Pattern Recognition 73 (2018) 259–274

Contents lists available at ScienceDirect

Pattern Recognition

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

Remote sensing image analysis by aggregation of segmentation-classification collaborative agents

Andrés Troya-Galvis [a,*], Pierre Gançarski [a], Laure Berti-Équille [b]

[a] Univeristy of Strasbourg, Icube UMR CNRS 7357, France
[b] Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar

ARTICLE INFO

Article history:
Received 8 March 2017
Revised 24 August 2017
Accepted 27 August 2017
Available online 30 August 2017

Keywords:
Segmentation
Classification
Collaborative approaches
Remote sensing

ABSTRACT

In this article we present two different approaches for automatic remote sensing image interpretation which are based on a multi-paradigm collaborative framework which uses classification in order to guide the segmentation process. The first approach applies sequentially many one-vs-all class extractors in a manner inspired by cascading techniques in machine learning. The second approach applies many collaborating one-vs-all class extractors in parallel. We show that the collaboration of the segmentation and classification paradigms result in a remarkable reduction of segmentation errors but also in better object classification in comparison to a hybrid pixel-object approach as well as a deep learning approach.

1. Introduction

Automatic interpretation of remote sensing images is a very challenging problem and it is rapidly becoming an indispensable requisite for many applications such as disaster management [1,2], forest mapping [3], urban planning [4,5], among others. Indeed, images of increasingly high spatial resolution are getting acquired more frequently in such a way that their treatment without any kind of computer assistance becomes intractable.

Object Based Image Analysis (OBIA) [6] techniques are generally used for dealing with Very High Spatial Resolution (VHSR) images. Indeed the object based representation allows for a better description of the image, so knowledge can be more easily extracted. From a ready-for-analysis image, OBIA methods generally use classical segmentation approaches to partition the image into homogeneous regions, hoping for a one-to-one mapping between those regions and geographic objects in the image. Many features are then computed to describe these segments. As a final step, classification approaches are employed to get an entirely labelled image which can be analysed by an expert.

One of the main drawbacks of OBIA is that the classification results are heavily dependant on the segmentation results [7], so classifiers often require ideal segments which perfectly match geographic objects in the image in order to give accurate predictions in the classification step. However, such perfect segments are rarely obtained regardless of the segmentation approach used. Indeed, it is commonly known that due to the variability and complexity of remote sensing images, it is extremely difficult –if not impossible– to find an algorithm (with its corresponding parameters) which produces a full mapping between the segments and the geographic objects in the image. Thus, classical segmentation approaches are not really suitable for OBIA since they tend to output a non-negligible number of over- and/or under-segmented regions and are strongly dependent on their parameters. Moreover, many objects commonly found in VHSR images are composed of several non-homogeneous regions; the roof of a house for example, is often composed of dark and light regions which not likely to be segmented together by a classical segmentation algorithm. Unfortunately, manual correction of such problematic segments is a tedious and time consuming task for the expert. Our idea to automatize this operation is to employ information implicitly encoded into one or more classification models trained offline, in order to guide a segmentation improving process. Our main hypothesis relies on the fact that when the classification of a given segment is confident, then this segment is likely to correspond to an actual geographic object of interest. Reciprocally, the more a given segment corresponds to a geographic object the higher its classification probability should be, provided that the classification model correctly encodes the semantic information related to the class of interest. In this article, we propose to aggregate many segmentation-classification collaborating agents, each one looking for a single class. More precisely, we propose a sequential approach inspired from machine learning techniques such as cascad-

* Corresponding author.
E-mail addresses: troyagalvis@unistra.fr, ac.troya@gmail.com (A. Troya-Galvis), gancarski@unistra.fr (P. Gançarski), lberti@qf.org.qa (L. Berti-Équille).

http://dx.doi.org/10.1016/j.patcog.2017.08.030
0031-3203/

ing [8] and boosting [9]; and a parallel approach inspired from collaborative clustering techniques such as SAMARAH [10].

The following sections are organized as follows. Section 2 presents some work related to our research. Section 3 presents the collaborative framework CoSC on which our proposition is based. Section 4 introduces a sequential scheme of segmentation and classification, and a parallel collaborative approach. Section 5 studies some properties of the proposed methods and presents a comparative study with a hybrid OBIA method as well as a deep learning method and discusses our results. Finally, Section 6 concludes and gives some research perspectives.

2. Related work

Several attempts to combine segmentation and classification methods have been proposed in order to improve the segmentation and classification processes. We can distinguish two main kinds of such methods.

2.1. Mono-class approaches

Mono-class extraction allows for focusing on class-specific features leading to a better extraction of objects from a given class [11]. Many efforts have been put into methods for extracting particular objects. For instance, buildings [12] or roads [13–16]. However, such methods usually are very specific and dependent on the application as well as on the type of object. More general approaches for mono-class extraction have also been proposed. For instance, in [17], the authors propose a method for learning fuzzy rules from expert-given examples in order to extract a given class of objects. In [18,19] the authors propose a generic collaborative framework called CoSC which employs an arbitrary chosen classifier (trained offline) to guide a segmentation modification algorithm. Finally, in [20], the authors propose a method that generates a large number of different meanshift segmentations, then a one-class svm classifier is applied to select the best segments based on their classification score.

2.2. Multi-class approaches

Many applications such as urban planning require a full land-cover map in order to carry out particular analysis [21–23]. Thus, multi-class extraction schemes have also been explored. For instance, a multi-agent system is proposed by Mahmoudi et al. in [24]. Their framework employs many class-specific agents which work concurrently to extract distinct classes, while a conflict-solving agent manages the global process. Similarly in [25], Hoffman et al. presented an agent-based image analysis framework, where each segment become a self-modifying agent seeking for correspondence to rules given by a domain ontology. Another example is given in [26], where Lizarazo and Elsner apply pixel-wise fuzzy classification algorithms for different classes to create what they call fuzzy segments, which are then merged by some given logical rules to construct a single labelled image. Furthermore, Tarabalka et al. in [27] present a hybrid method combining pixel- and object-based approaches; they refine a pixel-wise SVM classification by using a majority voting scheme in order to determine the class of segments obtained by EM segmentation. More recently, in [28] Csillik proposed an approach which is based on super-pixels, aiming to reduce the segmentation computing time by starting the segmentation process from compact groups of similar pixels that can be generated very fast, as opposed to classical segmentation approaches starting from raw pixels. Additionally, semi-supervised and unsupervised techniques for OBIA are also actively being explored. For instance, in [29] Sublime et al. develop a multi-scale version of the semantic-rich ICM algorithm that facilitate the mapping of clusters to semantic classes of interest by interpreting an affinity matrix which summarizes inter-cluster relationships. Another example can be found in [30] where Tuia et al. propose a method for semi-supervised labelling of a segmented image by using an active learning approach that does not require any previous training of a classifier.

These solutions have all their advantages and drawbacks. Multi-agent systems have accurate results but they are excessively complex or depend on very domain-specific knowledge which is in general hard to obtain or model. Pixel-object hybrid methods efficiently inject low level information from pixels into the OBIA framework but are highly dependent on the initial segmentation which is never reconsidered. In this article we propose a simple solution to reconsider and modify the initial segmentation which is easy to implement and can be adapted to many application schemes and which appears as robust to the initial segmentation.

3. Collaborative segmentation and classification: CoSC

Our proposition is built upon the CoSC framework presented in [18,19] which we briefly describe in order to make this article self-containing. CoSC is designed to extract a single thematic class from an image. Fig. 1 represents the mono-class CoSC process. It takes three input parameters: an image, an initial segmentation $\mathcal{S}$ and a one-vs-all classifier $\mathfrak{C}_{\mathcal{C}}$ trained to label segments corresponding to class $\mathcal{C}$ objects. $\mathfrak{C}_{\mathcal{C}}$ gives the probability $P_{\mathcal{C}}(R)$ that a segment $R \in \mathcal{S}$ belongs to class $\mathcal{C}$; a reject zone [31] on the classifier is defined by two thresholds $T_{\in}$ and $T_{\notin}$ over the probability $P_{\mathcal{C}}(R)$. We define as ambiguous, any segment R such that $T_{\notin} \leq P_{\mathcal{C}}(R) \leq T_{\in}$.

A CoSC instance is composed of the following elements: A segmentation agent noted $S_{\mathcal{C}}$ (red boxes in Fig. 1) which evaluates the quality of a given segment s_a and modifies it accordingly if necessary. A classification agent noted $C_{\mathcal{C}}$ (dark blue boxes in Fig. 1) which selects the most ambiguous segment s_a in S (i.e., the segment of which the classification probability is the farthest away from $T_{\in}$ and $T_{\notin}$) as candidate to be evaluated and modified by $S_{\mathcal{C}}$. $C_{\mathcal{C}}$ also evaluates intermediate solutions provided by $S_{\mathcal{C}}$ in order to verify if the process reached convergence or if there is still improvements going on. At last, the process outputs a modified segmentation as well as an image representing the probability of each pixel to belong to class $\mathcal{C}$. The CoSC process is formally given by Algorithm 1.

Algorithm 1: CoSC.

```
Data: Ready-for-analysis image I_A
Data: Initial segmentation S_0
Data: Classifier C
Data: Real number in [0, 1] T_∈
Data: Real number in [0, 1] T_∉
Result: Modified segmentation S'
Result: Probability image PI
begin
    while not convergence% tested by C_C
    do
        s_a = mostAmbiguous(S)% done by C_C
        q_s = evaluate(s_a)% done by S_C
        s_m = modify(s_a, q_s)% done by S_C
        S' = replace(s_a, s_m)% done by S_C
    foreach Segment s ∈ S' do
        foreach Point p ∈ s do
            PI[p] = 𝔆_C(s)
    return S', PI
```