Computers and Electrical Engineering 000 (2017) 1–17

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Object-oriented and multi-scale target classification and recognition based on hierarchical ensemble learning☆

Yang Liu[a,b], Feng-bin Zheng[a,b,*]

[a] College of Computer Science and Information Engineering, Henan University, Kaifeng 475004, PR China
[b] Laboratory of Spatial Information Processing, Henan University, Kaifeng 475004, PR China

ARTICLE INFO

Article history:
Received 7 June 2016
Revised 26 December 2016
Accepted 26 December 2016
Available online xxx

Keywords:
Target classification and recognition
Deep spiking convolutional neural network
Hierarchical latent Dirichlet allocation
Brain inspired computing
Multimedia neural cognitive computing
High resolution remote-sensing image

ABSTRACT

Target classification and recognition (TCR) of high resolution remote-sensing image is the important ability for earth observation system and unmanned autonomous system. It is difficult to improve the precision of TCR because of different imaging mechanism. In this paper, we propose a brain-inspired computing model for TCR using cognitive computing and deep learning. Accordingly, we have built an ensemble learning algorithm based on deep spiking convolutional neural network and hierarchical latent Dirichlet allocation. The hierarchical features were extracted from remote-sensing image. Then a TCR algorithm for small sample sizes and complex target was designed, which uses the incremental and reinforcement learning based on object-oriented and multi-scale data argumentation. Experimental results demonstrate that our algorithm has state-of-the-art performance on public data sets of optical remote-sensing image and synthetic aperture image. The model proposed can provide reference to explore an essential significance in brain-inspired intelligence, and has significant value in military and civil affairs.

1. Introduction

Automatic Targets Recognition (ATR) is an important intelligent information acquisition tool for the intelligent robotics and autonomous equipment. ATR is also an essential image interpretation method in high-resolution earth observing system, which has important value of theoretical and practical in military and civil affairs. ATR algorithms have different task levels, including target detection, target discrimination, target classification, target recognition, target confirmation, and target characterization. In general, target detection and classification reflect the fundamental problems of ATR. Target recognition and characterization aims to improve the performance of ATR. Further, these levels can be divided into three main stages. In the primary stage, target detection extracts regions containing potential targets from entire images. In the intermediate stage, target classification extracts coarse categorization of moderate resolution image based on the target discrimination. In the advanced stage, target recognition extracts fine categorization of high-resolution image based on the target classification. Specially, there is no essential difference between target classification and target recognition in algorithms. As that, we use Target Classification and Recognition (TCR) to mean both of them in this paper.

The definitions of ground objects depend upon the application purposes in TCR problems. It is primarily focused on these important targets such as ships, tanks, aircraft, oil depots, bridges, airports and harbors. ATR has the different imaging

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Asoociate Editor Dr. G. Botella.

* Corresponding author.
E-mail addresses: ly.sci.art@gmail.com (Y. Liu), zhengfb@henu.edu.cn (F.-b. Zheng).

http://dx.doi.org/10.1016/j.compeleceng.2016.12.026
0045-7906/

mechanism of sensors such as optics (e.g. visible-light, infrared and laser), radar and sonar, so TCR algorithm has to work with different kinds of images such as optical images (e.g. visible-light, hyperspectral and thermal infrared) and Synthetic Aperture Images (SAI) (e.g. synthetic aperture radar, synthetic aperture sonar and synthetic aperture ladar). Remote-sensing images reflect distinct information from a scene or target in spectrum, spatial and phase. All these facts make it difficult to improve the precision of TCR. Currently, most TCR algorithms are not robust, and many systems have problems with both false alarm and noise interference.

The objective of this paper is to design a high precision and robust TCR model and algorithm for optical remote-sensing image and SAI. Considering that semantic has a hierarchy in the remote-sensing image, we explore an effective integrated framework to address the TCR problem by brain-inspired computing (BIC). In addition, an ensemble TCR algorithm is proposed based on the hierarchical hybrid model by following Deep Spiking Convolutional Neural network (DSCN) and Hierarchical Latent Dirichlet Allocation (HLDA).

The remainder of the paper is organized as follows, Section 2 reviews the related work. In Section 3, we present the TCR framework based on hierarchical ensemble learning. Section 4 introduces incremental and reinforcement learning of TCR algorithm based on object-oriented and multi-scale data augmentation. Experimental results and analysis are given in Section 5, while conclusions and work-in-progress are reported in Section 6.

2. Related work

Essentially, TCR is a mapping between the original data space and high-level semantic space. There are many typical methods presented for optical image and SAI. Generally, the visual features (e.g., geometric size, gray statistical, edge shape, texture) are applied to optical images. Many methods have been designed for the TCR of the optical images such as fractal model, classification decision tree, fuzzy theory, rough set, probabilistic generative model, SVM, KNN, HDR, CNN [1] and RBF [2]. A number of algorithms have been employed for the TCR of the infrared image such as Adaboost, manifold learning [3], Fisher optimal discriminant with gray gradient, fractal and LBP. Unlike optical images, SAI needs unique methods. There are some methods presented the TCR of the Synthetic Aperture Radar (SAR) image, for instance, the template library matching, scatter graph matching, similarity of peak and shadow, Adaboost [4], HRS [5], sparse representation classification [6], 3D model with shape [7], texture, electromagnetic scattering, transform feature and local invariant. Similarly, some methods were proposed in SAS image TCR such as DS evidence theory, manifold learning [8], MCMC, BBA [9], shadow shape, geometric and active contour. The key objectives of TCR algorithms mostly focus on accuracy and speed. In traditional TCR methods, quite few studies have been conducted in optical images with complex backgrounds. It needs to seek effective and robust features for SAI. Recently, deep learning and cognitive computing have been applied to TCR methods.

2.1. Deep learning

Deep learning is a method of neural computing. It is a general feature learning process from data, and transforms the original data into high-level abstract representations [10]. Deep learning has achieved the breakthrough in image recognition, speech recognition and natural language understanding. Similar with LeNet, many deep learning models, such as AlexNet, ZFNet, VGGNet, GoogLeNet and ResNet, have achieved better results. There are more and more cognitive mechanisms, such as perception, attention, memory and emotion, have been employed in deep learning, and a lot of novel neural networks have been designed. In 2012, Hinton's research group has employed AlexNet and achieved top-5 test error rate of 15.32% on ImageNet [11]. They proposed many ticks such as the dropout, ReLU and data augmentation for network training. In 2014 CVPR, DeepID [12] and DeepFace [13] achieved 97.45% and 97.35% face recognition rate on the LFW. Recurrently, neural networks and long short-term memory are the most effective models for natural language processing and sequence signal analysis, for instance, handwriting recognition and speech recognition.

2.2. Cognitive computing

In general, cognitive computing belongs to probability methods of statistics learning theory. It is an advantaged tool that learns at scale, reason with purpose and interact with humans naturally, and can analyze and understand unstructured data. Probabilistic Graphical Model (PGM) is uncertain knowledge representation and processing method based on a graph model and the probabilistic theory. HLDA model [14] is a Hierarchical Topic Model (HTM) for natural language processing and information retrieval. HLDA is also a non-parametric generative model, which allows for a countable infinite number of topics, and describes the parent-children relationship between topic distribution and document with the hierarchical topic tree. Nodes of the topic tree have the hierarchical relationships from abstract to concrete in HLDA.

3. Hierarchical ensemble learning of TCR

The human brain is a complex system, and has a hierarchical modular organization. In fact, the hierarchy also exists in the natural signal such as, one-dimensional language and words, two-dimensional image and sound, and three-dimensional video and animation. Higher-level features are derived from lower-level features to form a hierarchical representation. For example, images have the hierarchical structure such as edges, motifs, parts and objects. Similarly, speech sounds have the