Abstract

This thesis develops new extensions to slow feature analysis (SFA) that solve supervised learning problems (e.g., classification and regression) on high-dimensional data (e.g., images) in an efficient, accurate, and principled way. This type of problems has been addressed by convolutional neural networks (CNNs) in the last decade with excellent results. However, additional approaches would be valuable, specially those that are conceptually novel and whose design can be justified theoretically.

SFA is an algorithm originally designed for unsupervised learning that extracts slow (i.e., temporally stable) features. Advantages of SFA include a strong theoretical foundation and that it might be intimately connected to learning in biological systems. One can apply SFA to high-dimensional data if it is implemented hierarchically, a technique called hierarchical SFA (HSFA). The extensions to SFA listed in the following allow the construction and training of deep HSFA networks, yielding competitive accuracy and efficiency.

Graph-based SFA (GSFA) is a supervised extension to SFA that introduces the concept of training graph, a structure in which the vertices are samples (e.g., images) and edges represent transitions between pairs of samples. Edges have weights that can be interpreted as desired output similarities of the corresponding samples. Compared to SFA, GSFA solves a more general optimization problem and considers many more transitions. Information about label (or class) similarities is encoded in the graph by the strength of the edge weights. Many training graphs are proposed to handle regression and classification problems. The efficacy of GSFA is demonstrated on a subproblem of face detection.

The exact label learning (ELL) method allows to compute training graphs where the slowest feature(s) one could extract would be equal to the label(s), if the feature space were unrestricted. In contrast to previously proposed graphs, the edge weights of the resulting ELL graphs are set precisely as needed, improving the label estimation accuracy. Moreover, ELL allows to learn multiple labels simultaneously using a single network, which is more efficient than learning the labels separately and often results in more robust features.

Hierarchical information-preserving GSFA (HiGSFA) improves the amount of label information propagated from the input to the top node in hierarchical GSFA (HGSFA). HiGSFA computes two types of features: slow features that maximize slowness, as usual, and reconstructive features that minimize an input reconstruction error, following an information-preservation goal. HiGSFA is evaluated on the problem of age estimation (along with gender and race) from facial photographs, where it yields a mean average error of 3.50 years, outperforming current state-of-the-art systems.

Among the proposed extensions, HiGSFA is the most promising. HiGSFA incorporates the other extensions and yields the best results, making this approach competitive, scalable, and robust. Moreover, HiGSFA is a versatile algorithm, allowing new technical applications and further principled extensions.