The Cross Language Image Retrieval Track: ImageCLEF 2007

Henning Müller[1], Thomas Deselaers[2], Michael Grubinger[3], Allan Hanbury[4], Jayashree Kalpathy-Kramer[6], Thomas M. Deserno[5], Bill Hersh[6], Paul Clough[7]

[1]University and Hospitals of Geneva, Switzerland
[2]RWTH Aachen University, Computer Science Dep. Germany
[3]Victoria University, Australia
[4]Vienna University, Austria
[5]RWTH Aachen University, Medical Informatics, Germany
[6]Oregon Health Science University
[7]Sheffield University, UK

ImageCLEF 2007

- General overview
  - Participation
  - Problems
- Photo retrieval task
- Medical image retrieval task
- Medical image annotation
- Object retrieval task
- Generalizations and conclusions

General participation and news

- 51 overall registrations from all continents
  - More than 30 groups submitted results

News:

- More realistic database for photo retrieval
- Larger database for medical retrieval
- Hierarchical classification of medical images
- New object retrieval task

Photographic Retrieval Task

- **ImageCLEFphoto 2007**
  - Evaluation of visual information retrieval from a generic photographic collection
  - IAPR TC-12 Benchmark (2nd year)
  - New subset this year: lightly annotated images
- **Research questions**
  - Are traditional text retrieval methods still applicable for such short captions?
  - How significant is the choice of the retrieval language?
  - How does the retrieval performance compare to retrieval from collections containing fully annotated images (compared to *ImageCLEFphoto 2006*)?
- **Additional Goal**
  - Attract more groups using content-based retrieval approaches

Image Collection

- **IAPR TC-12 image collection**
  - 20,000 generic colour photographs
  - taken from locations around the world
  - provided by an independent German travel organisation (viventura)
  - created as a resource for evaluation

- **Many images have similar visual content but varying**
  - illumination
  - viewing angle
  - background

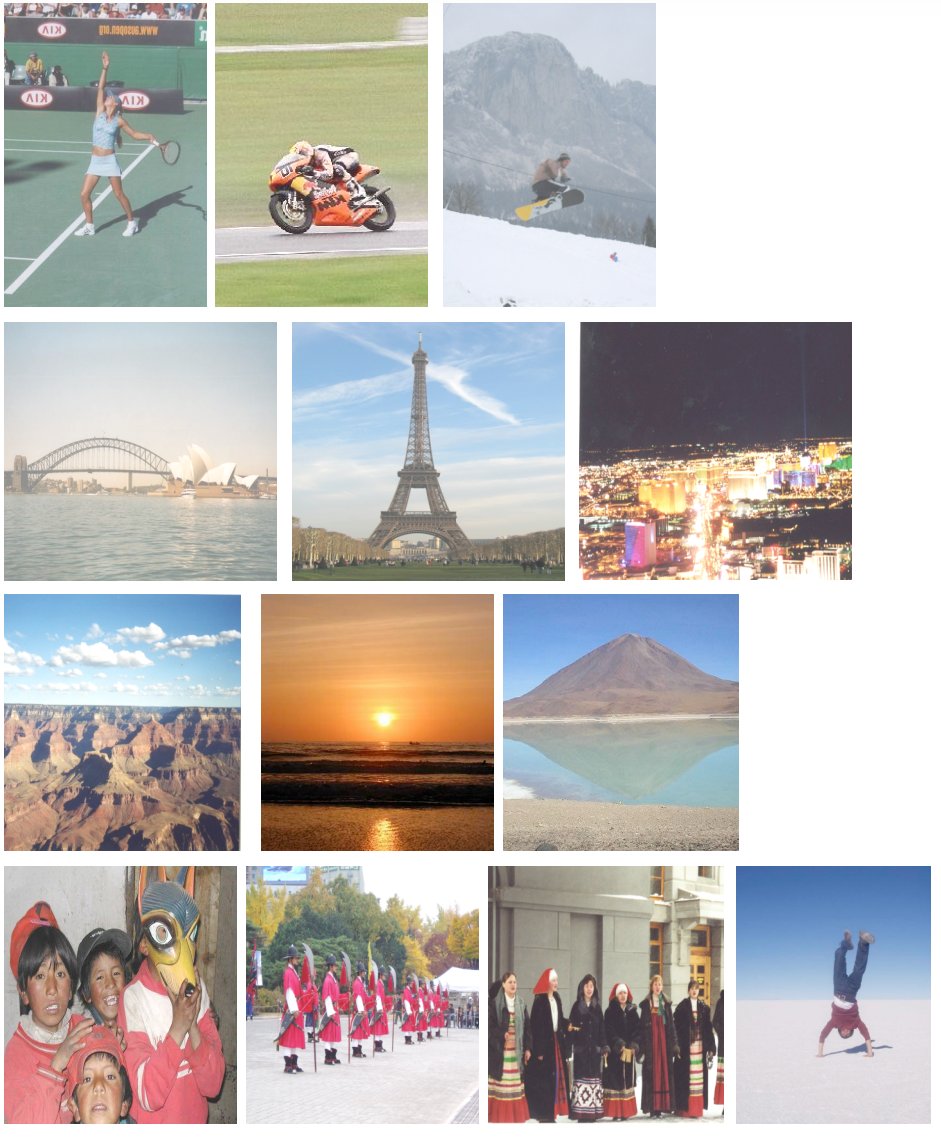


Image Captions

- **Accompanied by semi-structured captions:**
  - English
  - German
  - Spanish
  - Randomly chosen
- **Subset with "light" annotations**
  - title, notes, location and date provided
  - semantic descriptions NOT provided

```
<DOC>
<DOCNO>annotations/16/16019.eng</DOCNO>
<TITLE>Flamingo Beach</TITLE>
<DESCRIPTION> a photo of a brown sandy beach; the dark blue sea with small
breaking waves behind it; a dark green palm tree in the foreground on the left; a
blue sky with clouds on the horizon in the background; </DESCRIPTION>
<NOTES> Original name in Portuguese: "Praia do Flamengo"; Flamingo Beach is
considered as one of the most beautiful beaches of Brazil; </NOTES>
<LOCATION>Salvador, Brazil</LOCATION>
<DATE>2 October 2002</DATE>
<IMAGE>images/16/16019.jpg</IMAGE>
<THUMBNAIL>thumbnails/16/16019.jpg</THUMBNAIL>
</DOC>
```

Image Captions

- **Accompanied by semi-structured captions:**
  - English
  - German
  - Spanish
  - Randomly chosen
- **Subset with "light" annotations**
  - title, notes, location and date provided
  - semantic descriptions NOT provided

```
<DOC>
<DOCNO>annotations/16/16019.eng</DOCNO>
<TITLE>Flamingo Beach</TITLE>
<DESCRIPTION> a photo of a brown sandy beach; the dark blue sea with small
breaking waves behind it; a dark green palm tree in the foreground on the left; a
blue sky with clouds on the horizon in the background; </DESCRIPTION>
<NOTES> Original name in Portuguese: "Praia do Flamengo"; Flamingo Beach is
considered as one of the most beautiful beaches of Brazil; </NOTES>
<LOCATION>Salvador, Brazil</LOCATION>
<DATE>2 October 2002</DATE>
<IMAGE>images/16/16019.jpg</IMAGE>
<THUMBNAIL>thumbnails/16/16019.jpg</THUMBNAIL>
</DOC>
```

Query Topics

- **60 representative search requests**
  - reused topics from 2006
  - topic titles in 16 languages
  - narrative descriptions NOT provided
  - 3 sample images (removed from collection)
  - balance between realism and controlled parameters
- **Distribution**
  - 40 topics taken directly from log file (10 derived; 10 not)
  - 24 topics with geographical constraint
  - 30 topics semantic; 20 mixed and 10 visual
  - 4 topics rated as linguistically easy, 21 medium, 31 difficult; 4 very difficult

```
<top>
<num> Number: 1 </num>
<title> accommodation with swimming
pool </title>
<narr> Relevant images will show the
building of an accommodation facility
(e.g. hotels, hostels, etc.) with a
swimming pool. Pictures without
swimming pools or without buildings
are not relevant. </narr>
<image> images/03/3793.jpg </image>
<image> images/06/6321.jpg </image>
<image> images/06/6395.jpg </image>
</top>
```



Query Topics

- **60 representative search requests**
  - reused topics from 2006
  - topic titles in 16 languages
  - narrative descriptions NOT provided
  - 3 sample images (removed from collection)
  - balance between realism and controlled parameters
- **Distribution**
  - 40 topics taken directly from log file (10 derived; 10 not)
  - 24 topics with geographical constraint
  - 30 topics semantic; 20 mixed and 10 visual
  - 4 topics rated as linguistically easy, 21 medium, 31 difficult; 4 very difficult

```
<top>
<num> Number: 1 </num>
<title> accommodation with swimming
pool </title>
<narr> Relevant images will show the
building of an accommodation facility
(e.g. hotels, hostels, etc.) with a
swimming pool. Pictures without
swimming pools or without buildings
are not relevant. </narr>
<image> images/03/3793.jpg </image>
<image> images/06/6321.jpg </image>
<image> images/06/6395.jpg </image>
</top>
```



Result Generation & Participation

- **Relevance Judgments**
  - pooling method (n = 40)
  - average pool size: 2,299 images (max: 3237; min: 1513)
  - Interactive Search and Judge to complete with further relevant images
  - qrels(2007) UNION qrels(2006)
- **Performance Indicators**
  - MAP
  - P(20)
  - GMAP
  - BPREF
- **Participation and Submissions**
  - 32 groups registered (2006: 36)
  - 20 groups submitted (2006: 12, 9 new)
  - 616 runs (!!!) were submitted (2006: 157)
  - All runs were evaluated

ALICANTE, Alicante, Spain
BERKELEY, Berkeley, USA
BUDAPEST, Budapest, Hungary
CINDI, Montreal, Canada
CLAC, Montreal, Candada
CUT, Chemnitz, Germany
DCU-**UTA**, Dublin/**Tampere**, Ireland/**Finland**
GE, Geneva, Switzerland
IMPCOLL, London, UK
INAOE, Puebla, Mexico
IPAL, Singapore
MIRACLE, Madrid, Spain
NII, Tokyo, Japan
NTU, Hong Kong, China
NTU, Taipei, Taiwan
RUG, Groningen, The Netherlands
RWTH, Aachen, Germany
SIG-IRIT, Toulouse, France
SINAI, Jaen, Spain
XRCE, Meylan, France

Submission overview by topic and annotation languages

Query / Annotation	English	German	Spanish	Random	None	Total
English	**204 (18)**	18 (5)	6 (3)	11 (2)		**239 (18)**
German	31 (6)	**18 (5)**	1 (1)	11 (2)		**74 (9)**
Visual	1 (1)				52 (12)	**53 (12)**
French	32 (7)	1 (1)	**10 (2)**			**43 (7)**
Spanish	20 (5)		16 (7)	2 (1)		**38 (9)**
Swedish	20 (3)	12 (1)				**32 (3)**
Chinese (T+S)	28 (4)			1 (1)		**29 (4)**
Portuguese	19 (5)			2 (1)		**21 (5)**
Russian	17 (4)	1 (1)		2 (1)		**20 (4)**
Norwegian	6 (1)	12 (1)				**18 (1)**
Japanese	16 (3)					**16 (3)**
Italian	10 (4)			2 (1)		**12 (4)**
Danish		12 (1)				**12 (1)**
Dutch	4 (1)			2 (1)		**6 (1)**
Total	**408 (18)**	**88 (8)**	**33 (7)**	**32 (2)**	**52 (12)**	**616 (20)**

Results – Highest MAP

Languages	Run ID	MAP	P(20)	GMAP	BPREF
ENG – ENG	CUT/cut-EN2EN-F50	0.3175	0.4592	0.2984	0.1615
GER – ENG	XRCE/DE-EN-AUTO-FB-TXTIMG_MPRF	0.2899	0.3883	0.2684	0.1564
POR – ENG	Taiwan/NTU-PT-EN-AUTO-FBQE-TXTIMG	0.2820	0.3883	0.2655	0.1270
SPA – ENG	Taiwan/NTU-ES-EN-AUTO-FBQE-TXTIMG	0.2785	0.3833	0.2593	0.1281
RUS – ENG	Taiwan/NTU-RU-EN-AUTO-FBQE-TXTIMG	0.2731	0.3825	0.2561	0.1146
ITA – ENG	Taiwan/NTU-IT-EN-AUTO-FBQE-TXTIMG	0.2705	0.3842	0.2572	0.1138
ZHS – ENG	CUT/cut-ZHS2EN-F20	0.2690	0.4042	0.2438	0.0982
FRA – ENG	Taiwan/NTU-FR-EN-AUTO-FBQE-TXTIMG	0.2669	0.3742	0.2480	0.1151
ZHT – ENG	Taiwan/NTU-ZHT-EN-AUTO-FBQE-TXTIMG	0.2565	0.3600	0.2404	0.0890
NED – ENG	Taiwan/NTU-JA-EN-AUTO-FBQE-TXTIMG	0.2551	0.3675	0.2410	0.0937
JAP – ENG	INAOE/INAOE-NL-EN-NaiveWBQE-IMFB	0.1986	0.2917	0.1910	0.0376
SWE – ENG	INAOE/INAOE-SV-EN-NaiveWBQE-IMFB	0.1986	0.2917	0.1910	0.0376
VIS – ENG	INAOE/INAOE-VISUAL-EN-AN_EXP_3	0.1925	0.2942	0.1921	0.0390
NOR – ENG	DCU/NO-EN-Mix-sgramRF-dyn-equal-fire	0.1650	0.2750	0.1735	0.0573
SPA – SPA	Taiwan/NTU-ES-ES-AUTO-FBQE-TXTIMG	0.2792	0.3975	0.2693	0.1128
ENG – SPA	CUT/cut-EN2ES-F20	0.2770	0.3767	0.2470	0.1054
GER – SPA	Berkeley/Berk-DE-ES-AUTO-FB-TXT	0.0910	0.1217	0.0717	0.0080

Results – Highest MAP

Languages	Run ID	MAP	P(20)	GMAP	BPREF
GER – GER	Taiwan/NTU-DE-DE-AUTO-FBQE-TXTIMG	0.2449	0.3792	0.2386	0.1080
ENG – GER	XRCE/EN-DE-AUTO-FB-TXTIMG_MPRF_FLR0	0.2776	0.3617	0.2496	0.1121
SWE – GER	DCU/SW-DE-Mix-dictRF-dyn-equal-fire	0.1788	0.2942	0.1802	0.0707
DAN – GER	DCU/DA-DE-Mix-dictRF-dyn-equal-fire	0.1730	0.2942	0.1759	0.0733
FRA – GER	CUT/cut-FR2DE-F20	0.1640	0.2367	0.1442	0.0039
NOR – GER	DCU/NO-DE-Mix-dictRF-dyn-equal-fire	0.1667	0.2700	0.1653	0.0701
ENG – RND	DCU/EN-RND-Mix-sgramRF-dyn-equal-fire	0.1678	0.2850	0.1751	0.0683
GER – RND	DCU/DE-RND-Mix-sgram-dyn-equal-fire	0.1572	0.2817	0.1669	0.0644
FRA – RND	DCU/FR-RND-Mix-sgram-dyn-equal-fire	0.1409	0.2642	0.1476	0.0593
SPA – RND	INAOE/INAOE-ES-RND-NaiveQE-IMFB	0.1243	0.2275	0.1355	0.0266
NED – RND	INAOE/INAOE-NL-RND-NaiveQE	0.0828	0.1558	0.0941	0.0114
ITA – RND	INAOE/INAOE-IT-RND-NaiveQE	0.0798	0.1442	0.0864	0.0181
RUS – RND	INAOE/INAOE-RU-RND-NaiveQE	0.0763	0.1358	0.0848	0.0174
POR – RND	INAOE/INAOE-PT-RND-NaiveQE	0.0296	0.0425	0.0317	0.0006
VISUAL	XRCE/AUTO-NOFB-IMG_COMBFK	0.1890	0.3517	0.2009	0.1016

Retrieval Result Summary

- **Concept-based Image Retrieval**
  - bilingual retrieval performs as well as monolingual retrieval
  - choice of topic language almost negligible as many of the short captions contain proper nouns
  - combining concept- and content-based retrieval improves retrieval performance (MAP 24% higher than retrieval based on text only)
  - using query expansion and relevance feedback techniques can improve retrieval results (MAP) by almost 100%
  - results of concept-based techniques only slightly weaker than 2006, indicating an improvement of retrieval techniques

- **Content-based Image Retrieval**
  - Increased participation: 12 groups submitting 52 purely visual runs (compared to 3 groups submitting only 12 purely visual runs in 2006)
  - 53% of all retrieval approaches included CBIR (31% in 2006)
  - Retrieval results (MAP) 66% higher compared to 2006

Medical retrieval – news in 2007

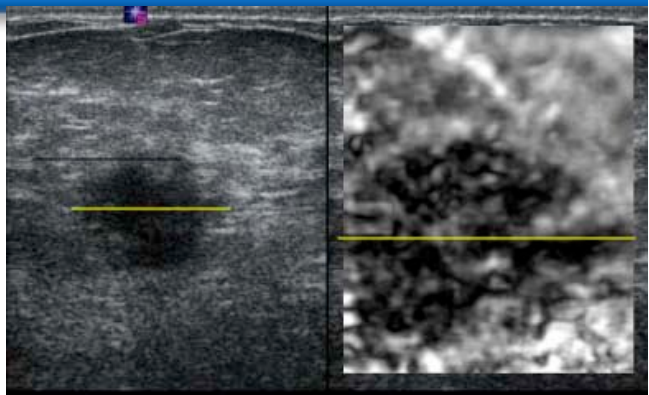
- Larger data set with almost 70,000 images
- Topics generated from medline queries
  - Medical literature database
  - Frequent queries with link to visual content
    - Automatic filtering and manual work
- 38 registrations
- 13 groups submitted results
- Relevance judgments paid by National Science Foundation (NSF) grant

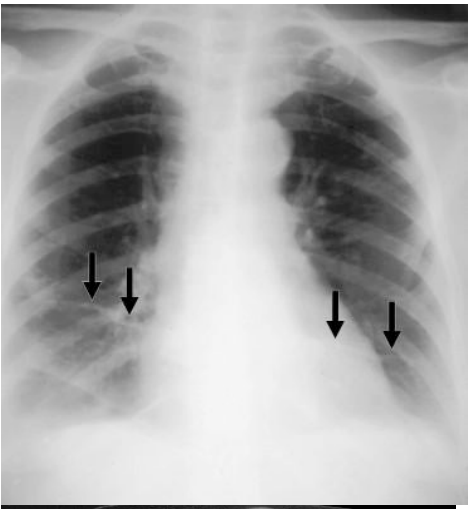
Databases used

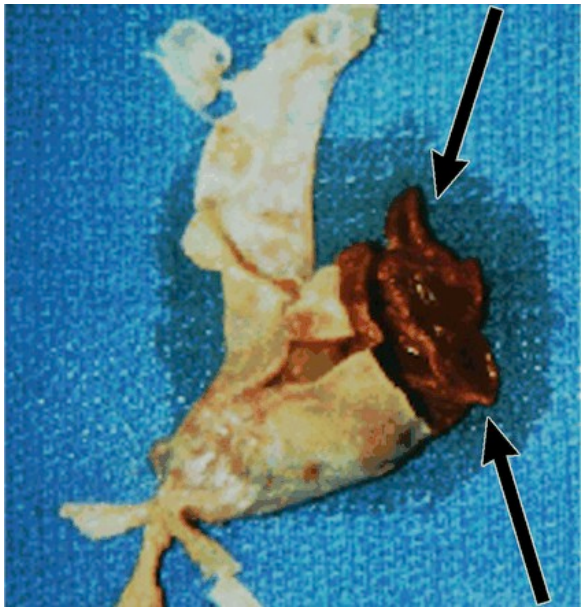
Collection	Predominant images	Cases	Images	Annotations	Size
Casimage	Mixed	2076	8725	English – 177 French – 1899	1.3 GB
Mallinckrodt Institute of Radiology (MIR)	Nuclear medicine	407	1177	English – 407	63 MB
Pathology Education Instructional Resource (PEIR)	Pathology	32319	32319	English – 32319	2.5 GB
PathoPIC	Pathology	7805	7805	German – 7805 English – 7805	879 MB
myPACS.net	Radiology	3577	15140	English - 3577	390 MB
CORI	Endoscopic	1496	1496	English - 1496	34 MB
Total		47680	66662	55485	5.2 GB

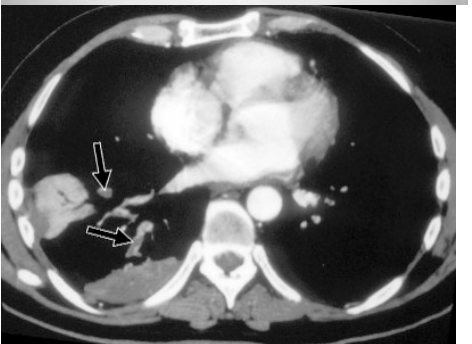
Example topics

Ultrasound with rectangular sensor.
Ultraschallbild mit rechteckigem Sensor.
Ultrason avec capteur rectangulaire.









Pulmonary embolism all modalities.
Lungenembolie alle Modalitäten.
Embolie pulmonaire, toutes les formes.

Participants in 2007

- CINDI, Concordia University, Montreal, Canada
- Dokuz Eylul University, Izmir, Turkey
- IPAL/CNRS joint lab, Singapore, Singapore
- IRIT-Toulouse, Toulouse, France
- MedGIFT, University and Hospitals of Geneva, Switzerland
- Microsoft Research Asia, Beijing, China
- MIRACLE, Spanish University Consortium, Madrid, Spain
- MRIM-LIG, Grenoble, France
- OHSU, Oregon Health & Science University, Portland, OR, USA
- RWTH Aachen Pattern Recognition, Aachen, Germany
- SINAI, University of Jaen Intelligent Systems, Jaen, Spain
- State University New York (SUNY) at Buffalo, NY, USA
- UNAL, Universidad Nacional Colombia, Bogota, Colombia

Runs submitted by category

	Visual	Textual	Mixed
Automatic	27	39	80
Manual	0	1	0
Feedback	1	0	1

Visual Results

Run	Relevant	MAP	R-prec	P10	P30	P100
RWTH-FIRE-ME-NT-tr0506	1613	0.2328	0.2701	0.4867	0.4333	0.2823
CINDI_IMG_FUSION	630	0.0333	0.0532	0.1267	0.1222	0.0777
RWTH-FIRE-NT-emp2	562	0.0280	0.0493	0.1067	0.0811	0.0587
miracleVisG	532	0.0186	0.0396	0.0833	0.0833	0.0470
UNALCO-nni_FeatComb	644	0.0082	0.0149	0.0200	0.0144	0.0143
miracleVisGFANDmin	165	0.0081	0.0225	0.0367	0.0478	0.0333
UNALCO-svmRBF_Tamura	375	0.0048	0.0109	0.0067	0.0100	0.0100
GE_4_8.treceval	292	0.0041	0.0192	0.0400	0.0322	0.0203
GE-GE_GIFT4	290	0.0040	0.0192	0.0400	0.0322	0.0203
DEU_CS-DEU_R2	277	0.0028	0.0052	0.0067	0.0022	0.0033
DEU_CS-DEU_R5	249	0.0014	0.0062	0.0000	0.0078	0.0077

Textual results

Run	Relevant	MAP	R-prec	P10	P30	P100
LIG-MRIM-LIG_MU_A	2347	0.3962	0.4146	0.5067	0.4600	0.3593
SinaiC100T100	2449	0.3668	0.3942	0.5467	0.5044	0.3457
LIG-MRIM-LIG_MU_L	2363	0.3643	0.3784	0.5033	0.4422	0.3183
miracleTxtENN	2294	0.3518	0.3890	0.5800	0.4556	0.3600
OHSU_as_out_1000rev1_c	2306	0.3453	0.3842	0.5300	0.4433	0.3033
UB-NLM-UBTextBL1	2244	0.3182	0.3306	0.5300	0.4756	0.3190
OHSU-OHSU_txt_exp2	1433	0.3135	0.3775	0.5867	0.4878	0.2893
IPAL-IPAL1_TXT_BAY_ISA0	1895	0.3057	0.3320	0.4767	0.4044	0.3163
SinaiC020T100	2028	0.2950	0.3138	0.4400	0.4389	0.2980
IPAL-IPAL4_TXT_BAY_ISA0	1831	0.2935	0.3177	0.4733	0.3978	0.3073
GE_EN	2170	0.2714	0.2989	0.3900	0.3356	0.2467
DEU_CS-DEU_R1	891	0.1694	0.2191	0.3967	0.3622	0.2533
GE_FR	1306	0.1557	0.1781	0.1933	0.2067	0.1520
UB-NLM-UBTextFR	1503	0.1184	0.1336	0.2033	0.1767	0.1320
miracleTxtDEN	724	0.0932	0.1096	0.1800	0.1356	0.0970
IRIT_RunMed1	1418	0.0660	0.0996	0.0833	0.1100	0.1023

Mixed media results

Run	Relevant	MAP	R–prec	P10	P30	P100
SinaiC100T80	2433	0.3719	0.4050	0.5667	0.5122	0.3517
ohsu_m2_rev1_c	2164	0.3461	0.3892	0.5567	0.4622	0.3287
UB-NLM-UBTI_1	2237	0.3230	0.3443	0.5167	0.4911	0.3317
RWTH-FIRE-ME-tr0506	1920	0.3044	0.3409	0.5267	0.4644	0.3410
miracleMixGENTRIGHTmin	2002	0.2740	0.2876	0.4500	0.3822	0.2697
RWTH-FIRE-emp	1809	0.2457	0.3123	0.4567	0.4467	0.3020
GE_VT1_4	2123	0.2425	0.2596	0.3533	0.3133	0.2253
SinaiC020T50	1973	0.2148	0.2500	0.4033	0.3422	0.2403
GE_VT10_8	1407	0.1937	0.2247	0.3600	0.3133	0.2157
CINDI_TXT_IMAGE_LINEAR	1053	0.1659	0.2196	0.3867	0.3300	0.2270
UB-NLM-UBmixedFR	1308	0.1201	0.1607	0.2100	0.2022	0.1567
OHSU-oshu_c_e_f_q	598	0.1129	0.1307	0.2000	0.1544	0.0837
ohsu_fire_ef_wt2_rev1_c	542	0.0586	0.0914	0.2000	0.1211	0.0760
7gift-3ohsu	1629	0.0181	0.0060	0.0033	0.0044	0.0073
miracleMixGFANDminENTLEFTmax	165	0.0081	0.0225	0.0367	0.0478	0.0333

Conclusions and announcements

- For the first time purely textual retrieval had the best overall run
  - But purely visual retrieval with learning was extremely good as well
- After having found inconsistencies in the judgements we are redoing some topics
  - Results should be out within 2-3 weeks
- Topics of 2005-2007 are combined for one large collection (new judgements are done)

Medical Image Annotation Task

- Purely visual task
- Given an image, find a textual description
- 2005:
  - 9,000 training images/1,000 test images
  - Assign one out of 57 possible labels to each image
- 2006:
  - 10,000 training images/1,000 test images
  - Assign one out of 116 possible labels to each image
- **2007:**
  - **11,000 training images/1,000 test images**
  - **Assign a textual label to each image**

Example of IRMA code

- Example: 1121-127-720-500

Technique: radiography, plain, analog, overview

Direction: coronal, AP, supine

Anatomy: abdomen, middle, unspec.

Biosystem: uropoetic system, unspec., unspec.

Aim:

Predict complete code
- as far as possible
- correctly

Evaluation Criterion

- incomplete codes 11__-12_-7__-5__
- not predicting a position: better than a wrong prediction
- incorrect prediction in one position invalidates all later predictions in this axis
- axes are independent
- early errors are worse than late

Examples

(for one axis): correct 318a

318a	0
318*	0.06
3187	0.12
31**	0.14
32**	0.52
8988	1.00

Example Images

11,000 train images

1,000 test images

116 complete labels

1121-120-200-700

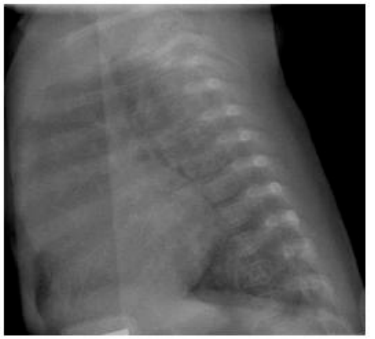
T: x-ray, plain radiography, analog, overview image
D: coronal, anteroposterior (AP, coronal), unspecified
A: cranium, unspecified, unspecified
B: musculosceletal system, unspecified, unspecified

1121-120-310-700

T: x-ray, plain radiography, analog, overview image
D: coronal, anteroposterior (AP, coronal), unspecified
A: spine, cervical spine, unspecified
B: musculosceletal system, unspecified, unspecified

1121-127-700-500

T: x-ray, plain radiography, analog, overview image
D: coronal, anteroposterior (AP, coronal), supine
A: abdomen, unspecified, unspecified
B: uropoietic system, unspecified, unspecified



1123-211-500-000

T: x-ray, plain radiography, analog, high beam energy
D: sagittal, lateral, right-left, inspiration
A: chest, unspecified, unspecified
B: unspecified, unspecified, unspecified

Participants

- Groups
  - 10 participations
- Runs:
  - In total 68 submitted
- Several groups participating the second or third time

- BIOMOD, Liege, Belgium
- **IDIAP, Martigny, Switzerland**
- medGIFT, Geneva Switzerland
- **CYU, Jung-Li, Taiwan**
- MIRACLE, Madrid, Spain
- OHSU, Portland, OR, USA
- RWTH from Aachen, Germany
- IRMA from Aachen Germany
- UFR from Freiburg, Germany
- DBIS from Basel, Switzerland

Results

Rank	Group	submissions	Best Score	Worst Score
1	BLOOM/IDIAP	7	26.8	72.4
6	RWTH	6	30.9	44.6
7	UFR	4	31.4	48.4
17	IRMA	3	51.3	80.5
19	UNIBAS	14	58.1	65.1
26	OHSU	2	67.8	68.0
30	BIOMOD	4	73.8	78.7
33	CYU	1	79.3	79.3
36	Miracle	30	158.8	505.6
63	medGIFT	3	375.7	391.0

Analysis of the Results

- performance of systems improved since last year:
  - the best system of last year is rank 10 this year
- large variety in submitted methods
  - image retrieval approaches
  - discriminative classification approaches
- large variety in used features
  - local features
  - global features
- only few groups used the hierarchy

CALL FOR PAPERS

Medical Image Annotation of ImageCLEF 2007

SPECIAL ISSUE IN PATTERN RECOGNITION LETTERS

All participating groups are encouraged to submit. Details will be announced.

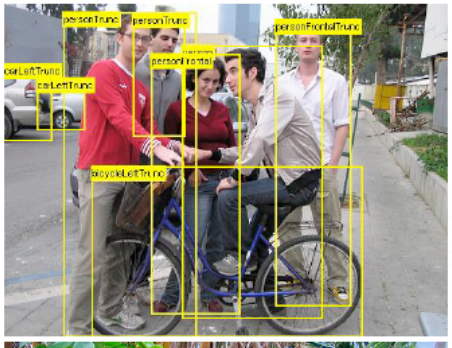
Object Retrieval Task

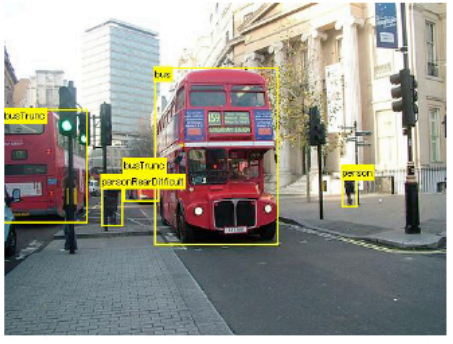
Find images showing

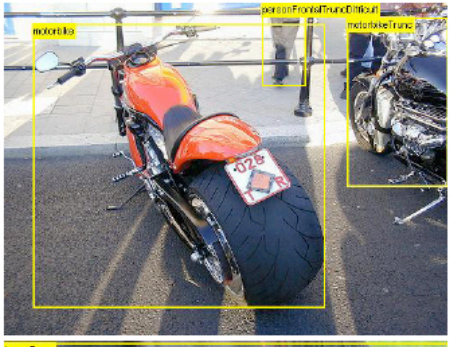
- Bicycles
- Busses
- Cars
- Motorbikes
- Cats
- Cows
- Dogs
- Horses
- Sheep
- Persons

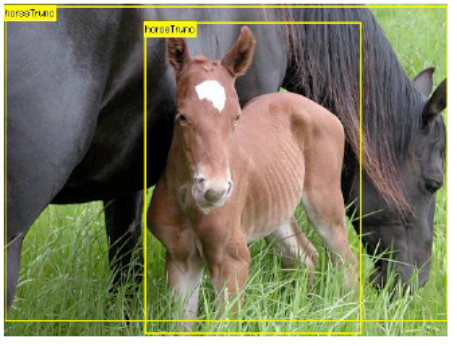
Using only the image, no text

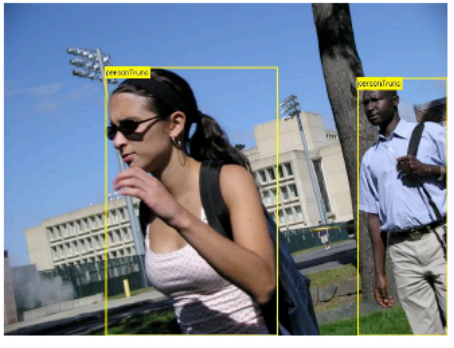
Example Images: Training Data





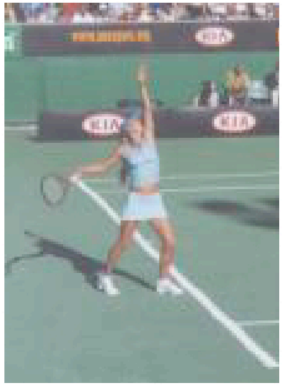


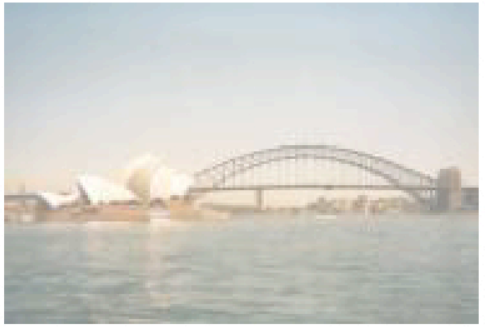


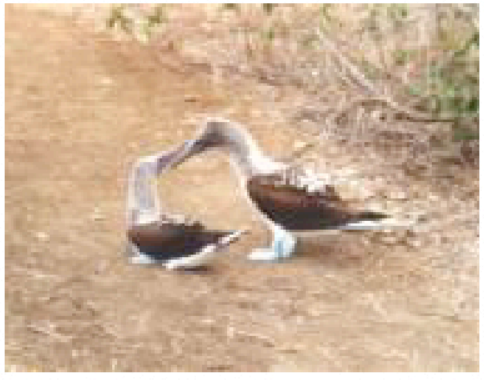


2,600 images, fully annotated

Example Images: Test Data








20,000 images

Participants

- groups
  - 7 participations
- runs:
  - in total 26 submitted

- Hung. Acad. Of Sciences, Budapest, Hungary
- Adaptive Informatics Research Centers, Helsinki, Finland
- INAOE, TIA Research Group, Tonantzintla, Mexico
- Microsoft Research, Asia, Beijing, China
- NTU: Nanyang Technological University, Singapore
- PRIP: Vienna University of Technology, Vienna, Austria
- RWTH Aachen University, Aachen, Germany

Results of the annotation process

Class	Rel. in pools	Rel. in ext. pools	Rel. in db.
Bicycle	66	254	655
Bus	23	69	218
Car	200	522	1268
Motorbike	7	28	86
Cat	5	18	7
Cow	7	23	49
Dog	9	22	72
Horse	13	94	175
Sheep	5	42	6
Person	554	3939	11248

Results of the Evaluation

- Bicycle:
  - Normal pools: HUT has best performance (map=21.3/next=13.0)
  - Extended: Budapest has best performance (9.1/7.2)
  - Full: Budapest clearly outperforms (28.3/4.1)
- Bus
  - Normal pools: RWTH has best performance (2.7/1.5)
  - Extended/full: HUT
- Car
  - HUT has best performance
- Motorbike
  - Normal: MSRA (3.5/1.5)
  - Extended/full: Budapest (6.2/3.8)/(18.5/1.8)
- Cat
  - Vienna (2.6/1.1)

Results of the Evaluation (cont'd.)

- Cow
  - HUT (1.5/1.0)
- Dog
  - HUT/PRIP (<1.0)
- Horse
  - HUT slightly better than others
- Sheep
  - Normal: HUT (20/3)
  - Extended full: HUT only slightly better than others
- Person
  - HUT, MSRA, RWTH

Interpretation of the Results

- HUT had very many runs (approx 50%) of the submissions
- Bias of the pools (favorable/unfavorable) for HUT
- Sheep/cat have so few relevant images that results do not tell much
- More images contain persons than were allowed to deliver
- for some of the queries the results are quite nice (in particular person)
- Mismatch of training/testing data is still a serious issue

Highlights of ImageCLEF 2007

- Photographic Retrieval
  - More visual runs
  - Limited annotation did not affect retrieval success
- Medical Retrieval
  - Purely visual/textual retrieval is very good
  - Combination not yet solved
- Medical Annotation
  - Hierarchy does not help
- Object Retrieval
  - Mismatch between training and test set is challenging

Breakout Session/Outlook

- Several Ideas for next year!
- What do you expect?
- What are our ideas?
- What data is available?

- **Breakout Session**:
  - Friday 11:00-12:00h