Tutorial on searching text and images in the medical domain

Henning Müller
MIE 2012, Pisa, Italy

Overview

- Improving search in the medical domain
  - Allan and Henning
- Searching for medical images
- Who searches for medical images and how?
- Combining text and visual search
- Challenges for search
  - Allan and Henning

2

Improving search in the medical domain

Scenario

- A close friend of yours has her 5-year-old daughter diagnosed with Leukemia
- As you are regarded as a medical professional in the largest sense they ask you to help find more information on the disease
  - How do you search for this information, what are the strategies?
  - What type of information are you targeting?
  - How can you assure that trustful information is being transferred?
  - How do you explain it to the 5-year-old?
  - What are difficulties and disadvantages of this approach?

Search target?

- Is a document searched for?
- Or an answer to a specific question?
- Maybe an expert in the domain?
- Educational material?
  - Maybe videos, or nice pictures also for children?
- Description of the disease vs. symptoms, treatments, chances of survival?
  - With the goal of making a choice of treatment, for example, comparing various options for the choice

5

Trustfulness

- How can we make sure that information is correct, or at least not totally wrong?
  - Sometimes differing opinions exist, knowledge changes over time
  - Cross check several sources, but this can also lead to wrong ways
  - Quality labels or certificates of trust can help
    - HONcode
- Classification of web pages into several classes
  - What are the characteristics
  - Success needs to be monitored closely (spam filtering)

6

Level of understanding

- Different target groups have a totally different level of information that is required
  - Also physicians between a specialist and a GP
- What about children and the elderly?
- Level of understanding changes the more we read about a specific topic, we become experts
  - Not a fixed thing …
- Users can give more information to get personalized results
- Cyberchondria

7

Document and page ranking

- Should wikis and blogs be rated highest?
  - Google does this but for medical information several studies show that this might not be the best strategy
- Most queries are short (1-2 terms) and thus not specific at all, how to deal with this?
- Is the search target known?
- Users want advanced search options but then these are rarely used in practice
- Go beyond ranking, explore the content space

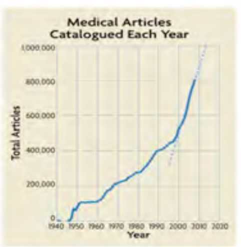
8

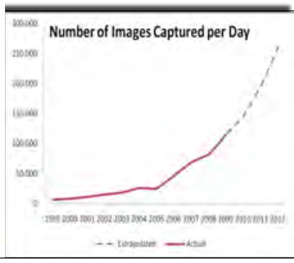
Searching for medical images

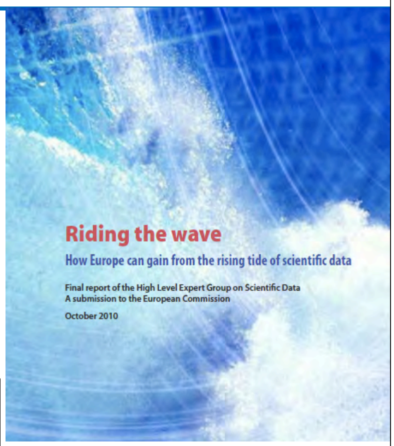
9

Motivation

- Medical imaging is estimated to occupy 30% of world storage capacity in 2010!
- Mammography data in the US in 2009 amounts to 2.5 Petabytes







Riding the wave – how Europe can gain from the rising tide of scientific data, *report of the European Commission*, 10/2010.

10

Retrieval of images

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Text retrieval of images
  - Is there any text attached to the images?
  - Doing this manually is expensive, subjective, language dependent, …
  - Take text close to the images (such as captions)
  - Semantic concepts could help in some cases
- Visual retrieval of images
  - Using automatically extracted visual features
  - Content-based image retrieval (CBIR)
  - Query by Image Example(s) (QBE)
- Multimodal retrieval (text+images)

11

Types of annotation

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- What is in the image?
- What is the image about?
- What does the image evoke (feelings)?

Content vs. context

- Most often text around images does not describe the image content itself
  - Unless specifically annotated for retrieval
  - Text often gives the context in which the images were taken (private, also medical)
- Image content is rarely described precisely and completely with text
  - Visual features describe the pure content
  - Low level of semantics
- Content and context are complementary for search

13

Age matters

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

14

Types of information needs/searches

- Known-item search (i.e. telephone number, …)
  - Question answering
- Exploration, exploitation, informational
  - Topic search or open ended search
- Comparison search (between things)
- Expert search, person search, entity search
- Geographical search
- Literature search
- Multimedia search (increasingly given as results)
- Browsing, no specific goal

15

Types of image searches

- Recent (time) pictures (journalists)
  - Date, given with camera
- Pictures of specific places, monuments
  - GPS in many cameras
- Pictures with particular persons (private search)
  - Face detection, recognition (Facebook, Picasa)
- Pictures with particular objects, types of images
- Pictures evoking specific feelings
  - Fear, joy, happiness, …
- Similarity search/browsing (Medicine, journalism)

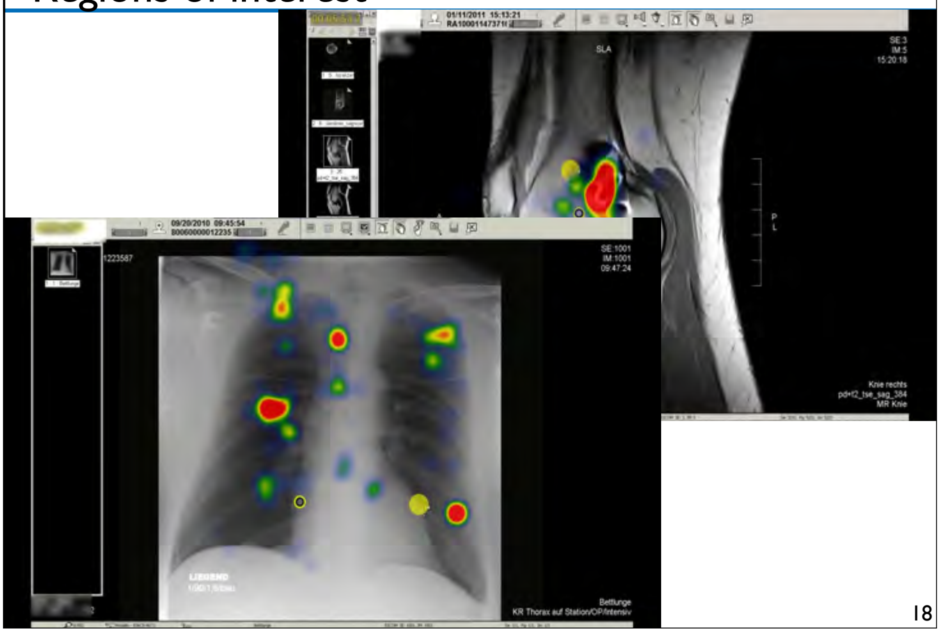
16

Visual information for retrieval

- Object detection
  - Then potentially mapping the objects to an ontology
  - Usually works well for a small number of objects
- Image classification
  - Training data, limited set of classes
  - Global classification of images vs. local classification of pixels, regions
- Similarity retrieval of images
  - Global image information, regions of interest (ROIs), small
  - No training data, relevance as criterion for quality

17

Regions of interest

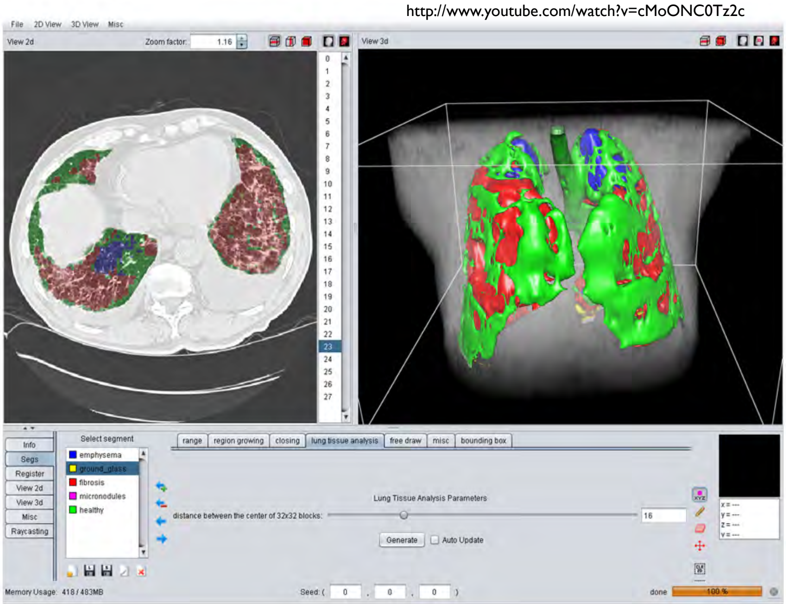


18

Visual classification and retrieval

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

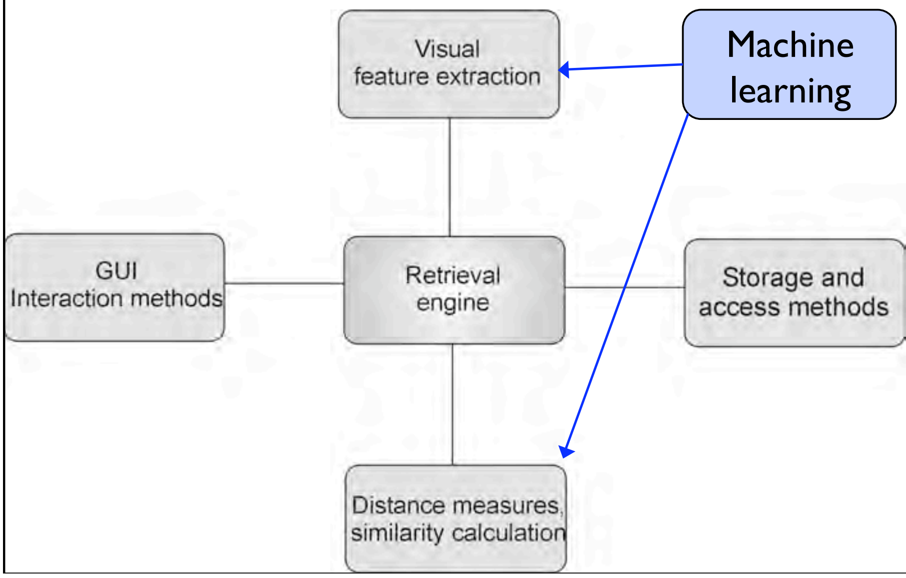
http://www.youtube.com/watch?v=cMoONC0Tz2c



19

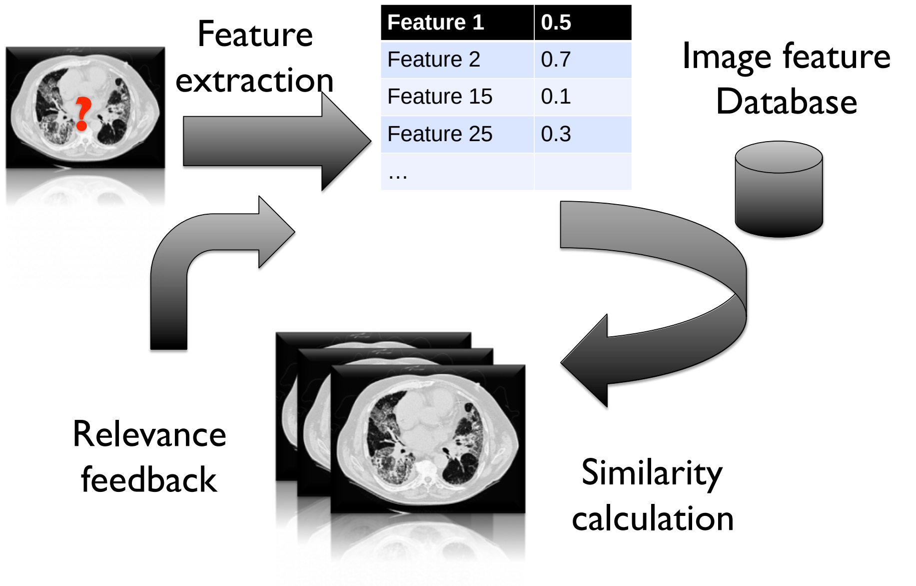
Components of image retrieval systems

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL



20

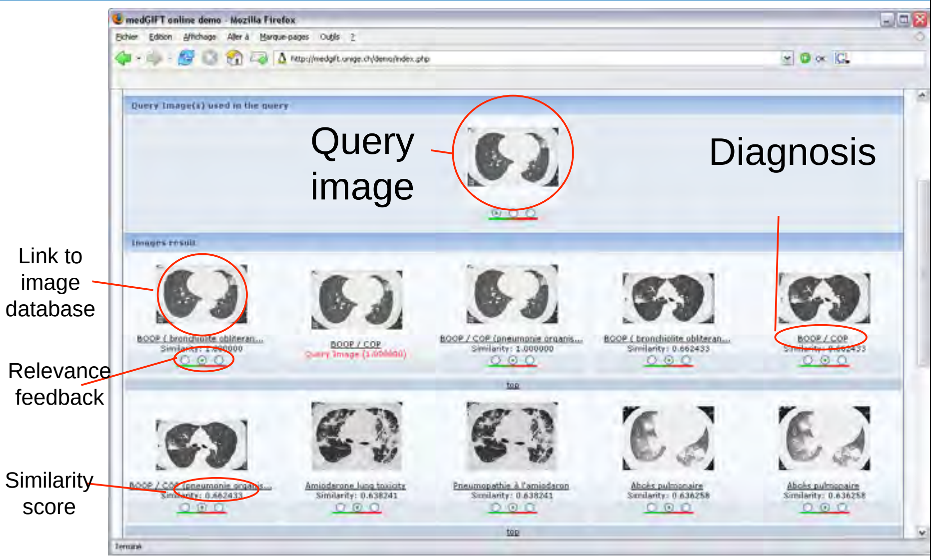
Model for CBIR



Steps of a retrieval system

- Image pre-processing
  - Normalization, background removal, …
- Salient region or point detection
- Visual feature extraction
  - Feature selection
  - Then feature modeling in the case of visual words
- Distance calculation or similarity measurement
- Results fusion (text, visual, …) for ranked list
- Results filtering (i.e. by modality)
- …

A typical (old) interface



23

Classifying visual features (Eakins)

- Level 1: primitive features
  - Color, texture, shape, spatial organization

Semantic gap

- Level 2: derived features
  - Individual objects or persons (Eiffel tower, Britney Spears)
  - Objects of a specific type (Volkswagen car)
- Level 3: abstract attributes
  - High level reasoning about meaning and purpose
  - Emotional or religious significance
  - Find images of “suffering”

24

Visual words

Salient regions

All pixes, grid, high gradient

Original feature space

- N-dimensional feature vector for each salient region

Quantization of the feature space

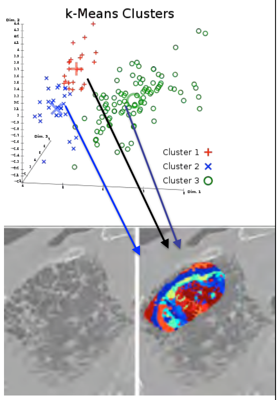
- Division of the feature space into M groups: visual words
- Clustering (k-means)
- Cluster centers are words

Visual words space

- Optimal number of words needs to be found
- For each image a histogram can be created
- Analogy to text words

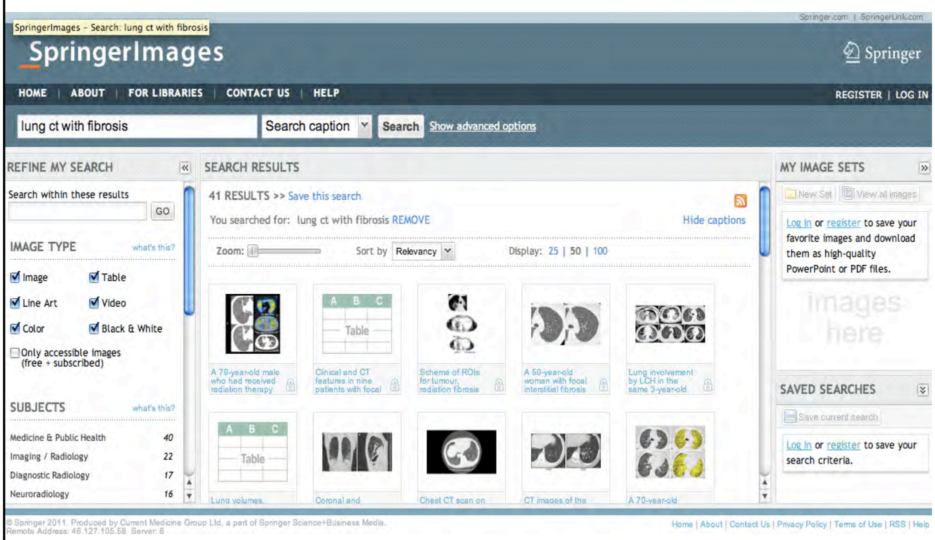
Bag of Visual Words

- Histogram of words for an image or an image region based of the salient points
- M-Dimensional feature vector



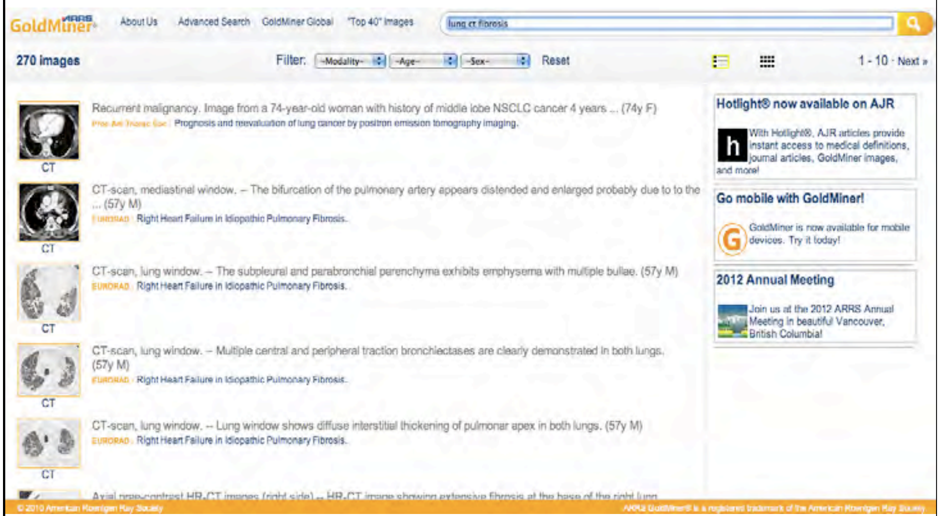
25

www.springerimages.com (3.3 mio)



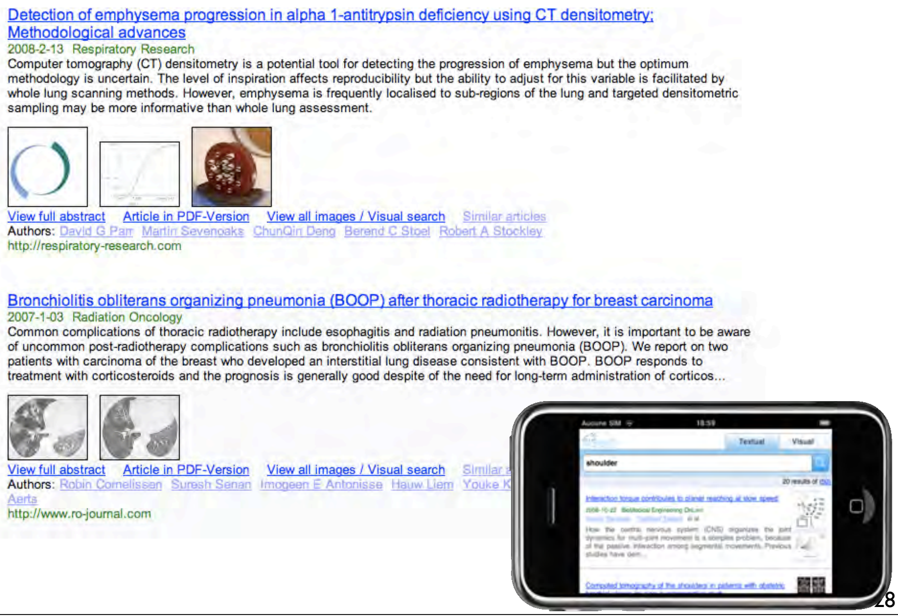
26

Goldminer.arrs.org (249,000 images)



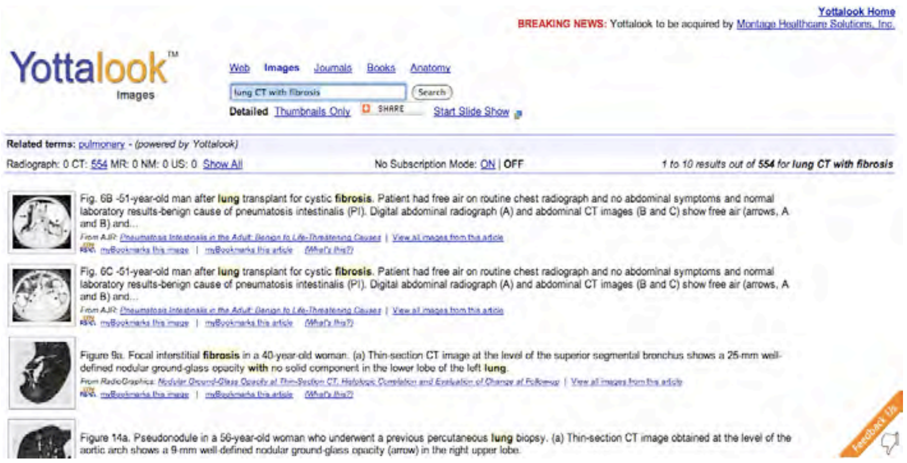
27

medgift.hevs.ch/, demos (300,000 images)



28

www.yottalook.com (70,000 images)



29

Who searches for medical images and how?

30

Background

khresmoi MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Health professionals' image search behavior has been subject to several surveys and log file analyses (MedLine, HONmedia, Goldminer)
- Goldminer log files of over 200'000 searches most comprehensive so far
- Khresmoi project performed a survey among radiologists to develop first prototypes
- Results change over time as user requirements change with the generation of physicians
  - Current medical students grow up with Google, Facebook and iPhones

31

Methods of Khresmoi survey

khresmoi MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Survey among radiologists, mainly in Geneva and Vienna University hospitals
  - Paper and electronic version
  - Survey filled in with a person explaining the goals
  - Survey took approximately 1 hour, research/clinical /teaching work separated with same questions
- 26 radiologists answered, 13 from Austria, 9 from Switzerland, form only sent on invitation
- 17 males, 9 females, mainly 26-35 years, few years of experience

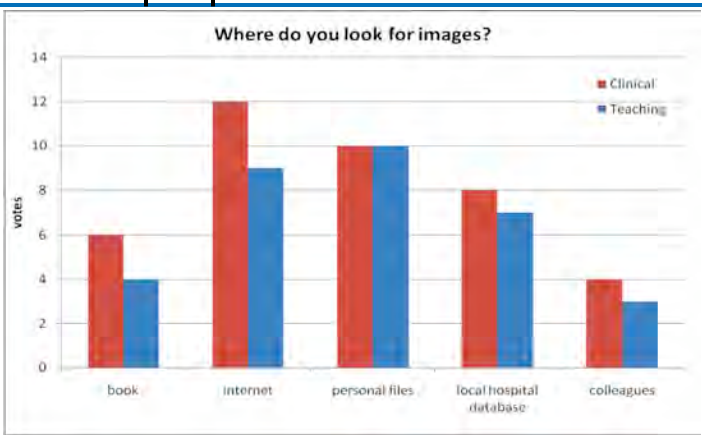
32

Reasons for image search

- Clinicians: Information on unclear cases, illustrate presentations, differential diagnosis were most frequent
- Teaching: Find similar cases, for example an easy, a medium and a tricky case for the same disease, problem-based learning
- Problem-based learning requires increased search skills for the students

33

Where people search

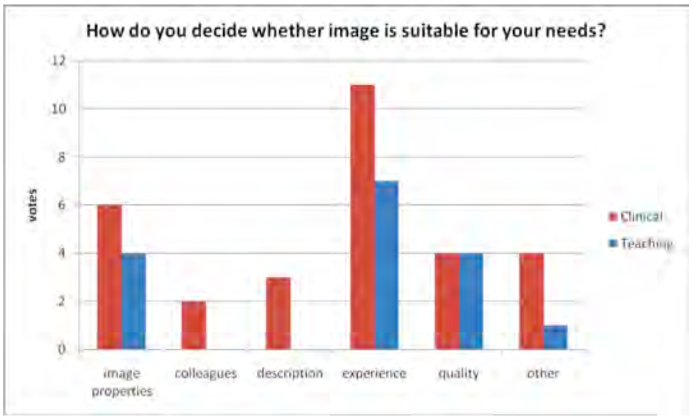


- The Internet replaces books and colleagues
- Personal files are not always optimal

34

Determination of relevance

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

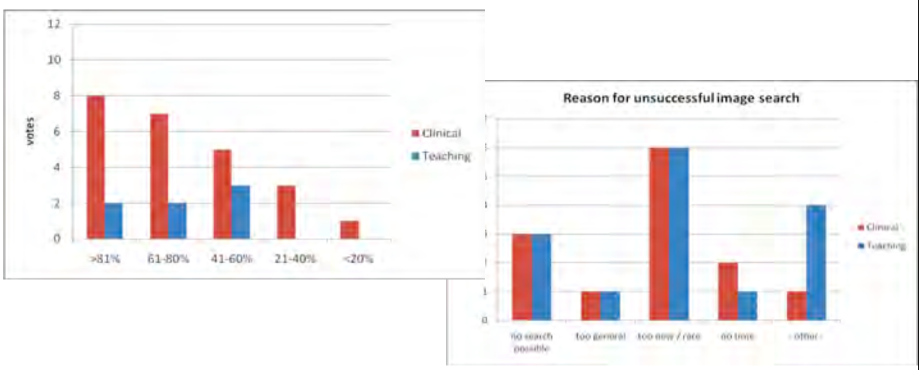


- Experience determines relevance
- Often additional proof such as biopsies or other clinical tests are requested

35

Success rates

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL



- For teaching success rates are higher
- Clinical work might have less well defined tasks, average of success at 60%

36

Search time

- 70% of successful searches less than ten minutes
- Failure often after over 15 minutes
- Less time available for clinical search than research/teaching
- Could a faster and more targeted image search system help?

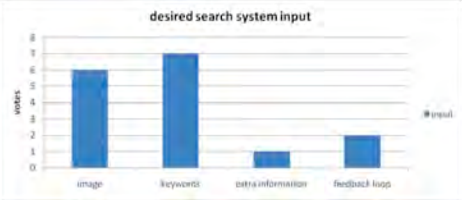
37

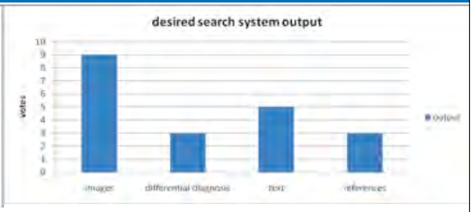
Useful additions

- Search by
  - pathology (13 times)
  - modality (10 times)
  - patient demography (6 times)
  - similar images (8 times)
- Other comments:
  - Multilingual retrieval
  - Pathology classification (using ontologies)
  - Search in 3D data
  - Confidence in the diagnosis

38

A perfect search system



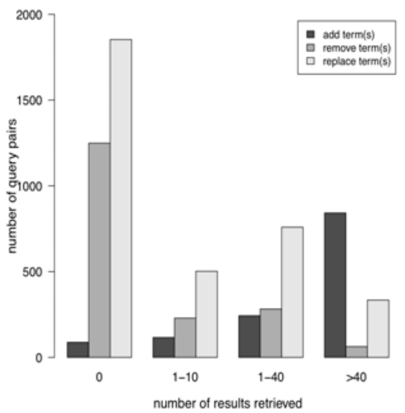


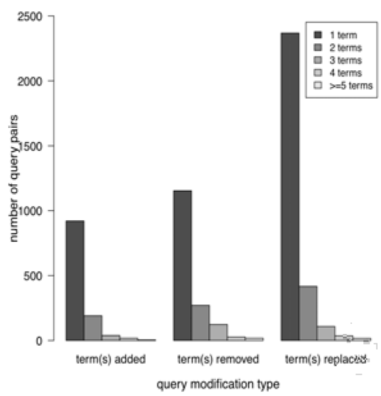
- Many free comments
  - Like Google but DICOM and text
  - Structure information and confidence in diagnosis
  - Search by regions of interest
  - Social networking, comments of other physicians
  - Search for similar cases
  - Quantification of structures (size, volume)

39

Goldminer log files

- Monday 12h15, talk on the analysis
- 25'000 consecutive queries





40

Combining text and visual search

41

Background

- ImageCLEF image retrieval benchmark has been run each year since 2004
- 12-20 research groups compare their tools and approaches on the same tasks and DBs
  - Visual, textual and combined approaches are used
  - Multilingual approaches are also possible
- Sometimes visual, sometimes textual and sometimes mixed approaches perform best
  - No clear outcome
  - Combination of results can be delicate, instable

42

Example search topic

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Show me x-ray images of a tibia with a fracture.
- Zeige mir Röntgenbilder einer gebrochenen Tibia.
- Montre-moi des radiographies du tibia avec fracture.



43

Visual vs. semantic vs. mixed searches

khresmoi
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Experts can predict what type of technique will most likely perform best
  - Can this prediction be modeled automatically?
  - If results were visually relatively homogeneous visual search can work
    - Same anatomic region, same view, same modality
  - If results are expected to be very different (no modality given), text would work best
  - Combinations often work best when some common aspects but some variability as same modality
- Visual and text retrieval are complementary

44

Early vs. late fusion

- Early fusion
  - Feature spaces are directly combined, so visual features and textual words treated in the same way
  - Number of features needs to be similar to avoid bias
- Late fusion
  - Results of systems are combined, not features
  - Each system can have a varying number of features
- For text/visual combinations late fusion is often simpler to employ and works better
  - When using visual words both could be used

45

Score based vs. rank-based fusion

- Score-based fusion
  - Score of the single systems is used for the combination of the results sets
  - Score needs to be normalized, potentially to have similar characteristics
- Rank-based fusion
  - The rank of an element is used to calculate fusion
    - Can be linear or logarithmic or in another form
  - Avoids the bias that very differing results sets of system can have
  - Often have better results when visual and textual systems are combined

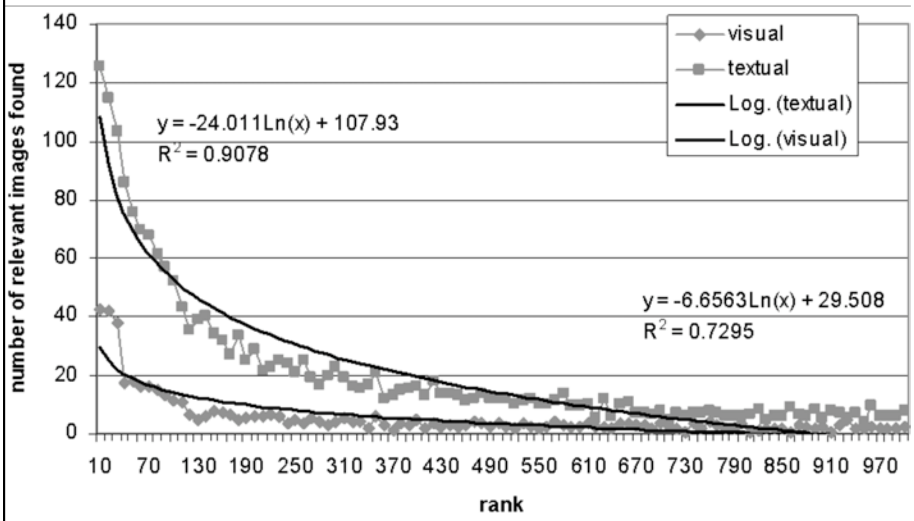
46

Types of fusion techniques

- Many types of fusion techniques exist
  - combSUM : $V_{combSUM}(i) = \sum_k V_k(i)$
  - combMAX : $V_{combMAX}(i) = \text{argmax}(V_k(i))$
  - combMNZ : $V_{combMNZ}(i) = f(i) \bullet V_{combSUM}(i)$
    - Where f(*i*) is the frequency of image *i* in the results
- At the ICPR 2010 conference a competition on fusion techniques was organized using the best ImageCLEF runs
  - Rank-based techniques using logarithmic decrease performed best in a variety of different approaches

47

Distribution of relevant documents



48

Other combinations

- Modality detection (using visual techniques or text of the captions) can work very well (80+%)
- Allowing to select the target modality can improve image search
  - Tests with all runs of ImageCLEF 2009
  - Many search engines allow for this such as Goldminer
    - This can be used for tabbed browsing as well
- Exclusion criteria for images can be chosen based on the text
  - Age group, gender, …

49

Combinations for case-based retrieval

- Mix of free text, structured data, images, and many other forms
- Interactions of the data can vary strongly between patients and diseases, also over time
- More complex combinations for images need to be found
  - Match images between case to match for similarity
- Currently text is better than most fusions
- Case description including images without diagnosis, find images for differential diagnosis

50

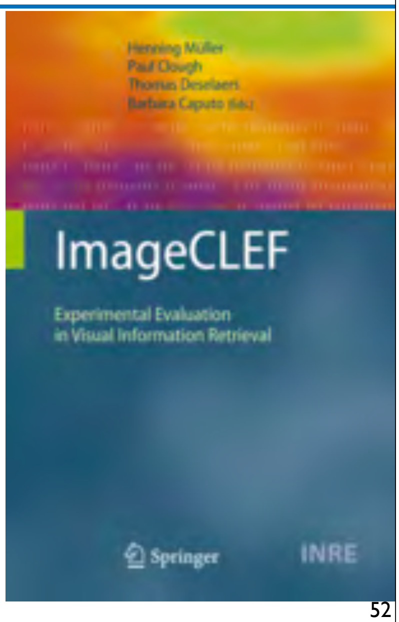
Some ImageCLEF lessons learned

- Text retrieval techniques are stable and deliver good results (i.e. Lucene is above average)
- Visual has had less evolution than text retrieval
  - GIFT (old!) has still relatively good results
    - Semantic gap is very present
  - Visual words-based approaches can be much better when using high quality training data
- Interactive retrieval can improve visual retrieval
- Many features combined deliver best results
- Mapping of images and text to ontologies can help
  - Improve semantic retrieval

51

Attention: advertisement

- ImageCLEF book
- All on image retrieval
  - Methods of evaluation
  - Task overviews
  - Participant reports
    - The best techniques
    - Industrial requirements
  - Industry perspectives
  - Specific techniques such as fusion



52

Challenges for search?

BIG data

- How can scalability be assured when treating extremely large amounts data?
  - 250'000 images per day in Geneva …
    - 150 TB of images in Catalonia archive
  - Extremely large scales allow solving many new problems
    - Rare diseases
    - Sufficiently large training data sets
- Hadoop/Mapreduce as is also used by Google, Yahoo or Facebook
- Use of cloud computing
  - Costs, confidentiality, also bandwidth …

Confidentiality

- Can patient records be made available?
  - Maybe partly, anonymized, only internally?
  - Could the data warehouse be used for this
  - Secondary use of data
- Availability for data mining not a specific very limited scenario (as ethics committees request)
- Can interoperability be assured using the same semantic standards
- How to link the literature with specific cases
  - Images not in the same quality, much more than just case descriptions, …

55

Search for medical cases not images

- Combine several data sources
  - Importance of each source is not fully clear
    - Interaction between content importance is complex
  - Different media from free text, structured data to semantics, signals or images including 3D, 4D
- Some data sources might be missing
  - Questions not asked, not documented, errors
  - How to deal with missing data
- What is a case exactly?
  - Limited period of time? Or on a patient basis including old data?

56

Diversity in the results

- Having a list of almost identical texts as result is not useful
  - Google filters out near duplicates
- Consistency vs. diversity have limits
- Some search systems cluster results and the present each cluster in a first step
- Diversity can favor data exploration and user relevance feedback
- Understand links between documents and content in the results

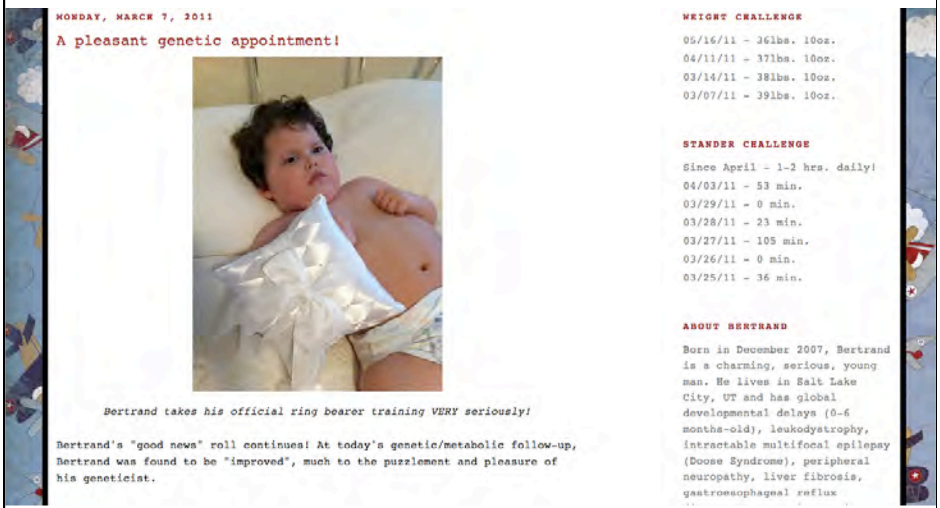
57

Retrieval from social networks

- People share data in social networks
  - Blogs, facebook, …
  - Sometimes more than any physician would share
    - People with rare diseases are sometimes desperate …
- People can comment on data
- Goals of comments are not always clear
  - Spam can create problems
- Some metadata is available, so free text, semi -structured metadata, images
  - Can semantics help with this?

58

A medical blog on movement disorder



59

Patientslikeme

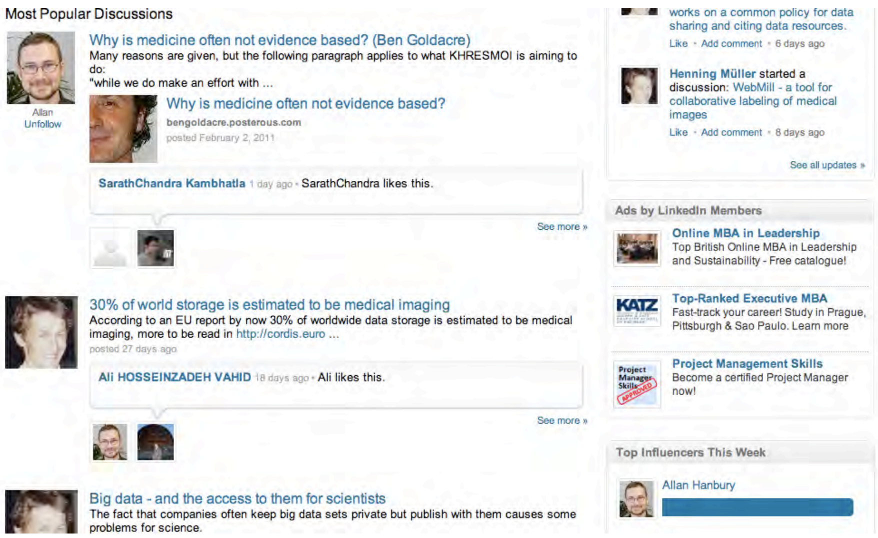
60

Visceral project

- VISual Concept Extraction challenge in RAdioLogy
  - Most likely as a MICCAI workshop in 2013 and 2014
- Two challenges
  - Identify organs in 3D data sets
  - Find similar cases using 3D data and rdiology reports
- Very large amounts of data (~10 TB)
  - Data distribution via the cloud
    - Participants will get a virtual machine for free
  - Creation of a gold and silver standard

61

LinkedIn Khresmoi group



62

A few references

- Eakins J, Graham M, Content-based Image Retrieval, JTAP-039, JISC report, 1999.
- Enser P, Visual image retrieval: seeking the alliance of concept-based and content -based paradigms, Journal of Information Science, 2000.
- Müller H, Michoux N, Bandon D, Geissbuhler A. A re-view of content-based image retrieval systems in medical applications-clinical benefits and future directions. Int J Med Inform. 2004 Feb ;73(1):1-23.
- Müller H, Despont-Gros C, Hersh W, Jensen J, Lovis C, Geissbuhler A, Medical professionals' image search and use behaviour, pages 24-32, Medical Informatics Europe, Mastricht, The Netherlands, 2006.
- Müller H, Kalpathy-Cramer J, Kahn CE Jr., Hatt W, Bedrick S, Hersh W. Overview of the ImageCLEFmed 2008 Medical Image Retrieval Task, Springer Lecture Notes in Computer Science 5706, 500-510, 2009.
- Smeulders AVM, Worring, Santini S, Gupta A, Jain R. Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Ma-chine Intelligence, Vol. 22(12), pages 1349-1380, 2000.Datta
- Vasconcelos N, Lippman A, A Probabilistic Architecture for Content-based Image Retrieval, IEEE International Conference on computer vision and pattern recognition, volume 1, pages 216-221, 1999.
- …

63

Conclusions

- Medical information search and access is an important technique in medicine
  - And this includes images!!
- Image information is most often complementary to text
- Visual information such as regions of interest can be used to formulate queries
  - Radiologists request this increasingly instead of searching in books and discussing with colleagues
- There is still much to be learning for combining visual and texual techniques

64

Questions?

- More information available from
  - http://www.imageclef.org/
  - http://khresmoi.eu/
  - http://medgift.hevs.ch/
  - Contact henning.mueller@hevs.ch

SEVENTH FRAMEWORK PROGRAMME

65