Spatial Localization using RNNs and Analysis of Emerging Firing Patterns

Bachelor Thesis Project (Stage-1)

by

Preetam Pinnada

170070042

Supervisor:

Prof. Udayan Ganguly

Department of Electrical Engineering

INDIAN INSTITUTE OF TECHNOLOGY BOMBAY

2020

Contents

Abstract

Neurons present in the Entorhinal Cortex (EC) of a mammalian brain produce interesting spatial firing patterns. Studies suggest that these spatial representations play a crucial role in understanding and navigating through the environment. These neurons are classified based on their firing patterns, viz. Grid cells, Place cells, Border cells, Band cells, Head direction cells, etc. It is believed that these representations encode the environment and perform path integration. Recent work in machine learning community supports this hypothesis. RNNs were trained to perform spatial localization and firing patterns in the hidden layers were observed to be strikingly similar to that of neurons in EC. As a part of BTP-1, I set out to build a spatial localization system to perform path integration for artificial agents. Further, I investigated model's hidden representations to observe the patterns and appreciate the necessary conditions for their occurrence.

Keywords: Entorhinal Cortex, Path Integration, Recurrent Neural Network

Chapter 1

Introduction

Through this investigation I attempt to understand – How do humans process and navigate in an environment? And how can this help in designing an efficient navigation system for artificial agents?

In 2005, a group of researchers at the Centre for the Biology of Memory (CBM) Norway, discovered a type of neurons in the Entorhinal cortex (EC) called the grid cells [1]. These neurons along with some other type of neurons such as the Place cells, Border cells, Displacement cells were found to constitute the positioning system in the brain.

In 2018 Christopher et al. [2] at the Columbia University reported a striking observation – they trained an RNN to perform spatial localization and observed a spontaneous emergence of grid-like representations in the model. This shows that artificial neural networks are a good proxy of mammalian brains for experimentation.

Path integration is the ability of an organism to calculate its updated position using idiothetic cues such as vision, motor neurons, directional cues, etc. I built an RNN to perform path integration and observed the representations learnt by the model to appreciate the spatial firing patterns. This model can be used with any artificial agent to perform spatial localization for navigation.

In chapter 2, the reviewed literature is discussed. In chapter 3, the details of model implementation and investigation is reported. Chapter 4 discusses the results obtained, and the Appendix (chapter 6) contains all the plots generated. Chapter 7 lists all the references, and chapter 8 links to the associated blogposts and code repositories.

Chapter 2

Literature Review

Christopher et al., 2018 [2] demonstrated a new way to understand the neural representations in the EC by reproducing them in ANNs. They trained an RNN to preform spatial localization and reported spontaneous emergence of grid-like spatial firing patterns in RNN. Translational speed and absolute direction are input to the model and the coordinates of the agent's position at the next time step is estimated. Some interesting observations are –

1. The nature of the grid-like patterns depended on the shape of the arena, unlike the exclusive hexagonal patterns observed experimentally
2. Neurons that fire selectively along the boundary typically emerge first. The grid-like responses with finer spatial tuning patterns only emerge later in training
3. The grid-like patterns did not appear in the absence of firing rate regularization
4. The grid-like patterns did not appear in the absence of noise in the state update
5. The grid-like patterns obtained by the model typically exhibit few periods, not as many as observed from experimental data of rodent's brain

Ingmar et al., 2016 [3] reported that no grid-like spatial firing patterns were observed when an RNN with LSTM units is trained to perform the following localization tasks –

1. Localization in a single familiar environment
2. Localization in novel environment
3. Localization in and classification of any of 100 familiar environments

Banino et al., 2018 [4] trained RNN (with LSTM units) to self-localize and observed the emergence of grid-like spatial representations that were independent of shape of the arena. Here instead of absolute direction, angular velocity is input to the model and is asked to predict both coordinates and the head direction at the next time step.

Chapter 3

Investigation

3.1 Model

The model consists of input layer of 2 neurons that accepts speed and absolute direction. Followed by a simple recurrent layer of 100 neurons to represent the hidden state, a gaussian noise layer, and a fully connected dense layer that maps the hidden state to an output (x, y) as shown in Fig 1(b). At every time step, speed and direction is input to the model and updated position at the next time step is estimated by the model.

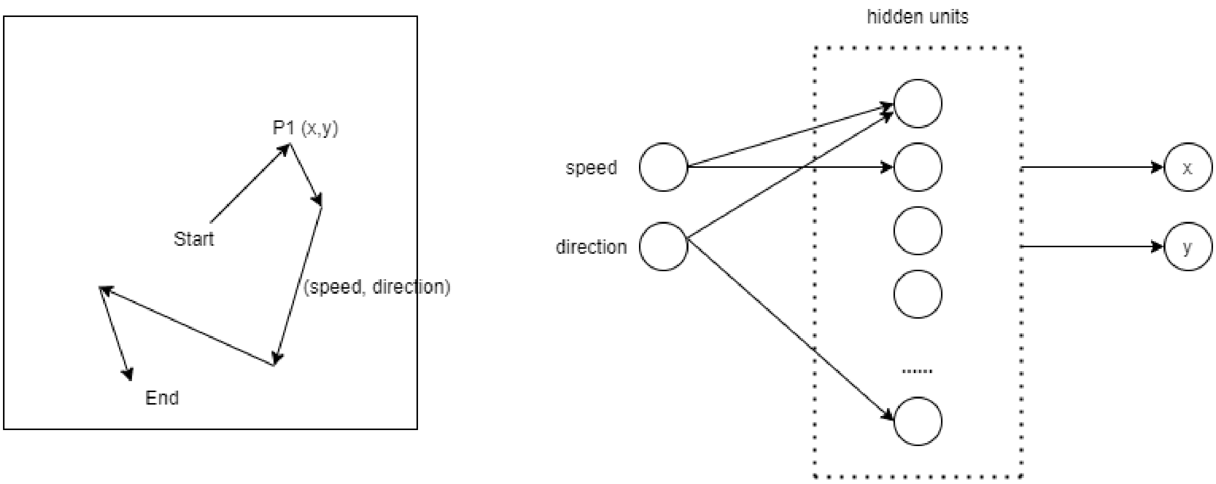


Fig 1. (a) Agent moving around in a rectangular environment; (b) Network architecture

State Update

$$\tau \frac{dx_i(t)}{dt} = -x_i(t) + \sum_{j=1}^{N} W_{ij}^{rec} u_j(t) + \sum_{k=1}^{N_{in}} W_{ik}^{in} I_k(t) + b_i + \xi_i(t)$$

$u_i(t) = tanh(x_i(t))$
$x_i(t)$: state of the i^{th} unit
$u_i(t)$: firing rate of the i^{th} unit
$\xi_i(t)$: noise

Output

$$y_j(t) = \sum_{i=1}^{N} W_{ji}^{out} u_i(t)$$

Cost function

$$E = \frac{1}{MTN_{out}} \sum_{m,t,j=1}^{M,T,N_{out}} (y_j(t,m) - y_j^{target}(t,m))^2$$

Regularization

1. $R_{L2} = \frac{1}{NN_{in}} \sum_{i,j=1}^{N,N_{in}} (W_{ij}^{in})^2 + \sum_{i,j=1}^{N,N_{out}} (W_{ij}^{out})^2$
2. $R_{FR} = \frac{1}{NTM} \sum_{i,t,m=1}^{N,T,M} u_i(t,m)^2$

3.2 Data

Two datasets are generated, one in a rectangular arena and other in a circular arena, to understand if there is any dependence of firing patterns on the shape of arena. Each dataset consists of 100000 paths with 10-20 steps in every path starting from the origin (0,0). Step size varies randomly between 0-10 units in a rectangular arena of size 100x100 units (or in a circular arena of radius 50 units). Every path generated is a random walk in 2D space within the boundary. An example path is shown in Fig 2.

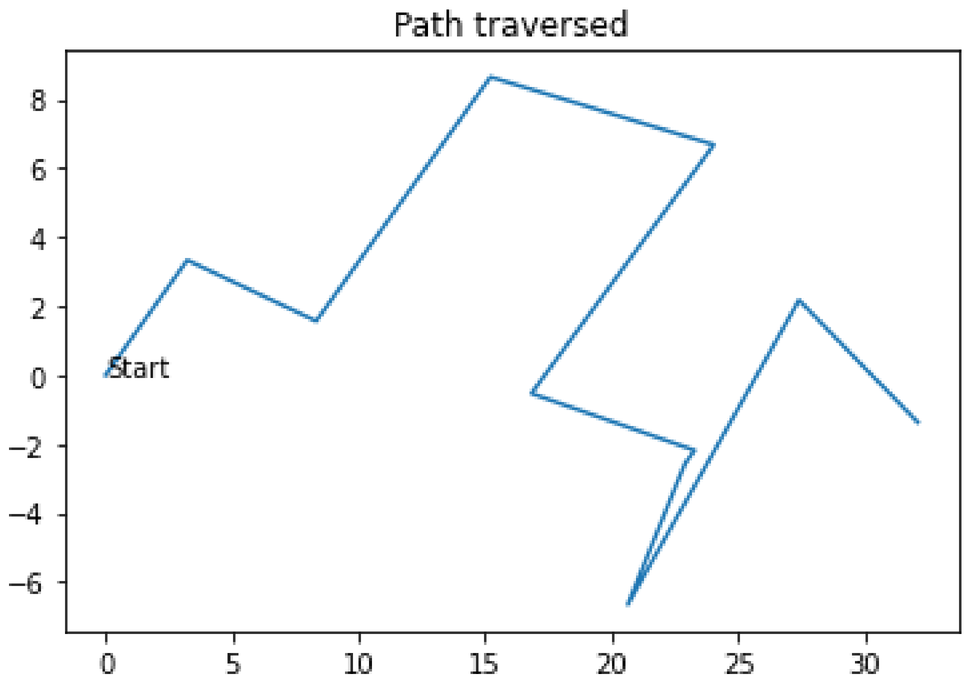


Fig 2. An example path from the recatngular arena dataset

3.3 Methodology

The model in trained on 80000 paths and is validated on 20000 paths in batches of 200 for 50 epochs. The model is then used to estimate the path traversed by the agent, and the hidden representations are obtained after every time step to plot spatial firing patterns.

Chapter 4

Results and Discussion

The model was able to successfully self-localize and path integrate. As the number of steps increased the error of the estimate increased. No convincing grid-like firing patterns were observed, because a smaller number of examples were used to plot the firing patterns due to limited computational capacity. But firing patterns resembling that of border cells, band cells, place cells were observed.

Recurrent and activity regularizations, and noisy state update were found to improve the path estimates and the firing patterns, but nothing can be commented conclusively about their impact on grid-like firing patterns.

The training process was stopped in between after 10, 20, 30, 40 epochs to observe the evolution of firing patterns. Initially the firing patterns were predominantly like place cells, but with progress in training border cells, band cells, and other complex patterns emerged. This is coherent with observations by Bjerknes et al., 2014 [6] about evolution of representations with brain development.

One test case each from rectangular and circular arenas are shown in Fig 3. Some interesting spatial firing patterns of the hidden layer of RNN are shown in Fig 4. Extensive results are reported in the Appendix (Chapter 6).

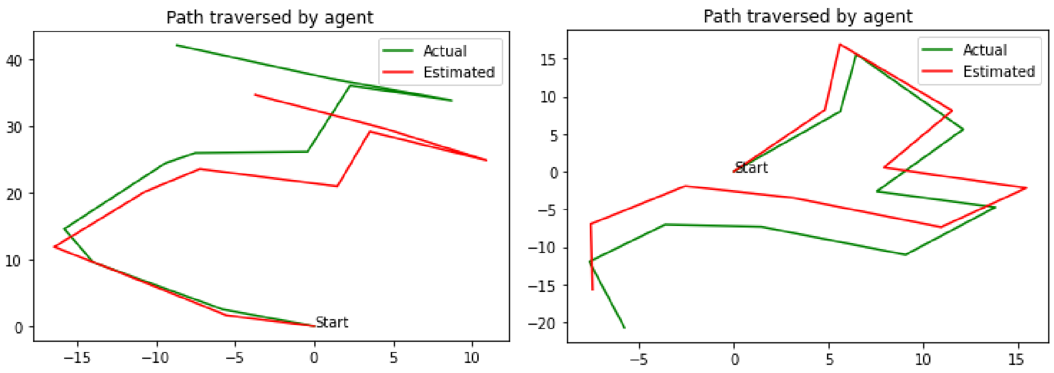


Fig 3. Green line is the actual path traversed by agent and red line is the path estimated by the model (a) Circular area; (b) Rectangular arena.

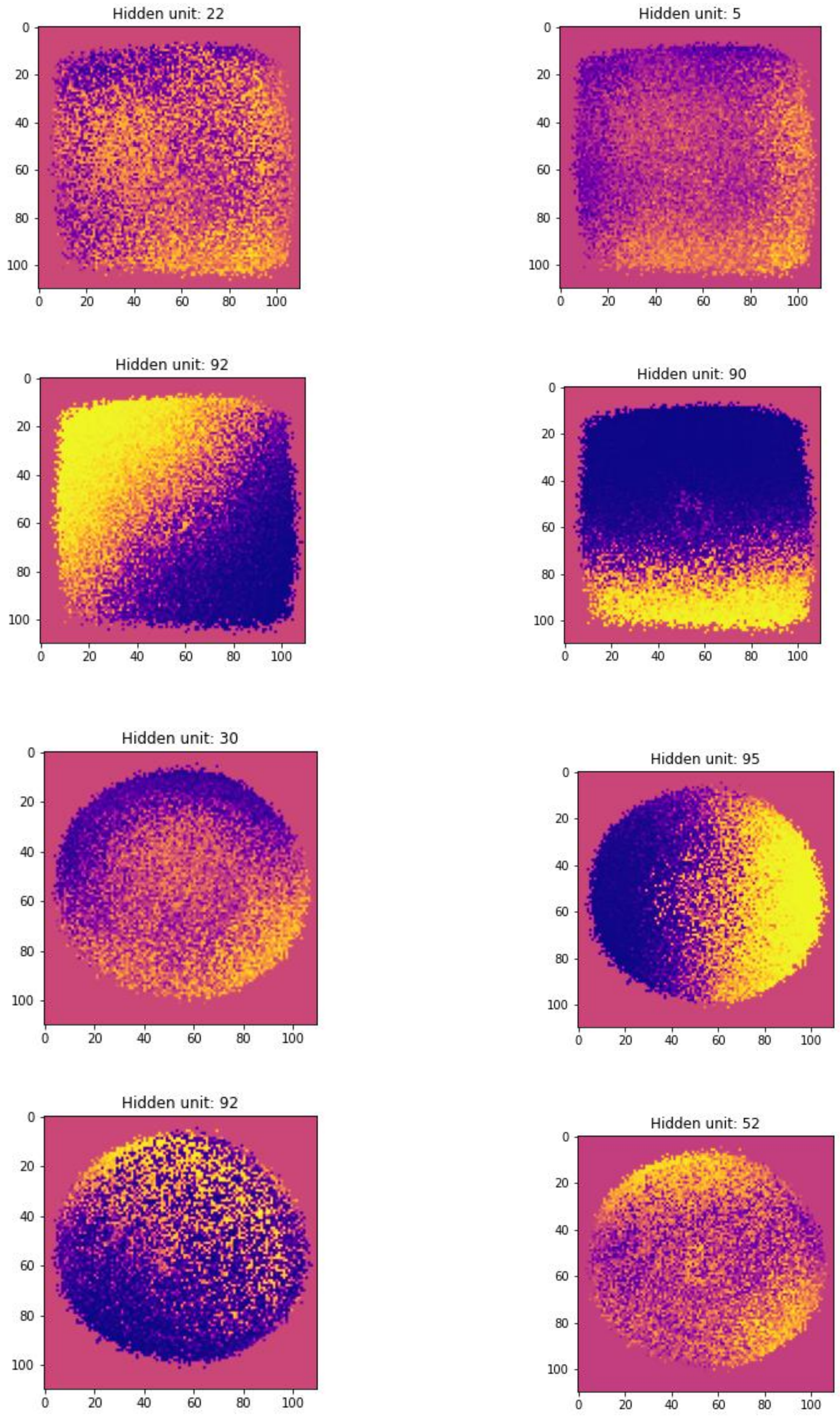


Fig 4. Rectangular arena - (i), (ii), (iii), (iv); Circular arena – (v), (vi), (vii), (viii)

Chapter 5

Summary

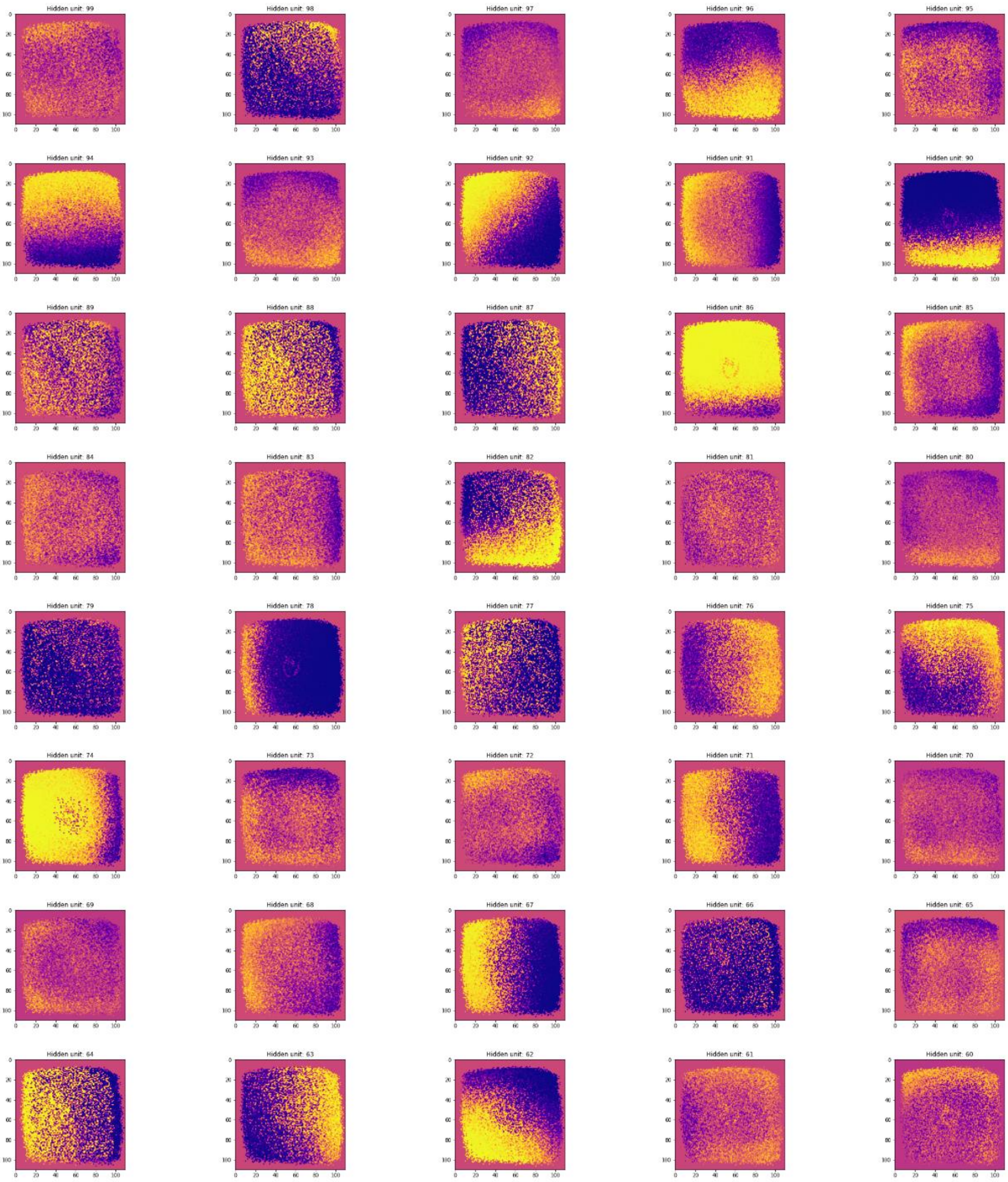
Spatial firing patterns similar to that in EC of a foraging mammalian brain have been observed in RNN trained to self-localize. The trained model can be easily integrated into an artificial agent to perform path integration. The evolution of firing patterns as training progressed supports the observations made by biologists about brain development. If not for the constraints on compute power, with greater number of steps in each path, the grid-like firing patterns would have become more evident.

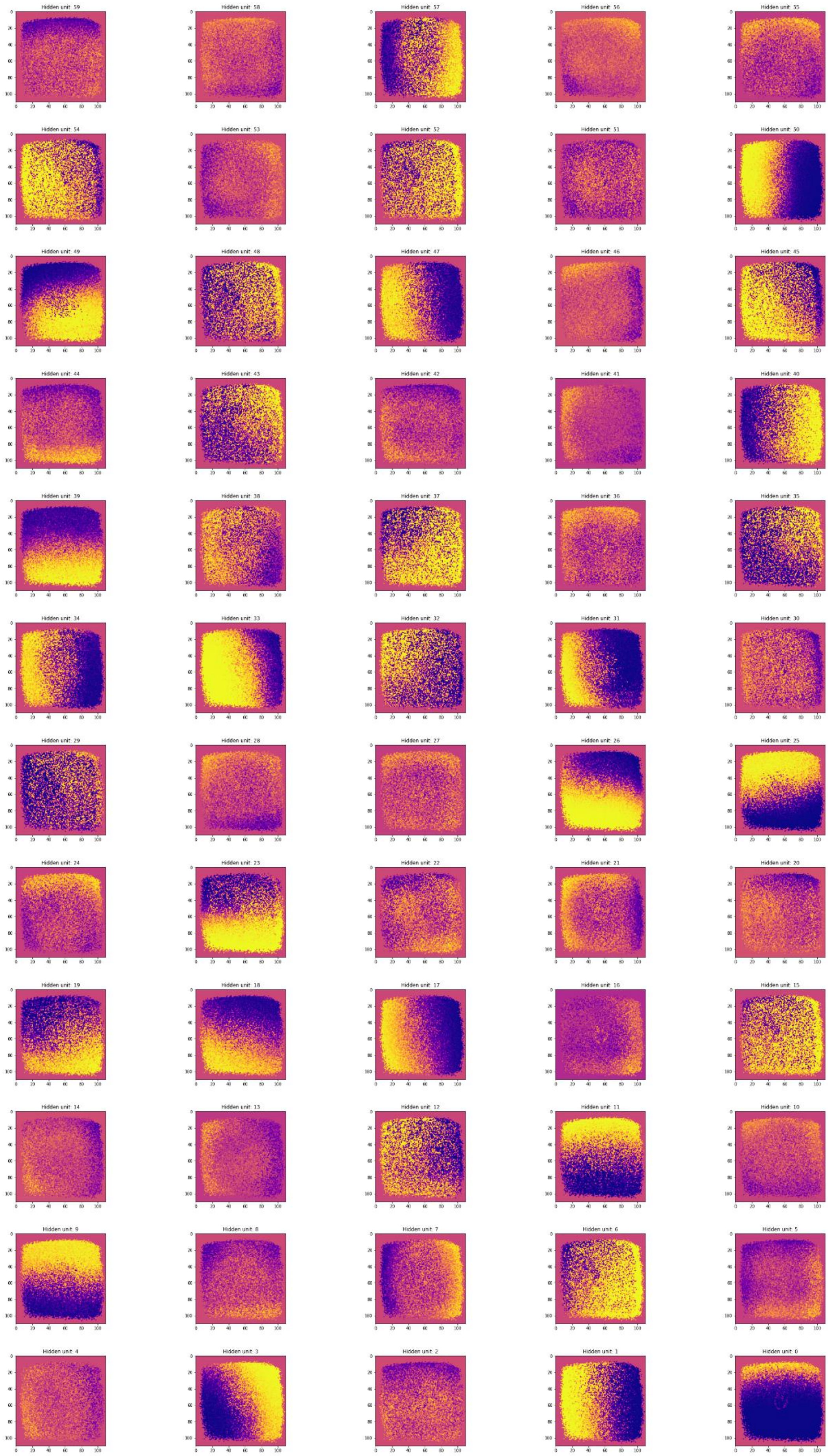
In stage-2 of my bachelor thesis I plan to investigate the framework for intelligence based on grid cells in the neocortex put forth by Hawkins et al., 2019 [5]. The framework suggests that just like the way grid cells helped in navigation, the grid cells present in neocortex might help with understanding objects around and their relationships. This investigation would help in designing a general intelligence system that can interpret objects in the environment.

Chapter 6

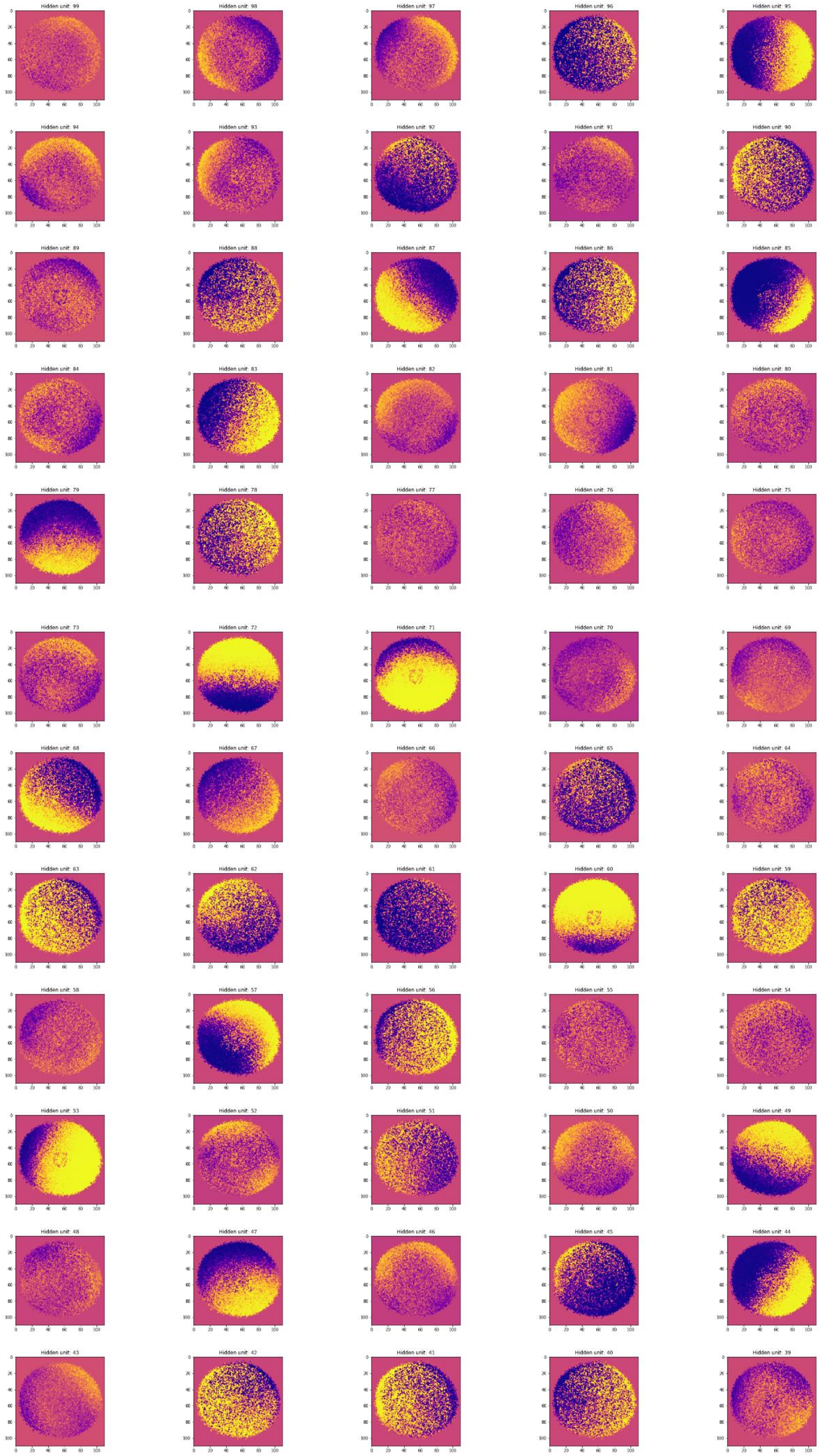
Appendix

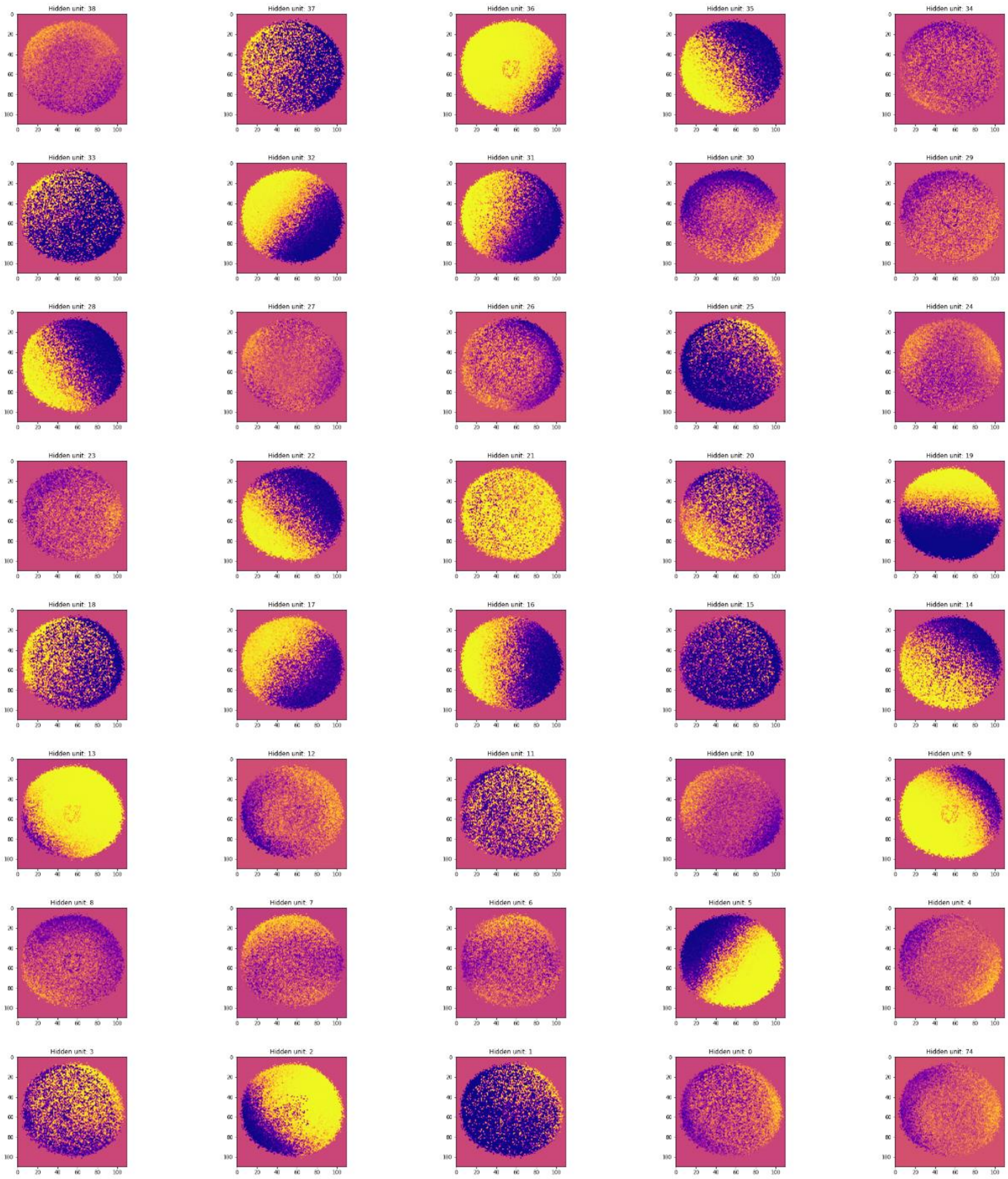
A.1 Firing patterns of all 100 hidden units trained in rectangular arena




Hidden unit: 59
Hidden unit: 58
Hidden unit: 57
Hidden unit: 56
Hidden unit: 55
Hidden unit: 54
Hidden unit: 53
Hidden unit: 52
Hidden unit: 51
Hidden unit: 50
Hidden unit: 49
Hidden unit: 48
Hidden unit: 47
Hidden unit: 46
Hidden unit: 45
Hidden unit: 44
Hidden unit: 43
Hidden unit: 42
Hidden unit: 41
Hidden unit: 40
Hidden unit: 39
Hidden unit: 38
Hidden unit: 37
Hidden unit: 36
Hidden unit: 35
Hidden unit: 34
Hidden unit: 33
Hidden unit: 32
Hidden unit: 31
Hidden unit: 30
Hidden unit: 29
Hidden unit: 28
Hidden unit: 27
Hidden unit: 26
Hidden unit: 25
Hidden unit: 24
Hidden unit: 23
Hidden unit: 22
Hidden unit: 21
Hidden unit: 20
Hidden unit: 19
Hidden unit: 18
Hidden unit: 17
Hidden unit: 16
Hidden unit: 15
Hidden unit: 14
Hidden unit: 13
Hidden unit: 12
Hidden unit: 11
Hidden unit: 10
Hidden unit: 9
Hidden unit: 8
Hidden unit: 7
Hidden unit: 6
Hidden unit: 5
Hidden unit: 4
Hidden unit: 3
Hidden unit: 2
Hidden unit: 1
Hidden unit: 0

A.2 Firing patterns of all 100 hidden units trained in circular arena




Hidden unit: 38
Hidden unit: 37
Hidden unit: 36
Hidden unit: 35
Hidden unit: 34
Hidden unit: 33
Hidden unit: 32
Hidden unit: 31
Hidden unit: 30
Hidden unit: 29
Hidden unit: 28
Hidden unit: 27
Hidden unit: 26
Hidden unit: 25
Hidden unit: 24
Hidden unit: 23
Hidden unit: 22
Hidden unit: 21
Hidden unit: 20
Hidden unit: 19
Hidden unit: 18
Hidden unit: 17
Hidden unit: 16
Hidden unit: 15
Hidden unit: 14
Hidden unit: 13
Hidden unit: 12
Hidden unit: 11
Hidden unit: 10
Hidden unit: 9
Hidden unit: 8
Hidden unit: 7
Hidden unit: 6
Hidden unit: 5
Hidden unit: 4
Hidden unit: 3
Hidden unit: 2
Hidden unit: 1
Hidden unit: 0
Hidden unit: 74

Chapter 7

References

1. Moser EI, Kropff E, Moser MB. Place cells, grid cells, and the brain's spatial representation system. Annu Rev Neurosci. 2008;31:69-89. doi: 10.1146/annurev.neuro.31.061307.090723. PMID: 18284371.
2. Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. ICLR 2018
3. Ingmar Kanitscheider and Ila Fiete. Training recurrent networks to generate hypotheses about how the brain solves hard navigation problems. arXiv 2016
4. Banino, A., Barry, C., Uria, B. et al. Vector-based navigation using grid-like representations in artificial agents. Nature 557, 429–433 (2018). https://doi.org/10.1038/s41586-018-0102-6
5. Hawkins Jeff, Lewis Marcus, Klukas Mirko, Purdy Scott, Ahmad Subutai. A Framework for Intelligence and Cortical Function Based on Grid Cells in the Neocortex. Frontiers in Neural Circuits 2019. DOI=10.3389/fncir.2018.00121
6. Tale L Bjerknes, Edvard I Moser, and May-Britt Moser. Representation of geometric borders in the developing rat. Neuron, 82(1):71–78, 2014.

Chapter 8

Other Resources

8.1 Code Repository

1. Rectangular arena

   https://colab.research.google.com/drive/1CORsN1IDF7Q1J2NdbWMhguZcSvgNUkYH?usp=sharing

2. Circular arena

   https://colab.research.google.com/drive/1rND4Iy998x9i_BYynfEJIH5x2wJHJXKV?usp=sharing

8.2 Blogposts

1. Emergence of grid-like representations in ANNs

   http://iitbnf.iitb.ac.in/udayanresearch/index.php/2020/09/09/emergence-of-grid-like-representations-in-anns/

2. Details of grid-like firing patterns in ANNs

   http://iitbnf.iitb.ac.in/udayanresearch/index.php/2020/10/08/details-of-grid-like-firing-patterns-in-anns/

Chapter 9

Acknowledgment

I sincerely appreciate the efforts put in by Prof. Udayan Ganguly, Vivek Saraswat, MeLoDe research group, and the Department of Electrical Engineering for making stage-1 Bachelor Thesis Project (BTP) possible by working remotely during Covid-19 pandemic.

This short stint as an undergraduate researcher was a steep learning curve. And am eagerly looking forward to beginning BTP stage-2 in the upcoming semester.