MA 598 Machine Learning Seminar
Summary 2

ID 5107

January 18, 2019

2 A Review of Convolutional Neural Networks for Inverse Problems

The paper begins with a brief introduction to inverse problems for image-processing and contrast the traditional approach with a more modern learning-based approach. The traditional approach is to solve the inverse problem by finding a suitable inverse transform via the objective function approach. More precisely, if H is the transform of interest, the reconstruction R is meant to model H and recovers an estimate of the image x from its transform y via the equation

$$R_{\text{obj}}(y) = \underset{x \in \mathscr{X}}{\arg\min}\, f(H(x), y), \tag{2.1}$$

where $f\colon \mathscr{Y} \times \mathscr{Y} \to \mathbb{R}^+$ is some appropriate measure of error. The inverse $\bar{H}^{-1}$ is usually found through filtered back projection (FBP) algorithm, or the back projection. However, these direct inverses show significant artifacts when the quality of the measurement decreases.

The proposed alternative is a learning approach, where a training set of ground truth images and their measurements $\{(x_n, y_n)\colon n = 1, ..., N\}$ is know, and a parametric reconstruction algorithm is solved by

$$R_{\text{learn}} = \underset{R_\theta, \theta \in \Theta}{\arg\min} \sum_{n=1}^{N} f(x_n, R_\theta(y_n)) + g(\theta), \tag{2.2}$$

where Θ is the set of all possible parameters, $f\colon \mathscr{X} \times \mathscr{X} \to \mathbb{R}^+$ the measure error, and $g\colon \Theta \to \mathbb{R}^+$ a regularizer whose purpose is to avoid parameter overfitting.

The paper tells of a successful application of CNN in the 2012 ImageNet Large Scale Visual Recognition Challenge. In the competition, the learning approach which achieved an error rate of 15.3% at the object localization and classification task, compared to a 26.2% error rate for the next closest method. In the subsequent competitions, CNN became the de facto method of solving the challenges, and said competitors managed to improve on the error rate of the classification task, getting it as far down as 2.6%. This is a significant improvement over more traditional methods.

Several architectures for designing the CNN are proposed including: a simple stack of series of convolutional layers and nonlinear functions; adapting CNN used for biomedical image segmentation to CT reconstruction; apply an iterative optimization algorithm, where each iteration becomes a layer in the CNN; and lastly, have the CNN learn only some part of an existing iterative method.

Some of the challenges met in constructing an effective CNN, lie in the lack of a unified way of describing the network architecture and training evaluation. Moreover, the problem (2.2) is non-convex, which means that the best solution we can hope to find is one of many local minima. Another major problem is the reliability of the reconstruction, since there is no clear way of measuring the expected error since CNN are treated as black-boxes.