Exploring Deep Recurrent and Spiking Neural Networks in Hand Gesture Recognition

Alwin Lor

Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota, USA

16 November 2024

Modern Prosthetics

- Modern science and engineering has led to more advanced prosthetics
- Uses electrical signals in the residual limb to identify a gesture

TASKA HandGen2
https://www.taskaprosthetics.com/products/taska-gen2

Methods Presented Today

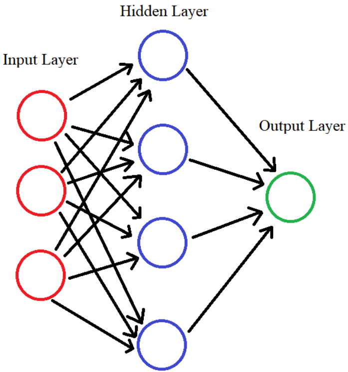
- Accomplish the same goal of gesture recognition
- Deep Recurrent Neural Networks with Long Short Term Memory as presented by Aliman et al. [1]
- Spiking Neural Networks as presented by Garg et al. [2]

Outline

1. Deep Recurrent Neural Networks
2. Spiking Neural Networks
3. Experiment Specifics
4. Results and Closing Thoughts

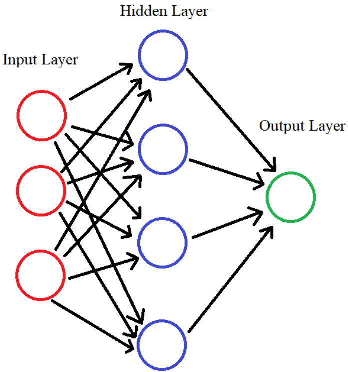
Neural Networks

- Inspired by how human brains process information
- Consists of an input layer, hidden layer(s), and an output layer
- Weights on edges influence information as it passes layer to layer



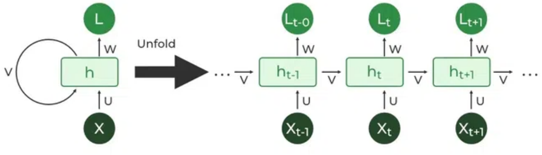
Neural Network Training

- Supervised training uses data where the outcome is known
- Backpropagation works backward from the output to the input to adjust weights
- Minimizing the average error on the dataset



Recurrent Neural Networks

- Variation of neural networks that work with sequential data
- Hidden nodes can loop back to themselves
- Uses backpropagation through time for training



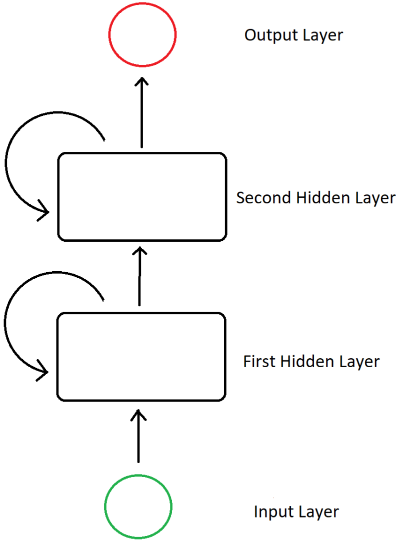
https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/

Long Short Term Memory (LSTM)

- Allows a Recurrent Neural Network to learn long-term relationships
- Uses memory cells in the hidden layer(s)
- A memory cell has a forget gate, input gate, and output gate

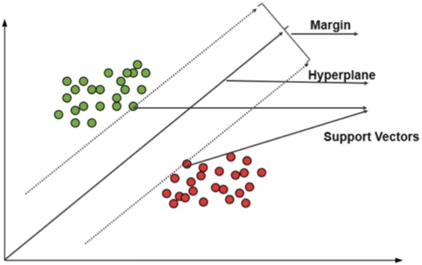
Deep Recurrent Neural Networks

- Is a Recurrent Neural Network with two or more hidden layers to process data
- Allows further learning of relationships between features within the data



Support Vector Machines (SVM)

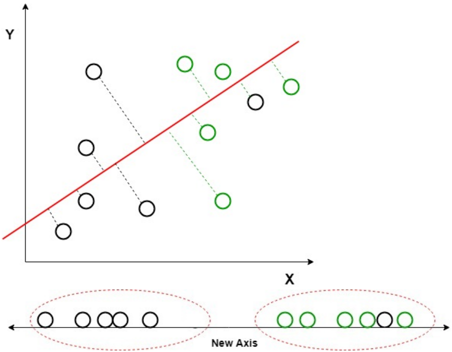
- Finds the hyperplane that maximizes the distance between classes
- *n*-dimensional hyperplanes can separate data in an $(n+1)$-dimensional space



https://www.ibm.com/topics/support-vector-machine

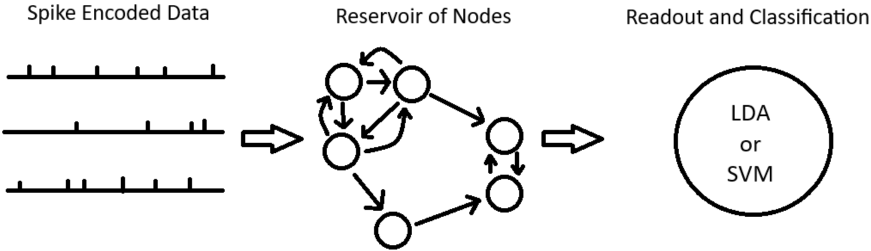
Linear Discriminant Analysis (LDA)

- Uses combinations of features and the labels of the data
- Projects the new combination of data points onto a lower dimension for classification



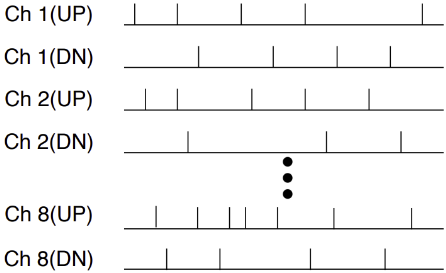
https://www.geeksforgeeks.org/ml-linear-discriminant-analysis/

Spiking Neural Network Overview



Spike Encoded data

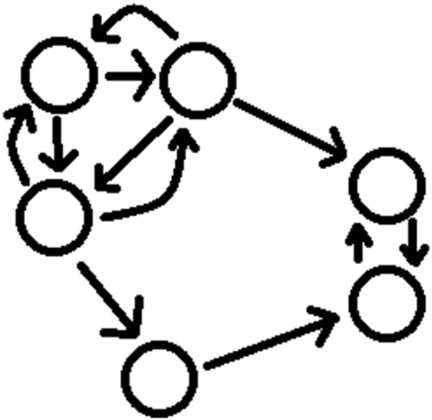
- A spike represents a point in the data where it exceeds a predefined threshold



Garg et Al. 2021
Signals to Spikes for Neuromorphic Regulated Reservoir Computing and EMG Hand Gesture Recognition

Reservoir of Nodes

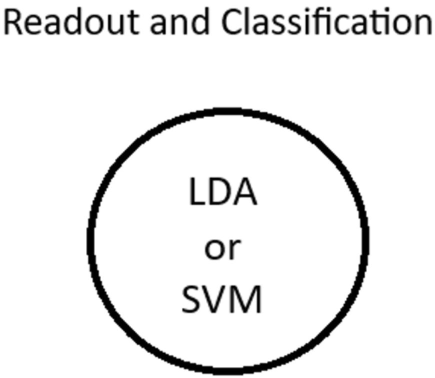
- Nodes update their internal state and wait to output
- Once the internal state reaches a threshold, the node will output
- Reservoirs have a topology that determines connections between nodes



Spiking Neural Networks

Readout and Classification of Data

- Uses the vector of the reservoir's state as output
- Uses a classifier such as SVM or LDA on the vector



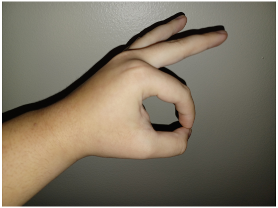
Data Collection

- Data for both experiments were collected using the Myo Armband
- Eight sensors that measure electromyographic signals through the skin

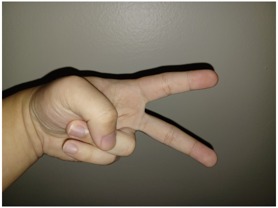
https://wearables.com/products/myo

Experiment Datasets

Okay hand gesture



Scissors hand gesture



Rock hand gesture

Paper hand gesture

Image Reference: Alwin Lor

Garg et al. SNN specifics

- As a node receives input, its threshold to fire off an output is increased. This is a Leaky Integrate-and-Fire node
- Weights are adjusted with Critical Plasticity. When a node outputs a signal, it essentially is output to only one other node. 80% of edges use Critical Plasticity
- Topology used is small-world-like

Result Metrics of the Deep Recurrent Neural Network

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{F1-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Results of the Deep Recurrent Neural Network

	Okay	Paper	Scissors	Rock
precision	0.98	0.93	0.95	0.92
recall	0.98	0.98	0.92	0.91
f1-score	0.98	0.96	0.94	0.92

Adapted from Aliman et al. [1]

Result metric of the Spiking Neural Network

$$\text{Accuracy} = \frac{\text{Correct}}{\text{Total}}$$

Results of the Spiking Neural Network

Dataset	Method	Readout classifier	Accuracy ($\mu \pm \sigma$)
Roshambo	Spike Enc. & Eval. Baseline	LDA	74.61 ± 2.44%
		SVM	85.44 ± 0.75%
	Reservoir (No plasticity)	LDA	76.44 ± 1.80%
		SVM	81.83 ± 0.93%
	Reservoir (CRITICAL)	LDA	83.06 ± 0.92%
		SVM	**88.00 ± 0.29%**

Garg et al. [2]

Closing Thoughts

- Difficult to directly compare the two methods due to differing metrics
- Both methods perform well according to their metrics
- The two may be closer in performance

Questions?

References

Hajar Y Aliman et al.
Deep Recurrent Neural Network Approach with LSTM Structure for Hand Movement Recognition Using EMG Signals
In Proceedings of the 2023 12th International Conference on Software and Information Engineering (Sharm El-Sheikh, Egypt) (ICSIE â23). Association for Computing Ma- chinery, New York, NY, USA, 58â65. https://doi.org/10.1145/3634848. 3634851

Nikhil Garg et al.
Signals to Spikes for Neuromorphic Regulated Reservoir Computing and EMG Hand Gesture Recognition
In International Conference on Neuromorphic Systems 2021 (Knoxville, TN, USA) (ICONS 2021). Association for Computing Machinery, New York, NY, USA, Article 29, 8 pages. https://doi.org/10.1145/3477145. 3477267