Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A REVIEW ON ONTOLOGY-BASED LABEL EXTRACTION FROM IMAGE DATA

[1]AHMAD ADEL ABU-SHAREHA, [2]MANDAVA RAJESWARI

[1] Department of Computer Science, Faculty of Information Technology, WISE University, Amman, Jordan
School of Computer Science, Universiti Sains Malaysia, Penang, 11800, Malaysia
[2] School of Computer Science, Universiti Sains Malaysia, Penang, 11800, Malaysia
E-mail: [1]abushareha@wise.edu.jo, [2]mandava@cs.usm.my

ABSTRACT

Ontology-based label extraction is extensively used to interpret the semantics found in image and video data. Particularly, ontology-based label extraction is one of the main steps in object class recognition, image annotation, and image disambiguation. These applications have important roles in the field of image analysis, and as such, a number of variations of the ontology-based label extraction used in these applications have been reported in the literature. These variations involve ontology development and utilization, and can affect the applicability (e.g., domain- and application-dependency) as well as the accuracy of the output. Unfortunately, the variability aspect of this variation has neither been established nor tracked. Thus, the variations were not configured. A review of the ontology-based label extraction based on the input data, the utilized technique, and the type of utilized ontology is presented in this paper. The ontology-based label extraction is categorized based on two aspects, namely, the type of input data and the type of ontology used. These two aspects determine the type of the label extraction technique to be used. As a result, the relative advantages and disadvantages of each category are determined. The gaps and future research directions in this field are also highlighted.

Keywords: *Label Extraction, Ontology, Semantics, Semantic Tagging, Knowledge Markup, Image Object Recognition, Image Disambiguation, Image Annotation*

1. INTRODUCTION

In image processing, recognition refers to the process of identifying the object(s) in a specific scene of an image or video. Object class recognition is used to identify an object in a given image. Image annotation recognizes, and sometimes spatially distinguishes, objects within a given multi-object image. The recognition process is implemented in two phases: training and prediction. The training phase creates a features-to-objects model, whereas the predicting phase uses the created model to predict or identify the object in the input image, given the extracted features [15]. Unfortunately, image features, which are the bases of these phases, are ambiguous in nature because the same object may have a wide range of features resulting from various image conditions, such as lights, noise, and viewing angle. Consequently, image disambiguation is commonly used as a follow-up procedure. Image disambiguation takes the set of ambiguous labels as input and outputs a refined set. The disambiguation process is also implemented with the training and predicting phases. The model created in the training phase and used in the predicting phase forms associations among objects based on the content information. The content information, in turn, is based on the hypothesis that objects appearing together during the training phase tend to appear together in the predicting phase [1, 17].

Overall, extracting true image labels in image annotation, class object recognition, and image disambiguation require trained models [2]. Recently, these applications have extensively used ontology to replace trained models, because ontology offers the required association between features and objects and has a well-established modeling procedure [1, 5, 14, 18, 20, 22]. Generally, these applications, despite having various input forms, all produce labels. Therefore, they share the process of ontology-based label extraction.

Ontology is a conceptual knowledge source, which mainly consists of concepts and their hierarchical relationships. A concept is a tag identified by a word, phrase or label, and describes

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

a real-world entity. Ontology may also have properties that describe the concepts and non-hierarchical relationships among the concepts of the ontology. Figure 1 illustrates an example of an ontology. Ontology may be used as a hierarchically-enabled browsing mechanism and can be employed in semantics extraction, the process of accessing ontology and inferring knowledge based on its concepts and relationships.

Ontology-based label extraction, a type of semantics extraction, produces true labels for an input image. Generally, given an input image, the ontology-based label extraction process has several steps. First, the input image features are projected and matched with concepts in the ontology through a process called mapping. The relationships connected to the matched concepts are then analyzed, and new concepts are identified sequentially until the final output is extracted. This process is called mining. An example of ontology-based label extraction is illustrated in Figure 2.

As shown in the example, concepts with labels identical to the input features (e.g., white and oval) are matched. New concepts are then identified through linkage with the matched concepts. In the example shown in Figure 2, the final semantics output is “natural scene.”

Unlike in the illustrated example, the actual application of the label extraction mechanism requires a more complex technique due to the ambiguous nature of the visual input features. Consequently, various techniques and ontologies are used, which affect the applicability (e.g., domain- and application-dependency) and accuracy of the output. Thus, a review of the technique used in association with the utilized ontology is required.

In this paper, a review on the existing label extraction mechanisms is conducted. First, the common characteristics of the existing methods have to be determined in Section 4. Second, the source of variability in the existing methods must be identified and analyzed, as given in Section 5, Section 6 and Section 7. Finally, the existing methods should be compared and analyzed based on the identified variability criteria, as given in Section 8. Eventually, the advantages, disadvantages, and the gap in the current methods can be identified as given in Section 9. Conclusion is given in Section 10.

2. RELATED WORK

Existing surveys mainly focus on a single application/problem (e.g., recognition, annotation, and disambiguation) and have reviewed the existing literature from several perspectives [6, 20]. Tousch et al. [43], for example, reviewed and analyzed the existing literature on image annotation based on the annotation level they provide (e.g., generic level “tower” or specific “Eiffel-tower”) and the method used in the annotation process (e.g., statistical, hierarchical, etc.). Hanbury [16] reviewed and compared studies on image annotation based on annotation level and user convenience. Liu et al. [25] and Zhang et al. [48] reviewed and analyzed the existing literature on image annotation based on the image features and the trained model used. Other surveys also looked into these applications from several other perspectives [8, 11, 39, 41, 34, 42, 50, 47]. Generally, the existing literature focuses on comparing and analyzing methods based on the characteristics of the output, with no linkage to the technique and type of ontology used. Thus, the connections among input, utilized technique, utilized ontology, and the output have yet to be established.

3. REVIEW SCOPE AND GOAL

The goal of this review is to review the existing research related to ontology based labels extractions. The review focuses on the forms of the input, the utilized techniques and the characteristics of the output. The papers included or may all compared to these included in this review are those used ontology-based labels extraction as a core, primary or secondary process. Existing research that uses other forms of labels extractions that do not use ontology are not involved. Also, other forms of ontology utilization also not involved.

4. COMMON CHARACTERISTICS

In this paper, the common characteristics of the reviewed papers are established as being an ontology-based, having label extraction process, having image-based/visual-based inputs.

5. VARIABILITY CRITERIA

Figure 3 illustrates the process of ontology-based label extraction and its related aspects. Generally, the methodology used in this type of extraction and the selected ontology control many other aspects related to the task and have major influence on the final output.

The input for ontology-based label extraction may be image features or object labels (maps) extracted using various image annotation techniques (Figure 3). The mapping procedure is constrained with the type of input and ontology

Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

characteristics. Features input, which have a wide range, require learning techniques. Meanwhile, maps can be mapped directly (e.g., using syntactic string matching). Ontology is task-independent and is developed by domain experts. Existing ontologies, such as WordNet [9] and Cyc [10], are upper-level ontologies that consist of a large number of concepts and their relationships. These ontologies may be used with various applications. However, an existing ontology may be customized depending on the task at hand and the desired output. Ontology customization usually involves extracting a specific part of the ontology, which includes the required concepts and some of their relationships. In addition, ontology-like knowledge may be developed if the required concepts or their relationships do not exist in the existing ontologies. The mining procedure depends greatly on the type of the output, that is, if the output is part of the input (i.e., image disambiguation), then a similarity technique is used; otherwise, a flooding procedure is implemented (i.e., image annotation).

In this paper, the reviewed literature (Figure 4) is categorized based on two criteria: image-related input data, which may consist of features or maps (labels), and the form of ontology used. These two criteria precisely determine the procedures to be used and categorize the extraction techniques into several approaches. The characteristics, advantages, and disadvantages of each approach are discussed in the following sections.

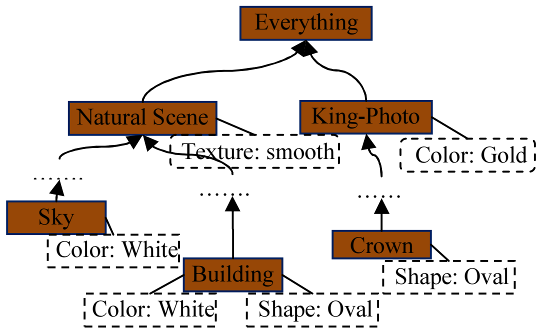
Note that the main focus of this paper is ontology-based label extraction in image annotation, recognition, and disambiguation. Thus, some pre-processing and post-processing steps might be ignored.

6. FEATURE BASED LABEL EXTRACTION

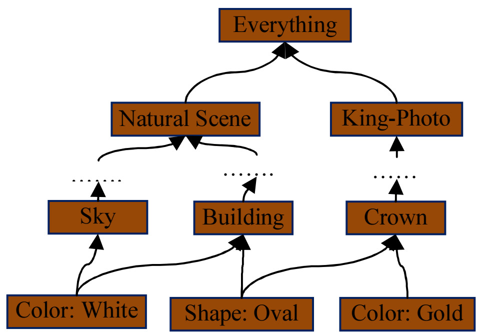
Image annotation and object recognition, as mentioned earlier, predict object(s) in a given scene based on the extracted features. Subsequently, these applications require an ontology that forms associations among features and labels for objects. Generally, existing ontologies do not include the visual properties of the described objects [9, 37]. Thus, feature-based label extraction uses customized ontologies or ontology-like knowledge developed for the task at hand.

The structure of these specific task ontologies depends on the task at hand and the desired output. Variations of this structure are reflected in the ways by which the required image features are represented. Image features may be represented in two ways: as properties associated with the ontology concepts (e.g., those shown Figure 1) or as bottom-level concepts (e.g., leaf concepts or those shown in Figure 2). Figure 5 illustrates the clear differences between these cases.

Mapping is implemented using a classification approach. Mining begins with a flooding process, which tracks the relationships among the matched concepts and the other concepts, in order to to identify the desired set of output concepts [36]. Flooding can be implemented bottom-up or top-down, depending on the way the visual properties are represented in the underlying ontology. If visual features are represented as bottom concepts, then bottom-up propagation is utilized. By contrast, if visual features are represented as properties, then either bottom-up or top-down propagation is utilized.



(A) Features As Properties



(B) Features As Concepts

Figure 5: Image features in ontology

In general, feature-based label extraction focuses on constructing an ontology that can efficiently map the low-level features of the images into concepts and allow a smooth and accurate mining procedure [12]. Once the ontology is constructed, the predicting process using classification and

Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

propagation is implemented. These procedures can be generally represented by Equations 1 and 2 below.

$$\{o\} = classification(inputFeatures|Ontology) \quad (1)$$

$$\{C\} = Flood(\{o\}|Ontology) \quad (2)$$

For Equation 1, {o} is the set of ontology components (either concepts or properties, depending on the utilized ontology) that match the input features using a classification method. For Equation 2, {C} is the ontology-based label extraction output that refers to the set of concepts extracted using the flooding process.

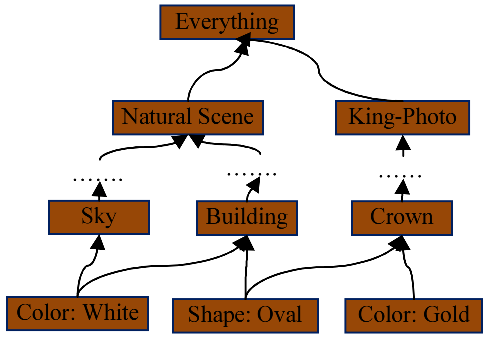
The feature-based label extraction approaches are categorized into two sub-approaches: task-oriented and general approaches.

6.1. Task-oriented

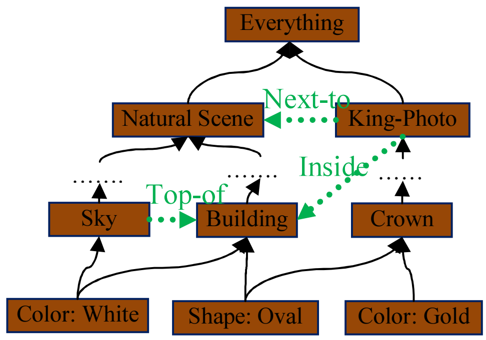
In the task-oriented category, ontology-like knowledge is developed to smoothly fit the task at hand. These ontologies, however, cannot be used elsewhere. Two main approaches, standard and advanced approaches, are then proposed. Their main differences are in the information conveyed by their ontologies, which require the use of different techniques.

In the standard approach, the ontology conveys the following information: object labels, hierarchical relationships, and low-level features. In the advanced approach, the ontology has an additional feature, i.e., the spatial relationships among concepts. Figure 6 illustrates the differences in the ontologies between the two approaches.

In the standard approach, Penta et al. [36] proposed a semantics extraction process for object recognition. During the ontological construction, concepts are created based on labels obtained from a dataset of labeled images. Then, another set of concepts with coarse granularity is manually created to facilitate the categorization principle of the ontology. Low-level features are then assigned as properties to each concept using a supervised machine learning process.



(A) Ontology Without Spatial Relationships



(B) Ontology With Spatial Relationships

Figure 6: Spatial information in ontology

In label extraction, features are extracted from the input image, labeled, and then mapped to properties in the ontology using a classification method. Mining is implemented as a propagation process, which transfers from one concept to another over the hierarchical relations in a top-down manner (from the concepts at the general level to the concepts at a specific level). The propagation process might be intermediate and have more classification processes, in order to filter out the concepts reached through the propagation process. Finally, the concepts obtained at the lowest level (i.e., leaf) of the propagation process are selected as the output. A label extraction for the birds' image retrieval was proposed by Liu et al. [26]. Figure 7 illustrates an example that involves all the processes in the standard approach.

In the advanced approach, Maillot and Thonnat [33] proposed a method for object recognition based on label extraction. They constructed an ontology with spatial relationships among the concepts. These relationships are defined using supervised learning based on a training image dataset. In this ontology, the visual properties are attached as leaf concepts with spatial, non-hierarchical relationships. The label extraction is implemented as a classification process wherein the features of the input image are matched to the leaf

Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

concepts, which is then propagated in a bottom-up manner. In general, spatial relationships may be used based either on a binary logic, by which objects/concepts that do not adhere to these relations may obtain a classification outcome equal to zero, or on a fuzzy logic, by which the objects or concepts that do not adhere to these relations are panelized rationally. Town [44] proposed adding co-occurrence probabilities to the spatial relationships among the concepts of the ontology. Similar methods have been proposed by Clouard et al. [5] and Ganea and Brezovan [12]. Figure 8 illustrates the common processes in the standard and advanced approaches for task-oriented, feature-based label extraction.

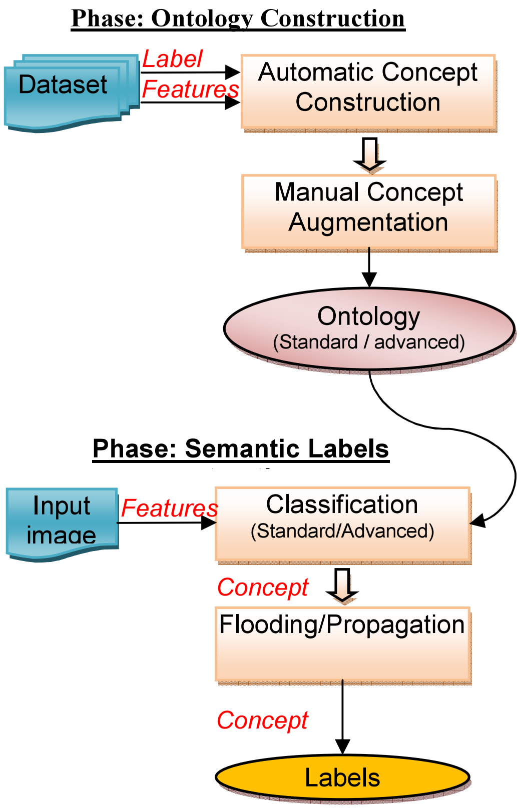


Figure 8: Common processes used in task-oriented, feature-based label extraction

Overall, task-oriented, feature-based label extraction uses classification and propagation processes to extract the desired labels, with reference to an ontology developed for the underlying task. In fact, both the standard and advanced approaches aim to reduce the ambiguities of the visual input features by conveying as much information as possible in the utilized ontology. Subsequently, the standard approach that uses the hierarchical relationships in the ontology has been extended in the advanced approach, which includes the spatial and co-occurrence relationships. As such, each variation of a visual component can be linked to the presence of other components in the same visual content.

The advantages of these approaches include their ability to convey various forms of information such as features, spatial relationships, and co-occurrence relationships. Theoretically, such information improves the accuracy of object recognition, controls the image low-level variation problem, and reduces the image ambiguity problem. The ambiguity problem still overlaps the label extraction process in image domain when these approaches are used. Nevertheless, another limitation of such an approach lies in the domain limitations of the constructed ontology, which is only created for a specific task. As such, the same ontology cannot be utilized for different datasets or with different types of image features [44].

6.2. Generic

In the generic category, ontology construction and label extraction phases are reported independently and separately. As the ontology is constructed, or at least its specifications are designed independently, a generic characteristic is established. Related studies on generic ontology construction focus on designing a widely applicable ontology for label extraction from a wide-range of image data.

The major contribution in this area is the recently produced Multimedia Ontology Web Language (M-OWL), a formal ontological language that supports media content description based on MPEG-7. The ontology in M-OWL encompasses all layers of content descriptions, namely, the low-level features based on MPEG-7, conceptual labels, hierarchical structure based on the Ontology Web Language (OWL), and concept-to-concept contextual probability based on the Bayesian theory [13, 14, 40]. Figure 9 illustrates the hierarchy of the top-level concepts of MPEG-7 ontology. Similar approaches for constructing a multimedia ontology language by integrating MPEG-7 and OWL have been proposed in previous works [[45, 46].

The ontology in the generic category is created with two goals, namely, to reutilize the ontologies and to extract labels from a wide range of images. However, given these two goals, the underlying ontologies resulted in complicated structures.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

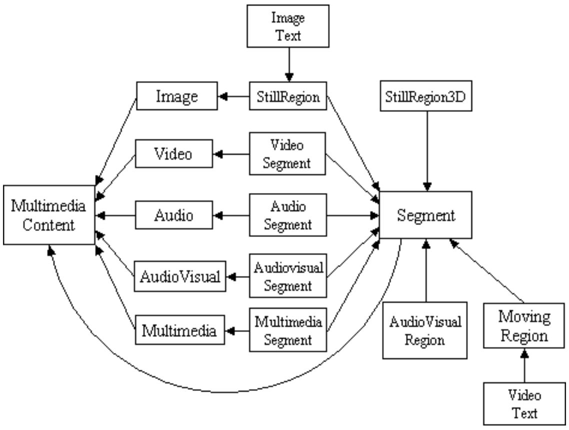


Figure 9: Hierarchy of the top-level concepts in MPEG-7 [4]

In label extraction, Mallik and Chaudhary [28] proposed an annotation method, which uses the MOWL specification to create a domain ontology. Label extraction is implemented using a classification process based on Bayesian theory and a propagation process associated with Bayesian probability. Previous works [29, 30] have proposed similar annotation methods based on MOWL. Overall, the generic feature-based label extraction follows the same techniques as the advanced approaches used in task-oriented label extraction. While still in the development stage, this approach aims to provide a widely accepted ontology, which can be adapted in several tasks. The common processes used in the generic feature-based label extraction are illustrated in Figure 10.

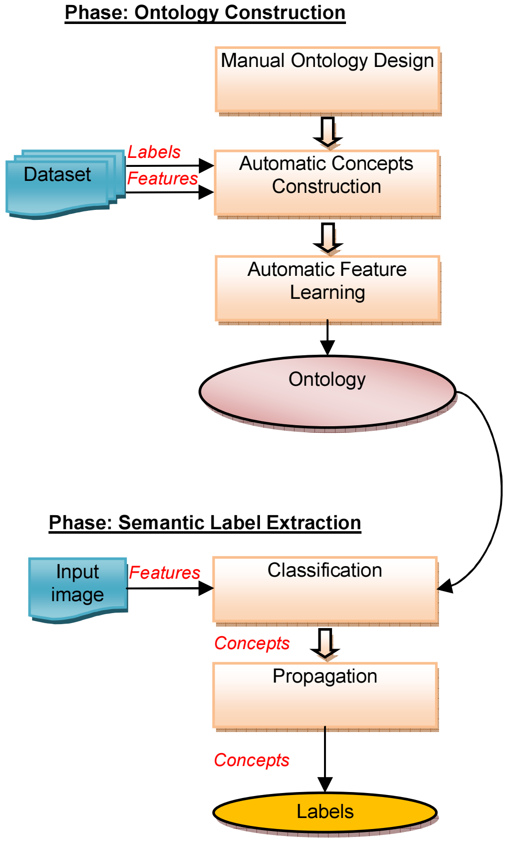


Figure 10: Common Processes Used In Generic Feature-Based Label Extraction

7. MAP-BASED LABEL EXTRACTION

Maps are labels that textually describe the contents of an image. Through an image annotation method, features of maps are extracted via pre-defined feature-to-map association [22, 17, 32].

As mentioned earlier, image disambiguation analyzes both input (a set of ambiguous labels) and output (a refined set), as well as requires an ontology that forms the associations among the objects. Existing ontologies provide a rich source of such associations. Map-based label extraction uses these existing ontologies or customized versions of these ontologies [49]. Examples of an ontology and its customized version are illustrated in Figure 11.

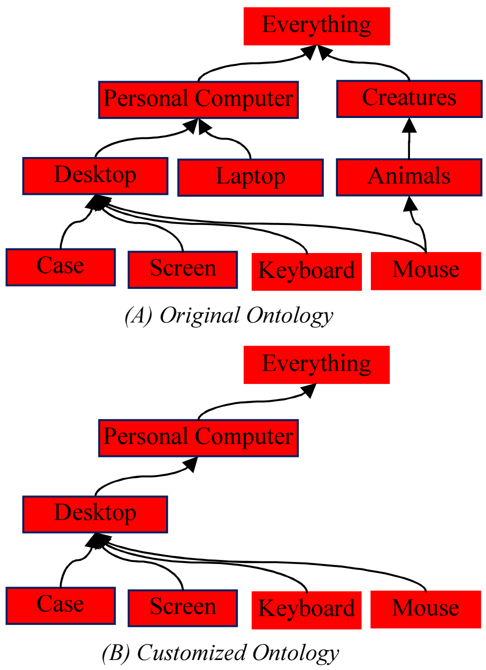


Figure 11: Customization of an ontology

Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Existing ontologies represent objects as concepts, and the associations among these objects are built through hierarchical relationships. Mapping is implemented using a string matching process, whereas mining is implemented using a semantic similarity process, which measures the strength of the relationships among the mapped concepts, in order to identify the desired set of output concepts [7].

Map-based label extraction uses string matching and semantic similarity processes. These processes are represented by Equations 3 and 4.

$$\{o\} = stringMatch(inputMaps|ConceptLabels) \quad (3)$$

$$\{C\} = similarity(\{o\}|hierarchicalRelationships) \quad (4)$$

In Equation 3, {o} is the set of ontology concepts that match the input maps using a string matching method. In Equation 4, {C} is the ontology-based label extraction output that refers to the set of concepts with the maximum similarity values.

The map-based category can be further classified into customized ontology-based and existing upper-level ontology-based categories based on the utilized ontology.

7.1. Customized Ontology

A customized ontology can be designed depending on the task at hand. Galleguillos et al. [11], Park and Lee [35], and Zlatoff et al. [51] created ontologies using object labels as concepts, hierarchical relationships, and co-occurrence relationships. Here, a machine learning approach is used to construct the desired ontologies. First, an ontology is constructed by transforming the image labels in the training images into concepts. The relationships among the concepts are then constructed based on the co-occurrence relationships between objects in the training images.

The label extraction begins with direct mapping and mining, which is a simple semantic similarity measure. Using the measured semantic similarity, each concept is classified as either relevant or irrelevant based on its similarity to the other concepts in the same visual content. As such, the initial set of concepts corresponding to the input labels are marked in the mapping process. Then the similarities between these concepts are calculated. The concepts with strong similarity to others are selected as the output while the rest are discarded.

Figure 12 illustrates the common processes used in the customization-based map-based label extraction. The drawback of this approach is its task and dataset dependency.

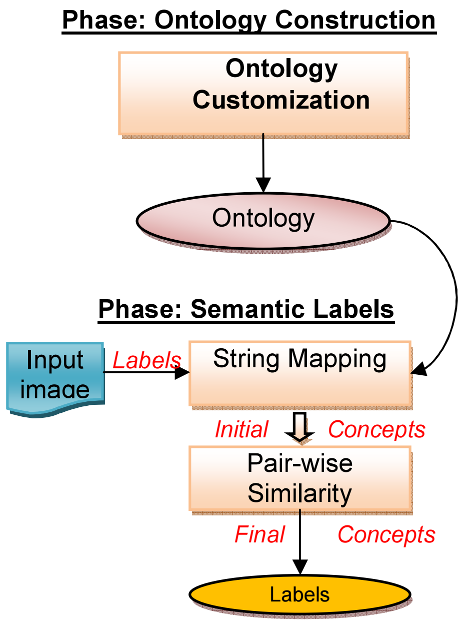


Figure 12: Common Processes Used In The Customization-Based Label Extraction

7.2. Existing Ontologies

WordNet [9] has been used widely in the existing ontology category. Jin et al. [20] used WordNet and a label extraction technique for a semantics-based image disambiguation. An input set of ambiguous image labels (label sets composed of the true labels and the ambiguous labels) is mapped directly to the concepts in the WordNet. Pair-wise relatedness measurements among these identified concepts are then conducted. Each concept is given a significant value based on the strength of its relationships with other concepts, noting that stronger relationships make the concept more significant. Finally, the concepts are classified as either relevant or irrelevant based on the value of their significance. Irrelevant concepts are discarded for all sets. Semantic similarities are measured using several methods, namely, the information content measures of Resnik [38], Jiang and Conrath [19], Lin [24] and Banerjee and Pedersen [3], as well as the edge-based measure of Leacock and Chodorow [23]. Similar approaches for semantics-based disambiguation based on various semantic similarity measures have been proposed, such as that of James and Hudelot [18], which uses structural semantic interconnections [31], and that

Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of Liu [27] which uses edge-based similarity. Figure 13 illustrates the common processes used in the existing-based map-based label extraction.

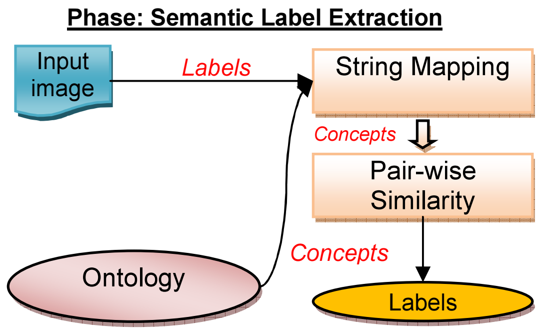


Figure 13: Common Processes Used In The Map-Based Label Extraction

Overall, the ontology construction phase in the map-based category is only implemented in the customized ontologies, following the same process used in the feature-based category but without utilizing the low-level features. The map-based label extraction depends mainly on an ontology that has concepts and hierarchy relationships. Spatial relationships are not included in the existing category because these relationships are usually not presented in existing general-purpose ontologies. However, the spatial and co-occurrence relationships included in the customized category are specifically developed using a number of training image dataset. Based on these prospective techniques, the input maps, which are string-based labels, are mapped directly to the concepts in the underlying ontology in both categories. Then, given the identified concepts, a pair-wise similarity is implemented to identify the final output labels.

8. COMPARISON

Ontology-based label extraction methods utilize either existing, developed, or customized (semi-automatically constructed) ontologies. The characteristics of the utilized ontology affect the applicability of the overall method in terms of the domain and task involved. Existing ontologies, which are developed by experts, are domain- and task-independent ontologies. By contrast, customized ontologies have limited capabilities, except for the generic sub-category in the feature-based label extraction, which is still not applicable in this field.

Based on these techniques, the feature-based category requires a classification process to map the extracted features into concepts in the utilized ontology, whereas the map-based category implements direct mapping. In the second step, all the methods belonging to different categories may use the same techniques even though they all operate on defined concepts. However, common feature-based methods use a propagation process, whereas map-based methods use pair-wise similarity calculation. Table 1 shows the comparisons among the discussed categories in terms of label extraction.

Based on the conveyed and employed information, methods used in different categories employed different types of information in the extraction process. Label extraction is surrounded by ambiguity resulting from the variation problem of the low-level features used for semantic extraction. This ambiguity is addressed by the previously discussed label extraction process using content and context information. By using content information, the feature variation of a given object can be mitigated, because each variation can be linked to the presence of other components in the same visual content (e.g., hierarchical relationships). By contrast, context information addresses the probability of having an object in the image using the presence or absence of some other objects, regardless of the content of the scene and its low-level features (e.g., non-hierarchical relationships). Figure 14 illustrates the process by which the discussed categories utilize content and context information.

The task-oriented, feature-based semantics approach utilizes content information only, whereas its generic counterpart covers a wide domain and utilizes both content and context information. However, this generic feature-based semantics approach has no actual implementation. The map-based approach utilizes contextual information, and leaves the contents to be conveyed using the annotation process, which produces the maps used as the input of the methods in this category.

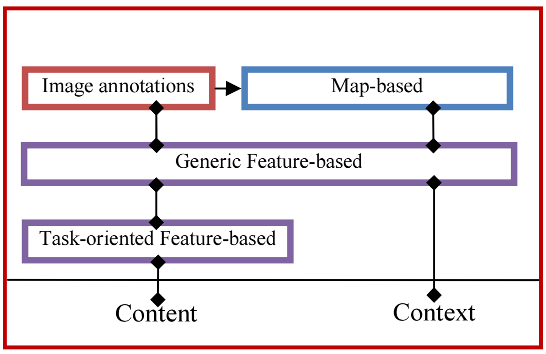


Figure 14: Content and the context information in the ontology-based label extraction

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

9. OPEN RESEARCH ISSUES

Overall, the existing ontology-based map-based approach has the best and widest applicability. Experimental evaluations have also confirmed the importance of labels in successfully executing this semantics approach in the image domain [21, 18]. The advantage of this approach lies in its independence from any given domain and in its implementation, which is based on upper-level ontologies. This approach can be implemented using different datasets, and depends mainly on contextual information. Measuring the similarities among the identified concepts (in terms of semantics) and the subsequent filtering process implicitly depend on identifying the common semantics among the input maps. Thus, only the maps that are closely related to the identified common semantics are retained, and the rest are discarded.

10. CONCLUSION

A novel categorization and comparison mechanism for the applications of ontology-based label extraction is proposed. A number of applications are categorized and analyzed based on the utilized ontology and the label extraction process. The existing approaches are categorized into feature- and map-based label extractions. Feature-based label extraction is further categorized into task-oriented and generic approaches, whereas map-based label extraction is categorized into existing and customized ontology-based approaches.

Based on the discussion, the task oriented feature-based sub-category utilizes content information only. The generic feature-based sub-category covers a wide domain and utilizes both content and context information. However, the latter approach has no actual implementation. In comparison, the map-based category utilizes contextual information and leaves the contents to be utilized via the annotation process. Most of the existing annotation techniques utilize content alone because this type of information does not require the use of any kind of knowledge.

Finally, among the reviewed approaches and categories, the existing ontology-based map-based category shows potential and wide applicability in addressing the image ambiguity problem. Moreover, the joint processing of the content elements in the image domain has been used efficiently in label disambiguation.

ACKNOWLEDGMENT

This work is supported by the Universiti Sains Malaysia's Research University grant "APEX DE (Grant No: 1002/PKOMP/910304)".

REFERENCES

[1] A. A. Abu-Shareha and R. Mandava, *Semantics Extraction in Visual Domain Based on WordNet*, *5th FTRA International Conference on Multimedia and Ubiquitous Engineering (MUE)*, IEEE, 2011, pp. 212-219.

[2] M. Atika, A. Akbar and A. Sultan, *Knowledge Discovery using Text Mining: A Programmable Implementation on Information Extraction and Categorization*, International Journal of Multimedia and Ubiquitous Engineering, 4 (2009).

[3] S. Banerjee and T. Pedersen, *Extended Gloss Overlaps as a Measure of Semantic Relatedness*, *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, Morgan Kaufmann Publishers Inc., Acapulco, Mexico, 2003, pp. 805-810.

[4] S.-F. Chang, T. Sikora and A. Purl, *Overview of the MPEG-7 standard*, IEEE Transactions on Circuits and Systems for Video Technology, 11 (2001), pp. 688-695.

[5] R. Clouard, A. Renouf and M. Revenu, *An Ontology-Based Model for Representing Image Processing Application Objectives.*, International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), 24 (2010), pp. 1181-1208.

[6] T. M. Deserno, S. Antani and R. Long, *Ontology of gaps in content-based image retrieval*, Journal of Digital Imaging, 22 (2009), pp. 202-215.

[7] H. J. Escalante, M. Montes and E. Sucar, *Multimodal indexing based on semantic cohesion for image retrieval*, Information retrieval, 15 (2012), pp. 1-32.

[8] X. Fan, *Contextual disambiguation for multi-class object detection*, *International Conference on Image Processing*, 2004, pp. 2873-2876.

[9] C. Fellbaum, *WordNet: An Electronic Lexical Database*, MIT Press (1998).

[10] D. Foxvog, *Cyc*, *Theory and Applications of Ontology: Computer Applications*, Springer, 2010, pp. 259-278.

[11] C. Galleguillos, A. Rabinovich and S. Belongie, *Object categorization using co-occurrence, location and appearance*, *IEEE*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008, 2008, pp. 1-8.

[12] E. Ganea and M. Brezovan, *An Object Oriented Graph Approach for Image Representation and Query based on Content*, International Journal of Computer Science & Applications, 7 (2010), pp. 45-60.

[13] H. Gaurav, C. Santanu and G. Hiranmay, *Using Multimedia Ontology for Generating Conceptual Annotations and Hyperlinks in Video Collections*, *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, IEEE Computer Society, Hong Kong, 2006, pp. 211-217.

[14] H. Ghosh, G. Harit and S. Chaudhury, *Ontology based Interaction with Multimedia Collections*, *International Conference on Digital Libraries*, New Delhi, 2006.

[15] M. Graf, *Form, Space and Object: Geometrical Transformations in Object Recognition and Categorization*, Wissenschaftlicher Verlag Berlin 2002.

[16] A. Hanbury, *A survey of methods for image annotation*, Journal of Visual Languages & Computing, 19 (2008), pp. 617-627.

[17] C. Hentschel, S. Stober, A. Nrnberger and M. Detyniecki, *Automatic Image Annotation Using a Visual Dictionary Based on Reliable Image Segmentation*, *5th International Workshop on Adaptive Multimedial Retrieval: Retrieval, User, and Semantics (AMR 2007)*, Springer-Verlag, Paris, France, 2008, pp. 45-56.

[18] N. James and C. Hudelot, *Towards Semantic Image Annotation With Keyword Disambiguation Using Semantic and Visual Knowledge In Proceedings of the IJCAI workshop on Cross-media Information Access and Mining (CIAM 2009)*, California, USA, 2009, pp. 35.

[19] J. J. Jiang and D. W. Conrath, *Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy*, *International Conference Research on Computational Linguistics (ROCLING X)*, Taiwan, 1997, pp. 900-908.

[20] Y. Jin, L. Khan and B. Prabhakaran, *Knowledge Based Image Annotation Refinement*, Journal of Signal Processing Systems, 58 (2010), pp. 387-406.

[21] Y. Jin, L. Khan, L. Wang and M. Awad, *Image Annotations By Combining Multiple Evidence & WordNet*, *Proceedings of the 13th Annual ACM International Conference on Multimedia*, ACM, Singapore, 2005, pp. 706-715.

[22] M. Johnson and R. Cipolla, *Improve Image Annotation and Labeling through Multi-Label Boosting*, *Proceedings of 2005 British Machine Vision Conference (BMVC '05)*, Oxford, UK, 2005.

[23] C. Leacock and M. Chodorow, *Combining Local Context and WordNet Similarity for Word Sense Identification*, in C. Fellbaum, ed., *WordNet: A Lexical Reference System and its Application*, MIT Press, 1998, pp. 305-332.

[24] D. Lin, *Using Syntactic Dependency as Local Context to Resolve Word Sense Ambiguity*, *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, Madrid, Spain, 1997, pp. 64-71.

[25] Y. Liu, D. Zhang, G. Lu and W.-Y. Ma, *A survey of content-based image retrieval with high-level semantics*, Pattern Recognition, 40 (2007), pp. 262-282.

[26] Y. Liu, J. Zhang, D. Tjondronegoro and S. Geve, *A Shape Ontology Framework for Bird Classification*, *Proceedings of the 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications*, IEEE Computer Society, 2007, pp. 478-484.

[27] Z. Liu, *A novel graph-based image annotation refinement algorithm*, *Proceedings of the 6th international conference on Fuzzy systems and knowledge discovery - Volume 5*, IEEE Press, Tianjin, China, 2009.

[28] A. Mallik and S. Chaudhury, *Using concept recognition to annotate a video collection*, *Pattern Recognition and Machine Intelligence*, Springer, 2009, pp. 507-512.

[29] A. Mallik, S. Chaudhury, S. Madan, T. Dinesh and U. V. Chandru, *Archiving mural paintings using an ontology based approach*, *Computer Vision-ACCV 2012 Workshops*, Springer, 2013, pp. 37-48.

[30] A. Mallik, P. Pasumarthi and S. Chaudhury, *Multimedia ontology learning for automatic annotation and video browsing*, *Proceedings of the 1st ACM international conference on Multimedia information retrieval*, ACM, 2008, pp. 387-394.

[31] R. Navigli and P. Velardi, *Structural Semantic Interconnections: A Knowledge-Based Approach to Word Sense Disambiguation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (2005), pp. 1075-1086.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[32] S. Nesic, F. Crestani, M. Jazayeri and D. Gasevic, *Concept-Based Semantic Annotation, Indexing and Retrieval of Office-Like Document Units*, *Proceeding of Adaptivity, Personalization and Fusion of Heterogeneous Information (RIAO '10)*, Paris, France, 2010, pp. 134-135.

[33] M. Nicolas Eric and T. Monique, *Ontology based complex object recognition*, Image Vision Comput., 26 (2008), pp. 102-113.

[34] K.-W. Park and D.-H. Lee, *Full-Automatic High-Level Concept Extraction from Images Using Ontologies and Semantic Inference Rules*, *ASWC*, 2006, pp. 307-321.

[35] K.-W. Park and D.-H. Lee, *Full-Automatic High-Level Concept Extraction from Images Using Ontologies and Semantic Inference Rules*, in R. Mizoguchi, Z. Shi and F. Giunchiglia, eds., *The Semantic Web*, Springer Berlin / Heidelberg, 2006, pp. 307-321.

[36] A. Penta, A. Picariello and L. Tanca, *Towards a definition of an Image Ontology*, *18th International Conference on Database and Expert Systems Applications*, 2007, pp. 74-78.

[37] D. Pisanelli, *Biodynamic ontology: applying BFO in the biomedical domain*, Ontologies in medicine, 102 (2004), pp. 20.

[38] P. Resnik, *Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language*, Journal of Artificial Intelligence Research, 11 (1999), pp. 95-130.

[39] K. Sanjiv and H. Martial, *A Hierarchical Field Framework for Unified Context-Based Classification*, *Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2*, IEEE Computer Society, 2005.

[40] R. Sharman, R. Kishore, R. Ramesh, H. Ghosh, S. Chaudhury, K. Kashyap and B. Maiti, *Ontology Specification and Integration for Multimedia Applications*, in R. Sharda and S. Voß, eds., *Ontologies*, Springer US, 2007, pp. 265-296.

[41] A. Singhal, J. Luo and W. Zhu, *Probabilistic Spatial Context Models for Scene Content Understanding*, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA, USA, 2003, pp. 1063-6919.

[42] C. Thies, H. Herzog, T. Schmitz-Rode and T. M. Deserno, *Bridging the semantic gap for object extraction from biomedical images by classification*, Biomedizinische Technik, 52 (2007).

[43] A.-M. Tousch, S. Herbin and J.-Y. Audibert, *Semantic hierarchies for image annotation: A survey*, Pattern Recognition, 45 (2012), pp. 333-345.

[44] C. Town, *Ontological Inference for Image and Video Analysis*, Machine Vision and Applications, 17 (2006), pp. 94-115.

[45] C. Tsinaraki, P. Polydoros and S. Christodoulakis, *Integration of OWL Ontologies in MPEG-7 and TV-Anytime Compliant Semantic Indexing*, in A. Persson and J. Stirna, eds., *Advanced Information Systems Engineering*, Springer Berlin / Heidelberg, 2004, pp. 143-161.

[46] C. Tsinaraki, P. Polydoros, F. Kazasis and S. Christodoulakis, *Ontology-Based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content*, Multimedia Tools and Applications, 26 (2005), pp. 299-325.

[47] Y. Wu, B. L. Tseng and J. R. Smith, *Ontology-based multi-classification learning for video concept detection*, *IEEE International Conference on Multimedia and Expo, ICME '04*, 2004, pp. 1003-1006.

[48] D. Zhang, M. M. Islam and G. Lu, *A review on automatic image annotation techniques*, Pattern Recognition, 45 (2012), pp. 346-362.

[49] S. Zinger, C. Millet, B. Mathieu, G. Grefenstette, P. Hède and P.-A. Moëllic, *Extracting an ontology of portrayable objects from WordNet*, *MUSCLE/ImageCLEF workshop on Image and Video retrieval evaluation*, 2005, pp. 17-23.

[50] N. Zlatoff, B. Tellez and A. Baskurt, *Image understanding and scene models: a generic framework integrating domain knowledge and Gestalt theory*, *Image Processing, 2004. ICIP '04. 2004 International Conference on*, 2004, pp. 2355-2358 Vol. 4.

[51] N. Zlatoff, B. Tellez and A. Baskurt, *Image Understanding and Scene Models: A Generic Framework Integrating Domain Knowledge and Gestalt Theory*, *International Conference on Image Processing (ICIP)*, 2004, pp. 2355-2358.

Journal of Theoretical and Applied Information Technology
20th January 2015. Vol.71 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

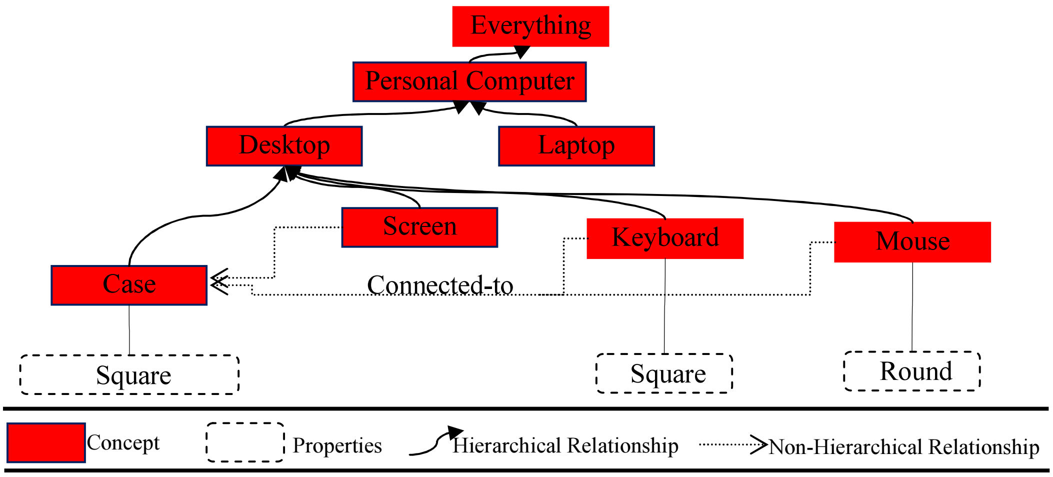


Figure 1: Example of an ontology

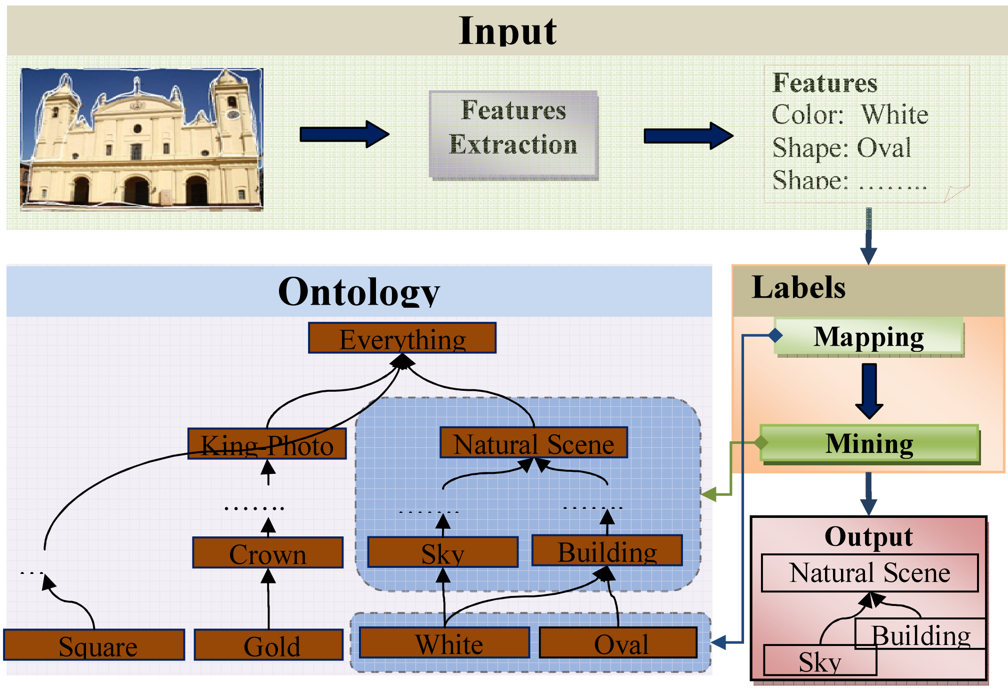


Figure 2: Example of the ontology-based label extraction process

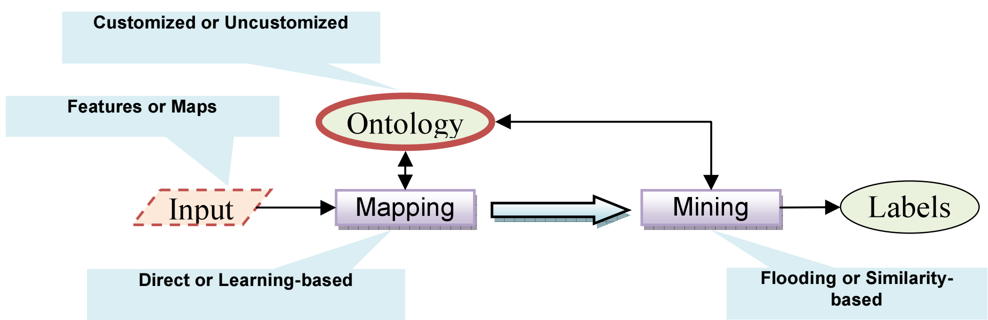


Figure 3: Variability criteria of ontology-based label extraction

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

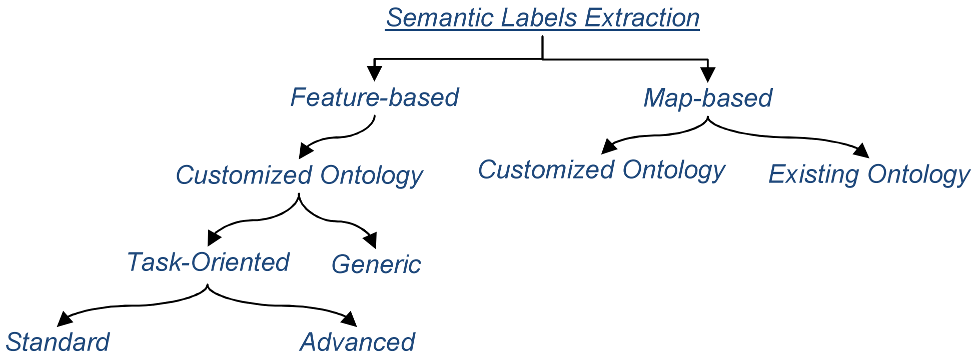


Figure 4: Categorization of label extraction

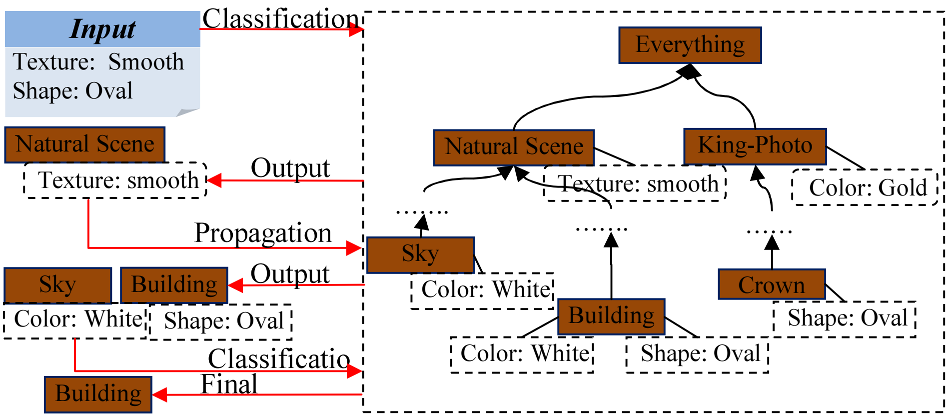


Figure 7: Processes used in the standard approach from the task-oriented category

Table 1: Comparison among the ontology-based label extraction mechanisms

	Feature-based/ Task-oriented	**Feature-based/ Generic**	**Map-based / Customized Ontology**	**Map-based/ Existing Ontology**
Domain	Dependent	Independent	Dependent	Independent
Task	Dependent	Independent	Dependent	Dependent
Mapping	Classification	None	Direct	Direct
Mining	Propagation	None	Pair-wise Similarity	Pair-wise Similarity