Scale- and Translation-Invariant Unsupervised Learning of Hidden Causes Using Spiking Neurons with Top-Down Attention

Youssef Kashef

7. August 2013

Contents

Chapter 1

Abstract

Nessler et al have demonstrated the ability of a spiking neuronal network governed by spike-timing-dependent-plasticity and a stochastic winner-take-all circuit to learn and predict causes from visual input. We aim to increase the computational power of the existing network through invariance to translation and scale. The visual system of the brain masters the recognition of objects wherever they appear in the visual scene and regardless of scale, orientation or even with partial occlusions. It achieves this through attention. Therefore, we turn to the pool of literature on modeling visual attention systems inspired from the brain. The architecture of the extended model is composed of the existing recognition module whose response modulates the attention module to be constructed in a top-down manner. This modulation will allow the attention module to alter the input window exposed for recognition. Attention is modeled as a network measuring for saliency in a scene by feature extraction with the use of hierarchies. The design and development of this extended model to achieve the required invariance using processes that approximate their biological counterparts is presented. Emphasis is put on making these approximations through computationally economic implementations. Evaluation of the model is based on its performance in a set of experiments as well as its computational efficiency. Experiments are constructed to scrutinize the behavior of the model, its ability to converge onto a sight within a scene that enables recognition. Artificial as well as natural images are used to further reveal the capabilities and limitations of our approach.

Chapter 2

Introduction

elaborated abstract with references.......[?]

Chapter 3

Object Recognition with Spike Expectation Maximization

3.1 Spike Expectation Maximization

Literature review of existing SEM model

Nessler et. al have articulated a bayesian model of how the brain analyszes sensory stimuli. The model demostrates the learning of hidden causes in visual stimuli emerging through by identifying correlations of a winner-take-all (WTA) network of spiking neurons that are activated continuously in the presence of their preferred stimulus. The model demonstrates utilizing Spike Time Dependent Plasticity (STDP) in WTA circuits as an approximation of Expectation Maximization.

3.2 SEM for learning features

3.2.1 Extending SEM by learning orientations

The current encoding of external variables accounts for the intensities of the spatial units (pixels) of a presented stimulus. The encoding of intensities is performed through a population coding by antognistic binary nodes per pixel that drive a poisson process [**?**]. Prallel to these intensity encoded nodes, we add a WTA circuit per pixel that determines the preferred orientation of this node relative to its spatial neighbors. This creates an orientation map of the presented stimulus. Whilst counter-intuitive with traditional learning models, SEM benefits from increasing the dimensionality of its feature space as this increases its resolution for detecting correlations between an output node z and input nodes y on a linear scale. Recalling the use of using population coding to encode in atagonistic (on-node, off-node) fashion, thus letting the WTA learn the likelihood of an input node firing, or not firing, explicitly, as shown by 3.1.

$$p(z=1|y) \propto y * p(y=1|z) + (1-y) * p(y=0|z) \tag{3.1}$$

As we introduce the orientation map we may add additional operands to 3.1 to account for the nodes preferred orientation.

$$\begin{aligned} p(z=1|y) \propto & \frac{y_I * p(y_I=1|z) + (1-y_I) * p(y_I=0|z)}{N_o - 2} \\ & + \frac{a}{N_o}[y_o * p(y_O = O_i) + (N_O - 1) * \sum_{i \neq o_i} p(y_O = o_i)] \end{aligned} \tag{3.2}$$

where

- y_I denotes an intensity node,
- y_I denotes an intensity node,
- N_O denotes the number of orientations available,
- y_O denotes an orientation node,
- O denotes the available orientation. Orientations can be defined discretly and arbitrarily (e.g. 30, 60,...180 degrees) or they can be learned [?],

We redesign the network with a cascade of hierarchical WTA circuits. The input layer is a matrix of WTA circuits per spatial. Each input WTA circuit decides on the preferred orientation and intensity of its input. We will experiment with configuring the input WTA circuit to only relay intensity, only orientation, or both information.

3.3 Extending SEM for learning hidden features

extending SEM by learning abstract features

We have seen the computational power of the SEM model as an unsupervised method for identifying hidden causes. So far the hidden causes have been used synonymously with predefined classes (e.g. digits[?]). We will extend the SEM model in a way that breaks this assumption. We insert an additional WTA circuit, responsible for learning hidden causes that depict abstract features of the object we're attempting to detect and recognize. This feature layer will contribute to the bottom-up learning as we expose it to the low-level input and and have it drive the WTA circuit already encountered in the original SEM architecture. With this additional feature-WTA circuit introduced we no longer require presentation of the entire stimulus but will restrict stimulus presentation to subregions within the space of a stimulus. These subregions may represent salient regions within a stimulus. The definition and method of selecting these subregions will be discussed in more detail as we discusss the object detection framework.

Chapter 4

Object detection

4.1 Attention

Attention is the ability to economize computational power and reduce its entropy.

4.2 Attention mechanisms

literature reivew of attention mechanisms

4.3 Bottom-up Attention

what we used from Itti's

4.4 Top down attention

describe attempt

Chapter 5

Achieving invariance

5.1 Model

5.2 Results

5.3 Discussion

5.4 Conclusion

Appendix A

The First Appendix

The `\appendix` command should be used only once. Subsequent appendices can be created using the Chapter command.

Appendix B

The Second Appendix

Some text for the second Appendix.

Bibliography

[1] Y. LeCun and L. Bottou. Gradient-based learning applied to document recognition. *Proceedings of the ...*, 86(11):2278–2324, 1998.

[2] Bernhard Nessler, Michael Pfeiffer, and Wolfgang Maass. STDP enables spiking neurons to detect hidden causes of their inputs. *In Proc. of NIPS 2009: Advances in Neural Information Processing Systems. MIT Press*, 22:1357–1365, 2010.

Appendix C

Afterword

That's all folks!

The back matter often includes one or more of an index, an afterword,

C.1 Acknowledgments

Michael Matthew Cook My family: Sahra, father My friends Malte Alf INI