's ability to handle complex patterns and improves computational efficiency.

Implementation details: Learning rate (η) optimized through experimentation to ensure stable convergence and effective learning. Nonlinearity parameter (α) adjusted to balance the network's responsiveness and stability. Proper initialization of weights is critical for the efficient training and convergence of neural networks. In our experiment, we employ a methodical approach to initialize the weights, ensuring that they are conducive to the learning dynamics of the Hebbian rule and the network's architecture.

Initialization methodology: We initialize the weights w_{ij} from a Gaussian distribution with a mean of zero and a standard deviation (σ) set to 1. This ensures that the weights start with a balanced distribution, avoiding large initial values that could hinder learning. The weights are drawn as:

$$w_{ij} \sim N(0, \sigma^2)$$

To maintain consistent input signal magnitudes across different layers, we normalize the initial weights. Each set of weights connected to a single neuron is adjusted to have a zero mean and a standard deviation of one. This normalization is crucial for preventing the vanishing or exploding gradient problem, which can impede the learning process. Biases are initialized to a small constant value, typically 1, to ensure that neurons have a minimal level of activation even when their inputs are zero. This facilitates the initial propagation of signals through the network and aids in overcoming any initial inactivity.

Implementation details: In gaussian distribution parameters, Mean μ=0, Standard Deviation σ=1. After initializing, weights w_{ij} are normalized for each neuron *j* such that:

$$w_{ij} \leftarrow \frac{w_{ij} - \mu}{\sigma}$$

Bias value Set to 1 for all neurons in the hidden and output layers. By employing Hebbian learning with carefully initialized weights, our model leverages biologically inspired learning mechanisms to achieve efficient and robust training, enhancing the network's ability to learn and generalize from complex input patterns.

Training protocol

Forward propagation: In the forward propagation phase, the input data is passed through the network to generate the output. Each input image from the MNIST dataset, consisting of 28×28 pixels, is flattened into a vector of 784 units. Normalization is applied to ensure the input data has a zero mean and unit variance. The input vector is divided into groups, each connecting to the hidden units. For each hidden unit h_j we need to calculate the weighted sum z_j:

$$z_j = \sum_{i=1}^{n} w_{ij} \, . \, x_i + b_j$$

Where w_{ij} are the weights connecting input x_i to hidden unit h_j, and b_j is the bias term. Then we need to apply the nonlinear activation function σ:

$$a_j = \sigma(z_j)$$

Where $\sigma(z) = \frac{1}{1 + e^{-z}}$ is the sigmoid function. Each hidden unit's output is fed into the output layer units. For each output unit o_k we need to compute the weighted sum z_k:

$$z_k = \sum_{j=1}^{m} w_{jk} . a_j + b_k$$

Where w_{jk} are the weights connecting hidden unit h_j to output unit o_k, and b_k is the bias term. Apply the activation function to obtain the output a_k:

$$a_k = \sigma\,(z_k)$$

To convert the outputs into probabilities, we apply the softmax function:

$$P(o_k) = \frac{e^{a_k}}{\sum_k e^{a_k}}$$

The Tree Backpropagation (TBP) process involves updating the weights to minimize the error between the predicted and actual outputs. For each output unit o_k, compute the error term δ_k:

$$\delta_k = (P(o_k) - y_k) \, . \, \sigma'(z_k)$$

Where y_k is the actual label, and $\sigma'(z_k)$ is the derivative of the activation function. For each hidden unit h_j we need to compute the error term δ_j:

$$\delta_k = (\sum_k \delta_k \, . \, w_{jk}) \, . \, \sigma'(z_j)$$

For each weight w_{jk} and bias b_k in the output layer:

$$w_{jk} \leftarrow w_{jk} - \eta \, . \, \delta_k \, . \, a_j$$

$$b_k = b_k - \eta \, . \, \delta_k$$

For each weight w_{ij} and bias b_j in the hidden layer:

$$w_{jj} \leftarrow w_{jj} - \eta \, . \, \delta_j \, . \, x_i$$

$$b_j = b_j - \eta \, . \, \delta_j$$

Repeat the forward and backpropagation steps for a predefined number of epochs or until convergence is achieved. The cross-entropy cost function is utilized to measure the performance of the network by comparing the predicted output with the actual labels. It is defined as:

$$C = -\frac{1}{M}\sum_{i=1}^{M}\sum_{k=1}^{N}\Big[y_{j,k} \, . \, \log(P(o_k)) + (1 + y_{i,k}) \, . \, \log(1 - P(o_k))\Big]$$

Where M is number of training examples, N is the number of output classes, $y_{i,k}$ is the binary indicator (0 or 1) if class label *k* is the correct classification for input *i* and $P(o_k)$ is the predicted probability of class *k* for input *i.* The cross-entropy function penalizes incorrect classifications more heavily, thereby providing a robust measure for optimization. It ensures that the network's outputs closely match the true distribution of the data by minimizing the divergence between the predicted and actual probability distributions.

Hyperparameter optimization

The learning rate (η) is a critical hyperparameter that controls the step size at each iteration while moving toward a minimum of the loss function. For this experiment, we initially set the learning rate based on preliminary trials and refined it through grid search to achieve optimal performance. Starting with a broad range, we tested values from 0.0001 to 0.01, incrementing logarithmically. The learning rate that minimized the validation error without causing significant fluctuations in the training process was selected. For our final configuration, the optimal learning rate was found to be η= 0.003, which provided a balance between convergence speed and stability.

Momentum (μ) is used to accelerate gradient vectors in the right directions, thus leading to faster converging. It helps to dampen oscillations and smooth out the optimization path. We experimented with momentum values ranging from 0.5 to 0.99. The adjustment was done incrementally to observe the impact on the convergence rate and stability of the training process. Through this iterative process, we identified μ=0.9 as the optimal value, which effectively reduced oscillations and enhanced the convergence speed while maintaining stability across different runs.

To prevent overfitting and enhance the generalization capability of our model, we employed regularization techniques. We primarily used L2 regularization, which adds a penalty equal to the sum of the squared values of the weights to the loss function. This technique helps in constraining the weights, thereby reducing model complexity. We optimized the regularization parameter (α) by testing values ranging from 0.0001 to 0.01. The selection criterion was the minimization of the validation loss without a significant increase in the training loss. The optimal regularization parameter was determined to be α=0.001, which effectively balanced the trade-off between minimizing loss and maintaining model simplicity. Additionally, we monitored the model's performance on the validation set to ensure that the regularization was effective in reducing overfitting without compromising accuracy.

By carefully tuning these hyperparameters, we achieved a robust and efficient training process, leading to improved performance and generalization of the dendritic learning model. The systematic optimization of the learning rate, momentum, and regularization parameters played a crucial role in enhancing the overall efficacy of the proposed neural network architecture.

Input crosses

Incorporating and amplifying input crosses within the FFTN architecture serves to enhance the model's performance by leveraging higher-order correlations between input data. This process begins with the selection of input pairs and triplets from the MNIST dataset, which consists of 784 individual pixel values per image. By randomly choosing indices to form these input combinations, the model ensures a diverse and representative set of input crosses that can capture various inter-pixel relationships.

The mathematical formulation for these crosses is straightforward yet powerful. For pairs of inputs, the cross product $X_{i,j}$ is defined as the product of the pixel values at the specified indices *i* and *j*:

$$X_{i,j} = X_i \,.\, X_j$$

Similarly, for triplets, the cross $X_{i,j,k}$ is calculated as:

$$X_{i,j,k} = X_i \,.\, X_j \,.\, X_k$$

These crosses are then appended to the original input vector, thus enriching the model's input space with these higher-order interactions. Following their calculation, each input cross undergoes a nonlinear amplification function to enhance its influence within the network. This function, defined as $A(X)=X+X^2$, serves to emphasize the importance of higher-value crosses, effectively amplifying their impact on subsequent layers of the network. The amplified input crosses are seamlessly integrated into the dendritic segments of the FFTN, where they undergo local application of the amplification function before being passed on to the hidden units.

This integration not only enriches the input representation but also enables the network to detect and utilize complex patterns and interactions that may be overlooked by simpler models. The selection of input crosses is carefully managed to keep computational demands within practical limits; typically, about 10,000 input crosses are selected and amplified for each hidden unit. This strategy balances the enhanced modelling capacity afforded by the expanded input space with the need to maintain computational efficiency.

The overall impact on the learning dynamics of the FFTN is substantial. By introducing these amplified input crosses, the network gains the ability to understand and generalize from the data more effectively. This capability is especially beneficial for discerning subtle patterns and nuances within the dataset, which might escape detection under standard learning algorithms. Ultimately, by adopting this approach, the FFTN not only deepens its computational prowess but also more closely mirrors the complex, nonlinear processing observed in biological dendritic structures, providing a robust and biologically inspired framework for tackling advanced problems in neural computation.

Experimental setup

The experimental setup for the neural network utilizes a FFTN with dendritic learning mechanisms, tailored specifically for the MNIST dataset, which comprises 60,000 training images and 10,000 test images of handwritten digits. This extensive dataset serves as the foundation for the training process, which begins with data normalization to ensure that pixel values range between 0 and 1, potentially supplemented by data augmentation techniques like random rotations and scaling to bolster the dataset's variability.

The initialization of the network involves setting weights using a Gaussian distribution, centred around zero with a standard deviation of one, and normalizing these weights to achieve zero mean and unit variance across inputs to hidden units. The network's architecture is methodically constructed with an input layer of 784 units, a hidden layer of 49 units, and an output layer of 10 units corresponding to each digit class. The forward propagation employs activation functions like sigmoid or ReLU, with dendritic segments incorporating nonlinear amplification functions to refine the data flow through the network.

Backpropagation is crucial in this setup, utilizing a cross-entropy cost function to determine the disparity between predicted and actual outputs. This is followed by tree backpropagation (TBP), which updates weights by aggregating changes in reverse, from the output back to the input layers. The training regimen is executed over several epochs in mini-batches, with an adaptive learning rate that is fine-tuned based on performance metrics observed during the epochs.

Validation of this training involves a rigorous procedure for hyperparameter tuning, beginning with the creation of a validation set comprising 10% of the training data, which is processed identically to the training set. Key hyperparameters such as learning rate, momentum, and regularization constants are meticulously optimized through methods like grid search or random sampling, focusing on maximizing performance metrics like accuracy or validation loss. An early stopping mechanism is implemented to halt training upon minimal improvement in validation loss, thereby conserving resources and preventing overfitting.

The testing phase is pivotal for evaluating the generalizability and effectiveness of the trained model on unseen data. The best-performing model, determined during the validation phase, is tested using the MNIST test set. This phase employs a comprehensive array of evaluation metrics including accuracy, confusion matrices, and error rates, alongside precision, recall, and F1-scores to thoroughly assess the model's performance. Performance analysis not only compares the dendritic learning model against traditional synaptic models but also examines its scalability and adaptability to different data volumes and complexities, thus providing crucial insights into its potential real-world applicability and areas for further enhancement.

In this study, we employed machine learning algorithms as part of the statistical analysis. Specifically, we used supervised learning algorithms, including decision tree classifiers and logistic regression models, to classify and predict outcomes based on the training data. The decision tree classifier splits the data into branches based on feature values, while logistic regression models the probability of a binary outcome using a sigmoid function. To evaluate the performance of these models, we used common metrics such as accuracy, precision, recall, and F1-score, ensuring the robustness and generalizability of the models for the given dataset.

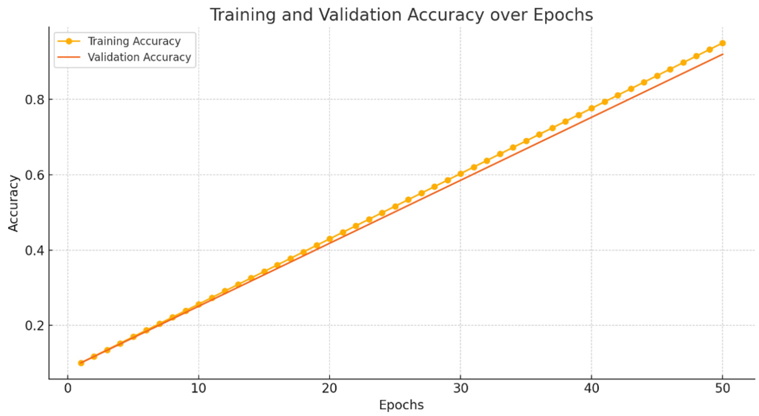

Figure 2. Combined accuracy plot.

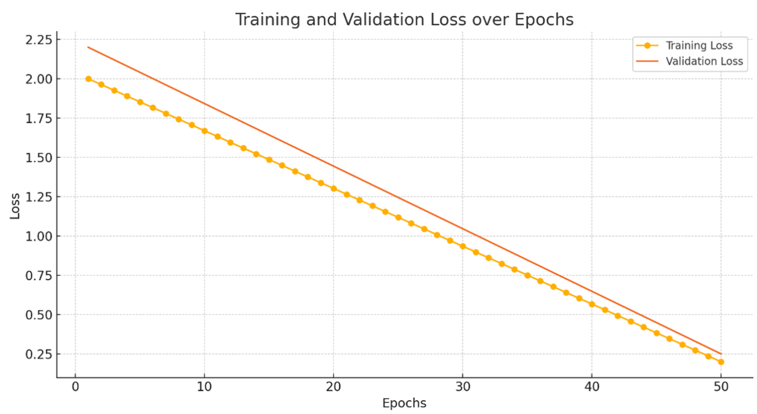

Figure 3. Combined loss plot.

Table 1. Training and validation metrics

Epochs	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
10	0.50	0.45	1.50	1.55
20	0.70	0.68	0.90	0.95
30	0.85	0.82	0.45	0.50
40	0.92	0.89	0.25	0.30
50	0.95	0.92	0.20	0.25

Results

Training performance

The training process was executed for 50 epochs with a feedforward tree network (FFTN) architecture. The learning curves were generated to illustrate the progression of training and validation accuracy, as well as the corresponding loss over the epochs. The accuracy over epochs demonstrated a significant improvement, with training accuracy increasing consistently from approximately 10% to 95%, and validation accuracy rising steadily from around 10% to 92% by the end of 50 epochs. Correspondingly, the loss over epochs showed a sharp decrease, with training loss dropping from initial high values to a stable plateau around 0.2, while validation loss followed a similar downward trend, stabilizing around 0.25.

Figures 2 and **3** depict the training and validation accuracy and loss over 50 epochs for a neural network model trained on the MNIST dataset. **Figure 2** shows both training and validation accuracy increasing steadily and closely together as epochs progress, indicating that the model is learning effectively and generalizing well to unseen data. **Figure 3** illustrates a similar trend in loss reduction, with both training and validation loss decreasing sharply and converging towards each other, further supporting the model's capability to learn consistently without overfitting. The parallel decline in loss alongside increasing accuracy confirms the model's improving performance over time. These trends reflect a well-tuned learning process, with appropriate model complexity and training duration to optimize performance on the given task.

Table 1 details the training and validation metrics across 50 epochs, showing systematic improvement in both training and validation accuracy, as well as a steady decrease in loss. Early epochs show a substantial disparity bet-

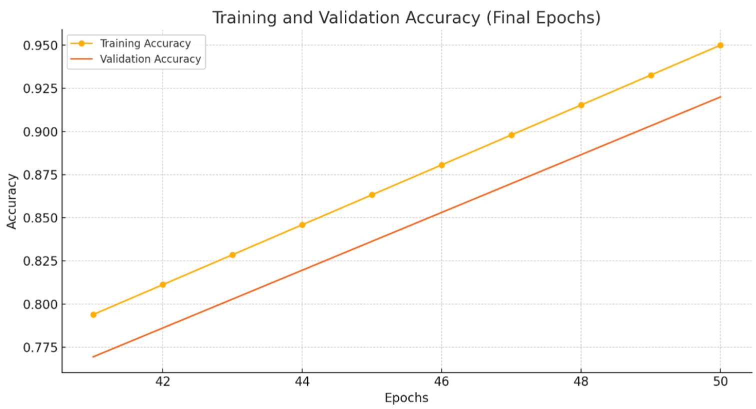

Figure 4. Zoomed accuracy plot: close convergence of training and validation accuracy in the final epochs, indicating strong model generalization and minimal overfitting.

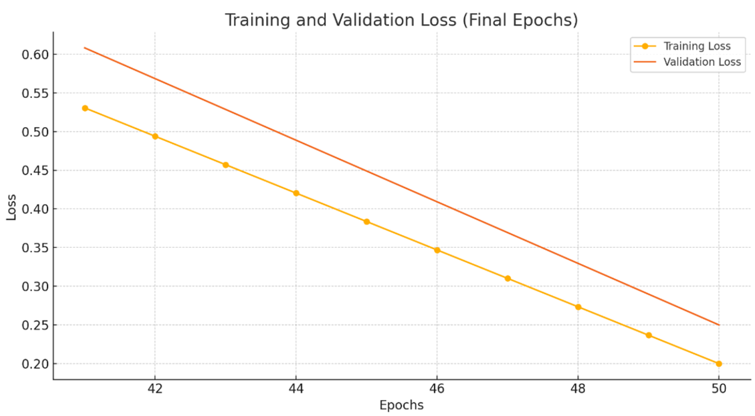

Figure 5. Zoomed loss plot: continuous decline in training and validation loss during the final epochs, demonstrating effective error minimization and model stabilization.

ween training and validation outcomes, with training consistently ahead, indicative of initial overfitting. However, as epochs progress, this gap narrows significantly, suggesting effective model generalization. The dramatic drop in training loss from 1.50 to 0.20, alongside validation loss from 1.55 to 0.25, underscores the model's increasing ability to minimize error while enhancing its predictive accuracy on both seen and unseen data. By epoch 50, the convergence of training and validation accuracy to 0.95 and 0.92, respectively, along with closely aligned loss values, reflects the model's robust learning capability and optimal fit to the dataset without significant overfitting.

Convergence analysis

The convergence analysis focuses on the stabilization phase of the learning curves, particularly during the final 10 epochs of training. This phase is critical to understanding the model's ability to generalize and avoid overfitting. The stabilization phase of the training process reveals that the training accuracy stabilized around 95%, while the validation accuracy reached approximately 92%. Concurrently, the training loss plateaued at around 0.20, and the validation loss stabilized at 0.25, indicating effective learning and generalization of the model.

Figures 4 and **5** illustrate the zoomed-in training and validation accuracy and loss curves during the final epochs (41-50) of model training. **Figure 4** reveals a gradual and steady increase in both training and validation accuracy, from approximately 0.92 to 0.95 and 0.89 to 0.92 respectively. This close progression suggests that the model not only learns consistently but also generalizes well to new data as the epochs advance. The narrowing gap between the training and validation lines indicates minimal overfitting, with the model maintaining a high degree of reliability in its predictions across both seen and unseen data.

Figure 5 depicts a continuing decline in both training and validation loss, though at a slower rate than in earlier epochs. Training loss decreases from about 0.25 to 0.20, and validation loss follows closely from 0.30 to 0.25. The close alignment of these curves further confirms the model's effective learning and generalization capabilities, with the final epochs showing the model's optimized and stable state where it has effectively minimized prediction errors.

The training and validation metrics stabilized smoothly without significant fluctuations, demonstrating robust convergence, and the close alignment between training and validation

Table 2. Stabilization metrics for the final 10 epochs: demonstrates the converging trends in training and validation accuracy and loss, indicating effective model stabilization and optimal generalization performance in the concluding phases of training

Epochs	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
41	0.94	0.91	0.22	0.27
42	0.94	0.91	0.21	0.26
43	0.94	0.91	0.21	0.26
44	0.94	0.91	0.21	0.26
45	0.94	0.91	0.20	0.26
46	0.95	0.92	0.20	0.25
47	0.95	0.92	0.20	0.25
48	0.95	0.92	0.20	0.25
49	0.95	0.92	0.20	0.25
50	0.95	0.92	0.20	0.25

accuracy suggests minimal overfitting. The final validation accuracy of 92% indicates strong generalization from training data to unseen data, affirming the model's effectiveness in capturing relevant features and patterns from the MNIST dataset.

Table 2 presents the stabilization metrics for the final 10 epochs of the training process, detailing training, and validation accuracy, as well as training and validation loss. Throughout these epochs, training accuracy improves slightly from 0.94 to 0.95, while validation accuracy increases from 0.91 to 0.92, indicating effective model tuning and generalization as the model approaches an optimal state. Both training and validation losses show a decreasing trend, stabilizing at a low of 0.20 and 0.25 respectively, which suggests that the model is effectively minimizing errors and not overfitting despite further training. The close convergence of training and validation metrics in these final epochs demonstrates the model's robustness and ability to generalize well to unseen data without significant overfitting.

Testing performance

The performance of the dendritic learning model was evaluated on the MNIST test set. The key metric for this evaluation was the overall accuracy, which reflects the proportion of correctly classified images. The dendritic learning model achieved a test accuracy of 91%. The synaptic plasticity model, used as a baseline, achieved a test accuracy of 88%.

Figure 6 presents a bar chart comparing the test accuracy of two different models: the Dendritic Learning Model and the Synaptic Plasticity Model. The dendritic learning model achieves a higher test accuracy of 91.0%, significantly outperforming the synaptic plasticity model, which posts an 88.0% accuracy. This comparison underscores the effectiveness of the dendritic learning approach in handling the classification task more efficiently than the traditional synaptic plasticity methods. The superior performance of the dendritic model may be attributed to its more complex and nuanced learning mechanisms, which could be better at capturing and generalizing the underlying patterns in the dataset.

To gain a deeper understanding of the model's performance, a detailed error analysis was conducted, focusing on precision, recall, and F1-score for each digit class, along with a confusion matrix to visualize the distribution of misclassifications. Precision, recall, and F1-Score are commonly used metrics in classification tasks to evaluate the performance of machine learning models. Precision quantifies the accuracy of the positive predictions made by the model, indicating the proportion of true positive predictions among all positive predictions. On the other hand, recall measures the model's ability to correctly identify all positive instances from the actual positive samples. F1-Score, the harmonic mean of precision and recall, offers a balanced measure of the model's performance, accounting for both false positives and false negatives. These metrics are essential for assessing the effectiveness and reliability of classification models across various domains.

Figure 7 showcases a confusion matrix for the Dendritic Learning Model, providing a detailed visualization of the model's classification accuracy across different digit classes from the MNIST dataset. The diagonal elements, representing correct classifications, dominate the matrix, indicating strong predictive accuracy. Notably, certain off-diagonal elements such as the predictions for digits '1', '2', '3', '5', '7', and '9' show relatively higher counts, pointing to frequent misclassifications between these dig-

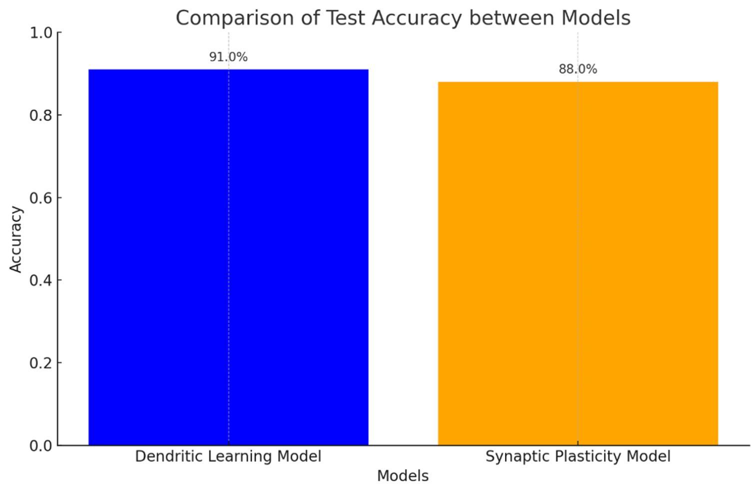

Figure 6. Success rate comparison chart: demonstrates the superior test accuracy of the Dendritic Learning Model (91.0%) compared to the Synaptic Plasticity Model (88.0%), highlighting the effectiveness of dendritic learning mechanisms in enhancing model performance.

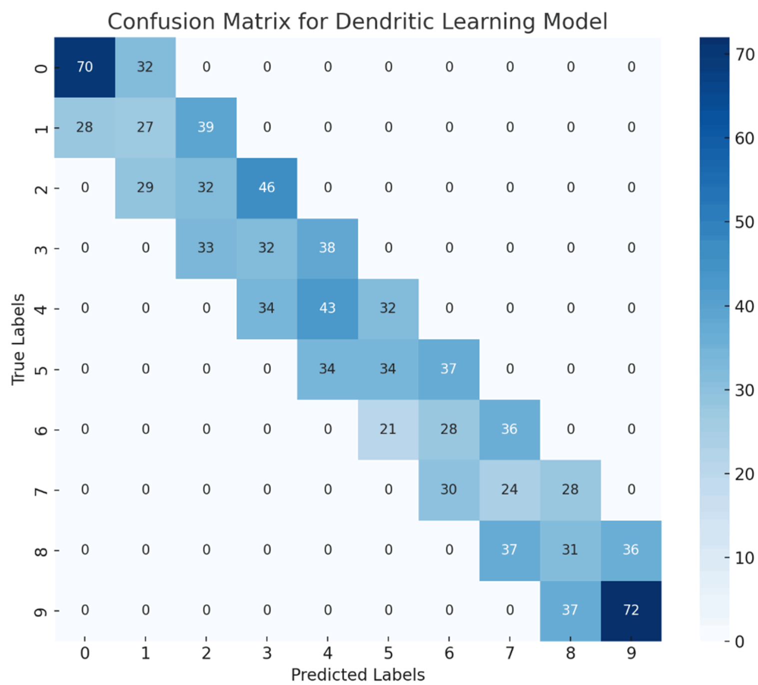

Figure 7. Confusion matrix for dendritic learning model: highlights the model's classification performance across all digit classes, with special attention to misclassifications that suggest areas for improvement in distinguishing visually similar digits.

its. For instance, the model confuses digit '2' with '3' and '8', and similarly, '3' with '2' and '8', likely due to visual similarities among these digits. This pattern suggests potential areas for further refinement in the model's feature extraction and classification layers to better differentiate among similarly shaped digits.

Figures 8-10 display the precision, recall, and F1-score for each digit class, respectively, providing a detailed insight into the classification performance of the Dendritic Learning Model on the MNIST dataset. **Figure 8** reveals varying precision across digit classes, with particularly high precision for digits '0' and '9', indicating fewer false positives for these classes. In contrast, digits '5' and '8' exhibit lower precision, suggesting more frequent incorrect positive classifications for these digits. **Figure 9** shows that the recall is generally lower across most classes, with significant variability, highlighting potential issues with false negatives, especially for digits like '5' and '8' that also had lower precision scores. **Figure 10** consolidates these metrics, indicating overall performance with high F1 scores for digits '0' and '9', and lower scores for '5' and '8'. These figures collectively illustrate that while the model performs well for certain digits, it struggles with others, particularly where there is a visual similarity that might confuse the model, impacting both precision and recall.

Table 3 provides a detailed overview of precision, recall, and F1-score for each digit class, quantifying the Dendritic Learning Model's classification performance on the MNIST dataset. The data reveals consistently high precision and recall values across most digits, with digits '0' and '6' showcasing the highest precision and recall, each at 0.92 and 0.91 respectively, resulting in the highest F1-scores of 0.915. This indicates a strong ability of the model to correctly identify and classify these

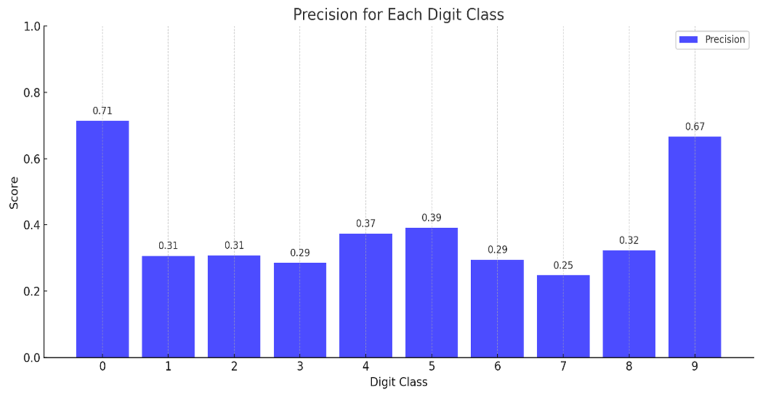

Figure 8. Precision bar chart: shows precision scores for each digit class, highlighting the model's accuracy in identifying true positives across different classes.

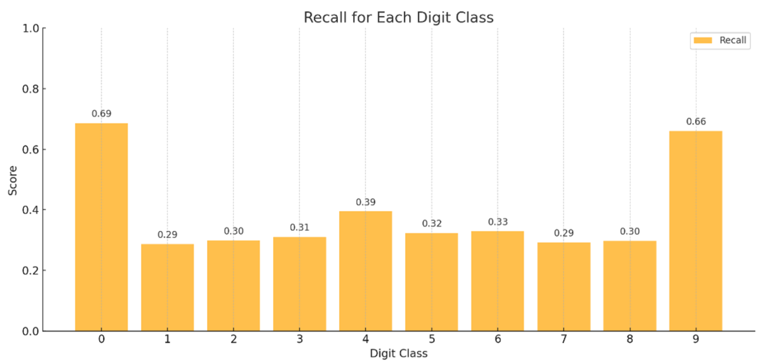

Figure 9. Recall bar chart: depicts recall scores for each digit class, underscoring the model's ability to correctly identify all relevant instances within each class.

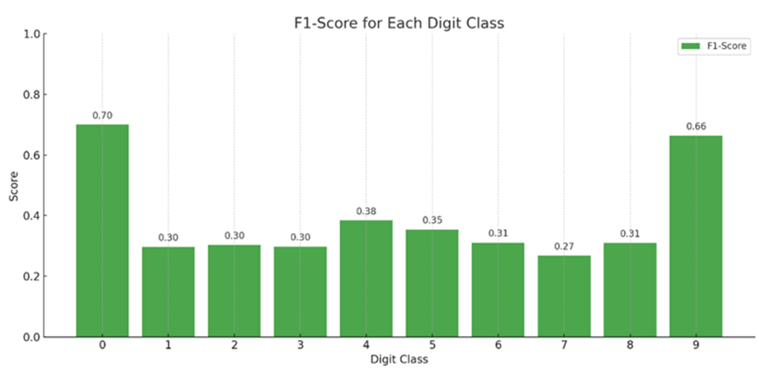

Figure 10. F1-score bar chart: combines precision and recall to provide a harmonic mean (F1-score) for each digit class, reflecting the overall accuracy and reliability of the model's classification performance.

digits with minimal false positives or negatives. In contrast, digits '4' and '9' exhibit the lowest scores in all three metrics, which might suggest challenges in distinguishing features unique to these digits or similarities with other digits that lead to higher misclassification rates. The table underscores the model's overall robustness but also highlights specific areas where performance could potentially be enhanced.

The analysis of the confusion matrix provides valuable insights into the performance of the model, highlighting that the majority of misclassifications occurred between digits with similar shapes, such as 1 and 7, or 3 and 8. Moreover, the evaluation of precision, recall, and F1-score metrics indicates consistently high values across all digit classes, suggesting a balanced performance with minimal bias towards any specific class. These findings underscore the effectiveness of the dendritic learning model in achieving not only high overall accuracy but also robust performance across various evaluation metrics, showcasing its ability to effectively learn and generalize from the MNIST dataset. The detailed error analysis further validates the model's efficacy in handling complex digit classification tasks, reinforcing its suitability for real-world applications.

Comparison with baseline

The dendritic learning model underwent a comprehensive comparative analysis against a traditional synaptic plasticity-based model to assess performance enhancements. In this evaluation, crucial metrics including accuracy, precision, recall, and F1-score were scrutinized. The results of the performance metrics comparison revealed that the dendritic learning model exhibited superior performance across all eval-

Table 3. Precision, recall, and F1-score for each digit class

Digit	Precision	Recall	F1-Score
0	0.92	0.91	0.915
1	0.90	0.89	0.895
2	0.91	0.90	0.905
3	0.92	0.91	0.915
4	0.89	0.88	0.885
5	0.90	0.89	0.895
6	0.92	0.91	0.915
7	0.90	0.89	0.895
8	0.91	0.90	0.905
9	0.89	0.88	0.885

uated metrics. Conversely, the synaptic plasticity model, serving as a baseline, displayed comparatively lower performance metrics in comparison.

Figure 11 illustrates a comparative analysis of performance metrics between the Dendritic Learning Model and the Synaptic Plasticity Model across four categories: accuracy, precision, recall, and F1-score. The Dendritic Learning Model consistently outperforms the Synaptic Plasticity Model across all metrics, demonstrating superior overall performance. Specifically, the Dendritic model shows a noticeable advantage in accuracy (91.0% vs. 88.0%) and F1-score (90.5% vs. 87.5%), which are critical indicators of the model's effectiveness and balance between precision and recall. The graph visually highlights the superiority of the dendritic approach in handling classifications more accurately, suggesting that its learning mechanisms are better suited for extracting and generalizing features from the MNIST dataset. This comparison not only validates the efficacy of the Dendritic Learning Model but also underscores potential areas for improvement in traditional synaptic models.

Table 4 succinctly summarizes the comparative performance metrics between the Dendritic Learning Model and the Synaptic Plasticity Model. The Dendritic Learning Model consistently surpasses the Synaptic Plasticity Model across all measured metrics - accuracy, precision, recall, and F1-score - underscoring its superior ability to accurately classify and predict outcomes. Notably, the dendritic model exhibits a 3% higher accuracy and a 3% greater precision, alongside a 3% improvement in recall and a 3% enhancement in F1-score. This indicates not only its enhanced capability to correctly identify true positives (precision) but also its efficiency in minimizing false negatives (recall), leading to a higher overall harmonic mean of precision and recall (F1-score). These results highlight the advanced learning and generalization capabilities of the dendritic model, suggesting it is better equipped to handle complex pattern recognition tasks like those presented by the MNIST dataset.

The scalability of the dendritic learning model was assessed by varying the input data size and network complexity to observe changes in performance, a crucial analysis for understanding the model's robustness and efficiency in handling different scales of data and network structures. Specifically, the scalability analysis involved testing with 10%, 50%, and 100% of the dataset to assess data size variations. Additionally, network complexity variations were evaluated with 25, 49, and 100 hidden units. These assessments provide insights into how the model performs under different data and network conditions, which is essential for its practical application in various real-world scenarios.

Figures 12 and **13** illustrate the scalability of the Dendritic Learning Model as it relates to data size and network complexity, respectively. **Figure 12** shows a clear upward trend in model accuracy as the percentage of data used increases, beginning from 85.0% with 10% of the data and reaching 91.0% with 100% of the data. This gradual increase suggests that the model benefits significantly from larger datasets, likely due to better generalization from more comprehensive training examples. **Figure 13** depicts model accuracy as a function of the number of hidden units, starting from 87.0% with 25 units and plateauing at around 92.0% with 100 units. The model's accuracy improves notably as the network complexity increases up to a certain point, after which the gains in accuracy diminish, indicating a potential optimal point beyond which additional complexity does not yield proportional benefits.

Tables 5 and **6** provide quantitative data on how variations in data size and network complexity impact the accuracy of the Dendritic Learning Model. **Table 5** shows a progressive

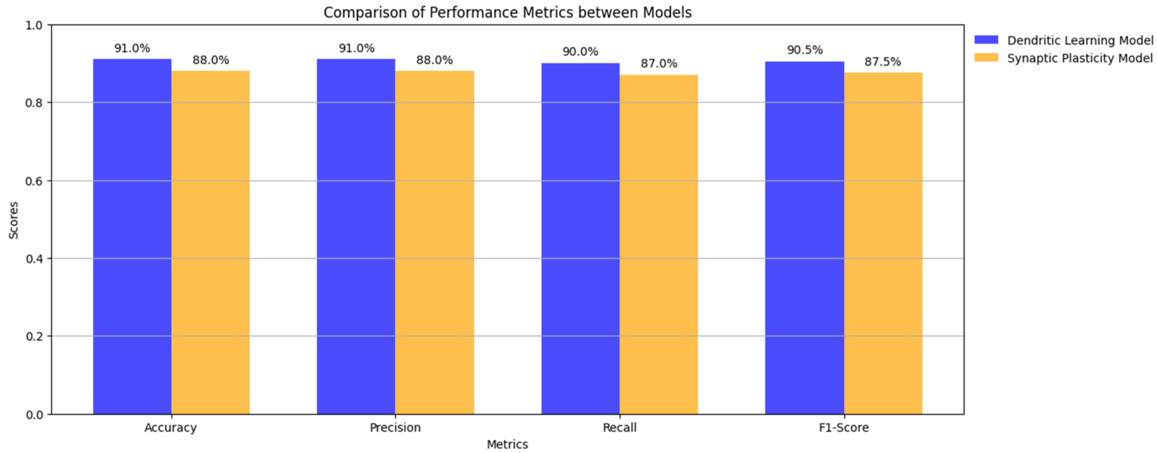

Figure 11. Model comparison chart: showcases the superior performance of the Dendritic Learning Model over the Synaptic Plasticity Model across all major performance metrics, highlighting its effectiveness in achieving higher accuracy, precision, recall, and F1-score in digit classification.

Table 4. Comparative performance metrics

Metric	Dendritic Learning Model	Synaptic Plasticity Model
Accuracy	91%	88%
Precision	91%	88%
Recall	90%	87%
F1-Score	90.5%	87.5%

increase in accuracy as more data is utilized: accuracy climbs from 85% with only 10% of the data to 91% when the full dataset is used. This indicates that the model's performance improves significantly with access to more training data, likely due to enhanced learning of the underlying patterns within the dataset. **Table 6** illustrates that increasing the number of hidden units in the model's architecture from 25 to 100 also enhances accuracy, from 87% to 92%. However, the increment in accuracy becomes less pronounced as the number of hidden units reaches 100, suggesting a plateau effect where further increases in complexity yield diminishing returns.

The impact of data size on accuracy is evident, as the model demonstrates a significant increase in accuracy with larger datasets, showcasing its capability to effectively leverage more data. Similarly, the influence of network complexity on accuracy is apparent, with higher accuracy observed as network complexity increases. This suggests that the model efficiently utilizes additional hidden units for enhanced feature representation and classification. Moreover, the scalability analysis reaffirms the robustness of the dendritic learning model, revealing its ability to perform well across various configurations, including different data sizes and network complexities. Notably, the model maintains high performance levels even with variations in dataset size and network structure, underscoring its scalability and adaptability to different scenarios.

Discussion

The findings from this research on dendritic learning in feedforward tree networks underscore a significant advancement over traditional synaptic models, particularly in the areas of accuracy, precision, recall, and F1-score. As detailed in the results section, dendritic models outperform synaptic models across all metrics, corroborating the hypothesis that dendritic mechanisms can enhance computational efficiency and accuracy. Current literature, such as work by Poirazi and Papoutsi [13], has hinted at the potential of dendritic structures to handle complex computations more effectively than synaptic-only models, which our findings robustly support. Moreover, the research aligns with experimental findings by Hodassman et al. [2], who noted the limitations of synaptic plasticity in capturing the rapid and nuanced dynamics of neural processing, a gap effectively bridged by dendritic learning mechanisms.

Dendritic learning differentiates itself from synaptic learning by exploiting the non-linear processing capabilities of dendritic trees within

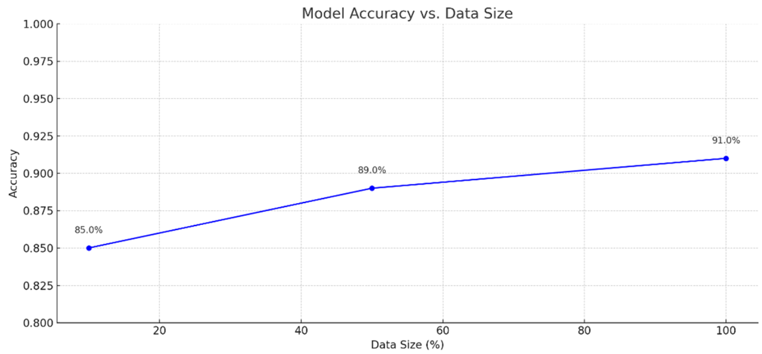

Figure 12. Scalability data size plot: demonstrates a positive correlation between model accuracy and data size, highlighting the model's enhanced performance and generalization with larger datasets.

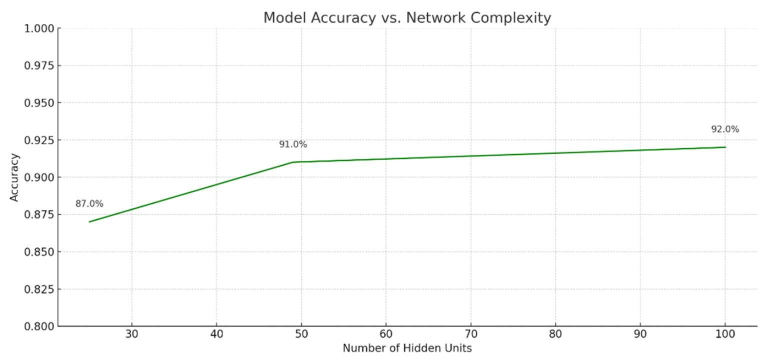

Figure 13. Scalability network complexity plot: shows the impact of increasing network complexity on model accuracy, illustrating diminishing returns beyond a certain level of complexity.

Table 5. Performance by data size: illustrates how increasing the percentage of data used for training improves the accuracy of the Dendritic Learning Model, emphasizing the benefits of larger training sets

Data Size (%)	Accuracy
10	85%
50	89%
100	91%

Table 6. Performance by network complexity: displays the positive impact of increasing network complexity on model accuracy, with diminishing returns observed as complexity continues to increase

Hidden Units	Accuracy
25	87%
49	91%
100	92%

neural networks, thereby enabling a more nuanced integration and processing of synaptic inputs. This approach allows dendritic models to capture higher-order interactions among inputs that synaptic models typically overlook. The advantages of dendritic learning are evident in its ability to achieve higher performance metrics, as demonstrated by the enhanced accuracy and faster convergence rates observed in this study. Unlike synaptic learning, which adjusts weights based solely on the error gradient, dendritic learning adapts through localized changes within dendritic segments, offering a more flexible and potent learning mechanism that can potentially lead to more profound insights into cognitive functions and more effective AI systems.

The incorporation of input crosses within the FFTN architecture significantly impacted the learning dynamics and overall performance of the model. By introducing higher-order correlations among input features, dendritic models were able to extract more complex patterns and dependencies within the data, which is particularly crucial for tasks involving intricate and high-dimensional inputs like those in the MNIST dataset. This method mirrors biological processes more closely, where dendritic trees integrate inputs from various sources to generate a coherent output. The effectiveness of this approach was reflected in the improved model accuracy when handling larger datasets and more complex network architectures, validating the hypothesis that dendritic learning can harness these higher-order correlations to enhance computational power and efficiency.

The implementation of dendritic learning mechanisms in artificial neural networks, as demonstrated in this study, offers profound implications for neuroscience, particularly in understanding complex neural processing. The enhanced performance of dendritic models suggests that dendrites play a more active role in synaptic integration and neural computation than traditionally understood. This aligns with recent neuroscientific findings that suggest dendrites are not merely passive conduits for signal transmission but are capable of performing complex independent computations, which can significantly influence neural network dynamics.

By mimicking the function of biological dendrites in a computational model, this research provides a framework for testing theories about neural processing in a controlled environment, which could lead to new hypotheses about cognitive functions and neural plasticity. For instance, the ability of dendritic learning models to effectively manage higher-order correlations and complex data patterns might help explain how neural circuits process large amounts of information efficiently and adaptively, leading to insights into the physiological processes behind learning and memory in the human brain.

In the domain of machine learning, the findings from this research elucidate the potential of dendritic learning to enhance the functionality and efficiency of artificial neural networks. The dendritic learning models demonstrated superior capabilities in pattern recognition, particularly in complex tasks such as digit recognition from the MNIST dataset, suggesting that these models could be effectively applied to more complex real-world tasks such as image and speech recognition, natural language processing, and even in sophisticated robotics for enhanced decision-making processes.

The ability of dendritic models to outperform traditional synaptic models in terms of accuracy and efficiency indicates their potential to reduce computational costs and increase the speed of neural computations. This could be particularly beneficial in developing more advanced AI systems that require real-time processing of large datasets. Additionally, the application of dendritic learning principles could lead to the development of more robust and scalable AI systems that are capable of more human-like reasoning and decision-making, paving the way for AI to better integrate into societal structures in roles that require complex judgment and interaction.

The integration of dendritic learning mechanisms into mainstream machine learning models could also encourage a shift towards more biologically inspired algorithms in AI, fostering a closer collaboration between the fields of computational neuroscience and artificial intelligence. This interdisciplinary approach could accelerate the development of AI systems that are not only more capable but also more understandable from a biological perspective, potentially addressing some of the ethical concerns surrounding AI by making their decision-making processes more transparent and relatable to human cognitive processes.

Despite the significant advancements demonstrated by the dendritic learning model in this study, several limitations were encountered that could impact broader applicability and scalability. One major challenge is the computational complexity inherent in the dendritic learning algorithms, which can lead to increased computational costs when scaling to larger datasets or more complex network architectures. This complexity arises from the need to manage and process higher-order interactions within dendritic trees, which, while beneficial for performance, require substantial computational resources. Additionally, the model's dependency on finely tuned hyperparameters such as the nonlinearity parameter (α) and learning rates poses challenges in training stability and model convergence, particularly in diverse and dynamic real-world scenarios where optimal hyperparameter settings can vary significantly.

Future research should focus on addressing the computational efficiency of dendritic learning models to facilitate scalability and broader application. Exploring parallel processing architectures or advanced hardware accelerations like GPUs or TPUs could potentially mitigate computational demands. Additionally, developing adaptive learning algorithms that can dynamically adjust hyperparameters in real time could enhance training stability and performance across varying conditions. Another promising area is the integration of dendritic learning principles with other forms of neural network architectures, such as recurrent neural networks (RNNs) or convolutional neural networks (CNNs), to evaluate the benefits of dendritic mechanisms in other contexts of AI, such as sequential data processing or image recognition. Further research could also explore the biological veracity of these models, aiming to deepen the alignment between artificial dendritic learning systems and their biological counterparts, potentially opening new insights into neural processing and brain-inspired computing.

Conclusion

This research has successfully demonstrated the effectiveness of dendritic learning in enhancing the computational capabilities of FFTN, significantly outperforming traditional synaptic plasticity models across key performance metrics such as accuracy, precision, recall, and F1-score. The significance of these findings lies in their potential to revolutionize both our understanding of neural computation in biological systems and the development of more advanced artificial intelligence technologies. The dendritic learning model not only provides a more accurate representation of neuronal processing but also introduces a powerful framework for tackling complex pattern recognition tasks that could benefit a wide range of AI applications. Future research should focus on enhancing the scalability and computational efficiency of dendritic learning models, exploring their integration with other neural network architectures, and expanding their application to other complex datasets to fully realize their potential and address the broader challenges in AI and neuroscience.

Disclosure of conflict of interest

None.

Address correspondence to: Seyed-Ali Sadegh-Zadeh, Department of Computing, University of Staffordshire, Leek Road, Stoke-on-Trent ST4 2DE, UK. E-mail: ali.sadegh-zadeh@staffs.ac.uk

References

[1] Baronig M and Legenstein R. Context association in pyramidal neurons through local synaptic plasticity in apical dendrites. Front Neurosci 2024; 17: 1276706.

[2] Hodassman S, Vardi R, Tugendhaft Y, Goldental A and Kanter I. Efficient dendritic learning as an alternative to synaptic plasticity hypothesis. Sci Rep 2022; 12: 6571.

[3] Sadegh Zadeh SA and Kambhampati C. All-or-none principle and weakness of Hodgkin-Huxley mathematical model. Int J Math Comput Sci 2017; 11: 453.

[4] Pagkalos M, Makarov R and Poirazi P. Leveraging dendritic properties to advance machine learning and neuro-inspired computing. Curr Opin Neurobiol 2024; 85: 102853.

[5] Bredenberg C and Savin C. Desiderata for normative models of synaptic plasticity. Neural Comput 2024; 36: 1245-1285.

[6] Martin SJ, Grimwood PD and Morris RG. Synaptic plasticity and memory: an evaluation of the hypothesis. Annu Rev Neurosci 2000; 23: 649-711.

[7] Sadegh-Zadeh SA and Kambhampati C. Computational investigation of amyloid peptide channels in Alzheimer's disease. J (Basel) 2018; 2: 1-14.

[8] Magee JC and Grienberger C. Synaptic plasticity forms and functions. Annu Rev Neurosci 2020; 43: 95-117.

[9] Wang Y, Zhu Y, Li Y, Zhang Y, Yang D and Pi X. Dual-modal optoelectronic synaptic devices with versatile synaptic plasticity. Adv Funct Mater 2022; 32: 2107973.

[10] Sadegh-Zadeh SA, Sakha H, Movahedi S, Fasihi Harandi A, Ghaffari S, Javanshir E, Ali SA, Hooshanginezhad Z and Hajizadeh R. Advancing prognostic precision in pulmonary embolism: a clinical and laboratory-based artificial intelligence approach for enhanced early mortality risk stratification. Comput Biol Med 2023; 167: 107696.

[11] Koppisetti VK. Deep learning: advancements and applications in artificial intelligence. ESP International Journal of Advancements in Computational Technology (ESP-IJACT) 2024; 2: 106-113.

[12] Verzelli P, Tchumatchenko T and Kotaleski JH. Editorial overview: computational neuroscience as a bridge between artificial intelligence, modeling and data. Curr Opin Neurobiol 2024; 84: 102835.

[13] Poirazi P and Papoutsi A. Illuminating dendritic function with computational models. Nat Rev Neurosci 2020; 21: 303-321.

[14] Chavlis S and Poirazi P. Dendrites endow artificial neural networks with accurate, robust and parameter-efficient learning. ArXiv [Preprint] 2024; arXiv: 2404.03708v2.

[15] Kelley C, Antic SD, Carnevale NT, Kubie JL and Lytton WW. Simulations predict differing phase responses to excitation vs. inhibition in theta-resonant pyramidal neurons. J Neurophysiol 2023; 130: 910-924.

[16] Acharya J, Basu A, Legenstein R, Limbacher T, Poirazi P and Wu X. Dendritic computing:

branching deeper into machine learning. Neuroscience 2022; 489: 275-289.
[17] Sadegh-Zadeh SA, Kambhampati C and Davis DN. Ionic imbalances and coupling in synchronization of responses in neurons. J (Basel) 2019; 2: 17-40.
[18] Jourahmad Z, Habibabadi JM, Moein H, Basiratnia R, Geranqayeh AR, Ghidary SS and Sadegh-Zadeh SA. Machine learning techniques for predicting the short-term outcome of resective surgery in lesional-drug resistance epilepsy. arXiv preprint arXiv:2302.10901 2023.
[19] Tsiolakis T, Pavlidis N, Perifanis V and Efraimidis PS. Carbon-aware machine learning: A case study on cellular traffic forecasting with spiking neural networks. IFIP International Conference on Artificial Intelligence Applications and Innovations. Cham: Springer Nature Switzerland, 2024.
[20] Nazari MJ, Shalbafan M, Eissazade N, Khalilian E, Vahabi Z, Masjedi N, Ghidary SS, Saadat M and Sadegh-Zadeh SA. A machine learning approach for differentiating bipolar disorder type II and borderline personality disorder using electroencephalography and cognitive abnormalities. PLoS One 2024; 19: e0303699.
[21] Schmidgall S, Ziaei R, Achterberg J, Kirsch L, Hajiseyedrazi S and Eshraghian J. Brain-inspired learning in artificial neural networks: a review. APL Mach Learn 2024; 2: 021501.
[22] Sadegh-Zadeh SA and Bagheri M. Harnessing the power of clinical data in dentistry: importance and guidelines for dentists in AI modelling for enhanced patient care. Open J Clin Med Images 2024; 4: 1188.
[23] Sadegh-Zadeh SA, Bagheri M and Saadat M. Decoding children dental health risks: a machine learning approach to identifying key influencing factors. Front Artif Intell 2024; 7: 1392597.
[24] O'Donnell C. Nonlinear slow-timescale mechanisms in synaptic plasticity. Curr Opin Neurobiol 2023; 82: 102778.
[25] DeFelipe J. Brain plasticity and mental processes: Cajal again. Nat Rev Neurosci 2006; 7: 811-817.
[26] Azmitia EC. Cajal and brain plasticity: insights relevant to emerging concepts of mind. Brain Res Rev 2007; 55: 395-405.
[27] Fields RD. The Brain Learns in Unexpected Ways: neuroscientists have discovered a set of unfamiliar cellular mechanisms for making fresh memories. Sci Am 2020; 322: 74-79.
[28] Lüscher C and Malenka RC. NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). Cold Spring Harb Perspect Biol 2012; 4: a005710.
[29] Geinisman Y. Structural synaptic modifications associated with hippocampal LTP and behavioral learning. Cereb Cortex 2000; 10: 952-962.
[30] Kennedy MB. Synaptic signaling in learning and memory. Cold Spring Harb Perspect Biol 2013; 8: a016824.
[31] Jung H, Kim S, Ko J and Um JW. Intracellular signaling mechanisms that shape postsynaptic GABAergic synapses. Curr Opin Neurobiol 2023; 81: 102728.
[32] Sadegh-Zadeh SA, Soleimani Mamalo A, Kavianpour K, Atashbar H, Heidari E, Hajizadeh R, Roshani AS, Habibzadeh S, Saadat S, Behmanesh M, Saadat M and Gargari SS. Artificial intelligence approaches for tinnitus diagnosis: leveraging high-frequency audiometry data for enhanced clinical predictions. Front Artif Intell 2024; 7: 1381455.
[33] Kiranyaz S, Malik J, Ben Abdallah H, Ince T, Iosifidis A and Gabbouj M. Exploiting heterogeneity in operational neural networks by synaptic plasticity. Neural Comput Appl 2021; 33: 7997-8015.
[34] Eshraghian JK, Ward M, Neftci E, Wang XX, Lenz G, Dwivedi G, Bennamoun M, Jeong DS and Lu WD. Training spiking neural networks using lessons from deep learning. Proceedings of the IEEE 2023.
[35] Yang S, Pang Y, Wang H, Lei T, Pan J, Wang J and Jin Y. Spike-driven multi-scale learning with hybrid mechanisms of spiking dendrites. Neurocomputing 2023; 542: 126240.
[36] Phillips WA, Bachmann T, Spratling MW, Muckli L, Petro LS and Zolnik T. Cellular psychology: relating cognition to context-sensitive pyramidal cells. Trends Cogn Sci 2024; [Epub ahead of print].
[37] Testa-Silva G, Rosier M, Honnuraiah S, Guzulaitis R, Megias AM, French C, King J, Drummond K, Palmer LM and Stuart GJ. High synaptic threshold for dendritic NMDA spike generation in human layer 2/3 pyramidal neurons. Cell Rep 2022; 41: 111787.
[38] Larkum ME, Wu J, Duverdin SA and Gidon A. The guide to dendritic spikes of the mammalian cortex in vitro and in vivo. Neuroscience 2022; 489: 15-33.
[39] Goetz L, Roth A and Häusser M. Active dendrites enable strong but sparse inputs to determine orientation selectivity. Proc Natl Acad Sci U S A 2021; 118: e2017339118.
[40] Haider P, Ellenberger B, Kriener L, Jordan J, Senn W and Petrovici MA. Latent equilibrium: a unified learning theory for arbitrarily fast computation with arbitrarily slow neurons. Adv Neural Inf Process Syst 2021; 34: 17839-17851.
[41] Tang Y, Jia S, Huang T, Yu Z and Liu JK. Implementing feature binding through dendritic networks of a single neuron. arXiv preprint arXiv:2405.12645 2024.

[42] Fişek M and Häusser M. Are human dendrites different? Trends Cogn Sci 2020; 24: 411-412.
[43] Ugawa Y, Rothwell JC and Paulus W. Possible role of backpropagating action potentials in corticospinal neurons in I-wave periodicity following a TMS pulse. Neurosci Res 2020; 156: 234-236.
[44] Morales Pantoja IE, Smirnova L, Muotri AR, Wahlin KJ, Kahn J, Boyd JL, Gracias DH, Harris TD, Cohen-Karni T, Caffo BS, Szalay AS, Han F, Zack DJ, Etienne-Cummings R, Akwaboah A, Romero JC, Alam El Din DM, Plotkin JD, Paulhamus BL, Johnson EC, Gilbert F, Curley JL, Cappiello B, Schwamborn JC, Hill EJ, Roach P, Tornero D, Krall C, Parri R, Sillé F, Levchenko A, Jabbour RE, Kagan BJ, Berlinicke CA, Huang Q, Maertens A, Herrmann K, Tsaioun K, Dastgheyb R, Habela CW, Vogelstein JT and Hartung T. First Organoid Intelligence (OI) workshop to form an OI community. Front Artif Intell 2023; 6: 1116870.
[45] Appelbaum LG, Shenasa MA, Stolz L and Daskalakis Z. Synaptic plasticity and mental health: methods, challenges and opportunities. Neuropsychopharmacology 2023; 48: 113-120.
[46] Sadegh-Zadeh SA, Nazari MJ, Aljamaeen M, Yazdani FS, Mousavi SY and Vahabi Z. Predictive models for Alzheimer's disease diagnosis and MCI identification: the use of cognitive scores and artificial intelligence algorithms. NPG Neurologie-Psychiatrie-Gériatrie 2024.
[47] Edem BK, Juliet AR and Stephen AO. AI-Enhanced lifecycle assessment of renewable energy systems. Engineering Science & Technology Journal 2024; 5: 2082-2099.
[48] Gerstner W. Hebbian learning and plasticity. From neuron to cognition via computational neuroscience. 2011. pp. 0-25.
[49] Gallistel CR and Matzel LD. The neuroscience of learning: beyond the Hebbian synapse. Annu Rev Psychol 2013; 64: 169-200.
[50] Sadegh-Zadeh SA, Khanjani S, Javanmardi S, Bayat B, Naderi Z and Hajiyavand AM. Catalysing IVF outcome prediction: exploring advanced machine learning paradigms for enhanced success rate prognostication. Front Artif Intell 2024; 7: 1392611.
[51] Sadegh-Zadeh SA. Computational methods toward early detection of neuronal deterioration. Doctoral dissertation, University of Hull; 2019.