Görüntülerden Bilgi Elde Etmek İçin Görüntü Açıklama Çalışmalarının Sistematik Haritalandırma İncelemesi

Araştırma Makalesi/Research Article

Arda SEZEN, Çiğdem TURHAN

Yazılım Mühendisliği Bölümü, Atılım Üniversitesi, Ankara, Türkiye
ardasezen@gmail.com, cigdem.turhan@atilim.edu.tr
(Geliş/Received:24.10.2019; Kabul/Accepted:23.09.2020)
DOI: 10.17671/gazibtd.637713

***Özet*—** Çeşitli görüntü türlerinden görüntü meta verilerini elde etmeyi, işlemeyi ve anlamlı sonuçlar çıkarmayı amaçlayan görüntü açıklaması kavramı, bilgi toplumunda görüntülerden bilgi edinme konusunda daha kritik hale gelmiştir. Bu konuda literatürde önerilen birçok yaklaşım vardır, ancak en uygun olanı seçmek kolay değildir. Bu çalışmada, görüntü açıklamalarındaki kilit yaklaşımlar, kapsamlı bir literatür taraması ve sistematik haritalama yoluyla incelenmiştir. Çalışmamızın yeniliği, literatürdeki görüntü açıklama çalışmalarını araştırmak, haritalamak ve analiz etmek ve seçilen çalışmaları sınıflandırmak, araştırma boşluklarını ortaya çıkarmak ve bu makalede sunulan araştırma soruları için görsel özet hazırlamak adına ilk denemeyi temsil etmesidir. Literatürde, ilgi alanını farklı açılardan araştırmak için sistematik haritalama yaklaşımı önerilmektedir. Bu amaçla, tespit edilen toplam 404 çalışma içerisinden 95 çalışma seçilmiştir. Alandaki literatürün incelenmesi, mevcut yöntemlerin / tekniklerin, araçların, metriklerin, süreçlerin veya diğer teknik yaklaşımların kendi başlarına eksiksiz bir çözüm üretmek için yeterli olmadığını göstermektedir. Önerilen teknikleri birleştirerek ve disiplinlerarası yaklaşımları dikkate alarak yeni çözümler üretmenin gerekliliği de önerilmiştir.

***Anahtar Kelimeler*—** anlamsal görüntü açıklaması, görüntü işleme, resim yorumlama

A Systematic Mapping Review of Image Annotation Studies for Obtaining Information Retrieval from Images

***Abstract*—** Image annotation concept which aims to obtain and process image metadata from various kinds of images, to achieve meaningful results has become more and more critical in the information society in retrieving information from images. There are many approaches proposed in the literature, however, choosing the most appropriate ones is not an easy task. In this study, the key approaches on image annotation have been investigated through a systematic mapping with an extensive literature review. The novelty of our study is that it represents the first attempt to explore, investigate, map and analyze image annotation studies in literature and to help classify the studies, reveal research gaps, and prepare a visual summary for the research questions presented in this paper. The literature recommends the systematic mapping approach to investigate an area of interest from different perspectives. For this purpose, 95 studies were selected from a total of 404 studies identified. The examination of the literature on the domain shows that the available methods/techniques, tools, metrics, processes, or other technical approaches are not enough to produce a complete solution on their own. The necessity of generating new solutions by combining the proposed techniques and considering interdisciplinary approaches are also suggested.

***Keywords*—** semantic image annotation, image processing, picture interpretation

1. INTRODUCTION

Image annotation is a systematic way of adding, obtaining and processing metadata to get meaningful results from digital images in various forms such as captioning, keywords or sentences. Most of the time, the term image annotation is mixed with the term of image classification, but they differ in terms of goals and functionality. Image classification does not intend to add metadata or to caption digital images, the purpose of image classification is to identify the content, and this can be used by image annotation to achieve successful annotations.

There are important topics that need to be examined in image annotation. The necessity of whether object recognition is required for image annotation is still an unclear issue in machines. There are two main questions Lavrenko addresses [1]:

What is the object in a Picture?

Where is it in the Picture?

When it comes to object detection or recognition, some specific objects such as faces or cars should be focused on. Lavrenko suggests that "the joint probability for different regions forming these objects" should be learned, then examples for these objects should be trained [1]. Finally, a "two-class classification" problem should be solved which is called binary classification [1].

Annotating images poses a few issues. The first of these issues is whether production or post-production is a more important issue for image annotation. For some aspects, it is easier to make productions earlier rather than annotating them later, because the required information is applicable during production time. In addition to this, it is cheaper and higher quality is achieved when metadata is added during production time rather than adding it later [2].

Second, there is a generic annotation versus task-specific annotation issue. It is ineffective and costly to annotate images without a clear aim, object and background. Also, in most situations, it is not possible to know which metadata will be used for the applications under development in the future, therefore disregarding irrelevant application-particular assumptions is the best option to an annotator [2].

Manual annotation versus automatic annotation issue and the "Semantic Gap" is another phenomenon. Experts believe that image descriptions can be provided by manual annotation "at the right level of abstraction" [2]. On the other hand, automatic annotation is faster and cheaper and provides more systematic feature extraction. As a result, automatic annotation produces a too low-level image description for many applications where the difference between the high-level content descriptions and the low-level feature is referred to as "Semantic Gap" [2].

Moreover, there are different types of metadata issues. Experts agree on the necessity of understanding the contrast between annotations portraying properties of the image, and those depicting the subject of the image, which is, the properties of the objects, people or ideas delineated by the image [2].

Finally, experts point out the lack of syntactic and semantic interoperability issues. For example, using the same metadata created by another tool is not possible because of a lack of interoperability since a different syntax may be used by a tool, and a different meaning or semantics may be assigned to the same annotation by a particular tool [2].

Various methods and techniques have been proposed to address the above challenges related to image annotation. While research and the number of related studies in the image annotation field increase, the need for a systematic classification of the current trends and debates in the academic literature also grows [3-5]. It is believed that a broader community of researchers and professionals will benefit from work on open research areas with the identification of future predictions in the field [3-6].

In this study, the information within the field of image annotation domain is analyzed and classified through a comprehensive systematic mapping (SM) [5]. For this purpose, a group of research questions is raised, inclusion and exclusion criteria are defined, and a classification schema is developed for the selected studies.

After the selection process which is detailed in the Research Method Section, 95 studies are included from a total of 404 studies that were published in the field of image annotation between the years of 2010 and 2018 [7]. Domain-specific trends such as types of input, research aspects, image annotation activities (e.g. research method), classification schemes, and types of evaluations in the primary studies are derived. This way, the precise formulation of the current studies in this domain is presented with the contribution of a methodical map that is fostered for the area of image annotation.

In this study, Section 2 discusses the general framework and related studies. Section 3 describes our methodology, including the SM procedure, the aim and research questions handled in this study, followed by the article selection process. Section 4 discusses the results of the systematic mapping. Lastly, conclusion and future studies are presented in Section 5.

2. BACKGROUND

In the image annotation domain, the term "annotation" is related to obtaining information from an image and presenting it in a meaningful way after an interpretation process. Although the definition is simple, the task is not because of the various issues mentioned in the previous section. Researchers work on alternative solutions for image annotation by combining various technological advancements. One such contributing area is the Semantic Web. The researchers utilize Semantic Web to create links meaningful to machines, classify both links and the target, and also provide some limited reasoning. With the vast amount of internet images available, this research area is becoming more valuable than ever for image annotation in terms of getting a hint about image content from links or maybe image file itself. Besides, ontology-based systems are used for defining representational primitives. More semantic results can be achieved when these systems are used together with the annotation of an image. These enhancements are applied during the different stages of the image annotation procedure to decrease the semantic gap.

The literature on image annotation provides a large number of studies, regarding the issues mentioned before. One such example is the lack of effective modeling method for the high-level semantics of images. Wang et al. proposed a method based on hot Internet topics which also addresses the lack of efficient dynamic update mechanisms for the training set in their study [8]. In terms of automatic image annotation, Bannour and Hudelot highlight the wide usage of machine learning techniques to provide a mapping function that allows classification using object features. At the same time, they also put forward the scalability issue of these approaches when dealing with broad content image datasets and propose a methodology for building fuzzy multimedia ontologies to address the issue in the image annotation domain to model image semantics [9]. In [10], the authors focus on the issue of subjectivity of human annotation and the ineffective time requirement for the manual annotation process. Fakhari and Moghadam propose an automatic image annotation approach based on semantic image retrieval for the high-level semantics within images regardless of their low-level features. In [11], the authors investigate the performance of capsule networks on the clothing classification task. The study of Gong and et al. shows that significant performance can be gained using deep convolutional architectures for multilabel annotation [12].

Another widely investigated concept in image annotation is multiple-instance learning, but multiple-instance learning has a drawback considering it can be solved by traditional supervised learning methods. Feature mapping usually overlooks the discriminative ability of the generated features. Hong et al. propose a multiple-instance learning method with discriminative feature mapping and feature selection in their study to address this certain drawback [13]. For discriminative image patches, a study conducted by Zhang et al. proposes a solution for annotating and retrieving Web-scale image data which is called ObjectPatchNet. Here, each vertex is defined as a collection of discriminative image patches annotated with object category labels [14].

Various methods, techniques and systems exist for semi-automatic annotation systems. Im and Park show an example of such a system by using semantic relations between social tags [15]. Seneviratne and Izquierdo propose another semi-automatic annotation approach to address the issue of the semantic annotation and tagging of multimedia contents. Their approach deviates from the conventional content-based image retrieval paradigm. The proposed approach uses a multifaceted mathematical model based on game theories to aggregate numbers of different key-paradigms [16]. Constructing high-quality image samples is another discussed problem in the image annotation domain due to labor intensiveness [17]. There are some proposed methods to solve this problem such as the negative sample image selection method. With this method, the highest accuracy can be achieved when a support vector machine is adopted, and if a uniform amount of negative sample images in the semantic hierarchical tree is selected [18]. In some papers, statistical methods are proposed together with alternative approaches for image annotation [19-21]. Liu et al. propose a sparse distribution attribute, local convergent assumption, and global convergent conjecture which are essential for keyword selection and image content understanding to overcome the semantic gap [19]. Mehmood et al. highlight the benefits of the weighted average of triangular histograms using a support vector machine which adds the image spatial contents to the inverted index of the bag-of-visual-words model [20].

These studies can be considered as some example sources of information about the research questions on image annotation since they present relevant information.

3. RESEARCH METHOD

In this section, the research method overview is introduced and the aim, and research questions are discussed.

3.1. Overview

This research is conducted through a systematic mapping (SM) method which is implemented following the guidelines presented in [3-5]. Several other SM studies such as [22, 23] were also taken into consideration while conducting the SM.

For this purpose, the SM procedure begins with identifying articles from different academic sources. At that point, a methodical guide is produced, and then, the outcomes are reported. The SM procedure is presented in detail in Figure 1 below (see Figure 1).

3.2. Goal and Research Questions

The main goal of the paper is to conduct a systematic mapping study of image annotation research and identify the goals reflected in the research questions (RQs) presented in this paper such as contribution aspect of the selected studies, type and number of RQs raised in the studies, domains which have been analyzed in the papers together with quantitative analysis to provide an overview of image annotation literature.

The aim of this study is; identifying the most advanced level of image annotation on the Internet, determining prospects for future work, identifying research gaps, preparing a visual summary about the current trends providing guidance in the area. For this purpose, the Goal, Question, Metric (GQM) methodology [24] is followed.

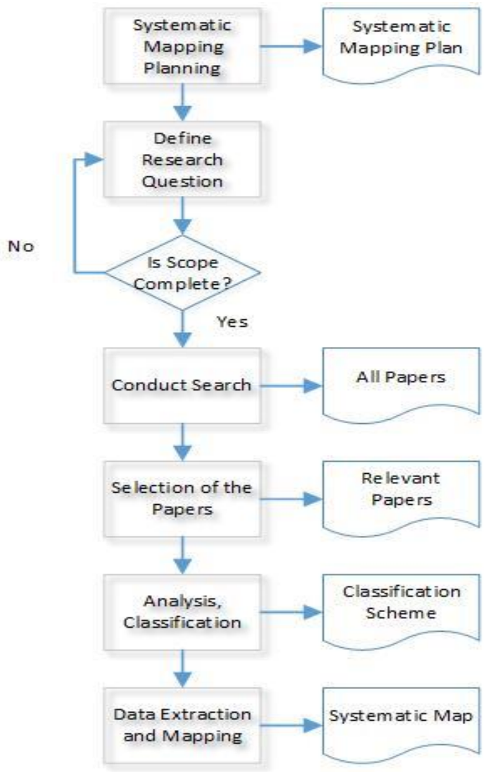

Figure 1. The systematic mapping procedure [25]

The following research questions (RQs) are raised based on the above goals.

RQ 1.1 – Research formulation by contribution aspect: How many works include image annotation strategies, methods, tools, models, measurements, or procedures?

These types of research commitments to lead methodical formulation research in software engineering is proposed by Petersen et al. [5] and this formulating approach has been used in most SM research, e.g., [26-28]. Responding to this RQ will let us determine whether there are possibilities of developing newer techniques or tools within the domain.

RQ 1.2 – Type and number of research questions raised in the empirical studies: What types of research questions have been addressed? Classification is performed using [29].

RQ 1.3 – Objects of study: What objects of study have been analyzed in the papers (e.g., datasets, domain, and other interesting aspects)?

Addressing this RQ will provide statistical information about study objects in the image annotation domain.

RQ 1.4 – How many studies present the image annotation process, image retrieval process, or both (e.g., automatic or semi-automatic image annotation, image retrieval, or both)?

RQ 1.5 – Which studies use ontologies?

RQ 2.1: What is the annual count of papers?

RQ 2.2: Which papers have been referenced the most?

The last two RQ's provide a quantitative analysis in the image annotation literature.

3.3. Search Strategy

To carry out an extensive search, some synonyms were selected. The studies which have been published in 2010 or after were chosen in conducting the SM. The resulting search-string was:

(semantic <or> ontology-based <or> high-level <or> top-level <or> upper-level <or> content-based <or> meaning-based <or> theme-based <or> large-scale <or> big-scale <or> massive-scale <or> tag-based <or> label-based <or> etiquette-based <or> automatic <or> robot <or> semi-automatic)

<and>

(image <or> picture <or> photo <or> painting <or> drawing <or> illustration <or> figure <or> shape)

<and>

(annotation <or> footnote <or> commentary <or> explanation <or> interpretation <or> classification <or> tagging <or> reranking)

<and>

(internet <or> web <or> semantic web).

3.4. Search Process

The main academic research engines on the internet used to find the relevant studies are: (1) IEEE Xplore, (2) ACM Digital Library, (3) SpringerLink, (4) ScienceDirect Elsevier, (5) Wiley Interscience, (6) Web of Science, (7) Scopus, (8) Google Scholar. Grey literature was not included since it is not cited properly, and not indexed comprehensively.

3.5. Inclusion and Exclusion Criteria

This research considers the criteria of the stage of extensiveness and assessment used in the studies and the criteria of whether a peer review is conducted. It is also decided that for multiple studies by the same author with the same title, the most recent one should be included.

The studies to be included should be written in English, be electronically available, and be published in 2010 or after.

Each candidate's study is carefully examined to see whether it is relevant to the field. The studies which lack technical depth about the field including different perspectives, strengths and weaknesses of the current proposals, are excluded. A study should include different academic perspectives within the literature and should provide adequate knowledge about the contribution aspects and feedback of the proposed system [7].

The decision of whether the articles in the initial pool would be in the category of inclusion or exclusion is made by the authors of the article. We first inspect the studies in the main pool, then perform a voting procedure on a 5-point scale for each one. '5' indicates a strong opinion for a study to be included, and '0' indicates a strong opinion for a study to be excluded. A 3-mark criteria is considered for the evaluation of the studies, and the studies which have a cumulative mark of 3 or less are excluded. Title, abstract and keywords of each paper are reviewed in the marking process. In case of a lack of substantial information in these resources, a more detailed evaluation is carried out. Grey literature is also excluded. As a result of the collaborative marking, 95 studies have been selected from a total of 404.

Table 1. The template column headings used for the classification scheme

Field Headings	Attributes
Paper Title	Title of the study
Database	Database Name
Bibliometric	Year, Venue Type, Venue Acronym, Author Affiliation, Number of Citation
RQ 1.1	Method/technique, Tool Used, Tool Developed, Model, Metric, Process
RQ 1.2	Exploratory, Base-rate, Relationship, Causality, Design, Other
RQ 1.3	Datasets, Domain, Other Interesting Aspects
RQ 1.4	Automatic, Semi-automatic, Retrieval, Both
RQ 1.5	Count
RQ 2.1	Annual count of papers
RQ 2.2	Citation numbers

3.6. Data Extraction

A template is developed to extract information about the papers including: (a) Paper ID, (b) Paper Title, (c) Publication Year, (d) Venue and Venue Acronym, (e) Author Affiliation, (f) Number of Citations, (g) Database in which the study was found. In Table 1, the field headings and their attributes are given.

4. RESULTS

4.1. Contribution Aspect

The top three contribution aspects are shown in Figure 2 which are method/technique (64,2%), model (20%) and tool developed (16,8%) which have been presented in 61 studies, 19 studies, and 16 studies, respectively. Following the figure, the references have also been given. The most cited research proposes a model that generates natural language descriptions of images and their regions [118].

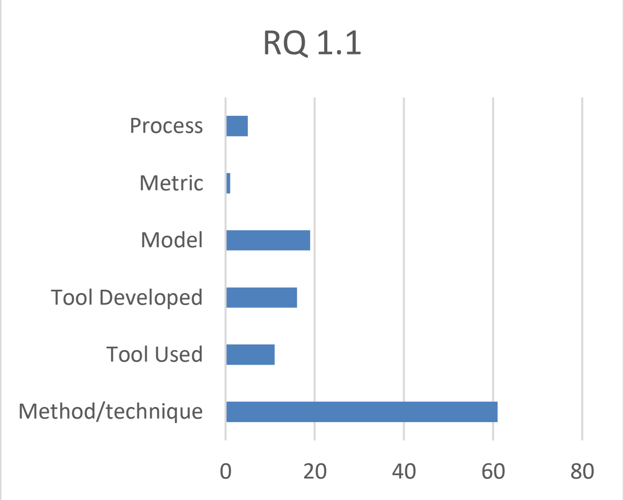

Figure 2. Contribution aspect distribution

Method / technique: [8-10, 12, 13, 18, 19, 32, 33, 35, 43, 45-48, 52-54, 56, 58-60, 63, 64, 66-73, 75-78, 80-85, 89, 92, 99, 102-113, 120]

Tool Used: [30, 37, 45, 49, 55, 62, 70, 77, 89, 90, 114]

Tool Developed: [14, 31, 34, 36, 38, 44, 47, 49, 50, 55, 57, 62, 74, 98, 100, 115]

Model: [15, 16, 39, 43, 48, 51, 52, 56, 61, 65, 75, 85, 86, 91, 116-119]

Metric: [81]

Process: [10, 55, 79, 82, 96]

4.2. Type and Number of Research Questions

The top three research question types are shown in Figure 3 which are design (38,9%), exploratory (30,5%), and causality (23,1%) which have been presented in 37 studies, 29 studies, and 22 studies, respectively. Following the figure, the relevant references are presented. One of the most cited research proposes an encoder-decoder pipeline that learns a multimodal joint embedding space with images and text [43].

4.3. Objects of Study

The domains mentioned in these studies are; Radiology, CORAL pictures, randomly selected images, satellite, real-world, gaming, internet images, images containing the packing cases of commercial products, medical images, images with social tags, plants, 3D content, news images, personal photographs, face images, human body, animals, endoscopic imaging, CT images, kidney biopsy images, and biomedical images. Figure 4 shows the distribution of the most widely used datasets in these studies. One of the most cited research presents the Visual Genome dataset to enable the modeling of interactions and relationships between objects in an image to achieve success at cognitive tasks [115]. NUS WIDE, Flickr30K and Flickr8K are the other datasets used by some of the most cited studies [12, 43, 118, 120].

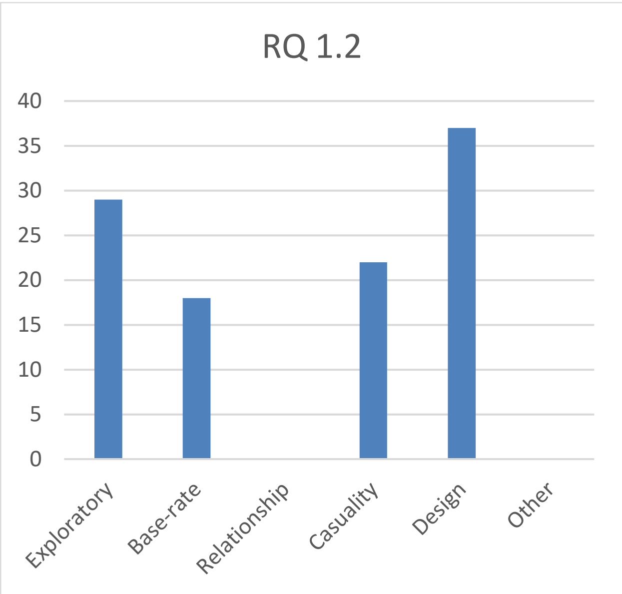

Figure 3. Research question type distribution

Base-rate: [30, 34, 37, 38, 43, 47, 50, 55, 59, 63, 66, 79, 84, 97, 98, 100, 104, 107]

Design: [8, 9, 14-16, 39, 43, 46, 48, 49, 51, 53, 56, 57, 60, 62, 65, 70-72, 74, 78, 80, 85, 86, 89, 102, 103, 109, 112, 113, 115-118, 120]

Causality: [10, 12, 15, 18, 35, 45, 52, 61, 63, 67, 73, 81, 82, 84, 93, 94, 96, 99, 105, 106, 110, 119]

Exploratory: [13, 19, 20, 30, 31-33, 36, 44, 49, 54, 58, 59, 64, 68, 69, 75-77, 82, 83, 85, 89-92, 108, 111, 114]

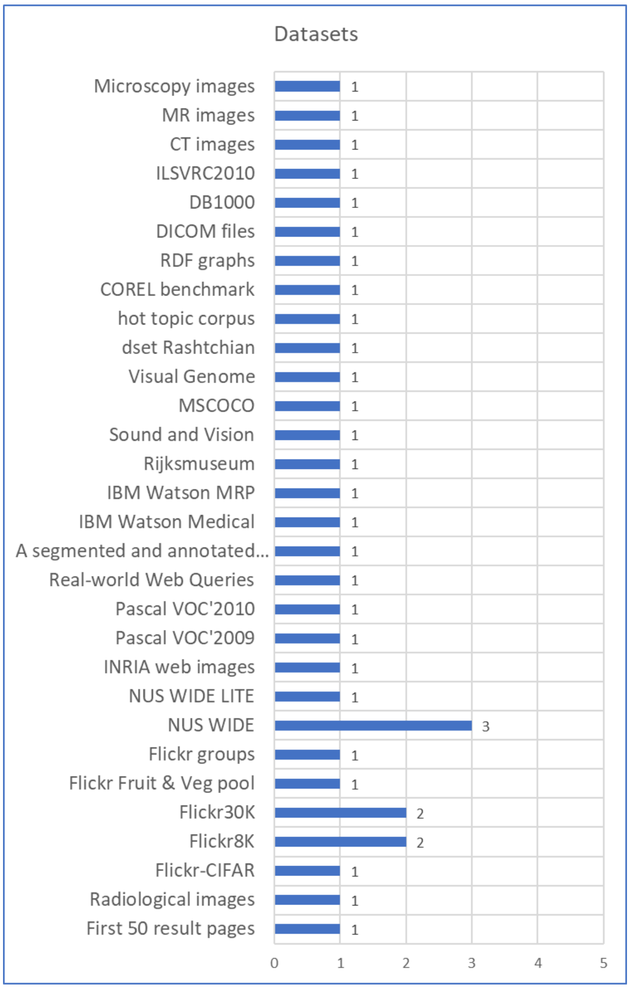

Figure 4. Datasets distribution shown in the studies

Flickr30K: [43], [118] RDF graphs: [30]
Flickr Fruit & Veg pool: [31] DICOM files: [31]
Flickr groups: [40] DB1000: [32]
NUS WIDE: [8], [12], [120] ILSVRC2010: [14]
NUS WIDE LITE: [13] CT images: [35]
First 50 result pages: [41] Visual Genome: [115]
Radiological images: [49] dset Rashtchian: [119]
Flickr-CIFAR: [120] hot topic corpus: [8]
Flickr8K: [43], [118] CORELbenchmark: [13]
INRIA web images: [120] MR images: [36]
Pascal VOC'2009: [9] Microscopyimages: [38]
Pascal VOC'2010: [9] Sound and Vision: [97]
Real-world Web Queries: [105] MSCOCO: [118]
"A segmented and annotated image dataset": [10]
IBM Watson Medical: [97] IBM Watson MRP: [97]
Rijksmuseum: [97]

4.4. The Studies Covering Image Annotation, Image Retrieval or Both

The top three results are shown in Figure 5 which are automatic (70,5%), both "automatic<or>semi-automatic <and> retrieval" (26,3%), and semi-automatic (24,2%) which have been presented in 67 studies, 25 studies, and 23 studies, respectively. The references are given following the figure. The studies which propose automatic annotation approaches are the most cited ones [43, 115, 118, 119].

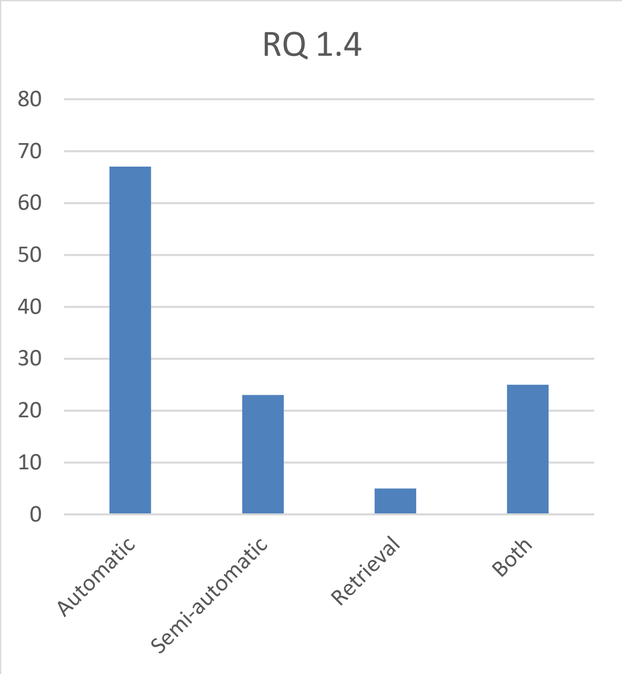

Figure 5. Study direction distribution

Automatic: [8-10, 12-14, 18, 19, 30, 32, 35, 38-40, 43, 44, 46, 48, 49, 51, 52, 54-56, 59, 63-73, 75-80, 82, 84, 85, 89, 91-93, 99, 102-107, 109-113, 115, 117-120]

Both: [10, 14, 15, 19, 30, 35, 39, 40, 47, 51-53, 55, 62, 75, 80, 82, 91, 96, 99, 100, 114, 115, 117, 120]

Semi-automatic: [16, 31, 34, 36, 39, 41, 46, 49, 50, 53, 57, 58, 60, 62, 83, 86, 90, 96, 98, 100, 103, 104, 114]

Retrieval: [45, 81, 88, 108, 116]

4.5. Number of Ontology Related Papers

The total number of studies that are directly or indirectly related to ontologies is 28 which amounts to 29,5%. The references are provided below. One of the most cited research proposes an object-based semantic classification method for high-resolution satellite imagery using an ontology that aims to take advantage of geographic object-based image analysis techniques [103].

[9, 30, 31, 35-37, 39, 47, 50, 55, 59-61, 66, 71, 72, 81, 83, 86, 90, 96, 100, 102-104, 107, 113, 114]

4.6. Annual Count of Papers

The count of papers between the years 2010 and 2018 is shown in Figure 6. First is 2014 (22,1%), the second is 2013 (20%) and third is 2015 (13,7%) which have appeared in 21 studies, 19 studies, and 13 studies, respectively. The first two of the most cited papers have been published in 2017 [115, 118].

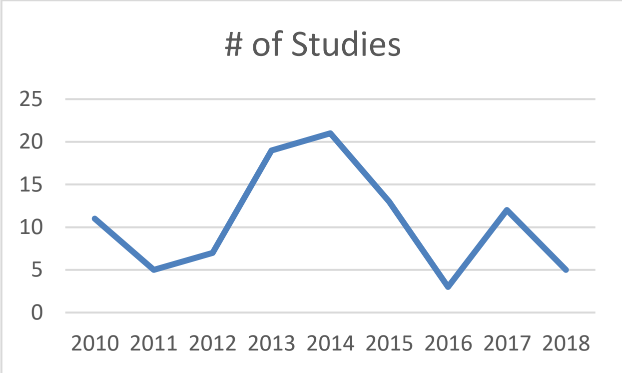

Figure 6. Annual Count of Papers between 2010 and 2018

4.7. Citation Number

The number of citations for the studies (for citation count > 100) is shown in Figure 7. By a large margin, the first place belongs to the study [118], second place belongs to [115] and third is [119] which have 1060 citations, 782 citations, and 341 citations, respectively.

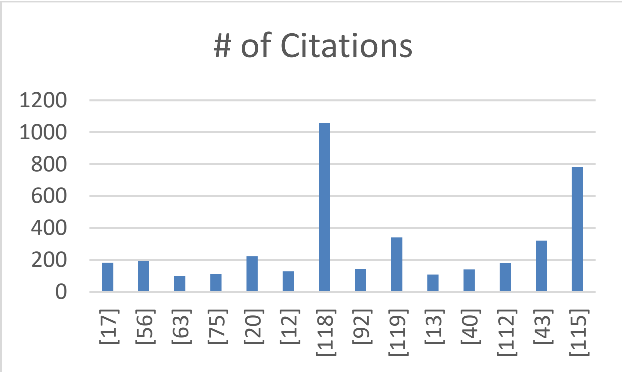

Figure 7. Citation Numbers of the Studies

5. CONCLUSION AND FUTURE WORK

Various technical approaches, research methods, research question types, and datasets have been identified from assorted studies associated with the related research questions in this study. A number of studies have been reviewed in the literature via a systematic mapping process to present the distributions among them according to the classification scheme. This study aims to support the decision-making process of researchers working in this field. In this systematic mapping research, 404 studies are identified from which 95 are selected to represent the sample set. Several studies are found to be highly focused on methods/techniques. Yet in terms of the most effective method, a consensus has not been reached. It is for this reason that a variety of methods, techniques, tools and processes coexist in the majority of the examined studies. Interdisciplinary approaches should also be considered in developing new solutions.

Some of the studies highlight tools. From this perspective, the newly developed tools for image annotation have majority when compared to the Commercial off-the-shelf (COTS) software. There are also problems in conducting a systematic mapping research in this domain, such as the utilized techniques. Some of them are adopted to cover just image annotation and some can be used for image retrieval as well. While classifying, these techniques may be misjudged. The majority of the papers belong to 2014, because of the accelerating technological development occurring in the image annotation domain in that year. For further analysis, there is a need for sub-classifications in some parts. For instance, other categories under the type of research method can be defined. There is also a need for categorization of the influencing image annotation factors because they are too generalized in the literature.

In future work, the findings of this study can be utilized to support sectoral image annotation projects to design new methods, techniques, tools, models, metrics, and processes based on the previous literature.

REFERENCES

[1] V. Lavrenko, R. Manmatha, J. Jeon, "A Model for Learning the Semantics of Pictures", **Advances in Neural Information Processing Systems 17**, Vancouver, Canada, 553-560, 13-18 December, 2004.

[2] Internet: Halaschek, C. Wiener, Image Annotation on the Semantic Web, https://www.w3.org/2005/Incubator/mmsem/XGR-image-annotation/#annot_intro, 03.05.2019.

[3] B. Kitchenham, S. Charters, **Guidelines for Performing Systematic Literature Reviews in Software Engineering**, Keele University, UK, 2007.

[4] B. Kitchenham, et al., "Systematic literature reviews in software engineering – A systematic literature review", *Information and Software Technology*, 51(1), 7-15, 2009.

[5] K. Petersen, et al., "Systematic Mapping Studies in Software Engineering", **12th International Conference on Evaluation and Assessment in Software Engineering**, Bari, Italy, 68-77, 26-27 June, 2008.

[6] J. Jill, L. Matheson, F. M. Lacey, **Doing Your Literature Review: traditional and systematic techniques**, SAGE Publishing, London, UK, 2011.

[7] Internet: A. Sezen, Ç. Turhan, SM – Classification Scheme, https://drive.google.com/open?id=1cnItNure_P7oolIcoVmDqP5EXJMim00-HPOweH_qPDA, 10.10.2019.

[8] X. Wang, et al., "High-level semantic image annotation based on hot Internet topics", *Multimedia Tools and Applications*, 74(6), 2055-2084, 2015.

[9] H. Bannour, C. Hudelot, "Building and using fuzzy multimedia ontologies for semantic image annotation", *Multimedia Tools and Applications*, 72(3), 2107-2141, 2014.

[10] A. Fakhari, A. M. E. Moghadam, "Combination of classification and regression in decision tree for multi-labeling image annotation and retrieval", *Applied Soft Computing*, 13(2), 1292-1302, 2013.

[11] F. Kınlı, F. Kıraç, "FashionCapsNet: Clothing Classification with Capsule Networks", *International Journal of Informatics Technologies*, 13(1), 87-96, 2020.

[12] Internet: Y. Gong, et al., Deep Convolutional Ranking for Multilabel Image Annotation, https://arxiv.org/abs/1312.4894, 03.05.2019.

[13] R. Hong, et al., "Image Annotation by Multiple-Instance Learning With Discriminative Feature Mapping and Selection", *IEEE Transactions on Cybernetics*, 44(5), 669-680, 2014.

[14] S. Zhang, et al., "ObjectPatchNet: Towards scalable and semantic image annotation and retrieval", *Computer Vision and Image Understanding*, 118(1), 16-29, 2014.

[15] D-H. Im, G-D. Park, "STAG: Semantic Image Annotation Using Relationships between Tags", **International Conference on Information Science and Applications**, Suwon, South Korea, 1-2, 24-26 June, 2013.

[16] L. Seneviratne, E. Izquierdo, "A mathematical approach towards semi-automatic image annotation", 19th European Signal Processing Conference, Barcelona, Spain, 559-563, 29 August-2 September, 2011.

[17] A. Ç. Seçkin, "Labeling System with EEG, EMG, and IMU for Visual Training of Autonomous Vehicles", *International Journal of Informatics Technologies*, 12(4), 299-305, 2019.

[18] S-B. Chan, H. Yamana, S. Satoh, "A Negative Sample Image Selection Method Referring to Semantic Hierarchical Structure for Image Annotation", **International Conference on Signal-Image Technology & Internet-Based Systems**, Kyoto, Japan, 162-167, 2-5 December, 2013.

[19] X. Liu, H. Yao, R. Ji, "Exploring statistical properties for semantic annotation: sparse distributed and convergent assumptions for keywords", **IEEE International Conference on Acoustics, Speech and Signal Processing**, Dallas, USA, 802-805, 14-19 March, 2010.

[20] Z. Mehmood, T. Mahmood, M. A. Javid, "Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine", *Applied Intelligence*, 48(1), 166-181, 2018.

[21] F. Bulut, "Performance Analysis of Ensemble Methods on Imbalanced Datasets", *International Journal of Informatics Technologies*, 9(2), 153-159, 2016.

[22] S. Ali, et al., "A Systematic Review of the Application and Empirical Investigation of Search-Based Test Case Generation", *IEEE Transactions on Software Engineering*, 36(6), 742-762, 2010.

[23] F. Paz, J. A. Pow-Sang, "Systematic Mapping Review of Usability Evaluation Methods for Software Development Process". International Journal of Software Engineering and Its Applications, 10(1), 165-178, 2016.

[24] V. R. Basili, Software Modeling and Measurement: The Goal/Question/Metric Paradigm, University of Maryland, USA, 1992.

[25] F. A. Silva, et al., "Benchmark applications used in mobile cloud computing research: a systematic mapping study", *The Journal of Supercomputing*, 72(4), 1431-1452, 2016.

[26] Internet: Y. Gong, et al., Deep Convolutional Ranking for Multilabel Image Annotation, https://arxiv.org/abs/1312.4894, 03.05.2019.

[27] I. Banerjee, et al., "Graphical user interface (GUI) testing: Systematic mapping and repository", *Information and Software Technology*, 55(10), 1679-1694, 2013.

[28] N. Bencomo, S. Hallsteinsen, E. S. de Almeida, "A View of the Dynamic Software Product Line Landscape", *IEEE Computer*, 45(10), 36-41, 2012.

[29] S. Easterbrook, et al., "Selecting Empirical Methods for Software Engineering Research", **Guide to Advanced Empirical Software Engineering**, Editor: F. Shull, J. Singer, D. I. K. Sjøberg, Springer, London, UK, 285-311, 2008.

[30] T. Kanimozhi, T., A. Christy, "Incorporating ontology and SPARQL for semantic image annotation", **IEEE Conference on Information & Communication Technologies**, Thuckalay, Tamil Nadu, India, 26-31, 11-12 April, 2013.

[31] D-H. Im, G-D Park, "Linked tag: image annotation using semantic relationships between image tags", *Multimedia Tools and Applications*, 74(7), 2273-2287, 2015.

[32] D. Wang, et al., "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation", *IEEE Transactions on Knowledge and Data Engineering*, 26(1), 166-179, 2014.

[33] T. Sjekavica, G. Gledec, M. Horvat, "Multimedia Annotation using Semantic Web Technologies", **7th European Computing Conference**, Dubrovnik, Croatia, 228-233, 25-27 June, 2013.

[34] Internet: G. Ciocca, P. Napoletano, R. Schettini, IAT - Image Annotation Tool: Manual, https://arxiv.org/abs/1502.05212, 17.05.2019.

[35] C. Kurtz, et al., "On combining image-based and ontological semantic dissimilarities for medical image retrieval applications", *Medical Image Analysis*, 18(7), 1082-1100, 2014.

[36] K. C. Wang, et al., "Ontology-based Image Navigation: Exploring 3.0-T MR Neurography of the Brachial Plexus Using AIM and RadLex", *RadioGraphics*, 35(1), 142-151, 2015.

[37] M. Grassi, et al., "Pundit: Creating, Exploring and Consuming Semantic Annotations", **3rd International Workshop on Semantic Digital Archives**, Valetta, Malta, 65-72, 26 September, 2013.

[38] D. K. Iakovidis, et al., "Ratsnake: A Versatile Image Annotation Tool with Application to Computer-Aided Diagnosis", *The Scientific World Journal*, 2014(286856), 1-12, 2014.

[39] A. C. Bukhari, M. Krauthammer, C. J. O. Baker, "SEBI: An Architecture for Biomedical Image Discovery, Interoperability and Reusability based on Semantic Enrichment", **International SWAT4LS Workshop**, Berlin, Germany, 1-6, 9-11 December, 2014.

[40] H. Chuanping, et al., "Semantic Link Network-Based Model for Organizing Multimedia Big Data", *IEEE Transactions on Emerging Topics in Computing*, 2(3), 376-387, 2014.

[41] S. Alcic, S. Conrad, "2-DOM: A 2-Dimensional Object Model towards Web Image Annotation", **Third International Workshop on Semantic Media Adaptation and Personalization**, Prague, Czech Republic, 3-8, 15-16 December, 2008.

[42] L. Nixon, R. Troncy, "Survey of Semantic Media Annotation Tools for the Web: Towards New Media Applications with Linked Media", **European Semantic Web Conference**, Anissaras, Crete, 110-114, 16 October, 2014.

[43] Internet: R. Kiros, R. Salakhutdinov, R. S. Zemel, Unifying Visual-Semantic Embeddings with Multimodal Neural Language Models, https://arxiv.org/abs/1411.2539, 17.05.2019.

[44] M. Bakalem, N. Benblidia, S. Ait-Aoudia, "A comparative image auto-annotation", **IEEE International Symposium on Signal Processing and Information Technology**, Athens, Greece, 000086-000091, 12-15 December, 2013.

[45] D. Zhao, W. Zou, G. Sun, "A fast image classification algorithm using Support Vector Machine", **2nd International Conference on Computer Technology and Development**, Cairo, Egypt, 385-388, 2-4 November, 2010.

[46] Irfanullah, et al., "A framework for high level semantic annotation using trusted object annotated dataset", **The 10th IEEE International Symposium on Signal Processing and Information Technology**, Luxor, Egypt, 491-495, 15-18 December, 2010.

[47] J. M. Almendros-Jiménez, L. Domene, J. A. Piedra-Fernández, "A Framework for Ocean Satellite Image Classification Based on Ontologies", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2), 1048-1063, 2013.

[48] X. Li, et al., "A Locality Sensitive Low-Rank Model for Image Tag Completion", *IEEE Transactions on Multimedia*, 18(3), 474-483, 2016.

[49] D. A. Moreira, et al., "3D Markup of Radiological Images in ePAD, a Web-Based Image Annotation Tool", **IEEE 28th International Symposium on Computer-Based Medical Systems**, Sao Carlos, Brazil, 97-102, 22-25 June, 2015.

[50] N. T. Lingutla, et al., "AISO: Annotation of Image Segments with Ontologies", *Journal of Biomedical Semantics*, 5(50), 2014.

[51] S. R. Kakade, N. R. Kakade, "A novel approach to link semantic gap between images and tags via probabilistic ranking", **IEEE International Conference on Computational Intelligence and Computing Research**, Enathi, India, 1-6, 26-28 December, 2013.

[52] C. Wang, et al., "A novel image tag saliency ranking algorithm based on sparse representation", **Visual Communications and Image Processing**, Kuching, Malaysia, 1-5, 17-20 November, 2013.

[53] H-K. Hong, K-W. Park, D-H. Lee, "A Novel Semantic Tagging Technique Exploiting Wikipedia-Based Associated Words", **IEEE 39th Annual Computer Software and Applications Conference**, Taichung, Taiwan, Vol. 3, 648-649, 1-5 July, 2015.

[54] J. Xu, S. Okada, K. Nitta, "A Semantic-Similarity-Based Method for Object Description and Clustering", **IEEE International Conference on Systems, Man, and Cybernetics**, Manchester, UK, 3669-3674, 13-16 October, 2013.

[55] G. K. Kyriazos, et al., "A semantically-aided approach for online annotation and retrieval of medical images", **International Conference of the IEEE Engineering in Medicine and Biology**, Boston, USA, 2372-2375, 30 August-3 September, 2011.

[56] P. Symeonidis, A. Nanopoulos, Y. Manolopoulos, "A Unified Framework for Providing Recommendations in Social Tagging Systems Based on Ternary Semantic Analysis", *IEEE Transactions on Knowledge and Data Engineering*, 22(2), 179-192, 2010.

[57] J-Y. Shih, I. Huang, S-W. Chan, "A Web-Based Annotation System to Conducting Learning Activities for Mammography", **International Conference on Educational Innovation through Technology**, Tainan, Taiwan, 218-221, 22-24 September, 2016.

[58] M. C. You, et al., "A web-based semi-automated method for semantic annotation of high schools in remote sensing images", **Third International Conference on Agro-Geoinformatics**, Beijing, China, 1-4, 11-14 August, 2014.

[59] H. Zarzour, M. Sellami, "Achieving consistency in collaborative image annotation systems", **5th International Conference on Information and Communication Systems**, Irbid, Jordan, 1-7, 1-3 April, 2014.

[60] M. Shi, H. Cai, L. Jiang, "An approach to semi-automatic semantic annotation on Web3D scenes based on an ontology framework", **12th International Conference on Intelligent Systems Design and Applications**, Kochi, India, 574-579, 27-29 November, 2012.

[61] Z. Kastrati, A. S. Imran, S. Y. Yayilgan, "An Improved Concept Vector Space Model for Ontology Based Classification", **11th International Conference on Signal-Image Technology & Internet-Based Systems**, Bangkok, Thailand, 240-245, 23-27 November 2015.

[62] Q. Wu, P. Boulanger, "An Unified Image Tagging System Driven by Image-Click-Ads Framework", **IEEE International Symposium on Multimedia**, Miami, USA, 369-372, 14-16 December, 2015.

[63] X-J. Wang, et al., "ARISTA - image search to annotation on billions of web photos", **IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, San Francisco, USA, 2987-2994, 13-18 June, 2010.

[64] A. S. Reddy, et al., "Automatic caption generation for annotated images by using clustering algorithm", **International Conference on Innovations in Information, Embedded and Communication Systems**, Coimbatore, India, 1-5, 19-20 March, 2015.

[65] Y. Feng, M. Lapata, "Automatic Caption Generation for News Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(4), 797-812, 2013.

[66] E. F. Luque, D. L. Rubin, D. A. Moreira, "Automatic Classification of Cancer Tumors Using Image Annotations and Ontologies", **IEEE 28th International Symposium on Computer-Based Medical Systems**, Sao Carlos, Brazil, 368-369, 22-25 June, 2015.

[67] M. Zachariasova, et al., "Automatic extraction of non-textual information in web document and their classification", **35th International Conference on Telecommunications and Signal Processing**, Prague, Czech Republic, 753-757, 3-4 July, 2012.

[68] F. Tian, et al., "Automatic Image Annotation Based on Sparse Representation and Multiple Label Learning", **International Conference on Virtual Reality and Visualization**, Qinhuangdao, China, 27-32, 14-15 September, 2012.

[69] G. Chu, K. Niu, B. Tian, "Automatic image annotation combining SVMS and KNN algorithm", **IEEE 3rd International Conference on Cloud Computing and Intelligence Systems**, Shenzhen, China, 13-17, 27-29 November, 2014.

[70] Y-F. Huang, H-Y. Lu, "Automatic Image Annotation Using Multi-object Identification", **Fourth Pacific-Rim Symposium on Image and Video Technology**, Singapore, Singapore, 386-392, 14-17 November, 2010.

[71] G. Ding, N. Xu, "Automatic semantic annotation of images based on Web data", **Sixth International Conference on Information Assurance and Security**, Atlanta, USA, 317-322, 23-25 August, 2010.

[72] J. Söderberg, E. Kakogianni, "Automatic tag generation for photos using contextual information and description logics", **International Workshop on Content Based Multimedia Indexing**, Grenoble, France, 1-7, 23-25 June, 2010.

[73] Y-Y. Chen, W. H. Hsu, H-Y. M. Liao, "Automatic Training Image Acquisition and Effective Feature Selection From Community-Contributed Photos for Facial Attribute Detection", *IEEE Transactions on Multimedia*, 15(6), 1388-1399, 2013.

[74] A. A. Olaode, G. Naghdy, C. A. Todd, "Bag-of-Visual Words Codebook Development for the Semantic Content Based Annotation of Images", **11th International Conference on Signal-Image Technology & Internet-Based Systems**, Bangkok, Thailand, 7-14, 23-27 November, 2015.

[75] H. Ma, et al. "Bridging the Semantic Gap Between Image Contents and Tags", *IEEE Transactions on Multimedia*, 12(5), 462-473, 2010.

[76] C. Hongjun, Z. Junfeng, "Classification of image based on semantic features and Bayesian networks", **International Conference on Consumer Electronics, Communications and Networks**, XianNing, China, 4858-4861, 16-18 April, 2011.

[77] J. Jing, et al., "Cognition-Based Semantic Annotation for Web Images", **Fourth International Conference on Big Data and Cloud Computing**, Sydney, Australia, 540-546, 3-5 December, 2014.

[78] M. Oujaoura, B. Minaoui, M. Fakir, "Combined descriptors and classifiers for automatic image annotation", **International Conference on Multimedia Computing and Systems**, Marrakech, Morocco, 270-276, 14-16 April, 2014.

[79] Y. Gong, S. Lazebnik, "Comparing data-dependent and data-independent embeddings for classification and ranking of internet images", **Conference on Computer Vision and Pattern Recognition**, Colorado Springs, USA, 2633-2640, 20-25 June, 2011.

[80] Y. Lu, et al., "Constructing Concept Lexica With Small Semantic Gaps", *IEEE Transactions on Multimedia*, 12(4), 288-299, 2010.

[81] V. Franzoni, et al., "Context-based image semantic similarity", **12th International Conference on Fuzzy Systems and Knowledge Discovery**, Zhangjiajie, China, 1280-1284, 15-17 August, 2015.

[82] M. Shiyamala, G. Kalaiarasi, "Contextual image search with keyword and image input", **International Conference on Information Communication and Embedded Systems**, Chennai, India, 1-5, 27-28 February, 2014.

[83] D. T. Parry, T-C. Tsai, "Crowdsourcing techniques to create a fuzzy subset of SNOMED CT for semantic tagging of medical documents", *Soft Computing*, 16(7), 1119-1127, 2012.

[84] Z. Ji, “Decoupling Sparse Coding with Fusion of Fisher Vectors and Scalable SVMs for Large-Scale Visual Recognition”, **IEEE Conference on Computer Vision and Pattern Recognition**, Portland, USA, 450-457, 23-28 June, 2013.

[85] M. Volpi, D. Tuia, “Dense Semantic Labeling of Subdecimeter Resolution Images With Convolutional Neural Networks”, *IEEE Transactions on Geoscience and Remote Sensing,* 55(2), 881-893, 2017.

[86] N. C. Kuicheu, et al., “Description logic based icons semantics: An Ontology for Icons”, **IEEE 11th International Conference on Signal Processing**, Beijing, China, 1260-1263, 21-25 October, 2012.

[87] S. P. Malewar, H. Dakhore, “Design approach of visual image detection in rescue robot system of urban search using Bayesian's logical algorithm”, **International Conference on Innovations in Information, Embedded and Communication Systems**, Coimbatore, India, 1-5, 19-20 March, 2015.

[88] S. Zhuo, et al., “Distributed File System and Classification for Small Images”, **IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing**, Beijing, China, 2231-2234, 20-23 August, 2013.

[89] X-J. Wang, L. Zhang, W-Y. Ma, “Duplicate-Search-Based Image Annotation Using Web-Scale Data”, Proceedings of the IEEE, 100(9), 2705-2721, 2012.

[90] D. K. Iakovidis, C. V. Smailis, “Efficient semantically-aware annotation of images”, **IEEE International Conference on Imaging Systems and Techniques**, Penang, Malaysia, 146-149, 17-18 May, 2011.

[91] L. Dai, et al., “Efficient Tag Mining via Mixture Modeling for Real-Time Search-Based Image Annotation”, **IEEE International Conference on Multimedia and Expo**, Melbourne, Australia, 134-139, 9-13 July, 2012.

[92] Z. Akata, et al., “Evaluation of Output Embeddings for Fine-Grained Image Classification”, **IEEE Conference on Computer Vision and Pattern Recognition**, Boston, USA, 2927-2936, 7-12 June, 2015.

[93] M. Morchid, R. Dufour, G. Linares, “Event detection from image hosting services by slightly-supervised multi-span context models”, **11th International Workshop on Content-Based Multimedia Indexing**, Veszprem, Hungary, 103-107, 17-19 June, 2013.

[94] M. Katsurai, M. Haseyama, “Exploring and visualizing tag relationships in photo sharing websites based on distributional representations”, **IEEE International Conference on Acoustics, Speech and Signal Processing**, Vancouver, Canada, 3617-3621, 26-31 May, 2013.

[95] V. Garousi, “Classification and trend analysis of UML books (1997-2009)”, Software & System Modeling (SoSyM), 11(2), 273-285, 2012.

[96] R. Gaur, D. K. Sharma, “Focused crawling with ontology using semi-automatic tagging for relevancy”, **Seventh International Conference on Contemporary Computing**, Noida, India, 501-506, 7-9 August, 2014.

[97] O. Inel, et al., “CrowdTruth: Machine-Human Computation Framework for Harnessing Disagreement in Gathering Annotated Data”, **International Semantic Web Conference**, Riva Del Garda, Italy, 486-504, 19-23 October, 2014.

[98] V. de Boer, et al., “DIVE into the event-based browsing of linked historical media”, *Journal of Web Semantics*, 35(3), 152-158, 2015.

[99] L. Wang, L. Khan, “Automatic image annotation and retrieval using weighted feature selection”, *Multimedia Tools and Applications*, 29(1), 55-71, 2006.

[100] Y. Chabane, et al., “Medical data management in the SYSEO project”, *ACM SIGMOD Record*, 42(3), 48-53, 2013.

[101] R. Fujimoto, et al., “Generation of Stratified Image Database with Web Image Sharing Service and Ontology”, **3rd International Conference on Advanced Applied Informatics**, Kitakyushu, Japan, 966-971, 31 August-4 September, 2014.

[102] T. Helmy, “A Generic Framework for Semantic Annotation of Images”, *International Journal of Image and Graphics*, 18(03), 1850013, 2018.

[103] H. Gu, et al., “An Object-Based Semantic Classification Method for High Resolution Remote Sensing Imagery Using Ontology”, *Remote Sensing*, 9(4), 329, 2017.

[104] D. Goel, et al., “An ontology-driven context aware framework for smart traffic monitoring”, **IEEE Region 10 Symposium**, Cochin, India, 1-5, 14-16 July, 2017.

[105] J. Yu, Y. Rui, D. Tao, “Click Prediction for Web Image Reranking Using Multimodal Sparse Coding”, *IEEE Transactions On Image Processing*, 23(5), 2019-2032., 2014.

[106] J. Wang, et al., “Image Semantic Segmentation Algorithm Based on Self-learning Super-Pixel Feature Extraction”, **International Conference on Emerging Internetworking, Data & Web Technologies**, Tirana, Albania, 773-781, 15-17 March, 2018.

[107] S. Baier, Y. Ma, V. Tresp, “Improving Visual Relationship Detection Using Semantic Modeling of Scene Descriptions”, **International Semantic Web Conference**, Vienna, Austria, 53-68, 21-25 October, 2017.

[108] C. Huang, et al., “Large-scale semantic web image retrieval using bimodal deep learning techniques”, *Information Sciences*, 430(2018), 331-348, 2018.

[109] M. Hu, et al., “Robust Web Image Annotation via Exploring Multi-Facet and Structural Knowledge”, *IEEE Transactions on Image Processing*, 26(10), 4871-4884, 2017.

[110] X. Wang, et al., “Semantic annotation for complex video street views based on 2D–3D multi-feature fusion and aggregated boosting decision forests”, *Pattern Recognition*, 62(2017), 189-201, 2017.

[111] P. Phursutkar, K. Wanjale, “Social re-ranking of image based on visual and semantic information”, **8th International Conference on Computing, Communication and Networking Technologies**, Delhi, India, 1-6, 3-5 July, 2017.

[112] Y. Wei, et al., “STC: A Simple to Complex Framework for Weakly-Supervised Semantic Segmentation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(11), 2314-2320, 2017.

[113] P. Wang, et al., “Training-free indexing refinement for visual media via multi-semantics”, *Neurocomputing*, 236(2017), 39-47, 2017.

[114] L. F. Sikos, “RDF-powered semantic video annotation tools with concept mapping to Linked Data for next-generation video indexing: a comprehensive review”, *Multimedia Tools and Applications*, 76(12), 14437-14460, 2017.

[115] R. Krishna, et al., “Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations”, *International Journal of Computer Vision*, 123(1), 32-73, 2017.

[116] Internet: P-M. Cheung, et al., Image annotation using aggregated page information from active and inactive indices, https://patents.google.com/patent/US9697236B2/en, 29.08.2019.

[117] L. Zheng, Z. Caiming, C. Caixian, “MMDF-LDA: An improved Multi-Modal Latent Dirichlet Allocation model for social image annotation”, *Expert Systems with Applications*, 104(2018), 168-184, 2018.

[118] A. Karpathy, L. Fei-Fei, “Deep Visual-Semantic Alignments for Generating Image Descriptions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 664-676, 2017.

[119] R. Socher, et al., “Grounded Compositional Semantics for Finding and Describing Images with Sentences”, *Transactions of the Association for Computational Linguistics*, 2(2014), 207-218, 2014.

[120] G. Yunchao, et al., “A Multi-View Embedding Space for Modeling Internet Images, Tags, and Their Semantics”, *International Journal of Computer Vision*, 106(2), 210-233, 2014.