



第189回GMSI公開セミナー／第13回CIAiSセミナー

On the Utility of Generic Image Representations from Deep Convolutional Networks

Associate Professor **Atsuto Maki**

Computer Vision and Active Perception Lab (CVAP),
Royal Institute of Technology (KTH), Sweden

日時: 2016年3月23日(水) 15:00~16:00

場所: 東京大学工学部2号館 3F 31B会議室

Global image descriptors given by Deep Convolutional Networks (ConvNets) have been recently found very effective in a range of visual recognition tasks. In this talk, we will discuss the utility of this form of ConvNet representation: in the common scenario, a ConvNet is trained on a large labeled dataset (source) and the feed-forward units activation at a certain layer of the trained network is used as a generic representation of an input image for a (target) task.

We will first visit several factors affecting the transferability of such representations. It includes learning factors of the source ConvNet such as its architecture, distribution of the training data, etc. and also the post-learning factors such as layer choice of the trained ConvNet or dimensionality reduction. By optimising all these factors, we show that significant improvements can be achieved on various (17) standard visual recognition tasks.

What information is really encoded in such representations then? Training of a ConvNet is expected to make class-related information explicit in the representations while disregarding auxiliary information. Exploring this by looking at earlier layers, however, we find strong spatial information implicit, e.g. in the first fully connected layer. We empirically demonstrate this with evidences in a few different tasks. This is joint work with the computer vision group of CVAP.