Exploiting Visual Similarities for Ontology Alignment

Charalampos Doulaverakis, Stefanos Vrochidis and Ioannis Kompatsiaris
Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece

Keywords: Ontology Alignment, Visual Similarity, ImageNet, Wordnet.

Abstract: Ontology alignment is the process where two different ontologies that usually describe similar domains are 'aligned', i.e. a set of correspondences between their entities, regarding semantic equivalence, is determined. In order to identify these correspondences several methods and metrics that measure semantic equivalence have been proposed in literature. The most common features that these metrics employ are string-, lexical-, structure- and semantic-based similarities for which several approaches have been developed. However, what hasn't been investigated is the usage of visual-based features for determining entity similarity in cases where images are associated with concepts. Nowadays the existence of several resources (e.g. ImageNet) that map lexical concepts onto images allows for exploiting visual similarities for this purpose. In this paper, a novel approach for ontology matching based on visual similarity is presented. Each ontological entity is associated with sets of images, retrieved through ImageNet or web-based search, and state of the art visual feature extraction, clustering and indexing for computing the similarity between entities is employed. An adaptation of a popular Wordnet-based matching algorithm to exploit the visual similarity is also proposed. Our method is compared with traditional metrics against a standard ontology alignment benchmark dataset and demonstrates promising results.

1 INTRODUCTION

Semantic Web is providing shared ontologies and vocabularies in different domains that can be openly accessed and used for tasks such as semantic annotation of information, reasoning, querying, etc. The Linked Open Data (LOD) paradigm shows how the different exposed datasets can be linked in order to provide a deeper understanding of information. As each ontology is being engineered to describe a particular domain for usage in specific tasks, it is common for ontologies to express equivalent domains using different terms or structures. These equivalences have to be identified and taken into account in order to enable seamless knowledge integration. Moreover, as an ontology can contain hundreds or thousands of entities, there is a need to automate this process. An example of the above comes from the cultural heritage domain where two ontologies are being used as standards, one is the CIDOC-CRM[1], used for semantically annotating museum content, and the other is the Europeana Data Model[2], which is used to semantically index and interconnect cultural heritage objects. While these two ontologies have been developed for different purposes, they are used in the cultural heritage domain and correspondences between their entities should exist and be identified.

In ontology alignment the goal is to automatically or semi-automatically discover correspondences between the ontological entities, i.e. their classes, properties or instances. An 'alignment' is a set of mappings that define the similar entities between two ontologies. These mappings can be expressed e.g. using the *owl:equivalentClass* or *owl:equivalentProperty* properties so that a reasoner can automatically access both ontologies during a query.

While the proposed methodologies in literature have proven quite effective, either alone or combined, in dealing with the alignment of ontologies, there has been little progress in defining new similarity metrics that take advantage of features that haven't been considered so far. In addition existing benchmarks for evaluating the performance of ontology alignments systems, such as the Ontology Alignment Evaluation Initiative[3] (OAEI) have shown that there is still room for improvement in ontology alignment.

In the last 5 years the proliferation of multime-

[1] CIDOC-CRM, http://www.cidoc-crm.org
[2] Europeana Data Model, http://labs.europeana.eu
[3] OAEI, http://oaei.ontologymatching.org

Doulaverakis, C., Vrochidis, S. and Kompatsiaris, I..
Exploiting Visual Similarities for Ontology Alignment.
In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015) - Volume 2: KEOD*, pages 29-37
ISBN: 978-989-758-158-8
Copyright © 2015 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

dia has generated several annotated resources and datasets that are associated with concepts, such as ImageNet[4] or Flickr[5] thus making their visual representations easily available and retrievable so that they can be further exploited, e.g. for image recognition.

In this paper we propose a novel ontology matching metric that is based on visual similarities between ontological entities. The visual representations of the entities are crafted by different multimedia sources, namely ImageNet and web-based image search, thus assigning each entity to descriptive sets of images. State of the art visual features are extracted from these images and vector representations are generated. The entities are compared in terms of these representations and a similarity value is extracted for each pair of entities, thus the pair with the highest similarity value is considered as valid. The approach is validated in experimental results where it is shown that when it's combined with other known ontology alignment metrics it increases precision and recall of the discovered mappings.

The main contribution of the paper is the introduction of a novel similarity metric for ontology alignment based on visual features. To the best of the authors knowledge this is the first attempt to exploit visual features for ontology alignment purposes. We also propose an adaptation of a popular lexical-based matching algorithm where lexical similarity is replaced with visual similarity.

The paper is organized as follows: Section 3 describes the methodology in detail, while Section 5 presents the experimental results on the popular OAEI conference track dataset. In Section 4 an metric that exploits the proposed visual similarity and lexical features is proposed and described. Related work in ontology alignment is documented in Section 2. Finally, Section 6 concludes the paper and a future work plan is outlined.

2 RELATED WORK

In order to accomplish the automatic discovery of mappings, numerous approaches have been proposed in literature that rely on various features. Of the most common are methods that compare the similarity of two strings, e.g. comparing *hasAuthor* with *isAuthoredBy*, are the most used and fastest to compute as they operate on raw strings. Existing string similarity metrics are being used, such as Levenshtein distance, Edit distance, Jaro-Winkler similarity, etc, while string similarity algorithms such as (Stoilos et al., 2005) have been developed especially for ontology matching. Other mapping discovery methods rely on lexical processing in order to find synonyms, hypernyms or hyponyms between concepts, e.g. *Author* and *Writer*, where Wordnet is most commonly used. In (Lin and Sandkuhl, 2008) a survey on methods that use Wordnet (Miller, 1995) for ontology alignment, is carried out. Approaches for exploiting other external knowledge sources have been presented (Sabou et al., 2006; Pesquita et al., 2014; Chen et al., 2014; Faria et al., 2014). Other similarity measures rely on the structure of the ontologies, such as the Similarity Flooding (Melnik et al., 2002) algorithm that stems from the relational databases world but has been successfully used for ontology alignment, while others exploit both schema and ontology semantics for mapping discovery. A comprehensive study of such methods can be found at (Shvaiko and Euzenat, 2005). In terms of matching systems, there have been proposed numerous approaches that combine matchers or include external resources of the generation of a valid mapping between ontologies. Most available systems have been evaluated in the OAEI benchmarks that are held annually. In (Jean-Mary et al., 2009) the authors use a weighted approach to combine several matchers in order to produce a final matching score between the ontological entities. In (Ngo and Bellahsene, 2012) the authors go a step further and propose a novel approach to combine elementary matching algorithms using a machine learning approach with decision trees. The system is trained from prior ground truth alignments in order to find the best combination of matchers for each pair of entities. Other systems, such as AML (Faria et al., 2013) and (Kirsten et al., 2011), make use of external knowledge resources or lexicons to obtain ground truth structure and entity relations. This is especially used when matching ontologies in specialized domains such as in biomedicine.

In contrast to the above we propose a novel ontology matching algorithm that corresponds entities with images and makes use of visual features in order to compute similarity between entities. To the authors knowledge, this is the first approach in literature where a visual-based ontology matching algorithm is proposed. Throughout the paper, the term “entity” is used to refer to ontology entities, i.e. classes, object properties, datatype properties, etc.

[4] ImageNet, http://www.image-net.org/

[5] Flickr, https://www.flickr.com/

(a) Images for "boat" (b) Images for "ship" (c) Images for "motorbike"

Figure 1: Images for different synsets. (a) and (b) are semantically more similar than with (c). The visual similarity between (a) and (b) and their difference with (c) is apparent.

3 VISUAL SIMILARITY FOR ONTOLOGY ALIGNMENT

The idea for the development of a visual similarity algorithm for ontology alignment originated from the structure of ImageNet where images are assigned to concepts. For example, Figure 1 shows a subset of images that is found in ImageNet for the words *boat*, *ship* and *motorbike*. Obviously, *boat* and *ship* are more semantically related than *boat* and *motorcycle*. It is also clear from Figure 1 that the images that correspond to *boat* and *ship* are much more similar in terms of visual appearance than the images of *motorbike*. One can then assume that it is possible to estimate the semantic relatedness of two concepts by comparing their visual representations.

In Figure 2 the proposed architecture for visual-based ontology alignment is presented. The source and target ontologies are the ontologies to be matched. For every entity in the ontologies, sets of images are assigned through ImageNet by identifying the relevant Wordnet synsets. A synset is a set of words that have the same meaning and these are used to query ImageNet. A single entity might correspond to a number of synsets, e.g. "track" has different meaning in transport and in sports as can be seen in Figure 3. Thus for each entity a number of image sets are retrieved. For each image in a set, low level visual features are extracted and a numerical vector representation is formed. Therefore for each concept different sets of vectors are generated. Each set of vectors is called a "visual signature". All visual signatures between the source and target ontology are compared in pairs using a modified Jaccard set similarity in order to come up with a list of similarity values assigned to each entity pair. The final list of mappings is generated by employing an assignment optimization algorithm such as the Hungarian method (Kuhn, 1955).

3.1 Assigning Images to Entities

The main source of images in the proposed work is ImageNet, an image database organized according to the WordNet noun hierarchy in which each node of the hierarchy is associated with a number of images. Users can search the database through a text-search web interface where the user inputs the query words, which are then mapped to Wordnet indexed words and a list of relevant synsets (synonym sets, see (Miller, 1995)) are presented. The user selects the desired synset and the corresponding images are displayed. In addition, ImageNet provides a REST API for retrieving the image list that corresponds to a synset by entering the Wordnet synset id as input and this is the access method we used.

For every entity of the two ontologies to be matched, the following process was followed: A preprocessing procedure is executed where each entity name is first tokenized in order to split it to meaningful words as it is common for names to be in the form of *isAuthorOf* or *is_author_of* thus after tokenization, *isAuthorOf* will be split to the words *is*, *Author* and *of*. The next step is to filter out stop words, words that do not contain important significance or are very common. In the previous example, the words *is* and *of* are removed, thus after this preprocessing the name that is produced is *Author*.

After the preprocessing step, the next procedure is about identifying the relevant Worndet synset(s) of the entity name and get their ids, which is a rather straightforward procedure. Using these ids, ImageNet is queried in order to retrieve a fixed number of relevant images. However trying to retrieve these images

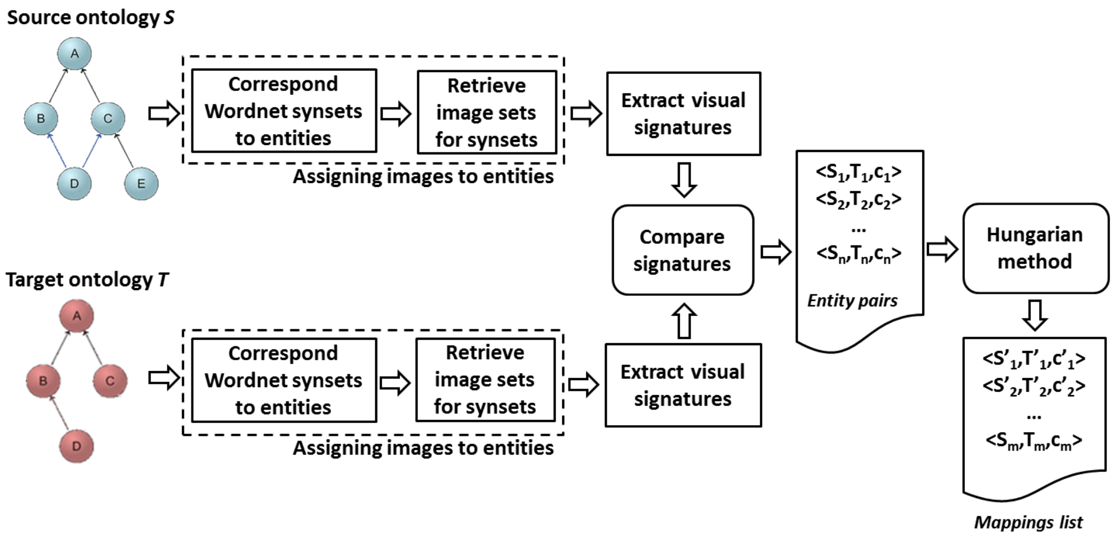


Figure 2: Architecture of the proposed ontology alignment algorithm.

(a) Images for "track (running)" (b) Images for "track (train)"

Figure 3: Images that correspond to different meanings of concept "track". Since we can't be certain of a word meaning (word sense), each concept is associated with all relevant synsets and corresponding image sets from ImageNet.

might fail, mainly due to two reasons: either the name does not correspond to a Wordnet synset, e.g. due to misspellings, or the relevant ImageNet synset isn't assigned any images, something which is not uncommon since ImageNet is still under development and is not complete. So, in order not to end up with empty image collections, in the above cases the entity name is used to query Yahoo™ image search[6] in order to find relevant images. The idea of using web-based search results has been employed in computer vision as in (Chatfield and Zisserman, 2013) where web image search is used to train an image classifier.

The result of the above-described process is to have each ontological entity C associated with n sets of images I_{iC}, with $i = 1, \dots, n$, where n is the number of synsets that correspond to entity C.

[6] Yahoo search, https://images.search.yahoo.com

3.2 Extracting the Visual Signatures of Entities

For allowing a visual-based comparison of the ontological entities, each image set I_{iC} has to be represented using appropriate visual descriptors. For this purpose, a state of the art approach is followed where images are represented as compact numerical vectors. For extracting these vectors the approach which is described in (Spyromitros-Xioufis et al., 2014) is used as it has been shown to outperform other approaches on standard benchmarks of image retrieval and is quite efficient. In short, SURF (Speeded Up Robust Features) descriptors (Bay et al., 2008) are extracted for each image in a set. SURF descriptors are numerical representations of important image features and are used to compactly describe image content. These are then represented using the VLAD (Vector of Locally Aggregated Descriptors) representation (Jégou et al., 2010) where four codebooks of size 128 each, were used. The resulting VLAD vectors are PCA-projected to reduce their dimensionality to 100 coefficients, thus ending up with a standard numerical vector representation v_j for each image j in a set. At the end of this process, each image set I_{iC} will be numerically represented by a corresponding vector set. This vector set is termed "visual signature" V_{iC} as it conveniently and descriptively represents the visual content of I_{iC}, thus $V_{iC} = \{v_j\}$, with $j = 1, \dots, k$ and k being the total number of images in I_{iC}.

The whole processing workflow is depicted in Figure 4.

Algorithm 1 outlines the steps to create visual signatures V_C of entities in an ontology.

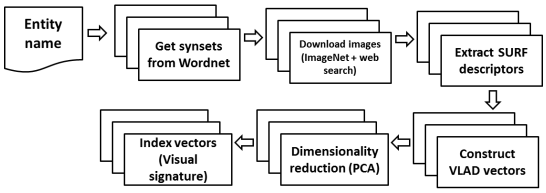


Figure 4: Block diagram of the process for extracting the visual signatures of an entity.

Algorithm 1: Pseudocode for extracting visual signature V_C of an entity C in ontology O.

Ensure: $V_C = \emptyset, C$ is an entity of ontology O
  $C_t \leftarrow removeStopWords(tokenize(C))$
  $W \leftarrow$ find Wordnet synsets of C_t
  for all synsets W_i in W **do**
    $I_{iC} \leftarrow$ download k images from ImageNet
    if $I_{iC} = \emptyset$ **then**
      download k images from web
    end if
    $V_{iC} \leftarrow \emptyset$
    for all images j in I_{iC} **do**
      $v_j \leftarrow extractVisualDescriptors(j)$
      $V_{iC} \leftarrow$ add v_j
    end for
    $V_C \leftarrow$ add V_{iC}
  end for
  return V_C

3.3 Comparing Visual Signatures for Computing Entity Similarity

Having the visual signatures for each entity, the next step is to use an appropriate metric in order to compare these signatures and estimate the similarity between image sets. Several vector similarity and distance metrics exist, such as cosine similarity or euclidean distance, however these are mostly suitable when comparing individual vectors. In the current work, we are interested in establishing the similarity value between vector sets so the Jaccard set similarity measure is more appropriate as it is has been defined exactly for this purpose. It's definition is

$$J_{V_{iCs},V_{jCt}} = \frac{|V_{iCs} \cap V_{jCt}|}{|V_{iCs} \cup V_{jCt}|} \quad (1)$$

where V_{iCs} and V_{jCt} are the i and j different visual signatures of entities C_s and C_t, $|V_{iCs} \cap V_{jCt}|$ is the intersection size of the two sets, i.e. the number of identical images between the sets, and $|V_{iCs} \cap V_{jCt}|$ is the total number of images in both sets. It holds that $0 \leq J_{V_{iCs},V_{jCs}} \leq 1$. For defining if two images A and B are identical, we compute the angular similarity of their vector representations.

$$AngSim_{A,B} = 1 - \frac{\arccos(cosineSim(A,B))}{\pi} \quad (2)$$

with $cosineSim(A,B)$ equal to

$$cosineSim(A,B) = \frac{\sum_{k=1}^{n=100} A_k \cdot B_k}{\sqrt{\sum_{k=1}^{n=100} A_k^2} \cdot \sqrt{\sum_{k=1}^{n=100} B_k^2}} \quad (3)$$

For $AngSim$, a value of 0 means that the two images are completely irrelevant and 1 means that they are identical. However, two images might not have $AngSim_{A,B} = 1$ even if they are visually the same but they are acquired from different sources due to e.g. differences in resolution, compression or stored format, thus we risk of having $|V_{iCs} \cap V_{jCt}| = \emptyset$. For this reason instead of aiming to find truly identical images we introduce the concept of "near-identical images" where two images are considered identical if the have a similarity value above a threshold T, thus

$$Identical_{A,B} = \begin{cases} 0 & \text{if } AngSim_{A,B} < T \\ 1 & \text{if } AngSim_{A,B} \geq T \end{cases} \quad (4)$$

T is experimentally defined. Using the above we are able to establish the Jaccard set similarity value of two ontological entities by corresponding each entity to an image set, extracting the visual signature of each set and comparing these signatures. The Jaccard set similarity value J_{V_i,U_j} is computed for every pair i, j of synsets that correspond to the examined entities, V, U. Visual Similarity is defined as

$$VisualSim(C_s,C_t) = \max_{i,j}(J_{V_{iCs},V_{jCt}}) \quad (5)$$

4 COMBINING VISUAL AND LEXICAL FEATURES

The Visual Similarity algorithm can either be exploited as a standalone measure or it can be used as complementary to other ontology matching measures as well. Since in order to construct the visual representation of entities Wordnet is used, one approach is to combine visual with lexical-based features. Lexical-based measures have been used in ontology matching systems in recent OAEI benchmarks, such as in (Ngo and Bellahsene, 2012)where, among others, the Wu-Palmer (Wu and Palmer, 1994) Wordnet-based measure has been integrated. The Wu-Palmer similarity value between concepts C_1 and C_2 is defined as

$$WuPalmer_{C_1,C_2} = \frac{2 \cdot N_3}{N_1 + N_2 + 2 \cdot N_3} \quad (6)$$

where C_3 is defined as the least common superconcept (or hypernym) of both C_1 and C_2, N_1 and N_2 are the number of nodes from C_1 and C_2 to C_3, respectively, and N_3 is the number of nodes on the path from C_3 to root. The intuition behind this metric is that since concepts closer to the root have a broader meaning which is made more specific as one moves to the leaves of the hierarchy, if two concepts have a common hypernym closer to them and further from the root, then it's likely that they have a closer semantic relation.

Based on this intuition we have defined a new similarity metric that takes into account the visual features of both concepts and of their least common superconcept. Using the same notation and meaning for C_1, C_2, C_3, the measure we have defined is expressed as

$$LexiVis_{C_1,C_2} = \frac{V_3}{3 - (V_1 + V_2)} \quad (7)$$

where V_3 is the visual similarity value between C_1 and C_2 and V_1,V_2 are the visual similarity values between C_1,C_3 and C_2,C_3 respectively. V_1,V_2 and V_3 are calculated according to Eq. 5. In all cases, $0 \leq LexiVis_{C_1,C_2} \leq 1$. The intuition behind this measure is that semantically related concepts will be each other highly visually similar to each other and also highly similar visually with their closest hypernym. The incorporation of the closest hypernym in the overall similarity estimation of two concepts will allow for corrections in cases where concepts might be visually similar but semantically irrelevant, e.g. "boat" and "hydroplane" pictures depict an object surrounded by a body of water, however when they are visually compared against their common superconcept, in the previous example it is the concept "craft", their pair-wise visual similarity value will be low thus lowering the concepts' similarity. This example is depicted in Figure 5.

Boat Hydroplane Craft

Boat-Hydroplane=0.49, Boat-Craft=0.24,
Hydroplane-Craft=0.35

LexiVis (Boat-Hydroplane) = 0.20

Figure 5: Visual similarity values between the concepts "Boat" and "Hydroplane" which are semantically irrelevant but visually similar. Their common hypernym is "Craft". The LexiVis measure, by taking advantage of lexical features, lowers their similarity value.

5 EXPERIMENTAL RESULTS

For analyzing the performance of the Visual Similarity ontology matching algorithm we ran it against the Ontology Alignment Evaluation Initiative (OAEI) Conference track of 2014 (Dragisic et al., 2014)[7]. The OAEI benchmarks are organized annually and have become a standard in ontology alignment tools evaluation. In the conference track, a number of ontologies that are used for the organization of conferences have to be aligned in pairs. The conference track was chosen as, by design, the proposed algorithm requires meaningful entity names that can be visually represented. Other tracks, such as benchmark and anatomy, weren't considered due to this limitation which is further discussed in Section 6. Reference alignments are available and these are used for the actual evaluation in an automated manner. The reference alignment that was used is "ra1" since this was readily available for the OAEI 2014 website.

The *VisualSim* and *LexiVis* ontology matching algorithms were integrated in the Alignment API (Euzenat, 2004) which offers interfaces and sample implementations in order to integrate matching algorithms. The API is recommended from OAEI for participating in the benchmarks. In addition, algorithms to compute standard information retrieval measures, i.e. precision, recall and F-measure, against reference alignments can be found in the API, so these were used for the evaluation of the tests results. In these tests we changed the threshold, i.e. the value under which an entity matching is discarded, and registered the precision, recall and $F1$ measure values.

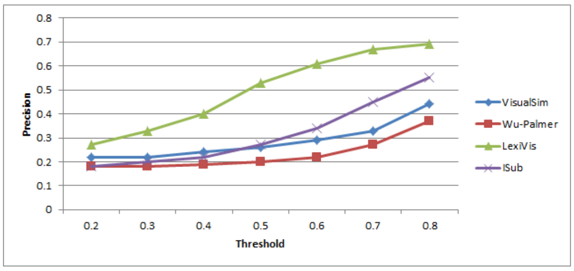
In order to have a better understanding of the proposed algorithms we compared it against other popular matching algorithms. Ideally the performance of these would be evaluated against other matching algorithms that make use of similar modalities, i.e. visual or other. This wasn't feasible as the proposed algorithms are the first that makes use of visual features, so we compare it with standard algorithms that exploit traditional features such as string-based and Wordnet-based similarity. For this purpose we implemented the ISub string similarity matcher (Stoilos et al., 2005) and the Wu-Palmer Wordnet-based matcher which is described in Section 4. These matchers have been used in the YAM++ ontology matching system (Ngo and Bellahsene, 2012) which was one of the top ranked systems in OAEI 2012.

All aforementioned algorithms, ISub, Wu-Palmer, VisualSim and LexiVis, are evaluated using Precision, Recall and $F1$ measure, with

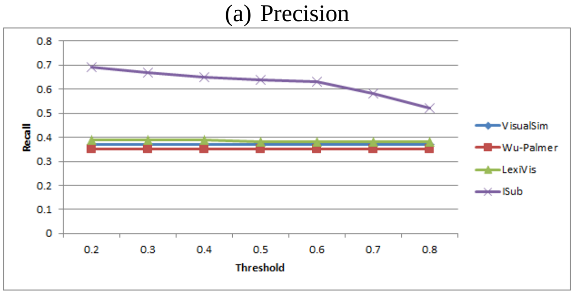
[7]OAEI 2014,http://oaei.ontologymatching.org/2014/

Table 1: Performance of the LexiVis matching algorithm in combination with other matching algorithms (ISub, Name Equality, Similarity Flooding (Melnik et al., 2002)), and how the performance is compared to matching systems that participated in OAEI 2014 conference track.

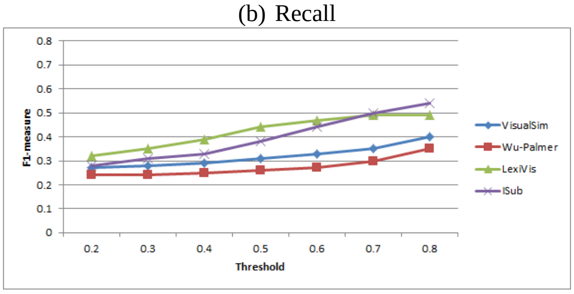
System	**Precision**	**Recall**	**F1-measure**
AML	0.85	0.64	0.73
LogMap	0.80	0.59	0.68
LogMap-C	0.82	0.57	0.67
XMap	0.87	0.49	0.63
NameEq + ISub + SimFlood + LexiVis	*0.71*	*0.53*	*0.60*
NameEq + ISub + SimFlood	0.81	0.47	0.59
OMReasoner	0.82	0.46	0.59
Baseline (NameEq)	0.80	0.43	0.56
AOTL	0.77	0.43	0.55
MaasMtch	0.64	0.48	0.55



(a) Precision



(b) Recall



(c) F1 measure

Figure 6: Precision, Recall and F1 diagrams for different threshold values using the conference track ontologies of OAEI 2014.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (8)$$

The results of this evaluation are displayed in Figure 6.

It can be seen from Figure 6 that *VisualSim* and the *LexiVis* algorithms performs better in all measures than the Wu-Palmer alignment algorithm which confirms with our initial assumption that the semantic similarity between entities can be reflected in their visual representation using imaging modalities. This allows a new range of matching techniques based on modalities that haven't been considered so far to be investigated. However, the string-based ISub matcher displays superior performance, which was expected as string-based matchers are very effective in ontology alignment and matching problems, which points out that the aforementioned new range of matchers should work complementary to the existing and established matchers as these have proven their reliability though time.

An additional performance factor that should be mentioned is the computational complexity and overall execution time for the Visual based algorithm which is much greater than the simpler string-based algorithms. Analyzing Figure 4, of all the documented steps by far the most time consuming are the image download and visual descriptor extraction. However, ImageNet is already offering visual descriptors which are extracted from the synset images and are freely available to download[8]. The range of images that have been processed is not yet complete but as ImageNet is still in development, the plan is to have the whole image database processed and have the visual descriptors extracted. This availability will make the calculation of the proposed visual-based ontology alignment algorithms faster.

5.1 In Combination with other Ontology Alignment Algorithms

As a further test, using the Alignment API we integrated the LexiVis matching algorithm and aggre-

[8]ImageNet visual features download, http://image-net.org/download-features

gated the matching results with other available matching algorithms in order to have an understanding on how it would perform in a real ontology matching system. We used the LexiVis algorithm as it was shown to perform better than the original Visual Similarity algorithm (Figure 6). The other algorithms that were used are the ISub and Similarity Flooding matchers in addition to the baseline NameEq matcher. These were used in order to have a combination of matchers that exploit different features, i.e. string, structural and visual. The matchers were combined using an adaptive weighting approach similar to (Cruz et al., 2009). For this test we again used the conference track benchmark dataset of OAEI 2014. For this dataset, results regarding the performance of the participating matching systems are published in OAEI's website and in (Dragisic et al., 2014). It can be seen from Table 1, in the line denoted with italic font, that the inclusion of the LexiVis ontology matching algorithm in the matching system results in better overall performance than running the system without it. The added value of 0.01 in $F1$ results in an overall $F1$ value of 0.60 which brings our matching system in the top 5 performances. The rather small added value of 0.01 is mainly due to the fact that the benchmark is quite challenging as can be seen from the results of Table 1. For example the XMap system, which is ranked 4^{th}, managed to score 0.07 more in $F1$ than the baseline NameEq matcher which simply compares strings and produces a valid pair if the names are equal. Even this small increase of $F1$ just by including the LexiVis algorithm proves that it can improve results in such a challenging benchmark thus showing its benefit.

6 CONCLUSIONS

In this paper a novel ontology matching algorithm which is based on visual features is presented. The algorithm exploits ImageNet's structure which is based on Wordnet in order to correspond image sets to the ontological entities and state of the art visual processing is employed which involves visual feature descriptors extraction, codebook-based feature representation, dimensionality reduction and indexing. The visual-based similarity value is taken by calculating a modified version of the Jaccard set similarity value. A new matcher is also proposed which combines visual and lexical features in order to determine entity similarity. The proposed algorithms have been evaluated using the established OAEI benchmark and has shown to outperform Wordnet-based approaches. A limitation of the proposed visual-based matching algorithm is that since it relies of visual depictions of entities, in cases where entity names are not words, e.g. alphanumeric codes, then its performance will be poor as images will be able to be associated with it. A way to tackle this is to extend the approach to include other data, such as *rdfs:label*, which are more descriptive. Another limitations of this approach would be the mapping of concepts that are visually hard to express, e.g. "Idea" or "Freedom", however this is partly leveraged by employing web-based search which likely retrieves relevant images for almost any concept.

The current version of the algorithm only uses entity names Future work will focus in optimizing the processing pipeline in order to have visual similarity results in a more timely manner using processing optimizations and other approaches such as word sense disambiguation in order to reduce the image sets that correspond to each entity.

ACKNOWLEDGEMENTS

This work was supported by MULTISENSOR (contract no. FP7-610411) and KRISTINA (contract no. H2020-645012) projects, partially funded by the European Commission.

REFERENCES

Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3):346–359.

Chatfield, K. and Zisserman, A. (2013). VISOR: Towards on-the-fly large-scale object category retrieval. In *Asian Conference of Computer Vision – ACCV 2012*, pages 432–446. Springer Berlin Heidelberg.

Chen, X., Xia, W., Jiménez-Ruiz, E., and Cross, V. (2014). Extending an ontology alignment system with bioportal: a preliminary analysis. In *Poster at Intl Sem. Web Conf.(ISWC)*.

Cruz, I. F., Palandri Antonelli, F., and Stroe, C. (2009). Efficient selection of mappings and automatic quality-driven combination of matching methods. In *ISWC International Workshop on Ontology Matching (OM) CEUR Workshop Proceedings*, volume 551, pages 49–60. Citeseer.

Dragisic, Z., Eckert, K., Euzenat, J., Faria, D., Ferrara, A., Granada, R., Ivanova, V., Jimenez-Ruiz, E., Kempf, A., Lambrix, P., et al. (2014). Results of the ontology alignment evaluation initiative 2014. In *International Workshop on Ontology Matching*, pages 61–104.

Euzenat, J. (2004). An API for ontology alignment. In *The Semantic Web–ISWC 2004*, pages 698–712. Springer.

Faria, D., Pesquita, C., Santos, E., Cruz, I. F., and Couto, F. M. (2014). Automatic background knowledge selection for matching biomedical ontologies. *PLoS ONE*, 9(11):e111226.

Faria, D., Pesquita, C., Santos, E., Palmonari, M., Cruz, I. F., and Couto, F. M. (2013). The agreementmakerlight ontology matching system. In *On the Move to Meaningful Internet Systems: OTM 2013 Conferences*, pages 527–541. Springer.

Jean-Mary, Y. R., Shironoshita, E. P., and Kabuka, M. R. (2009). Ontology matching with semantic verification. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3):235–251.

Jégou, H., Douze, M., Schmid, C., and Pérez, P. (2010). Aggregating local descriptors into a compact image representation. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 3304–3311. IEEE.

Kirsten, T., Gross, A., Hartung, M., and Rahm, E. (2011). Gomma: a component-based infrastructure for managing and analyzing life science ontologies and their evolution. *J. Biomedical Semantics*, 2(6).

Kuhn, H. W. (1955). The hungarian method for the assignment problem. *Naval research logistics quarterly*, 2(1-2):83–97.

Lin, F. and Sandkuhl, K. (2008). A survey of exploiting wordnet in ontology matching. In *Artificial Intelligence in Theory and Practice II*, pages 341–350. Springer.

Melnik, S., Garcia-Molina, H., and Rahm, E. (2002). Similarity flooding: A versatile graph matching algorithm and its application to schema matching. In *Data Engineering, 2002. Proceedings. 18th International Conference on*, pages 117–128. IEEE.

Miller, G. A. (1995). Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Ngo, D. and Bellahsene, Z. (2012). YAM++: A multi-strategy based approach for ontology matching task. In *Knowledge Engineering and Knowledge Management*, pages 421–425. Springer.

Pesquita, C., Faria, D., Santos, E., Neefs, J.-M., and Couto, F. M. (2014). Towards visualizing the alignment of large biomedical ontologies. In *Data Integration in the Life Sciences*, pages 104–111. Springer.

Sabou, M., d'Aquin, M., and Motta, E. (2006). Using the semantic web as background knowledge for ontology mapping. In *In Proc. of the Int. Workshop on Ontology Matching (OM-2006*.

Shvaiko, P. and Euzenat, J. (2005). A survey of schema-based matching approaches. In *Journal on Data Semantics IV*, pages 146–171. Springer.

Spyromitros-Xioufis, E., Papadopoulos, S., Kompatsiaris, I., Tsoumakas, G., and Vlahavas, I. (2014). A comprehensive study over VLAD and product quantization in large-scale image retrieval. *IEEE Transactions on Multimedia*, 16(6):1713–1728.

Stoilos, G., Stamou, G., and Kollias, S. (2005). A string metric for ontology alignment. In Gil, Y., editor, *Proceedings of the International Semantic Web Conference (ISWC 05)*, volume 3729 of *LNCS*, pages 624–637. Springer-Verlag.

Wu, Z. and Palmer, M. (1994). Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pages 133–138. Association for Computational Linguistics.