Image understanding with deep architectures

Mots clés :

- Directeur de thèse : Matthieu Cord
- Co-encadrant(s) :
- Unité de recherche : Laboratoire d'informatique de Paris 6
- Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal : Divers

Résumé du projet de recherche (Langue 1)

A deep architecture is hierarchical structure composing of multiple levels of nonlinear operations. Recent theoretical studies in machine learning have suggested that a deep architecture is necessary for learning complex tasks with high levels of abstraction, such as vision and language. Due to its complexity, training a deep architecture is a difficult optimization task. However, a recently developed deep learning algorithm is successful in tackling this problem. Many computer vision solutions either require the tedious task of feature design or improvise by using existing features that might not be optimized for the task. As a result, one of the motivations of deep architectures is to automatically discover abstractions from the lowest-level features to the highest-level concepts, with as little human effort as possible. This multiple-layered distributed processing is similar to the multiple levels of abstraction that humans naturally describe the world. Interestingly, research has shown that this multi-level hierarchical structure also corresponds to the organization of neuronal encoding by the visual system of our brain. OBJECTIVE: Within the framework of a deep architecture, the objective of this PhD thesis is to develop novel methods to automatically form image representations that are similar to the visual encoding of the human visual system. This new initiative represents cross-disciplinary research between statistical machine learning, computer vision and computational neuroscience. APPROACH: The project shall be divided into three parts: 1. Sparse Visual Coding Experiments in neuroscience have shown that neurons in the human visual system maximize information efficiency and minimize metabolic cost by performing sparse visual coding. This result in a neuron's firing rates distribution to be fitted by either an exponential distributions or a gamma distribution. Methods to recreate sparse codes in deep architectures are currently ad-hoc in design and do not consider the distribution of the firing rates. Our aim is to develop a method to achieve sparse visual coding of images by modeling the distributions of neural firing rates. The study of the phenomenon of temporal sparseness and population sparseness in the framework of deep architectures will also pioneered. 2 Topographical Feature Maps Neurons in sparse feature maps of the human visual system are grouped according their orientation preference, forming “pinwheels” around “orientation centres”. We want to study the automatic formation of such feature maps. Two approaches will be explored. The first approach optimizes a pulling layer after every layer in the deep architecture, and the other approach takes inspiration from neuroscience experiments by learning local lateral connections within each layer of the deep architecture. 3 Transformation Invariant Representations In the image formation process, an object may experience transformations such as translations, rotations and scaling. This is one of the main problems of computer vision, yet our visual system is able to handle such variations with ease. As a result, a large portion of this project will be dedicated to studying and solving this problem. Furthermore, there is currently very little work in deep architectures to handle this problem. Existing methods attempt to tackle the problem through an engineering approach but it is computationally expensive and non-elegant.

Résumé du projet de recherche (Langue 2)

This research objective is to design machine learning algorithms for deep architectures, leading to scalable solutions for broad competence computer vision.

Informations complémentaires (Langue 1)

Bilateral project with Singapore (I2R institute, A*star)

Informations complémentaires (Langue 2)

Joo-Hwee Lim (I²R, A*STAR, Singapore) is co-Advisor