

Guest Editorial: Deep Learning

Marc'Aurelio Ranzato · Geoffrey Hinton ·
Yann LeCun

Published online: 24 April 2015
© Springer Science+Business Media New York 2015

Deep Learning methods aim at learning feature hierarchies. Applications of deep learning to vision tasks date back to convolutional networks in the early 1990s. These methods have been the subject of a recent surge of interest for two main reasons: when labeled data is scarce, unsupervised learning algorithms can learn useful feature hierarchies. When labeled data is abundant, supervised methods can be used to train very large networks on very large datasets through the use of high-performance computers. Such large networks have been shown to outperform previous state-of-the-art methods on several perceptual tasks, including category-level object recognition, object detection and semantic segmentation.

In “Stacked Predictive Sparse Decomposition for Classification of Histology Sections” (doi:[10.1007/s11263-014-0790-9](https://doi.org/10.1007/s11263-014-0790-9)) the authors propose the use of an unsupervised feature learning algorithm for the analysis of biological tissue imagery. Biomedical applications is one domain where unsupervised learning can be very useful because of the paucity of labeled images available at training time.

M. Ranzato (✉) · Y. LeCun
Facebook AI Research, Menlo Park, USA
e-mail: ranzato@fb.com

G. Hinton
Department of Computer Science, University of Toronto,
Toronto, USA
e-mail: geoffrey.hinton@gmail.com

G. Hinton
Google Inc, Mountain View, CA, USA

Y. LeCun
Courant Institute of Mathematical Sciences, New York University,
New York, USA
e-mail: yann@cs.nyu.edu

Another important application is human pose estimation. In “Heterogeneous Multi-task Learning for Human Pose Estimation with Deep Convolutional Neural Network²” (doi:[10.1007/s11263-014-0767-8](https://doi.org/10.1007/s11263-014-0767-8)) the authors propose a convolutional neural network approach with two key ingredients. First, they jointly infer the pose and identity of human body parts, and they show that the multi-task learning setting helps to improve generalization. Second, they feed the classifier with features belonging to different layers in order to leverage information at different spatial scales and levels of abstraction. We believe that these two simple ideas are rather general and could help many other applications as well.

The last application paper is “Discriminative Deep Face Shape Model for Facial Point Detection” (doi:[10.1007/s11263-014-0775-8](https://doi.org/10.1007/s11263-014-0775-8)). Here, the authors propose a gated RBM for inferring the position of facial key points which are useful for head pose estimation, alignment, etc. This work illustrates an interesting use of a structured graphical model to jointly represent key points, shape features and latent variables encoding how these interact with pose and facial expression.

A potential drawback of deep learning methods is their high computational cost, which can hinder their use in real-time applications on portable devices. In “Spiking Deep Convolutional Neural Networks for Energy-Efficient Object Recognition” (doi:[10.1007/s11263-014-0788-3](https://doi.org/10.1007/s11263-014-0788-3)) the authors present a method to convert a convolutional neural network, without any significant decrease of accuracy, into a spiking network which can then be run on ultra-low power, spike-based neuro-morphic devices.

Finally, “A Neural Autoregressive Approach to Attention-based Recognition” (doi:[10.1007/s11263-014-0765-x](https://doi.org/10.1007/s11263-014-0765-x)) tackles another very important problem: how to scale computation to higher resolution images and achieve robustness to clutter. Instead of processing images by applying the same

processing stages to all image locations, the authors propose a method that learns where to look in images and makes predictions by integrating information across the whole sequence of glimpses. In this paper, the authors report improved accuracy by using a fully gradient-based approach.

The five papers in this special issue cover a wide range of topics, methods and applications, thereby appealing to both the experts in the field as well as to those who want a snapshot of the current breadth of deep learning research.