Deep Convolution Networks as Geometric Image Representations

Stéphane Mallat
École Normale Supérieure, France

Abstract

Until recently, computer vision algorithms mostly resulted from geometric considerations over shapes, textures, motion and 3D perspective projections. Deep learning algorithms seem to ignore these geometric considerations, while providing state of the art classification results on complex image data bases. It is thus time to wonder what type of image information is extracted by these deep neural networks and why do they work so well.

We shall explain why deep convolution networks provide remarkable architectures to represent geometric image information, including shapes, textures and complex structured objects. They compute descriptors having appropriate invariance and stability to geometric transformations. The properties of these deep representations have strong similarities with Gestalt perception. However, the network filters do not need to be learned and can adapted to prior information on geometry, with multiscale wavelets. This brings us back from learning everything to more traditional computer vision approaches. Beyound low-level geometry, supervised learning seems to play an important role to build high level representations adapted to compositional properties of images in each class. Classification results will be shown on multiple image data bases, and comparisons with audio classification will also be discussed.

Keywords
Deep Convolution Networks, Geometric Image Representations