Universitatea POLITEHNICA din București

Facultatea de Automatică și Calculatoare,
Departamentul de Calculatoare

LUCRARE DE DIPLOMĂ

Deep Learning în jocuri

Conducător Științific:
As. Drd. Ing. Tudor Berariu

Autor:
Florentina-Ștefania Bratiloveanu

București, 2015

University POLITEHNICA of Bucharest

Faculty of Automatic Control and Computers,
Computer Science and Engineering Department

BACHELOR THESIS

Playing games with Deep Learning

Scientific Adviser:
As. Drd. Ing. Tudor Berariu

Author:
Florentina-Ștefania Bratiloveanu

Bucharest, 2015

Abstract

Deep Learning is a new, interesting and a fast-growing field of Machine Learning. It combines the classical model of multilayer perceptrons with layers of feature extraction inspired from visual cortex of the brain. Moreover, dealing with the curse of dimensionality is another advantage when we talk about Convolutional Neural Networks. This paper proposes an alternative to the classical reinforcement learning techniques, such as Q-Learning and SARSA. The question which arises is this: what if we make an agent and allow it to play a game based on information from visual frames and rewards recieved during the game?

Contents

Chapter 1

Introduction

1.1 Motivation

Humans have always felt the need to explore and find a way to a better life. Starting with stone tools (2.6 million years ago) and continuing with personal computers are only some proof in human evolution. Lately, we have discovered Machine Learning which have the purpose to allow machines for learning to do different tasks.

Deep learning is a subfield of machine learning area and it is inspired by how the human brain works. Through deep learning, we are trying to solve some of the most pressuring human problems as cancer classification (benign or malignant tumor)[1], self-driving cars based on pedestrian detections[2] or recognizing digits taken by Google Street View[3], all using unsupervised learning of features.

This paper propese an in-dept look of solving reinforcement learning problems, using convolutional neural networks. More exactly, we want to design an agent which is capable of learning to play a game[4] seeing only the pixels from the frames and being rewarded after finishing the game.

This ideea was brought to attention in 2013, when Alex Graves et al. from DeepMind[1] came with the ideea of creating a convolutional neural network capable of playing different Atari 2600 games, being as good as a game tester. The motivation behind this experiment is not only connected with the fact that we have a machine capable of playing games, but the more important problem is that they achieved the generalization of a machine that could learn several types of games, which in fact is the starting of a new revolution in machine learning. If we are capable of making one unique algorithm that can solve multiple tasks, we are capable of simulating the real human brain and capable of reducing complex problems to another ones, much more simple and thus, we can reduce programming burden and solving problems as cancer classification.

1.2 Project description

This paper presents two different algorithms in creating a good agent which can learn playing games. It is worth mentioning that the agent does not know the rules and can not infere with some rules at the beginning of the game. The whole topic has been splitted in three parts:

[1] http://deepmind.com/

running Q-Learning on Hanoi Towers game, learning values predicted by Q-Learning using a convolutional neural network and connecting the first two parts together.

First of them is the classical Q-Learning, which is searching the optimal policy for taking actions from each state of the game. We will discus about exploration vs exploitation problem, how to determine the learning rate suitable for our purposes, how many episodes we have to play until we determine a policy close to the optimal one.

The second topic we approach in this paper is the possibility of predicting continous output from image. Practically, we take the dataset from Q-Learning which contains image as input and a table of four values corresponding to the four possible actions in Hanoi Tower games: up, down, left and right.

The last part is the one we combine the first two modules.

1.3 Technologies

All algorithms described in this paper have been implemented using Torch7[1], a deep learning framework written in Lua[2], a scripting language based on a C API. For creating the game, generating frames or modifying images LOVE platform[3] was the close alternative which is also written in Lua. For interactive computing(image/filter visualization) iTorch[4] has been used.

Why Lua and Torch? They provide a fast environment as opposed to another ones[5], multiple modules with functions already implemented such as transfer functions(tanh, sigmoid), loss functions(Mean Squared Error, Negative Log Likelihood) or convolutional layers(SpatialMaxPooling, SpatialSubSampling)

1.4 Structure of this paper

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

1.5 Copyright infringement

[1]http://torch.ch/
[2]http://www.lua.org/about.html
[3]https://love2d.org/wiki/Main_Page
[4]https://github.com/facebook/iTorch

Chapter 2

State of the art

Being a large domain, deep learning imposes a research in multiple subdomains, neuroscience, reinforcement learning, neural networks and convolutional neural networks. That being said, this chapter is dedicated to gathering information on each of them.

2.1 Reinforcement learning

2.1.1 History

The history of reinforcement learning starts with the pavlovian experiment[6] in 1903, when Pavlov tried to demonstrate that there is a connection between conditioned and unconditioned stimulus. The unconditioned response of salivation of a dog is fired up by bringing an unconditioned stimulus as food. If we try to make the dog to salivate in the presence of a neutral stimulus(the sound of a bell) we won't succeed, but if we use in the same time the unconditioned and neutral stimulus we can make the dog salivate and transform the bell and salivation in conditioned stimulus and conditioned response.

2.1.2 Q-Learning. Sarsa

Q-Learning

Chapter 3

System design and implementation

Sys

Chapter 4

Results

Results

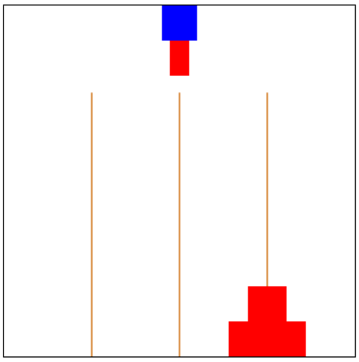

Figure 4.1:
A really Awesome Image

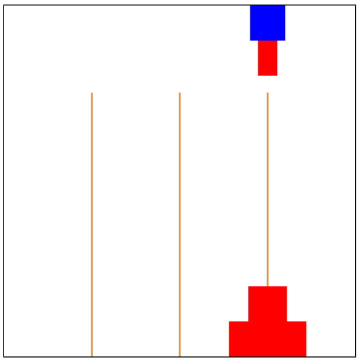

Figure 4.2:
A really Awesome Image

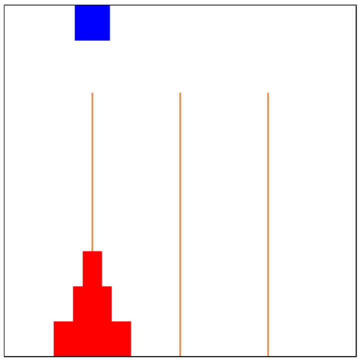

Figure 4.3:
97,6530
100,0000
93,8538
92,5261

Chapter 5

Conclusions

Conclusions

Chapter 6

Future work

Future work

Bibliography

[1] DanC. Cireşan, Alessandro Giusti, LucaM. Gambardella, and Jürgen Schmidhuber. Mitosis detection in breast cancer histology images with deep neural networks. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 411–418. Springer Berlin Heidelberg, 2013.

[2] Yonglong Tian, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Pedestrian detection aided by deep learning semantic tasks. *CoRR*, abs/1412.0069, 2014.

[3] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y. Ng. Reading digits in natural images with unsupervised feature learning. In *NIPS Workshop on Deep Learning and Unsupervised Feature Learning 2011*, 2011.

[4] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin A. Riedmiller. Playing atari with deep reinforcement learning. *CoRR*, 2013.

[5] Ronan Collobert, Koray Kavukcuoglu, and Clément Farabet. Torch7: A matlab-like environment for machine learning. In *BigLearn, NIPS Workshop*, 2011.

[6] Ivan Petrovich Pavlov. *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press: Humphrey Milford, 1927.