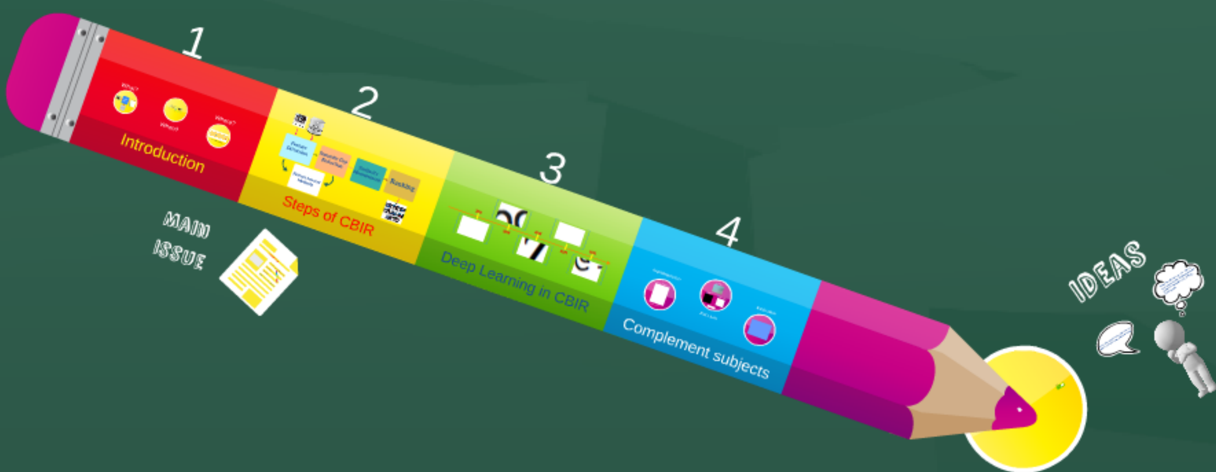
Content Based Image Retrieval

Study of Deep Learning Approach

Seminar Report

November 2015



By: Mina Ameli

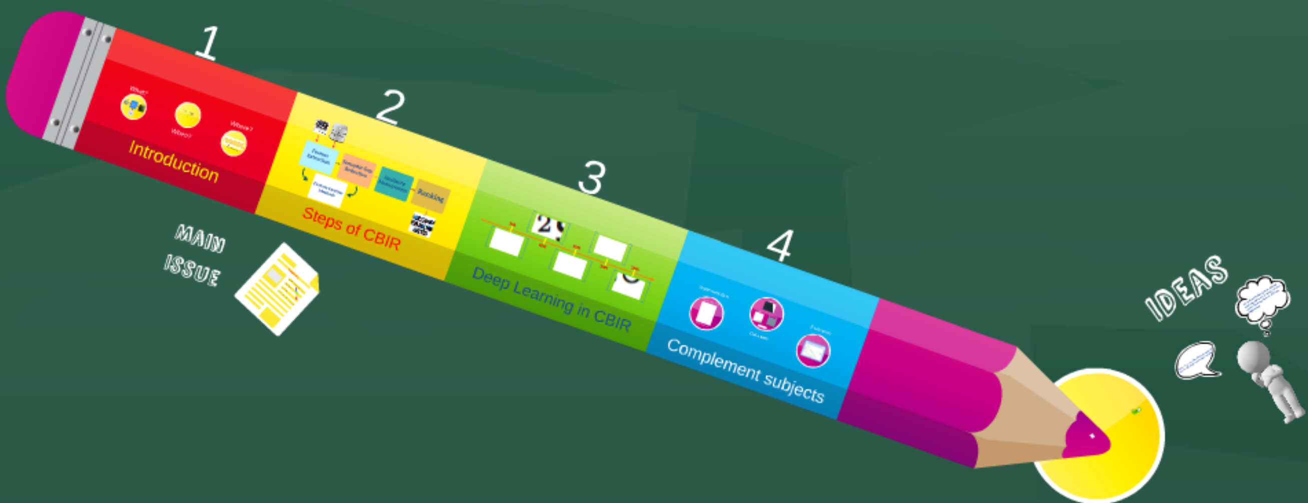
Under supervision:
Dr. Shanbehzadeh

Content Based Image Retrieval

Study of Deep Learning Approach

Seminar Report

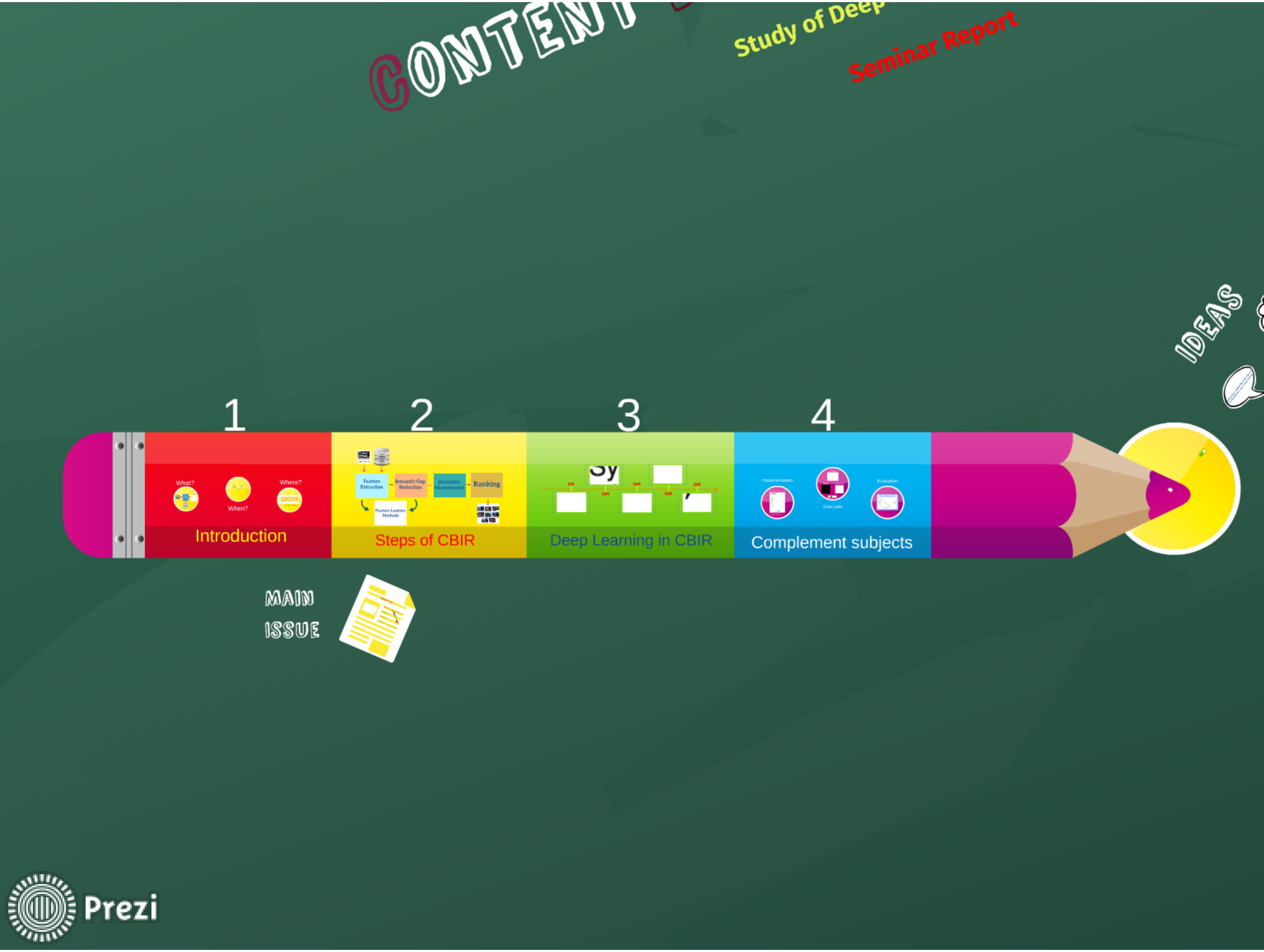
November 2015



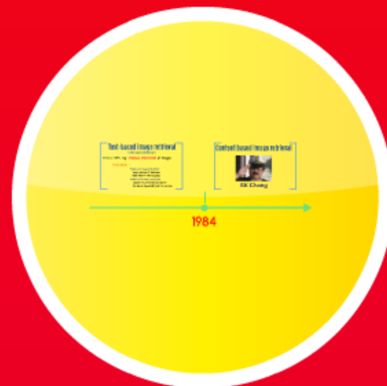
By: Mina Ameli

Under supervision:
Dr. Shanbehzadeh

Prezi

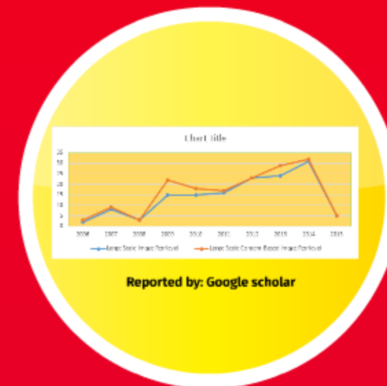

CONTENT
Study of Deep
Seminar Report
IDEAS
1
2
3
4
What?
When?
Where?
Introduction
Feature Extraction
Semantic Gap Reduction
Similarity Measurement
Ranking
Feature Learner Methods
Steps of CBIR
Deep Learning in CBIR
Complement subjects
MAIN
ISSUE
Prezi

What?

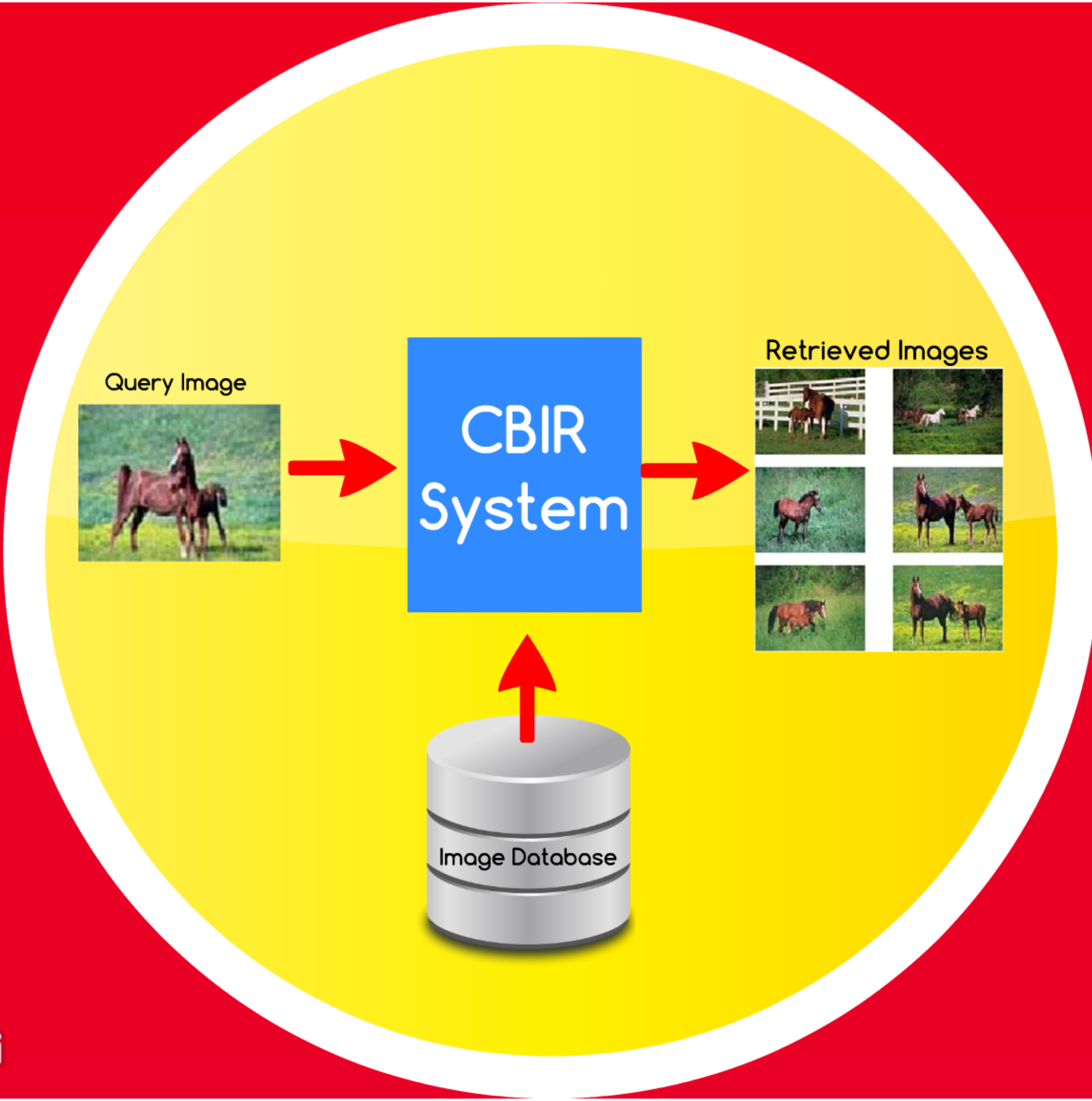


When?

Where?



Introduction



Prezi

Text-based image retrieval

(Context based/annotation based)

Before 1984, by: Manual annotation of images.

Drawbacks :

Problem of image annotation

- Large volumes of databases
- Valid only for one language

Problem of human perception

- Subjectivity of human perception
- Too much responsibility on the end-user

Content based image retrieval

SK Chang

1984

Text-based image retrieval

(Context based/annotation based)

Before 1984, by: Manual annotation of images.

Drawbacks :

Problem of image annotation

- Large volumes of databases
- Valid only for one language

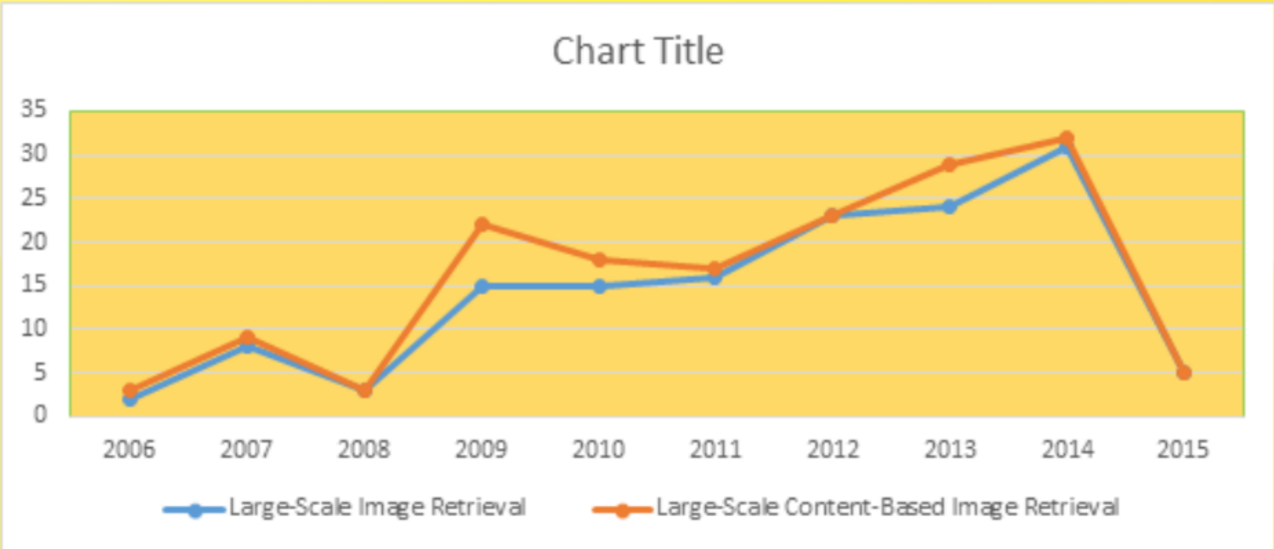
Problem of human perception

- Subjectivity of human perception
- Too much responsibility on the end-user

Content based image retrieval

SK Chang

Prezi

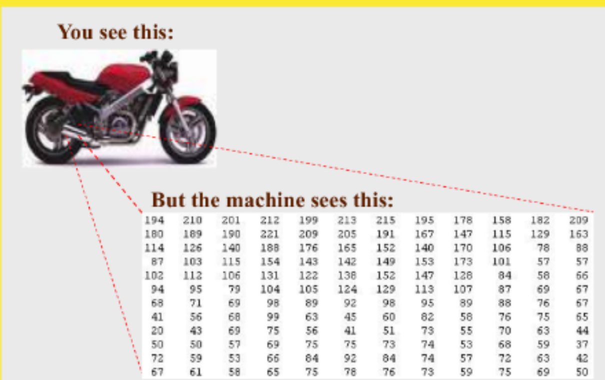


Reported by: Google scholar

MAIN ISSUE


Semantic Gap
Gap between low level features and high-level concepts
Human in the loop
Interactive systems
Retrieval accuracy
and
Retrieval speed

Semantic Gap

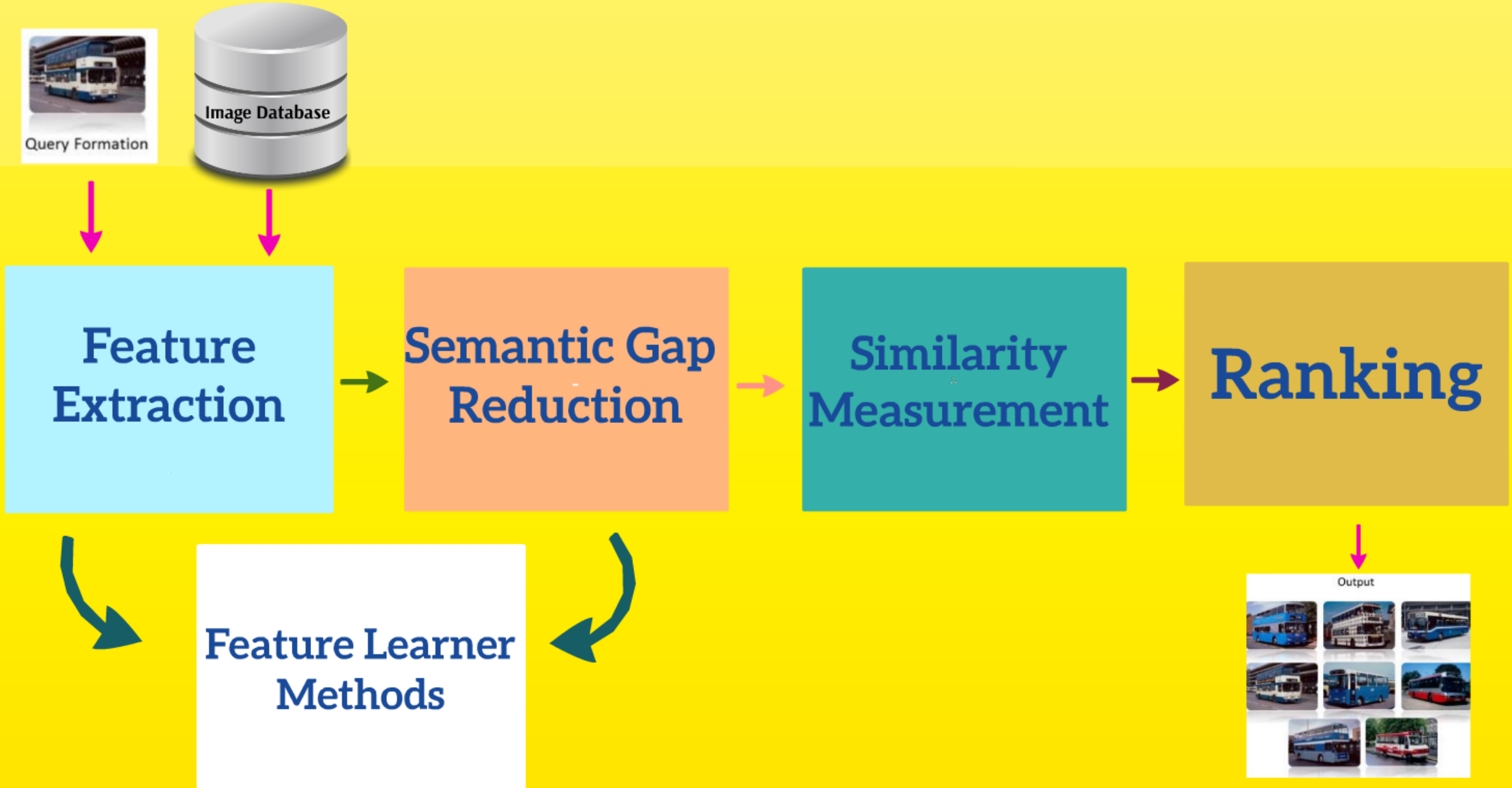


Gap between low level features and high-level concepts

Human in the loop

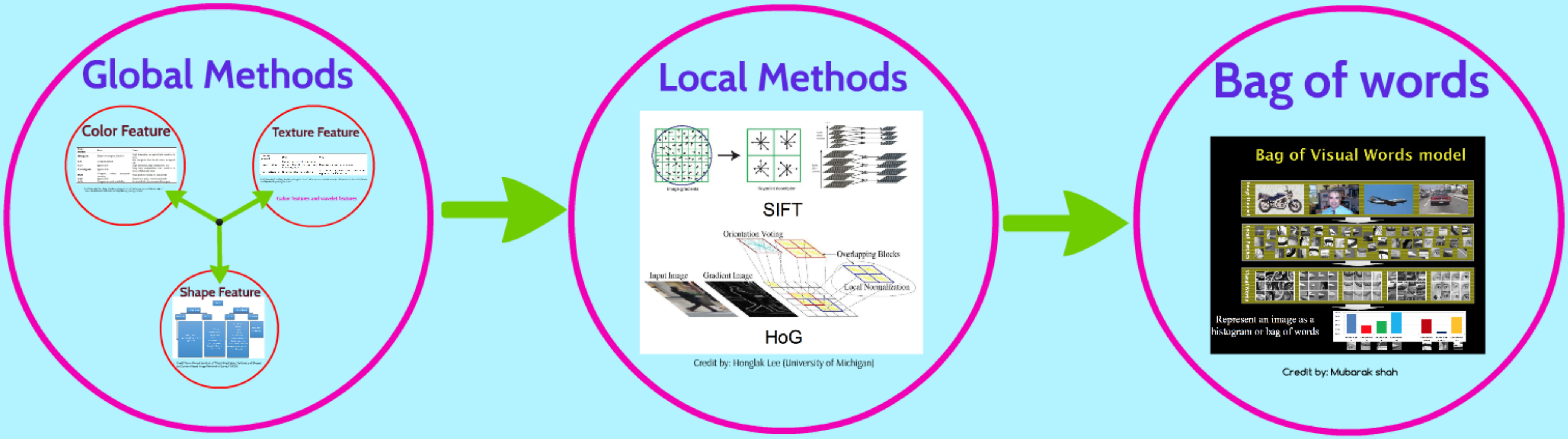
Interactive systems

Retrieval accuracy and Retrieval speed



Steps of CBIR

Feature Extraction


Global Methods
Color Feature
Texture Feature
Shape Feature
Local Methods
SIFT
Orientation Voting
Overlapping Blocks
Input Image
Gradient Image
Local Normalization
HoG
Credit by: Honglak Lee (University of Michigan)
Bag of words
Bag of Visual Words model
Represent an image as a histogram or bag of words
Credit by: Mubarak shah

Global Methods

Color Feature

Color method	Pros.	Cons.
Histogram	Simple to compute, intuitive	High dimension, no spatial info, sensitive to noise
CM	Compact, robust	Not enough to describe all colors, no spatial info
CCV	Spatial info	High dimension, high computation cost
Correlogram	Spatial info	Very high computation cost, sensitive to noise, rotation and scale
DCD	Compact, robust, perceptual meaning	Need post-processing for spatial info
CSD	Spatial info	Sensitive to noise, rotation and scale
SCD	Compact on need, scalability	No spatial info, less accurate if compact

Credit from: Jing, Tian, Dong. "A review on image feature extraction and representation techniques." International Journal of Multimedia and Ubiquitous Engineering 8.4 (2013)

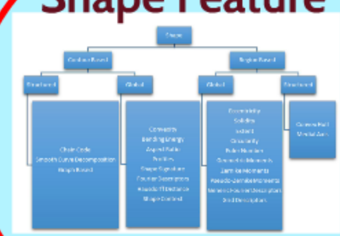
Texture Feature

Texture method	Pros.	Cons.
Spatial texture	Meaningful, easy to understand, can be extracted from any shape without losing info.	Sensitive to noise and distortions
Spectral texture	Robust, need less computation	No semantic meaning, need square image regions with sufficient size

Credit from: Jing, Tian, Dong. "A review on image feature extraction and representation techniques." International Journal of Multimedia and Ubiquitous Engineering 8.4 (2013)

Gabor features and wavelet features

Shape Feature



Credit from: Ahmad, Jamil, et al. "Describing Colors, Textures and Shapes for Content Based Image Retrieval-A Survey." (2015).

Color Feature

Color method	Pros.	Cons.
Histogram	Simple to compute, intuitive	High dimension, no spatial info, sensitive to noise
CM	Compact, robust	Not enough to describe all colors, no spatial info
CCV	Spatial info	High dimension, high computation cost
Correlogram	Spatial info	Very high computation cost, sensitive to noise, rotation and scale
DCD	Compact, robust, perceptual meaning	Need post-processing for spatial info
CSD	Spatial info	Sensitive to noise, rotation and scale
SCD	Compact on need, scalability	No spatial info, less accurate if compact

Credit from: ping Tian, Dong. "A review on image feature extraction and representation techniques." International Journal of Multimedia and Ubiquitous Engineering 8.4 (2013)

Texture Feature

Texture method	Pros.	Cons.
Spatial texture	Meaningful, easy to understand, can be extracted from any shape without losing info.	Sensitive to noise and distortions
Spectral texture	Robust, need less computation	No semantic meaning, need square image regions with sufficient size

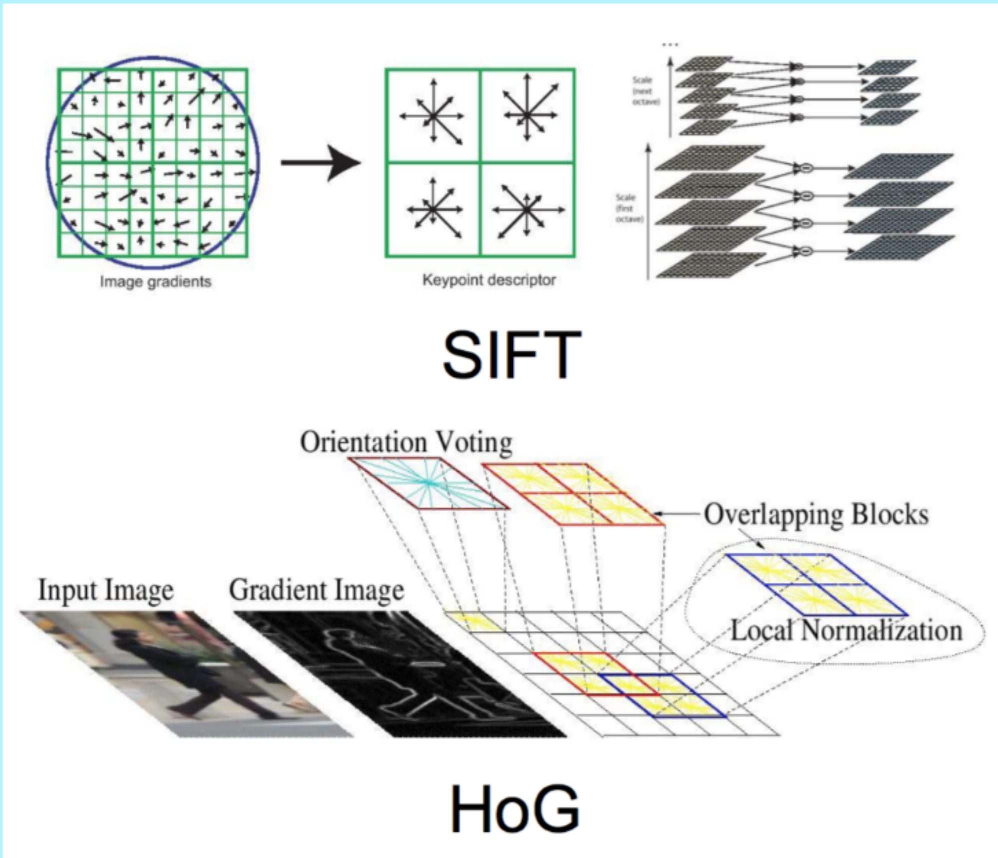
Credit from: ping Tian, Dong. "A review on image feature extraction and representation techniques." International Journal of Multimedia and Ubiquitous Engineering 8.4 (2013)

Gabor features and wavelet features

Shape Feature

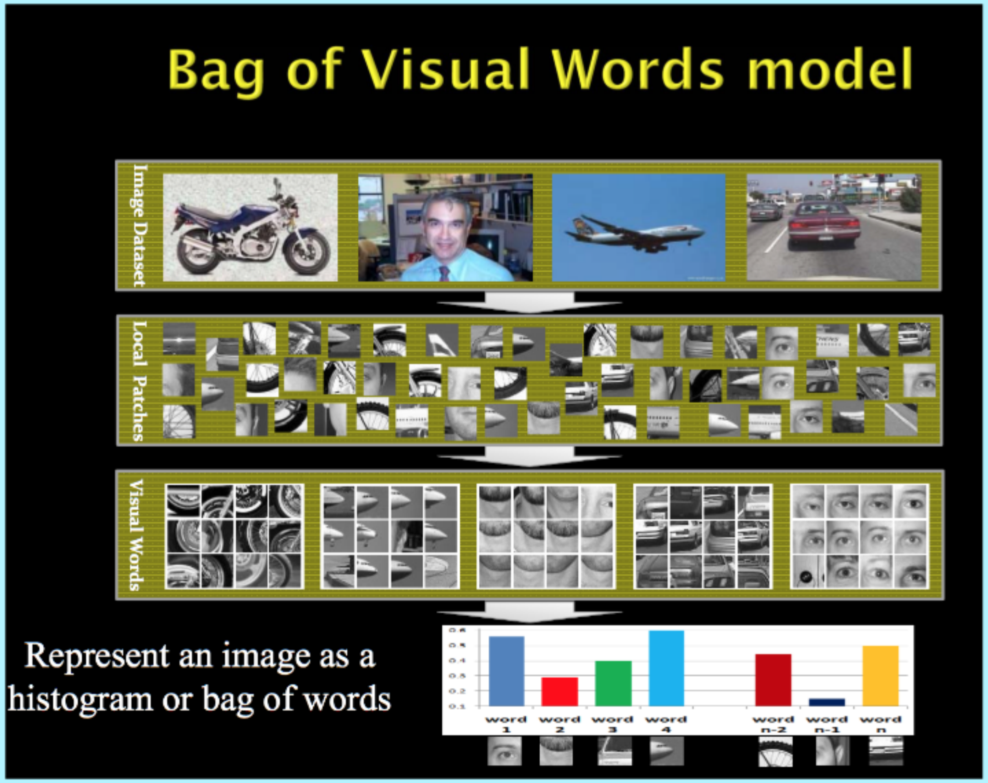
Credit from: Ahmad, Jamil, et al. "Describing Colors, Textures and Shapes for Content Based Image Retrieval-A Survey." (2015).

Local Methods



Credit by: Honglak Lee (University of Michigan)

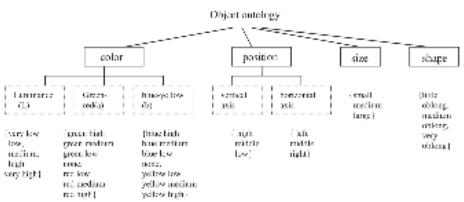
Bag of words



Credit by: Mubarak shah

Semantic Gap Reduction

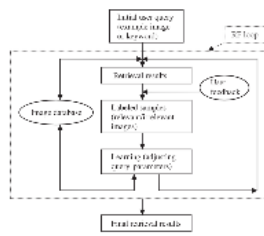
Object ontology



V. Mezaris, I. Kompatsiaris, M.G. Strintzis, An ontology approach to object-based image retrieval, Proceedings of the ICIP, vol. II, 2003.

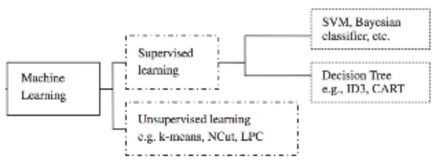
Relevance feedback

during mid 1990

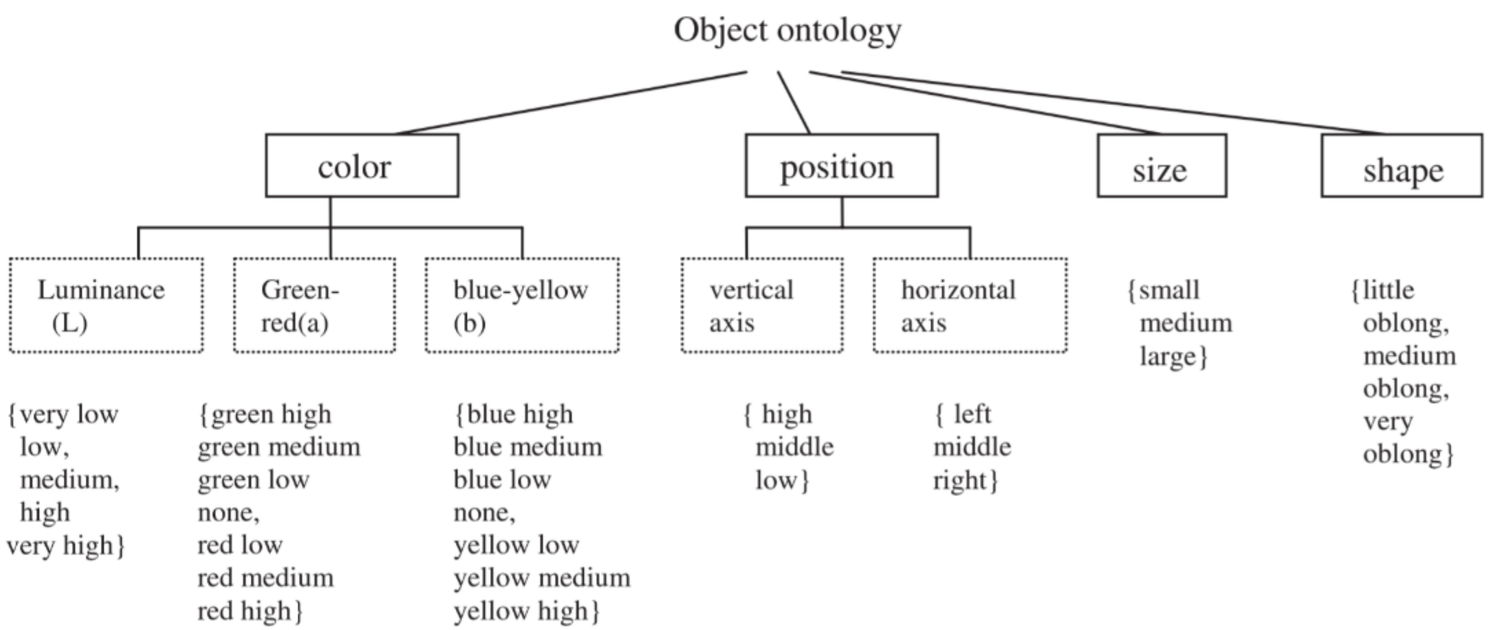


credit by: Liu, Ying, et al. "A survey of content-based image retrieval with high-level semantics."

Machine learning methods



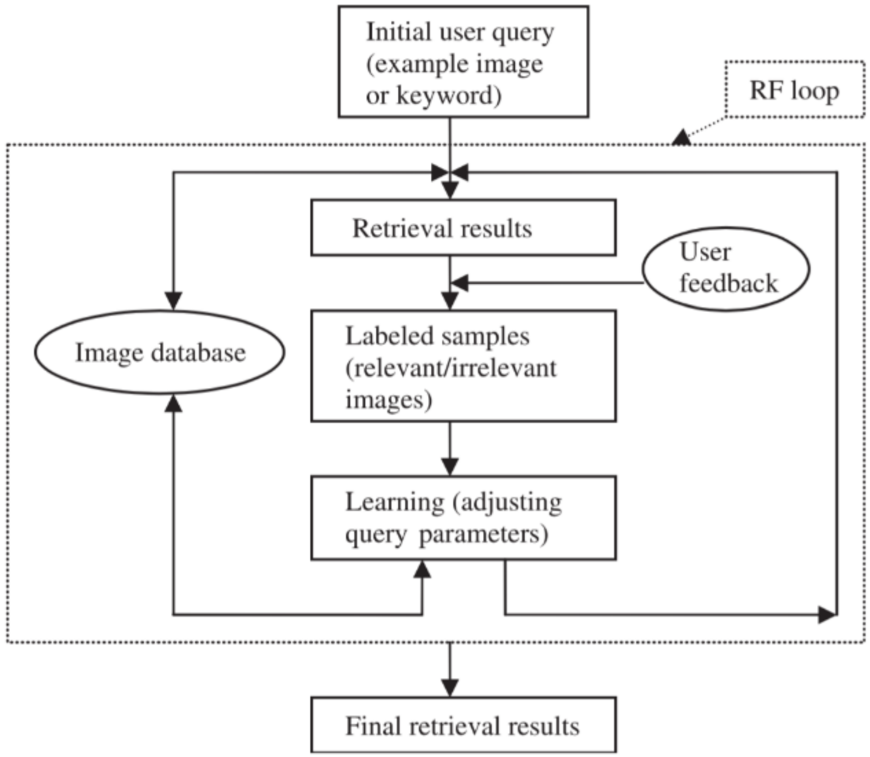
Object ontology



V. Mezaris, I. Kompatsiaris, M.G. Strintzis, An ontology approach to object-based image retrieval, Proceedings of the ICIP, vol. II, 2003.

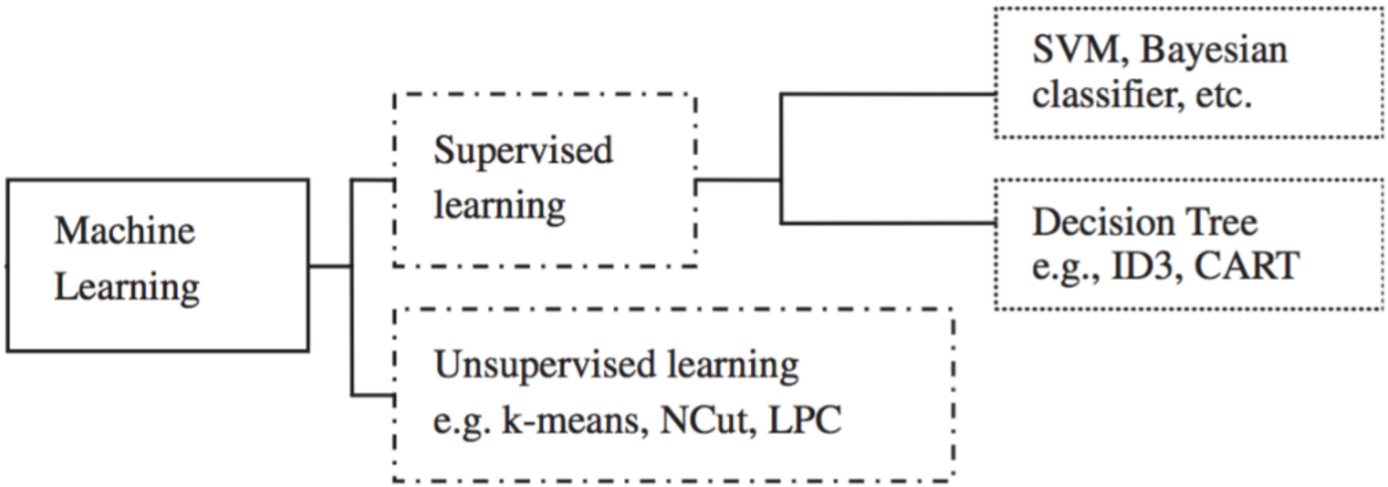
Relevance feedback

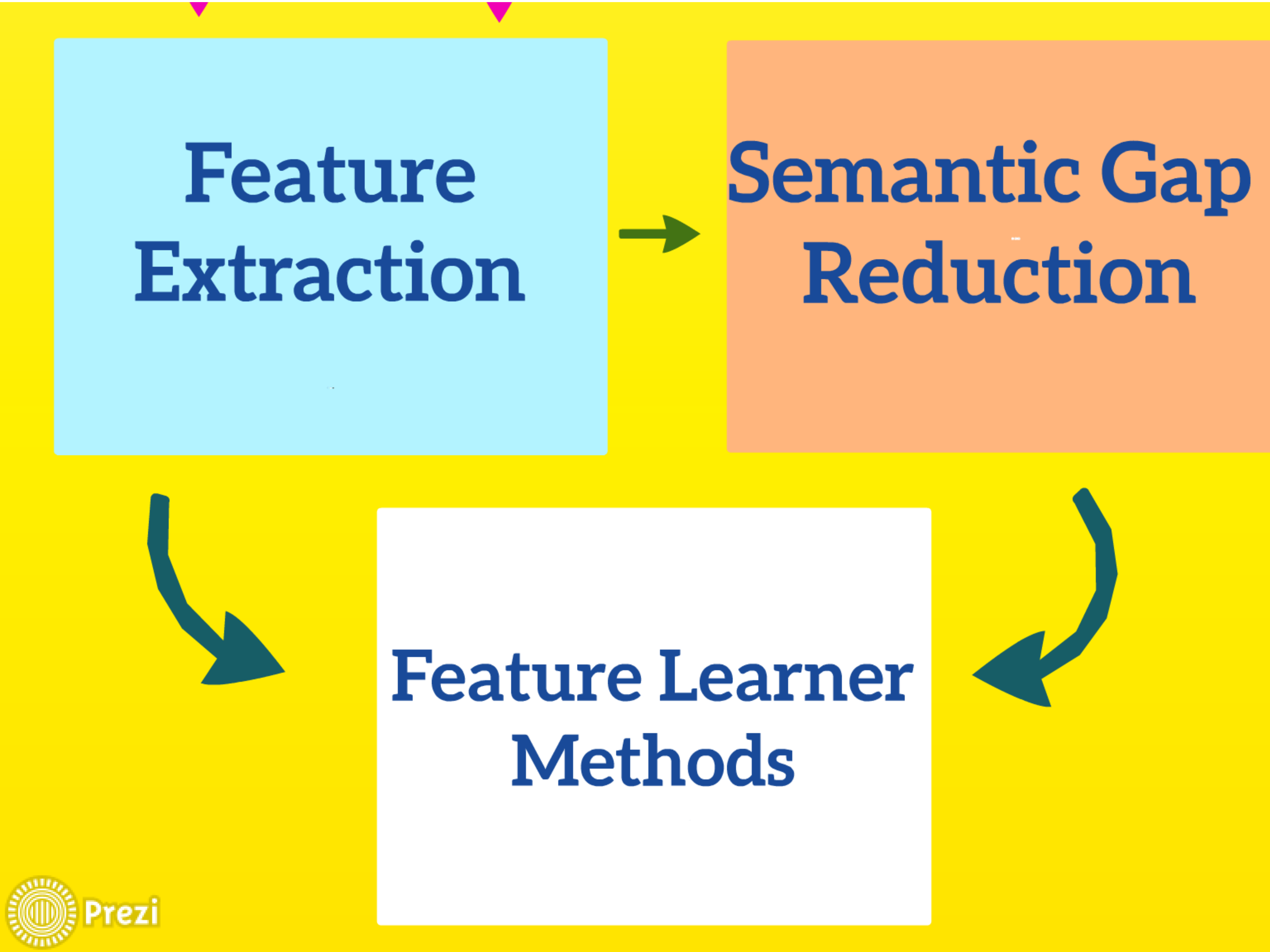
during mid 1990



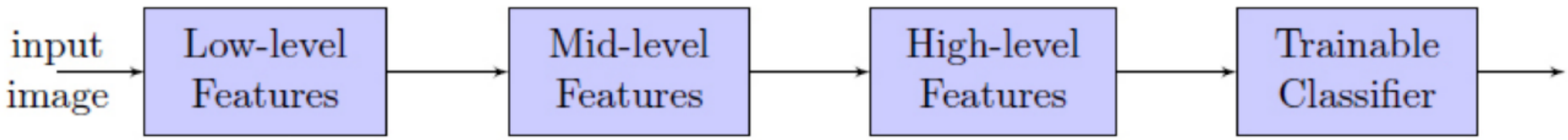
credit by: Liu, Ying, et al. "A survey of content-based image retrieval with high-level semantics."

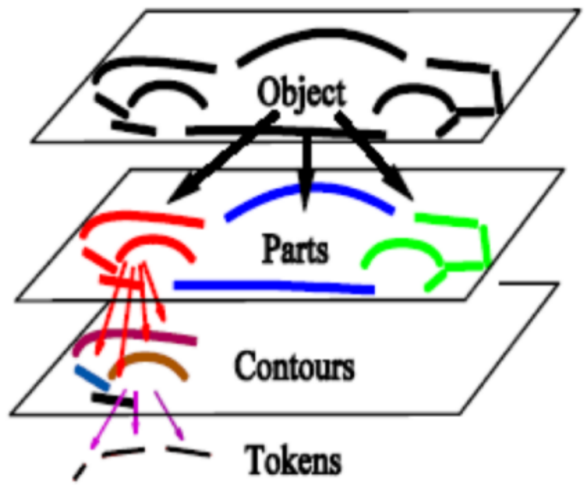
Machine learning methods




Feature Extraction
Semantic Gap Reduction
Feature Learner Methods
Prezi

Hierarchie





Slides: Dr.Qassabi

Deep Learning

Reasons of popularity from 2012:

- The availability of large number of labeled data
- The prevalent use of high end GPU
- Lower cost of computing hardware
- Advances in machine learning and signal/ information processing research

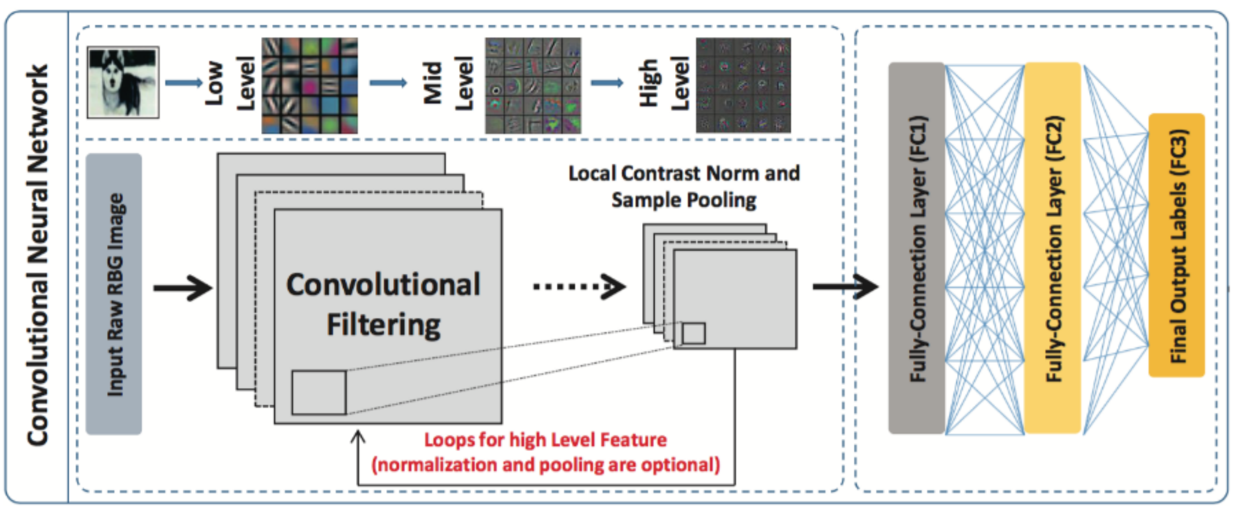
deep versus shallow:

Sharing parameters is good

- computational complexity

Efficient representation:

- no redundancy



Similarity Measurement

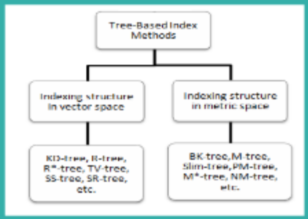
Prezi

Distance

On low level features
Minkowski-type metric, like:

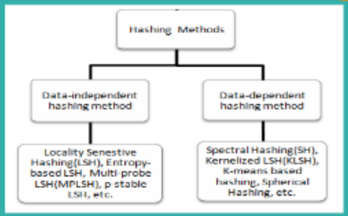
- Euclidean distance
- Cosine similarity

Indexing



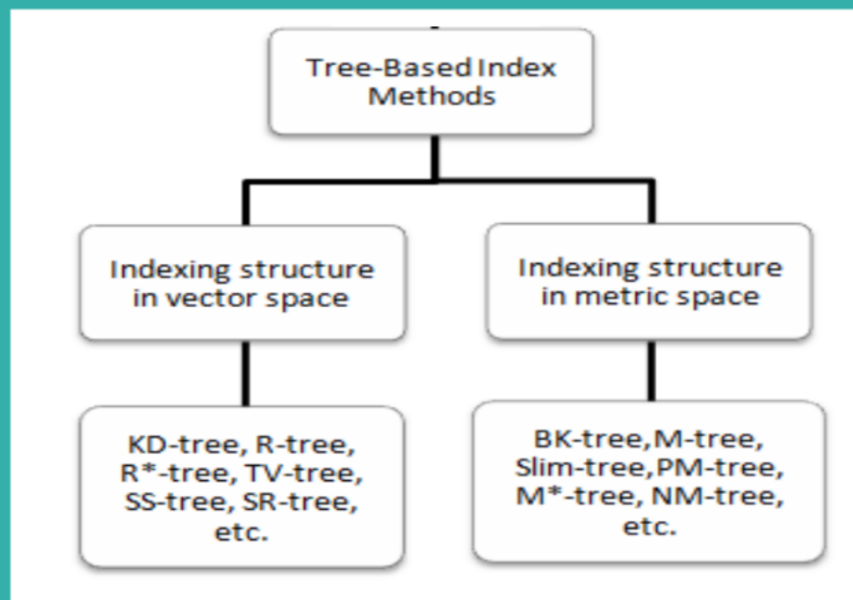
Credit from: Tikle, Ashwini N., Chandu Vaidya, and Prashant Dahiwale. "A Survey of Indexing Techniques For Large Scale Content-Based Image Retrieval." (EESCO), 2015 International Conference on. IEEE. (2015)

Hashing



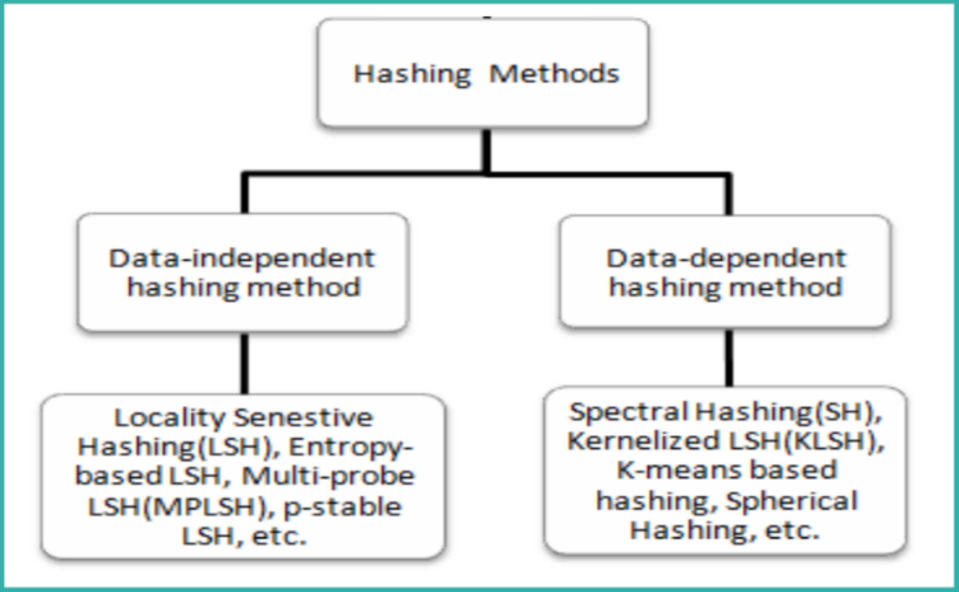
Credit from: Tikle, Ashwini N., Chandu Vaidya, and Prashant Dahiwale. "A Survey of Indexing Techniques For Large Scale Content-Based Image Retrieval." (EESCO), 2015 International Conference on. IEEE. (2015)

Indexing



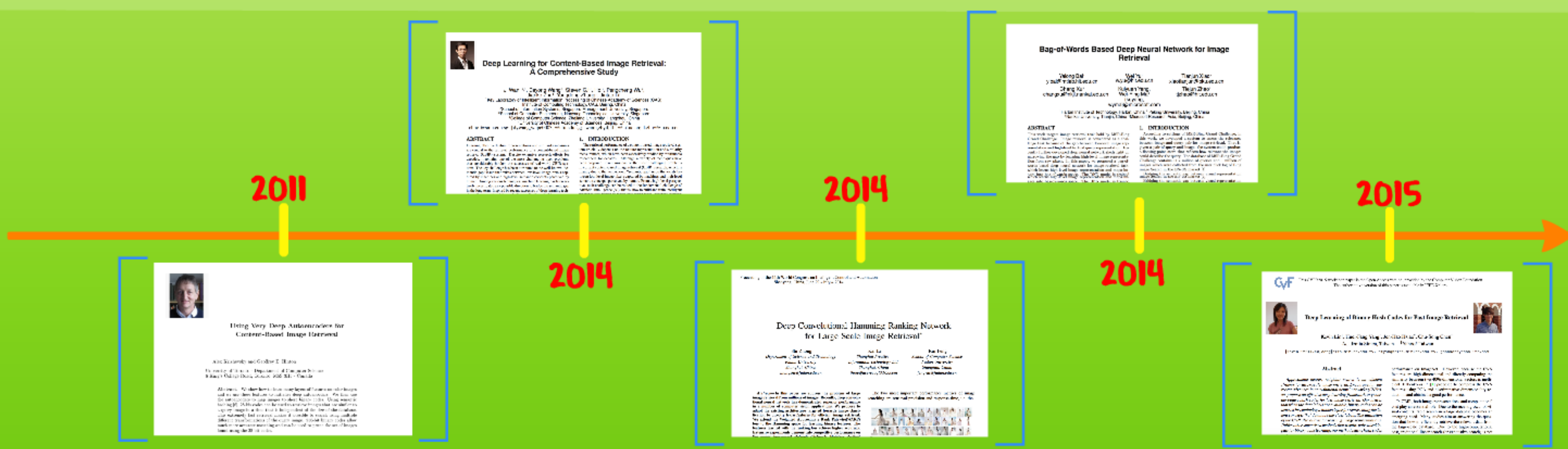
Credit from: Tikle, Ashwini N., Chandu Vaidya, and Prashant Dahiwale. "A Survey of Indexing Techniques For Large Scale Content-Based Image Retrieval." (EESCO), 2015 International Conference on. IEEE, (2015)

Hashing



Credit from: Tikle, Ashwini N., Chandu Vaidya, and Prashant Dahiwale. "A Survey of Indexing Techniques For Large Scale Content-Based Image Retrieval." (EESCO), 2015 International Conference on. IEEE, (2015)

Ranking



Deep Learning in CBIR

Using Very Deep Autoencoders for Content-Based Image Retrieval

Alex Krizhevsky and Geoffrey E. Hinton

University of Toronto - Department of Computer Science
6 King's College Road, Toronto, M5S 3H5 - Canada

Abstract. We show how to learn many layers of features on color images and we use these features to initialize deep autoencoders. We then use the autoencoders to map images to short binary codes. Using semantic hashing [6], 28-bit codes can be used to retrieve images that are similar to a query image in a time that is independent of the size of the database. This extremely fast retrieval makes it possible to search using multiple different transformations of the query image. 256-bit binary codes allow much more accurate matching and can be used to prune the set of images found using the 28-bit codes.

Prezi

Deep Learning for Content-Based Image Retrieval: A Comprehensive Study

Ji Wan[1,2,5], Dayong Wang[3], Steven C.H. Hoi[2], Pengcheng Wu[3], Jianke Zhu[4], Yongdong Zhang[1], Jintao Li[1]
[1]Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing, China
[2]School of Information Systems, Singapore Management University, Singapore
[3]School of Computer Engineering, Nanyang Technological University, Singapore
[4]College of Computer Science, Zhejiang University, Hangzhou, China
[5]University of Chinese Academy of Sciences, Beijing, China
chhoi@smu.edu.sg; {dywang,wupe0003}@ntu.edu.sg; {wanji,zhyd,jtli}@ict.ac.cn; jkzhu@zju.edu.cn

ABSTRACT

Learning effective feature representations and similarity measures are crucial to the retrieval performance of a content-based image retrieval (CBIR) system. Despite extensive research efforts for decades, it remains one of the most challenging open problems that considerably hinders the successes of real-world CBIR systems. The key challenge has been attributed to the well-known "semantic gap" issue that exists between low-level image pixels captured by machines and high-level semantic concepts perceived by human. Among various techniques, machine learning has been actively investigated as a possible direction to bridge the semantic gap in the long term. Inspired by recent successes of deep learning tech-

1. INTRODUCTION

The retrieval performance of a content-based image retrieval system crucially depends on the feature representation and similarity measurement, which have been extensively studied by multimedia researchers for decades. Although a variety of techniques have been proposed, it remains one of the most challenging problems in current content-based image retrieval (CBIR) research, which is mainly due to the well-known "semantic gap" issue that exists between low-level image pixels captured by machines and high-level semantic concepts perceived by human. From a high-level perspective, such challenge can be rooted to the fundamental challenge of Artificial Intelligence (AI), that is, how to build and train intelligent

Proceeding of the 11th World Congress on Intelligent Control and Automation
Shenyang, China, June 29 - July 4 2014

Deep Convolutional Hamming Ranking Network for Large Scale Image Retrieval*

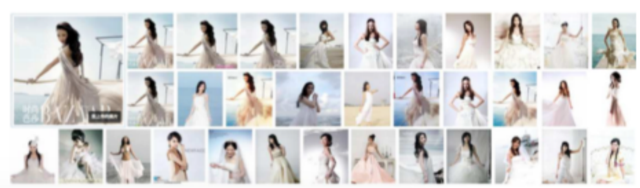
Shi Zhong
Department of Science and Technology
Fudan University
Shanghai, China
zhongshi@fudan.edu.cn

Kai Li
Shanghai Freative
Information Technology Ltd.
Shanghai, China
threedfacerecog@163.com

Rui Feng
School of Computer Science
Fudan University
Shanghai, China
fengrui@fudan.edu.cn

***Abstract*—In this paper we address the problem of large image retrieval from millions of images. Recently, deep convolutional neural network has demonstrated superior performance in a number of computer vision applications. We propose to adapt the existing architecture targeted towards image classification to directly learn features for efficient image retrieval. We extend the Weighted Approximate Rank Pairwise(WARP) loss to the Hamming space for learning binary features. The features learned with the ranking loss achieve higher accuracy. Extensive experiments demonstrate competitive performance on five public benchmark datasets UKbench, Holidays, Oxford**

The two most important performance metrics of image searching are retrieval precision and response time, i.e. the



Bag-of-Words Based Deep Neural Network for Image Retrieval

Yalong Bai[1]
ylbai@mtlab.hit.edu.cn

Wei Yu[1]
w.yu@hit.edu.cn

Tianjun Xiao[2]
xiaotianjun@pku.edu.cn

Chang Xu[3]
changxu@nkjl.nankai.edu.cn

Kuiyuan Yang,
Wei-Ying Ma[4]
{kuyang,
wyma}@microsoft.com

Tiejun Zhao[1]
tjzhao@hit.edu.cn

[1]Harbin Institute of Technology, Harbin, China [2] Peking University, Beijing, China
[3]Nankai University, Tianjin, China [4]Microsoft Research Asia, Beijing, China

ABSTRACT

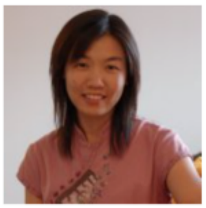
This work targets image retrieval task hold by MSR-Bing Grand Challenge. Image retrieval is considered as a challenge task because of the gap between low-level image representation and high-level textual query representation. Recently further developed deep neural network sheds light on narrowing the gap by learning high-level image representation from raw pixels. In this paper, we proposed a bag-of-words based deep neural network for image retrieval task, which learns high-level image representation and maps images into bag-of-words space. The DNN model is trained
which learns high-level image representation and maps images into bag-of-words space. The DNN model is trained

1. INTRODUCTION

According to settings of MSR-Bing Grand Challenge, in this work, we developed a system to assess the relevance between image and query pair for image retrieval. That is, given a pair of query and image, the system could produce a floating-point score that reflects how relevant the images could describe the query. The database of MSR-Bing Grand Challenge contains 11.7 million of queries and 1 million of images which were collected from the user click log of Bing image Search in the EN-US market [4].

Bridging the semantic gap between visual representation

image Search in the EN-US market [4].
Bridging the semantic gap between visual representation



Deep Learning of Binary Hash Codes for Fast Image Retrieval

Kevin Lin†, Huei-Fang Yang†, Jen-Hao Hsiao‡, Chu-Song Chen†
†Academia Sinica, Taiwan ‡Yahoo! Taiwan
{kevinlin311.tw,song}@iis.sinica.edu.tw, hfyang@citi.sinica.edu.tw, jenhaoh@yahoo-inc.com

Abstract

Approximate nearest neighbor search is an efficient strategy for large-scale image retrieval. Encouraged by the recent advances in convolutional neural networks (CNNs), we propose an effective deep learning framework to generate binary hash codes for fast image retrieval. Our idea is that when the data labels are available, binary codes can be learned by employing a hidden layer for representing the latent concepts that dominate the class labels. The utilization of the CNN also allows for learning image representations. Unlike other supervised methods that require pair-wised inputs for binary code learning, our method learns hash codes

performance on ImageNet. However, because the CNN features are high-dimensional and directly computing the similarity between two 4096-dimensional vectors is inefficient, Babenko *et al.* [1] proposed to compress the CNN features using PCA and discriminative dimensionality reduction, and obtained a good performance.

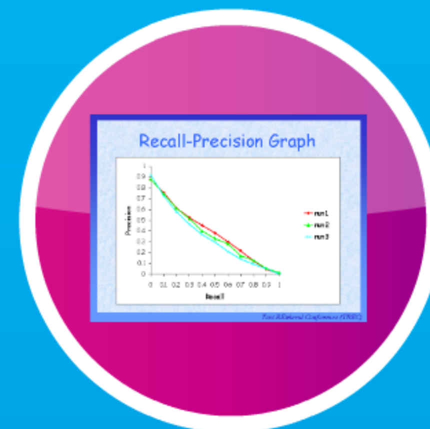
In CBIR, both image representations and computational cost play an essential role. Due to the recent growth of visual contents, rapid search in a large database becomes an emerging need. Many studies aim at answering the question that how to efficiently retrieve the relevant data from the large-scale database. Due to the high-computational cost, traditional linear search (or exhaustive search) is not

Prezi

Implementation

Data sets

Evaluation



Complement subjects

100 Best GitHub: Deep Learning

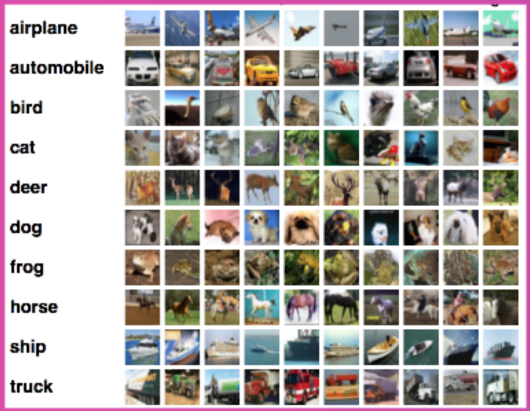
Languages

Python	186
Matlab	78
C++	33
Cuda	24
Java	22
C	18
TeX	15
Lua	15
JavaScript	15
Scala	9

Deep Learning

Meta-Guide.com

CFAR10



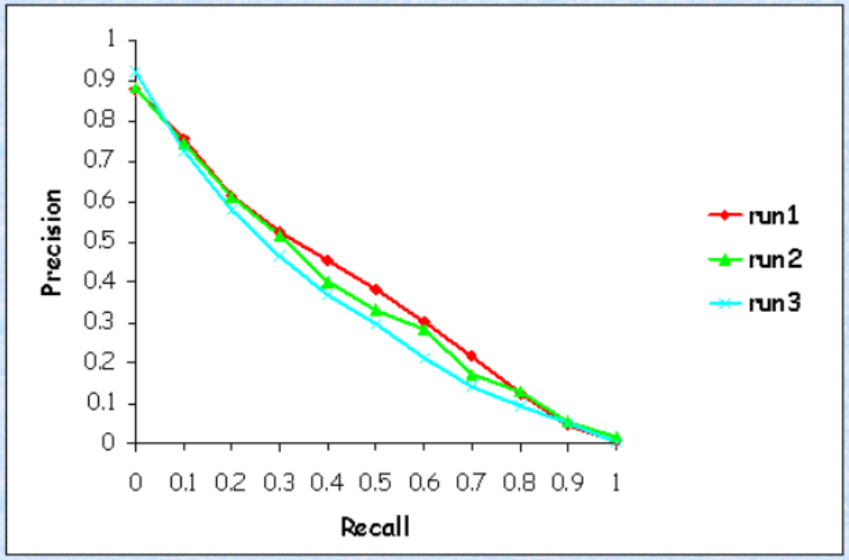
MNIST

Caltech256

IMAGENET

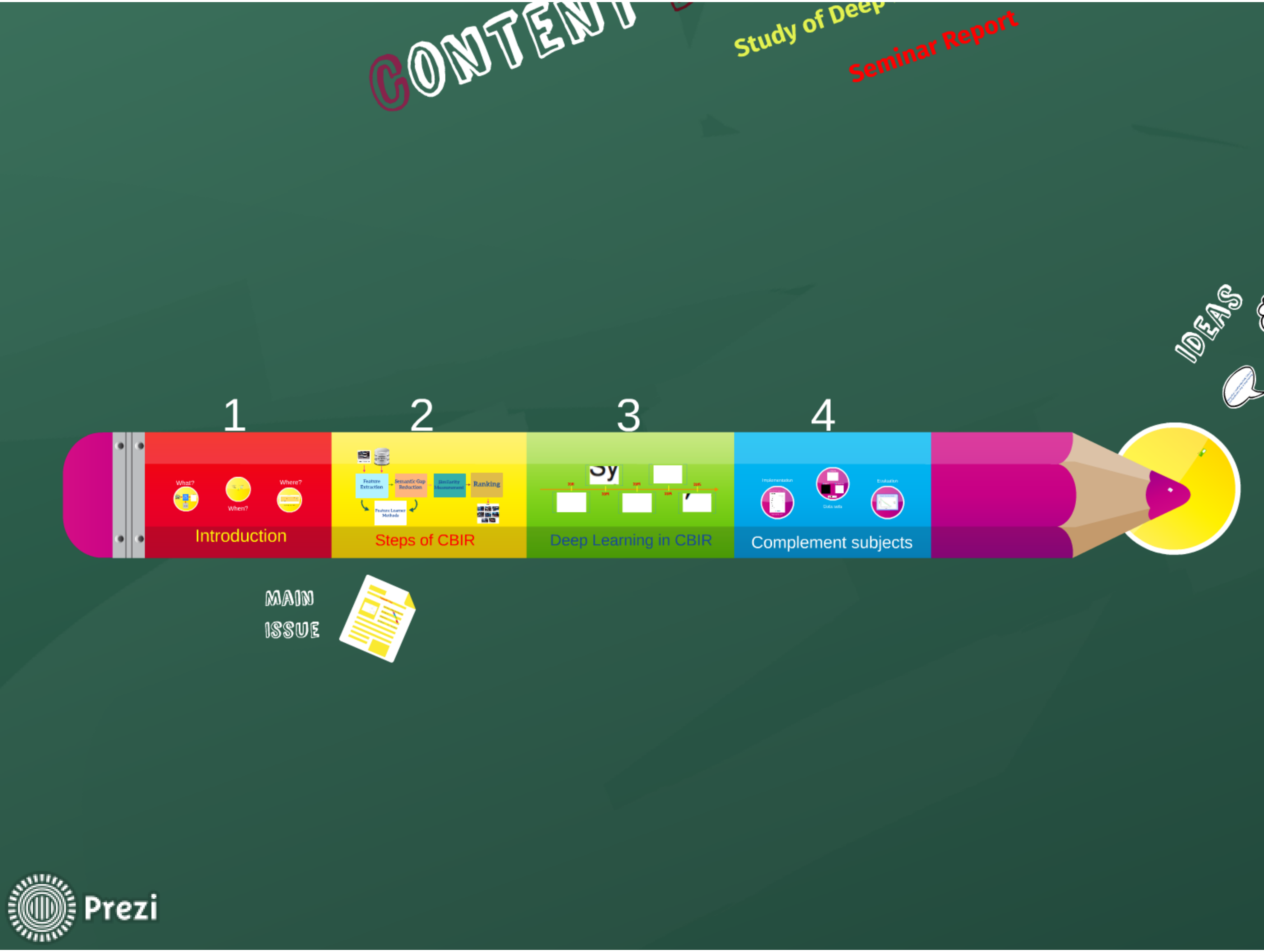
Prezi

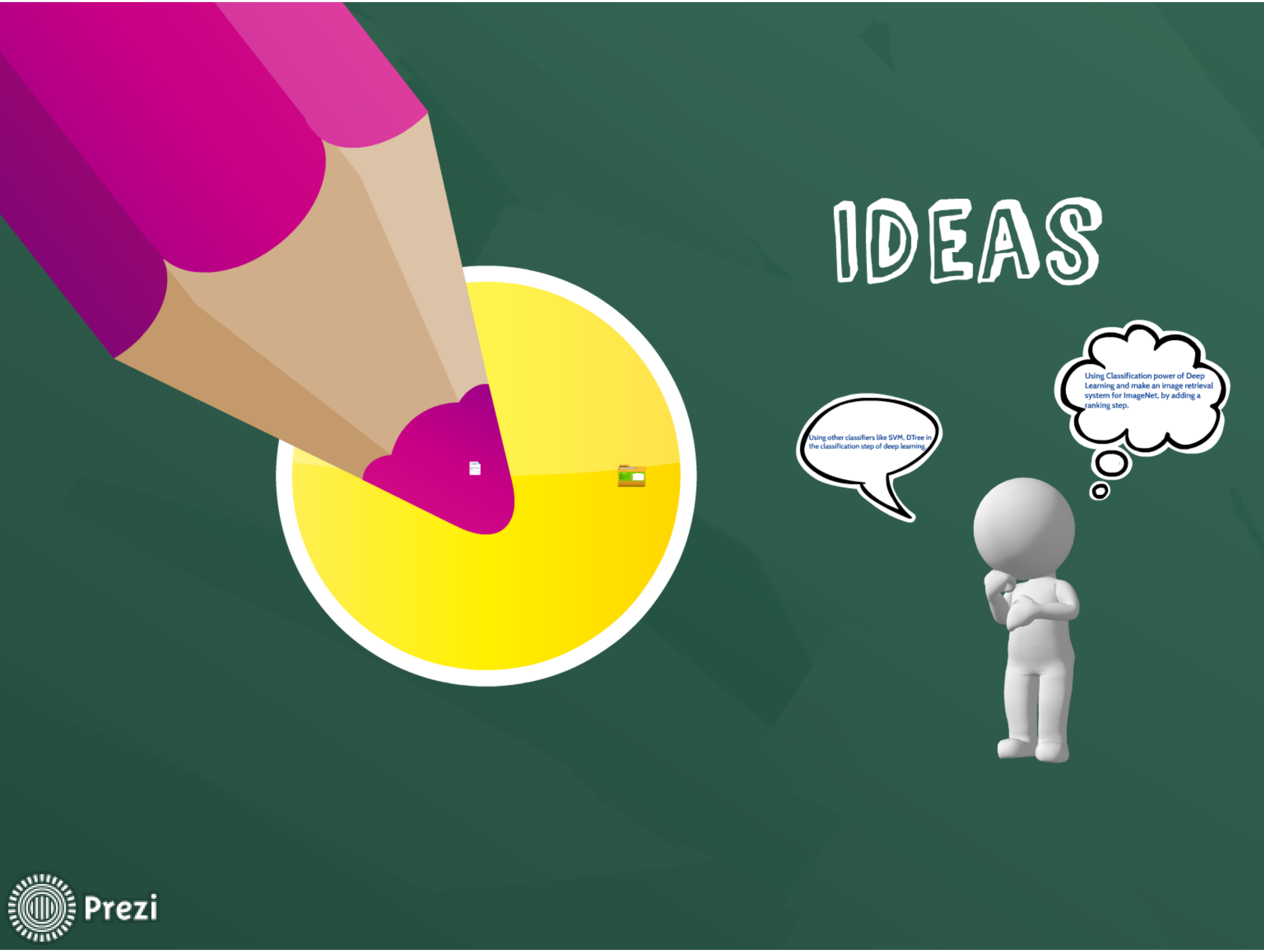
Recall-Precision Graph



Text REtrieval Conference (TREC)

Prezi


CONTENT
Study of Deep
Seminar Report
IDEAS
1
2
3
4
What?
When?
Where?
Introduction
Feature Extraction
Semantic Gap Reduction
Ranking
Feature Learner Methods
Steps of CBIR
Deep Learning in CBIR
Complement subjects
MAIN
ISSUE
Prezi


IDEAS
Using other classifiers like SVM, DTree in the classification step of deep learning.
Using Classification power of Deep Learning and make an image retrieval system for ImageNet, by adding a ranking step.
Prezi

Using Classification power of Deep Learning and make an image retrieval system for ImageNet, by adding a ranking step.

Using other classifiers like SVM, DTree in the classification step of deep learning.

Conclusion

We talk about history, definition, Process of CBIR and introduce methods of each steps.

Review Deep Learning methods and it's application in CBIR.

It's the start of Deep learning in CBIR but ...

Rerfrences:

Papers:

[1] Tikle, Ashwini N., Chandu Vaidya, and Prashant Dahiwale. "A Survey of Indexing Techniques For Large Scale Content-Based Image Retrieval." (EESCO), 2015 International Conference on. IEEE, (2015)

[2] Wang, Jingdong, et al. "Hashing for similarity search: A survey." arXiv preprint arXiv:1408.2927 (2014)

[3] Wan, Ji, et al. "Deep learning for content-based image retrieval: A comprehensive study." Proceedings of the ACM International Conference on Multimedia. ACM, (2014)

[4] Deng, Li. "A tutorial survey of architectures, algorithms, and applications for deep learning." APSIPA Transactions on Signal and Information Processing 3 (2014)

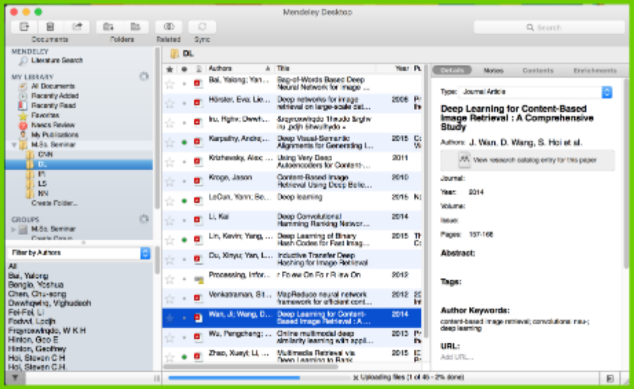
[5] CVPR (2014) workshop tutorials on deep learning

[6] ping Tian, Dong. "A review on image feature extraction and representation techniques." International Journal of Multimedia and Ubiquitous Engineering 8.4 (2013)

[7] Bengio, Yoshua. "Learning deep architectures for AI." Foundations and trends® in Machine Learning 2.1 (2009)

[8] Liu, Ying, et al. "A survey of content-based image retrieval with high-level semantics." Pattern Recognition 40.1 (2007)

Mendeley:



THANKS

Special thanks to Dr.Shanbehzadeh,
because of his wonderful guidance.

Question Or Comment ???