SPATIAL ENHANCEMENT OF REMOTELY SENSED IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Ugo Palatucci

July 2020

Abstract

Deep Learning techniques are becoming increasingly important and some difficult tasks can be easily addressed with it. It is characterized by Neural Networks with usually more than two layers that can learn exclusively from data instead of hand-crafted features hard to tune. Computer vision has received a huge boost from these techniques surpassing for a lot of tasks the human ability like in object detection and indexing, segmentation and self-driving cars. Indeed, Convolutional Neural Networks (CNN) have proven to be perfect at extracting features from raw images and they are widely adopted in computer vision. Following this wave, the use of deep learning in remote sensing from a couple of years is becoming more predominant. Remote sensing has its unique problems that are translated into new challenges for Deep Learning. Indeed, light scattering mechanisms and sensors acquisition behaviour are arduous to model with mathematical formulas as involve object shapes, atmospheric effects and platform vibrations that can be unique for every frame. However, a Neural Network can, in theory, overcome those complex issues with appropriate dataset and strategy. Even if numerous satellites and planes register remote sensing data constantly for several years and the different space agency have collected exabyte of data, the images acquired are expensive. For this reason, the common problem is to build a large enough dataset to achieve discrete performance. Recently a new Neural Network used for super-resolution was specialized for pansharpening (called PNN) and this became the starting point to accost the pansharpening with Deep Learning techniques. For the limited dataset used, it was essential to build a network with few weights to train. Pansharpening is a particular case of super-resolution in which two images, one multispectral and one panchromatic, are fused to combine the properties of both. The main goal is to reach, in a new image, resolutions that cannot be obtained by a single sensor. This constitutes an important challenge for Deep Learning as there is not a target available for the training of the network. In previous researches, to obtain a training target, the images have been downsampled. The downsampled images have been used as training inputs and the original ones as the targets. Our research aim is to explore a different approach called in Machine Learning Unsupervised Learning in which a target is not necessary for the training. This requires the definition of a no-reference differentiable function

and two no-reference quality assessment indexes have been tested for this purpose: QNR and HQNR. Furthermore, we had to implement new training and validation processes to compare different methods. It was demonstrated that the QNR index as full resolution loss is not appropriate and, despite we performed some approximations to make the index differentiable, the HQNR shows better results related to the previous approach. Tests on an ideal training scheme based on the target (Ground truth) show that there is room for further improvements. For these reasons, the current research lay the foundation of future development.