



# Topic creation for medical image retrieval benchmarks

Image  
CLEF

ImageCLEF/MUSCLE workshop,  
Alicante, 19.9.2006

Henning Müller,  
Bill Hersh



UNIVERSITÉ DE GENÈVE

HUG    
Hôpitaux Universitaires de Genève



## Overview

- Image retrieval benchmarking and applications
  - Components
- Medical image retrieval
- Finding out more on **information needs**
  - Analysis of the content of our dataset
  - Surveys among professional users
  - Log file analysis (foundation health on the net)
- Examples
- Conclusions

## Image retrieval and evaluation

- Retrieval vs. classification
  - Nothing is known on a retrieval dataset
- In other domains **standard datasets** have existed for a long time
  - Text retrieval, segmentation, character recognition, ...
- Image retrieval starts getting better
  - Benchathlon
  - TRECVID
  - ImageCLEF
  - ImageEval, ...



## Components of a benchmark

- A **dataset**
  - Large (! Problems are different !)
  - Realistic with respect to a certain user model
    - Annotation, etc.
- Query **topics** based on real information needs
- **Participants** for comparison
- Ground truth/Relevance **judgments**
- Performance **measures**
- **Workshop**
  - Foster discussions, not a pure competition



## Medical image retrieval

- **Research** domain
  - Users are often technophobe
- Frequently proposed as important (potential) but never really used in practice
  - A single study on diagnostic use
- Most users work with Google but do not know anything about visual retrieval
  - Problems and possibilities
- Use on varied dataset vs. Diagnostic aid (very specific databases)

## Motivation

- Find out more on the **behavior** of medical professionals regarding the use of images
  - How is searched for images?
  - What can be useful in the future?
- **Educate** them on the techniques available and their possibilities
- **Stimulate** creativity to learn about potentially good applications
  - Brainstorming
- Goal is **multimodal** image retrieval (visual included)
  - For ImageCLEFmed

- Query topics were **images, only**
  - Radiologist familiar with the database choose them
  - Represent the database well with its variability
  - Text could be used for subsequent steps
- Goal was to retrieve images similar/same in anatomic region, modality, and view
- Well defined task ... but is this realistic?
- **User model MD**
  - Would they search with an image only?
  - How to get the image?



## Surveys among medical professionals

- In Portland and Geneva
- Separated by **function**
  - Librarian
  - Student
  - Lecturer
  - Researcher
  - Clinician
- Get **typical search tasks** as examples
  - From various departments
- **Qualitative**, not too time consuming



## Questions at the survey

- What kind of **tasks** do you perform in your daily work where images are useful for you?
- For each of these tasks, can you give us an **example** of what kind of image you are searching for?
- For each of these tasks, **where** do you search for the images? (Ordered by preference)
- When you search for images, **how** do you search for them?
- When you find an image, how do you decide whether one or another corresponds to your **needs**?
- What search **tools** or functions would be useful for you to search for images in addition to what is currently used?

## Some results

- Search tasks vary strongly between functions
- Clinicians often do not have much choice
  - Access per patient and by patient id
- Several people did not know about visual retrieval
- Retrieval for **pathology** was regarded as most important
  - And currently not possible
- Retrieval of **similar cases** was proposed as very useful several times



## Log file analysis of a medical media search

- Health On the Net (<http://www.hon.ch/>)
- 35'000 query terms of a one year query log
  - HONmedia search for medial images and videos
  - Spelling errors
  - Several languages
- Calculate **frequencies** of term combinations
- **Removal** of media types (images, photos, videos, ...) from the queries
- Removal of frequent spelling errors
- Change of word order (alphabetic)

## Some results

- Half of the queries are **unique!!**
- Almost the majority of queries contains one word
- Queries are most often **not specific** at all
  - Risk to have thousands of results!!
    - Heart
    - Lung
    - Images/videos
- People do not only search for health subjects
- Few very specific questions
  - But these were very specific!

## Other sources ...

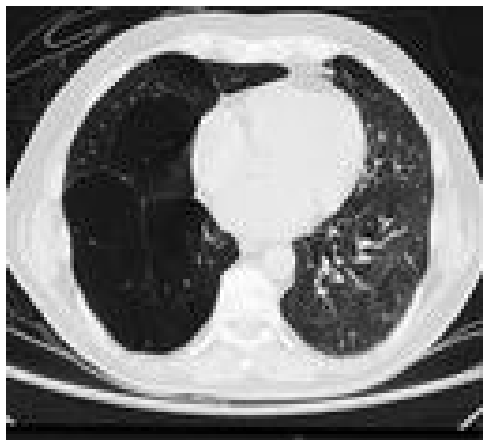
- Content of the **data base** needs to be taken into account to have varied queries
- **Frequent** causes of death are most important (CDC)
- Develop variety along four **axes**
  - Modality
  - Anatomic region
  - Pathology
  - Visual observation

## ... and constraints

- Number of relevant items needs to be **limited**
  - Otherwise we would miss many relevant
- There should be **at least a few** relevant items
- How to choose images for the queries
  - From collection, modified, from the web?
- We would like visual, mixed and semantic queries
  - Satisfy all participants
- Create **candidates** and then reduce number
- Create **unambiguous** topics!
  - A negative description for judges can help

## Exemples 2005

Show me x-ray images with fractures of the femur.  
Zeige mir Röntgenbilder mit Brüchen des  
Oberschenkelknochens.  
Montre-moi des fractures du fémur.



Show me chest CT images with emphysema.  
Zeige mir Lungen CTs mit einem Emphysem.  
Montre-moi des CTs pulmonaires avec un emphysème.

Show me any photograph showing malignant melanoma.  
Zeige mir Bilder bösartiger Melanome.  
Montre-moi des images de mélanomes malignes.



## Example 2006

3.6  
Show me x-ray images of bone cysts.  
Zeige mir Röntgenbilder von Knochenzysten.  
Montre-moi des radiographies de kystes d'os.





## Example 2006 (2)

1.4

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.



## Conclusions

- Topic creation is extremely important for benchmarks
  - Need to be useful for user model, not purely academic
- Several sources can be used even if no real use of system is available
- **Discussions** with professionals can bring up many good ideas (and educate your users)
- A development in several steps helps to correspond to all **constraints**
  - Define constraints in advance
  - Start with a larger number and then reduce
    - But: robustness



## Questions?

[henning.mueller@sim.hcuge.ch](mailto:henning.mueller@sim.hcuge.ch)

<http://www.sim.hcuge.ch/medgift/>

<http://ir.shef.ac.uk/imageclef/>

<http://ir.ohsu.edu/image/>