NLM at ImageCLEF 2017 Caption Task

Asma Ben Abacha, Alba G. Seco de Herrera, Soumya Gayen, Dina Demner-Fushman, and Sameer Antani

Lister Hill National Center for Biomedical Communications,
National Library of Medicine, Bethesda, USA.
asma.benabacha@nih.gov, albagarcia@nih.gov, soumya.gayen@nih.gov,
ddemner@mail.nih.gov, santani@mail.nih.gov

Abstract. This paper describes the participation of the U.S. National Library of Medicine (NLM) in the ImageCLEF 2017 caption task. We proposed different machine learning methods using training subsets that we selected from the provided data as well as retrieval methods using external data. For the concept detection subtask, we used Convolutional Neural Networks (CNNs) and Binary Relevance using decision trees for multi-label classification. We also proposed a retrieval-based approach using Open-i image search engine and MetaMapLite to recognize relevant terms and associated Concept Unique Identifiers (CUIs). For the caption prediction subtask, we used the recognized CUIs and the UMLS to generate the captions. We also applied Open-i to retrieve similar images and their captions. We submitted ten runs for the concept detection subtask and six runs for the caption prediction subtask. CNNs provided good results with regards to the size of the selected subsets and the limited number of CUIs used for training. Using the CUIs recognized by the CNNs, our UMLS-based method for caption prediction obtained good results with 0.2247 mean BLUE score. In both subtasks, the best results were achieved using retrieval-based approaches outperforming all submitted runs by all the participants with 0.1718 mean F1 score in the concept detection subtask and 0.5634 mean BLUE score in the caption prediction subtask.

Keywords: Concept Detection, Caption Prediction, Convolutional Neural Networks, Multi-label Classification, Open-i, MetaMapLite, UMLS

1 Introduction

This paper describes the participation of the U.S. National Library of Medicine[1] (NLM) in the ImageCLEF 2017 caption task [1]. ImageCLEF [2] is an evaluation campaign organized as part of the CLEF[2] initiative labs. In 2017, the caption task consisted of two subtasks including concept detection and caption prediction. A detailed description of the data and the task is presented in Eickhoff et al. [1].

[1] http://www.nlm.nih.gov
[2] http://clef2017.clef-initiative.eu

The concept detection subtask consists of identifying the UMLS® (Unified Medical Language System)[3] Concept Unique Identifiers (CUIs). To solve this first challenge of detecting CUIs from a given image from the biomedical literature, we propose several approaches based on multi-label classification and information retrieval. For the multi-label classification, Convolutional Neural Networks (CNNs) and Binary Relevance using Decision Trees (BR-DT) are applied. The information retrieval approach is based on the Open-i Biomedical Image Search Engine[4] [3].

The caption prediction subtask aims to recreate the original image caption.To predict the captions of the images, we proposed a retrieval-based approach using Open-i and a second approach based on the retrieved CUIs and the UMLS® to find the associated terms and groups.

The rest of the paper is organized as follows. Section 2 describes the data provided for the two subtasks and our method to select training subsets. Then we present the proposed approaches for concept detection in Section 3 and caption prediction in Section 4. Section 5 provides a description of the submitted runs. Finally Section 6 presents and discusses our results.

2 Data Analysis and Selection

Training, validation and test datasets were provided containing 164,614, 10,000 and 10,000 biomedical images respectively. The images were extracted from scholarly articles on PubMed Central[5] (PMC).

For the concept detection subtask, a set of CUIs was provided for each image. For the caption prediction subtask, captions were provided. Figure 1 shows an example from the provided data.

2.1 Analysis of Concept Detection Data

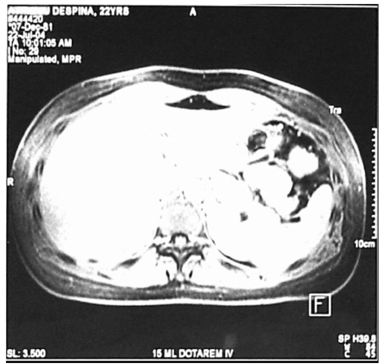
We analyzed the task data in order to study the types of methods that could be applied for concept detection subtask and whether it is needed to select training data and remove the less frequent CUIs. Also we studied whether it is relevant to build rule-based methods and construct patterns for the caption prediction subtask based on the recognized CUIs.

For the concept detection subtask:

– Training data includes 164,614 images associated with 20,463 CUIs. 19,145 CUIs have less than 100 images, including 6,251 CUIs with only one image.
– Validation data includes 10,000 images associated with 7,070 CUIs. 6,981 CUIs have less than 100 images, including 3,247 CUIs with only one image.

[3] https://www.nlm.nih.gov/research/umls
[4] http://Open-i.nlm.nih.gov
[5] http://www.ncbi.nlm.nih.gov/pmc

Image:



Concepts:

– C0016911: Gadolinium
– C0021485: Injection of therapeutic agent
– C0024485: Magnetic Resonance Imaging
– C0577559: Mass of body structure
– C1533685: Injection procedure

Caption: Magnetic resonance imaging. After intravenous injection of adolinium, the mass showed a progressive, heterogeneous, and delayed enhancement.

Fig. 1. Example of an image and the associated CUIs and caption from the training set of the ImageCLEF 2017 caption task.

Table 1 presents the most frequent CUIs in the training data and their UMLS®[6] terms and semantic types.

Table 1. Most frequent CUIs in the training data of the 2017 concept detection subtask.

CUI	UMLS Term	UMLS Semantic Type	# Associated Images
C1696103	Image-dosage form	Intellectual Product	17,998
C0040405	X-Ray Computed Tomography	Diagnostic Procedure	16,217
C0221198	Lesion	Finding	14,219
C1306645	Plain x-ray	Diagnostic Procedure	10,926
C0577559	Mass of body structure	Finding	9,769
C0027651	Neoplasms	Neoplastic Process	9,570
C0441633	Scanning	Diagnostic Procedure	9,289

[6] We used the UMLS 2017AA release.

2.2 Training Data Selection for Concept Detection

The heterogeneous distribution of CUIs in the training data is not adapted for multi-label classification, therefore we studied data selection methods.

Cho et al. [4] applied deep learning to medical image classification and focused on determining the ideal training data size to achieve high classification accuracy. They trained a CNN using different sizes of training data and tested the models on 6000 computed tomography (CT) images. Using 200 training samples, the classification accuracy was already near or at 100%. Based on these experiments, we fixed a threshold of 200 training images for each CUI.

In addition to the number of examples for each CUI, some CUIs are a lot more frequent than others in the datasets (the number of training images for each CUI varies from 1 to 17,998). Therefore we built two different training subsets targeting the most frequent CUIs:

- **Subset 1 [92 CUIs with frequency>=1,500]**: We selected CUIs having at least 1,500 training examples. This subset corresponded to 92 distinct CUIs. For each CUI, we selected randomly 200 training examples from the provided training images.
- **Subset 2 [239 CUIs with frequency>=400]**: We selected CUIs having at least 400 training examples. This subset corresponded to 239 CUIs. For each CUI, we selected randomly 200 training examples.

We used these two subsets to train our machine learning (ML) methods.

3 Concept Detection Methods

For the concept detection subtask, each image can be associated with one or multiple CUIs. We approached the problem in two ways, (1) applying multi-label classification methods and (2) using a retrieval-based approach.

In the multi-label classification approach we consider the CUIs in the training set as the labels to be assigned. Thus each image will be assigned one or multiple labels from the predefined label set. Two methods for multi-label classification were applied: Convolutional Neural Networks (CNNs) and Binary Relevance using Decision Trees (BR-DT). To train our ML models, we utilized the high-performance computational capabilities of the Biowulf Linux cluster at the U.S. National Institutes of Health[7].

In the information retrieval approach, we used Open-i to retrieve the most similar images and their associated labels and CUIs.

3.1 Multi-label classification with Convolutional Neural Networks (CNNs)

Deep learning methods have been widely applied to image analysis. In particular, CNNs achieved excellent results for image classification [5, 6].

[7] http://biowulf.nih.gov

We applied CNNs for multi-label classification and tested different neural networks such as the GoogleNet network [7]. GoogLeNet won the classification and object recognition challenges in the 2014 ImageNet LSVRC competition (ILSVRC2014[8]). In our experiments on the training sets, the GoogleNet network provided better results compared to AlexNet [8] and LeNet [9].

We ran the CNNs using NVIDIA Deep Learning GPU Training System (DIGITS)[9]. DIGITS is a Deep Learning (DL) training system with a web interface that allows designing custom network architectures and evaluating their effectiveness. It also allows the design of new models by providing the details of optimization and network architecture. DIGITS can be used for image classification, segmentation and object detection tasks.

In our final runs, we used the GoogLeNet network. We applied stochastic gradient descent (SGD) and performed 100 training epochs. We used the two training subsets associated respectively to 92 and 239 CUIs to train the network (see Section 2.2).

3.2 Multi-label Classification with Binary Relevance using Decision Trees (BR-DT)

The Meka project [10][10] is based on the Weka machine learning library [11], and provides an open source implementation of methods for multi-label classification. It contains several algorithms, such as Binary Relevance (BR) or Label Powerset.

Similar to [12] we used BR-DT as implemented in Meka (J48). BR methods create an individual model for each label, thus each model is a simply binary problem. We used Decision Trees (DT) as a base classifier because DT are able to capture relations between labels. For the experiments we extract from the images one visual descriptors commonly used for image classification Colour and Edge Directivity Descriptor (CEDD) [13]. The descriptor was provided as input to Meka.

Before submitting the runs we carried out some experiments on the training data using also Fuzzy Colour and Texture Histogram (FCTH) [14] as a visual descriptor. However using CEDD provided better results.

3.3 Retrieval and Annotation Approach with Open-i and MetaMapLite

The Open-i service of the NLM enables search and retrieval of abstracts and images (including charts, graphs, clinical images) from the open source literature, and biomedical image collections. Open-i provides access to over 3.7 million images from about 1.2 million PubMed Central articles, 7,470 chest x-rays with 3,955 radiology reports, 67,517 images from NLM History of Medicine collection,

[8] http://image-net.org/challenges/LSVRC/2014/eccv2014
[9] http://github.com/NVIDIA/DIGITS
[10] http://meka.sourceforge.net

2,064 orthopedic illustrations and 8084 medical case images from MedPix[11]. Open-i combines text processing, image analysis and machine learning techniques to retrieve relevant images from an input image-query.

We submitted each query image to the Open-i search API and selected 10 result images with captions. For each retrieved image, we annotated its caption with MetaMapLite[12] (3.1-SNAPSHOT version) to recognize CUIs. MetaMapLite recognizes named entities using the longest match as well as associated CUIs. It also allows restricting the CUIs with UMLS Semantic Types. We did not use any restriction as CUIs in the provided data have heterogeneous semantic types.

4 Caption Prediction Methods

To predict image captions, we used two different methods based on UMLS® and Open-i.

4.1 UMLS-based Method

We used the CUIs recognized in the first concept detection subtask to generate the associated UMLS terms and semantic types. We then grouped the recognized UMLS terms using the UMLS groups of their semantic types. The UMLS Semantic Network includes 15 groups: Activities & Behaviors, Anatomy, Chemicals & Drugs, Concepts & Ideas, Devices, Disorders, Genes & Molecular Sequences, Geographic Areas, Living Beings, Objects, Occupations, Organizations, Phenomena, Physiology and Procedures.

The following are examples of four captions and their corresponding image IDs, generated using the UMLS-based method:

1. **1471-2342-10-23-4**: Procedures: diagnostic computed tomography, imaging pet. Anatomy: armpit. Disorders: metastasis. Physiology: uptake.
2. **iej-04-20-g007**: Procedures: h&e stain. Chemicals & Drugs: haematoxylin, 11445 red, eosin. Disorders: proliferation.
3. **13014_2015_335_Fig1_HTML**: Procedures: brain mri, diffusion weighted imaging, bodies weight. Concepts & Ideas: rows. Chemicals & Drugs: gadolinium.
4. **fonc-04-00350-g002**: Procedures: antineoplastic chemotherapy regimen. Disorders: abnormally opaque structure, condition response. Anatomy: left lung, anterior thoracic region.

4.2 Open-i-based Method

For each input image, the Open-i biomedical image search engine returns a list of similar images. In our experiments we performed several tests with the caption, mention, Medical Subject Headings (MeSH®) terms, three outcomes and

[11] As of September 2016.
[12] https://metamap.nlm.nih.gov/MetaMapLite.shtml

medical problems from the retrieved images. In our final runs, we used only the captions of the first and second retrieved images.

The following are two examples of results provided by Open-i:

1. **1MS-10-20646-g003**: Open-i provides the following relevant results:
   - Caption: Laryngostenosis in patient with laryngeal tuberculosis. Tracheostomy.
   - Problem(s): tuberculoses.
   - Concept(s): laryngostenoses; laryngeal tuberculoses.
   - Outcomes: (i) Within the group of patients with lymph node tuberculosis in 15 cases there were infected lymph nodes of the 2(nd) and 3(rd) cervical region and in 11 infected lymph nodes of the 1(st) cervical region. (ii) In 5 cases of laryngeal tuberculosis there was detected coexistence of cancer. (iii)Chest X-ray was performed in all cases and pulmonary tuberculosis was identified in 26 (35.6%) cases.
   - Mention: Moreover, histopathological examination revealed in 5 cases coexistence of planoepithelial carcinoma with tuberculosis. In all 5 cases total laryngectomy was performed. Chest X-ray was performed in all patients and the evidence of lung tuberculosis was confirmed in 14 (70%) cases. Tuberculin skin test was positive in 10 (66.6%) out of 15 tests performed. Contact history with active tuberculosis was detected in 3 (15%) cases (Figures 2 and 3).
2. **110.1177_2324709614529417-fig1**: Open-i provides the following results:
   - Caption: Magnetic resonance imaging after the onset of isolated adrenocorticotropic hormone deficiency. Magnetic resonance imaging showed no space-occupying lesions in the pituitary gland or hypothalamus.
   - Problem(s): isolated adrenocorticotropic hormone deficiency.
   - Concept(s): isolated adrenocorticotropic hormone deficiency.
   - Outcomes: (i) Although the neutrOpen-ia and fever immediately improved, he became unable to take any oral medications and was bedridden 1 week after admission. (ii) His serum sodium level abruptly decreased to 122mEq/L on the fifth day of hospitalization. (iii) Hydrocortisone replacement therapy was begun at 20mg/day, resulting in a marked improvement in his anorexia and general fatigue within a few days.
   - Mention: CT and magnetic resonance imaging showed no space-occupying lesion or atrophic change in his pituitary gland or hypothalamus (Figure1).

5 Runs

This section provided a detailed description of the runs submitted to ImageCLEF 2017 caption task. The methods used to implement these runs are described in previous Sections 3 and 4.

5.1 Concept Detection

As specified by the task guidelines, a maximum of 50 UMLS concepts per figure is accepted. Therefore, if the limit of 50 CUIs per image is reached, we took only the first 50 CUIs for each image. We submitted the following runs to the Concept Detection subtask:

- DET 1. **DET_run_1_Open-i_MetaMapLite_1**: We used Open-i to find similar images and then extracted CUIs from their captions using MetaMapLite. In this first run, we used the caption of the most similar image according to Open-i. The returned CUIs are all the CUIs recognized by MetaMapLite.
- DET 2. **DET_run_1_baseline**: The same DET 1 run with the exclusion of test images if they are retrieved by Open-i.
- DET 3. **DET_run_2_Open-i_MetaMapLite_2**: The same as *DET 1* except that we took only the first CUI recognized by MetaMapLite for each term.
- DET 4. **DET_run_3_Open-i_MetaMapLite_3**: Similar to *DET 1* except that we used the captions from the first and second best images retrieved by Open-i.
- DET 5. **DET_run_5_Meka_CEDD**: Multi-label Classification method using MEKA software to applied binary relevance method. CEED is used as a visual descriptor for the images. Subset 1 of 92 CUIs is used for training.
- DET 6. **DET_run_6_CNN_GoogLeNet_92Cuis**: Multi-label classification with a convolutional neural network. We trained the GoogLeNet network using subset 1 of 92 CUIs.
- DET 7. **DET_run_7_CNN_GoogLeNet_239Cuis**: We trained the GoogLeNet network using subset 2 of 239 CUIs.
- DET 8. **DET_run_8_comb1_CNN2**: Fusion of the runs *DET 6* and *DET 7*)
- DET 9. **DET_run_9_comb2_CNN2Meka**: Fusion of the runs *DET 5*, *DET 6* and *DET 7*)
- DET 10. **DET_run_10_comb3_CNN2MekaOpen-i**: Fusion of the runs *DET 1*, *DET 5*, *DET 6* and *DET 7*)

5.2 Caption Prediction

We submitted the following runs to the Caption Prediction subtask:

- PRED 1. **PRED_run_1_Open-iMethod**: We used Open-i Biomedical Image Search Engine to find similar images. In this run, we used the caption of the first retrieved image.
- PRED 2. **PRED_run_1_baseline**: Same as *PRED 1*, except we excluded the test images if they are retrieved by Open-i.
- PRED 3. **PRED_run_2_CNN_92**: We used the CUIs recognized by the CNN (CRun *DET 6*) and the UMLS semantic groups to generate the captions.

PRED 4. **PRED_run_3_CNN_239**: We used the CUIs recognized by the CNN (run *DET 7*) and the UMLS semantic groups to generate the captions.

PRED 5. **PRED_run_4_CNN_comb**: We used the CUIs recognized by the CNN (run *DET 8*) and the UMLS semantic groups to generate the captions.

PRED 6. **PRED_run_5_comb_all**: We used the CUIs recognized by the hybrid method (run *DET 10*) and the UMLS® to generate the captions.

6 Official Results

In this section we describe and discuss the results obtained by the submitted runs.

6.1 Concept Detection Results

Table 2 shows our official results in the concept detection subtask and their ranks compared with all the 37 runs submitted by the 9 participating teams.

Table 2. Results of our submitted runs to the concept detection subtask and their ranks in comparison with the 37 submitted runs by 9 groups.

Run	Mean F1 Score	Ranking
DET 1	**0.1718**	1
DET 3	0.1648	2
DET 10	0.1390	10
DET 4	0.1228	13
DET 8	0.0880	18
DET 9	0.0868	20
DET 6	0.0811	22
DET 7	0.0695	23
DET 2 (baseline)	0.0162	34
DET 5	0.0012	36

The best overall results were obtained by run *DET 1* followed by run *DET 3*; both approaches are based on Open-i retrieval system. To better understand the results, Table 3 shows the efficiency of the Open-i system on the test set by presenting how many times the query image itself was retrieved and ranked in the first 10 positions when searching on the full Open-i collection (3.7 million images). We analyze only the first 10 because it is the maximum number of retrieved images that we used in our experiments.

Open-i was able to find the image in the first top 10 results in 61% of the cases, and extract the relevant information from the image itself.

For comparison, we performed a second run called *DET 2*, which is equivalent to run *DET 1* but with the exclusion of test images if they are retrieved by Open-i. For run *DET 2* the mean F1 score decreased to 0.0162, which we consider as baseline result. The best results using Open-i based approaches were obtained when using all the CUIs associated with the first retrieved image.

Table 3. Number of times Open-i retrieves the query image itself and its rank. Images belong to the test set of the ImageCLEF 2017 caption task.

Ranking	1	2	3	4	5	6	7	8	9	10	Total
# matches	3847	756	409	280	230	182	124	99	112	77	6116

Without using external resources the results were poorer. One of the reasons could be that not all the CUIs in the test set were contained in the training and validation sets. Also, we only considered the most frequent CUIs in the training set. With CNNs, up to 0.0880 mean F1 score was achieved and only 0.0012 when applying BR-DT (BR-DT detected at least one CUI on 2046 images only).

Table 2 also shows the performance of three hybrid methods: run *DET 8*, run *DET 9* and run *DET 10*.

6.2 Caption Prediction Results

Table 4 shows our official results in the caption prediction subtask and their ranks compared with the 34 runs submitted by the 5 participating teams.

Table 4. Results of our submitted runs to the caption prediction subtask and their ranks in comparison with the 34 submitted runs by 5 groups.

Run	Mean BLUE Score	Ranking
PRED 1	**0.5634**	1
PRED 6	0.3317	2
PRED 2 (baseline)	0.2646	4
PRED 5	0.2247	11
PRED 4	0.1384	18
PRED 3	0.1131	19

The best results were achieved by run *PRED 1* using Open-i with 0.5634 mean BLUE score and was ranked first. As baseline, we proposed run *PRED 2*, similar to run *PRED 1* but without including test images if they are retrieved by Open-i. Run *PRED 2* obtained 0.2646 mean BLUE score and was the 4th best run out of 34 submitted runs by the participating teams.

CNN approaches achieved good results with 0.2247 mean BLUE score despite the limited number of CUIs used for training and the simple UMLS-based patterns built for caption generation. Two hybrid methods were also presented: run *PRED 5* and run *PRED 6*. In this subtask, run *PRED 6* was ranked second.

7 Conclusions

This paper describes our participation in ImageCLEF 2017 caption task. We proposed and compared different approaches for concept detection and caption prediction. Our retrieval methods using Open-i obtained the best results with

0.1718 mean F1 score in the concept detection subtask and 0.5634 mean BLUE score in the caption prediction subtask. We proposed baseline results by excluding test images if they are found by Open-i. Open-i baseline was ranked 4th with 0.2646 mean BLUE score in the caption prediction subtask.

We also performed multi-label classification of CUIs with CNNs and BR-DT. Both methods used selected subsets from the training data. CNNs provided acceptable results with regards the limited number of CUIs used for training. CNNs method achieved 0.2247 mean BLUE score in the caption prediction subtask.

Future improvements can tackle Open-i method as it does not support images with panels. One better way would be to perform panel segmentation before the search. Open-i also has size limitations on images of 2MB. A better approach would be to resize the image if needed before submitting to Open-i API. Also, MetaMapLite provided CUIs that are different from the gold standard even if the labels retrieved by Open-i are correct. Moreover, we only used the fusion to combine the results of our different methods for concept detection (the intersection gave very few CUIs). More sophisticated combination methods could be used to improve the results of the hybrid methods.

Acknowledgments

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC).

References

1. Eickhoff, C., Schwall, I., García Seco de Herrera, A., Müller, H.: Overview of ImageCLEFcaption 2017 - the image caption prediction and concept extraction tasks to understand biomedical images. CLEF working notes, CEUR (2017)
2. Ionescu, B., Müller, H., Villegas, M., Arenas, H., Boato, G., Dang-Nguyen, D.T., Dicente Cid, Y., Eickhoff, C., García Seco de Herrera, A., Gurrin, C., Islam, Bayzidul and, K.V., Liauchuk, V., Mothe, J., Piras, L., Riegler, M., Schwall, I.: Overview of ImageCLEF 2017: Information extraction from images. In: CLEF 2017 Proceedings. Lecture Notes in Computer Science, Dublin, Ireland, Springer (September 11-14 2017)
3. Demner-Fushman, D., Antani, S., Simpson, M.S., Thoma, G.R.: Design and development of a multimodal biomedical information retrieval system. Journal of Computing Science and Engineering **6**(2) (2012) 168–177
4. Cho, J., Lee, K., Shin, E., Choy, G., Do, S.: Medical image deep learning with hospital PACS dataset. CoRR **abs/1511.06348** (2015)
5. Roth, H.R., Lee, C.T., Shin, H.C., Seff, A., Kim, L., Yao, J., Lu, L., Summers, R.M.: Anatomy-specific classification of medical images using deep convolutional nets. In: ISBI, IEEE (2015) 101–104
6. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A.C., Bengio, Y., Pal, C., Jodoin, P., Larochelle, H.: Brain tumor segmentation with deep neural networks. CoRR **abs/1505.03540** (2015)

7. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015. (2015) 1–9
8. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q., eds.: Advances in Neural Information Processing Systems 25. Curran Associates, Inc. (2012) 1097–1105
9. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11) (November 1998) 2278–2324
10. Read, J., Reutemann, P., Pfahringer, B., Holmes, G.: MEKA: A multi-label/multi-target extension to Weka. Journal of Machine Learning Research **17**(21) (2016) 1–5
11. Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann (2016)
12. Tanaka, E.A., Nozawa, S.R., Macedo, A.A., Baranauskas, J.A.: A multi-label approach using binary relevance and decision trees applied to functional genomics. Journal of Biomedical Informatics **54** (2015) 85–95
13. Chatzichristofis, S.A., Boutalis, Y.S.: CEDD: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In: Lecture notes in Computer Sciences. Volume 5008. (2008) 312–322
14. Chatzichristofis, S.A., Boutalis, Y.S.: FCTH: Fuzzy color and texture histogram: A low level feature for accurate image retrieval. In: Proceedings of the 9th International Workshop on Image Analysis for Multimedia Interactive Service. (2008) 191–196