

Research article

Investigation of Memristor-based neural networks on Pattern recognition

Gayatri Routhu ^{1*}, Ngangbam Phalguni Singh¹, Selvakumar R¹, Eppala Shashi Kumar Reddy¹

routhugayatri0207@gmail.com, phalsingh@gmail.com, selvakumar.r@kluniversity.in, shashieppala150@gmail.com

¹Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, KL Deemed to be University, Vaddeswaram, Andhra Pradesh, India

Abstract: Mobile phones, laptops, computers, digital watches, and digital calculators are some of the most much-using products in our daily life. In the background, to make these gadgets work as per our desire there are many simple components necessary for electronics to function like resistors, capacitors, inductors, etc., these are three basic circuit elements. The Memristor is one such component. This paper provides simulation results of the memristor circuit and its V-I characteristics at different functions as an input signal. A well-trained ANN can able to recognize images with higher precision. To enhance the properties like accuracy, precision, and efficiency in recognition memristor characteristics are introduced to the neural network but previous devices experience some non-linearity issues causing conductance tuning problems. At the same time to be as advanceable in some applications, ANN requires a huge amount of vector-matrix multiplication based on in-depth network expansion. An ionic floating gate (IFG) device with the characteristics of a memristive device can solve these problems. This work proposes a fully connected ANN using the IFG model, the simulation results of the IFG model are given as synapses in deep learning. We use algorithms like the Gradient-descent model, Forward and Backward propagation for network building, and weight setting in neural networks to enhance their ability to recognize images. A well-trained network is formed by tuning those memristive devices to an optimized state. The synaptic memory which is obtained from the IFG device will be used in other deep neural networks to increase recognition accuracy. To be an activation function in the neural network sigmoid functions were used but later it is replaced by the ReLu function to avoid vanishing gradients. Here in this paper, we can see how images were recognized by their front, top, and side views.

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

Citation: To be added by editorial staff during Published: date production.

Academic Editor: Firstname Lastname



Keywords: Memristor; Memristive device; Ionic Floating Gate (IFG); Artificial neural network (ANN); Gradient descent model; Forward and Backward propagation; ReLu function.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

1. Introduction

A memristor is a simple passive two-terminal structure first coined by professor Leon Chua in 1971. It is the concatenation of a “Memory resistor”, which is a fourth fundamental circuit element besides resistor, capacitor, and inductor, this is the missing pair link that states the relationship between charge and flux over four basic circuit variables[1]. Depending on the input signal applied, there are two types of memristors, current-controlled and voltage-controlled [2]. The basic memristor is given in figure (1).

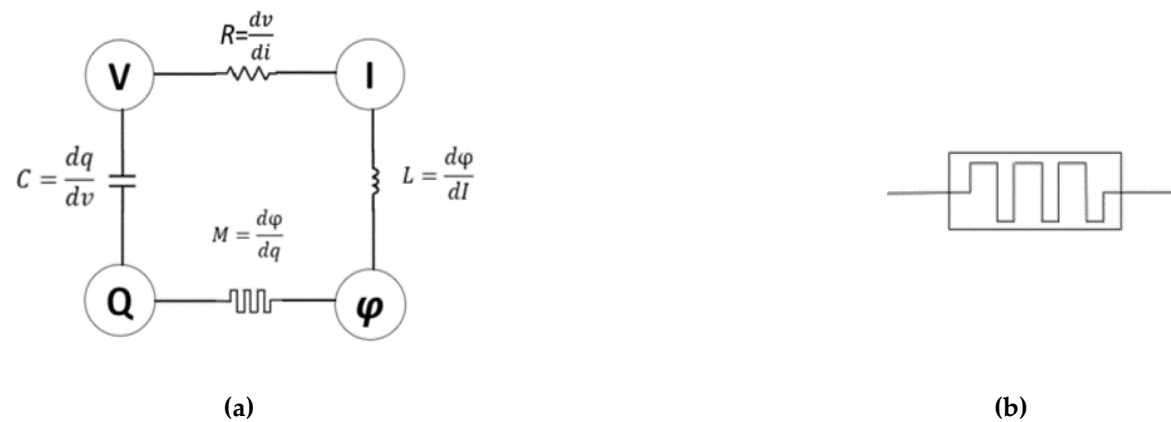


Figure 1. (a). The link between the Voltage and Current is shown by a resistor, Voltage, and Charge by a capacitor, and current and flux by an inductor; (b) Memristor basic symbol.

By equations (1) & (2) we can find the value memristance or memductance. The properties of the memristor made it useful for non-volatile memory and storage technology applications as it can theoretically develop multiple states. It can also operate at a very low voltage level. For the operation of a neural network as a synapse, these properties of memristor could be useful. The below equations (1) and (2) show the relationship between flux and charge concerning voltage and current [1].

1

2

3

4

5

6

7

8

9

10

11

$$\mathbf{M}(\mathbf{q}) = \frac{d\Psi}{d\mathbf{q}} \quad (1)$$

$$\mathbf{M}(\mathbf{q}(t)) = \frac{d\Psi}{dt} \cdot \frac{dt}{dq} = \frac{V(t)}{I(t)} \quad (2)$$

ANNs are interconnected groups of nodes that are composed of artificial neurons. An artificial neural network is a non-linear and self-adapting computational model which retains the biological concept of neural networks as human/animal brains[3][1]. An ANN has three layers, an input layer, a hidden layer, and the output layer. Every node of the input layer connects through links with every node of the next preceding layer, called a fully connected network[4][2]. It receives inputs, combines them, performs required computations based on a predefined activation function, and delivers the output. The devices with ANN using two terminals suffer from non-linearity and asymmetric conductance tuning problems. Due to this, their scalability was also affected. A memristive device with IFG (Ionic Floating Gate) model will be introduced to solve these problems. Such that an ANN with memristive properties gives Memristor-based Neural Network (MNN). To implement this a cadence Verilog A model will be used along with deep learning concepts. A memristor device has a simple structure with two terminals, its conductance can be modified by simple positive and negative pulses while representing synaptic weight[1][5]. These main functions of the memristor made it suitable for realizing the synaptic weight in an artificial neural network. By using a memristor device output as a synapse, a memristive neural network is expanded to a multi-layer with modified memristor-based backpropagation and gradient descent learning rules[2][6].

2. Memristor-based neural networks

For the design of MNN building the network and weight loading both takes a crucial role. Thereafter the IFG model will act as a conductor. For the pattern recognition to be verified by MNN a dataset called MNIST is used at input neurons. Also, we use resistors to build a fully connected MNN. The optimization of weights is done by software[1][7]. The following equation (3) is given for the output voltage.

$$\mathbf{Vout} = \mathbf{Rlimit} \cdot \sum_{i=1}^{\Psi} V Ri \cdot \frac{1}{Rij} \quad (3)$$

Here R_{ij} is the reciprocal of optimized conductance weights W_{ij4e} , which are obtained by software training. To train and test the neural network, we need a massive dataset, fortunately, the MNIST (modified national institute of standards and technology) database exists which contains 60,000 training images and 10,000 testing images. In this dataset, picture pixels will be normalized by greyscale numbers divided from 0 to 255 values[2][8].

2.1. Memristor-based IFG model

This IFG (Ionic Floating Gate) memristive device has three terminals, it is a combination of a redox transistor and a non-volatile CBM, to make it a non-volatile synaptic memory[1][9]. A redox transistor was developed in memristor techniques to find a way to the limitations like low writing efficiency, vanishing gradients, and limited accuracy. It contains three layers, the first layer is PEDOT: PSS film which is made of poly (3,4)- ethylene-dioxythiophene: polystyrene sulfonate, followed by the Nafion layer, the last layer is PEI (poly(ethylenimine)) layer which is partially reduced by PEDOT: PSS film[2][10]. Connecting to this a conductive bridge memory is designed with a layer of Ag in between Pt electrodes. The IFG model can be used as a memory storage device to memorize the last operation. Memory operates in two modes that are read operation and write operation. Figures (2) and (3) shows the internal view of the IFG memory[1][11].

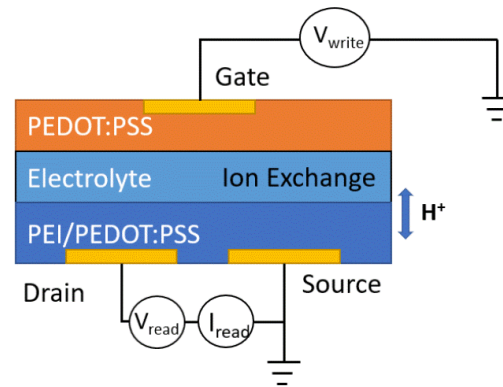


Figure 2. A polymer-based redox transistor



1
2
3
4
5
6
7
8
9

10

1
2

Figure 3. Complete IFG model showing connections between redox transistor and CBM
(a) Write operation; (b) Read operation.

By taking the above two models as a reference including some modifications in them, a compact IFG model characterized in Verilog-A 3
code is developed, and importing this Verilog-A code file to cadence generates the IFG model[2][12]. The cadence model is shown in 4
figure [4]. Here we change the selector position to the electrolyte and middle of the gate terminal. These are suitable for the design of 5
memristor-based synaptic circuits. The memristor IFG model conductance values are taken as synaptic arrays. For the recognition of 6
patterns/images, we use a three-layer artificial network that is fully connected[1][13]. All weight connections hold their fixed conductance 7
between the layers of the neural network which is stored as IFG's memristance 'G' and mapped into the source-drain conductance of 8
the synaptic array[2][14]. Here, the change in conductance is proportional to the flux, its gate voltage is greater than the threshold voltage, 9
and tunes the gate to the source voltage[15]. The schematic circuit of IFG is shown in (4) which is implemented in cadence. 10

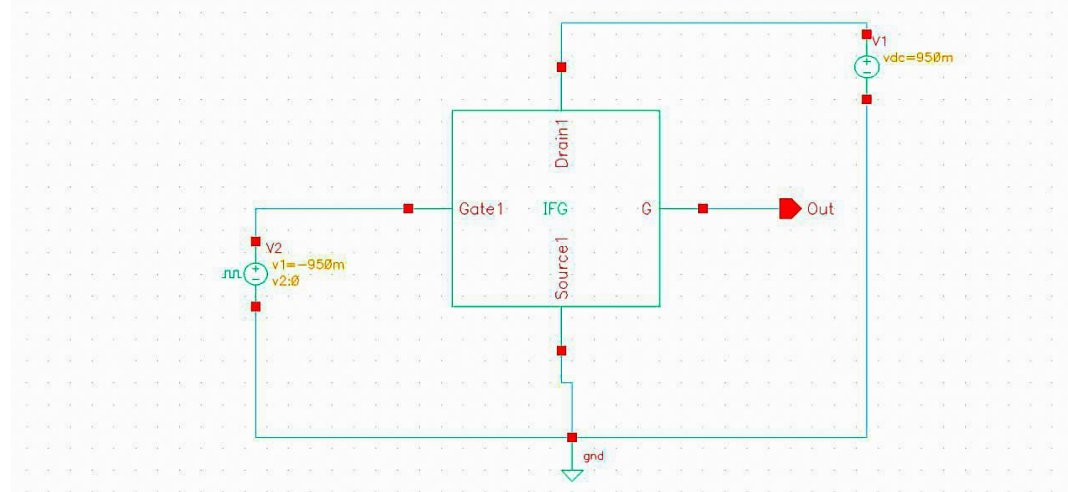


Figure 4. IFG model equivalent circuit. In the IFG model, the voltage applied at the "Gate" terminal is 'V_w' which is the write voltage, and at the "Drain" terminal is 'V_r' which is the read voltage

11

The voltage applied at the gate node must be greater than the threshold voltage (V_{th}). The voltage level should be ± 0.95 v voltage which 12
is given to the gate and source node. 13

14
15

2.2 Weight setting through gradient descent model and Backpropagation

Specific learning algorithms are useful to fasten the training of neural networks. Here we adopt the Gradient descent model and backpropagation along with memristor-based IFG circuit. It takes less than a few seconds to run the ANN[1][16]. A text-based “Spectra” is generated for the connections in the ANN circuit with help of python, the input for this python code is the database from cadence simulation (Synapse)[2][17]. First normal weighted synapse is given from output of the IFG model to neural networks, after getting the optimized values from the neural networks, again these optimized values should be applied to IFG and observe whether our activation function is optimized. Here the weights are optimized by using gradient descent and backpropagation models. The differences are backpropagated to the neural networks. At epoch 0.68 the accuracy increases by 0.8% as compared with the traditional method. The circuit-level design and implementation of gradient and backpropagation learning architectures help to get optimized results. These results are compared with the original one. The below given (5) circuit shows the simple operation of ANN.

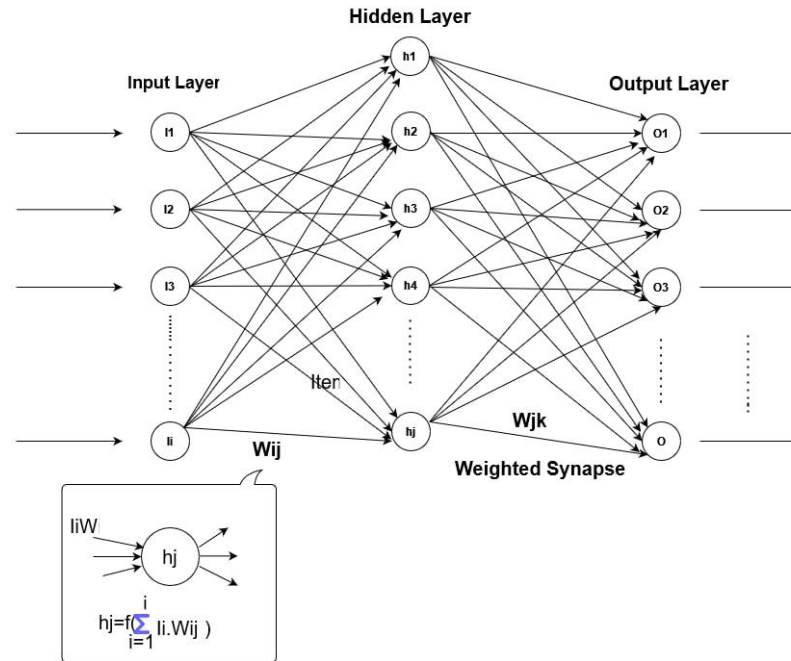


Figure 5. A simple circuit affirmation from the input layer to the output layer

To achieve Image pattern recognition, the optimized values are again given as weighted synapses in the neural networks along with the images as input. Here ReLu and Adams activation functions are used. W_{ij} are weighted neurons, h_j is the summation of input neurons and weights. The summation is fed to the activation function, which is shown in (6).

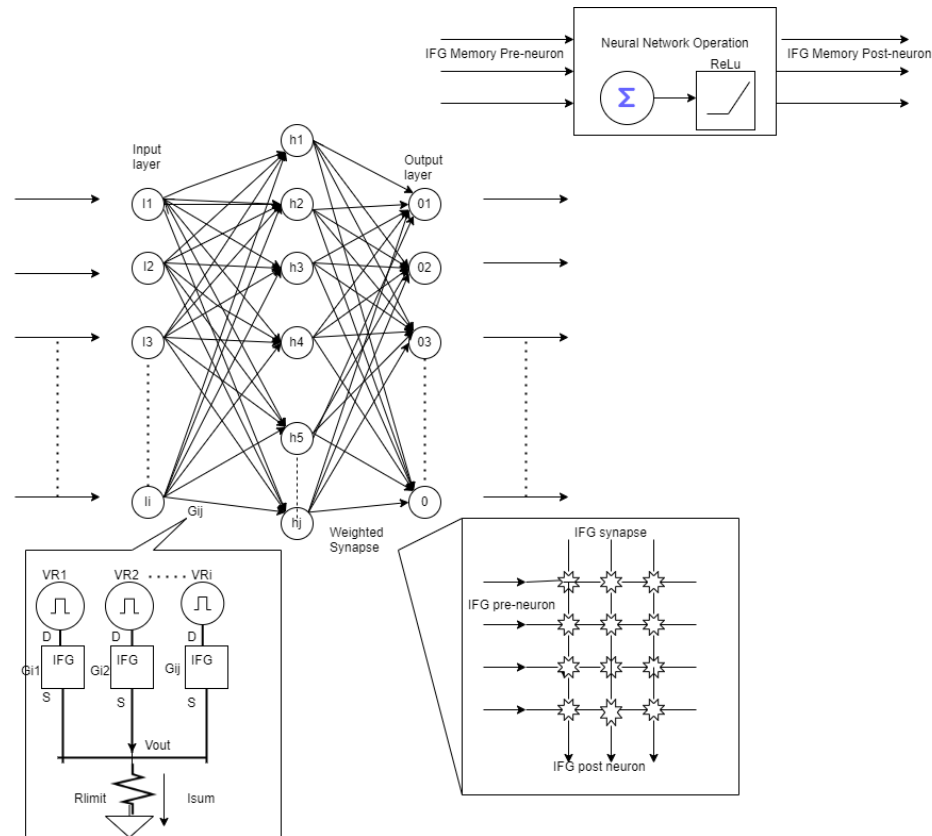


Figure 6. Operation of neural networks with IFG memory synapse and implementing image processing using gradient descent algorithm and backpropagation

3. Results and Analysis

To check basic memristor characteristics the sine function is given as input harmonic with voltage (V) of 1V amplitude and 1 Hz frequency ($V(t) = V_0 \sin(\omega t)$ for $\omega = 2\pi f$), the V-I characteristics shows for the negative voltage value the current increase, and for the positive voltage value the current drops resulting in the formation of a hysteresis loop and also, we can see the linear relationship between charge and flux.

1
2
3
4
5
6

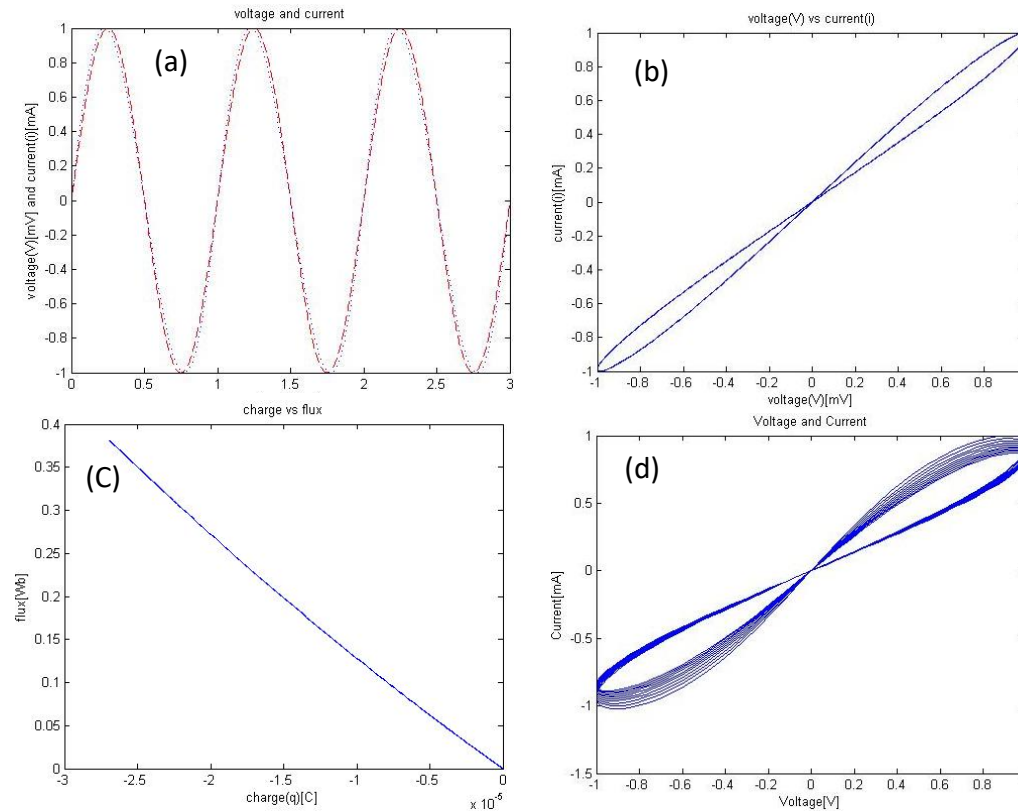


Figure 7. MATLAB simulation results of basic Memristor. (a) Input sinewave signal (b) V-I characteristics of Memristor (c) Linear relation between charge and flux (d) Pinched hysteresis loop (Non-Linear characteristics)

In the below-given result (8), for a positive pulse, the conductance values decrease and for negative values, the conductance values increase as like memristor characteristics. These are the simulation results of the IFG model in cadence for which the voltage value is ± 0.95 V and the threshold voltage is 0.4 V. The voltage value is given in between the gate and source i.e., input pulse, threshold voltage is given to the dc source which is connected in between drain and source.

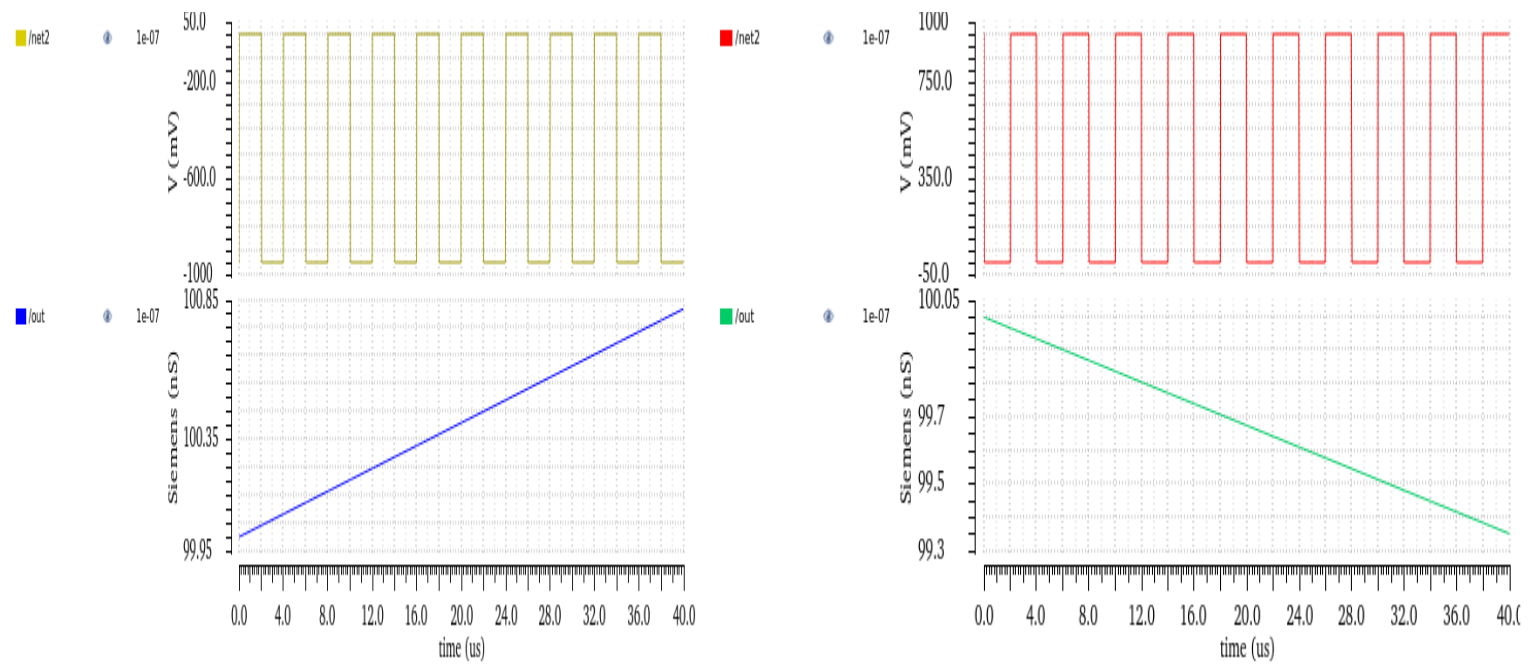


Figure 8. Two sets of graphs, interactive 4 is the positive pulse response and interactive 7 is the negative pulse response

1
2
3
4
5
6
7
8

Below given figure (9) shows the post-simulation values for the optimized values from the IFG model

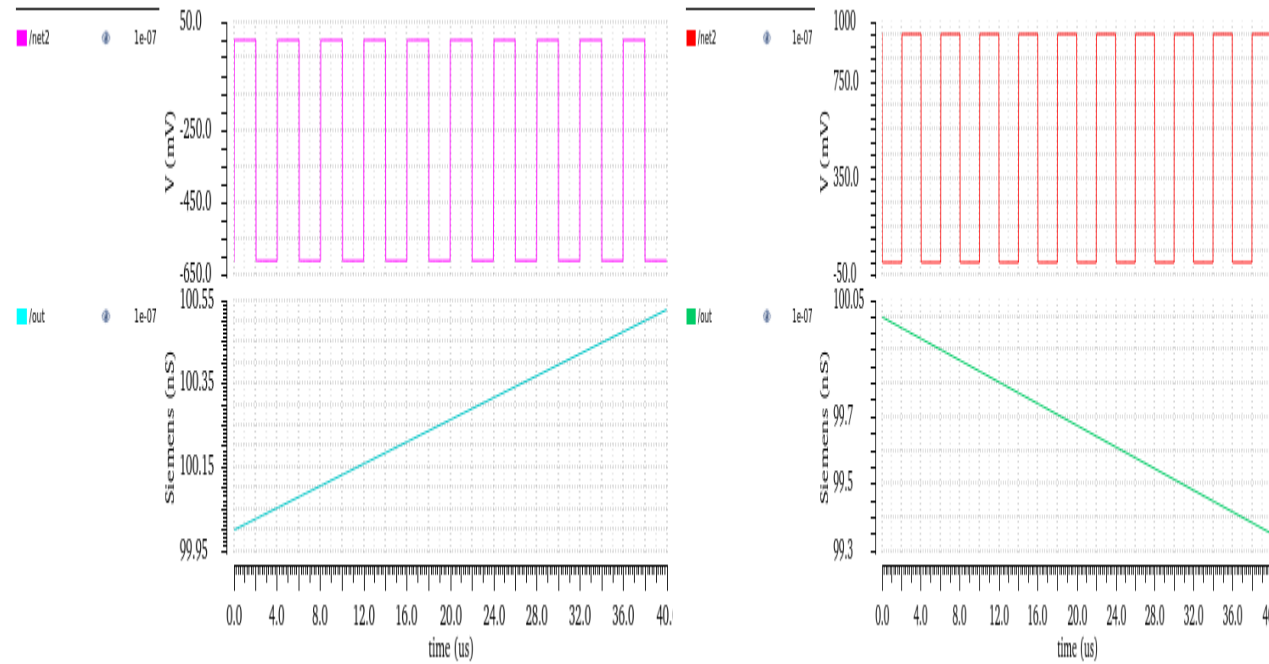


Figure 9. Post-simulation results for optimized parameters

After applying the optimized values to the IFG device, the same operation is repeated as in previous case but with increased accuracy by 0.8%. Here the optimized parameters are $X = 0.150$ V and $Y = -0.612$ V.

Table 1. Parametric values of the current, voltage, flux and charge for Hysteresis Curve w.r.t figure (1)

1

Measurement #	Input Voltage (V)	Input Current (I)	Flux(Ψ)	Charge(q)
1.	0.018mV	0.022mA	3.3×10^{-9} (Wb)	-2.8×10^{-9} C
2.	0.037mV	0.044mA	1.3×10^{-4} (Wb)	-1.1×10^{-8} C
3.	0.056mV	0.066mA	1.5×10^{-4} (Wb)	-2.5×10^{-8} C
4.	0.075mV	0.088mA	1.8×10^{-4} (Wb)	-4.5×10^{-8} C

Table 2. Parametric value analysis when the input pulse response=(+0.95V)

2

Measurement #	Input Voltage (V)	Siemens (ns)	
		Gmax=1e+07(X)	Gmax=1e+07(Y)
1.	0.95 V	0	1.0×10^{-7}
2.	0.60 V	2.0×10^{-8}	1.0×10^{-7}
3.	0.31 V	3.0×10^{-7}	1.0×10^{-7}
4.	0.19 V	4.02×10^{-6}	9.9×10^{-8}
5.	0	6.02×10^{-6}	9.9×10^{-8}

Table 3. Parametric value analysis when the input pulse response=(-0.95V)

3

Measurement #	Input Voltage (V)	Siemens (ns)	
		Gmax=1e-07 (X)	Gmax=1e-07(Y)
1.	0	0	1.0×10^{-7}
2.	-0.19 V	2.0×10^{-8}	1.0×10^{-7}
3.	-0.31 V	3.0×10^{-7}	1.0×10^{-7}
4.	-0.60 V	4.0×10^{-6}	1.0×10^{-7}
5.	-0.95 V	6.0×10^{-6}	1.0×10^{-7}

Here in table (1) and (2) mentioned some values of how current varies with a change in voltage and its sign also the same hysteresis characteristics of both the memristor and IFG model. The various parameters stated here for basic memristor are current (I), Voltage (V), Flux(Ψ), and charge (q), for the IFG model the Gmax (conductance) values are given as a relationship between voltage and siemens.

Table 4. Comparison between analytical values and optimized values

Model #	Activation function	Accuracy
Traditional model	Sigmoid function	93.8
Proposed model	ReLu function and Adams function	94.6

By observing the above table (4), after giving the optimized parameters from IFG to the neural networks, now by applying the Adams activation function the accuracy of recognition is increased and updated by 0.8%. So with proposed IFG model the accuracy increases.

4. Conclusion

Artificial neural networks place a vital role in unsupervised deep learning models to implement application based on pattern recognition and also in many ares of our daiy life. This paper provides us knowledge of developing an IFG device using cadence with the characteristics of memritior based circuit. To demonstrate the capability of IFG device in pattern recognition the values are optimized by gradient descent model and the resulted optimized parameters are compared with the existing parameters. The testing ability of neural network increases by using IFG model by 0.8%. Hence accuracy of original network is 93.8 % which is later increaases to 94.6%. Here ReLu and Adams activation helps in faster optimization and by this result we can conclude by using IFG-based memristor networks in neural networks accuracy will be increase precisely. Images are accurately recognized.

References

- [1] Creswell, "UC Santa Cruz UC Santa Cruz Electronic Theses and Dissertations Title," 2020. [Online]. Available: <https://escholarship.org/uc/item/0jx2107r>
- [2] Dindha Amelia, "Application of memristive device arrays for pattern recognition-Title," 2020. [Online]. Available: <http://mpoc.org.my/malaysian-palm-oil-industry/>
- [3] W. Zhang *et al.*, "Hardware-Friendly Stochastic and Adaptive Learning in Memristor Convolutional Neural Networks," vol. 2100041, 2021, doi: 10.1002/aisy.202100041.

-
- [4] O. Krestinskaya, K. N. Salama, and A. P. James, "Learning in memristive neural network architectures using analog backpropagation circuits," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 66, no. 2, pp. 719–732, 2019, doi: 10.1109/TCSI.2018.2866510. 1
2
- [5] O. Krestinskaya, T. Ibrayev, and A. P. James, "Hierarchical Temporal Memory Features with Memristor Logic Circuits for Pattern Recognition," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 37, no. 6, pp. 1143–1156, 2018, doi: 10.1109/TCAD.2017.2748024. 3
4
- [6] M. Chu *et al.*, "Neuromorphic Hardware System for Visual Pattern Recognition with Memristor Array and CMOS Neuron," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2410–2419, 2015, doi: 10.1109/TIE.2014.2356439. 5
6
- [7] A. Aginsbte and T. Memristor, "Synaptic Characteristics of," vol. 64, no. 4, pp. 1–6, 2017. 7
- [8] R. Yan, Q. Hong, C. Wang, J. Sun, and Y. Li, "Multilayer Memristive Neural Network Circuit Based on Online Learning for License Plate Detection," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 41, no. 9, pp. 3000–3011, 2022, doi: 10.1109/TCAD.2021.3121347. 8
9
- [9] B. Abdoli, A. Amirsoleimani, J. Shamsi, K. Mohammadi, and A. Ahmadi, "A Novel CMOS-Memristor Based Inverter Circuit Design," no. May, 2014, doi: 10.1109/IranianCEE.2014.6999567. 10
11
- [10] M. R. Azghadi, B. Linares-Barranco, D. Abbott, and P. H. W. Leong, "A Hybrid CMOS-Memristor Neuromorphic Synapse," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 2, pp. 434–445, 2017, doi: 10.1109/TBCAS.2016.2618351. 12
13
- [11] D. Querlioz, O. Bichler, P. Dollfus, and C. Gamrat, "Immunity to device variations in a spiking neural network with memristive nanodevices," *IEEE Trans. Nanotechnol.*, vol. 12, no. 3, pp. 288–295, 2013, doi: 10.1109/TNANO.2013.2250995. 14
15
- [12] H. Ran *et al.*, "Memristor-Based Edge Computing of Blaze Block for Image Recognition," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 5, pp. 2121–2131, 2022, doi: 10.1109/TNNLS.2020.3045029. 16
17
- [13] X. Xu, W. Xu, B. Wei, and F. Hu, "Memristor-based neural network circuit of delay and simultaneous conditioning," *IEEE Access*, vol. 9, pp. 148933–148947, 2021, doi: 10.1109/ACCESS.2021.3122973. 18
19
- [14] Y. Zhang, X. Wang, Y. Li, and E. G. Friedman, "Memristive Model for Synaptic Circuits," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 64, no. 7, pp. 767–771, 2017, doi: 10.1109/TCSII.2016.2605069. 20
21
- [15] Q. Duan *et al.*, "Spiking neurons with spatiotemporal dynamics and gain modulation for monolithically integrated memristive neural networks," *Nat. Commun.*, vol. 11, no. 1, pp. 1–13, 2020, doi: 10.1038/s41467-020-17215-3. 22
23
- [16] D. Sacchetto, P. E. Gaillardon, M. Zervas, S. Carrara, G. De Micheli, and Y. Leblebici, "Applications of multi-terminal memristive devices: A review," *IEEE Circuits Syst. Mag.*, vol. 13, no. 2, pp. 23–41, 2013, doi: 10.1109/MCAS.2013.2256258. 24
25
- [17] S. N. Truong, "Single Crossbar Array of Memristors with Bipolar Inputs for Neuromorphic Image Recognition," *IEEE Access*, vol. 8, pp. 69327–69332, 2020, doi: 10.1109/ACCESS.2020.2986513. 26
27
28