ENHANCING THE ASTROCYTE SPIKING DEEP NEURAL NETWORKS AND CALCIUM BASED NEURAL NETWORK MODEL FOR CLASSIFICATION

BY

BASSAM ABDUL-RAHMAN ABED

A thesis submitted in fulfilment of the requirement for the
degree of Doctor of Philosophy in Computer Science

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

DECEMBER 2020

ABSTRACT

The ongoing challenge in machine learning is to advance general and biological inspired artificial models of neural networks which are compatible with the spatial and temporal constraints of the brain. For instance, a new term has recently been emerged to describe the communication between two neurons and a single astrocyte, the other type of cells in the brain, called tripartite synapse. The communication between astrocytes and astrocytes in a network called astrocytic syncytium is explored. The calcium dynamics in the brain is also considered as a significant player in brain information processing. Therefore, the study proposes to generalize mathematical models for tripartite synapse and astrocytic syncytium to advance new Tripartite Synapse Model (TSM), Artificial Astrocytic Syncytium (AAS) model and Calcium Based Artificial Neural Network (caANN) to mimic the calcium dynamics in the brain. Moreover, the study utilizes the proposed models of TSM and AAS within the architecture of the deep neural networks such as convolutional neural networks and deep belief neural networks. The simulation results of incorporating the real astrocyte in spiking response model (RSM) and Recurrent-Simple Neural Network (RSNN) has shown that astrocyte increases the postsynaptic potential and in turn, improves the RSM. Besides, the simulations of TSM in two neuron models, leaky integrate and fire (LIF) and Izhikevich model, has shown that using TSM has changed the spiking behavior (rate and firing pattern) of these models. The simulation related to the AAS by probability distribution and K-L divergence comparison has shown that the gap junction channels can be opened by the higher probability astrocytes. Finally, the result of the simulations of TSM in CNN and DBN showed that incorporating astrocytes' dynamics, properties and roles in deep learning networks compete and sometimes outperform the standard architectures of deep neural networks in terms of training and validation accuracy.

خلاصة البحث

يتمثل التحدي المستمر في التعلم الآلي في تطوير النماذج الاصطناعية العامة والمستوحاة من الشبكات العصبية المتوافقة مع القيود المكانية والزمنية للدماغ. على سبيل المثال ، ظهر مؤخراً مصطلح جديد لوصف الاتصال بين اثنين من الخلايا العصبية وخلايا نجمية واحدة ، وهو النوع الآخر من الخلايا في الدماغ ، يسمى المشبك الثلاثي. بالإضافة إلى ذلك ، الاتصال بين الخلايا النجمية والخلايا النجمية في شبكة تسمى الخلايا النجمية. وأخيرًا ، تم اعتبار ديناميكيات الكالسيوم في الدماغ لاعبًا مهمًا في معالجة معلومات الدماغ. لذلك ، نقترح تعميم النماذج الرياضية للمشابك الثلاثي والمزامنة النجمية لتطوير نموذج المشبك الثلاثي الجديد (TSM) ، ونموذج Synocytic Astrocytic الاصطناعي (AAS) والشبكة العصبية الاصطناعية القائمة على الكالسيوم (CaANN) لمحاكاة ديناميكيات الكالسيوم في الدماغ. علاوة على ذلك ، استخدمنا النماذج المقترحة من TSM و AAS داخل بنية الشبكات العصبية العميقة مثل الشبكات العصبية التلافيفية والشبكات العصبية العميقة. أظهرت نتائج المحاكاة لدمج الخلايا النجمية الحقيقية في نموذج استجابة الطفرة (RSM) والشبكة العصبية المتكررة البسيطة (RSNN) أن الخلايا النجمية تزيد من إمكانات ما بعد المشبكية وبالتالي تحسن RSM. بالإضافة إلى ذلك ، أظهرت محاكاة TSM في نموذجين من الخلايا العصبية ، الاندماج المتسرب والنار (LIF) ونموذج Izhikevich ، أن استخدام TSM قد غيّر سلوك السنبلة (معدل ونمط الإطلاق) لهذه النماذج. أظهرت المحاكاة المتعلقة بـ AAS من خلال التوزيع الاحتمالي ومقارنة الاختلاف بين K-L أنه يمكن فتح قنوات تقاطع الفجوة بواسطة الخلايا النجمية ذات الاحتمالية الأعلى. وأخيرًا ، أظهرت نتائج المحاكاة لـ TSM في CNN و DBN أن دمج ديناميكيات الخلايا النجمية وخصائصها وأدوارها في شبكات التعلم العميق تتنافس وتتفوق أحيانًا على البنى القياسية للشبكات العصبية العميقة من حيث التدريب ودقة التحقق.

.

APPROVAL PAGE

The thesis of Bassam Abdul-Rahman Abed has been approved by the following:

Amelia Ritahani Bt. Ismail
Supervisor

Normaziah Binti Abdul Aziz
Co-Supervisor

Imad Fakhri Alshaikhli
Internal Examiner

Azlan Mohd Zain
External Examiner

Radwan Jamal Elatrash
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Bassam Abdul-Rahman Abed

Signature Date 11/ 12 / 2020

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

ENHANCE THE ASTROCYTE SPIKING DEEP NEURAL NETWORKS AND CALCIUM BASED NEURAL NETWORK MODEL FOR CLASSIFICATION

I declare that the copyright holders of this thesis are jointly owned by the student and IIUM.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.

3. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Bassam Abdul-Rahman Abed

	11 /12 / 2020
Signature	Date

ACKNOWLEDGEMENTS

Foremost, I am grateful to Allah for the good health that was necessary to accomplish this work. I would like to dedicate this work to the soul of my father, my mother, my wife and my children Abdulrahman and Jad; thank you for your support and patience.

I would like to express my sincere gratitude to my friends especially Mohammad El-Jammal who provided his time, effort and support for this thesis, thank you for sticking with me.

Finally, a special thanks to Professor Amelia Ritahani Ismail for her aspiring guidance, indispensable support, friendly encouragement and constructive leadership during this research period, and for that, I will be forever grateful.

TABLE OF CONTENTS

LIST OF TABLES

Table No. Page No.

LIST OF FIGURES

LIST OF ABBREVIATIONS

2-AG	Endocannabinoid 2-Arachidonyl Glycerol
[Ca2+]	Intracellular Calcium Concentration
AAS	Artificial Astrocytic Syncytium
ACNN	Astrocyte-CNN
ADBN	Astrocyte-DBN
AI	Artificial Intelligence
ANGN	Artificial Neuron-Glial Network
ANN	Artificial Neural Networks
AP	Action Potentials
ATP	Adenosine triphosphate
ca^{2+}	Calcium
CaANN	Calcium-based ANN
CB1R	Cannabinoid Receptor
CD	contrastive divergence
CICR	Calcium Induced Calcium Release
CNN	Convolutional Neural Networks
CNS	Central Nervous System
DBN	Deep Belief Network
DMS	Destexhe Model of Synapse
DNN	Deep Neural Network
DSE	Depolarization-induced Suppression of Excitation
EPSP	Excitatory Post Synaptic Potential
ER	Endoplasmic-Reticulum
FDF	Fire-Diffuse-Fire
GA	Genetic Algorithm
GLU	Glutamate
GMM	Gaussian Mixture Model
GMDH	Group Method of Data Handling
GPU	Graphics Processing Unit
IICR	IP3 Induced Calcium Release
InsP3Rs	inositol 1,4,5-trisphosphate receptors
IP3	Inositol trisPhosphate
IP3R	Inositol trisPhosphate Receptor
IPSP	Inhibitory Post Synaptic Potential
HH	Hodgkin–Huxley
K+	Potassium
KL	Kullback-Leibler
LIF	Leaky-Integrate-and-Fire
LTD	Long Time Potentiation
LSM	Liquid State Machine
LSTM	Long Short-Term Memory
LTP	Long Time Depression
mGluRs	Metabotropic Glutamate Receptors
MNIST	Modified National Institute of Standards and Technology database
MLE	Maximum Likelihood Estimation
MLP	Multi-Layer Perceptron
MSE	Mean Square Error
SGD	Stochastic Gradient descent

Na+	Sodium
NLP	Natural Language Processing
NMDA	N-methyl-D-aspartate
PDF	Joint Probability Distribution
PSP	Postsynaptic Potential
PR	Probability of Release
RBM	Restricted Boltzmann Machine
ReLU	Rectified linear unit
RNN	Recurrent Neural Network
SERCA	Sarco- Endoplasmic-Reticulum Ca2+-ATPase
SICs	Slow Inward Currents
SNAN	Spiking Neuron-Astrocyte Network
SNN	Spiking Neural Network
STDP	Spike Timing Dependent Plasticity
SOM	Self-Organizing Map
SRM	Simple Spike Response Model
TSM	Tripartite Synapse Model
VB	Variational Bayesian
VOCs	Voltage-Operated Channels
Wm	Weight Limit

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

Artificial Neural Networks (ANNs) are powerful tools to perform modelling and to solve non-linear problems inspired by the biological neural networks that compose brains. Indeed, ANNs have been used on sophisticated tasks, such as classification, prediction, optimization, and regression. However, McCulloch & Pitts (1943) proposed the first computational model for the ANNs called threshold logic. In recent years, Deep Neural Networks (DNN), such as Convolutional Neural Networks (CNN) and Deep Belief Networks (DBNs) have substantially achieved state-of-the-art results in numerous contests in machine learning with relevant real-world tasks such as image recognition, and voice recognition. DNN architectures are utilizing abstract features biologically inspired from the deep hierarchical arrangement of the visual cortex which is the primary cortical region of the brain that receives, integrates, and processes visual information. Hence, the brain is organized in a deep architecture with multiple levels of abstraction, each level equivalent to a distinct area of cortex.

DNN can build hierarchical representations from unlabelled data by feature detector units in their deep hierarchal layers where the lower layers detect simple features and feed them into higher layers which in turn detect more complex features and so forth. However, the foremost DNN of the Multi-Layer Perceptron (MLP) called the Group Method of Data Handling (GMDH) was proposed by Ivakhnenko (1965). The GMDH contained eight hidden layers with polynomial activation function learned by regression. After a while, Ivakhnenko (1971) introduced some theoretical proposals for learning DNNs with Stochastic Gradient Descent (SGD).

Astrocytes, the predominant glial cell type in the brain, were traditionally considered as merely passive supportive cells without any important roles in synaptic information processing. In contrast, the contemporary view has given rise to show that astrocytes play active roles such as synaptic neurotransmission (Fellin & Carmignoto, 2004; Fields & Stevens-Graham, 2002; Perea & Araque, 2009), information processing (Newman, 2003; Perea, Sur, & Araque, 2014; Santello & Volterra, 2009; Verkhratsky, Rodríguez, & Parpura, 2012), Long Time Potentiation / Depression (LTP/LTD) (Henneberger et al., 2010), Short Term Potentiation/Depression (STDP) (Wade et al., 2011) and learning or synaptic plasticity (Araque, 2008; Barker & Ullian, 2010; Haydon & Nedergaard, 2016; Parpura et al., 1994; Schummers, Yu, & Sur, 2008).

Information processing in the Central Nervous System (CNS) is usually determined by the electrical excitability nature of neurons, their impact on the synaptic activities and their ability to release chemical neurotransmitters. Likewise, astrocytes show some of the excitability by dynamic calcium (ca^{2+}) signalling (Charles et al., 1991; Cornell-Bell et al., 1990; Dani, Chernjavsky, & Smith, 1992), they are involved in synaptic activities which have solely been considered as neuron's responsibilities (Fellin, 2009; Halassa, Fellin, & Haydon, 2009; Volterra, Liaudet, & Savtchouk, 2014), and finally they release gliotransmitter in response to calcium (ca^{2+}) elevations (Parpura et al., 1994; Ventura & Harris, 1999). Therefore, two concepts including the Artificial Neuron-Glial Networks (ANGNs) and Spiking Neuron-Astrocyte Network (SNAN) have come forth to incorporate astrocyte in conventional ANN and Spiking Neural Network (SNN), respectively.

However, the current study proposes mathematical models based on kinetic dynamics to mimic the communication between astrocytes-to-neurons (*tripartite synapse*) and astrocytes-to-astrocytes (*astrocytic syncytium*). The study further suggests the implementation of these biological inspired mathematical models in the state-of-the-art DNN for both architectures CNN and DBN. The receptive waves of calcium release from intracellular stores are often termed “calcium spikes” (Skupin et al., 2008) or “spatial spikes” by analogy to neuronal spike trains (Deitmer, 1998; Navarrete et al., 2012). Moreover, calcium spikes are reminiscent of neuronal Action Potentials (APs) and their importance for intracellular communication is predicted to be equivalent to the importance of communication AP in CNS (Jaffe, 1993). Calcium spikes are considered as Aps to enable cells to use digital logic in transducing signals (Navarrete et al., 2012). By encoding the signal in temporal and spatial patterns, information is often transmitted as a calcium wave which is much the same as using the electric voltage or current signals in information technology. Falcke (2004) demonstrated that the calcium dynamics are another illustration for the flexible applicability of the FitzHugh–Nagumo model, and that calcium dynamics exceed this illustration and quite similar to the approach which looks for progress in the theory of pattern formation and nonlinear systems. Therefore, to emulate the intracellular-to-intercellular calcium signalling, a novel model called calcium-based ANN (caANN) based on Fire-Diffuse-Fire (FDF) model (Dawson, Keizer, & Pearson, 1999), calcium toolkit (Berridge, Lipp, & Bootman, 2000) and polarization property of transverse waves that specifies the geometrical orientation of the oscillations was proposed.

1.2 STATEMENT OF THE PROBLEM

Artificial Neural Networks are very important and have been used in many applications to solve problems such as classification but are still evolving based on recent discoveries in neurobiology to find artificial networks that imitate the biological networks in the brain. In spite of the fact that artificial neuron (or perceptron) was discovered in the 1950s to mimic the biological processes of neurons such as the certain components of the neuron (dendrites, cell bodies and axons) and despite that the artificial neural networks were inspired by the connections between biological neurons based on simplified mathematical models that had powerful performance in some classification tasks, these artificial networks, however, neglected the creation of the connection, destruction of the connection and timing of the signals (Pfeiffer & Pfeil, 2018). SNN incorporated the concept of time in their architectures in which the time of firing is important when neurons reach a specific threshold. Recent concepts have been emerged to represent the roles of astrocytes (one type of glial cells) in brain information processing and the significance to incorporate such roles in the architectures of conventional ANN called Artificial Neural Glial Networks (ANGNs) and in SNN called Spiking Neuron-Astrocyte Networks (SNANs). Nevertheless, ANGNs and SNANs have consolidated only one role of astrocytes and added them in last layer (output layer) of their architectures and have not included the astrocytic network nor the calcium dynamics in the brain.

On the other hand, the DNNs have intimated outstanding success in classification and recognition problems as a result of the availability of Graphics Processing Unit (GPU) and the large-scale datasets for training. Ferré, Mamalet, and Thorpe (2018) claimed that DNNs suffer from the lack of biological plausibility. Bengio et al., (2015) discussed that Boltzmann machines are the most biologically plausible learning algorithms for deep architectures, but they also encounter several biological implausibility such as the weight transport problem to attain symmetric weights, and the positive-phase vs negative-phase synchronization (Diehl et al., 2015; Fatahi et al., 2018; Neil, Pfeiffer & Lie, 2016; Panda & Roy, 2016). The top layer of CNNs is trained by the supervised back-propagation algorithm which is biological implausible (Kheradpisheh et al., 2018). Moreover, implementation of DNN in conventional digital platforms suffers from the substantial cost of training and resource-consuming process (Diehl et al., 2015; Fatahi et al., 2018; Neil, Pfeiffer, & Liu, 2016; Tavanaei et al., 2018). As a consequence, SNNs have been introduced to incorporate in DNN's architectures in different ways to solve such shortcomings. For instance, CNN architectures were converted into SNNs (Esser et al., 2016; Rueckauer et al., 2017), the standard version of SNN learning rule called Spike Timing Dependent Plasticity (STDP) was converted to convolutional version with Leaky-Integrate-and-Fire (LIF) neurons in SNNs (Kheradpisheh et al., 2018; Lee et al., 2018; Panda, Srinivasan, & Roy, 2017), deep spiking MLP network was proposed by O'Connor and Welling (2016), spiking Restricted Boltzmann Machine (RBM) (Neftci, Augustine, Paul, & Detorakis, 2017), Spiking DBN was introduced by Stromatias et al. (2015), and finally spiking deep networks with LIF Neurons was presented by Hunsberger and Eliasmith (2015). In conclusion, despite that DNN and spiking DNN are biologically plausible architecture in terms of the hierarchy representation of the visual cortex, DNN nor spiking DNN overlooked the functions of other brain structures or the other cell types such as glial cells in imitation. Moreover, calcium spikes are reminiscent of neuronal APs and their importance for intracellular communication is predicted to be equal to the importance of communication AP in CNS (Jaffe, 1993).

In summary, ANN neglected the concept of timing and the roles of other cell types of the brain, astrocytes, and their network-syncytium. Hence, SNAN incorporated the concept of timing and consolidated only one role of astrocyte which is the self-repair of the neuron network and ignored the astrocytic network in their architectures. ANGN utilized the feed-forward ANN and added astrocyte to their layers but ignored the concept of tripartite synapse nor the astrocytic syncytium in their architectures. Spiking

DNNs finally included the concept of timing but ignored the roles of astrocytes nor their network to introduce more biological inspiring networks.

1.3 RESEARCH QUESTIONS

1- What are the architectures of the artificial neural networks that are incorporated, simulated and modelled by astrocytes and their networks?

2- What is the role of astrocytes and calcium dynamics in brain information processing in terms of tripartite synapse and astrocytic syncytium?

3- What are the existing mathematical models that could be utilized in the proposed models for the roles of astrocytes, astrocytic syncytium and calcium dynamics?

4- How can the spiking neural networks and simple recurrent network be utilized to propose new models for real-time astrocytes, the kinetic models of a single synapse to create a new model of tripartite synapse, and both the tripartite synapse model and the kinetic model of multiple synapses to create a new model for astrocytic syncytium?

5- How can the fire-diffuse-fire model, the idea of polarization as the receptive field, the hierarchy of calcium signalling and the calcium signalling toolkit be used to create a novel model of calcium-based ANN (caANN)?

6- How can the TSM and the AAS be utilized to create a new model of astrocytic CNN (ACNN) and astrocytic DBN (ADBN)?

7- What are the results of simulations of ACNN and ADBN in handwritten digit recognition problem in the MNIST database?