

The digital Earth Observation Librarian: a data mining approach for large satellite images archives

Mihai Datcu, Alexandru-Cosmin Grivei, Daniela Espinoza-Molina, Corneliu Octavian Dumitru, Christoph Reck, Vlad Manilici & Gottfried Schwarz

To cite this article: Mihai Datcu, Alexandru-Cosmin Grivei, Daniela Espinoza-Molina, Corneliu Octavian Dumitru, Christoph Reck, Vlad Manilici & Gottfried Schwarz (2020): The digital Earth Observation Librarian: a data mining approach for large satellite images archives, Big Earth Data, DOI: [10.1080/20964471.2020.1738196](https://doi.org/10.1080/20964471.2020.1738196)

To link to this article: <https://doi.org/10.1080/20964471.2020.1738196>



© 2020 The Author(s). Published by Taylor & Francis Group and Science Press on behalf of the International Society for Digital Earth, supported by the CASEarth Strategic Priority Research Programme.



Published online: 16 Apr 2020.



Submit your article to this journal [↗](#)



Article views: 323



View related articles [↗](#)



View Crossmark data [↗](#)

The digital Earth Observation Librarian: a data mining approach for large satellite images archives

Mihai Datcu^{a,b}, Alexandru-Cosmin Grivei^b, Daniela Espinoza-Molina^a,
Corneliu Octavian Dumitru^a, Christoph Reck^c, Vlad Manilici^c and Gottfried Schwarz^a

^aGerman Aerospace Center, Remote Sensing Technology Institute, Wessling, Germany; ^bResearch Center for Spatial Information, University Politehnica of Bucharest, Bucharest, Romania; ^cGerman Remote Sensing Data Center, German Aerospace Center, Wessling, Germany

ABSTRACT

Throughout the years, various Earth Observation (EO) satellites have generated huge amounts of data. The extraction of latent information in the data repositories is not a trivial task. New methodologies and tools, being capable of handling the size, complexity and variety of data, are required. Data scientists require support for the data manipulation, labeling and information extraction processes. This paper presents our Earth Observation Image Librarian (EOLib), a modular software framework which offers innovative image data mining capabilities for TerraSAR-X and EO image data, in general. The main goal of EOLib is to reduce the time needed to bring information to end-users from Payload Ground Segments (PGS). EOLib is composed of several modules which offer functionalities such as data ingestion, feature extraction from SAR (Synthetic Aperture Radar) data, meta-data extraction, semantic definition of the image content through machine learning and data mining methods, advanced querying of the image archives based on content, meta-data and semantic categories, as well as 3-D visualization of the processed images. EOLib is operated by DLR's (German Aerospace Center's) Multi-Mission Payload Ground Segment of its Remote Sensing Data Center at Oberpfaffenhofen, Germany.

ARTICLE HISTORY

Received 19 November 2019
Accepted 18 February 2020

KEYWORDS

Earth observation;
TerraSAR-X; data mining
system

1. Introduction

The Earth observation data, gathered throughout the past decades, fall very well in the big data 3Vs (Volume, Velocity, and Variety) paradigm (Laney, 2001). In order to bring the end-users closer to the knowledge hidden within the huge volumes of data, new tools are required, which are capable of finding and retrieving the sought-for information. A multitude of domains, such as agriculture, ecology, forestry, urban planning, security, and many others can benefit from the hidden potential of EO data. A general framework, which could serve all of them, has to overcome the differences which occur due to the multitude of various sensor data acquired throughout the years. Currently, EO image data are stored in data warehouses, and depending on the source mission/sensor, they can have different spatial resolutions, number of spectral bands, and were obtained through active or passive methods.

Manual annotation of the gathered EO data is not feasible because of the overwhelming volumes of data being involved. Computers can ease this process, but several issues must be tackled. The difference between human and machine knowledge, and the way each of them understands data is a potential impediment. Humans rely on visual perception (color, texture, and shape), while computers require appropriate numerical representations in order to find logical correlations in the data. Thus, the question of how to represent the data arises. Another impediment is the overall volume of the data. Even if all information is correctly extracted, finding it in the raw data and within a reasonable amount of time is quite a challenge. Thus, defining an efficient method for storing and indexing imaging data with supplementary information may be the most challenging task as any proposed solution must be highly scalable, since ongoing EO missions provide increasing amounts of data.

We introduce a new paradigm for Earth observation: Data Knowledge Discovery. This paradigm defines the entire chain “data-information-knowledge-value” and relies on its meaningful EO content extraction, i.e. the semantic and knowledge aspects.

We develop user-invariant and domain-compensatory EO methods for the individual users and for subjective domain biases. The derived models generate a shareable knowledge body as a means to enable the communication of fragmented knowledge learned from meta-data, image data, and other data in synergy with the domain expertise of EO users.

This paper describes the full functionality of EOLib – an Earth Observation Image Librarian, a software framework which tries to overcome the 3Vs of big EO data, working directly with the PGS data (Wolfmuller et al., 2009), in an effort to provide information each time when data are produced. In [Section 2](#) we present several EO data mining systems which offer solutions for some of the impediments encountered when tackling large volumes of EO data. The main data source for EOLib, the TerraSAR-X test dataset, is described in [Section 3](#), together with all the associated pre-processing functions. [Section 4](#) describes the overall architecture of the system, and each of the individual, specialized modules. The algorithms used in some of the key components of the system are presented in [Section 5](#). A step-by-step presentation of the system functionality is given in [Section 6](#), while [Section 7](#) is reserved for conclusions and potential future improvements.

2. Related work

2.1. State of the art

Throughout the years, several Image Information Mining (IIM) systems and concepts have been developed. These tried to solve some of the Earth Observation data mining issues through different methods. A solution for data classification and fast queries of compound structures from very high-resolution optical EO data is proposed in (Gueguen, 2015). Here, a bag of features with spectral and shape representation of multi-scale segmented images (tiles) is employed. Each tile is represented by the distribution of visual words it fully contains; MinTree (Salembier, Oliveras, & Garrido, 1998) is used to produce the necessary multi-scale segmentations. The bag of features employs nine spectral descriptors and four shape descriptors. For test purposes, the authors of Salembier et al. (1998) used the UC Merced 21-class dataset (Yang & Newsam, 2010) and two WorldView-2 (8-band) images for feature definition.

feature vectors which were grouped into five categories: clouds, flowers, foliage, mountains, and water.

Wang and Song proposed the SBRSIR (Semantics Based Remote Sensing Image Retrieval) system (Wang & Song, 2013) based on semantic spatial relationship detection. The proposed scheme is composed of a series of steps: image decomposition, segmentation, object-based classification, spatial relationship reasoning, and scene modeling based on semantic spatial relationship detection. Each image is decomposed into blocks using a quin-tree structure. An image segmentation method divides the image blocks into parcels. Each parcel has its VF (visual features: texture and color) extracted and classified into objects. The area, orientation, and the semantic topology features are calculated and stored for the labeled parcels. Each generated parcel is vectorized and its features (spectral features – spectral mean and standard deviation, shape features – such as parcel area, perimeter, and length/width ratio, texture features – GLCM (Popescu, Gavat, & Datcu, 2012) and Gabor coefficients) are calculated and stored. The test dataset for the SBRSIR method consists of three SPOT-5 images, five GeoEye images, and six ALOS images acquired in 2000 and 2011. These were split into 580 (1,024 × 1,024 pixel) sub-images and processed using PCA, and decomposed into quin-tree sub-blocks (2,900 second level blocks with 512 × 512 pixels). The level-2 sub-blocks were decomposed into 16 × 16 feature blocks and their visual features were calculated and stored. The parcels were classified using an SVM with an RBF kernel (Chang & Lin, 2011) (with cost=25 and gamma=40 parameters). During testing, one to five semantic categories were selected, and the search time on the entire database ranged from 46 to 66 s.

Other mentionable systems are KIM (Knowledge driven content-based image Information Mining) (Datcu et al., 2003) which relies on the stochastic modeling of probabilities providing an application-independent EO data mining system; and GeoIRIS (Shyu et al., 2007) which tries to manage the huge volumes of data through automatic pre-processing and indexing of image data. Both systems incorporate content-based retrieval methods and use features to represent the image content. Data processing requires a high level of expertise from the end-users which can limit the access for some of them. A solution for this limitation is presented in (Rasiwasia, Moreno, & Vasconcelos, 2007), which uses semantics-based queries in order to reduce the level of expertise required from the end-users. The latest, more complex systems, such as (Espinoza-Molina & Datcu, 2013), use multiple sources of information to optimize the search patterns.

Today's EO paradigms and technologies are largely domain-oriented. As an example, ESA's Thematic Exploitation Platforms (TEPs) (Thematic exploitation platform [TEP], 2019) are designed and focused on Coastal Areas, Forest, Urban Areas, and other application domains, integrating a standard processing chain that has low user interaction. Within the currently developed Copernicus system, the Data and Information Access Services (DIAS) (Data and information access services [DIAS], 2019) are a major achievement but still in a "classic" paradigm technology.

2.2. Benefits of the EOLib system

In an effort to obtain a complete EO mining system, we have developed and implemented Earth Observation Image Librarian (EOLib) system. It is a latest-generation active Image Information Mining (IIM) system, which has novel functions for image content exploration and knowledge discovery. EOLib is integrated and operated within the Payload Data

Ground Segment (PDGS) of the German TerraSAR-X mission (TerraSAR-X, 2019). It allows a more complete exploration of SAR data by the use of large-scale information mining functions. The system is capable of closing the “data-information knowledge-value” chain, containing modules which extract information from both new data and existing databases. Its modular architecture enables it to grow as new EO missions produce new data. The EOLib software framework has been tested on a dataset totaling more than 700 GB, and the results are detailed in [Section 6](#).

The system offers to the users a list of feature extraction methods (Blanchart et al., 2011b), and the exploration of the image content (e.g., classification, annotation) is made with very few examples due to the developed algorithm (Dumitru & Datcu, 2013) which is also capable of extracting a large variety of categories from the images.

Another important point of the system is the unique hierarchical semantic annotation schema (Dumitru, Schwarz, & Datcu, 2016) that initially has been developed for high-resolution images (e.g., TerraSAR-X), and was now upgraded also for medium-resolution and low-resolution images.

The output of the system can be a classification map, a list of query results (combining metadata with semantic labels), or some statistics.

3. Data overview

For our tests, from the TerraSAR-X archives (TerraSAR-X, 2013), we prepared a dataset that mainly covers urban and industrial areas together with their infrastructure from all over the world. The dataset consists of 1,100 full scenes with a distribution presented in [Figure 1](#).

The TerraSAR-X imaging parameters are detailed in [Table 1](#). The scenes were selected based on their availability, their content, the typical diversity of country-specific land covers, and their acquisition parameters. For storage and indexing purposes, each image is divided into patches of different sizes (e.g., 200×200 and 160×160 pixels). Feature vectors are extracted from each patch and the information was stored in a database.

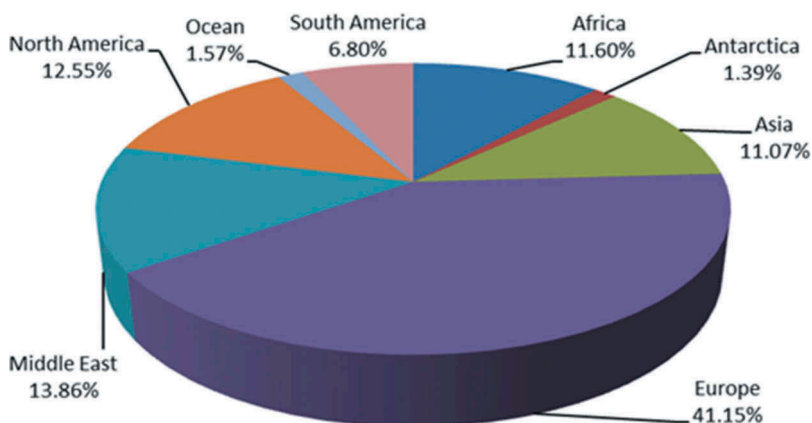


Figure 1. Geographical distribution among continents of the TerraSAR-X dataset used to test the system.

Table 1. TerraSAR-X imaging parameters.

SAR instrument	TerraSAR-X
SAR band	X-band
Instrument mode	High Res Spotlight
Polarization	HH or VV
Orbit branch	Ascending or Descending
Incidence angle	Between 20° . . . 50°
Product type	MGD – Multi-Look Ground Range Detected (amplitude data)
Enhancement	Radiometrically enhanced
Ground range resolution	2.9 m
Pixel spacing	1.25 m
Full image sizes (cols x rows)	4,200 x 6,400
Selected patch sizes (cols x rows)	200 x 200/160 x 160 pixels

Given the number of images, the dimensions (in pixels) of each individual image, and the number of segmentation levels, the database then contains 9.24 million vectors (per each type of image descriptor).

4. EOLib system architecture

EOLib (our Earth Observation Image Librarian) is an operational Image Information Mining (IIM) system with a modular structure integrated within the Payload Ground Segment (PGS) of the TerraSAR-X mission, at DLR's premises at Oberpfaffenhofen, Germany. It enables content interpretation and helps create a correlation between semantic labels and object categories. EOLib offers multiple functionalities such as: data model generation, visual data mining and knowledge discovery in databases (including semantic annotation), queries, and epitome production. The system is capable of understanding TerraSAR-X data and their meta-data in order to improve the entire processing chain consisting of pre-processing, feature extraction, semantic annotation of the EO data, and the creation of product catalogs. Through EOLib, we attained a more complete and comprehensive exploration of the satellite image data. The system is intended as a long-term exploitation of the PGS component by Data Mining (DM) and Knowledge Discovery in Database (KDD) elements. The architecture allows for continuous evolution, migration and scalability. The system has applications both for SAR and optical image data. It also provides the infrastructure for further development of algorithms, methods, and sub-components for DM and KDD. EOLib manages to provide the following high-level functionality:

- Operates in real time (once the data from the PGS have been processed and ingested into the database)
- Operates on very large data volumes
- Extends the functionality of the PGS with new Knowledge Discovery components
- Integrates novel components into the PGS architecture.

The system currently runs on a single server machine at DLR's premises. The server is composed of 8 CPUs (each with 10 cores and 20 threads), 256 GB of RAM, 22 TB of HDD storage, and runs under Ubuntu Linux Operating System. The EOLib software is based on Java and uses MonetDB (MonetDB, 2019) as its Database Management System.

The architecture of the EOLib system together with its main components and their interactions is shown in Figure 2. Rectangles with rounded corners are used to represent the main components (PGS components are presented in blue, while EOLib components are given in orange). Arrows are used to represent interfaces which interconnect the components. The orientation of the arrows presents the data flow direction, while the data passing through the interface are presented as a rectangle above or near the arrow. The components which require the interaction of the user are presented by a user icon on their upper-right corner.

We will further describe the EOLib components given in orange in Figure 2. The processing chain starts with the pre-processing of TerraSAR-X products and ends with the semantic annotation. Figure 3 outlines the main modules of EOLib together with the existing interface with the PGS, the end-user, and the data flows between modules. Components are shown as rectangles, while functional sub-components are represented by folders. Dashed arrows represent used-by relations, while full arrows are used to represent data flows. The main components of EOLib which will be described in the following subsections are: Data Model Generation, Data-Mining Database, Query Engine, Content-based Image Retrieval, Visual Data Mining, Knowledge Discovery in Databases, Epitome Generation, and System Evaluation.

Operation of the EOLib system by the users: The relation between the modules, and the steps of their use are presented below before detailing all the modules. Due to our concept that the data have already been extracted from the archive and processed via Data Model Generation (DMG), the system can be operated in three ways:

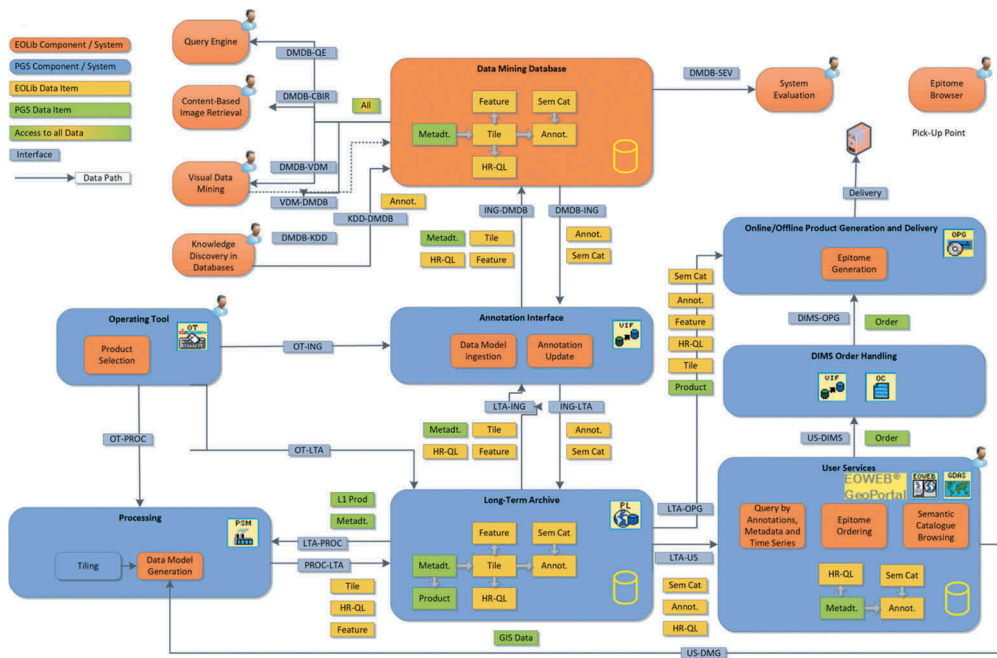


Figure 2. EOLib – system architecture and interfaces with the Payload Ground Segment (PGS). A full description and details of each component from the figure can be found in (EOLib, 2019).

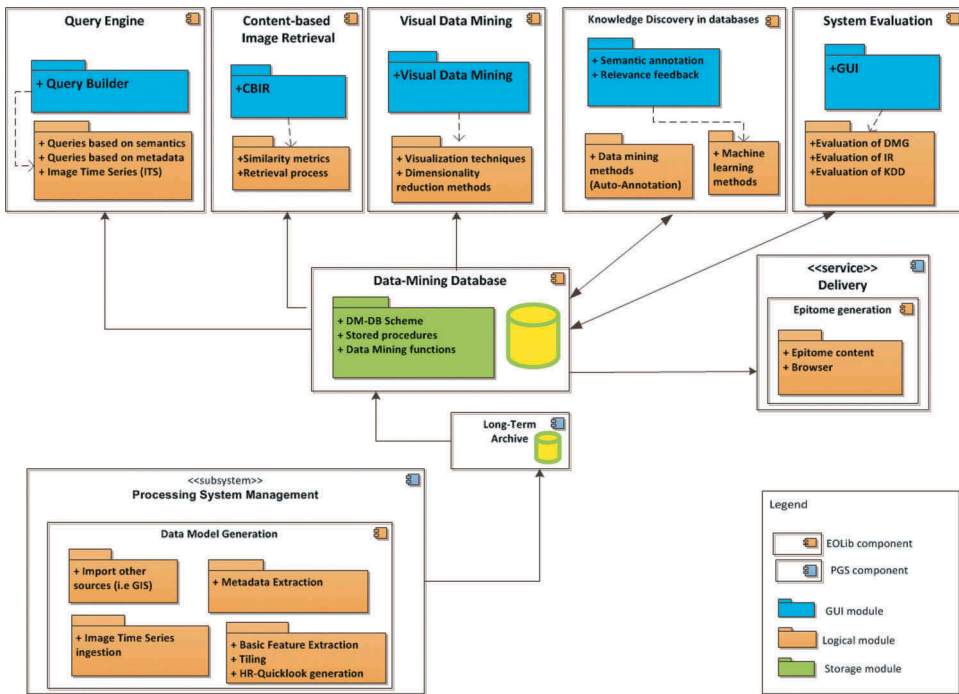


Figure 3. EOLib – core components, overall functionality, and dependencies.

- (1) *Visual Data Mining*: The users navigate through the data (based on the extracted features) and can understand the content of the data as well as potential cluster-groupings.
- (2) *Image Mining*: The users run a machine learning tool via its interactive GUI (Graphical User Interface) based on an active learning module (Blanchart et al., 2011b) which is exploiting all actionable information. Its functions are: search, browse, and query for image patches (the full image was already tiled into square sub-images with a side length defined before in the DMG). The discovered relevant structures are semantically annotated and stored into the Data-Mining Database. For this, we use the features extracted from the patches.
- (3) *Data Mining*: Via the Query Engine, the users perform SQL searches, queries, and browse or extract the data analytics information. Data Mining uses image features, image semantics, and selected satellite product meta-data.

4.1. Data model generation

The Data Model Generation (DMG) module is the first element of the TerraSAR-X product analysis chain. As can be seen in Figure 3, it is part of the Processing System Management (PSM), which is a sub-component of the DIMS system (Wolfmuller et al., 2009). DMG processes the Level-1b (L-1b) TerraSAR-X products. The TerraSAR-X products are composed of a GeoTIFF image and are accompanied by XML meta-data files. The processing steps (shown in Figure 4) of the DMG module are:

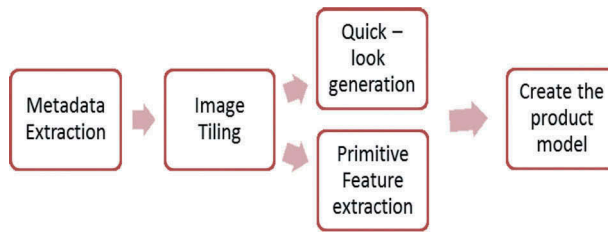


Figure 4. DMG – processing steps and workflow.

- Meta-Data extraction
- Tiling of the image content at multiple resolutions
- Basic feature extraction (BFE)
- High-resolution quick-look (HR-QL) generation
- Ingestion of image time-series (ITS)
- Import of other sources as, for example, GIS information.

During the meta-data extraction phase, relevant entries such as location, angles, and mission information are obtained from the XML file which accompanies the image data. A multi-resolution tile pyramid (several grid levels with tiles of different size) and quick-looks are created in order to improve visual data manipulation and to aid the machine learning process. BFE is applied to the previously generated tiles which are considered as distinct entities, using the relations between pixels within each patch and no single pixel information (Popescu et al., 2012). Gabor filters (Manjunath & Ma, 1996), (Singh & Datcu, 2013) and Weber Local Descriptors (WLDs) (Cui, Dumitru, & Datcu, 2013) are used as feature extraction methods in the DMG module. In order to aid the data exploration process, in this phase, a quick-look image (a compressed RGB or gray-scale image, either optical or SAR) is created and stored in the database.

The final product model of the DMG is an XML file which contains the previously extracted meta-data, the grid levels, the tile information, and data vectors representing the extracted image descriptors derived by the Gabor and WLD algorithms. The descriptors are used later in the classification process of the knowledge discovery module. In addition, the XML file is ingested into the Data Mining Database which is described below.

4.2. Data mining database

The products generated by the DMG are mapped into a relational database. The Data-Mining Database (DM-DB) module manages data handling, storage, administration, and some of the image processing for all EOLib components. It is composed of data, a database schema, and stored procedures (which support the data mining operations). The stored information is composed of meta-data, image parameters, tile parameters, quick-looks, feature vectors, and semantic labels. Thus, all information about the TerraSAR-X products are stored into a table-based scheme, which implements the proper relationships between the tables and indices for optimized processes. This module is supported by a relational Data Base Management System (DBMS), which represents a core component of the system, and interacts with all

the user-oriented components and supports their functionality. This database also stores the semantic annotations made by the end-users and provided through the Knowledge Discovery component.

4.3. Query Engine

The Query Engine (QE) consists of a GUI (Graphical User Interface) named Query Builder, and the Query Engine Core which uses the queries received through the GUI to retrieve data from the DM-DB. The end-user can use the QE to find and retrieve desired scenes with specific content. The image database can be queried using meta-data, semantic definitions, and multi-temporal parameters. In order to generate semantics-based queries, the end-user can select one or several labels from the options being available in the semantic catalog. The most complex queries can combine meta-data and semantic definitions in order to produce more specific results. The QE GUI (standalone or web-based) helps the user to access the query services of the image archives. The queries read part of the data model stored within the data mining database (DM-DB) and the output is a list of data items like image patches or full images.

4.4. Content-based image retrieval

The Content-based Image Retrieval (CBIR) consists of a GUI which interacts with the user at one end, and receives images from the Query Engine Core. The latter reads the user-designed queries, and uses them to retrieve the data from the DM-DB. This component is based on the query-by-example (QbE) concept, in which the user provides an external or internal image, and retrieves images which have similar content. The semantic labels are determined through one or multiple examples, and the results are returned based on the similarity of features extracted from the query samples, and the previously ingested images. The results are ordered in decreasing order of their similarity.

4.5. Visual data mining

The Visual Data Mining (VDM) module provides a graphical and intuitive method to search highly complex, and nonvisual data sets being stored within the database. The selection of different images in 2-D or 3-D space is achieved through visualization techniques, data reduction methods, and similarity metrics to group the retrieved images into relevance clusters (Singh & Datcu, 2013). VDM provides browsing, querying, and zooming, which enable the end-user to navigate the data mining database. Each 2-D and 3-D projection of the database is obtained through a dimensionality reduction algorithm called t-distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten & Hinton, 2008). It uncovers heterogeneous hidden structures in the data, exposes natural clusters, and smoothes non-linear variations along the feature dimensions.

The VDM module consists of a GUI which allows the user to navigate through the data, and a VDM Core which retrieves the data from the DM-DB and adapts them in order to be displayed through the GUI.

4.6. Knowledge discovery in databases

Manual semantic annotations are time-consuming and user-intensive tasks. The process of applying semantic labels to the ingested data can be optimized through supervised or un-supervised machine learning methods.

The Knowledge Discovery in Databases (KDD) module consists of a GUI which receives the end-user input, and a KDD Core which processes the user's input, and applies it to retrieve data from the DM-DB. This module is used to create semantic definitions of the image content. This is achieved through semi-supervised machine learning methods, and uses a relevance-feedback loop in order to include human expertise in the annotation process and the semantic category definition.

The active learning algorithms are iterative sampling schemes where a classifier is tuned at each iteration by providing it with newly labeled samples. In turn, this helps achieve greater accuracy with fewer training labels (Espinoza-Molina, Datcu, & Balint, 2014).

These concepts for the labeling of image content are implemented in the KDD component (see Figure 5) as follows:

- The algorithm starts by providing the end-user, through the help of a GUI, with a set of image tiles and their source image
- The user gives positive and negative examples by clicking on the corresponding tiles
- The training data set is provided interactively by the end-user who generates a list of positive and negative examples
- The training set is passed to the classifier which makes predictions and retrieves results with feature vectors similar to the positive samples.

The active learning component is based on a Support Vector Machine (SVM) with cascaded learning (Blanchart, Ferecatu, Cui, & Datcu, 2014). The end-user visually analyzes the results and refines the training data by giving more examples and by correctly re-labeling the erroneous results. The actions are supported by the Relevance Feedback (RF) interface contained in the GUI, which generates an automatic ranking of the returned images. The active learning loop is stopped when the user considers the results to be satisfactory. The process can end with a definition of a semantic category for the ranked tiles, which is further stored within the data mining database catalog.

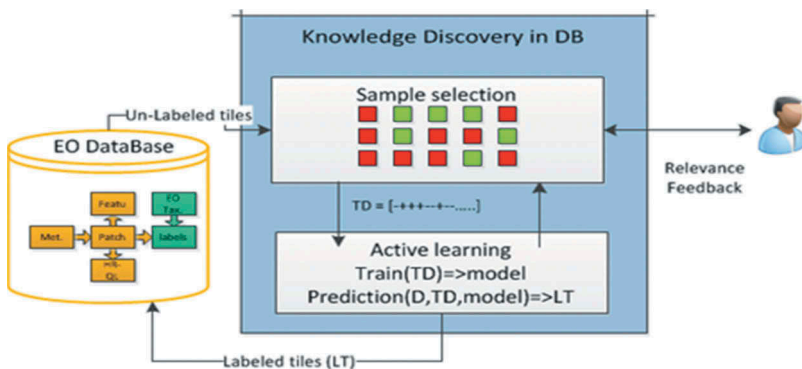


Figure 5. Knowledge discovery in databases workflow.

The two goals of the active learning phase are to learn the targeted image category as accurately and exhaustively as possible, and to minimize the number of iterations in the relevance feedback loop.

4.7. Epitome generation

An epitome is a new product type, for future satellite ground segments, consisting of a semantic label generated through a feature-based description of the image product content. It includes the semantic labels of the image tiles and is produced through image feature extraction and semantic annotation. The generated information can be delivered independently or together with a standard EO product. An epitome represents an enriched and added-value product that can be exploited interactively with appropriate end-user applications. The epitome is generated in the Delivery PGS component and is composed of meta-data, high-resolution quick-looks, basic features, and all the annotations. The epitome generation component reads the information from the Data Mining Database and transfers it to the delivery service (Espinoza Molina & Datcu, 2016).

4.8. System evaluation

The System Evaluation (SE) module consists of a GUI which interacts with the end-user, and receives its input from the SE Core, which exploits the user's input to retrieve and process data from the DM-DB. The SE component collects information and statistical data about the activity of the EOLib components. Quality metrics and evaluation procedures are defined according to the evaluated element.

This component requires pre-defined test data sets, and is also in charge of the quality evaluation of individual EOLib components, as well as system validation.

5. EOLib algorithms

EOLib uses different feature extraction methods according to the input data. The feature extraction algorithms read the original L-1b EO products (e.g., for detected and geo-coded TerraSAR-X products, data are represented by unsigned 16-bit values, while for QuickBird unsigned 11-bit values are used). The algorithms are applied to the patches generated from a full image, according to the analyzed window (and its patch size, as a parameter). The main feature extraction algorithms implemented in EOLib are: Weber Local Descriptor (WLD) (Chen et al., 2010), Non-Linear Spectral Feature Description (Popescu et al., 2012), Gabor Linear Moments (Manjunath & Ma, 1996), Adaptive Weber Local Descriptor (AWLD) (Cui et al., 2013), and Multi-Temporal Similarity Metric Based on Wavelet Modeling (Cui & Datcu, 2012). The currently implemented classifiers are Semi-Automatic Semantic Annotation Based on a Support Vector Machine with Relevance Feedback (SVMRF) (Dumitru, Cui, Schwarz, & Datcu, 2015), and Cascaded Active Learning for Object Retrieval (CALOR) (Blanchart, Ferecatu, & Datcu, 2011a).

5.1. Weber Local Descriptor (WLD)

This descriptor consists of two components: differential excitation, and orientation. It is inspired by Weber's Law, which is a psychological law. It states that the change of a stimulus (such as sound, lighting) that will be just noticeable is a constant ratio of the original stimulus. When the change is smaller than this constant ratio of the original stimulus, humans would recognize it as background noise rather than a valid signal. Motivated by this point, for a given pixel, the differential excitation component of the proposed Weber Local Descriptor (WLD) is computed based on the ratio between two terms: the relative intensity brightness differences of a current pixel against its neighbors (e.g., a 3×3 square box); and the brightness of the current pixel. With the differential excitation component, we attempt to extract the local salient patterns in the input image. In addition, we also compute the gradient orientation of the current pixel. That is, for each pixel of the input image, it is possible to compute two components of the WLD feature (i.e., differential excitation and gradient orientation). By combining the WLD feature per pixel, one can represent an input image (or a region) with a histogram, which is called WLD histogram hereinafter. In (Chen et al., 2010), the WLD features are computed pixel-wise as a dense descriptor.

5.2. Non-linear spectral feature description

A non-linear short-time Fourier transform analysis is based on the principle of stationary short-time signals as proposed in (Popescu et al., 2012). The method extracts six non-linear features: the first two features are based on statistical properties of the spectrum, and the next four ones are motivated from timbre features used for music genre classification. The computed results are: mean and variance (of the coefficients), spectral centroid in range and azimuth, and spectral flux in range and azimuth.

5.3. Gabor linear moments

A Gabor filter is a linear filter used in image processing (Manjunath & Ma, 1996). Frequency and orientation representations of a Gabor filter are similar to those of the human visual system, and were found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2-D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The Gabor filters are self-similar and can be generated from one mother wavelet by dilation and rotation. This implementation of the Gabor filter (Manjunath & Ma, 1996) convolves an image with a lattice of possibly overlapping banks of Gabor filters at different scales, orientations, and frequency. The scale is the scale of the Gaussian being used to compute the Gabor wavelet. The texture parameters are computed from the first-order statistics of the Gabor filter: mean and variance for different scales, orientations, and frequencies.

5.4. Adaptive Weber Local Descriptor (A-WLD)

An adaptive WLD is proposed for texture characterization, and contains two components: differential excitation and orientation. Based on these two terms, a joint histogram can be

constructed, followed by a conversion to a one-dimensional histogram, which is the WLD descriptor. The advantage of this feature is that it considers not only the local contrast but also the structure information represented by a gradient histogram. Although Weber's law is a theory from visual perception, and there is no visual criterion for SAR images, the principle of feature extraction is applicable to SAR images, too. However, in the case of SAR images, multiplicative speckle noise dramatically decreases its discriminative ability for image indexing. To combat speckle, a solution is to replace the gradient in WLD by the ratio of mean differences. Based on our newly adapted components (orientation and excitation), WLD for SAR images can be defined as a joint histogram (Cui et al., 2013). Following the same strategy as WLD, this joint histogram is converted to a one-dimensional histogram. The adapted WLD includes not only local statistics but also local structure information, resulting in an improved performance in SAR image indexing.

5.5. Multi-temporal similarity metric based on wavelet modeling

In the context of multi-temporal SAR image change detection for Earth monitoring applications, one critical issue is to generate accurate change maps. A common method to generate change maps is to apply a logarithm to the ratio image. However, due to the speckle effect and without taking contextual information into account, logarithms are usually not efficient for accurate change detection applications. In (Cui & Datcu, 2012), an unsupervised change detection method in the wavelet domain based on statistical wavelet sub-band modeling is proposed. The motivation is to capture textures efficiently in the wavelet domain. A wavelet transformation is applied to decompose an image into multiple scales. A probability density function of the coefficient magnitudes of each sub-band, assumed to be generalized Gaussian distributions (GGD), are obtained by fast parameter estimation.

A closed-form expression of Kullback-Leibler divergence between two corresponding sub-bands of the same scale is computed and used to generate the change map. This approach is comprehensively evaluated and compared using different parameter setting, scales, window sizes, and estimators. The proposed SAR change detection in the wavelet domain shows promising results as texture can then be better characterized than in the spatial domain. Through this study, the authors of (Cui & Datcu, 2012) conclude that the accuracy depends heavily on the estimation methods, although the model is also important.

5.6. Semi-automatic semantic annotation based on a support vector machine with relevance feedback (SVMRF)

Support Vector Machine (SVM) methods have been intensively used for classification and learning in Content-based Image Retrieval (CBIR) systems, with very good results during the last years. These methods are kernel-based learning machines and they have the capability of performing high-accuracy classifications for different types of data and with a small number of examples (patches) for each category. Of high importance in retrieval systems is the problem of Relevance Feedback (RF), which is a form of query-free retrieval (e.g., based on Euclidian distances (Costache & Datcu, 2007)), and which allows that images be retrieved according to a similarity measure and a given set of sample images. The goal of the system is to retrieve images based on the user's decision and on their interest. Thus, in the RF loop, the user labels the images as being either relevant or irrelevant. SVM is a tool that is well suited to implement

the RF method (Costache, Maitre, & Datcu, 2006). RF based on an SVM can be seen as a two-step loop process.

The first step is the learning step where, using a small number of examples (patches), the system learns how to separate data which belong to different categories. In the RF concept, there are only two categories: relevant samples, and irrelevant samples. In the second step, the classification of the remaining samples (patches) is made in order to decide which samples are relevant or irrelevant (Costache et al., 2006).

An SVM-RF tool is used in (Dumitru et al., 2015) to automatically retrieve the existing categories from a large dataset and to semantically annotate them. Before starting the annotation, in (Dumitru et al., 2015), an evaluation of the four feature extraction methods is performed. After selecting the ones with the best performance, based on the precision-recall metric, the tool is used to automatically retrieve the categories from the dataset. In order to annotate and find an appropriate semantic meaning for each retrieved category, human interaction is needed using a reference dataset (e.g., from Google Earth). Our dataset covers different areas of the Globe (e.g., 39 locations), where 320 categories were identified.

5.7. Cascaded Active Learning for Object Retrieval (CALOR)

CALOR (Blanchart et al., 2011a) is based on a hierarchical top-down processing scheme for object retrieval in high-volume high-resolution optical satellite repositories. It learns via a multistage active learning process, with a cascade of classifiers working each at a certain scale on a patch-based representation of an image. On each stage, we seek to eliminate large parts of images considered as being non-relevant; the purpose is to focus on the finest scales and on more promising and as spatially limited as possible areas. The scheme is based on the fact that by reducing the size of the analysis window (i.e., the size of the patch), we are able to better capture the properties of the targeted object. The cascaded hierarchy is introduced to compensate for the extra computational burden induced by diminishing the size of the patches, which causes an exponential growth of the number of patches to be processed. Blanchart (Blanchart, 2011) proposed a cascaded active learning strategy to build a classifier at each level of the hierarchy and we provided a new Multiple Instance Learning Algorithm to automatically propagate the training samples from one level of the hierarchy to the next one.

6. EOLib – functionality

In this section, the full functionality of the EOLib system will be presented together with examples. These are: Query by Meta-Data and Semantics, Content Based Image Retrieval, Visual Data Mining, Knowledge Discovery in Databases, and System Evaluation.

6.1. Querying meta-data and semantics

From the various query types that are implemented, we demonstrate three typical queries referring to meta-data and semantic annotation. The first group of examples shall find typical TerraSAR-X scenes.

Figure 9 shows the product quick-looks matching a criterion (bottom half) together with their meta-data:

- Incidence angle between the lower and upper bound of the specified range (see Figure 6)
- Ascending and descending orbit branch and looking direction (see Figure 7)
- Latitude and longitude between the range values of each continent (see Figure 1)
- Acquisition dates between 2007 and 2015 (see Figure 8).

A second example is to find all patches within our database that contain “Storage tanks” as a semantic label and to display them together with their meta-data. The GUI which offers such information is shown in Figure 9. The meta-data is displayed in the top-half part of the interface, while the visual representation is given in the bottom half. An element selection in either half will produce a highlight in the other half. A similar example can be to find “Skyscrapers” within the database (Figure 10) and to display the percentage of retrieved “Skyscrapers” for several given North American cities (e.g., Calgary, Ciudad Juarez, North San Diego, Poway, South San Diego, Sun Lakes, Tijuana, Tucson, Santa Clarita, Reno, Vancouver, Washington DC, and Ottawa) (see Figure 11).

6.2. Content-based image retrieval

We carried out some experiments using semantically annotated TerraSAR-X data. We searched our database for patches with different content and we compared the retrieved results with the reference data (the ground truth data). Finally, we computed the precision/recall metrics. Precision is defined as the fraction of the retrieved images which is relevant, while recall is defined as the fraction of relevant images which has been retrieved (Han, Kamber, & Pei, 2012). From our TerraSAR-X dataset, we extracted a subset that covers German cities such as: Bonn, Oldenburg, Munich, Kiel, Cologne, Berlin, Lindau, Karlsruhe, Mannheim, Stuttgart, and Garmisch-Partenkirchen. These images were divided into patches, and about 50 semantic categories were identified

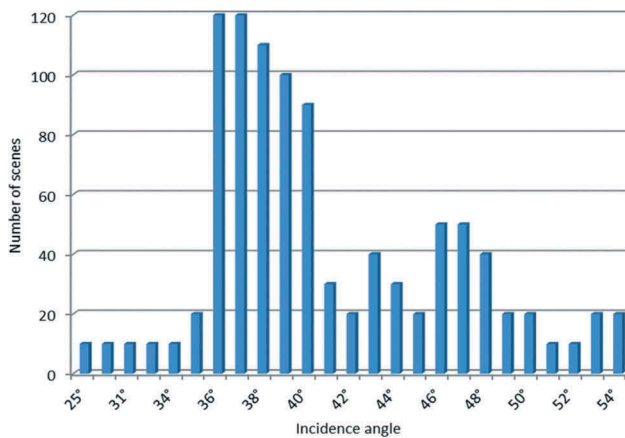


Figure 6. Number of scenes distributed between the lower and upper limits of the incidence angle.

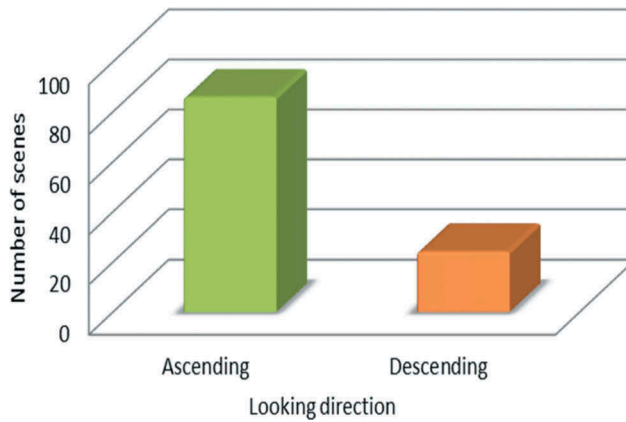


Figure 7. Percentage of scenes acquired along an ascending or descending orbit branch.

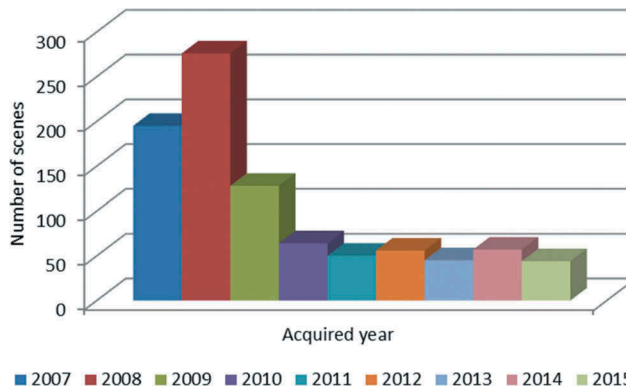


Figure 8. Number of scenes acquired from the TerraSAR-X archives and ingested in our database.

(e.g., “Agricultural land”, “Bridges”, “Channels”, “Broad-leaf forest”, “Mixed forest”, “High-density residential areas”, “Medium-density residential areas”, “Low-density residential areas”, “Mountains”, “Lakes”, “Roads”, “Railways”, “Rivers”, and “Storage tanks”). In [Tables 2 and 3](#), we show the performance metrics for the results obtained from two cities, namely Munich and Berlin, Germany using Gabor linear moments as a feature extraction algorithm. The average precision is 88% in the case of Munich, and 90% in the case of Berlin, while the recall amounts to 85% in the case of Munich, and 60% in the case of Berlin. The results of the CBIR are summarized in [Figure 12](#) for a number of important categories (e.g., “High-density residential areas”, “Channels”, “Broad-leaf forest”, “Sports terrain”, and “Railway tracks”) selected from the two cities.

6.3. Visual Data Mining

Visual Data Mining is a tool which allows the user to interactively visualize feature spaces of large repositories of images with a functionality supporting browsing, querying, and zooming (Espinoza-Molina, Datcu, Teleaga, & Balint, 2014). The

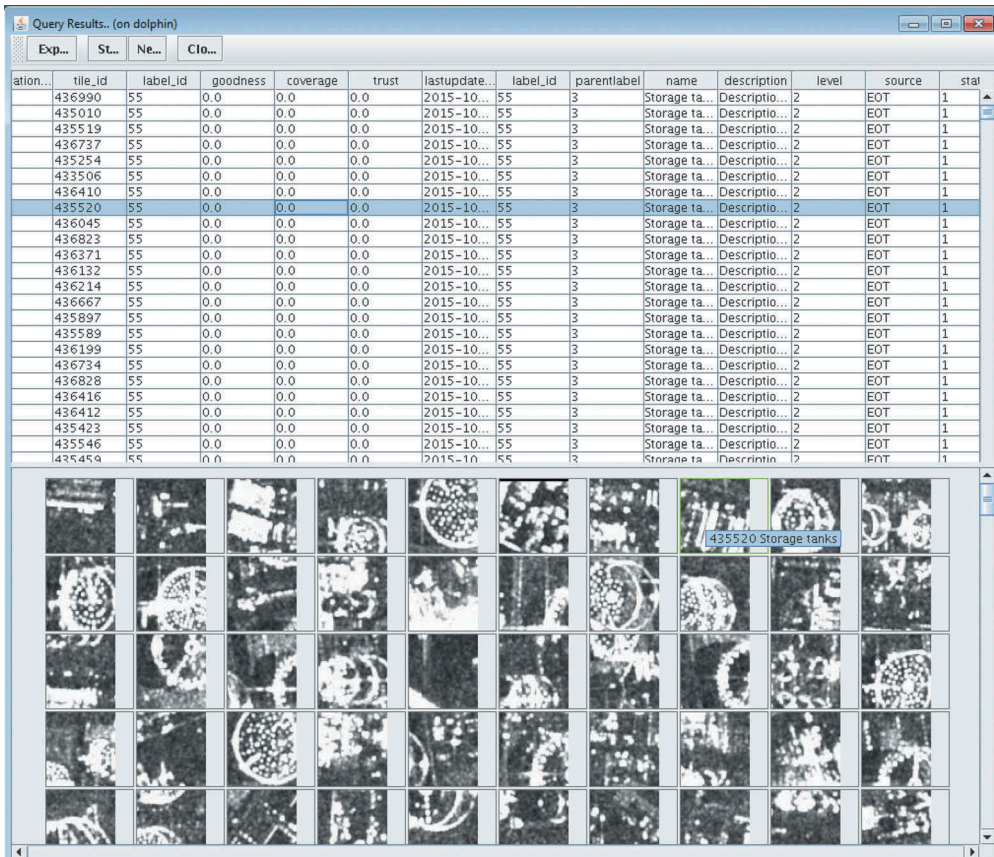


Figure 9. GUI interface used to display query results. It shows the product quick-looks matching this criterion (bottom-half) together with their meta-data (top-half).

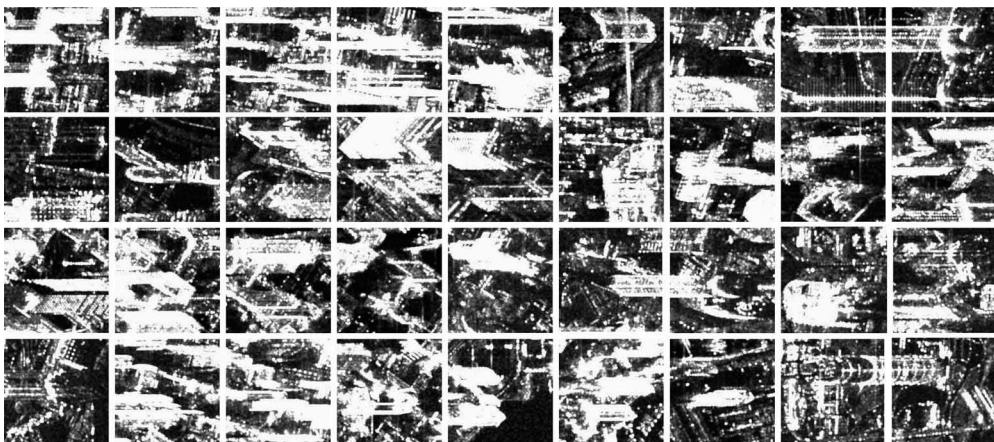


Figure 10. First patches of the category "Skyscrapers" retrieved into the database using a similar query criterion.

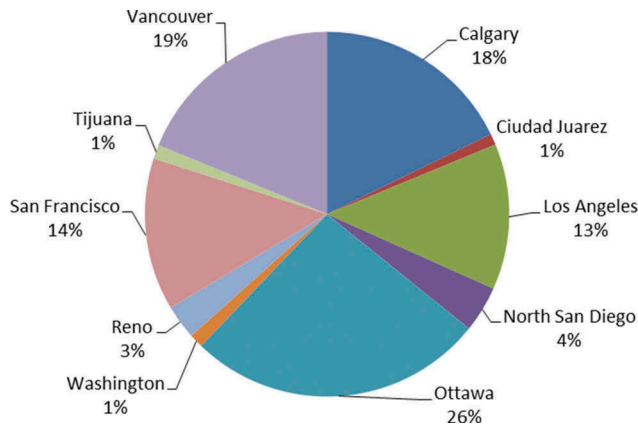


Figure 11. Percentage of "Skyscrapers" retrieved for several given North American cities.

Table 2. Precision/Recall for the TerraSAR-X Image of Munich.

Semantic categories	Precision (%)	Recall (%)
<i>Amusement parks</i>	100.00	100.00
<i>Mixed forest</i>	89.56	78.50
<i>Grassland</i>	100.00	100.00
<i>High-density residential areas</i>	82.35	80.00
<i>High-density residential areas and Roads</i>	87.90	85.00
<i>Medium-density residential areas</i>	84.62	84.62
<i>Railway tracks</i>	87.50	92.11
<i>Rivers</i>	82.92	78.96
<i>Bridges</i>	75.00	69.50

Table 3. Precision/Recall for the TerraSAR-X Image of Berlin.

Semantic categories	Precision (%)	Recall (%)
<i>Mixed forest</i>	100.00	78.57
<i>Channels</i>	94.34	67.57
<i>Railway tracks</i>	100.00	58.33
<i>High-density residential areas</i>	86.37	66.67
<i>Skyscrapers</i>	80.10	68.06
<i>Medium-density residential areas</i>	65.22	44.12
<i>Roads</i>	90.00	39.58
<i>Sports fields (terrain)</i>	100.00	55.56

loading time for more than 120,000 patches with a size of 160×160 pixels is less than two minutes, but it should be noted that by increasing the patch size, the loading time will increase. This tool allows interactive exploration and analysis of very large, high-complexity, and non-visual data sets stored within the database. It provides the end-user with an intuitive tool for Data Mining. By using a graphical interface, where the selection of different images and/or image content in 2-D or 3-D space is achieved through visualization techniques, data reduction methods, and similarity metrics to group the images, efficient data manipulation at a higher scale can be achieved. Samples can be annotated and stored into the database and can be augmented using Google Earth.

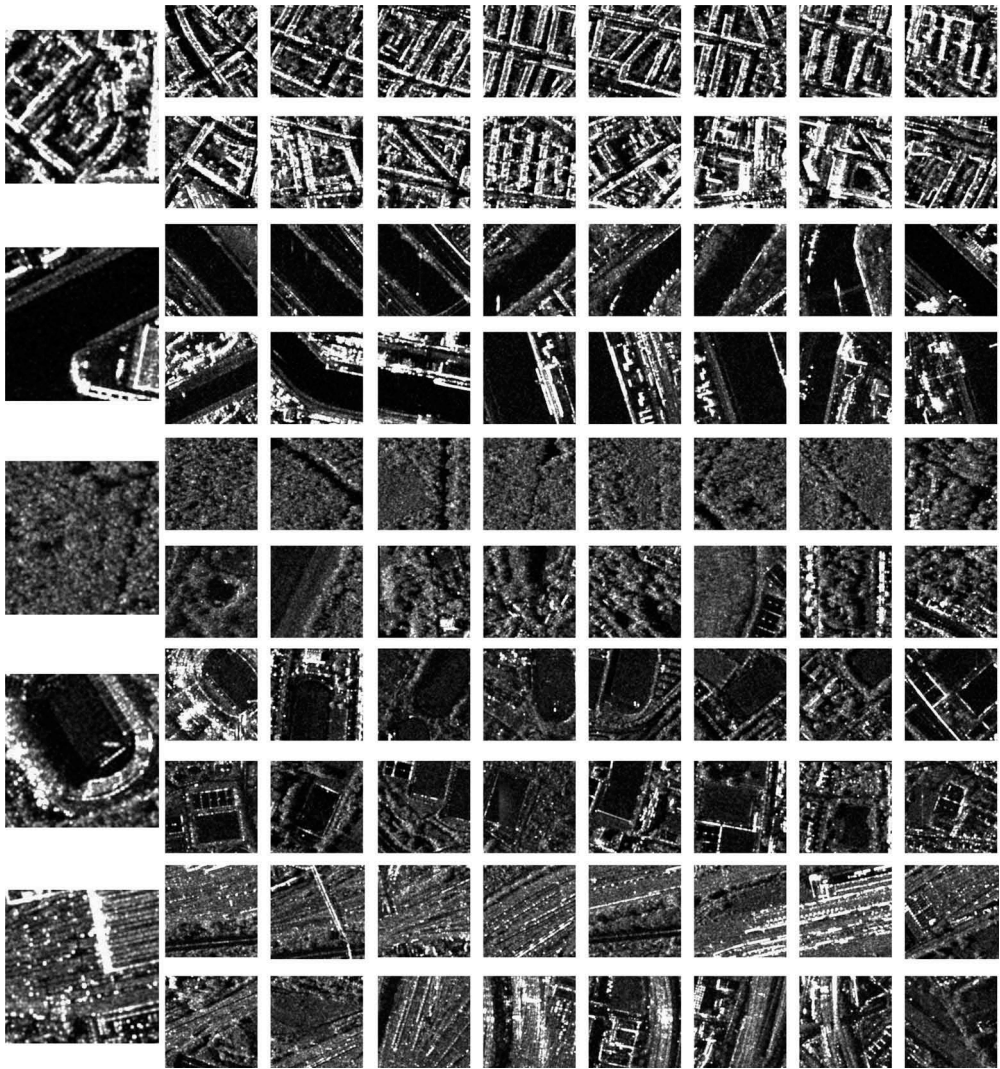


Figure 12. CBIR results for five semantic categories: “High-density residential areas”, “Channels”, “Broad-leaf forest”, “Sports terrain”, and “Railway tracks”. The left column shows the example, while the right side presents the retrieved patches similar to the provided query example.

An experiment was performed using TerraSAR-X scenes from Germany comprising a total of 7,000 patches. Each patch created a Gabor feature vector with 48 dimensions. Dimensionality reduction was applied and the feature vector of 48 dimensions was converted into a three-dimensional vector. Most of the patches are grouped into one big cluster.

In [Figure 13](#), we selected a group of patches to be exported to Google Earth. The next step is the mapping of the projected patches into a three-dimensional space, and an analysis of their content and possible outliers. The patches can also be annotated and exported to Google Earth in order to be validated by comparing them with terrain information

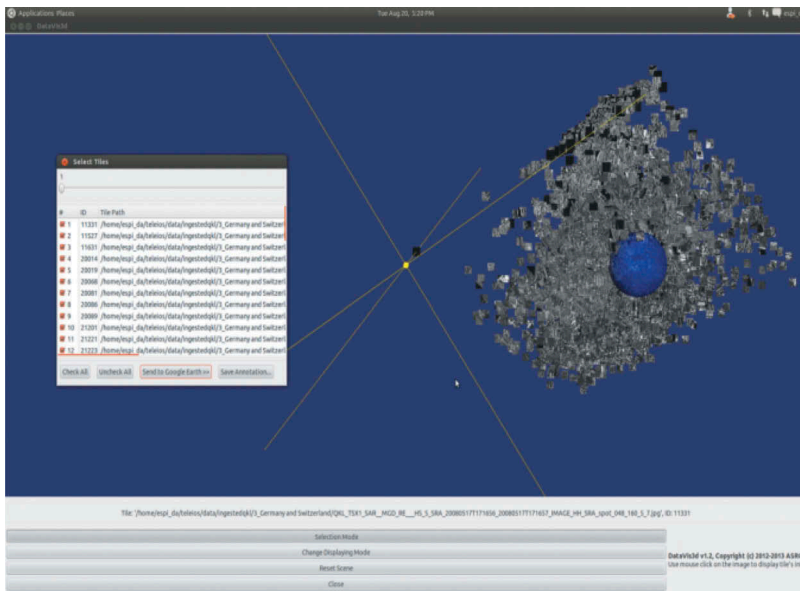


Figure 13. Selection of a group of patches for exporting them to Google Earth.

6.4. Knowledge discovery in databases

The KDD module is used to search for image content and to create semantic annotations of the ingested images. This implementation is based on the Cascaded Active Learning for Object Retrieval (CALOR) algorithm, which contains a Support Vector Machine (SVM) as active learning and relevance feedback method in order to include human expertise in the annotation. The definition of image semantics is achieved by using an interactive loop where human expertise is required to terminate the loop and to define the semantic category.

The idea behind active learning is that a machine learning algorithm can achieve higher accuracy with fewer training labels if it's allowed to choose the data from which it learns. The main functionality of the KDD module is two-fold:

- Exploration of the data set in order to see the evolution of the content of a category along the retrieved patches
- Semantic annotation of the dataset after the user has an idea about the content of the data set, and semantic meaning can be added to the data.

The experiments were performed on a dataset consisting of 200 images out of the 1,100 TerraSAR-X acquisitions which cover a wide variety of the Earth's surface. These images include categories such as: "Skyscrapers", "Storage tanks", "Ships/Boats", "Harbor infrastructure", "Airports", and "Inhabited built-up areas".

Once a user has an idea about the content of the data set, the user can semantically annotate the data with the explored categories. These semantic categories are stored into the DBMS and can be queried later. In [Figure 14](#), we show the steps that need to be followed by the user in order to make the annotation and save the information into the DBMS. In this example, 33 TerraSAR-X images have been used together with three panels.

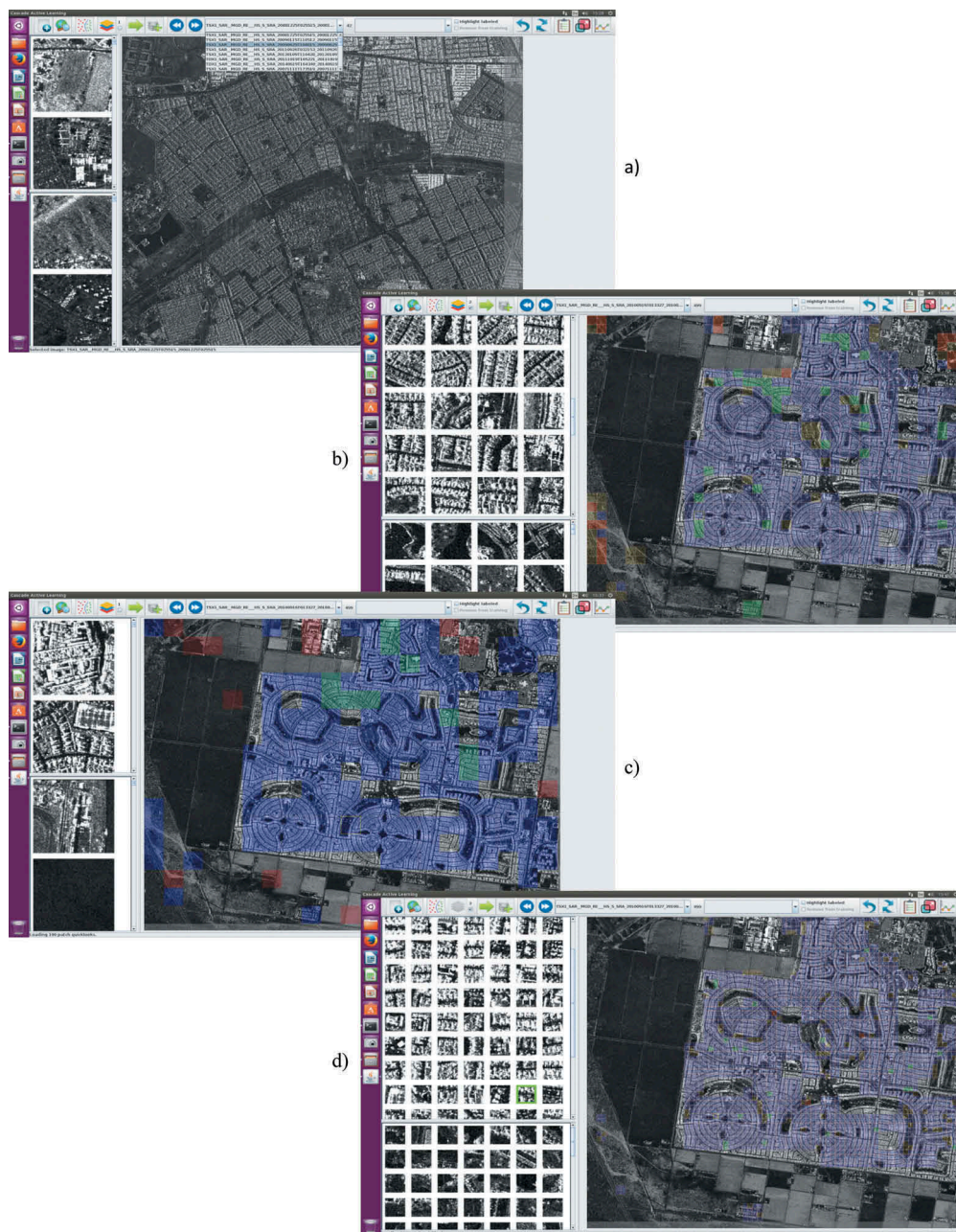


Figure 14. Process of semantic annotation of the category "Mixed urban areas" using the KDD module of EOLib. Figures (a) to (d) are from top to bottom. The positive and negative patches are marked in green and red, while the retrieved patches by the KDD are marked in blue. Figures (e) to (h) are from top to bottom. The positive and negative patches are marked in green and red. The retrieved patches by the KDD are marked in blue, while the patches stored from one level to another are marked with purple. Figure h) shows the outline of the patches in blue for "Mixed urban areas" category and in green for "Sand".

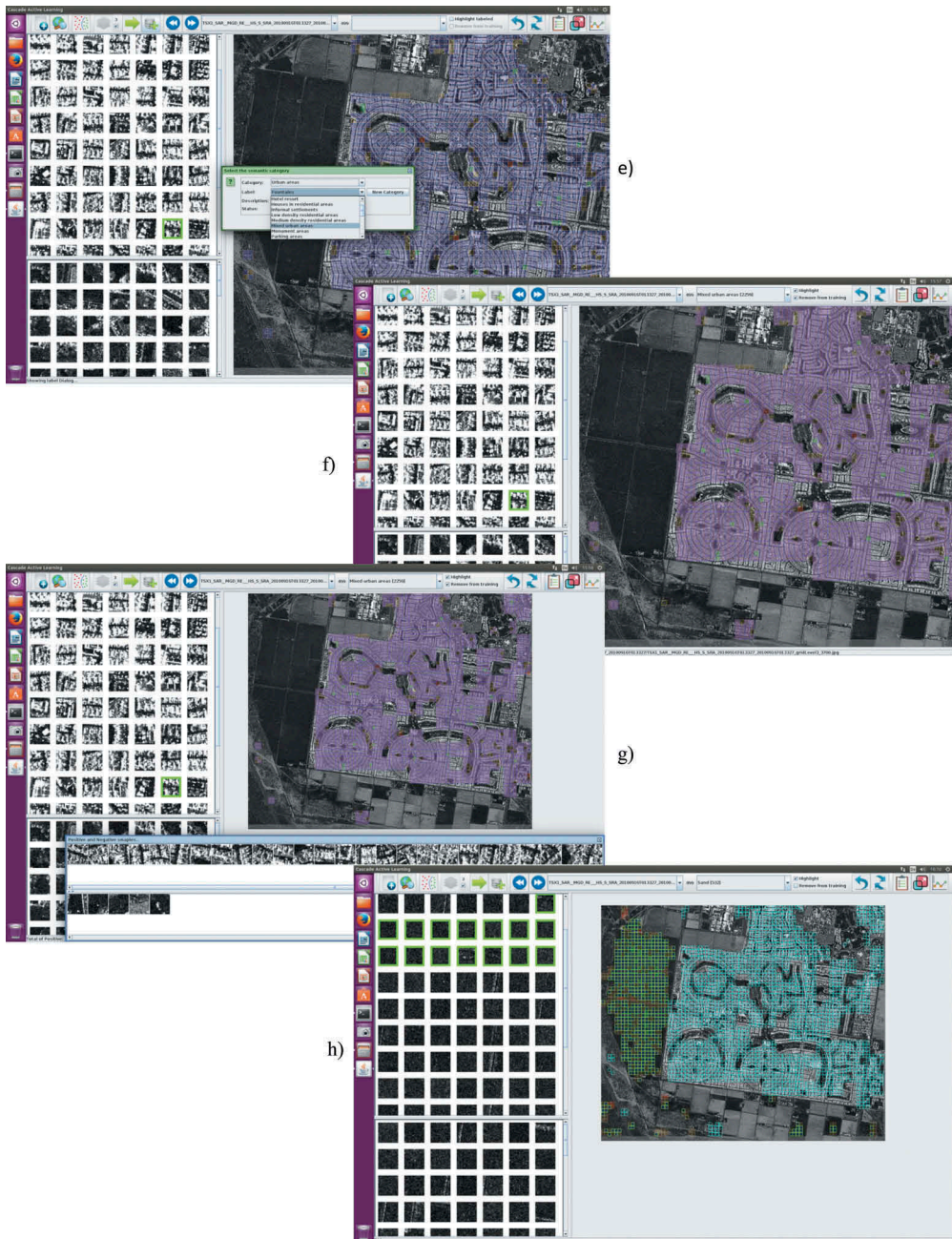


Figure 14. (Continued.)

In Figure 14(b) the GUI interface shows the results after the training, in which we are looking to find those patches with similar content. In this case, we are looking for the “Mixed urban areas” category. The retrieved patches in level 1 are marked in blue and projected on the image. Figure 14(c) shows the retrieved patches in level 2 and projected onto the image. Figure 14(d) shows the retrieved patches in level 3 and projected onto

the image. In [Figure 14\(e\)](#) the user can now semantically annotate the retrieved patches by choosing the correct label from an existing or user-defined list. For each patch, a label is allocated and saved into the database. [Figure 14\(f\)](#) presents the annotated area in reddish together with the label and the number of annotated patches with this label (see the GUI bar interface on its center-left side). In [Figure 14\(g\)](#) the user can see the positive and negative patches that he/she selected during the training (in the last level). The user can see the selected positive and negative patches during the training (in the last level). These are given in the new GUI interface on the bottom of the figure, where the first line shows the positive examples and the second line shows the negative examples selected by the user. [Figure 14\(h\)](#) shows the results of the “Mixed urban areas” category in blue and another category, namely “Sand” in green. The annotation can continue until the full image is annotated. Other labels to be retrieved can be “Ploughed agricultural land”, “Industrial buildings”, “Roads”, etc. while the sought-for category is “Mixed urban areas”. The GUI interface is composed of a top-left panel which shows the relevant retrieved patches, and a bottom-left panel which shows the irrelevant retrieved patches. The large right panel shows the image/images (one can see the list of images on the top-center GUI bar together with their TerraSAR-X unique id) that are being processed, benefiting from the zoom function. Through this panel, a user can see the distribution of the retrieved patches, and all the training samples. The users can also verify the selected training samples by checking the surrounding context or displaying it like in the bottom of [Figure 14\(g\)](#). Training samples can be selected from any of the three panels.

6.5. System evaluation

6.5.1. Evaluations based on performances

We evaluated the system using our TerraSAR-X data set (using the Gabor linear moments algorithm and the cascaded learning approach). The data were divided into data sub-sets, one for each continent. Here, we chose for demonstration our North American data set that contains 10 cities: Tijuana, Tucson, North and South San Diego, Poway, Sun Lakes, San Francisco, Ciudad Juarez, Calgary, and Ottawa.

We evaluated the performance of the KDD/cascaded learning algorithms using this data sub-set in order to verify the reduction of the computational effort. This approach helped reduce the number of training samples (i.e., patches) from one level to another level and finally, at the lowest level, to obtain an accurate semantic annotation. Based on this method, we can also generate a “bag of objects” for subsequent classification activities. [Table 4](#) contains for eight given general categories (Dumitru et al., 2016), for each level, the initial number of patches for each category and, after classification, the number of patches that contain this category being used further for the next level. Based on the cascaded learning algorithm, only the positive patches (this means the patches that contain the selected category we are looking for) are kept and used for annotation. [Table 5](#) shows the percentages of the data kept from the entire data set. For example, for the “Water bodies” category, we are using only 11.99% from the entire amount of patches, while the rest of the patches are assigned to other categories. These “positive” patches are split again, classified, and the residues that do not belong to the desired category are removed (we keep 65.57% of all patches). On the last level, we repeated the previous procedure and we were finally annotating 94.17% of the patches with the category we are

Table 4. Performance evaluation for ten American cities.

Semantic category	Level 1 (3,627 patches)		Level 2 (14,508 patches)		Level 3 (58,032 patches)	
	Total no. of patches	Positive patches	Total no. of patches	Positive patches	Total no. of patches	Positive patches
<i>Water bodies</i>	3,627	435	435*4 = 1,740	1,141	1141*4 = 4,564	4,298
<i>Natural vegetation</i>	3,627	59	59*4 = 236	135	135*4 = 540	495
<i>Settlements</i>	3,627	2,318	2318*4 = 9,272	7,536	7536*4 = 30,144	26,241
<i>Bare ground</i>	3,627	423	423*4 = 1,692	952	952*4 = 3,808	3,515
<i>Transport</i>	3,627	180	180*4 = 732	418	418*4 = 1,672	1,499
<i>Agriculture</i>	3,627	0	0	0	0	0
<i>Military facilities</i>	3,627	464	464*4 = 1,856	1,123	1123*4 = 4,492	3,157
<i>Industrial production areas</i>	3,627	295	295*4 = 1,180	792	792*4 = 3,168	1,578

Table 5. Percentage of categories preserved after each level (see positive patches of Table 4).

Semantic category	Level 1	Level 2	Level 3
<i>Water bodies</i>	11.99%	65.57%	94.17%
<i>Natural vegetation</i>	1.63%	57.2%	91.67%
<i>Settlements</i>	63.91%	81.28%	87.05%
<i>Bare ground</i>	11.66%	56.26%	92.31%
<i>Transport</i>	4.96%	57.1%	89.65%
<i>Military facilities</i>	0.00%	0.00%	0.00%
<i>Agriculture</i>	12.79%	60.51%	70.28%
<i>Industrial production areas</i>	8.13%	67.12%	49.81%

looking for. This procedure was repeated for all categories (Dumitru, Schwarz, & Datcu, 2018a). Similar results were obtained for all other data sub-sets.

The performances of the EOLib KDD annotation tool is demonstrated, in Figure 15 above, using five metrics, namely Precision, Recall, Accuracy, F-measure, and Specificity. The definition of each evaluation metric is described in detail in (Manning, Raghavan, &

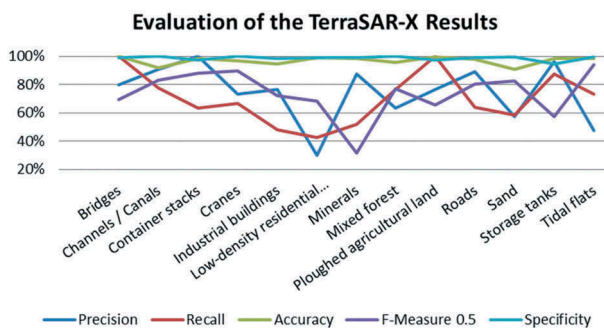
**Figure 15.** Attainable classification evaluation metrics (Precision, Recall, Accuracy, F-Measure 0.5, and Specificity) obtained by Gabor feature vectors for all retrieved categories.

Table 6. Evaluation Metrics of our TerraSAR-X Data.

Semantic categories	Precision	Recall	Accuracy	F-Measure			Fall-out	Specificity
				0.5	1	2		
<i>Boats</i>	0.70	0.65	0.98	0.69	0.68	0.66	0.01	0.99
<i>Bridges</i>	0.80	1.00	0.99	0.83	0.89	0.95	0.00	0.99
<i>Channels/Canals</i>	0.91	0.78	0.92	0.88	0.84	0.80	0.03	0.97
<i>Container stacks</i>	1.00	0.64	0.99	0.90	0.78	0.69	0.00	1.00
<i>Cranes</i>	0.73	0.67	0.97	0.72	0.70	0.68	0.02	0.98
<i>Industrial buildings</i>	0.77	0.48	0.94	0.68	0.59	0.52	0.01	0.99
<i>Low-density residential areas</i>	0.30	0.43	0.99	0.32	0.35	0.39	0.01	0.99
<i>Minerals</i>	0.88	0.52	0.98	0.77	0.65	0.56	0.00	0.99
<i>Mixed forest</i>	0.63	0.77	0.9	0.66	0.69	0.74	0.03	0.97
<i>Ploughed agricultural land</i>	0.77	1.00	0.99	0.80	0.87	0.94	0.01	0.99
<i>Roads</i>	0.89	0.64	0.98	0.83	0.75	0.68	0.00	0.99
<i>Sand</i>	0.57	0.58	0.90	0.57	0.58	0.58	0.05	0.95
<i>Storage tanks</i>	0.96	0.87	0.99	0.94	0.91	0.89	0.00	0.99
<i>Tidal flats</i>	0.48	0.73	0.98	0.51	0.58	0.66	0.01	0.99

Schutze, 2008) and (Powers, 2011). For TerraSAR-X, we chose images for which the ground truth already existed (Dumitru, Schwarz, & Datcu, 2018b) and then we computed the evaluation metrics. The primitive features used to characterize the image patches (with a size of 160×160 pixels) were Gabor filters with 5 scales and 6 orientations. Table 6 summarizes the results obtained for EOLib.

Figure 14: Process of semantic annotation of the category “Mixed urban areas” using the KDD module of EOLib. Figures (e) to (h) are from top to bottom. The positive and negative patches are marked in green and red. The retrieved patches by the KDD are marked in blue, while the patches stored from one level to another are marked with purple. Figure h) shows the outline of the patches in blue for “Mixed urban areas” category and in green for “Sand”.

6.5.2. Evaluations based on user feedback

In 2016, an EOLib User Workshop was organized by the German Aerospace Center (DLR) in order to conclude the EOLib project, to evaluate the latest version of the EOLib system, and to collect user feedback and future requirements for the validation of the system. This workshop was attended by users from the European Space Agency, Italy; Airbus Defence and Space, Germany; the National Research Council, Italy; DEIMOS Engenharia, Portugal; Thales Alenia Space, France; and the DLR, Germany. It proposed applications based on TerraSAR-X and other satellite data (e.g., Copernicus).

The first part of the User Workshop was a presentation of the EOLib system, where the users could interactively ask questions and make observations and comments. In the second part, the overall functionalities of the system was demonstrated in order to show the users how the system looks like, and the functions of the different components. The third part was training and a hands-on session.

During the hands-on interactive session, the users could familiarize with the system to define some semantic categories, and to annotate several the images. Here, Knowledge Discovery in Databases and Query by Meta-Data and Semantics having been tested.

In the end, a round table discussion was organized where the users exposed their opinions, comments, and suggestions about the evaluated system. The inputs provided by the users during the user workshop told us that the users were very satisfied with the system and its offered functionality, its response times, and the obtained results. The users

also proposed a number of additional applications such as ship detection, coastline detection, rapid mapping, etc. They also addressed a change detection scenario highlighting the locations where changes occurred by using the query component of the system. These aspects were already taken into account in Dumitru et al. (2016, 2018a), Dumitru, Cui, Faur, and Datcu (2015), etc.

The users also suggested to add additional components: the integration of evaluation metrics, and the generation of automated classification maps and statistical analytics. All these observations were taken into account and integrated with the latest version of the system described in (H2020 Candela project, 2019).

7. Conclusions

In this paper, we introduced the modular EOLib software framework which is capable to perform data mining and knowledge discovery within the TerraSAR-X Payload Ground Segment of the German Aerospace Center. EOLib can serve as a model for the next generation of Image Information Mining systems. The main goal of EOLib is to create a domain-invariant data mining solution, directly attached to the Payload Ground Segment. It supplies the end-user content-enriched imagery, in the form of taxonomies and meta-data.

Each module of EOLib can be used as a standalone application. We described the main components of the modular EOLib framework covering functions such as ingestion and feature extraction from EO imagery, meta-data extraction, semantic definition of the image content based on machine learning and data mining methods, advanced queries of the image archives utilizing content, meta-data and semantic categories, as well as 3-D visualizations of our huge and complex image archives.

EOLib is interfaced and operated in the Multi-Mission Payload Ground Segment (PGS) of DLR's Remote Sensing Data Center. The dataset size shows that the presented software framework is capable of handling the 3 Vs of big data.

In general, the performances of EOLib modules are reliable, however, some water-related categories remain difficult to classify for the KDD. This can be overcome by implementing new feature extraction algorithms and classifiers which better suit this specific class.

The experiments show that for TerraSAR-X images the best feature extraction method is Gabor Linear Moments (e.g., with 5 scales and 6 orientations) for man-made infrastructure categories such as urban and industrial areas, transportation, while for natural categories such as agriculture, forest, vegetation the Adaptive Weber Local Descriptor is the best performer (e.g., with 8 orientations and 18 excitation levels). A comparison between different feature extraction methods is already described in Dumitru et al. (2018b), Dumitru and Datcu (2013), and Cui, Dumitru, and Datcu (2014). For multi-spectral images the best feature extraction method is the Weber Local Descriptor (e.g., with 8 orientations and 18 excitation levels). For EOLib, the optimal classifier is based on Cascaded Active Learning (Blanchart et al., 2011a).

An extension of EOLib aims at the development of new algorithms and a validation of the system using Copernicus and selected third-party mission data, the demonstration of leading edge concepts, and methods for information content exploration and utilization for Earth observation data, mainly for Sentinel-1, Sentinel-2 but also other third-party Earth observation image data. This new version is used and validated (Dumitru, Schwarz, Castel, Lorenzo, & Datcu, 2019) in the CANDELA project (H2020 Candela project, 2019).

Acknowledgments

We thank DLR's TerraSAR-X Science Service System for providing the SAR images (Proposal MTH1118). The work was supported by EOLib—an ESA technological project ESA EOLib project, 2019.

Data availability statement

Data not available due to TerraSAR-X data policy restrictions. The distribution of the TerraSAR-X data should follow the German legislation and is based on data proposals (for more information see TerraSAR-X (2019)).

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

- Blanchart, P., "Fast learning methods adapted to the user specificities: Application to Earth observation image information mining," in Paris: Ecole Doctorale d'Informatique, Telecommunications et Electronique de Paris and Telecom ParisTech, 2011.
- Blanchart, P., Ferecatu, M., Cui, S., & Datcu, M. (2014, April). Pattern retrieval in large image databases using multiscale coarse-to-fine cascaded active learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(4), 1127–1141.
- Blanchart, P., Ferecatu, M., & Datcu, M. (2011a, September). Cascaded active learning for object retrieval using multiscale coarse to fine analysis. In *2011 18th IEEE International Conference on Image Processing* (pp. 2793–2796), Brussels, Belgium.
- Blanchart, P., Ferecatu, M., & Datcu, M. (2011b, April). *Indexation of large satellite image repositories using small training sets*. ESA-EUSC-JRC 2011 Conference (pp. 37–40), Ispra, Italy.
- Chang, C. C., & Lin, C. J. (2011, April). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27.
- Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X., & Gao, W. (2010, September). WLD: A robust local image descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9), 1705–1720.
- Costache, M., & Datcu, M. (2007, July). *Learning - unlearning for mining high resolution EO images*. IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 4761–4764), Barcelona, Spain.
- Costache, M., Maitre, H., & Datcu, M. (2006, July) *Categorization based relevance feedback search engine for Earth observation images repositories*. IEEE International Symposium on Geoscience and Remote Sensing (IGARSS) (pp. 13–16), Denver, USA.
- Cui, S., & Datcu, M. (2012, August). Statistical wavelet sub-band modeling for multi-temporal SAR change detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4), 1095–1109.
- Cui, S., Dumitru, C. O., & Datcu, M. (2013, September). Ratio-detector-based feature extraction for very high resolution SAR image patch indexing. *IEEE Geoscience and Remote Sensing Letters*, 10(5), 1175–1179.
- Cui, S., Dumitru, C. O., & Datcu, M. (2014). Semantic annotation in Earth Observation based on active learning. *International Journal of Image and Data Fusion*, 5(2), 152–174.
- Data and information access services (DIAS). 2019 [Online]. Retrieved from <https://www.copernicus.eu/en/access-data/dias>
- Datcu, M., Daschiel, H., Pelizzari, A., Quartulli, M., Galoppo, A., Colapicchioni, A., ... D'Elia, S. (2003, December). Information mining in remote sensing image archives: System concepts. *IEEE Transactions on Geoscience and Remote Sensing*, 41(12), 2923–2936.

- Dumitru, C. O., Cui, S., Faur, D., & Datcu, M. (2015). Data analytics for rapid mapping: Case study of a flooding event in Germany and the Tsunami in Japan using very high resolution SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1), 114–129.
- Dumitru, C. O., Cui, S., Schwarz, G., & Datcu, M. (2015, April). Information content of very-high-resolution SAR images: Semantics, geospatial context, and ontologies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(4), 1635–1650.
- Dumitru, C. O., & Datcu, M. (2013, August). Information content of very high resolution SAR images: Study of feature extraction and imaging parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 51(8), 4591–4610.
- Dumitru, C. O., Schwarz, G., Castel, F., Lorenzo, J., & Datcu, M. (2019, September 5). Artificial intelligence data science methodology for Earth Observation. *Advanced Analytics and Artificial Intelligence Applications*, IntechOpen Press, (online), pp. 1–20.
- Dumitru, C. O., Schwarz, G., & Datcu, M. (2016, June). Land cover semantic annotation derived from high-resolution SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(6), 2215–2232.
- Dumitru, C. O., Schwarz, G., & Datcu, M. (2018a, July). *Evaluation of retrieved categories from a TerraSAR-X benchmarking data set*. IEEE International Symposium on Geoscience and Remote Sensing (IGARSS) (pp. 1–4), Valencia, Spain.
- Dumitru, C. O., Schwarz, G., & Datcu, M. (2018b, May). SAR image land cover datasets for classification benchmarking of temporal changes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1571–1592.
- EOLib Architecture concept. (2019). [Online]. Retrieved from <http://wiki.services.eoportal.org/tiki-index.php?page=EOLib>
- ESA EOLib project. (2019). [Online]. Retrieved from <http://wiki.services.eoportal.org/tiki-index.php?page=EOLIB+Project>
- Espinoza Molina, D., & Datcu, M. (2016, October). The Earth-Observation epitome: A new interactive value-added product. *IEEE Geoscience and Remote Sensing Letters*, 13(10), 1438–1442.
- Espinoza-Molina, D., & Datcu, M. (2013, November). Earth-observation image retrieval based on content, semantics, and metadata. *IEEE Transactions on Geoscience and Remote Sensing*, 51(11), 5145–5159.
- Espinoza-Molina, D., Datcu, M., & Balint, C. (2014). *Application of visual data mining for Earth observation use cases*. ESA-EUSC-JRC-2014 Image Information Mining Conference: The Sentinels Era (pp. 111–114), Bucharest, Romania.
- Espinoza-Molina, D., Datcu, M., Teleaga, D., & Balint, C. (2014, March). *Application of visual data mining for Earth Observation use cases*. The 9th Conference on Image Information Mining Conference: The Sentinels Era (pp. 111–114), Bucharest, Romania.
- Gueguen, L. (2015, April). Classifying compound structures in satellite images: A compressed representation for fast queries. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4), 1803–1818.
- H2020 Candela project. (2019). [Online]. Retrieved from <http://www.candela-h2020.eu/>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining concepts and techniques*. Waltham, MA: Morgan Kaufmann.
- Haralick, R. M., & Kelly, G. L. (1969, April). Pattern recognition with measurement space and spatial clustering for multiple images. *Proceedings of the IEEE*, 57(4), 654–665.
- Laney, D. (2001, February). 3D data management: Controlling data volume, velocity, and variety. *Application Delivery Strategies*.
- Manjunath, B. S., & Ma, W. Y. (1996, August). Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8), 837–842.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. New York, NY: Cambridge University Press.
- MonetDB. (2019). [Online]. Retrieved from <https://www.monetdb.org/>
- Pesaresi, M., & Benediktsson, J. A. (2001, February). A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(2), 309–320.

- Popescu, A. A., Gavati, I., & Datcu, M. (2012, January). Contextual descriptors for scene classes in very high-resolution SAR images. *IEEE Geoscience and Remote Sensing Letters*, 9(1), 80–84.
- Powers, D. M. W. (2011). Evaluation: From precision, recall and f-measure to ROC. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Rasiwasia, N., Moreno, P. J., & Vasconcelos, N. (2007, August). Bridging the gap: Query by semantic example. *IEEE Transactions on Multimedia*, 9(5), 923–938.
- Rathore, U., Paul, A., Ahmad, A., Chen, B. W., Huang, B., & Ji, W. (2015, October). Real-time big data analytical architecture for remote sensing application. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(10), 4610–4621.
- Salembier, P., Oliveras, A., & Garrido, L. (1998, April). Anti-extensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, 7(4), 555–570.
- Sentinel application platform (SNAP). (2019). [Online]. Retrieved from <http://step.esa.int/main/toolboxes/snap/>
- Sheikholeslami, G., Chang, W., & Zhang, A. (2002, September). Semquery: Semantic clustering and querying on heterogeneous features for visual data. *IEEE Transactions on Knowledge and Data Engineering*, 14(5), 988–1002.
- Shyu, C. R., Klaric, M., Scott, G. J., Barb, A. S., Davis, C. H., & Palaniappan, K. (2007, April). GeoIRIS: Geospatial information retrieval and indexing system - content mining, semantics modeling, and complex queries. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4), 839–852.
- Singh, J., & Datcu, M. (2013, December). SAR image categorization with log cumulants of the fractional Fourier transform coefficients. *IEEE Transactions on Geoscience and Remote Sensing*, 51(12), 5273–5282.
- TerraSAR-X (2013). TerraSAR-X basic product specification document, issue 1.9. Retrieved from <http://sss.terrasar-x.dlr.de/pdfs/TX-GS-DD-3302.pdf>
- TerraSAR-X (2019). [Online]. Retrieved from <https://sss.terrasar-x.dlr.de/>
- Thematic exploitation platform (TEP). (2019). [Online]. Retrieved from <https://tep.eo.esa.int/>
- van der Maaten, L., & Hinton, G. (2008, November). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.
- Wang, M., & Song, T. (2013, May). Remote sensing image retrieval by scene semantic matching. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 2874–2886.
- Wolfmuller, M., Dietrich, D., Sireteanu, E., Kiemle, S., Mikusch, E., & Bottcher, M. (2009, January). Data flow and workflow organization - the data management for the TerraSAR-X payload ground segment. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1), 44–50.
- Yang, Y., & Newsam, S. (2010, November 3–5). *Bag-of-visual-words and spatial extensions for land-use classification*. 18th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, San Jose, CA.