Abstract

Andrea Bordone Molini

April 5, 2021

In this thesis we present novel deep learning methods to tackle two inverse problems in imaging i.e., super-resolution and denoising. These enhancement tasks are often used as a pre-processing step by many pattern recognition and analysis algorithms as they can leverage an image reconstruction with enriched spatial information and details that eases image understanding, thereby improving their performance.

Recently, convolutional neural networks (CNN) have been successfully applied to many remote sensing problems. However, deep learning techniques for multi-image super-resolution from multi-temporal unregistered imagery have received little attention so far. In the first part of this thesis we propose a novel CNN-based technique that exploits both spatial and temporal correlations to combine multiple images. This novel framework integrates the spatial registration task directly inside the CNN, and allows to exploit the representation learning capabilities of the network to enhance registration accuracy. The entire super-resolution process relies on a single CNN with three main stages: shared 2D convolutions to extract high-dimensional features from the input images; a subnetwork proposing registration filters derived from the high-dimensional feature representations; 3D convolutions for slow fusion of the features from multiple images. The whole network is trained end-to-end to recover a single high resolution image from multiple unregistered low resolution images. As opposed to the vast majority of the work in literature that use synthetic datasets, the training procedure is carried out through a set of real-world low resolution observations and the corresponding high resolution image for the same scene, captured from the same platform. This method is the winner of the PROBA-V super-resolution challenge issued by the European Space Agency.

The second contribution of this thesis is a deep learning method to tackle a denoising task in the field of synthetic aperture radar (SAR) remote sensing. Information extraction from SAR images is heavily impaired by speckle noise, hence despeckling is a crucial preliminary step in scene analysis algorithms. The recent success of deep learning envisions a new generation of despeckling techniques that could outperform classical model-based methods. However, current deep learning approaches to despeckling require supervision for training, whereas clean SAR images are impossible to obtain. In the literature, this issue is tackled by resorting to either synthetically speckled optical images, which exhibit different properties with respect to true SAR images, or multi-temporal SAR images, which are difficult to acquire or fuse accurately. In this paper, inspired by recent works on blind-spot denoising networks, we propose a self-supervised Bayesian despeckling method. The proposed method is trained employing only noisy SAR images and can therefore learn features of real SAR images rather than synthetic data. Experiments show that the performance of the proposed approach is very close to the supervised training approach on synthetic data and superior on real data in both quantitative and visual assessments.