Semi Supervised Clustering Guided and Interpreted by Ontology Reasoning Application to Remote Sensing Images Classification

Hatim Chahdi*†, Nistor Grozavu†, Isabelle Mougenot*, Laure Berti*‡ and Youns Bennani†
*UMR Espace-Dev , IRD - Université de Montpellier
Maison de la télédétection, 500 Rue J.F. Breton, 34093 Montpellier, FRANCE
Email: first.last@ird.fr
†UMR 7030 LIPN, CNRS - Universit Paris 13
99, avenue Jean-Baptiste Clment, 93430 Villetaneuse, FRANCE
Email: first.last@lipn.univ-paris13.fr
‡Qatar Computing Research Institute, Hamad bin Khalifa University
Doha, Qatar

***Abstract*—Recent studies have shown that the use of *a priori* knowledge can significantly improve the results of unsupervised classification. However, capturing and formatting such knowledge as constraints is not only very expensive requiring the sustained involvement of an expert but it is also very difficult because some valuable information can be lost when it cannot be encoded as constraints. In this paper, we propose a new constraint-based clustering approach based on ontology reasoning for automatically generating constraints and bridging the semantic gap. The use of ontology as *a priori* knowledge has many advantages that we leverage in the context of satellite image interpretation. The experiments we conduct have shown that our proposed approach can deal with incomplete knowledge while completely exploiting the available one.**

I. Introduction

In the last decade, increasingly large volumes of remote sensing images have been made publicly available. The analysis and interpretation of these images are no longer manually feasible but they are mandatory to find actionable solutions to today's environmental and societal issues.

One of the most important step of satellite image processing is segmentation, which is essentially based on clustering algorithms. Classically used in exploratory and unsupervised settings, clustering aims at partitioning large volumes of non-labelled instances into groups of data based on similarity, density and proximity. However, in many cases, the quality of such partitioning is relative and highly depends on the user's points-of-interest and/or expertise. Two experts (e.g., hydrologist and urban geographer) who do not have the same thematic of interest will classify differently the same image segmentation output. Another important point is related to the interpretation and the semantics to associate to the clustering output in a user-defined context is a very challenging task. In this context, the introduction of knowledge in the process becomes essential, either for guiding the clustering task or for helping the interpretation of the clustering results. Furthermore, in the field of knowledge engineering, ontologies have shown their effectiveness especially in facilitating symbols and semantics anchoring, expressing more and more complex knowledge, as well as performing advanced deductive reasoning. However, the formalization of knowledge remains the bottleneck and some concepts remain difficult or even impossible to define precisely. As an illustrative example, in the context of remote sensing image analysis, experts can more easily define concepts related to vegetation or water rather than concepts related to buildings.

In this paper, we propose a new approach for data semantic labeling, automatically combining expert knowledge and clustering with generated constraints. This approach allows to guide and to improve the semi-supervised clustering process based on ontology reasoning and it offers multiple advantages:

- Generalization: The approach can be adapted to different domains;
- Minimal user involvement: Experts are involved only for building the domain ontology modeling the knowledge of the domain and the constraints are automatically generated;
- Uncertainty management: it can cope with incomplete and uncertain knowledge bases;
- Adaptive Clustering: Clustering is automatically adapted to get as closer as possible to the vision of the expert.

The next sections are organized as following, we will first present some related work in section II. Secondly, we present some core concepts of ontology and description logics reasoning, followed with a description of our approach in section III. The fourth section will give the details of the application of our approach on remote sensing images and summarize the obtained results. A conclusion and future works conclude the paper.

II. Related Work

Several studies were conducted to exploit the available domain's knowledge. When labeled data is not available or insufficient, two approaches can be used: knowledge based classification approaches and semi-supervised clustering approaches. Although these approaches share the same objective to use the available knowledge to increase there efficiency, they proceed differently. They are also often used at different levels in the knowledge extraction process. The main difference between these two approaches remains in the type of the reasoning they adopt. Most knowledge based systems relies on a deductive reasoning, while the semi-supervised clustering approaches are essentially inductive.

A. Knowledge based systems

Knowledge based systems have been widely used to reduce the semantic gap and/or to provide high level of semantic interpretation. Forestier et al. [1] proposed a method that uses formalized domain knowledge to label the objects of satellite images using the concepts formalized in a knowledge base. First, Forestier et al. uses a segmentation algorithm to obtain the objects from the image. Then, a matching process compute the similarity between the characteristics of the objects and the concepts in the **KB**. The object is then labeled with the concept having the highest similarity score. Falomir et al. [2] and Andres et al. [3] used description logic reasoning to label the extracted objects from the images. They also perform a manually parametrized segmentation over the images to extract the objects.

The works cited above use expert knowledge in different ways, but none of these approaches is used during a clustering or segmentation process. The main use of these methods is the semantic interpretation of the extracted objects. Forestier et al. proposed a method which uses similarity measures on semantic descriptors of objects extracted from the satellite image, but have not exploited the logic reasoning. Falomir and Andres have used the reasoning, but on previously extracted objects from images and always for the interpretation and not to guide the clustering process. Overall, only few state of art methods has applied the logic reasoning to satellite images, which cannot be used to strengthen the clustering process.

B. Semi supervised clustering

In the literature, a lot of work has been proposed to introduce and leverage *a priori* knowledge in clustering [4], [5]. Typically, several ways have been explored for integrating expert knowledge and supervision into the clustering process. Constraint-based clustering at the instance level is known to be very efficient to guide the cluster formation. Initially introduced by [6], knowledge is expressed as two types of links: *must-link* and *cannot-link*. The constraint *must-link* $ml(x_i, x_j)$ specifies that two instances, noted x_i and x_j have to be in the same final cluster, whereas *cannot-link* $cl(x_i, x_j)$ indicates that the two instances should not belong to the same cluster. The both *must-link* and *cannot-link* constraints are transitive.

Several constrained clustering variants have been proposed in the literature, that can be summarized in three types:

- Change of the update step for assigning the instances to the final clusters [7], [8];
- Adjustment of the initialization step of the clustering [9];
- Adaptation of the objective function of the clustering [4].

COP-KMEANS [7] has been the first algorithm that integrates constraints. An instance is assigned to a cluster only if no constraint is violated. Other techniques modify the initialization of the clustering algorithm. In a variant of hierarchical clustering based on constraints proposed by Davidson et Ravi [9], transitive closures are computed from the constraints for producing connected instances that will be used later on by the clustering algorithm.

Algorithms proposed by [7]–[9] have shown that clustering results can be improved by the use of constraints guiding the clustering. However, Some clustering variants adopt a *strict enforcement* approach (or *hard constrained*) where the algorithm has to find the most feasible clustering output that respects **all** the constraints. Experiments made by Davidson et al. [10] have shown that these algorithms are very sensitive to noise and have issues with inconsistent constraints. Other work has proposed algorithms for *partial enforcement* (or *soft constrained*) of the constraints, trying to find the best clustering output that respects the maximum number of constraints. Most of these approaches rely on the modification of the objective function of the clustering adding some penalty weight in case of constraint violation, e.g., *CVQE* [9], *PCKmeans* [11].

Although constraint-based clustering has received lot of attention in the last years, little work has been focused on automating the constraint generation for clustering. Current methods rely on manually defined constraints by the expert/user. In a lot of applications, setting constraints can be very expensive and requires a deep knowledge about data. Another inconvenience about encoding knowledge using only constraints is the lost of classes semantics when available. In this context, backing the constrained clustering by formalized knowledge will permit **automatic generation** of constraints relative to the user/expert points-of-interests.

III. Back-end Ontology for Improving Constraint-based Clustering

A. Background and Preliminaries

We introduce in this section some important elements of the Web Ontology Language (**OWL 2**) and the Description Logic (**DL**). OWL [12] is a standard language introduced and maintained by the World Wide Web Consortium (**W3C**). The aim of OWL is to give users a simple way to represent rich and complex knowledge, while facilitating the sharing and the publication of this knowledge over the Web. OWL introduces standardised elements which have a precise meaning and formal semantics. The formal part of OWL is mainly based on DL. DL [13] are a family of knowledge representation

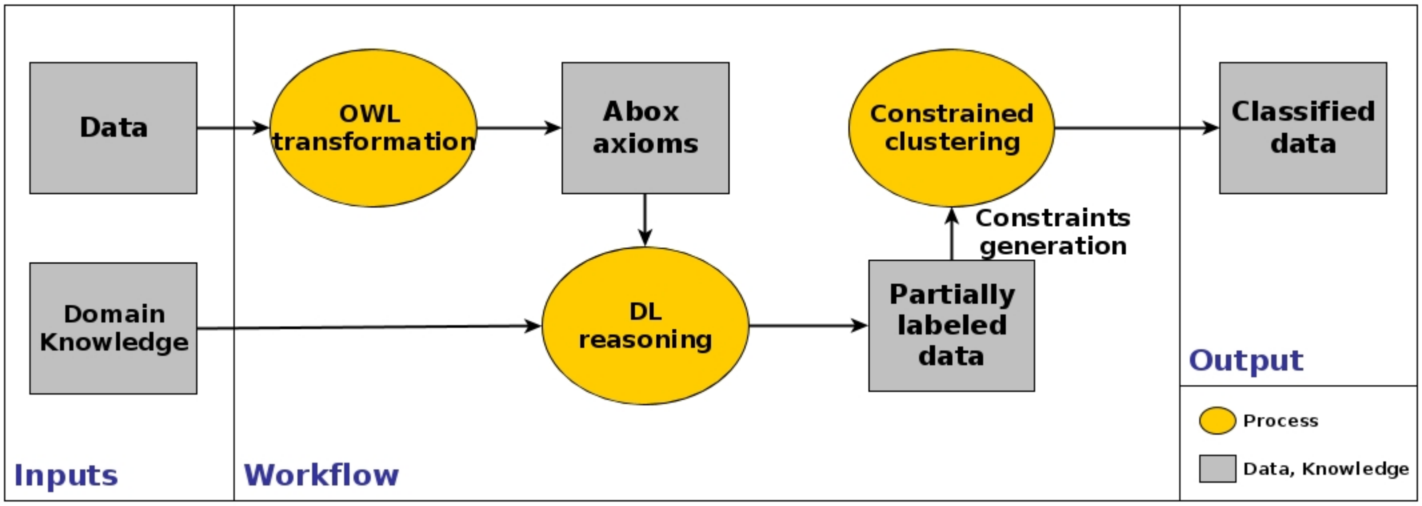


Fig. 1. Approach overview

languages that have been used before OWL to capture in a structured way a representation of domain knowledge.

In the following, we define an ontology[1] $\mathcal{O}$ as a set of axioms (facts) describing a particular situation in the world from a specific domain point of view. Formally, an ontology consists of a three sets: classes (concepts) denoted $\mathcal{N_C}$, properties (roles) denoted $\mathcal{N_P}$ and instances (individuals) denoted $\mathcal{N_I}$. Conceptually, it is often divided into two parts: TBox $\mathcal{T}$ and ABox $\mathcal{A}$, where the TBox contains axioms about classes (Domain knowledge) and ABox contains axioms about instances (data), i.e:

$$\mathcal{O} = \ < \mathcal{T}, \mathcal{A} > \ = \ < \mathcal{N_C}, \mathcal{N_P}, \mathcal{N_I} > \tag{1}$$

The formalization of the knowledge using formal semantics allows its automatic interpretation by the machines. This is done by computing the logical consequences of the explicitly stated axioms in $\mathcal{O}$ to infer new knowledge. An interpretation $\mathcal{I}$ of an ontology $\mathcal{O}$ consists of (Δ^I,$.^I$), where Δ^I is called the domain of I ,and $.^I$ the interpretation function of I that maps every class to a subset of Δ^I and every property to a subset of $\Delta^I \times .^I$.

The interpretation of the ontology is calculated using DL reasoners, which provide a set of inference services where each inference service represents a specific reasoning task. This capability makes OWL very powerful for knowledge modelling and processing.

B. Overview of the proposed approach

In this paper, we propose a novel method combining deductive reasoning based on DL and semi-supervised clustering based on automatically generated constraints. The key idea in our method is to reason over the available domain knowledge in order to obtain semantically labeled instances, and use those labeled instances to generate constraints that will guide and enhance clustering. To enable the automatic interpretation of the expert knowledge, we use OWL to formalize the domain knowledge. OWL enables the exploitation of DL reasoners to predict instances types, while keeping our approach generic and modular. For the semi supervised clustering part, the constraints are generated automatically from the reasoning results and without using any labeled data. The proposed approach consists of the following steps (Figure 1):

1) *Semantic labeling based on the reasoning*
   - Conceptualization and formalization of the domain knowledge as the TBox of the ontology
   - Transformation of the data instances to ABox axioms
   - Reasoning over the constructed Knowledge Base (KB) for instance checking (Semantic classification)
2) *Semi Supervised clustering guided by generated constraints*
   - Constraints generation based on the results of the semantic labelling
   - Constrained clustering
   - Affinity of clustering results and reasoning results

The figure 1 shows the steps followed to guarantee an efficient use of the available knowledge and the automatically generated constraints. The rectangles correspond to the structural components of the process. The inputs are unlabeled data $X = \{x_i\}_{i=1}^n \in R^d$ where each instance x_i is described by a set of attributes $V = \{v_j\}_{j=1}^d$, and a domain knowledge formalized as the TBox of the OWL ontology. Our proposition can be applied for any problem where domain knowledge, even incomplete, is available.

In the next two subsections, we will describe in detail the different steps of our approach. We will also illustrate the effects of this approach when applied on a dataset(Figure 2).

1) Semantic labeling based on DL reasoning: As shown in the figure 1, we use OWL to formalize this knowledge. Based

[1]Ontology is considered here equivalent to Knowledge Base in DL literature

on DL, OWL reasoners can provide a set of inference services. An important aspect about the formalized domain knowledge that can be used in our method is its ability to bridge the semantic gap. The semantic gap is a known issue in image analysis, it's due to the difficulty when we want to extract high level semantics from low level features describing the image. Many works have been done to tackle this problem and ontologies have shown their efficiency to reduce the semantic gap. Using the elements introduced by the OWL language, such as complex concept definition, data properties restrictions and logical operators, the OWL ontology will be used as support for the formalized knowledge.

Algorithm 1 Semi Automatic Transformation of Data to ABox assertions

Inputs:

Data $X = \{x_i\}_{i=1}^{n} \in \mathbb{R}^d$ described by $V = \{v_j\}_{j=1}^{d}$

Domain Knowledge : $\mathcal{T} =< \mathcal{N_C}, \mathcal{N_P} >$

Output:

ABox : $\mathcal{A} = \{a_i\}_{i=1}^{n}$

Method:

```
for all p_k in N_P and v_i in V do
    Boolean Query = Does p_k correspond to v_i
    if Query.isTrue() then
        map(N_P, V).add(p_k, v_i)
    end if
end for
for all x_i in X do
    a_i := createOWLInstance();
    for all p_k in map(N_P, V) do
        a_i.addProperty(p_k)
        a_i.setPropertyType(p_k, T.getPropertyType(p_k))
        a_i.setPropertyValue(p_k, x_i.getValueOf(v_k))
    end for
    return a_i : OWL representation of x_i
    A.add(a_i)
end for
```

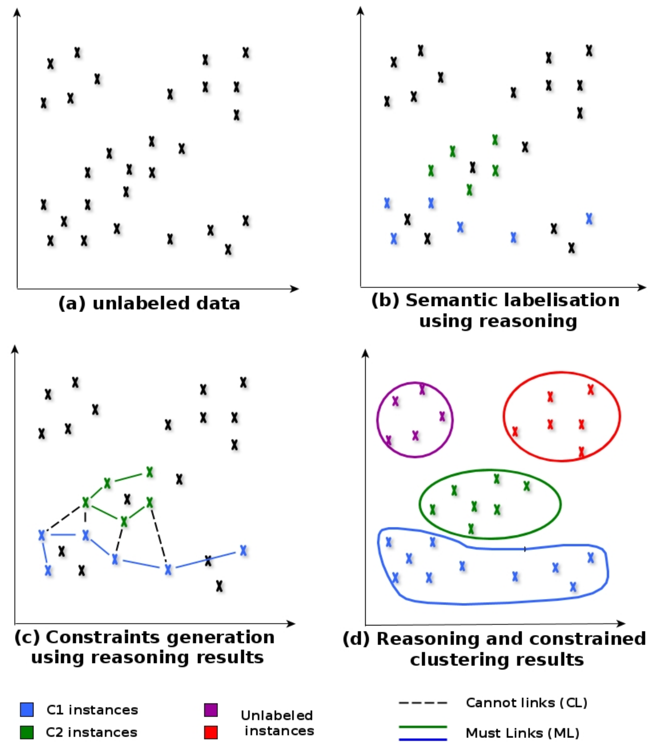


Fig. 2. Illustrative example (Should be analyzed in color mode)

Once the domain knowledge formalized, the second step is the projection of data instances to OWL axioms. After the completion of this process, each instance from the input data will be represented by a set of axioms. To perform this transformation, a dedicated semi automatic process has been developed. This process takes as inputs the TBox of the ontology and the data instances to transform. Based on the properties of the TBox and the variables describing the data, the process suggests to the user to map between them. Once the mapping is established, our process generates OWL axioms representing the data instances. Each instance is described by the properties available in the TBox which are getting their values from the data. At this point, all the required components are prepared to build up the Knowledge Base(KB).

In fact, the **KB** is the combination of the TBox and the ABox, where the TBox represents the formalized domain knowledge, while the ABox corresponds to the axioms describing the instances. As the OWL language is based on description logic, it allows the use of deductive reasoning to infer new knowledge possible. A set of inferences services can be used, such as *concept satisfiability*, *classification*, *realization*, etc.

In our method, we are interested in the *realization*, which attempts to find the most specific concept that a given instance belongs to. Performing this reasoning task over the constructed knowledge base will allow us to retrieve the instances of the concepts formalized in the TBox. Description logic reasoning operates under *open world assumption*, DL were designed to deal with incomplete information. This means that not all the data will be labeled, which is quite understandable as only the instances that fit completely with the definition of the concepts will be typed. Reasoning over the ontology will produce a set of labeled instances (Figure 2.b). The figure 2 represents a set of points that will be used to illustrate the effect of applying the different steps of the method. In this example, we suppose that the expert is interested in identifying 4 classes and that he has a formalized knowledge containing the definition of two concepts *C1* and *C2*. The two other concepts are not available (unknown or hard to formalize).

2) Semi supervised clustering guided by generated constraints: As shown in the figure 2(b), we obtain a set on labeled instances using the *realization* inference service. Based on those results, we generate a set of *ML* constraints to link the identified instances from the same concept between them and *CL* constraints for the instances identified as belonging to a different ones. The generated constraints will be used to enforce the clustering.

As mentioned above, two variants of constrained clustering exist, hard and soft constrained algorithms. In our case, The constraints are automatically generated based on reasoning over the available knowledge. In this automatized process, encoding the results of reasoning as soft constraints is the only way to guarantee the consistency of our approach as the knowledge can contain some approximations and produce some errors. In this step, we choose to use PCKMeans [11] as constrained algorithm, but our method can be applied using most soft constrained algorithms. Compared to classical KMeans, the objective function of PCKMeans is weighted by the ML and CL constraints :

$$R_{pckm} = \frac{1}{2}\sum_{x_i \in \chi} \|x_i - \mu_{li}\|^2 + \sum_{(x_i,x_j)\in M} w_{ij}1[l_i \neq l_j] + \sum_{(x_i,x_j)\in C} \overline{w}_{ij}1[l_i = l_j]$$

Where l_i ($l_i \in h_{h=1}^k$) is the cluster assignment of the instance x_i, and $w_{ij}1[l_i \neq l_j]$ and $\overline{w}_{ij}1[l_i = l_j]$ correspond to the cost of the violation of constraints $ml(x_i, x_j) \in M$ and $cl(x_i, x_j) \in C$. Note also that 1 is an indicator function with $1[true] = 1$ and $1[false] = 0$. x_i represents the instance affected to the partition χ_{li} with the centroid μ_{li}. The algorithm 2 show the adapted PCKMeans with automatically generated constraints.

IV. Experiments

In this section, we describe some elements of our implementation, the data we used to experiment our approach and the results we obtained. We apply our approach to the real world application of classification of satellite images. The images used are Landsat 5 TM images, with a spatial resolution of 30 meters and seven spectral bands. The size of each image is of 760x680 pixels. The images concern the region of the river Rio Tapajos in the Amazon, Brazil.

In our experiments, the only inputs are the TBox of the ontology containing the expert knowledge, i.e the formalization of two concepts: Vegetation and Water, and the pixels

Algorithm 2 Semi Supervised clustering with generated constraints

Inputs:
Dataset : $X = \{x_i\}_{i=1}^n \in \mathbb{R}^d$
Sub-dataset of labeled instance: $X_L = \{(x_i, C_l)\}_{i=1}^m$
Where $C_l \in \mathcal{N}_\mathcal{C}$ the set of Classes of the TBox
k : number of clusters

Method:
Generate the ML and CL constraints from X_L
$\lambda = size(\mathcal{N}_\mathcal{C})$
if $\lambda \geq k$ **then**
initialize $\{\mu_h^{(0)}\}_{h=1}^k$ with centroids of $\{N_p\}_{p=1}^k$
else if $\lambda < k$ **then**
initialize $\{\mu_h^{(0)}\}_{h=1}^\lambda$ with centroids of $\{N_p\}_{p=1}^\lambda$
if $\exists$ point x cannot-linked to all neighborhoods $\{N_p\}_{p=1}^\lambda$ initialize $\mu_{\lambda+1}^{(0)}$ with x
Initialize remaining clusters randomly
end if
Repeat until *convergence*
assign_cluster: assign each $x_i \in X$ to the cluster h^*, for $h^* = argmin(\frac{1}{2}\|x_i - \mu_h^{(t)}\|^2 + w\sum_{(x_i,x_j)\in ML} 1[l_i \neq l_j] + w\sum_{(x_i,x_j)\in CL} 1[l_i \neq l_j]$
estimate means :
$\{\mu_h^{(t+1)}\}_{h=1}^k = \{\frac{1}{\|X_h^{(t+1)}\|}\sum_{x\in X_h^{(t+1)}} x\}_{h=1}^k$
$t = t+1$

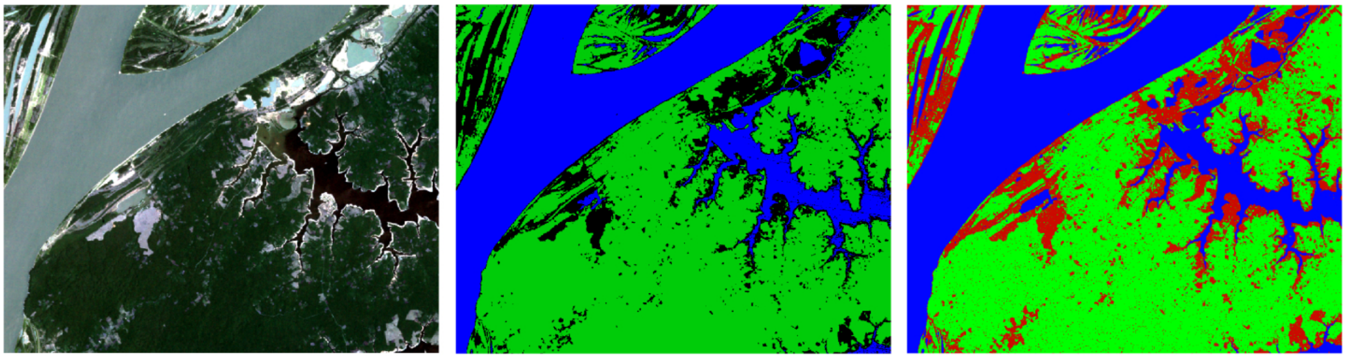
of the images. We do not use any labeled data. To build the corresponding TBox, several spectral bands and indices were used. The concepts were defined using the seven bands : TM1,...,TM7 and the spectral indices : ndwi, ndvi[2] [14]. for example, the water concept is defined as follows :
$Water\_Pixel \equiv Pixel \wedge ((\exists TM4. < 0.05 \wedge \exists ndvi. < 0.01) \vee (\exists TM4. < 0.11 \wedge \exists ndvi. < 0.001))$

Several frameworks have been used to implement our approach. A dedicated process for the pre-treatment and the transformation of the images have been developed using the *Orfeo ToolBox* library. Concerning the semantic side, the transformation of the data to OWL individuals is ensured by a Java program that uses the *OWL API* and a semi-automatic mapping. Pellet is the DL reasoner that have been chosen to perform the realization task and materialize the deduced type of the pixels. Pellet has a datatype oracle that can reason with XML Schema based datatypes, which allow us to use concrete domains to define our concepts. Finally, constraint generation and the constrained clustering PCKMeans have been also implemented with Java.

A. Results and discussions

The experiments we conducted have multiple objectives. First, they show the feasibility of our approach and the advantages that comes from the simultaneous exploitation of the ontology and the generated constraints.

[2]Normalized Difference Vegetation Index



(a) The raw image

(b) The result of the classification based on ontology reasoning

(c) The result of our approach (constrained clustering + ontology)

#Water #Vegetation Unknown Cluster 1

Fig. 3. Application of our approach to a Landsat satellite image

Images	Clustering		Ontology			Proposed approach	
	Prec.	F-Mes.	% labeled	Prec.	F-Mes.	Prec.	F-Mes.
Image 1	0.8899	0.8764	83,6	0.8359	0.8360	**0.9445**	**0.9296**
Image 2	0.8701	0.8598	81,26	0.8125	0.8126	**0.9271**	**0.9181**
Image 3	0.8889	0.9241	90,33	0.9031	0.9020	**0.9299**	**0.9304**

TABLE I
EXPERIMENTS RESULT ON IMAGES OF THE AMAZONIAN REGION

Figure 3 shows the results of applying our approach to one of the images used in the experiments. For a better understanding, the figure 3 should be analyzed in color mode. The figure 3.a represents the raw image in true colors, the figure 3.b shows the intermediate step of ontology reasoning and the final result of the approach is illustrated in the figure 3.c. We can visually see that our approach improves the results obtained only with reasoning over the ontology. Two important elements are shown in this figure. The first point we can notice is the labeling of water present in the top left corner of the images. Those pixels have been semantically labeled using constrained clustering. This shows how our approach can complete the knowledge about the concepts of the ontology. Here, the definition of the experts have not been sufficient to label those pixels, but using the constrained clustering, those pixels have been correctly labeled. The second element is the apparition of the new cluster, which have been identified with the clustering. This cluster, representing the bare soil, is not specified in the ontology but have been detected by the constrained clustering. This shows how our approach can deal with different paradigms and produces labeled and unlabeled clusters simultaneously. When we apply K-Means and our approach on the same image (Figure 4), we can see the improvements made by the ontology. The first difference is the automatic semantic interpretation. The second observation concerns the confusion between water and vegetation when we use K-Means. The expert explains that those errors are due to the nature of the Amazonian forest, where some vegetation grows on wet soils. The ontology is very useful in this case, where the available domain knowledge helps our approach to distinguish the instances of the two concepts (vegetation and water). To evaluate the quality of the results, we calculate the Precision and the F-Measure metrics based on a reference classification made by the expert. We also compare the scores of our approach with those obtained by K-Means and Ontology reasoning. We have to point out that the evaluation of K-Means is made after the intervention of the expert to label each cluster, which constitute an important difference with our approach where the semantic labeling is automatic. Table 1 shows the performances of the different methods on the three images provided by the expert. Reasoning over the ontology to label the pixels is operated under the open world assumption (OWA) and the TBox contains the formalization of the two concepts of Water and Vegetation. This configuration leads to a partial labeling of the instances (83.6 % for the image 1). If we consider only the labeled pixels, the precision of the ontology is very high (between 0.97 and 0.99). However, the ontology is incapable to label the pixels with the third concept (as no definition is given). It is also incapable of labeling all the vegetation and water pixels as they do not fit exactly with the specifications of the concept. This disadvantage of the ontology is also a motivation to use our approach, where all the pixels are classified. The table also shows the improvements of the clustering results in our approach when compared to K-Means, which demonstrates that, in addition to the semantic interpretation, our approach has better performance.

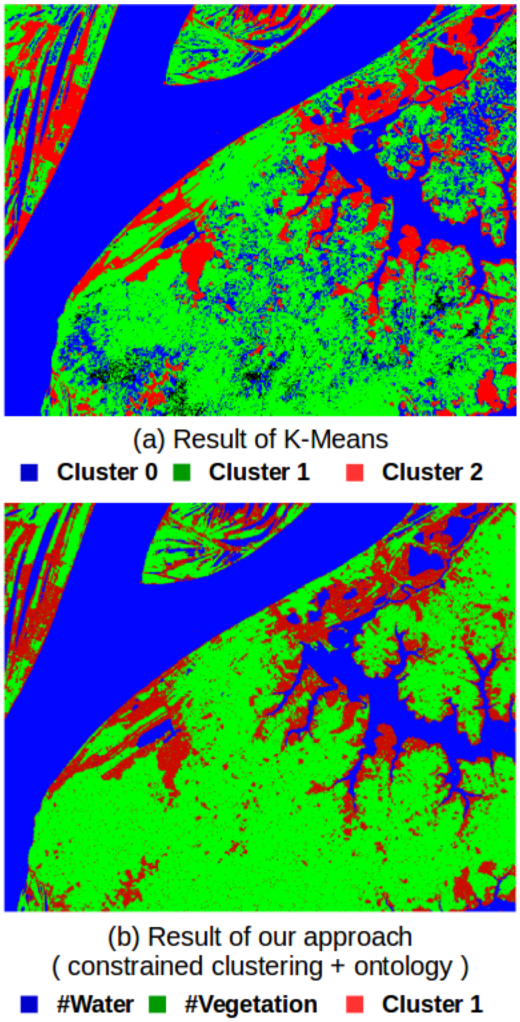


Fig. 4. The results of K-Means (a) and our approach (b) applied to the same image

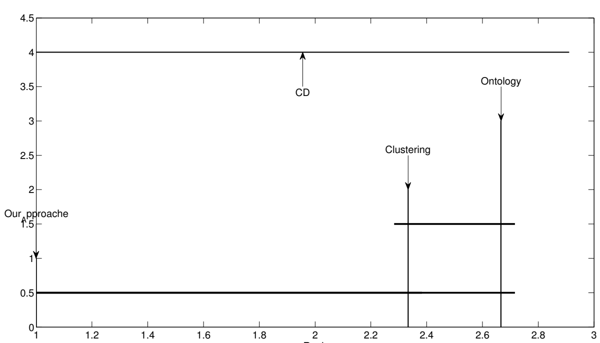


Fig. 5. Friedman test for Precision results

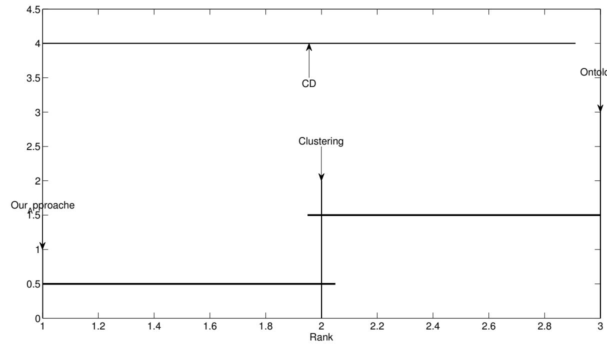


Fig. 6. Friedman test for F-Measures results

We also evaluate the obtained results using the Friedman test. A critical Friedman diagram represents a projection of average ranks of classifiers on enumerated axis. The classifiers are ordered from left (the best) to right (the worst) and a thick line which connects the classifiers were the average ranks are not significantly different (for the 5% level of significance). From this two tests (Figures 5 and 6), it can be observed that our approach outperforms the classical clustering and the ontology based classification as it is situated on the left side of the both figures.

V. Conclusion

We have presented in this paper a new hybrid approach combining reasoning over an ontology and clustering, guided with automatically generated constraints. By proposing an approach that is both deductive and inductive, our method can exploit the available knowledge even if it is incomplete. We have applied our approach to the real world problem of satellite image classification. The results have shown that our approach improves the quality of the clustering while introducing an automatic semantic labeling.

In the future, we plan to extend our approach by introducing the selection of most important constraints as in this method, all the constrains have the same weight. Another perspective concerns the enrichment of the ontology by adding new thematic concepts.

Acknowledgment

This work was supported by the French Agence Nationale de la Recherche under Grant ANR-12-MONU-0001.

References

[1] G. Forestier, A. Puissant, C. Wemmert, and P. Gançarski, "Knowledge-based region labeling for remote sensing image interpretation," *Computers, Environment and Urban Systems*, vol. 36, no. 5, pp. 470–480, 2012.

[2] Z. Falomir, E. Jiménez-Ruiz, M. T. Escrig, and L. Museros, "Describing images using qualitative models and description logics," *Spatial Cognition & Computation*, vol. 11, no. 1, pp. 45–74, 2011.

[3] S. Andres, D. Arvor, and C. Pierkot, "Towards an ontological approach for classifying remote sensing images," in *Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on*. IEEE, 2012, pp. 825–832.

[4] S. Basu, M. Bilenko, and R. J. Mooney, "A probabilistic framework for semi-supervised clustering," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 59–68.

[5] I. Davidson and S. Basu, "A survey of clustering with instance level constraints," *ACM Transactions on Knowledge Discovery from Data*, pp. 1–41, 2007.

[6] K. Wagstaff and C. Cardie, "Clustering with instance-level constraints," 2000.

[7] K. Wagstaff, C. Cardie, S. Rogers, S. Schrödl *et al.*, "Constrained k-means clustering with background knowledge," in *ICML*, vol. 1, 2001, pp. 577–584.

[8] N. Shental, A. Bar-Hillel, T. Hertz, and D. Weinshall, "Computing gaussian mixture models with em using equivalence constraints," *Advances in neural information processing systems*, vol. 16, no. 8, pp. 465–472, 2004.

[9] I. Davidson and S. Ravi, "Agglomerative hierarchical clustering with constraints: Theoretical and empirical results," in *Knowledge Discovery in Databases: PKDD 2005*. Springer, 2005, pp. 59–70.

[10] I. Davidson, K. L. Wagstaff, and S. Basu, *Measuring constraint-set utility for partitional clustering algorithms*. Springer, 2006.

[11] S. Basu, A. Banerjee, and R. J. Mooney, "Active semi-supervision for pairwise constrained clustering." in *SDM*, vol. 4. SIAM, 2004, pp. 333–344.

[12] W. O. W. Group *et al.*, "Owl 2 web ontology language document overview," 2009.

[13] F. Baader, *The description logic handbook: theory, implementation, and applications*. Cambridge university press, 2003.

[14] R. DeFries and J. Townshend, "Ndvi-derived land cover classifications at a global scale," *International Journal of Remote Sensing*, vol. 15, no. 17, pp. 3567–3586, 1994.