A collaborative framework for joint segmentation and classification of remote sensing images

Andrés Troya-Galvis, Pierre Gançarski and Laure Berti-Équille

Abstract In this article, we present a collaborative framework for joint segmentation and classification (CoSC). The framework is guided by and aware of the quality of each segment at every stage; it allows the consideration of both homogeneity based criteria as well as implicit semantic criteria to extract the objects belonging to a given thematic class. We apply CoSC to vegetation extraction in a very high spatial resolution image of Strasbourg. We compare our results to a pixel-based method, an object-based method and a hybrid segmentation-classification method. The experiments show that CoSC manages to reach good classification results while remarkably improving the segmentation results.

1 Introduction

Automatic interpretation of remote sensing images is a difficult but crucial task in a wide range of applications. Since the apparition of Very High Spatial Resolution (VHSR) imagery, the Object Based Image Analysis (OBIA) paradigm has been preferred over pixel oriented approaches [Blaschke, 2010]. Indeed, at VHSR a single pixel is not informative enough for the classification task since the types and complexity of identifiable objects increase considerably. Thus, image segmentation is performed first in order to obtain higher level objects called segments which allow a better description of the image.

Andrés Troya-Galvis
iCube UMR CNRS 7357. University of Strasbourg, e-mail: troyagalvis@unistra.fr

Pierre Gançarski
iCube UMR CNRS 7357. University of Strasbourg e-mail: gancarski@unistra.fr

Laure Berti-Équille
Qatar Computing Research Institute, Hamad Bin Khalifa University, e-mail: lberti@qf.org.qa

Figure 1 illustrates the complete workflow of OBIA. First (V)HSR images are acquired via satellite sensors. Atmospheric and geometric corrections are applied as pre-processing to produce ready-for-analysis images. Segmentation is then applied and the image is now described by segments, which are sets of similar pixels satisfying a given homogeneity criterion. A vectorization stage consisting in representing each segment as a vector of descriptive features such as shape, textural, or sophisticated radiometric indexes is then applied. Next, classification techniques which may use external information such as examples or ontologies, are applied to obtain a full land-cover map of the image. As a final step, the resulting image is validated either by an expert or by objective metrics in order to assess the quality of the results.

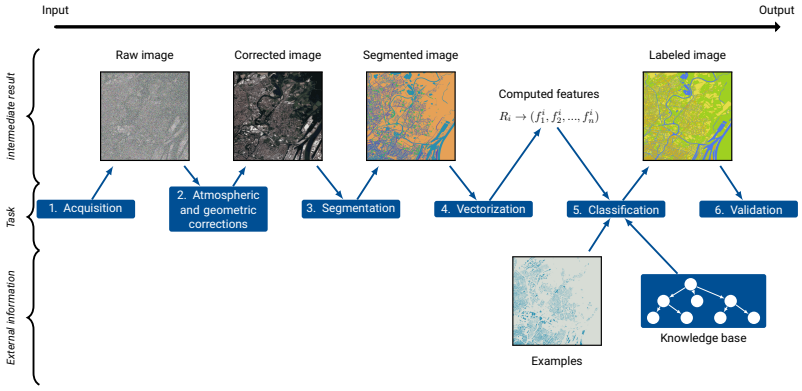


Fig. 1: Typical workflow of OBIA methods

The segmentation and classification are – in our opinion – the most critical steps. Many classification techniques such as [di Sciascio et al., 2013], have proven to be successful when the segmentation approaches a one-to-one mapping between the segments and geographic objects. Unfortunately, there is no single segmentation algorithm leading to such perfect mapping in all cases. Our intuition suggests that (1) the improvement of image segmentation should lead to the improvement of the classification; and (2) classification results may be useful to improve the quality of the segmentation. Moreover, while full interpretation is often desirable, the extraction of a single thematic class is generally sufficient for multiple tasks such as forest mapping [Räsänen et al., 2013], deforestation tracking [Duveiller et al., 2008], landslide risk management [Promper et al., 2014], urban planning [Pham et al., 2011], etc. Thus, we propose a collaborative framework for joint segmentation and classification (CoSC) which is quality-centric and attempts to simultaneously improve both segmentation and classification for a given thematic class by the interaction of these two paradigms which are closely related in the OBIA context.

The rest of this article is structured as follows: Section 2 introduces the context and related work. Section 3 presents our collaborative framework for joint segmentation and classification. Section 4 presents an experimental study demonstrating the

applicability of the proposed framework. Finally, in Section 5 we conclude and give some research perspectives.

2 Related work

Many work can be found in the literature about combining image segmentation and classification methods in a remote sensing context. For instance, [Lizarazo and Elsner, 2011] proposed a fuzzy segmentation approach in which the image pixels are classified by fuzzy classifiers trained for every desired class, thus obtaining what they call fuzzy segments; the fuzzy segments are defuzzyfied and merged together following a set of logical rules in order to obtain a fully labelled image. [Derivaux et al., 2010] proposed a supervised segmentation algorithm which adapts the parameters of the watershed segmentation algorithm by exploiting examples given by an expert and applying evolutionary algorithms. [Kurtz et al., 2012] developed a hierarchical multi-resolution approach for the extraction of complex urban objects. They take advantage of multi-source images which allows to have different views on the same data; using low resolution images to extract coarse objects such as a whole city and higher resolution images to extract finer objects like a district or individual buildings. [Mahmoudi et al., 2013] proposed a multi-agent system for remote sensing image classification where each agent is specialized in the extraction of a given class. A particular agent is charged to manage conflict resolution. [Tarabalka et al., 2009] presented Spectral-Spatial Classification (SSC), a hybrid approach mixing pixel- and object-based approaches; it uses EM clustering as segmentation algorithm, then a pixel-wise SVM classification is performed and a voting scheme is employed to decide the label for each segment. A final filtering step is applied to remove small noisy segments. Recently, [Hofmann et al., 2014] formalised an Agent Based Image Analysis (ABIA) framework, in which the idea is to let every segment be an independent agent which modifies itself trying to improve and modify classification rules given by a domain ontology.

Our work is mainly inspired by the wrapper-based segmentation framework [Farmer, 2009], which aims at segmenting complex objects by the integration of a semantic context to the segmentation process. This integration is done by wrapping a classifier inside a segmentation algorithm and use it as a quality metric to optimize. First, an initial segmentation is computed, this segmentation has to be over-segmented in order to the wrapper based framework to be effective. Indeed, the further steps consist in the filtering of irrelevant segments, the remaining segments are likely to be part of the object of interest. Then an optimization algorithm which alternates between adding or removing a segment to the current segmentation, and the classification of the latter. When the classification probability reaches an acceptable threshold, the best segmentation found and the classified objects are returned. The use of the classification to guide the segmentation process eliminates the homogeneity constraint on the segmentation allowing the extraction of complex objects.

Nevertheless, the wrapper framework supposes that the number of objects to extract is known to be relatively low and that their position can be easily determined to filter out unnecessary regions. These assumptions are generally not satisfied within a remote sensing context. Indeed, in a real-life scenario, objects of interest may vary from a few tens to a many hundred, or may not be present at all in the image. Moreover, objects of a same class may vary in shape and sizes and can be all over the image so it is not trivial to determine their possible locations.

We propose a generalization of the wrapper framework which allows the modification of segments in a more flexible manner making possible to extract objects of a given class regardless of their shape or position.

3 Proposition

The segmentation and classification paradigms pursuit different goals, yet they are closely related in a remote sensing context. Indeed, segmentation tries to partition the image spatially, mainly based on color properties (radiometric responses); while classification aims at partitioning data based on some knowledge which can be explicitly modelled or implicitly injected as training examples. Our intuition is that proper interactions between both approaches should lead to the improvement of both segmentation and classification results simultaneously, as the ideal segmentation is the one allowing to best classify the image and the perfect classification result in a partitioning with a one-to-one mapping between image regions and geographic objects.

Studies have shown the benefits of treating a single thematic class at a time [Musci et al., 2013]. In fact, it is easier to set the parameters of a segmentation algorithm so that the resulting segments fit one class without taking care of the rest. As a direct consequence, the classification of objects from this class is improved. Thus, our collaborative framework for joint segmentation and classification (CoSC) is devoted to the extraction of the objects belonging to one thematic class from the image. Let C be such a class of interest, we assume the existence of the following elements:

- An initial segmentation $S = \{R_i \mid 0 < i < W\}$ where W is the number of segments; and where each segment $R_i = \{(x_k^i, y_k^i) | 0 < k < M\}$ with M is the number of pixels in R_i.
- A 1-vs-all classifier $\mathbb{C}_C$ allowing to discriminate objects of class C from the rest. Note that this classifier has to be properly trained offline avoiding extreme over- and under-fitting since no further learning is performed during the CoSC process.

3.1 Definitions

Let us pose some definitions which are necessary in order to formally describe the proposed framework.

Definition 1. A specialized segmenter, noted S_C, is a segmentation agent capable of locally evaluating and modifying a given segment R.

Definition 2. A class extractor, noted C_C, is a classification agent capable of determining by using $\mathbb{C}_C$, the probability $P_C(R)$ that a given segment $R \in S$ belongs to the class C.

Definition 3. A CoSC agent, noted SC_C, is an agent capable of managing the collaboration between a specialized segmenter and a class extractor.

Definition 4. Let $T_\in$ (resp. $T_\notin$) be a threshold over the probability of being (resp. not being) a C class object. Thus, $T_\in$ and $T_\notin$ define a reject zone [Chow, 1970] for the classifier $\mathbb{C}_C$.

Definition 5. An ambiguous segment is a segment R such that $T_\notin \leq P_C(R) \leq T_\in$.

3.2 Global CoSC process

The collaborative process is described in Algorithm 1. After the selection of an initial segmentation, the process involve interactions between classification and segmentation approaches. In fact, the following steps are iterated until convergence:

1. A segment selection step relying on a classification-based criterion.
2. A segment modification step relying on local segmentation evaluation and segmentation operators.
3. An evaluation step based on classification probability.

Thus, the first step aims at selecting a poorly classified segment; the second step attempts to correct segmentation errors around the selected segment which in turn should improve its classification; finally, the evaluation step allows the process to quantify the improvement (or deterioration) of the current solution and checking for convergence. In the following subsection we describe in detail one possible implementation of these steps.

3.2.1 Initialization

S_C begins with an initial segmentation S which can be arbitrarily chosen. Nevertheless, over-segmentation is preferable to under-segmentation, since the latter is harder to correct. Our experiments show that the closer the size of the segments are close to the size of C class objects, the quicker collaboration converges.

Algorithm 1: CoSC process

Input : Specialized Segmenter: S_C,
Class extractor: C_C
Output: Segmentation: S_{best}

begin
  $S \leftarrow S_C$.initialSegmentation()
  $S_{best} \leftarrow S$
  $S_{current} \leftarrow S$
  while *not* $C_C.convergence()$ **do**
    $R_a \leftarrow C_C.candidate(S_{current})$
    $R_m \leftarrow S_C.SegmentModifications(R_a)$
    Update $S_{current}$ with R_m
    $C_C.evaluate(S_{current})$
    if $S_{current}$ *is better than* S_{best} **then**
      Update S_{best} with $S_{current}$
  return S_{best}

3.2.2 Candidate segment selection

C_C selects a candidate segment R_a to be modified. It is essential to chose a good candidate segment, in other words, a segment which may lead to the improvement of the classification after its modification. Different strategies can be used, such as selecting a random segment, or selecting the segment which is closest to $T_{\in}$ or $T_{\notin}$. In our experiments (Section 4) we select the most ambiguous segment, as follows:

$$R_a = \arg\min_{R_i} |P_C(R_i) - T_{avg}| \tag{1}$$

where $T_{avg} = \frac{T_{\in} + T_{\notin}}{2}$.

3.2.3 Segment modification

S_C evaluates the quality of the candidate segment and then tries to improve it's segmentation consequently. Following a given quality criterion, S_C associates a segment with one of the following states: over-segmentation, under-segmentation or undetected segmentation error.

A segment modification operator is a function $O : \mathfrak{D}_i \rightarrow \mathfrak{D}_i$ where $\mathfrak{D}_i = R_i \cup N_{R_i}$ and N_{R_i} denotes the set of points belonging to the segments which are adjacent to R_i (i.e., neighbour segments). Thus, a list containing at least one function O is associated to each segmentation state. For over-segmentation, the list O_L should contain operators which may make the segment bigger. For under-segmentation, the list U_L should contain operators which may make the segment smaller. For undetected miss-segmentation, the list M_L should contain a variety of mixed operators which may help to improve the classification of the segment. The modifications are applied

as depicted in Algorithms 2 and 3. A random operator is applied from one of the three lists to the candidate segment in function of its quality state.

In our implementation (Section 4) three different O functions are used.

The first one, is intended to correct heavy over-segmentation, it is a merging operator, noted FN, which merges segment R_i with its closest neighbour R_j:

$$Z = \underset{R_j \in N_{R_i}}{\arg\min}\ \delta_{L*a*b}(R_i, R_j) \tag{2a}$$

$$FN(R_i) = (R_i \cup Z) \cup (N_{R_i} \setminus Z) \tag{2b}$$

where $\delta_{L*a*b}(R_i, R_j)$ is the Euclidean distance between segments R_i and R_j in the $L*a*b$ colorspace [Chen and Wang, 2004].

The other two, are intended to correct slight over- and under-segmentation, they are based on morphological operators which may be easily applied to grow or shrink a segment.

The growing operator, noted Gr, is defined by:

$$Gr(R_i) = R_i \oplus \boxdot_1 \cup N^-_{R_i} \tag{3}$$

where $R_i \oplus \boxdot_1$ denotes the morphological dilation [Haralick et al., 1987] of R_i by a 3×3 square; and $N^-_{R_i} = N_{R_i} \setminus ((R_i \oplus \boxdot_1) \setminus R_i)$.

The shrinking operator, noted Sh, is defined by:

$$Sh(R_i) = R_i \ominus \boxdot_1 \cup N^+_{R_i} \tag{4}$$

where $R_i \ominus \boxdot_1$ denotes the morphological erosion [Haralick et al., 1987] of R_i by a 3×3 square; and $N^+_{R_i} = N_{R_i} \cup (R_i \setminus (R_i \oplus \boxdot_1))$. Exceeding pixels resulting from the erosion are allocated to the closest neighbour $R_j \in N_{R_i}$.

Note that if the application of an operator O implies modifying the topology of resulting segments, then the segment is returned unmodified.

We set $O_L = \{FN(R_i)\}$; $U_L = \{Sh(R_i)\}$; and $M_L = \{Gr(R_i), Sh(R_i), FN(R_i)\}$. In this case the candidate segment modification strategy is equivalent to the following rules:

- If R_a is over-segmented, it is merged with its most similar neighbour.
- If R_a is under-segmented, it is shrunk by an morphological erosion.
- Otherwise, three operations (merging, growing and shrinking) are tested randomly and the first successful (i.e., actually modifying R_a) transformation is kept.

3.2.4 Evaluation

C_C evaluates the quality of the current solution $S_{current}$ by using an objective function based on classification criteria. Thus, under the assumption that the classifier is well trained and capable of discriminating C class objects from other kind of objects, we

Algorithm 2: ApplyModifications

```
Input  : Segment: R_a,
         List of O: L
Output: Segment: R_m
begin
    Shuffle(L)
    while L.iterator.has_next() do
        R_m ← modify(R_a, L.iterator.next())
        if R_m ≠ R_a then
            return R_m
    return R_a
```

Algorithm 3: SegmentModification

```
Input  : Segment: R_a,
Output: modified Segment: R_m,
begin
    switch LocalSegmentEvaluation(R_a) do
        case oversegmentation
            R_m ← ApplyModifications(R_a, O_L)
        case undersegmentation
            R_m ← ApplyModifications(R_a, U_L)
        otherwise
            R_m ← ApplyModifications(R_a, M_L)
    return R_m
```

can evaluate the solution by looking at how well the classifier separates objects of interest from the rest. In other words, the goal is to reduce the number of ambiguous segments (c.f., def 5). The evaluation function Q_{cs} can be defined as follows:

$$Q_{cs} = \frac{1}{W}\left(\sum_{i|P_C(R_i)>T_{\in}} P_C(R_i) + \sum_{i|P_C(R_i)<T_{\notin}} (1 - P_C(R_i))\right) \tag{5}$$

where W is the total number of segments in the current segmentation. The idea behind this formula is to reward those segments for which the classifier has made a decision proportionally to the confidence of the classifier. The more the classifier is sure about the class of a given segment R_i, the more R_i is contributing to the score. Plus, ambiguous segments are penalized as they do not contribute to the score, but they are taken into account by the $\frac{1}{W}$ factor. Thus, in the ideal case where the classifier has perfectly classified all of the segments with 100% confidence then $Q_{cs} = 1.0$; in the worst case in which all of the segment are ambiguous then $Q_{cs} = 0.0$. At each iteration S_{best} is updated if $S_{current}$ has a higher Q_{cs} score.

3.2.5 Convergence

A simple strategy is to let the collaboration continue as long as it improves the Q_{cs} score, but it would probably lead to a premature convergence problem, thus to a suboptimal result. To address this issue a variant of simulated annealing is used to drive the optimization process.

3.3 Output

The output of each iteration of a modified segmentation S_{best} as well as a probability image associated to S_{best} in which each pixel $(x_k^i, y_k^i) \in R_i$ correspond to the probability $P(R_i)$.

It is then possible to evaluate these results in terms of classification, or in terms of segmentation as we show in Section 4.2.

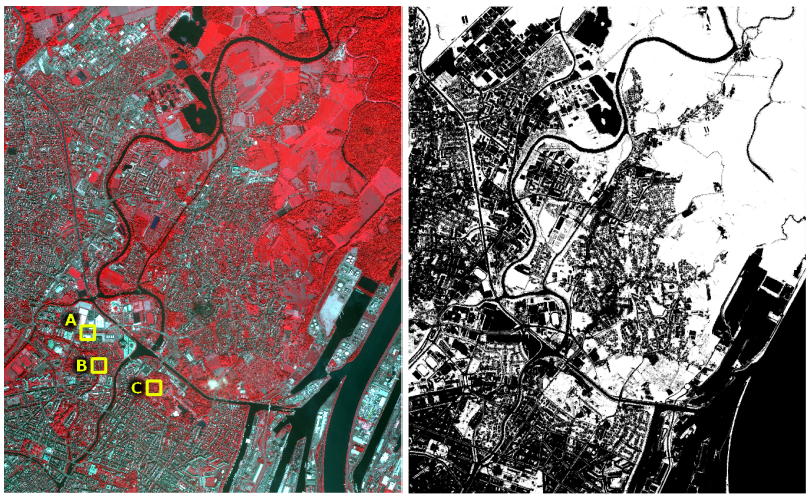
4 Case study

The performance of the classifier $\mathbb{C}_C$ has a considerable impact on the results of CoSC. To limit the bias induced by $\mathbb{C}_C$, we chose to apply CoSC for the extraction of vegetation zones. Indeed, in this case we could easily define and train a well-performing classifier, allowing us to better validate our proposition.

The studied image is a 9211×11275 (10^8 pixels) VHSR image of Strasbourg. It is a Pleiades pansharpened image at 50cm resolution, with 4 spectral bands: red (R), green (G), blue (B), and near infra red (NIR). Figure 2a shows the studied image in false colors (i.e., the red was replaced by the NIR, as vegetation is known to have high radiometric responses in this region of the spectrum). For ease of computation, the image was cropped into 1620 tiles of 256×256 pixels. Each tile was processed with the same parameters. In this paper we analyse in detail the results of three tiles: the first one corresponds to a zone of industrial facilities; the second one corresponds to an urban area; and the third one corresponds to a public park. They are labelled A, B and C respectively on Figure 2a and shown in detail on Figures 3c–3b. The resulting probability images are compared to available reference-data shown in Figure 2b.

4.1 Instantiation

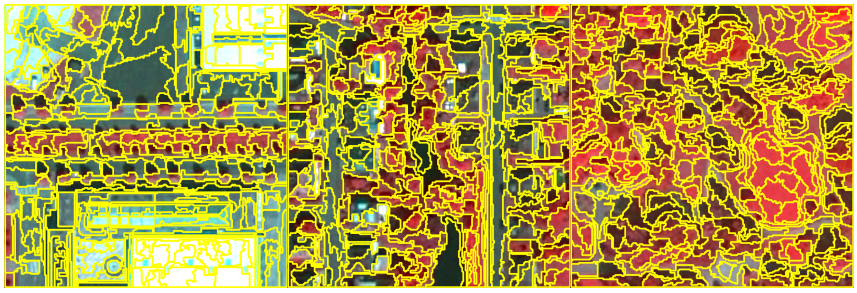
The framework was instantiated as follows.



(a) (b)

Fig. 2: Left: studied image. Right: available reference data.

(a) Tile A: industrial facilities (b) Tile B: urban area (c) Tile C: public park



(d) Init. segmentation of tile A (e) Init. segmentation of tile B (f) Init. segmentation of tile C

Fig. 3: Top: tiles A, B and C at a detailed scale. Bottom: the initial mean-shift segmentations computed for those tiles.

4.1.1 Specialized segmenter (Def. 1)

The UOA_{L_2} metric [Troya-Galvis et al., 2015] was used as segmentation quality metric; the homogeneity index used was the entropy [Zhang et al., 2003] and the threshold value was learnt from some explicitly given examples of vegetation. The local evaluation function of UOA is used as well to determine whether a segment is over-, under- or well-segmented.

The mean-shift algorithm was applied to generate the initial segmentations (Figures 3d-3f). Its parameters were selected by evaluating 100 different parameter combinations with the UOA_{L_2} metric. Remark that it is possible to apply any other method to select the initial segmentation. Different initial segmentations would surely lead to different results; however a complete sensitivity analysis on this aspect is out of the scope of this article.

4.1.2 Class extractor (Def. 2)

A linear regression model was learnt using the Weka API [Hall et al., 2009]. The training set is a subset of the fuzzy reference data containing both vegetation and non-vegetation segments chosen randomly. The feature set employed was composed of 32 features, including radiometrical attributes such as mean and standard deviation values for each band [Rougier and Puissant, 2014], as well as geometrical attributes such as the area, shape, orientation and solidity indexes among others. When the predicted values are out of the accepted bounds, they are simply truncated to 1.0 or 0.0 as needed.

Remark that the definition of the classification thresholds can be misleading. Indeed, the classifier is a parameter of CoSC and it has to be trained offline. Thus, the classification thresholds (def. 4) within CoSC are merely a tool to "push" the algorithm to look for very confident classifications. Then, these thresholds should have little or no impact on the final classification results. In order to verify this fact, we conducted the experiment for $T_{\in} = 0.9$, $T_{\notin} = 0.1$; $T_{\in} = 0.8, T_{\notin} = 0.2$; and $T_{\in} = 0.7, T_{\notin} = 0.3$, using the same initial segmentations. The results are summarized in table 1, we can see that the choice of the classification thresholds has almost no impact on the classification results, however they have some influence on the computation times. In Section 4.2 we study in detail the results for $T_{\in} = 0.9$ and $T_{\notin} = 0.1$.

Table 1: Quality measures for different $T_{\in}$ and $T_{\notin}$ values

$T_{\in}$	$T_{\notin}$	Acc	Pr	Re	F1	κ	Am%	Q_{cs}
0.7	0.3	0.932	0.952	0.945	0.949	0.846	0.469	0.526
0.8	0.2	0.930	0.952	0.943	0.948	0.844	0.466	0.529
0.9	0.1	0.930	0.952	0.943	0.947	0.844	0.464	0.531

The CoSC process takes in average 17 seconds per tile, and the computation time for the whole image ranged from 5 to 10 hours depending on the parameters.

A badly chosen initial segmentation S leads to increased computation times and aberrant cases may result in bad segmentation and classification results. Finally, our first experiments suggest that the classification model $\mathbb{C}_C$ have no significant influence on computation time, but is crucial to segmentation and classification results. However this behaviour has to be verified by a more formal study which is out of the scope of this article. The impact of these parameters is briefly summarized in Table 2.

Table 2: Impact of CoSC parameters on CPU time, segmentation and classification quality

Parameter	CPU time	segmentation	classification
$T_{\in}$	medium	very low	very low
$T_{\notin}$	medium	very low	very low
S	high	medium	medium
$\mathbb{C}_C$	low	medium	high

4.2 Results

In order to validate the applicability of CoSC, we compare our results to a pixel-based method, a classic OBIA approach and a hybrid segmentation-classification approach.

The pixel-based method consists in computing the normalized difference vegetation index (NDVI) [Rouse et al., 1974] which is a well-known radiometric index for the extraction of vegetation zones. The NDVI image is then normalized in $[0,1]$ so that we can see each pixel value as the probability of belonging to the vegetation class. In the rest of the article we denote this NDVI-based approach simply as NDVI. The classic OBIA method consists in classifying the initial segmentations with the same linear regression model that we used for CoSC. The hybrid segmentation-classification approach is a variant of the Spectral-Spatial Classification (SSC) approach described in Section 2. We evaluate the results in terms of classification and segmentation separately.

Classification

We computed both crisp and fuzzy quality metrics against the reference data (Figure 2b). A threshold of 0.5 was used to *defuzzify* these data to compute the following crisp measures : accuracy (Acc), Precision (Pr), Recall (Re), F-measure (F1), and Cohen's kappa (κ). Table 3 shows the crisp results. We observe that the NDVI

achieves great accuracy but the percentage of rejected or ambiguous pixels is 89%. The OBIA classic approach as well as the proposed CoSC reach acceptable Acc, Pr, Re, F1 and κ values. We remark that they reduced the percentage of ambiguous pixels to 52% and 46% respectively which is a good trade-off. SSC achieves high accuracy results without reporting ambiguous segments as the SVM model has binary output.

Table 3: Crisp quality measures

Image	Acc	Pr	Re	F1	κ	Am%	Q_{cs}
NDVI	0.997	0.999	0.988	0.994	0.992	0.894	0.103
OBIA	0.931	0.953	0.948	0.950	0.834	0.516	0.479
SSC	0.894	0.872	0.912	0.892	0.787	0.000	1.000
CoSC	0.930	0.952	0.943	0.947	0.844	0.464	0.531

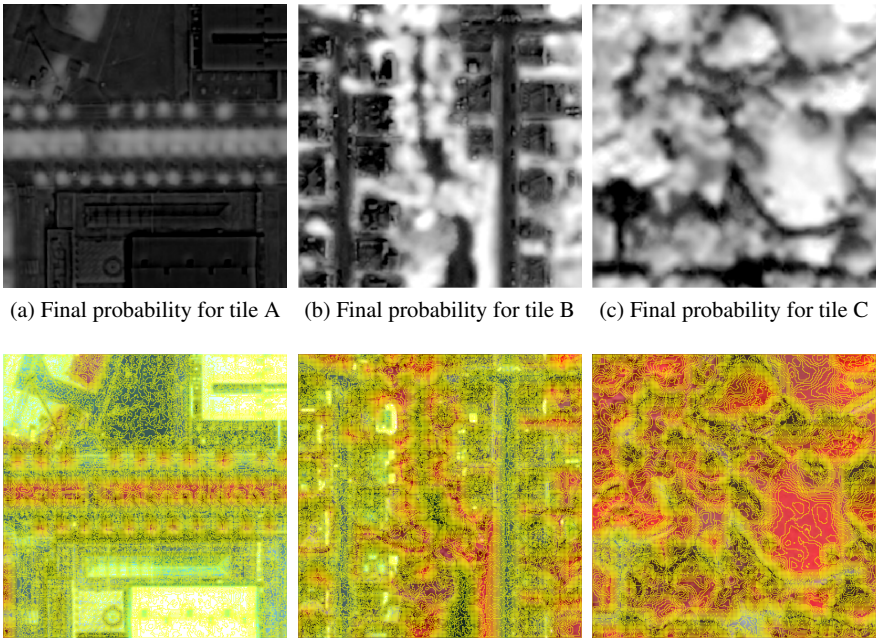
We also computed the fuzzy error matrix [Binaghi et al., 1999] and derived from it the fuzzy $\widetilde{Acc}$, $\widetilde{Pr}$, $\widetilde{Re}$, and $\widetilde{F1}$. Fuzzy metrics allows to take into account the intrinsic uncertainty of the reference data and have a more accurate quality assessment. Table 4 shows the fuzzy results for the compared methods. We observe that the NDVI method is globally outperformed by the other methods, and SSC achieves better classification results. There is no significant difference between the OBIA and CoSC methods, and they still have an acceptable performance.

Table 4: Fuzzy quality measures

Image	$\widetilde{Acc}$	$\widetilde{Pr}$	$\widetilde{Re}$	$\widetilde{F1}$
NDVI	0.774	0.791	0.720	0.754
OBIA	0.809	0.745	0.915	0.830
SSC	0.869	0.849	0.886	0.867
CoSC	0.825	0.768	0.912	0.834

Segmentation

Let's take a closer look at segmentation results. Figure 4, 5, 6 and 7 illustrate in detail the results of the NDVI method, OBIA, SSC, and CoSC respectively for the three tiles highlighted in Figure 2a. The first row shows the probability images resulting from each method. The second row shows the boundaries of the segmentation corresponding to the flat regions (i.e., adjacent pixels having the same values) in the probability image.



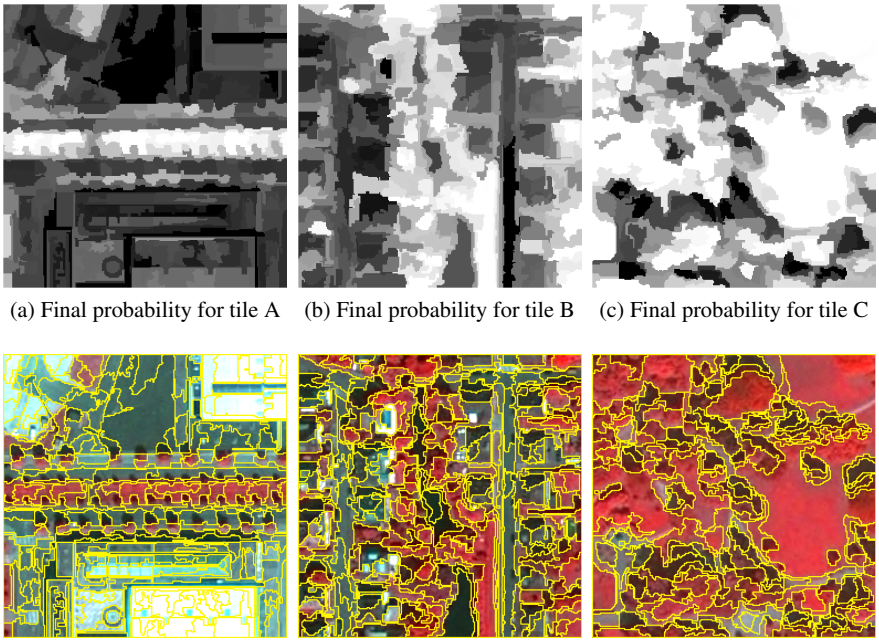
(a) Final probability for tile A (b) Final probability for tile B (c) Final probability for tile C

(d) Final segmentation for tile A (e) Final segmentation for tile B (f) Final segmentation for tile C

Fig. 4: Detailed view of the results obtained with the NDVI.

We observe that the NDVI method is able to highlight vegetation zones quite accurately. However, the probability image is noisy so the flat-region based segmentation is extremely over-segmented; this shows one of the disadvantages of purely pixel-based approaches. Indeed, almost each pixel has a different NDVI value which results in the extreme over-segmentation observed in Figure 4. More sophisticated post-processing techniques involving thresholds or clustering for example are required in order to obtain a good quality segmentation based on the NDVI approach.

The classic OBIA method achieves also a good discrimination of vegetation areas. However the associated segmentation is still over-segmented at many places. Indeed, in tile A, we observe that the different buildings are as over-segmented as in the initial segmentation (Fig 3d); in tile B we observe many elongated and irregular segments at the boundaries of some vegetation objects, also, the houses are still divided in many different segments; finally in tile C, we observe a lot of very small segments which could be merged to form medium sized trees.



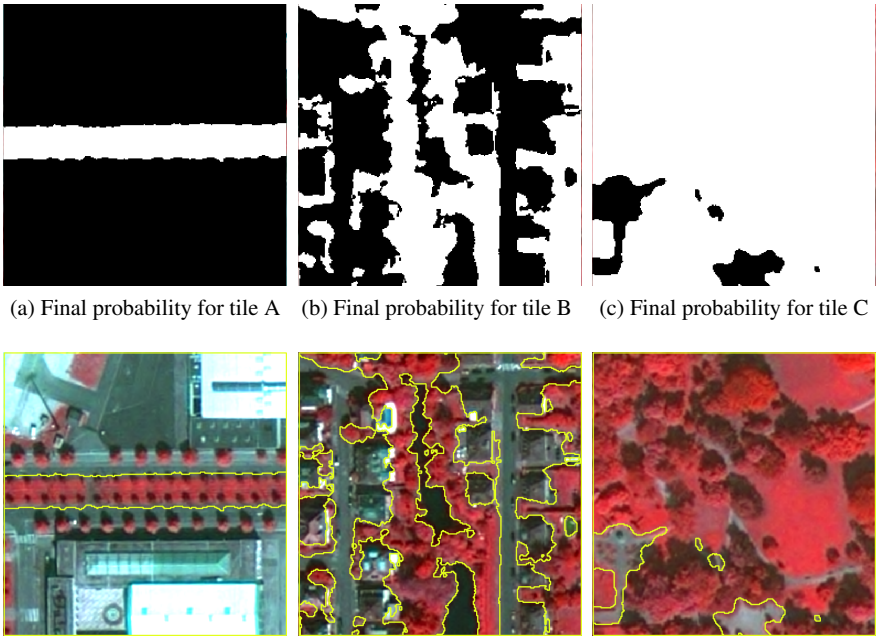
(a) Final probability for tile A (b) Final probability for tile B (c) Final probability for tile C

(d) Final segmentation for tile A (e) Final segmentation for tile B (f) Final segmentation for tile C

Fig. 5: Detailed view of the results obtained with Classic OBIA.

The SSC method results in very under-segmented results. Indeed, the binary nature of this method as well as the post-processing filtering result in a very rough distinction between vegetation and non-vegetation objects losing all of the spatial information concerning higher level objects such as individual trees or buildings for instance.

By contrast, CoSC is able to give high probabilities to vegetation objects while giving very low probabilities to non-vegetation objects. Moreover, we observe that the associated segmentation is remarkably good. Indeed, there are almost no evident over- or under-segmented regions. Another surprising result is that the segmentation of non-vegetation objects was also improved as a side-effect. This can be explained by the fact that over-segmentation is mainly related to radiometrical homogeneity and the segmenter takes this into account to modify candidate segments, thus improving segmentation quality globally.



(a) Final probability for tile A (b) Final probability for tile B (c) Final probability for tile C

(d) Final segmentation for tile A (e) Final segmentation for tile B (f) Final segmentation for tile C

Fig. 6: Detailed view of the results obtained with Spectral-Spatial Classification.

In order to objectively evaluate the segmentation results we employed the UOA_{L_2} metric. Recall that this metric is defined by:

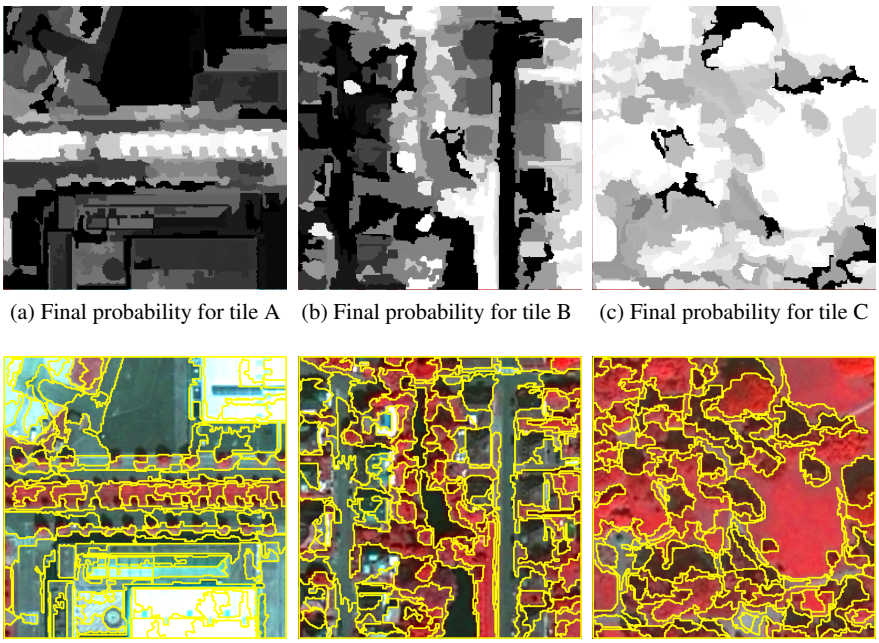
$$\Psi = \sum_{R_i|\phi_\delta(R_i)=-1} \omega(R_i)$$

$$\Theta = \sum_{R_i|\phi_\delta(R_i)=1} \omega(R_i)$$

$$UOA_{L_2} = \sqrt{\Psi^2 + \Theta^2} \tag{6}$$

where Ψ and Θ represent the under- and over-segmentation rates respectively; and $\phi_\delta(R_i)$ is a local evaluation function which estimates if a segment is over-segmented, under-segmented or well segmented in in function of a given homogeneity index H and a threshold δ.

For each tile we evaluated the results of the 4 tested approaches with UOA_{L_2}. We used the segment entropy as the homogeneity index and δ varying from 0 to 1 by 0.01 steps.

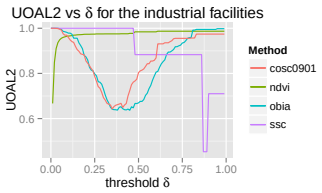


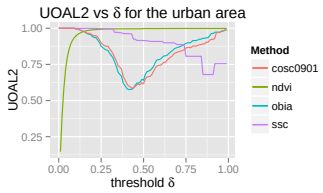
(a) Final probability for tile A (b) Final probability for tile B (c) Final probability for tile C

(d) Final segmentation for tile A (e) Final segmentation for tile B (f) Final segmentation for tile C

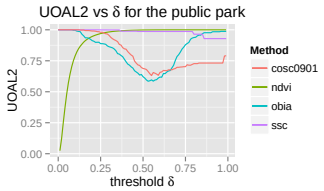
Fig. 7: Detailed view of the results obtained with CoSC.

Figure 8 reports the results of these computations. Remark that UOA_{L_2} ranges from 0 (best) to 1 (worse); thus the lower the area under the UOA_{L_2} curve, the better the segmentation. For tile A and B, we observe that as expected the NDVI has the greatest area under the UOA_{L_2}, followed by SSC; CoSC and OBIA have very similar curves the area under the curve of OBIA is slightly lower than that of CoSC for tile A while for tile B we observe the inverse result. For tile C, the SSC method has the worst result, followed by the NDVI, the OBIA method and finally by CoSC which has a clearly lower area under the curve for this tile. These results confirm our previous observations, CoSC has indeed globally improved the segmentation, which is a very interesting result since the produced segmentation could be re-used for further processing and analysis in multi-class or multi-scale applications for example.





(a) UOA_{L2} for the segmentation of tile A(b) UOA_{L2} for the segmentation of tile B



(c) UOA_{L2} for the segmentation of tile C

Fig. 8: Plots 8c, 8a, and 8b show the variation of the UOA_{L_2} metric over threshold values in $[0,1]$ for the three tiles.

5 Conclusion

In this article we presented CoSC, a collaborative framework for joint segmentation and classification. It aims at the extraction of a given thematic class of objects from a remote sensing image, and allows the exchange of implicit information between these different but closely related paradigms in order to simultaneously improve both segmentation and classification results. We presented the extraction of vegetation areas from a large image of Strasbourg as case-study to demonstrate the applicability and pertinence of the CoSC framework. Experiments show that the CoSC is able to give accurate classification results while remarkably improving the segmentation of the whole image, not only for the objects of interest but for all types of objects. In short-term we will study different strategies for choosing the candidate segment as well as different classification models and their influence on the results. The CoSC framework is designed for the extraction of a single thematic class; our future research will be focused on the study of collaborative strategies between many of such CoSC processes in order to achieve automatic and full (multi-class) interpretation of remote sensing images.

Acknowledgements

The research leading to these results has received funding from the French *Agence Nationale de la Recherche* (Grant Agreement ANR-12-MONU-0001).

References

[Binaghi et al., 1999] Binaghi, E., Brivio, P. A., Ghezzi, P., and Rampini, A. (1999). A fuzzy set-based accuracy assessment of soft classification. *Pattern Recognition Letters*, 20(9):935–948.

[Blaschke, 2010] Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65:2–16.

[Chen and Wang, 2004] Chen, H.-C. and Wang, S.-J. (2004). The use of visible color difference in the quantitative evaluation of color image segmentation. In *IEEE International Conference on Acoustics, Speech and Signal Processing – ICASSP*, pages 593–596.

[Chow, 1970] Chow, C. (1970). On optimum recognition error and reject tradeoff. *IEEE Transactions on Information Theory*, 16(1):41–46.

[Derivaux et al., 2010] Derivaux, S., Forestier, G., Wemmert, C., and Lefèvre, S. (2010). Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation. *Pattern Recogn Letters*, 31:2364–2374.

[di Sciascio et al., 2013] di Sciascio, C., Zanni-Merk, C., Wemmert, C., Marc-Zwecker, S., and de Beuvron, F. d. B. (2013). Towards a semi-automatic semantic approach for satellite image analysis. *Procedia Computer Science*, 22:1388–1397.

[Duveiller et al., 2008] Duveiller, G., Defourny, P., DesclÃľe, B., and Mayaux, P. (2008). Deforestation in central africa: Estimates at regional, national and landscape levels by advanced processing of systematically-distributed landsat extracts. *Remote Sensing of Environment*, 112(5):1969 – 1981.

[Farmer, 2009] Farmer, M. E. (2009). *Application of the Wrapper Framework for Robust Image Segmentation For Object Detection and Recognition*. INTECH Open Access Publisher.

[Grozavu et al., 2015] Grozavu, N., Rogovschi, N., Cabanes, G., Troya-Galvis, A., and Gancarski, P. (2015). VHR satellite image segmentation based on topological unsupervised learning. In *International Conference on Machine Vision Applications – MVA*, pages 543–546.

[Hall et al., 2009] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18.

[Haralick et al., 1987] Haralick, R., Sternberg, S. R., and Zhuang, X. (1987). Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(4):532–550.

[Hofmann et al., 2014] Hofmann, P., Lettmayer, P., Blaschke, T., Belgiu, M., Wegenkittl, S., Graf, R., Lampoltshammer, T. J., and Andrejchenko, V. (2014). ABIA - A conceptional framework for agent based image analysis. *South-Eastern European Journal of Earth Observation and Geomatics*, 3(25):125–129.

[Kurtz et al., 2012] Kurtz, C., Passat, N., Gançarski, P., and Puissant, A. (2012). Extraction of complex patterns from multiresolution remote sensing images: A hierarchical top-down methodology. *Pattern Recognition*, 45:685–706.

[Lizarazo and Elsner, 2011] Lizarazo, I. and Elsner, P. (2011). Segmentation of remotely sensed imagery: Moving from sharp objects to fuzzy regions. *INTECH Open Access Publisher*.

[Mahmoudi et al., 2013] Mahmoudi, F. T., Samadzadegan, F., and Reinartz, P. (2013). Object oriented image analysis based on multi-agent recognition system. *Computers & Geosciences*, 54:219–230.

[Musci et al., 2013] Musci, M., Feitosa, R., and Costa, G. (2013). An object-based image analysis approach based on independent segmentations. In *Urban Remote Sensing Event – JURSE*, pages 275–278.

[Nadeem et al., 2010] Nadeem, M. S. A., Zucker, J.-D., and Hanczar, B. (2010). Accuracy-rejection curves (ARCS) for comparing classification methods with a reject option. In *Machine Learning in Systems Biology*, pages 65–81.

[Pham et al., 2011] Pham, H. M., Yamaguchi, Y., and Bui, T. Q. (2011). A case study on the relation between city planning and urban growth using remote sensing and spatial metrics. *Landscape Urban Plan*, 100:223–230.

[Promper et al., 2014] Promper, C., Puissant, A., Malet, J.-P., and Glade, T. (2014). Analysis of land cover changes in the past and the future as contribution to landslide risk scenarios. *Applied Geography*, 53:11–19.

[Räsänen et al., 2013] Räsänen, A., Rusanen, A., Kuitunen, M., and Lensu, A. (2013). What makes segmentation good? A case study in boreal forest habitat mapping. *International Journal of Remote Sensing*, 34:8603–8627.

[Rougier and Puissant, 2014] Rougier, S. and Puissant, A. (2014). Improvements of urban vegetation segmentation and classification using multi-temporal pleiades images. In *International Conference on Geographic Object-Based Image Analysis*, page 6.

[Rouse et al., 1974] Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Harlan, J. C. (1974). *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation*. Texas A & M University, Remote Sensing Center.

[Tarabalka et al., 2009] Tarabalka, Y., Benediktsson, J., and Chanussot, J. (2009). Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(8):2973–2987.

[Troya-Galvis et al., 2015] Troya-Galvis, A., Gançarski, P., Passat, N., and Berti-Équille, L. (2015). Unsupervised quantification of under and over segmentation for object based remote sensing image analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(5):1936–1945.

[Zhang et al., 2003] Zhang, H., Fritts, J. E., and Goldman, S. A. (2003). An entropy-based objective evaluation method for image segmentation. In *Electronic Imaging 2004*, pages 38–49.