

DEVELOPING A MODEL SEMANTIC-BASED IMAGE RETRIEVAL BY COMBINING KD-TREE STRUCTURE WITH ONTOLOGY

Thanh The Van¹, Thanh Le², and Nguyen Thi Dinh²

¹Ho Chi Minh City University of Education

²Hue University University of Sciences

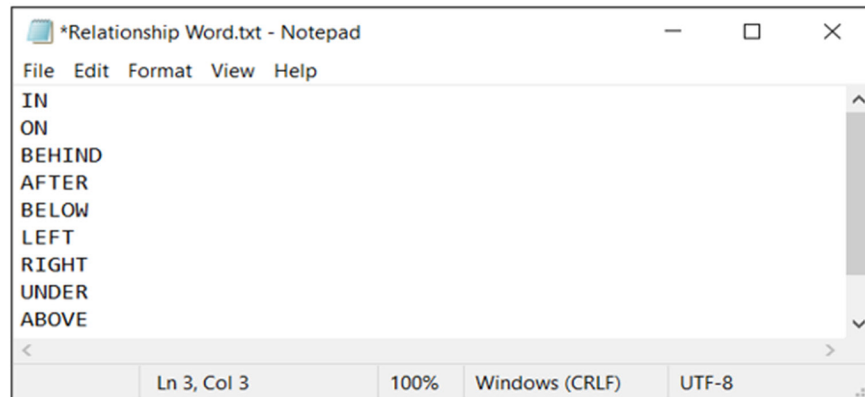
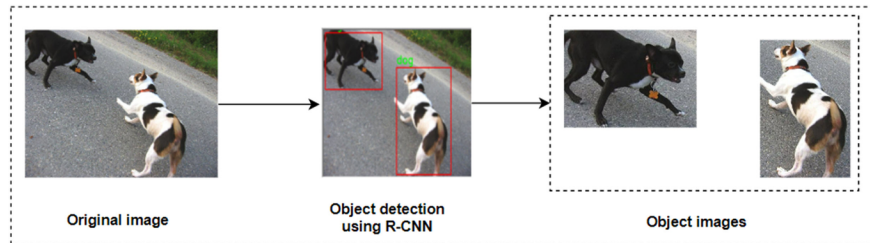
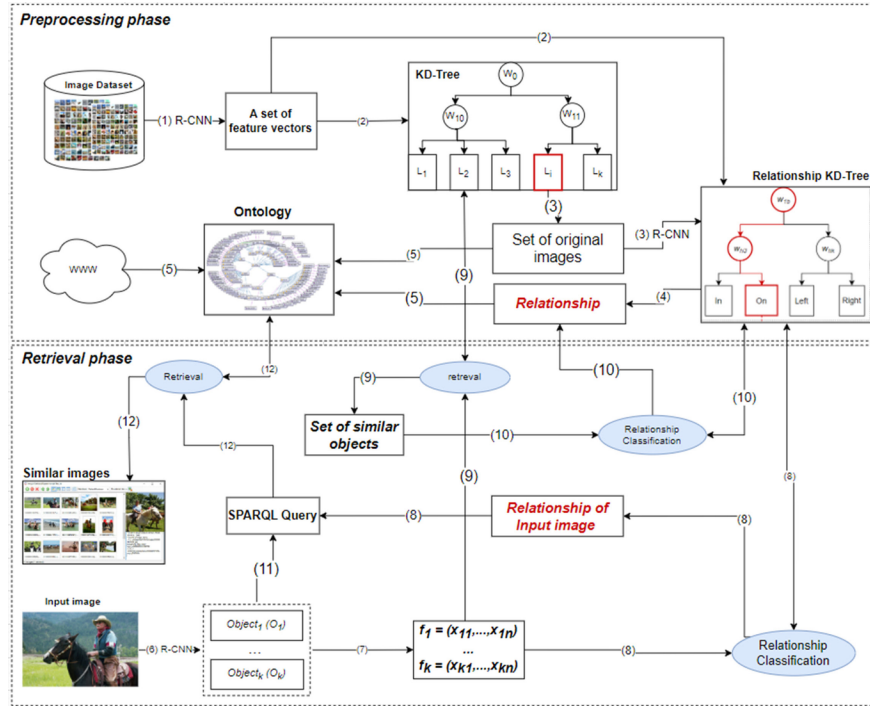
March 27, 2023

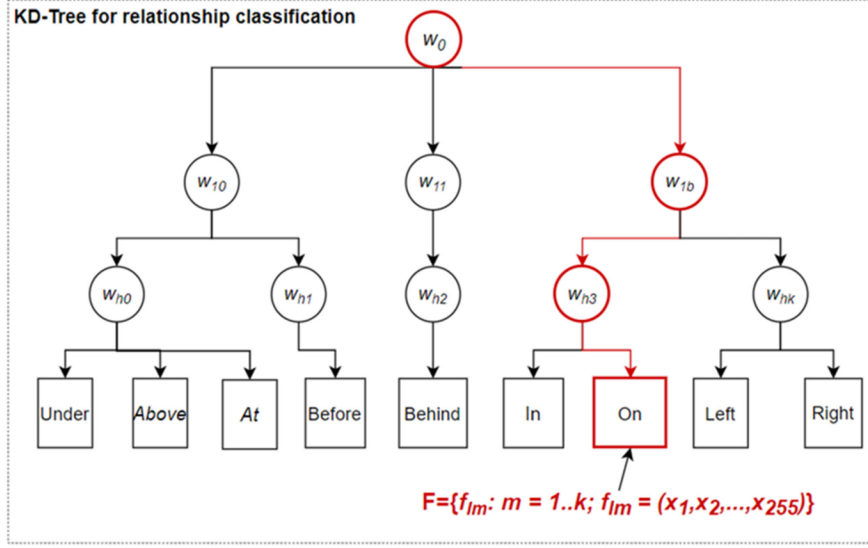
Abstract

The paper proposes an alternative approach to improve the performance of image retrieval. In this work, a framework for image retrieval based on machine learning and semantic retrieval is proposed. In the preprocessing phase, the image is segmented objects by using Graph-cut, and the feature vectors of objects presented in the image and their visual relationships are extracted using R-CNN. The feature vectors, visual relationships, and their symbolic labels are stored in KD-Tree data structures which can be used to predict the label of objects and visual relationships later. To facilitate semantic query, the images use the RDF data model and create an ontology for the symbolic labels annotated. For each query image, after extracting their feature vectors, the KD-Tree is used to classify the objects and predict their relationship. After that, a SPARQL query is built to extract a set of similar images. The SPARQL query consists of triple statements describing the objects and their relationship which were previously predicted. The evaluation of the framework with the MS-COCO dataset and Flickr showed that the precision achieved scores of 0.9218 and 0.9370 respectively.

Hosted file

ES_DEVELOPING A MODEL SEMANTIC-BASED IMAGE RETRIEVAL BY COMBINING KD-TREE STRUCTURE WITH ONTOLOGY.docx available at <https://authorea.com/users/600380/articles/631998-developing-a-model-semantic-based-image-retrieval-by-combining-kd-tree-structure-with-ontology>

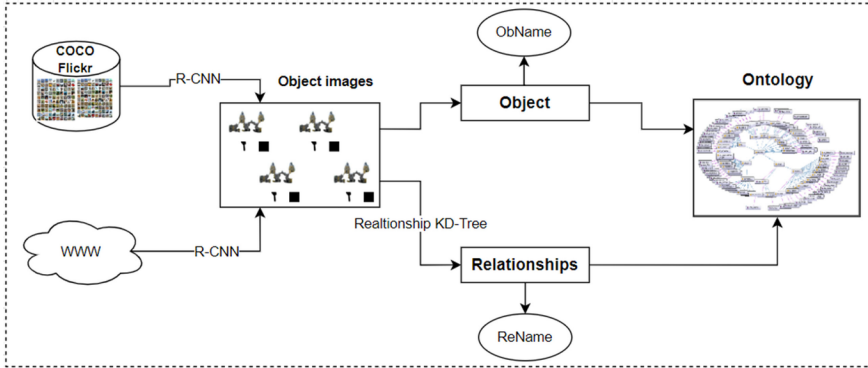


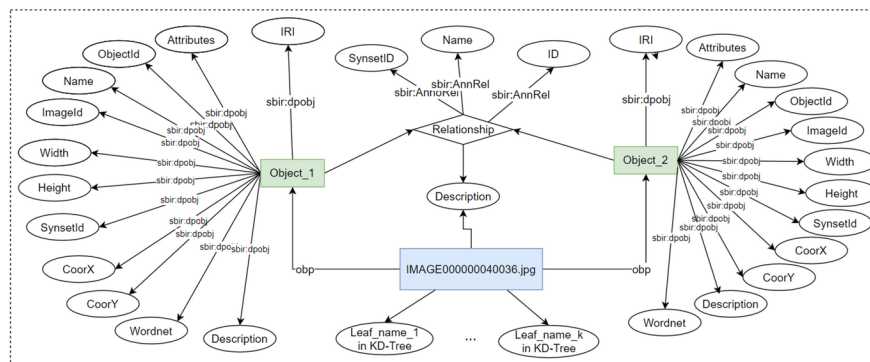


```

1 "categories": [{"supercategory": "person", "id": 1, "name": "person"},
2 {"supercategory": "vehicle", "id": 2, "name": "bicycle"},
3 {"supercategory": "vehicle", "id": 3, "name": "car"},
4 {"supercategory": "vehicle", "id": 4, "name": "motorcycle"},
5 {"supercategory": "vehicle", "id": 5, "name": "airplane"},
6 {"supercategory": "vehicle", "id": 6, "name": "bus"},
7 {"supercategory": "vehicle", "id": 7, "name": "train"},
8 {"supercategory": "vehicle", "id": 8, "name": "truck"},
9 {"supercategory": "vehicle", "id": 9, "name": "boat"},
10 {"supercategory": "outdoor", "id": 10, "name": "traffic light"},
11 {"supercategory": "outdoor", "id": 11, "name": "fire hydrant"},
12 {"supercategory": "outdoor", "id": 13, "name": "stop sign"},
13 {"supercategory": "outdoor", "id": 14, "name": "parking meter"},
14 {"supercategory": "outdoor", "id": 15, "name": "bench"},
15 {"supercategory": "animal", "id": 16, "name": "bird"},
16 {"supercategory": "animal", "id": 17, "name": "cat"},
17 {"supercategory": "animal", "id": 18, "name": "dog"},
18 {"supercategory": "animal", "id": 19, "name": "horse"},

```





```

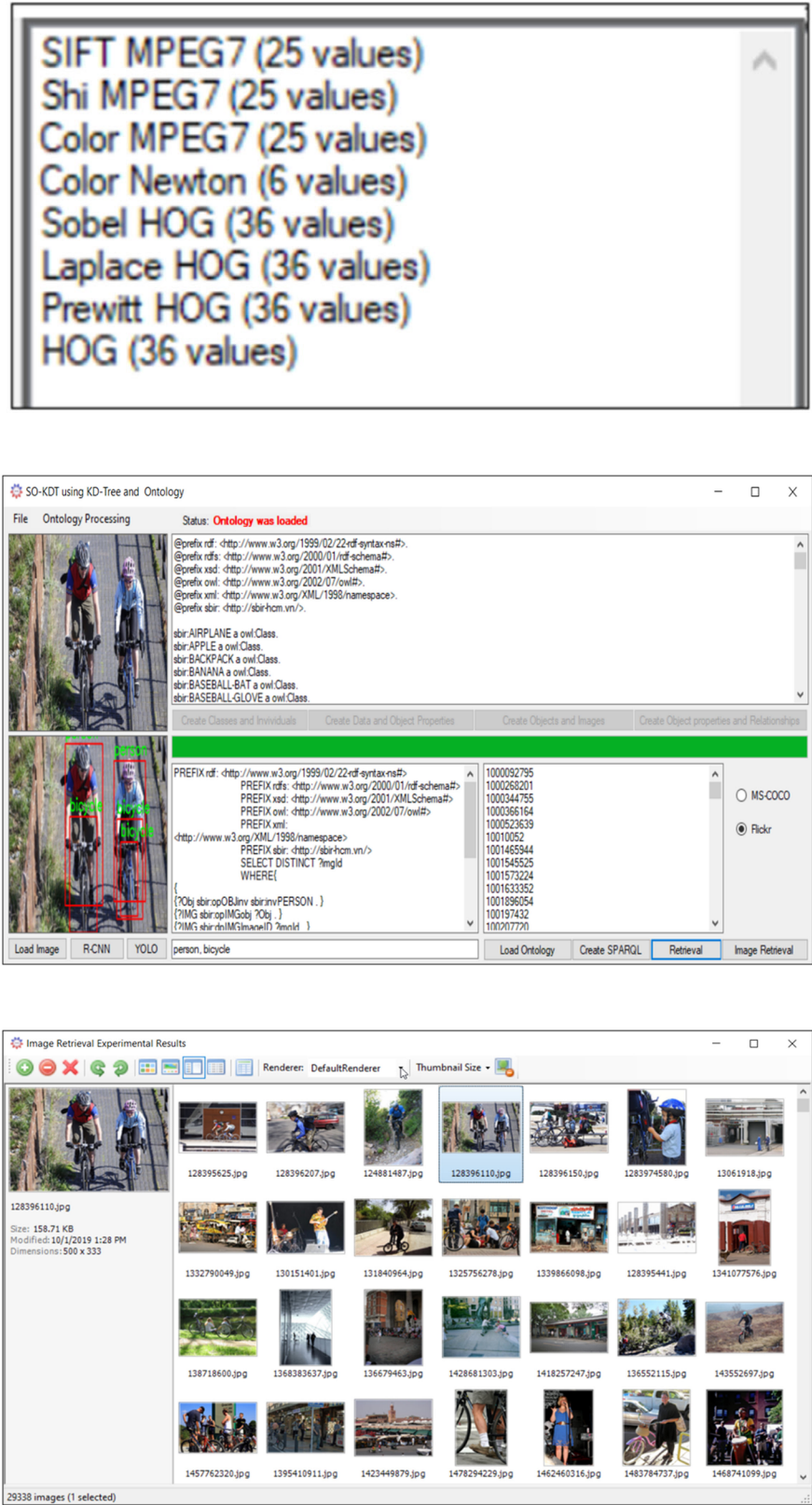
1 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
2 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
3 @prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
4 @prefix owl: <http://www.w3.org/2002/07/owl#>.
5 @prefix xml: <http://www.w3.org/XML/1998/namespace>.
6 @prefix sbir: <http://sbir-hcm.vn/>.
7
8 sbir:IMG000000000001 sbir:dpIMGImageID "000000000001";
9 sbir:opIMGobj sbir:OBJ000000000001_1,
10 sbir:OBJ000000000001_2,
11 sbir:OBJ000000000001_3;
12 a owl:NamedIndividual.
13 sbir:IMG000000000009 sbir:dpIMGImageID "000000000009";
14 sbir:opIMGobj sbir:OBJ000000000009_1,
15 sbir:OBJ000000000009_2,
16 sbir:OBJ000000000009_3,
17 sbir:OBJ000000000009_4,
18 sbir:OBJ000000000009_5;
19 a owl:NamedIndividual.
20 sbir:IMG000000000016 sbir:dpIMGImageID "000000000016";
21 sbir:opIMGobj sbir:OBJ000000000016_1,
22 sbir:OBJ000000000016_2,
23 sbir:OBJ000000000016_3;
24 a owl:NamedIndividual.
25 sbir:IMG000000000019 sbir:dpIMGImageID "000000000019";
26 sbir:opIMGobj sbir:OBJ000000000019_1,
27 sbir:OBJ000000000019_2;
28 a owl:NamedIndividual.
29 sbir:IMG000000000025 sbir:dpIMGImageID "000000000025";
30 sbir:opIMGobj sbir:OBJ000000000025_1,
31 sbir:OBJ000000000025_2;
32 a owl:NamedIndividual.

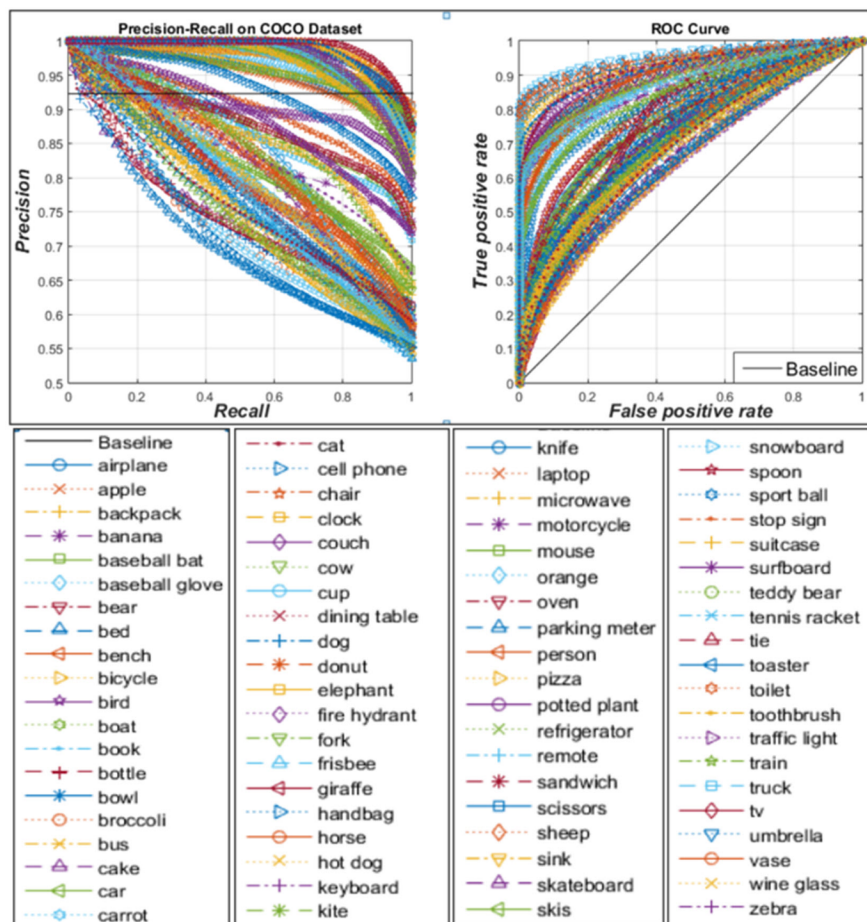
```

```

*SPARQL.txt - Notepad
File Edit Format View Help
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xml: <http://www.w3.org/XML/1998/namespace>
PREFIX sbir: <http://sbir-hcm.vn/>
SELECT DISTINCT ?imgId
WHERE{
    {
        {?Obj sbir:opOBJInv sbir:invVASE . }
        {?IMG sbir:opIMGobj ?Obj . }
        {?IMG sbir:dpIMGImageID ?imgId . }
    }
    LEFT
    {
        {?Obj sbir:opOBJInv sbir:invPOTTED-PLANT . }
        {?IMG sbir:opIMGobj ?Obj . }
        {?IMG sbir:dpIMGImageID ?imgId . }
    }
}
Ln 23, Col 2 100% Windows (CRLF) UTF-8

```



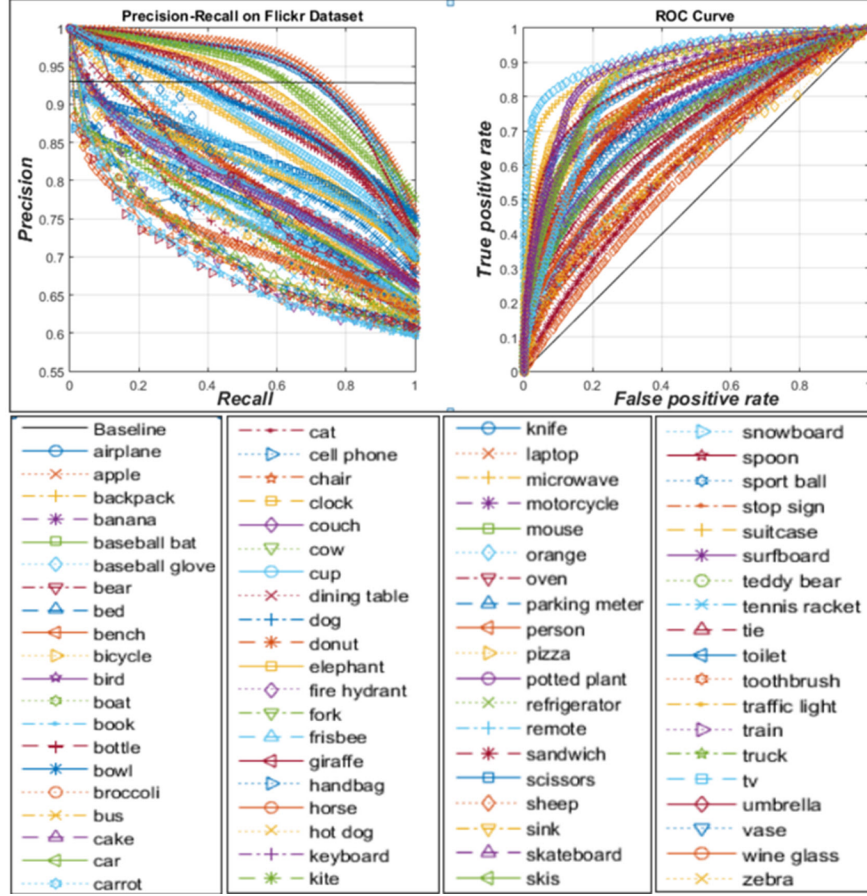


Table I. Describe experimental image data sets

<i>Data sets</i>	<i>Number of images</i>	<i>Training images</i>	<i>Testing images</i>	<i>Validation images</i>
MS-COCO	163,957	118,287	40,670	5,000
Flickr	31,783	29,000	1,783	1,000

Table II. Experimental results of image retrieval using KD-Tree and Ontology

<i>Data sets</i>	<i>Validation images</i>	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>	<i>Time query (ms)</i>
MS-COCO	5,000	0.9218	0.8372	0.8774	105.90
Flickr	1,000	0.9370	0.8939	0.9149	89.20

Table III. Comparison of mean average precision (*MAP*) of methods on MS-COCO dataset

Methods (Authors)	TopK	MAP
CN MAX, (Icarte, R. T.), TopK = 4, 2017 (Icarte, 2022)	10	0.3910
CAMP, (Wang, liu et al.), 2019 (Wang Z., 2019)	10	0.6890
CNN-RNN, (Wang, J., et al.), 2016 (J. Wang, et al., 2016)	-	0.6120
Hamming Ranking using AlexNet, (Zhiwei Zhang, et al.), 2021	3	0.7640
Resnet, (Wen, S., et al.), 2020 (Wen, 2020)	-	0.8110
D-MVE-Hash, (Chenggang Yan et al.), 2020 (Yan, 2020)	-	0.8892
Vision Transformer based Hashing, (Shiv Ram Dubey et al.), 2022 (Dubey, 2022)	-	0.9110
SO-KDT (Re KD-Tree, RF KD-Tree, Ontology)	-	0.9218

Table IV. Comparison of mean average precision (*MAP*) of methods on Flickr dataset

Methods (Authors)	TopK	MAP
IRSGS-GCN, (Yoon, 2021)	10	0.5670
CAMP, 2019 (Wang Z., 2019)	10	0.7710
BGAN- 48 bit, (Song, 2018)	-	0.7030
Unsupervised Deep Discriminative Hashing, (Hu, 2017)	-	0.8530
LocNar using text-only queries, (Changpinyo)	-	0.8970
SO-KDT (Re KD-Tree, RF KD-Tree, Ontology)	-	0.9370