Pattern Recognition ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Pattern Recognition

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

Discriminant deep belief network for high-resolution SAR image classification

Zhiqiang Zhao*, Licheng Jiao*, Jiaqi Zhao, Jing Gu, Jin Zhao

Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, International Research Center for Intelligent Perception and Computation, Joint International Research Laboratory of Intelligent Perception and Computation, Xidian University, Xi'an Shaanxi Province 710071, China

ARTICLE INFO

Article history:
Received 26 October 2015
Received in revised form
19 May 2016
Accepted 24 May 2016

Keywords:
Discriminant feature learning
Deep belief network
SAR image classification
Ensemble learning
Similarity measurement

ABSTRACT

Classification plays an important role in many fields of synthetic aperture radar (SAR) image understanding and interpretation. Many scholars have devoted to design features to characterize the content of SAR images. However, it is still a challenge to design discriminative and robust features for SAR image classification. Recently, the deep learning has attracted much attention and has been successfully applied in many fields of computer vision. In this paper, a novel feature learning approach that is called discriminant deep belief network (DisDBN) is proposed to learning high-level features for SAR image classification, in which the discriminant features are learned by combining ensemble learning with a deep belief network in an unsupervised manner. Firstly, some subsets of SAR image patches are selected and marked with pseudo-labels to train weak classifiers. Secondly, the specific SAR image patch is characterized by a set of projection vectors that are obtained by projecting the SAR image patch onto each weak decision space spanned by each weak classifier. Finally, the discriminant features are generated by feeding the projection vectors to a DBN for SAR image classification. Experimental results demonstrate that better classification performance can be achieved by the proposed approach than the other state-of-the-art approaches.

1. Introduction

With inherent properties, such as high-resolution, wide converge and day and night systematic imaging capabilities, the synthetic aperture radar (SAR) images have been widely used in a variety of civil and military applications. They contain quite different information with images which are acquired from optical sensors [1,2]. They have been utilized in a variety of earth observation applications, such as land-cover classification [3–5], urban detection and extraction [6,7], and environment risk evaluation [8–10]. As dielectric and geometric properties become relevant for the backscattering in the microwave region, the radar image emphasize the relief and morphological structure of the observed terrain as well as the ground conductivity. Due to the limited spatial resolution of the first generation of SAR data (such as ERS-1/2 satellites and ENVISAT), the number of elementary scatters present in a single resolution cell was normally very large. The speckle are seriously as a result of the constructive and destructive interferences of the fields backscattered by the individual scatter [11]. In recent decades, many high-resolution space-borne and air-borne SAR sensors have been launched, such as TerraSAR-X/TanDEM-X satellites, COSMO-SkyMed satellite, ALOS PALSAR-II satellite, F-SAR, UAVSAR, and GeoSAR [12–14]. With the increase of the spatial resolution, it is more common to find that a resolution cell contains only a few or even just one single domain scatter. It makes the statistics of SAR data with a low and median resolution which is no longer applicable to model the content of the high-resolution SAR images. Therefore, the major work of this paper focuses on how to extract discriminative and robust features for high-resolution SAR image classification.

In recent years, several features have been developed to characterize the content of SAR images. The texture descriptors [15,16] have been widely utilized to model the intensity and amplitude value of SAR images. They are some statistics computed from the gray level co-occurrence matrix (GLCM), such as energy, entropy, inertia, contrast, correlation, and roughness. By modeling the global properties of SAR image, texture descriptors have played an important role in many fields of SAR image processing. However, the complex structures in a local content are becoming more critical for SAR image processing with the increase of the ground resolution. In order to model local details, several multi-scales filters, such as Gabor filter bank [17–19] and wavelet transform [20,21], have been utilized to extract local features of SAR images.

* Corresponding authors.
E-mail addresses: zhq_zhao@yeah.net (Z. Zhao), lcjiao@mail.xidian.edu.cn (L. Jiao).

http://dx.doi.org/10.1016/j.patcog.2016.05.028
0031-3203/

However, it is confusing to determine an appropriate number of orientations and scales of these filters to exploit enough information for SAR image classification. In addition, some high-level features, such as bag-of-words (BoW) [22–24] and sparse representation [25,26], have also been introduced to characterize the content of SAR images. All of these high-level feature descriptors have excellent capacity for SAR image representation, but they do not have enough discriminative information for classification tasks [27]. In order to characterize SAR image discriminatively, several features that appeared in the literature of computer vision can also be employed to extract the discriminant information, such as metric learning [28,29] and Fisher linear discriminative analysis [30–32]. However, it is impossible to obtain enough labeled data for supervised learning with the increasing volume of SAR images [33,34]. Therefore, it is eagerly to obtain an effective approach that can learn discriminative features for SAR images in an unsupervised manner.

Fortunately, the deep learning [35–38] has been attracting great attention in recent years, as they could learn high-level features and representations from a large amount of unlabeled data. The deep architectures attempt to learn a hierarchical structure by learning simple concepts first and then successfully building up more complex concepts [39]. In particular, the deep belief network (DBN) [40,41] has been widely used in the community of deep learning. It can be treated as a multi-layer generative graph model, in which the statistical dependencies among each units in the lower layer are encoded in the higher layer. The DBN can be greedily trained layer-by-layer in an unsupervised manner, where each layer is typically made of a restricted Boltzmann machine (RBM) [42–44]. The RBM is an undirected, generative energy-based model with an input layer and single hidden layer, in which the connections only exist between the visible units of the input layer and the hidden layer. The DBN has been successfully applied to learn high-level features in a wide variety of domains, such as handwritten digits recognition [40], image segmentation [45,46], and hyperspectral image classification [47]. As the DBN has excellent capacity to learn high-level features, it has also been exploited to learn high-level features for high-resolution SAR images classification in this work.

Inspired by the prototype theory [48,49] of cognitive science, the discriminative information of a SAR image patch can be exploited by comparing some prototypes that are learned from the training SAR image patches. Recently, an ensemble projection (EP) approach has been introduced by Dai et al. [50,51] to learn high-level features based on the prototype theory and ensemble learning. Specifically, a set of weak classifiers is constructed for ensemble learning, in which each one is trained by some samples that are captured with both the local-consistency and the exotic-inconsistency assumption. Then, an image is projected to each decision space generated by each weak classifier to produce several projection vectors. All of these projection vectors are concatenated together to produce the high-level feature of the considered image. However, with the increasing number of learned prototypes and the trained weak classifiers, a high-dimensional feature can be resulted by this simple concatenation. It may induce many difficulties for SAR image classification, such as heavy computational burden and redundancy information. Meanwhile, the complementary information between each weak classifier is not fully exploited by this approach.

In this work, a discriminant deep belief network which is denoted as DisDBN is proposed to learn high-level discriminative features to characterize the SAR image patches by combining the ensemble learning and DBN. The pipeline of the proposed approach is shown in Fig. 1. It consists of two major parts of the proposed approach, which are weak classifiers training and high-level feature learning. In the first part, a set of weak classifiers is trained based on some prototypes which are learned in an unsupervised manner. While in the second part, a DBN is utilized to learn the high-level feature descriptors from all of the projection vectors which are generated by projecting each SAR image patch onto all of weak decision spaces defined by each weak classifier. Specifically, the major contributions of this work are listed as follows.

1. A DisDBN is proposed in this work, which combines the ensemble learning and DBN to learn high-level discriminative feature of the SAR image patch in an unsupervised manner.
2. Both the *clustering*-based and *instance*-based prototypes learning approaches are investigated in this work to produce prototypes of SAR image patches. Based on these learned prototypes, several weak decision spaces are constructed to fully exploit the discriminative information of the specified SAR image patch by projecting it into each weak decision space.
3. All of these projection vectors will be employed to synthesis a high-level feature of the specified SAR image patch in a hierarchy manner. It can fully exploit the complementary information between all of the weak classifiers.

Experimental results of several SAR images demonstrate the effectiveness and robustness of the proposed discriminative feature learning approach.

The remainder of this paper is organized as follows. The method of discriminative deep belief network is introduced in Section 2. Several experiments are presented in Section 3 to evaluate the performance of the proposed approach. In the last section, some conclusions and future works are presented.

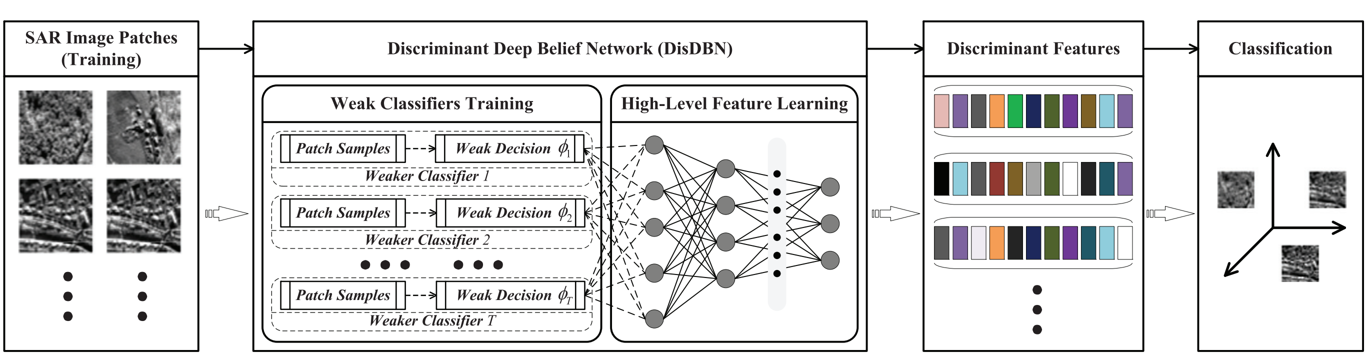

Fig. 1. Pipeline of the proposed discriminative feature learning approach (DisDBN). It consists of two parts, which are weak classifier training and followed by a deep belief network (DBN) for high-level feature learning.

Please cite this article as: Z. Zhao, et al., Discriminant deep belief network for high-resolution SAR image classification, Pattern Recognition (2016), http://dx.doi.org/10.1016/j.patcog.2016.05.028