



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكرو فيلم

بسم الله الرحمن الرحيم



HANAA ALY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم



شبكة المعلومات الجامعية التوثيق الإلكتروني والميكروفيلم



HANAA ALY



شبكة المعلومات الجامعية
التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغييرات



يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار



HANAA ALY



Knowledge-based Image Representation and Retrieval

A thesis submitted to the Department of Information Systems, Faculty of
Computer and Information Sciences, Ain Shams University. In partial
fulfilment of the requirements for the degree of Master of Science in Computer
and Information Sciences

By:

Haitham Samih Atea

Supervised by:

Prof. Tarek F. Gharib

Head of Information Systems Department
Faculty of Computer and Information
Sciences Ain Shams University

Prof. Manal Ahmed Ismail

Computer Engineering and Systems Department
Faculty of Engineering
Helwan University

Dr. Sherine Rady Abdel Ghany

Associate Professor in Information Systems
Department
Faculty of Computer and Information Sciences
Ain Shams University

**Ain Shams University, Cairo, Egypt
December, 2020**



Knowledge-based Image Representation and Retrieval

A thesis submitted to the Department of Information Systems, Faculty of Computer and Information Sciences, Ain Shams University. In partial fulfilment of the requirements for the degree of Master of Science in Computer and Information Sciences

By:

Haitham Samih Atea

Supervised by:

Prof. Tarek F. Gharib

Head of Information Systems Department Faculty
of Computer and Information Sciences
Ain Shams University

Prof. Manal Ahmed Ismail

Computer Engineering and Systems Department
Faculty of Engineering
Helwan University

Dr. Sherine Rady Abdel Ghany

Associate Professor in Information Systems Department
Faculty of Computer and Information Sciences
Ain Shams University

December, 2020

To my dear family...

Acknowledgements

In the beginning, I bow head thanking to “Allah” who paved the way and only by HIS will everything can be achieved.

*I would like to express my deeply grateful and respect to **Prof. Tarek F. Ghraib** for giving me this opportunity to work under his esteemed guidance, motivation and support. I also wish to express profound gratitude for his constant encouragement and critical discussions throughout this research program and during the preparation of this thesis.*

*Deep thanks and sincere gratitude with appreciation to **Dr. Sherine Rady**. This work would not have been possible without her guidance, support and encouragement. For her effective supervision, helpful comments, and the extensive time she devoted to this work, providing many facilities during preparation, experimental measurements, and scientific supervision that helped to accomplish this research.*

*Deep thanks and sincere gratitude with appreciation to **Prof. Manal Ismail** for her continuous advices, useful comments and scientific supervision throughout this research.*

Also, thanks to my friends and my colleagues in the Egyptian E-Learning University for their fruitful help and support.

Finally, no words can describe how thankful I’m to my family, my lovely wife for their help, encouragement and inspiration to accomplish all this work.

Abstract

Analyzing complex natural language queries through image/multimedia search engines remains a big challenge. Traditional text-based retrieval systems associate textual descriptions with each image, based on subjective human perception. These descriptions are next matched lexically against the user interrogated queries. Such annotation-based paradigm does not achieve the best results, since the lexical comparison is not sufficient for matching sentences in a semantic manner. Combining image retrieval processing with rich semantics and knowledge-based modeling provide promising solutions towards better image search engines.

This thesis proposes a knowledge-based image representation and retrieval which integrates external knowledge sources for obtaining a higher-level inference that can both handle complex natural language queries and increase the number of relevant retrievals for image search engines.

The thesis presents two solution approaches for the purpose of enhancing image retrieval. The first solution proposes a semantic framework for image representation and retrieval that can efficiently handle complex human-wise queries. The second solution proposes a semantic evaluation for auto-generated image annotations based on similarity measurement.

In the proposed image representation and retrieval framework, semantics are integrated by employing external knowledge sources and query expansion in the retrieval process. A set of developed and off the shelf parsing tools are used to obtain a full semantic understanding for relating the natural language queries and image annotations. The user query is parsed and next fused with the external knowledge sources in a query expansion process to infer supplemental knowledge about the terms of the query and hence increasing the

searchability of the query over image captions while increasing the average Recall. For providing relevance, a relation similarity metric is proposed to rank retrieved images based on their similarity scores.

For the purpose of evaluation of auto-generated image annotations, a semantic evaluation metric that measures the similarities of the generated annotations towards a set of human-written reference annotations is used.

Experiments are conducted on the Flickr datasets using a large set of natural language queries. The proposed solutions have been compared versus existing related systems for different evaluation measures: Number of Retrievals (Recall), Retrieval Accuracy, expansion rule hits, Pearson's, and Kendall's correlations.

The results show that the proposed image representation and retrieval framework that integrates external sources and inference for a semantic rule-based query expansion outperforms related systems with 40% increase in the system Recall at approximately 100% retrieval accuracy. This is due to the high usage of the expansion rules which are 88% and 61% for event and entity rules respectively.

Regarding the proposed semantic evaluation for auto-generated annotations, the results show that the proposed annotation evaluation approach outperforms existing automatic annotation evaluation metrics by achieving a Pearson's ρ correlation of 0.73 and Kendall's τ rank correlation of 0.49 for system and caption-level correlations, respectively.

Table of Contents

Acknowledgements	i
Abstract	ii
List of Figures	vii
List of Tables	viii
List of Abbreviations	xi
Publications	xii
1 Introduction	1
1.1 Research Problem	5
1.2 Research Motivation.....	6
1.3 Research Objective	7
1.4 Research Contribution	7
1.5 Thesis Structure	8
2 Related Work	9
2.1 Annotation-based image retrieval approaches.....	10
2.2 Query expansion approaches	14
2.3 Evaluating image annotation systems	16
3 A Framework for Semantic image representation and retrieval	20
3.1 Preliminaries	20
3.1.1 Commonsense-knowledge bases	20
3.1.2 Knowledge parser (K-parser)	21

3.2	Proposed Framework	23
3.3	Image Annotation	25
3.4	Semantic query generation	26
3.4.1	Query expansion	27
3.4.2	Query parsing	30
3.4.3	Query matching.....	33
3.4.4	Result Ranking	36
4	Semantic Evaluation for Auto-generated Image Annotations	40
4.1	Proposed Annotation Evaluation	42
4.2	Graph Generation	43
4.2.1	Natural Language Understanding (NLU)	45
4.2.2	Entity and Predicate Mapping.....	50
4.3	Graph Evaluation.....	52
5	Experiments, Results and Discussion	55
5.1	Datasets.....	54
5.2	Evaluation Metrics.....	55
5.3	Experimenting the Proposed Image Representation and Retrieval Framework.....	57
5.3.1	Evaluation of the Query Expansion.....	57
5.3.2	Evaluation of the Proposed Image Representation and retrieval Framework on larger data set	62
5.3.3	Evaluation of the Quality of Expansion Rules and Relation Similarity Metric	70

5.4	Experimenting the Proposed Semantic Evaluation for Auto-generated Image Annotations :.....	71
5.4.1	Evaluating the correlation of the proposed semantic evaluation with human judgment ...	72
5.4.2	Evaluation of the quality of the proposed semantic evaluation for auto-generated Annotations.....	73
6	Conclusion and Future Work	79
	References	84

List of Figures

3.1 Example for concepts and their relations ConceptNet.....	20
3.2 The output of K-parser for "A dog swim in water" sentence.....	22
3.3 Architecture of the proposed framework.....	24
3.4 RDF generated triplet annotations from the K-parser output.....	26
3.5 SPARQL query translation.....	31
3.6 An example of query generator run and the generated triplet outputs with their knowledge types.....	35
4.1 Proposed Semantic Graph Representation and Evaluation.....	43
4.2 Graph Generator Module Architecture.....	45
4.3 An Example of Stanford Parse Tree.....	46
4.4 An Example of NLU Run.....	51
4.5 An Example of DSG Generation.....	52
5.1 Comparison of retrieval results for background queries in terms of (a) No. of Retrievals (b) Accuracy.....	66
5.2 Comparison of retrieval results on different query categories in terms of (a) Average No. of Retrievals (b) Average Accuracy.....	69

List of Tables

3.1	Examples of Expanded Terms for Different Knowledge Categories.	29
3.2	Parsing Rules.....	32
3.3	Semantic Expansion Rules.....	33
3.4	Semantic Relation Similarities for Two SRGs.....	39
4.1	Information Extraction Rules.....	47
4.2	Examples of Extracted Relations for Different Natural Language Sentence.....	50
5.1	Categories of Natural Language Queries.....	60
5.2	Natural Language Queries Set.....	60
5.3	Comparison of the Retrieval Results -Proposed Work versus Related Work.....	62
5.4	Query Samples Generated for Different Query Categories in the Second Experiment.....	64
5.5	Comparison of the Proposed Framework with Reference Approach for Different Query Categories.....	65
5.6	Semantic Expansion Rules Effect on the Different Query Categories in the Second Experiment.....	67
5.7	Comparison of the Accuracy at k - proposed Versus Related Works.....	71

5.8	Correlation between Automatic Annotation Evaluation Metrics and Human Judgment Scores.....	73
5.9	Comparison of the Retrieval Accuracies-Proposed Versus Related Work for Different Similarity Measures.....	74