

**STACKED CAPSULE AUTOENCODER BASED  
GENERATIVE ADVERSARIAL NETWORK**

Galagama Arachchige Chatura. Madhusanka

(189386G)

Degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa

Sri Lanka

October 2020

# **STACKED CAPSULE AUTOENCODER BASED GENERATIVE ADVERSARIAL NETWORK**

Galagama Arachchige Chatura. Madhusanka

(189386G)

Thesis/Dissertation submitted in partial fulfilment of the requirements for the degree of  
Master of Science in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa  
Sri Lanka

October 2020

# Declaration

I declare that this dissertation does not incorporate, without acknowledgment, any material previously submitted for a Degree or a Diploma in any University and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organization.

Name of Student

G.A.C. Madhusanka

Signature of Student

Date:

Supervised by

Dr. K.S.D. Fernando

Signature of Supervisor

Date:

# **Dedication**

I dedicate this thesis to my parents and all my teachers.

# **Acknowledgement**

I acknowledge my supervisor Dr. K.S.D. Fernando for her insight to make my research successful. Without the support from supervisor it would be hard to complete this project. I thank for her advices and techniques which helped me a lot in the process of developing this research.

Also I should thank to all the staff members of Department of Computational Mathematics including lecturers and non-academic staff.

Again I must thank to family members including my mother and farther for their huge support.

# Abstract

Convolutional neural network based generative adversarial networks have become the dominant generative model in the field of generative deep learning. But limitations of convolutional neural networks affect generative adversarial networks also, since most of the current generative adversarial networks are based on convolutional neural networks. The main limitation of convolutional neural networks is that they are invariant. In other words, convolutional neural networks can't preserve spatial information of features in an image. In contrast, capsule networks gained attention in recent years due to their equivariant architecture which preserves spatial information.

Stacked capsule autoencoder is a type of capsule networks that is able to overcome the limitations that convolutional neural networks suffer from. Stacked capsule autoencoder is an equivariant model which preserves spatial, relational, geometrical information between parts and objects in an image. So in this research we implemented a generative adversarial network which uses stacked capsule autoencoder as the discriminator of it, by replacing the conventional convolutional neural network discriminator.

Then we evaluated the implementation of stacked capsule autoencoder based generative adversarial network using MNIST images. As the qualitative evaluation we observed the visual quality of generated images. Quality and diversity of the generated images are acceptable. Then we evaluated our model quantitatively using inception score for MNIST. Findings of this research show that, the stacked capsule autoencoder can be used as the discriminator of a generative adversarial network instead a convolutional neural network and its performances are plausible.

# Table of Contents

	<b>Page</b>
<b>Chapter 1 – Introduction</b>	<b>1</b>
1.1 Prolegomena	1
1.2 Aim and Objectives	2
1.3 Background and Motivation	2
1.4 Problem in Brief	3
1.5 Stacked Capsule Autoencoder based GAN	3
1.6 Structure of the thesis	4
1.7 Summary	4
<b>Chapter 2 – Literature Review</b>	<b>5</b>
2.1 Introduction	5
2.2 Deep Learning	5
2.3 Generative models	6
2.4 Generative Adversarial Networks	7
2.5 Capsule Networks	8
2.6 Limitations of CNNs	9

2.7	Capsule Generative Adversarial Networks	10
2.8	Problem definition	12
2.9	Summary	13
<b>Chapter 3 – Technology Adapted</b>		<b>14</b>
3.1	Introduction	14
3.2	Architecture of Generative Adversarial Networks	14
3.3	Deep Convolutional GAN	15
3.4	Architecture of Capsule Networks	15
3.5	Stacked Capsule Autoencoders	16
3.6	Summary	19
<b>Chapter 4 – Approach</b>		<b>20</b>
4.1	Introduction	20
4.2	Hypothesis	20
4.3	Inputs to the system	20
4.4	Outputs from the system	21
4.5	Process	21
4.6	Features	21
4.7	Users	22



4.8	Summary	22
-----	---------	----

<b>Chapter 5 – Design</b>	<b>23</b>
---------------------------	-----------

5.1	Introduction	23
-----	--------------	----

5.2	Architecture of proposed GAN	23
-----	------------------------------	----

5.3	Architecture of generator	25
-----	---------------------------	----

5.4	Architecture of discriminator	26
-----	-------------------------------	----

5.5	Part Capsule Autoencoder	27
-----	--------------------------	----

5.6	Object Capsule Autoencoder	28
-----	----------------------------	----

5.7	Summary	29
-----	---------	----

<b>Chapter 6 – Implementation</b>	<b>30</b>
-----------------------------------	-----------

6.1	Introduction	30
-----	--------------	----

6.2	Dataset preparation	30
-----	---------------------	----

6.3	Generator implementation	31
-----	--------------------------	----

6.4	Discriminator implementation	32
-----	------------------------------	----

6.5	Part Capsule Autoencoder implementation	32
-----	---	----

6.6	Object Capsule Autoencoder implementation	33
-----	---	----

6.7	Training of GAN	34
-----	-----------------	----

6.8	Summary	35
<b>Chapter 7 – Evaluation</b>		<b>36</b>
7.1	Introduction	36
7.2	Evaluation procedure	36
7.3	Visual quality evaluation of generated images	36
7.4	Losses of the GAN	37
7.5	Inception Score for MNIST	39
7.6	Summary	40
<b>Chapter 8 – Conclusion</b>		<b>41</b>
8.1	Introduction	41
8.2	Conclusion	41
8.3	Limitations and Future Work	42
8.4	Summary	43
<b>References</b>		<b>44</b>
<b>Appendix A Codes</b>		<b>47</b>
<b>Appendix B Results</b>		<b>51</b>

# List of Figures

	<b>Page</b>
Figure 2.1: CNN can't differentiate these two images ...	9
Figure 3.1: DCGAN architecture	15
Figure 3.2: Architecture of capsule networks.	16
Figure 3.3: SCAEs can learn object classes ...	17
Figure 3.4: At PCAE stage it learns parts and poses ...	18
Figure 4.1: Sample of MNIST data	21
Figure 5.1: High level design diagram of proposed GAN	24
Figure 5.2: Architecture of Generator	26
Figure 5.3: SCAE discriminator architecture	28
Figure 7.1: Generated MNIST images	37
Figure 7.2: Generator loss	38
Figure 7.3: Discriminator loss	38
Figure B.1: Qualitative comparison of generated images	51

## List of Tables

	<b>Page</b>
Table 2.1: Comparison of capsule networks based GANs.	11
Table 5.1: Comparison of the architectures ...	25
Table 7.1: Inception means and standard deviations ..	39

## Abbreviation

Convolutional Neural Network	CNN
Generative Adversarial Network	GAN
Deep Convolutional Generative Adversarial Network	DCGAN
Stacked Capsule Autoencoders	SCAE