ISPRS Journal of Photogrammetry and Remote Sensing xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Skipping the real world: Classification of PolSAR images without explicit feature extraction

Ronny Hänsch *, Olaf Hellwich

Technische Universität Berlin, Computer Vision & Remote Sensing, Germany

ARTICLE INFO

Article history:
Received 17 March 2017
Received in revised form 28 November 2017
Accepted 30 November 2017
Available online xxxx

Keywords:
Random Forest
PolSAR
Classification
Feature learning

ABSTRACT

The typical processing chain for pixel-wise classification from PolSAR images starts with an optional preprocessing step (e.g. speckle reduction), continues with extracting features projecting the complex-valued data into the real domain (e.g. by polarimetric decompositions) which are then used as input for a machine-learning based classifier, and ends in an optional postprocessing (e.g. label smoothing). The extracted features are usually hand-crafted as well as preselected and represent (a somewhat arbitrary) projection from the complex to the real domain in order to fit the requirements of standard machine-learning approaches such as Support Vector Machines or Artificial Neural Networks. This paper proposes to adapt the internal node tests of Random Forests to work directly on the complex-valued PolSAR data, which makes any explicit feature extraction obsolete. This approach leads to a classification framework with a significantly decreased computation time and memory footprint since no image features have to be computed and stored beforehand. The experimental results on one fully-polarimetric and one dual-polarimetric dataset show that, despite the simpler approach, accuracy can be maintained (decreased by only less than 2% for the fully-polarimetric dataset) or even improved (increased by roughly 9% for the dual-polarimetric dataset).

1. Introduction

Synthetic Aperture Radar (SAR) is an active air- or space-borne sensor that emits a microwave signal and measures amplitude and phase (thus a complex number) of the echo which is backscattered at the ground. As active sensor it is independent of daylight and thus contrasts optical and hyperspectral sensors. Due to the electromagnetic properties of the used microwave it is less influenced by weather conditions and can penetrate clouds, dust, and to some degree even vegetation. Polarimetric SAR (PolSAR) transmits and measures in different polarisations which enables it to provide information contained neither in single channel SAR nor in other remote sensing data. The measured echoes depend on several properties including moisture, surface roughness, as well as object geometry and are therefore highly correlated to specific object categories. Due to these advantages there are nowadays many modern sensors available that provide PolSAR data, i.e. images that contain complex-valued vectors in each pixel (see Section 2).

The increasing amount of data, but also methodological problems of manual interpretation such as loss of information during visualisation and image effects human operators are unaccustomed with, led to a dire need of automatic procedures for PolSAR image analysis. The generation of semantic maps of land use/cover by pixelwise classification is one of the most typical and most important tasks of automatic interpretation of remote sensing images. It is usually solved within a supervised machine learning framework, where the internal parameters of a generic model are adjusted based on training data which contains the desired class labels alongside with the image data. The corresponding literature can be coarsely divided into two groups: Approaches that directly work on the (Pol)SAR data by modelling its statistical properties and methods that apply general purpose classifiers to extracted features. The first group has seen a lot of success in the early years of (Pol)SAR image classification and continues to propose new models that are better able to cope with the challenges of modern data. These mostly parametric approaches include models based on distributions such as Rayleigh (Kuruoglu and Zerubia, 2004), generalized Gaussian (Moser et al., 2006), generalized Gamma (Li et al., 2011), Weibull (Taravat et al., 2014), and Fisher (Tison et al., 2004) as well as models such as Finite Mixture Models

* Corresponding author.
E-mail address: r.haensch@tu-berlin.de (R. Hänsch).

https://doi.org/10.1016/j.isprsjprs.2017.11.022

Please cite this article in press as: Hänsch, R., Hellwich, O. Skipping the real world: Classification of PolSAR images without explicit feature extraction. ISPRS J. Photogram. Remote Sensing (2017), https://doi.org/10.1016/j.isprsjprs.2017.11.022

(Krylov et al., 2011) and approaches based on the Mellin transform (Nicolas and Tupin, 2016) that combine and generalize different distributions. These methods usually make parametric assumptions about the underlying data distributions, which tend to fail within heterogeneous regions or with contemporary high-resolution SAR data. If the ultimate goal is to derive a classification decision, discriminative approaches have been shown to be easier trained and to be more robust and accurate as such generative models. The general approach is to extract (often hand-crafted and class-specific) image features, which are descriptive enough for the classification task at hand, and use them as input to typical classifiers such as Support Vector Machines (SVMs, e.g. in Mantero et al., 2005), Multi-Layer Perceptrons (MLPs, e.g. in Bruzzone et al., 2004), or Random Forests (RFs, e.g. in Hänsch and Hellwich, 2010b). However, most of such classifiers are designed for real-valued input only and cannot easily be extended to the complex domain. This problem is solved by extracting real-valued features which are then used by the classifier. This feature extraction step states a couple of problems: Firstly, the extracted features are mostly hand-crafted and preselected for a specific classification task. The design and preselection of hand-crafted features usually involves manual trial and error. The creation of a descriptive feature set for typical tasks such as land use classification continues to be an active field of research. Despite an abundance of available features capturing polarimetric (such as polarimetric decompositions see e.g. Cloude and Pottier, 1996) or textural (e.g. as in He et al., 2013) information, it is unclear which combination yields the best results within the applied classification framework. Secondly, the computation of a diverse and descriptive set of features increases the computational load of the whole processing chain. On the one hand does the computation time increase. The features need to be computed from the PolSAR image and the computation time of many classifiers increases with a higher dimensional input space. On the other hand does the memory footprint of the method increase as well, since all the computed features need to be kept in memory to be accessible by the classifier. Thirdly, the more-or-less arbitrary projection of complex-valued data into the real-valued domain causes at best a dependency of the classification result on the used projection, i.e. feature extraction method. In the worst case it means a loss of valuable information and thus sub-optimal results.

In principle, there are two approaches to cope with the aforementioned problems: Either avoiding the extraction of real-valued features and aiming to model the relation between the PolSAR data itself and the class label directly, or using an exhaustive feature set which at least potentially contains all information necessary to solve the classification task.

Besides generative statistical models as discussed above, only very few approaches belong to the first group and aim to work directly on the complex-valued PolSAR data instead of extracting real-valued features. One example are complex-valued MLPs which have been applied to the task of land use classification e.g. in Hänsch (2010). Within these networks, input, weights, as well as output are modelled as complex-valued numbers. Thus, they can directly be applied to complex-valued data such as PolSAR images at the cost of a slightly more complicated training procedure as MLPs usually use bounded and analytical functions which do not exists for the complex domain. Another possibility are SVMs with complex-valued kernels tailored to the analysis of PolSAR images as in Moser and Serpico (2014). However, setting the corresponding hyperparameters is a non-trivial task and often results in time consuming grid search within the parameter space.

Approaches of the second group that are based on quasi-exhaustive feature sets usually aim to reduce this set to a meaningful subset which is then used by a standard classifier. Typical examples are principle component analysis (Licciardi et al., 2014), independent component analysis (Tao et al., 2015), linear discriminant analysis (He et al., 2014), or genetic algorithms (Haddadi et al., 2011). Other works rely on classifiers such as Random Forests (RFs) which are able to handle high-dimensional and partially undescriptive feature spaces due to an inbuilt feature selection. Random Forests have been extensively applied to the classification of remotely sensed data in general and PolSAR images in particular (see e.g. Belgiu and Dragut, 2016 for a recent review). The work in Hänsch (2014) extracts hundreds of real-valued features from a given PolSAR image and uses a RF to focus on the most descriptive ones. The work in Tokarczyk et al. (2015) pushes this idea even further by extracting thousands of (easy to compute) features and relying on boosted decision stumps to select relevant features for land cover classification from optical images. As exhaustive feature sets are less likely to be biased towards specific classification tasks and more likely to contain useful information for the task at hand, those approaches are very generic in the sense that they achieve remarkable results for various classification problems (with retraining, but without redesigning the processing chain i.e. implementing a different feature extraction). However, the large feature sets represent a considerable burden regarding memory footprint and computation time.

The solution to this problem is to parametrize feature extraction and include it into the optimization problem during the training of the classifier. This feature learning works on the data itself (or very basic features) which allows to compute task-optimal features on the fly and only when and where needed. Convolutional Networks (ConvNets) are the most well known examples of modern feature learning approaches and have been applied to close range images (e.g. Ranzato et al., 2007) as well as to remote sensing images (e.g. in Mnih and Hinton, 2012). As ConvNets are - similar to MLPs - originally designed for real-valued data, most methods (e.g. Zhou et al., 2016) extract real-valued features in order to classify PolSAR images, while only few works use complex-valued networks which are tailored to the complex values of PolSAR images (e.g. in Hänsch and Hellwich, 2010a; Zhang et al., 2017). Despite the potential power and large success of ConvNets in other computer vision domains, their sensitivity to training parameters as well as their need for large amount of (labelled) data often lead to results that are barely competitive to conventional approaches to classification of remote sensing images (Tokarczyk et al., 2012).

Another example of modern feature learning approaches are RFs where the internal node tests of the individual trees are especially designed for image data: Instead of treating the input samples as n-dimensional (feature) vectors in $\mathbb{R}^n$, image or feature patches (i.e. elements of $\mathbb{R}^{w \times w \times c}$ with patch size w and c channels) are used. The corresponding node tests are then defined over those patches, e.g. by performing comparisons of random pixel pairs within the patch. These RFs are heavily used in the computer vision community to analyse close-range optical images (see e.g. Criminisi and Shotton, 2013; Fröhlich et al., 2012), but also have found their way into remote sensing (e.g. Fröhlich et al., 2013). The work of Hänsch (2014) applies these kind of Random Forests in a sophisticated multi-stage framework for generic object classification from image data with a focus on PolSAR images: A first step extracts several low-level features from the provided image data. As this feature set might contain redundant as well as non-descriptive features for a specific classification task, Random Forests are used which on the one hand reject meaningless features due to their built-in feature selection and on the other hand are able to perform a spatial analysis of the provided feature maps. This first classification is then further processed by a semantic segmentation and a second classification step based on a similar Random Forests which exploits segment-based features.

Please cite this article in press as: Hänsch, R., Hellwich, O. Skipping the real world: Classification of PolSAR images without explicit feature extraction. ISPRS J. Photogram. Remote Sensing (2017), https://doi.org/10.1016/j.isprsjprs.2017.11.022