



Dual Prism Tri-Modal 3D Pathed Neural Networks (DPTM-3DPNN's) by *Ben Darlington*

1. Introduction

The ***Dual Prism Tri-Modal Pathed Neural Network (DPTM-3DPNN)*** is an advanced neural network architecture designed to handle multi-modal inputs (*text, audio, and image data*) and fuse them into a unified output through a fusion and diffusion process. The network is characterized by its ***path-based neuron activation*** in a ***3D grid***, with neurons that have health indicators and dynamic creation properties. The model offers a highly adaptive approach to learning across multiple modalities, with significant reductions in computational complexity due to its selective neuron activation.

2. Key Features

- 1. 3D Neural Network:** The network for each modality is structured as a 3D grid, where inputs traverse through this space and activate specific neurons.
- 2. Tri-Modal Input:** The network processes *text, audio, and image* data simultaneously through dedicated 3D networks for each modality.
- 3. Fusion and Diffusion Layers:** A ***fusion layer*** combines the outputs of the three 3D input networks, while a ***diffusion layer*** splits the fused output back into three separate modality-specific outputs.
- 4. Path-Based Activation:** Only a subset of neurons are activated based on the input's path through the network, reducing computational overhead and making the network more efficient.
- 5. Neuron Health:** Each neuron has a health score that affects its learning rate and whether it should be pruned (removed) or reborn (reinitialized).
- 6. Dynamic Neuron Creation:** Neurons can be dynamically created when a path exceeds the network's current bounds, allowing the network to grow and adapt over time.
- 7. Model Saving and Loading:** The model can be saved as a JSON file, allowing for easy persistence and reloading of trained networks.

3. Architecture Overview

3.1 3D Neural Network Structure

Each modality (*text*, *audio*, *image*) has its own **3D neural network**, which consists of neurons arranged in a grid of dimensions `(depth, width, height)`.

- **Text Network:** Responsible for processing textual input data.
- **Audio Network:** Responsible for processing audio input data.
- **Image Network:** Responsible for processing image input data.

Each network has:

- A grid of neurons, where each neuron processes part of the input data and forwards it to the next neuron along a path.
- Neurons have weights, learning functions, and health indicators.

3.2 Neuron Structure

Each neuron is represented by:

-Weights:

A 3D vector that determines how much influence the neuron has on the next neuron in the path.

-Health:

A scalar value representing the "fitness" of the neuron. Neurons with low health may be pruned (removed) from the network, and neurons can be "reborn" if the network needs more capacity.

-Learning Function:

The activation function for the neuron, which could be ReLU, Sigmoid, Tanh, or other functions.

4. Fusion and Diffusion

4.1 Fusion Layer

After processing through the 3D networks for text, audio, and image, the outputs of these networks are **fused** together into a combined representation. The fusion is performed by taking the outputs of the three modalities and averaging them to create a unified representation in 3D space.

4.2 Diffusion Layer

Once the fused representation is computed, it is **diffused** back into separate outputs for each modality (*text*, *audio*, *image*). The diffusion process splits the fused output back into three 3D outputs, which are then fed into the respective output networks.

5. Path-Based Processing

Instead of activating every neuron in the network, **path-based processing** ensures that only neurons on the specific path from the input to the output are activated. This drastically reduces the number of neurons involved in each computation, resulting in more efficient training and inference.

5.1 Path Traversal

Each input (*text, audio, or image*) follows a unique path through its corresponding 3D network. The path is determined by the weights of the neurons, which dictate the direction of traversal in 3D space. If a path goes out of bounds (e.g., the path extends beyond the current size of the network), new neurons are created dynamically.

5.2 Path-Based Backpropagation

During training, only the neurons on the active path are updated. This allows the network to focus on the specific neurons that contributed to the output, making training more efficient and reducing the risk of overfitting.

6. Neuron Health and Dynamic Creation

6.1 Neuron Health

Each neuron has a **health score**, which is a measure of how successful it has been in previous tasks. Neurons that consistently perform well will have high health, while those that perform poorly will lose health over time.

- **Health Gain:** If a neuron contributes positively to the network's output, it gains health.
- **Health Loss:** If a neuron contributes negatively, it loses health.
- **Pruning:** If a neuron's health falls below a certain threshold, it is pruned from the network, meaning it is no longer part of the active model.
- **Rebirth:** If a path requires a neuron that has been pruned (or the path extends beyond the current network bounds), a new neuron is created with random weights and learning functions.

6.2 Dynamic Neuron Creation

If the path exceeds the network's current size, new neurons are dynamically created. This allows the network to **grow organically** over time, adding capacity as needed based on the complexity of the task.

7. Learning Process

7.1 Feedforward

During feedforward, the input is passed through the 3D network by activating a path of neurons. Each neuron on the path adjusts the input according to its weights and forwards it to the next neuron.

7.2 Backpropagation

After the feedforward pass, the network calculates the error between the predicted output and the actual target output. The error is backpropagated through the active path, and the weights of the neurons on the path are adjusted to minimize the error. Only neurons on the active path are updated during training, making the process more efficient.

8. Model Persistence (Saving and Loading)

The entire model, including the weights, health of neurons, and structure of the 3D grid, can be saved to a JSON file. This allows the model to be reloaded later for further training or inference. Each network (*text*, *audio*, and *image*) is saved separately, along with the fusion and diffusion layers.

9. Computational Efficiency

One of the primary advantages of **DPTM-3DPNN** is its computational efficiency:

- ***Path-Based Activation***: Only a small subset of neurons are activated during each inference, reducing the overall computation needed compared to fully-connected networks.
- ***Neuron Pruning and Rebirth***: Neurons that perform poorly are pruned, reducing the size of the network, while dynamic neuron creation ensures that the network has enough capacity for complex tasks.
- ***Selective Backpropagation***: Since only neurons on the active path are updated, training is faster and more focused.

10. Use Cases

10.1 Multi-Modal Data Fusion

DPTM-3DPNN is ideal for tasks that require the fusion of multiple modalities (text, audio, image). For example:

- ***Sentiment Analysis***: Combining text sentiment, audio tone, and image context to provide a unified sentiment score.
- ***Multimodal Translation***: Translating between different languages using text, audio, and images to provide context.

10.2 Dynamic Environments

The dynamic neuron creation feature makes **DPTM-3DPNN** well-suited for tasks that require adaptability and scalability, such as:

Online Learning: Continuously learning from new data as the environment changes.

Adaptive Control Systems: Adjusting the network structure based on real-time feedback in autonomous systems or robotics.

11. Challenges and Future Work

11.1 Hyperparameter Tuning

Tuning the parameters of the network, such as the depth, width, and height of each modality's network, the learning rates, and the neuron health thresholds, is critical for optimizing performance.

11.2 Path Exploration

Choosing the optimal paths through the network remains a challenge. Future work may focus on developing more sophisticated pathfinding algorithms to better explore the 3D space and discover optimal neuron activation sequences.

11.3 Neuron Function Diversity

Currently, neurons are initialized with random learning functions. Future work could explore the use of more diverse activation functions, or allow neurons to evolve their functions based on task-specific performance.

12. Conclusion

The **Dual Prism Tri-Modal Pathed Neural Network (DPTM-3DPNN)** is a highly adaptive, multi-modal neural network architecture that uses 3D path-based processing to efficiently handle text, audio, and image data. Its dynamic neuron creation and health-based learning allow the network to grow and evolve as needed, making it a powerful tool for tasks that require multi-modal data fusion and real-time adaptability.

The combination of path-based processing, neuron health management, and dynamic growth ensures that the ***DPTM-3DPNN*** is both **computationally efficient**** and **calable**, capable of handling complex real-world tasks with minimal computational overhead.