A knowledge organization system for image classification and retrieval in petroleum exploration domain

Mara Abel[1], Eduardo S. L. Gastal[1], Cassiana Michelin[2], Luiza Gonçalves Maggi[2], Bruno Eduardo Firnkes[1], Felix Eduardo Huaroto Pachas[1], Renata dos Santos Alvarenga[2]

[1]Informatics Institute, [2]Geosciences Institute - Universidade Federal do Rio Grande do Sul (UFRGS) PO 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{marabel,eslgastal,bruno.firnkes}@inf.ufrgs.br, cassiana.michelin@ufrgs.br

Abstract. *This paper describes a knowledge organization scheme for types of images in the domain of petroleum exploration based on ontological criteria. The classification separates the characteristics of representation, visualization, and storage from the semantics of the content, where each of these features has its proper organization system. The representation and visualization classes optimize the effort of image annotation by grouping the instances of images according to a set of criteria described in this paper, which are more directly identified by automatic classification methods. These classes keep a relationship with the geological entities depicted in the images. In this way, we can separate training methods for identifying the type of representation and for those that learn about what geological entity this representation is. In this ongoing project, we apply the knowledge organization for information retrieval over a set of 1927 images related to petroleum exploration. Preliminary results show good accuracy in simple classification tasks and indicate the need for improved classifiers in complex tasks, where the proposed ontological system will be fundamental for organizing the image datasets.*

1. Introduction

The recent growth in technologies for image creation and processing has greatly simplified the production of visual records and pictorial representation of pieces of geological evidence that supports exploration activities. This evolution results in a growing amount of digital visual content available in unstructured and non-indexed forms on data repositories that support the interpretation and decision taking by geologists and reservoir engineers. In the production of geological models, the geologist needs to access the previously produced maps, photographs, and many types of graphical visualization of analogic measures in order to support and explain her/his process of interpretation. The selection of relevant material should consider their type, scale of analysis, generator activity, and semantic content. However, even in the petroleum industry, the evolution of the image organization systems have not provided the adequate conceptual tools to help in the development of the image-retrieval computing systems to cope with the variety of types of geological visual content. The machine-learning techniques did lack a supporting classification system that organizes the sets of images and helps in reducing the number of samples required for learning these types.

An image classification system fitted to geological content would improve the nowadays techniques of image indexing based on manual annotation or content-based retrieval. The manual annotation approach associates keywords to the image in the time it is stored, and it shows the best results for semantic retrieval. However, the approach has intrinsic limitations. The first is that human annotation is subjective, so it does not offer a homogeneous classification system for indexing that allows the user to recover similar images. The images in Figure 1 are both about normal geological faults. The Figure 1(a) is a 3D diagram block that schematically represents a geological normal fault, while the Figure 1(b) is a picture of a geological normal fault in Thingvellir rift, Iceland. A person can annotate the Figure 1(a) as *3D diagram* (a type of visualization) or *geological normal fault* (the represented entity) or a divergent zone (an interpretation about the cause of this geological occurrence). Figure 1(b) can be annotated as a *photography*, a *landscape* or a *geological normal fault*, a *lake*, etc. When very large sets of images (in the order of millions of images) are available for training, linear scalable machine learning algorithms shows some impressive results, as the Google image retrieval system [G. Chechik 2009]. However, in expertise interpretation, the number of collected images is low considering the variety of represented entities and, moreover, the user wishes to retrieve visual objects in a more abstract level that has no immediate translation for physical attributes of the image [Abel, Mastella, et al. 2004].

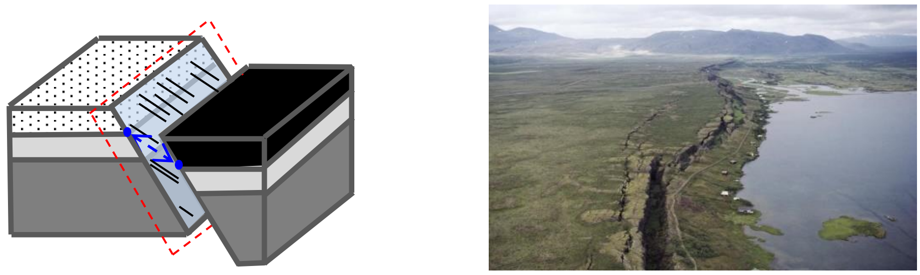

Figure 1 (a) A 3D diagram block depicting a normal fault. (b) A photograph that shows a normal fault (picture of National Geographic Society).

Our classification system offers to the person the possibility of classifying the types of visual content in separate of the visual content itself. We aim to provide a uniform way of annotating images and producing a previous classification of the visual content based on image features that allow the development of fitted algorithms of image processing and machine learning that can learn over a small set of images and figures. The approach will be further used to retrieve the large set of unlabeled images stored in corporate databases.

The further sections of this paper describe, in section 2, a review of related works in ontology-based image indexing, while section 3 reviews the conceptual basis for modeling of visual content. Section 4 presents the basis of our classification system and details the main classes and dimensions utilized for indexing. Finally, section 5 describes the preliminary results of automatic image classification that applies our system, ending with conclusions and next planned steps.

2. Related work on ontology-base retrieval systems

Several recent works are exploring the use of ontologies to help the visual content extraction in information retrieval for imagistic domains. Pandey and colleagues [Pandey, Khanna et al. 2016] offer a relevant review on techniques for combining image extraction and semantic processing to deal with the user intention in recovering visual content. Also, Zin and colleagues present a systematic review on content-based image retrieval specifically for the medical domain [Mohd Zin, Yusof et al. 2018]. According to them, a content-based image retrieval system applies one of the following approaches to deal with the semantic gap between image and meaning of content: (1) a domain ontology that reduces the search space, (2) machine learning algorithms, for large databases with the uniform type of images; (3) relevance feedback of user; (4) semantic template generation; (5) combined textual and visual content of images. According to them, the combination of image extraction and conceptual models is the most challenging approach nowadays and strongly depends on the quality of metadata. The works of [Tian 2016; Chen and Chen 2017; Gonçalves, Guilherme et al. 2018; Kuang, Yu et al. 2018] apply the merge of techniques. We understand that for restricted knowledge-intensive domains, such as petroleum geology, the support of separate domain ontologies and image feature ontologies may increase the accuracy and relevance of the retrieved content. The followed works also apply this approach.

Sharm and Siddiqui [Sharma and Siddiqui 2016] describe an ontology-based framework for retrieval of museum artifacts. As we do, the authors propose a domain ontology with the representation and visualization aspects of the domain - *analysis ontology* – in separate of the domain ontology that defines the museum artifacts – *domain ontology*. The approach starts by segmenting the input image and extracting low-level descriptors of the segments and their relation to the concepts of domain ontology. The researchers have tested the approach with 1200 images from 11 categories. The images were pre-processed to uniformize their type, size, contrast and noise content.

Santos Neto [Neto 2013] proposes the ontology ONTOLIME to support information retrieval with medical images. As we propose in our approach, the author models the physical aspects of the images, such as technical capturing process, color, and texture, as subclasses aspects of the *Image* concept, while model medical concepts (such as *Anatomy*) has its proper subclasses related to the knowledge domain.

Even though there is a large effort in developing domain ontologies for the subdomains of petroleum exploration and the importance of visual content in this domain, few works were published reporting the use of Geology ontologies for image retrieval. In her dissertation [Barreiros 2010], Barreiros proposes an ontological model for outcrop description that relates the image content produced by a geologist in the fieldwork. However, the property and types of images are not detailed in the ontology, and neither are the methods used for image retrieval. More relevant is the effort of industry and organizations in producing image-based retrieval systems for geology. Endeeper has produced the portal PetrographicPedia[1] for image retrieval of petrographic images [Castro 2012] and the content-based image indexing system RockViewer[2] using the Petrography ontology of Petroledge [Abel, Goldberg et al. 2013]. In both systems, the geologist describes the image in the moment of capturing using a controlled vocabulary based on the ontology. The CPRM (Geological Service of Brazil) offers access to a large set of geographical maps and geological images through its exploration and production database BDEP, where a local thesaurus of geosciences [Nascimento and Freire 2005] supports the content organization. In addition, the C&C Reservoir Company[3] keeps a large database of classified images, based in an organization system that mixes content, interpretation

[1] www.endeeper.com/petrography-pedia-en-us/Main_Page

[2] www.endeeper.com/products

[3] www.ccreservoirs.com

and image artifact types.

3. Previous work on conceptual modeling of visual content entities

Even being intrinsically connected, images and visual content are disjoint concepts. In order to understand the distinction, we refer to the work of Lorenzatti, which proposed a distinction on existent objects, concepts, representations and visualizations [Lorenzatti, Abel et al. 2009]. An *object* is supposed to be an existent entity in reality, which can be an abstract entity (such as an *emotion*, or the number *5*), a social entity (an *enterprise*) or a concrete instance (a *dog*). A *concept*, otherwise, is a mental abstraction of a portion of reality, emphasizing the aspects of entities that are the interest of the human observer. A *representation* is one of the many possibilities in wha person externalize the concepts to share her/his conceptualization among the community. A representation serves to the communication process between a sender, the interpreter of reality, and a receiver that is part of the community. Finally, *visualization* refers to the process of creating a pictorial expression of the concept to help the receptor to gain insight or understanding of the sender.

Moreover, the work of [Perrin, Rainaud et al. 2013] defines, as shown in Table 1, the meaning of *model*, *representation*, and *visualization* in the context of petroleum geological modeling. We consider these meta-types of information to propose our approach and separate the modeling of geological objects from those of its representation and visualization artifacts. Representations and visualization are informational entities.

In YAMATO, [Mizoguchi and Toyoshima 2006] consider that a representation artifact is composed of (representation) form and content, and a representing thing is composed of a representation asset and a representation medium. Visualizations are the representation form of some content that express a geological concept or its individuals.

The Information Artifact Ontology (IAO) [Smith and Werner 2015] offers a domain-neutral resource to represent information content entities, such as documents, databases and digital images. The authors derive the ontology from BFO definitions [Arp, Smith et al. 2015]. For IAO, an *information content entity* is generically dependent on some *material entity* (in BFO sense) that retain a relation of *aboutness* to some entity. According to the authors, *aboutness* can be considered a reference relation that includes the relations of cognitive and intentional directedness involved in the capturing of information. Our model will refer to these uplevel concepts.

Table 1 – Definition of the geological modeling artifacts extracted from [Perrin, Rainaud et al. 2013]

Model	Abstracts a portion of reality according to the conceptualization of some observer. Is conceived according to some explicit theory known by the modeler.
Representation	Rests on a definite symbolism restricted by a representation language. Is connected to one model. The theory that underlines the model is incorporated into the representation. Many representations can be associated with a single model.
Visualization	Relies on the human visual system to perceive the modeled information Includes some parameters associated with the observer or to the conditions of observation. Many visualizations can be associated to one representation.

We believe that the understanding of the domain to propose the organization of image and content ontologies can improve the indexing and retrieval of domain images, then we based our organization system in the understanding of the domain methods and information for geological interpretation. This analysis is in section 4.

4. Criteria for image classification in petroleum exploration

The main contribution of this work is to clarify the criteria applied for visual content organization in petroleum domain and build the domain ontology for labeling or image artifact indexing. For a question of space, only the main classes of the ontology were described here while the quality attributes derived from the listed criteria and the *is about* relations were omitted.

In the following of this section, we describe the criteria applied for the visual content organization for exploration domain in our work. We propose the dimensions of analysis of the data and describe their meaning in Tables 2 to 7. On section 5, we formalize the entities deriving from IAO ontology.

4.1 Criterion Scale of Analysis

In petroleum exploration, the scale of analysis is one of the more important organization dimension. The scale of analysis considers both the dimension of the object of investigation, as well as, the scope of geological study. Our proposal is based on the works of [Fávera 2001; Jarna, Bang-Kittilsen et al. 2015]. The investigation of new economic assets starts on the study of the continental situation of the sedimentary basins at a very large scale, where the object of analysis spreads around 10^7 meters and proceeds with growing detail until the image acquisition of rock samples in nanoscale (10^{-6} meters). The data in continental and regional scale come from Government, academy or public agencies and support the decision of the target areas that will be licensed by companies. From this license on, the detailed studies that will produce data and images will be carried on by the company. Table 2 describes the scales of analysis and target objects in the exploration activities.

Table 2 – Indexing nomenclature for the scale of analysis

Name of scale	Order of magnitude	Geological/engineer concepts	Visualizations
Continental scale	10^7 meters	Continent, tectonic plate, tectonic structure	Continental map, diagram of plate tectonics, structural map
Basin scale	10^5 meters	Sedimentary basin Depositional system Tectonic system	Regional map, paleogeography map, diagram of tectonic and sedimentation interpretations
Field scale	10^3 meters	Petroleum field, regional stratigraphy, and geological formations, depositional environment	Local map, geological map, stratigraphy diagram, seismic section
Reservoir scale	10^0 meters	Geological formation, Regional Stratigraphy, geological formations tectonic structures	Geological 3D model, diagram block, seismic cube and section, stratigraphy diagrams, grid model, flow model
Outcrop scale	10^{-1} meters	Sedimentary facies, vertebrate fossil, rock layer, tectonic structure, geological formation, depositional geometry	Geophysics well log, lateral well imaging, diagram block, geological 3D model, columnar diagram, well pictorial descriptions, pictures

Macroscopic scale	10^{-3} meters	Lithology, sedimentary structure, fossil. Well cores and hand samples	Pictures, diagrams, graphical plots
Microscopic scale	10^{-6} meters	Lithology, sedimentary structure, microfossil, grain and minerals, chemical composition. Thin section, cuttings, earth material for chemical analysis	Pictures, diagrams, graphical plots
Nanoscale	10^{-9} meters	Grain and crystal aspects, mineral composition, disaggregate rock samples	Pictures, diagrams, graphical plots

4.2 Criterion Type of Visualization

Each of these scale analysis produces a large variety of types of visualizations for the geological objects. The Table 3 list here the type of visualizations found in petroleum exploration and their description. We consider that our classification system covers all the visualization found in the geology domain and then can be used to segment the whole set of available visualizations in a corporate database to support specific methods of content extraction.

Table 3 – Types of visualization for geological objects

	Visualization type	Description
1	Map	Diagrammatic visualization of the geographic distribution of some measure or information in plant
2	Cross Section	Diagrammatic visualization of the vertical slices of the Earth in some particular scale of analysis.
3	Profile	Diagrammatic visualization of the vertical slices of limited lateral portions of the terrain. A profile emphasizes the vertical variation of the geological features, which distinguishes it from cross-sections.
4	Geological Model	It is a 2D or 3D visualization of the interpreted distribution of rocks and geological structures in the subsurface. It is conceived to be produced by computer systems.
5	Photograph	An image captured by some equipment that can register the variances of light over the object or the scene.
6	Chart	Combined visualization of several distinct geological data related to a single variable (usually geological time) to help the geologist to get an integrated comprehension of the data.
7	Diagram	A schematic or simplified graphical representation of some geological feature
8	Graph	A graphical representation of a data set showing the relationship between two or more variables
9	Sketch	Hand-made draw that tries to keep some spatial correlation with an existing geological feature or scene

4.3 Criterion Methods for information acquisition

Besides scale and types of visualization, the geology method that produces the images also provides useful information for retrieval. Table 4 lists the main techniques utilized to produce information (textual and visual content) to support petroleum exploration.

Table 4 – Techniques and methods for information acquisition in exploration

Information acquisition technique	Scale of analysis	Visualizations
Aerophotogrametry/ Satellite image capturing	Continental	Aerial photograph, physical geography map, structural map
Radar	Continental	Radar image, Physical geography map, Structural map
Gravimetry study	Continental	Isopach map
Magnetic study	Continental	Isopach map
Geological regional mapping	Basin	Geological map, stratigraphic chart, photograph
Field studies	Field	Geological map, stratigrafic chart, photograph, paleontology chart, outcrop sketch
Seismic exploration	Field /reservoir	Subsurface map, cross section, profiles, cube, structural subsurface map and cross section, stratigraphic chart
Well perforation	Reservoirs/ Outcrop	Geophysic log, stratigraphic chart, profile, stratigraphic chart, core description, well-cutting log, borehole imaging, core box photograph, sample photograph
Seismic exploration	Field /reservoir	Subsurface map, cross section, profiles, cube, structural subsurface map and cross section, stratigraphic chart
Well perforation	Reservoirs/ Outcrop	Geophysic log, stratigraphic chart, profile, stratigraphic chart, core description, well-cutting log, borehole imaging, core box photograph, sample photograph
Geochemical analysis	Macroscopic	Graphic plot, compositional map
Petrographic analysis	Microscopic	Lithology description, microphotograph, graphic plot, lithology chart
Tomography/ spectrography analysis	Nanoscale	Photograph, graphic plot

4.4 Criterion entity accessibility

The professional can capture the data from direct and indirect observation (Table 5). Direct observation refers to any of the representations and their visualizations produced by a person who was able to have direct sensorial access to the geological object to produce a representation. Indirect observation refers to the representations and visualizations produced from data capture by analog devices, such as seismic, well logs, spectrograph, and tomography analysis.

Table 5 – Type of observation of the geological entity.

Type of observation	Geological/engineer concepts	Visualizations
Direct observation	Outcrop, core, rock sample	Photographs, core description, petrographic plots, aerial photograph
Indirect observation	Structure, geological unit, reservoir	Seismic profile and sections, seismic cubes, geophysical logs, radar images, isopach maps

4.5 Criterion location

Since Geology deals only with concrete entities, the location dimension is of central importance in information organization. Several distinct reference systems index information contents regarding location. Considering altitude/depth reference (Table 6), the data can come from the surface (maps, gravimetrical measures, geographic and geological maps, pictures) or subsurface data (seismic sections and maps, and geophysical well logs). In some situations, examples of visualizations may repeat themselves in surface and subsurface location. The more important index for general geological information is the geological and geographic location of the geological entities. Considering the maturity of the ontology-based systems for location, we index the visual content in this work with support of the previously existing systems for geographical position (such as, latitude and longitude, UTM coordinates, and other reference coordinates), geographical and geological location organized by the Open Geospatial Consortium (OGC) [OGC 2015]. For the geological location, we only consider the possible locations of petroleum relevant occurrences, which means sedimentary terrains. For the geoeconomic location, we adopt the Glossary of the Brazilian Petroleum Agency (ANP) [ANP 2016]. Finally, a particular case for geological location is the chronostratigraphic scale, which studies the position of rock bodies in relation to time. Table 7 summarizes the geolocation systems that we adopted in this work.

Table 6 – Indexing nomenclature for geological object location in surface/subsurface.

Vertical	Geological/engineer concepts	Visualizations
Surface	Continent, region, sedimentary basin, field, outcrop, isolate rock sample	Maps, columnar sections, gravimetric records, stratigraphic and compositional charts, diagrams and photomicrography
Subsurface	Geological units (host rocks and reservoir) , well cores	Seismic profile and sections, seismic cubes. Well logs, wall imaging of the well, diagrams and photomicrography

Table 7 – Adopted standard system and indexing nomenclature for geological object location in surface

Geolocation systems (horizontal location)	Adopted standard or nomenclature
Geographical position	The OGC standard for geographic information coordinate systems [OGC 2015]
Geographical	The OpenGIS Standard of OGC , for places.
Geological location	Continent, craton, basin, formation, layer and their subdivisions. OGC
Geoeconomic location	Sedimentary basin, petroleum or gas field, block, reservoir. The ANP Glossary [ANP 2016]
Chronostratigraphic location of geological units	Eonothem, Erathem, System, Series, and Stage

Besides technique and context, we need to classify the visual content according to the organizational origin of the data. Which sector of the company or external agent has produced the data and what is the organizational function of these agents? The geologist complements the exploration investigation with the studies of analogous, i.e., description of other reservoir occurrences in the world that keeps similarities with the target area and can bring useful lessons. In this work, we will not detail the classification system for organizational and external agents, since they are tightly related to the managerial organization of the companies.

Our work proposes an image classification system that considers each of the above aspects of information, which complements the geological domain ontology that describes the content of the images and pictorial representations.

5. Ontology modeling

We modeled the described entities by specializing the information content entity of the Information Artifact Ontology (IAO) [Smith and Werner 2015] available in Ontobee[4] [Xiang, Mungall et al. 2011]. Figure 1 intends to show what concepts were derived from IOF in order to cover the whole classification systems described in the previous sections. The derived nine geological visual entities have the names underlined in Figure 1. Each one of these entities keeps an *aboutness* relationship with predefined geological entities. The geological entity will be of one of the following types of material entities- *rock, geological object* – or generically dependent continuant - *geological structure* or specifically dependent continuant - *geological contact.* For reasons of space, it is not included in this paper the classes of geological objects and engineering objects that keep the aboutness relationship with the information content entities. In this stage of the research, only the visual content classes that we explain in this article were applied to classify the images and help the user to retrieve the image by the criteria described in section 4.2 to 4.5.

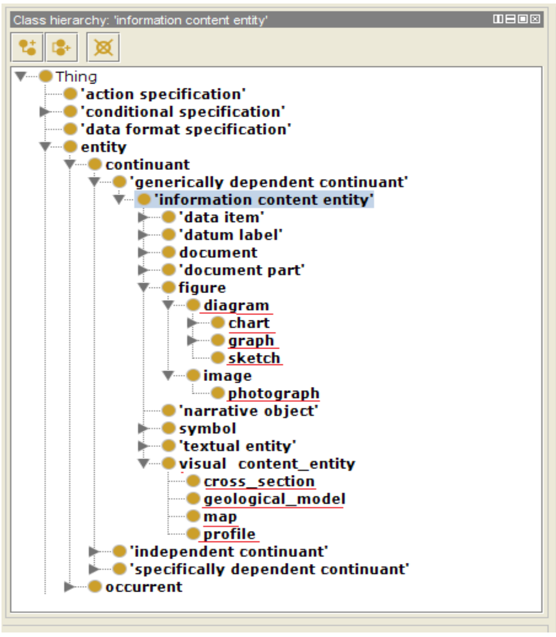

Figure 1 – Specializations of the Information dependent entity of IAO.

6. Content extraction

As a starting point, we have manually annotated, using the classification system, a dataset containing 261 images. The initial set includes images on two categories: (a) 128 images of type "Photograph" containing a variety of sub-types that include visual textural information of rocks (including photos of outcrops and cores); and (b) 133 images of type

[4] http://www.ontobee.org/ontology/IAO?iri=http://purl.obolibrary.org/obo/IAO_0000030

"Maps" containing examples of many of its sub-types. Figure 2 shows examples of photographs, and Figure 3 contains examples of maps from this dataset.

For each image, we computed a feature number based on the mean-squared-error deviation from a straight line of the image's log-histogram of horizontal derivatives (approximated by a convolution with the filter [1 -1]). Figure 4 shows preliminary results for each type of visualization. We then defined a binary classifier to divide automatically the images onto the two categories described above by hard thresholding. The threshold was selected by a simple brute-force search, where the objective function was defined as to maximize the accuracy of the classifier. This simple proof-of-concept obtained a precision of 97% (98 true positives and 3 false positives) with a recall rate of 76%.

Figure 2 – Photographs of well cores, outcrops, and sedimentary structures. Images from [d'Avila et al. 2004; Castro et al. 2005; Szatmari 2005; Magnavita et al. 2005; Moraes et al. 2006; Santos et al. 2008].

With this promising initial result, we moved to more challenging instances of the problem. We manually annotated a dataset of 1927 images from the domain of petroleum exploration using the proposed ontology. This dataset is composed of eight classes: seismic sections (248 images), geological maps (685 images), 3D block diagrams (100 images), profiles (99 images), thin sections (100 images), lithostratigraphic charts (170 images), portraits (164 images), and photographs of rocks (361 images).

For each image in this dataset, we computed a 253-dimensional real vector of features, composed of: *(i)* the 125 values of the image's RGB color histogram, computed by dividing the 3D RGB space in 5 bins in each dimension ($5 \times 5 \times 5 = 125$); and *(ii)* the first 128 values of the histogram of the image's horizontal derivatives magnitudes (that is, the absolute value was taken after applying the filter [1 -1] described above). Both histograms were normalized by dividing each bin count by the maximum bin count.

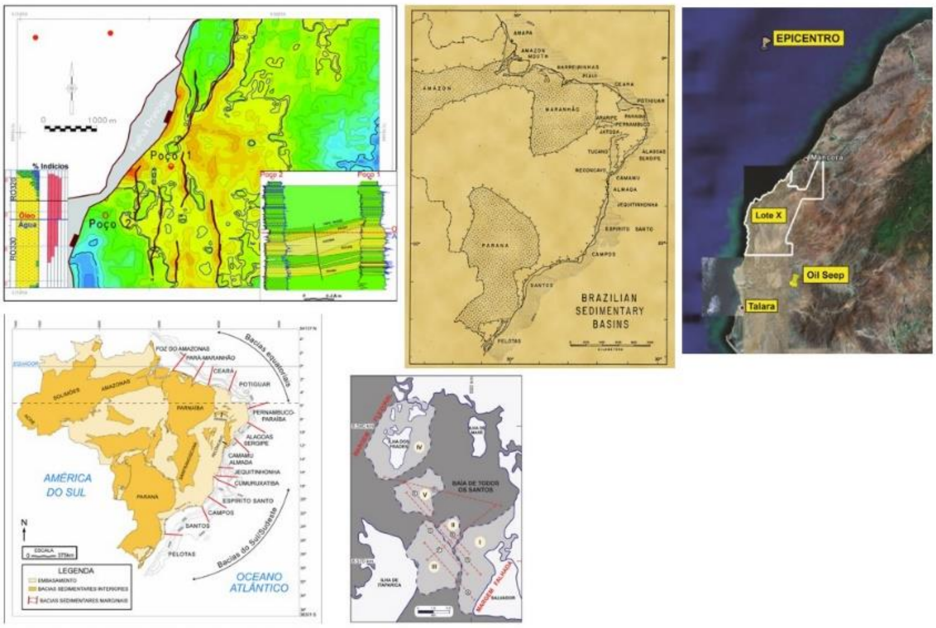

Figure 3 – Instances of maps in several scales and techniques of analysis. Images from [Júnior et al. 2004; Arai and Lana 2004; Ponte and Asmus 2004; Cupertino and Bueno 2005; Daudt et al. 2009]

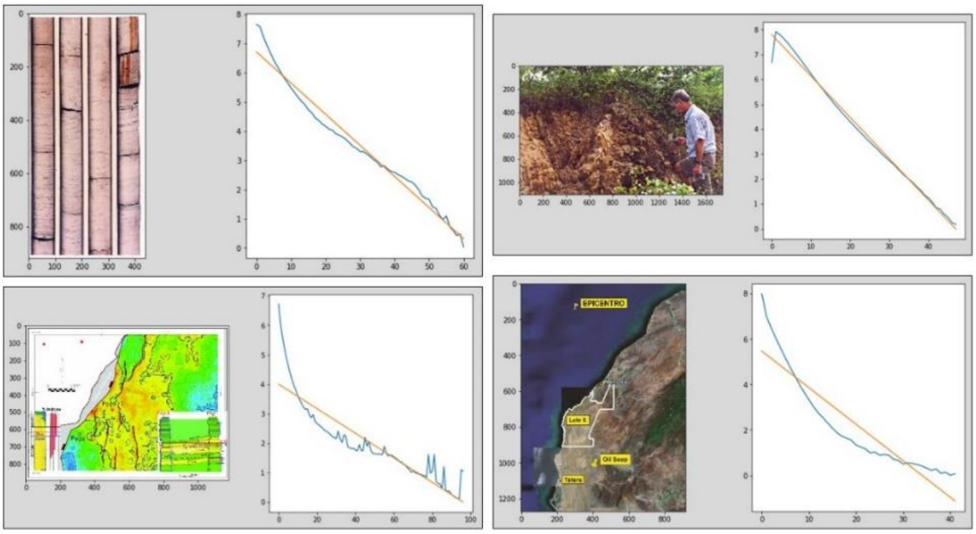

Figure 4 – Four examples of visualization with the respective log-histogram of horizontal derivatives that shows a particular pattern for each visualization type.

With the described features, we trained a Logistic Regression multi-class classifier with a simple L2 penalty. For the loss function, we balanced the weight given to each class based on the total number of images per class, to avoid biasing the classifier to the most frequently occurring classes. We performed 50 runs using 10-fold cross validation, with random permutations of the dataset, and obtained a mean accuracy of 59% in the classification (a good improvement over the expected accuracy of 12.5% for a uniformly random classifier, assuming a uniform prior on the class distribution). Minimum/Average/Maximum precision and recall rates are, respectively, 15%/54%/83% and 34%/57%/75%.

These numbers show that there is a lot of room for improvement in the classification. We are now exploring non-linear classifiers based on Neural Networks, where a large and correctly annotated dataset becomes essential. This method will be complemented by the annotation of the geological objects on the images and exploration of the *aboutness* relation on the automatic classification algorithm. In this direction, the proposed ontological organization system is a key part for simplifying and guiding the annotation process.

7. Conclusion

We presented here a knowledge organization system for the classification and indexing of visualizations of geological content in petroleum exploration domain. We investigate the several types of visual content utilized in petroleum exploration, and we propose a limited number of classification criteria and the associated visual content classes to support image organization in corporate systems. Our approach has the benefit of the ontological analysis of the object of representation (the visual artifact), as well as the understanding of types and techniques for expressing the semantic content (the geological object that the artifact is about). We consider that this multi-classification system will reduce the need for a large number of images in training sets on the visual content extraction in the domain. Preliminary results in visual-content extraction showed that, with support of the organization system, we reduce some complex steps of image analysis that allows applying basic image processing algorithms.

Acknowledge: We carry on this project in collaboration with Petrobras. We acknowledge the support of Brazilian sponsor agencies CAPES and CNPQ.

References

Abel, M., K. Goldberg, et al. (2013). "Ontology-based rock description and interpretation", Shared Earth Modeling. M. Perrin and J.-F. Rainaud. Paris, Editions Technip. 1: 268-271.

Abel, M., L. S. Mastella, et al. (2004). "How to model visual knowledge: a study of expertise in oil-reservoir evaluation", Database and Expert Systems Applications. F. Galindo, M. Takizawa and R. Traunmüller. Zaragoza, Spain, Springer-Verlag GmbH & Company KG, Berlin, Germany. 3180: 455-464.

ANP. (2016). "Glossario ANP", http://www.anp.gov.br/glossario#gloss-C, February 27, 2019.

Arai, M., C. C. Lana. (2004). "Retrospective of fossil dinoflagellate studies in Brazil: their relationship with the evolution of petroleum exploration in the Cretaceous of continental margin basins", B. Geoci. Petrobras, Rio de Janeiro, v. 12, n. 1, p. 175-189, nov. 2003/maio.

Arp, R., B. Smith, Spear, A.D. (2015). "Building Ontologies with Basic Formal Ontology", MIT Press.

Barreiros, C. C. M. (2010). "Uma ontologia sobre descrição de afloramento para auxiliar geólogos nas atividades de campo", Master Dissertation, UFRJ

Castro, E. S. E. D. (2012). "PetrographypediA. The portal of petrography.", http://www.petrographypedia.com/, July 15, 2015.

Castro, J. C., L. C. Weinschütz, et al. (2005). "Estratigrafia de sequências das formações Taciba e Rio Bonito (Membro Triunfo) na região de Mafra/SC, leste da Bacia do

Paraná.", B. Geoci. Petrobras, Rio de Janeiro, v. 13, n. 1, p. 27-42, nov. 2004/maio.

Chen S.H. & Chen Y.H. (2017). “A Content-Based Image Retrieval Method Based on the Google Cloud Vision API and WordNet”. In: Intelligent Information and Database Systems. ACIIDS 2017 (eds. Nguyen N, Tojo S, Nguyen L & Trawiński B). Springer Kanazawa, Japan.

Cupertino, J. A., G. V. Bueno. (2005). "Architecture of stratigraphic sequences developed in the deep lake phase of the Recôncavo Rift", B. Geoci. Petrobras, Rio de Janeiro, v. 13, n. 2, p. 245-267.

Daudt, J. A. B., A. Benedicto, and Pozo, E.G. (2009). "Seismicity induced petroleum migration: field and subsurface observations in Talara Basin (northwestern Peru)", B. Geoci. Petrobras, Rio de Janeiro, v. 17, n. 2, p. 371-374.

D'Avila, R. S. F., Souza Cruz, C. E. , and Oliveira filho, J.S.. (2004). "Fácies e modelo deposicional do Canyon de Almada, Bacia de Almada, Bahia", B. Geoci. Petrobras, Rio de Janeiro, v. 12, n. 2, p. 251-286, maio/nov.

Fávera, J. C. D. (2001). “Fundamentals of Modern Stratigraphy” (original in Portuguese). Rio de Janeiro, Federal University of Rio de Janeiro State.

Chechik, G.; Shalit, U., Sharma, V; Bengio, S. (2009). “An Online Algorithm for Large Scale Image Similarity Learning”, Advances in Neural Information Processing Systems (NIPS). Vancouver, NIPS Foundation.

Gonçalves, F. M. F., I. R. Guilherme, et al. (2018). “Semantic Guided Interactive Image Retrieval for plant identification”, Expert Systems with Applications 91: 12-26.

Jarna, A., A. Bang-Kittilsen, et al. (2015). “3-Dimensional geological mapping and modeling activities at the geological survey of Norway”, Joint International Geoinformation Conference 2015 Kuala Lumpur, Malaysia.

Machado Júnior, D. L., Silva Coelho D. F. da, Selbach, H.S. et al. (2004). “Structural reconstruction and oil prediction in the Roncador Field, Campos Basin, Brazil.”, B. Geoci. Petrobras, Rio de Janeiro, v. 12, n. 1, p. 89-101, nov. 2003/maio.

Kuang, Z., J. Yu, et al. (2018). “Integrating multi-level deep learning and concept ontology for large-scale visual recognition”, Pattern Recognition 78: 198-214.

Lorenzatti, A., M. Abel, et al. (2009). “Ontology for Imagistic Domains: Combining Textual and Pictorial Primitives”, Advances in Conceptual Modeling - Challenging Perspectives. C. A. Heuser and G. Pernul. Gramado, Brazil, Springer Berlin / Heidelberg. 5833: 169-178.

Magnavita, L. P., R. R. Silva, et al. (2005). “Field trip guide of the Recôncavo basin, NE Brazil”, B. Geoci. Petrobras, Rio de Janeiro, v. 13, n. 2, p. 301-334.

Moraes, M. A. S., P. R. Blaskovski, et al. (2006). “Arquitetura de reservatórios de águas profundas”, B. Geoci. Petrobras, Rio de Janeiro, v. 14, n. 1, p. 7-25.

Ponte, F. C., H. E. Asmus. (2004). “The Brazilian marginal basins: current state of knowledge”, B. Geoci. Petrobras, Rio de Janeiro, v. 12, n. 2, p. 385-420.

Santos, J. P. P., C. Bettini, et al. (2008). “Spatial representation of turbidite lobes and channels in outcrops of Apiúna region, Itajaí Basin, Santa Catarina.”, B. Geoci. Petrobras, Rio de Janeiro, v. 16, n. 1, p. 69-85.

Szatmari, P.. (2005). “Entrevista Peter Szatmari”, B. Geoci. Petrobras, Rio de Janeiro, v. 13, n. 1, p. 105-121.

Mizoguchi, R. and F. Toyoshima (2006). "YAMATO: Yet-Another More Advanced Top-level Ontology", Applied Ontology 3: 1-3.

Mohd Zin, N. A., R. Yusof, et al. (2018). "Content-Based Image Retrieval in Medical Domain: A Review", Journal of Physics: Conference Series 1019: 012044.

Nascimento, F. M. d. and T. Freire (2005). "GEODESC - Vocabulário controlado em Geociências". Rio de Janeiro CPRM/DIDOTE.

Neto, M. F. d. S. (2013). "Ontolime: modelo de ontologia de descrição de imagens médicas", Master Dissertation, UNESP.

OGC (2015). "Geographic information - Well-known text representation of coordinate reference systems", Open Geospatial Consortium. http://docs.opengeospatial.org/is/12-063r5/12-063r5.html, July 17, 2019.

Pandey, S., P. Khanna e H. Yokota. (2016) "A semantics and image retrieval system for hierarchical image databases", Inform. Processing & Management, v.52, n.4, p.571-591.

Perrin, M., J.-F. Rainaud, et al. (2013). "Earth models as subsurface representations", Shared Earth Modeling M. Perrin and J.-F. Rainaud. Paris, Editions Technip: 3-24.

Sharma, M. K. and T. J. Siddiqui (2016). "An Ontology-Based Framework for Retrieval of Museum Artifacts", Procedia Computer Science 84: 169-176.

Smith, B. and C. Werner (2015). "Aboutness: Towards Foundations for the Information Artifact Ontology", Sixth International Conference on Biomedical Ontology (ICBO), CEUR-WS. 1515: 1-5.

Tian, Y. (2016). "Integrating Textual Ontology and Visual Features for Content-Based Search in an Invertebrate Paleontology Knowledgebase", Master Dissertation, University of Kansas.

Xiang, Z., C. Mungall, et al. (2011). "Ontobee: A Linked Data Server and Browser for Ontology Terms", 2nd International Conference on Biomedical Ontologies (ICBO). Buffalo, NY: 279-281.