KIT

Karlsruhe Institute of Technology

Title of thesis
Second title line

Master Thesis of

Forename Surname

Department of Computer Science
Institute for Anthropomatics
and
FZI Research Center for Information Technology

Reviewer: Prof. Dr.–Ing. J. M. Zöllner
Second reviewer: Prof. Dr.–Ing. R. Dillmann
Advisor: Dipl.–Inform. Max Mustermann

Research Period: XX. Monat 20XX   –   XX. Monat 2016

Title of thesis - Second title line

by
Forename Surname

Master thesis
in Monat 2016

FZI

Affirmation

Ich versichere wahrheitsgemäß, die Arbeit selbstständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde.

Karlsruhe,   *Forename Surname*
in Monat 2016

Contents

1. Introduction

1.1. Motivation

In 2012 by winning the Imagenet Large Scale Visual Recognition Challenge 2012 convolutional neural network gained a big rise in popularity. Now they are becoming popular for their powerful abstraction mechanism in the fields of image and video classification and description and speech recognition. This can be contributed to compositional structure in which the world can be perceived and to their ability to extract high level features on spatial and/or temporal conditioned data.

Generating discriminative high level features extractors allows the system to dynamically adapt to the input data and work on various kinds of data. In addition the features extractors do not need to have any semantic representation and can be more complex to the manually build feature extractors. This also removes the labor-intensive and time consuming task of manually building feature extractors. Consequently they recently got adapted to solve robotic problems like grasp planning, drone navigation and autonomous driving.

One precursor of those are the deep belief networks (DBNs), which are built up of restricted boltzmann machines (RBMs). DBNs have shown excellent performance on image classification tasks in the early 2000s.

Compared to classical CNNs, DBNs allow recurrent connections and are trained in an unsupervised manner and do not need labeled data. They have been described as "probably the most biologically plausible learning algorithm for deep architectures we currently know". DBNs can be used as generative model as well, which means they can sample data according to a learned distribution, e.g. find the most probable completion for a partially erased image.

Adding convolution to DBNs increased the performance of DBNs on image classification tasks, caught up to state of the art results and made the system more similar to the primate visual cortex than a standard RBM.

All those approaches use scalar values between neurons at discrete time slices to propagate information. This proposes some difficulties and is not biologically plausible, since biological neuron interleave linear and nonlinear operations, they communicate by stochastic binary values and are not synchronized. Spiking neural networks (SNNs), designed to simulate the communication between neurons with action potentials/ spikes, work in continuous time by design and do not suffer from the aforementioned limitations.

1.2. Problem Statement

To our best knowledge, up to today, there exists no system which utilizes the benefits of all those approaches. The main objective of this thesis is to realize such a spiking network, which integrates convolution and can be easily trained utilizing the RBM learning algorithm, to extract high level features. Two approaches are described. The first approach trains convolutional RBMs on discretized input data, to build up a DBN which is then converted to a SNN. The next approach directly works on continuous (event-driven) input-data and realizes a STDP learning algorithm with shared weights to train spiking convolutinal RBMs directly. Both approaches will learn to extract high level features, which can be further used to classify an object or directly generate a grasp id.

1.3. The Human Brain Project

Heiko:

This thesis is under the scope of the research of the SP10 (sub-project 10) of the Human Brain Project. The Human Brain Project is an European Commission Future and Emerging Technologies Flagship and a large ten-year research project which aims to create a collaborative research infrastructure across national borders to progress the knowledge in neuroscience, computational neuroscience, medicine, scientific computation, and robotics. In the project over 120 institution from across Europe collaborate in 12 sub-projects.

The sub-project 10 of the Human Brain Project develops the Neurorobotics Platform which permits researchers to simulate robotic experiments with regard to neurorobotics. Apart from the platform, research is focused on the development of applications in robotics based on insights from neuroscience. One focus of the research group at the FZI Karlsruhe is the development of computational models for neurobiological inspired robotic grasping.

1.4. Overview

This thesis describes the approach and implementation of a spiking convolutional deep belief networks to extract high level features on continuous visual input. The thesis is structures as follows:

Chapter 2 introduces sine background information, which is used in chapter 3 to describe state-of-the-art research used in this thesis. Chapter 4 will describe the different approaches to build such a spiking convolutional deep belief networks. In chapter 5 the different implementation steps and the architecture will be described. Chapter 6 outlines and compares the performance of the networks. Chatper 7 will conclude the gathered insight of this thesis, state its limitations and give suggestions for further improvements and research.

2. Background

2.1. Probabilistic Graphical Models

Neural networks can be seen as graphs

Using PGMs can be used to infer properties of some Neural Nets

PGMs are used to structure a model of the input and can be used for different task: Density estimation, Denoising, Missing value imputation, Sampling.

2.1.1. Bayesian network

Bayesian networks or belief networks are directed acyclic graphs, in which random variables are represented by nodes and their causal dependencies are represented by (directed) edges/connections.

If theres an connection from Node Xi to Xj then Xi is referred to as a parent of Xj and, similarly, Xj is referred to as the child of Xi.

In addition to the DAG structure, which is often considered as the "qualitative" part of the model, one needs to specify the "quantitative" parameters of the model. The parameters are described in a manner which is consistent with a Markovian property, where the conditional probability distribution (CPD) at each node depends only on its parents. FORMULARS

A BN reflects a simple conditional independence statement. Namely that each variable is independent of its nondescendents in the graph given the state of its parents. This property is used to reduce, sometimes significantly, the number of parameters that are required to characterize the JPD of the variables. This reduction provides an efficient way to compute the posterior probabilities given the evidence. Such a reduction provides great benefits from inference, learning (parameter estimation), and computational perspective.

GRAPH EXAMPLE

Given some observed variables, the bayes net can be used to infer the most likely states of the other hidden variables. This process of computing the posterior distribution of variables given evidence is called probabilistic inference. A Bayesian network can thus be considered a mechanism for automatically applying Bayes' theorem to complex problems, which is defined through the "qualitative" and "quantitative" parameters of the net.

2.1.2. Markov Random Field

In contrast to Bayesian networks, Markov random fields are undirected graphical models, in which random variables are represented by nodes and edges/ connections indicate conditional dependencies.

For all cliques in the graph a "Clique Potential" / "Likelihood" can be given, which indicates how likely it is for the variables in the clique to take given values.

A clique in a graph is a group of nodes which are in themselves fully connected, i.e. each node has each other node of the graph a as direct neighbour.

The clique potentials can be used to calculate an unnormalized probability distribution p(x) FORMULAR

EXAMPLE

2.1.3. Energy-Based Models

One way to model the unnormalized probability distribution p(x) is to use an Energy function E(x): Formular

This results in an energy-base model (EBM).

Since exp(a)*exp(b) = exp(a+b) , where a and b are the Energy of a subgraph/clique and the Energy of the complete graph is the sum of all cliques, it is apparent that EBMs are a sub cathegory of MRF.

Because exp(z) > 0 , there is probability of each state if greater than zero.

2.1.4. Sampling

Sampling is concerned with the selection of a subset of individuals from within a statistical population to estimate characteristics of the whole population.

Sampling provides a flexible way to approximate many sums and integrals at reduced cost, which could only have been calculated using costly operations or were intractable altogether.

Graphical models also facilitate the task of drawing samples from a model.

Ancestral Sampling For directed graphical models is that a simple and efficient procedure called ancestral sampling can produce a sample from the joint distribution represented by the model. The basic idea is to sort the variables xi in the graph into a topological ordering,so that for all i and j, j is greater than i if xi is a parent of xj. The variables can then be sampled in this order.

The topological sorting operation guarantees that the conditional distributions are valid and one can sample from them in order.

Markov Chain Monte Carlo If pmodel(x) is represented by an undirected model, Markov Chain Monte Carlo methods can be used. MCMC methods interpret the model as a Markov chain, and work best, when no state gets zero probability assigned be the model.

The basic idea is to begin a state x with some arbitrary value. Then for a (infinite) time x is repeatedly randomly updated using the by the model given transition distribution T(x,x). Eventually x becomes a fair sample from p(x)/ from the stationary distribution of the Markov chain.

To get more than one sample, one can run more Markov chains in parallel, each initialized with a random starting state. Another method is to run just one Markov chain, run it form some burn in/ mixing time, which allows the Markov chain to reach its equilibrium, and than take samples at

different timesteps. Those approaches need the Markov chain to reach its equilibrium distribution, which is usually done, by letting it run for some burn in time. But there is no guaranty, that the Markov chain has settled in the given timespan. Another problem with the second approach is, that since it can be hard to escape probable states, and when not run for an infinite timespan, more likely states can be over represented and less likely states under represented, if they did not occur or over represented if they did occur and the number of samples it not big enough to scale the estimate.

Gibbs sampling   Gibbs sampling is a commonly used MCMC algorithm. The basic idea in Gibbs sampling to perform the transition from one state to another in accordance with T(x,x) is, to select a single variable xi and sample it conditioned on its neighbours. Several variables can be sampled at the same time as long as they are conditionally independent given all of their neighbours.

2.2. Neural Networks

In this section at first the neuroscience foundations of Neural Networks will be explained. After that different models will be described, starting with models working at discrete time steps and then describing models, which work in continuous time.

2.2.1. Natural

Brain

Neuron

Learning

2.2.2. Time discrete

Perceptron

Model

Perceptron Learning

Decision Fuction

Mutlilayer-Perceptron

Model architecture

Activation functions

Backpropagation

Convolutinal Neural Networks

Convolution

Architecture

Training

Hopfield Nets

Model

Properties / Advantages / Disadvantages

Boltzmann Machines

Model

RBMs

CD-Training

Deep Belief Networks

Stacking Up RBMs

Fine-tuning

2.2.3. Continuous

Neuron Models

LIF

Hodgkin-Huxley

Poisson

Neural Coding

Learning

STDP

Symetric STDP

3. Related Work

3.1. Convolutional RBM

3.2. Sampling in SNNs

3.3. NN -> SNN conversion

3.4. eCD and Sampling Machines

4. Approach

4.1. Conversion

4.1.1. Conv DBNs

4.1.2. Conversion

4.2. eCD

4.2.1. Convolution

4.2.2. DBNs

5. Implementation

5. Implementation

6. Experiments&Results

6. Experiments&Results

7. Conclusion and Outlook

A. Some appendix

A.1. (if needed)

Liste der noch zu erledigenden Punkte

B. List of Figures

C. List of Tables

D. Bibliography

[1] G. Berndes, M. Hoohwijk, and R. van den Broek. The contribution of biomass in the future global energy supply: a review of 17 studies. *Biomass and Bioenergy 25*, pages 1–28, 2003.

[2] W. Feng, H. J. van der Kooi, and J. de S. Arons. Phase equilibrium for biomass conversion processes in subcritical and supercritical water. *Chem. Eng. J. 98*, pages 105–113, 2003.

[3] X. H. Hao, L. Guo, X. M. Mao, and H. J. Chen. Hydrogen production from glucose as a model compound of biomass gasified in supercritical water. *Int. J. Hydrogen Energy*, pages 53–64, 2003.

[4] R. A. H. III. The age of energy gases. *Int. J. Hydrogen Energy 27*, pages 1–9, 2002.

[5] I.-G. Lee, M.-S. Kim, and S.-K. Ihm. Gasification of glucose in supercritical water. *Ind. Eng. Chem. Res. 41*, pages 1182–1185, 2002.

[6] Y. Matsumura and T. Minowa. Biomass gasification in near- and super-critical water: Status and prospects. *Biomass and Bioenergy 29*, pages 269–292, 2005.

[7] M. Osada, T. Sato, M. Watanabe, T. Adschiri, and K. Arai. Low temperature catalytic gasification of lignin and cellulose with a ruthenium catalyst in supercritical water. *Energy & Fuels 18*, pages 327–333, 2004.

[8] I. T. and Y. Matsumura. Gasification of cellulose, xylan and lignin mixtures in supercritical water. *Ind. Eng. Chem. Res. 40*, pages 5469–5474, 2001.

[9] S. T, S. Kurosawa, R. L. Smith, T. Adschiri, and K. Arai. Water gas shift reaction kinetics under noncatalytic conditions in supercritical water. *J. Sup. Fluids 29*, pages 113–119, 2004.

[10] M. O. T. Sato, M. Watanabe, M. Shirai, and K. Arai. Gasification of alkylphenols with supported noble metal catalyst in supercritical water. *Ind. Eng. Chem. Res. 42*, pages 4277–4282, 2003.

[11] J. D. Taylor, C. M. Herdman, B. C. Wu, K. Wally, and S. F. Rice. Hydrogen production in a compact supercritical water reformer. *Int. J. of Hyd. Ene. 28-11*, pages 1171–1178, 2003.

[12] T. Yoshida, Y. Oshima, and Y. Matsumura. Gasification of biomass model compounds and real biomass in supercritical water. *Biomass&Bioenergy 26*, pages 71–78, 2004.

[13] Y. Yoshida, K. Dowaki, Y. Matsumura, R. Matsuhashi, D. Li, H. Ishitani, and H. Komiyama. *Comprehensive comparison of efficiency and CO2 emissions between biomass energy conversion technologies-position of supercritical water gasification in biomass technologies, Biomass and Bioenergy.* 2003.

[14] D. Yu, M. Aihara, and M. J. Antal. Hydrogen production by steam reforming glucose in supercritical water. *Energy & Fuels, 7*, pages 574–577, 1993.