Overview of ImageCLEF 2016

Mauricio Villegas
(on behalf of all organisers)

ImageCLEF history

- Started in 2003 with a photo retrieval task
  - 4 participants submitting results
- In 2009 we had 6 tasks and 65 participants
- Previous editions, 2014 and 2015
  - 4 tasks, 13 and 25 working notes, respectively
- Current edition, 2016
  - 3 tasks: medical figures, web images, scanned handwriting
  - 98 registrations, 54 signed EUA
  - 19 participants, 15 working notes papers

ImageCLEF objectives

- Automatic image annotation and retrieval for a wide range of source images and annotation objectives
  - General imagery but also domain specific
- Annotating to allow language-independent indexing and retrieval from image collections
- Use of multimodal information: textual, visual, 3D, etc.
- Evaluation of machine learning and natural language processing approaches

Tasks organised in 2016

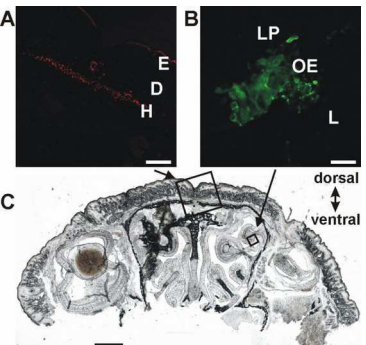
- Medical - ImageCLEFmed (12th edition)
  - Figures from biomedical literature
    - Identification, multi-label classification and separation of compound figures
    - Caption prediction (new subtask)
- Scalable Concept Image Annotation (5th edition)
  - Concept detection and localization in web images
  - Generation of sentence descriptions of the content
  - Illustrating text with images
- Handwritten Scanned Document Retrieval (1st edition)
  - Retrieval of image segments for multi-word queries

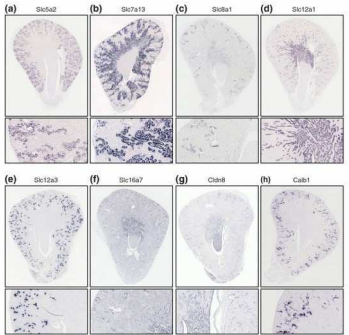
The Medical Task

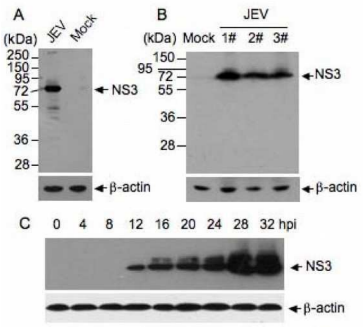
ImageCLEFmed

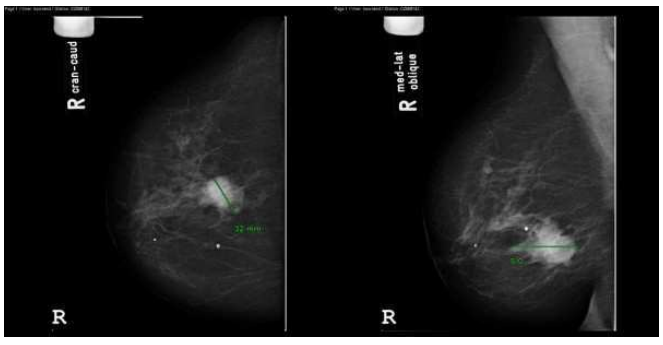
Compound figure separation and multi-label classification; figure caption prediction

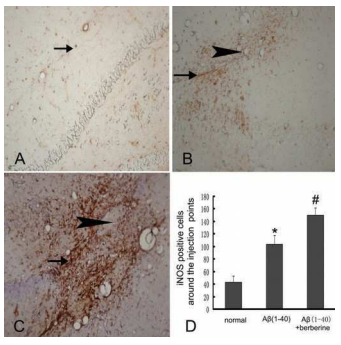
(12th edition)





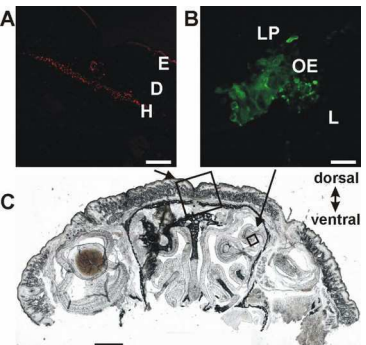


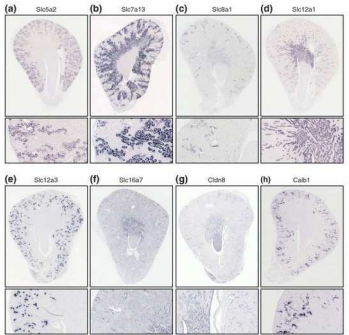


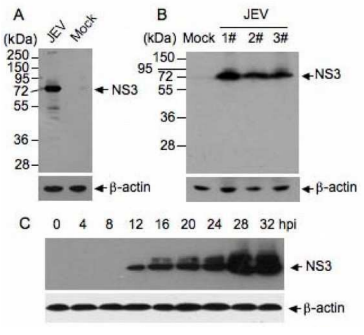


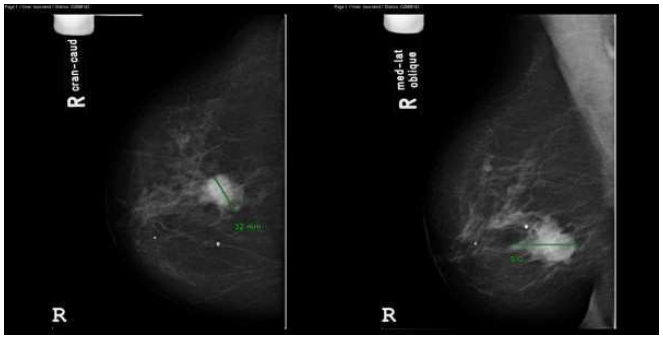
Motivation and aim

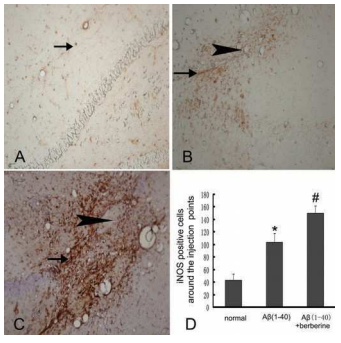
- Motivation:
  - About 40% of figures in PubMed Central are compound figures
  - In digital articles often compound figures are available as a single image block without information of its parts
  - The compound figures need to be handled by image indexing and retrieval systems
- Aim:
  - Develop techniques to handle compound figures and predict figure captions in biomedical literature retrieval







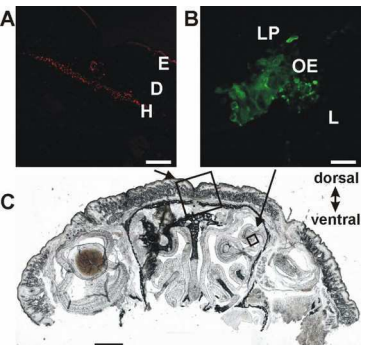


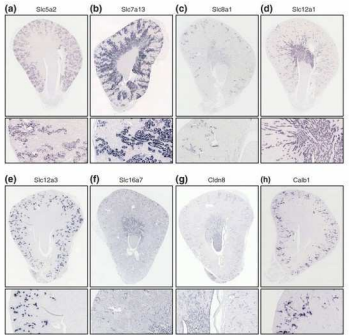


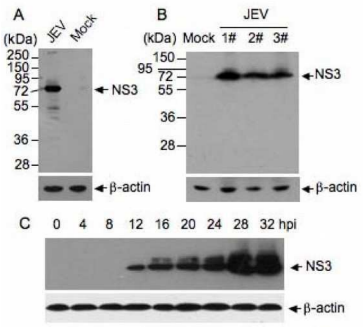
Data and task

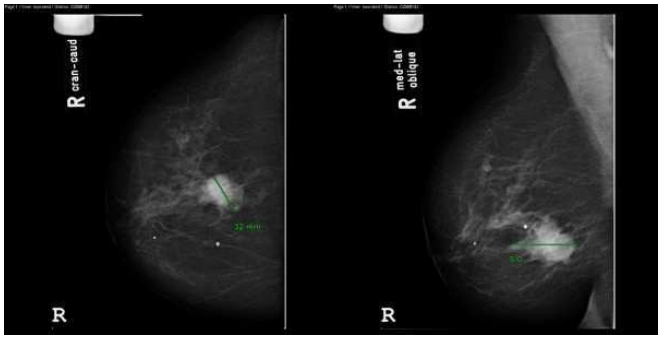
- Tasks:
  - Compound figure detection (26,456 compound and single figures)
  - Compound figure separation (8,397 compound figures)
  - Multi-label classification (2,675 compound figures)
  - Subfigure classification (10,942 sub-figures)
  - Caption prediction (20,000 non-compound diagnostic figures)

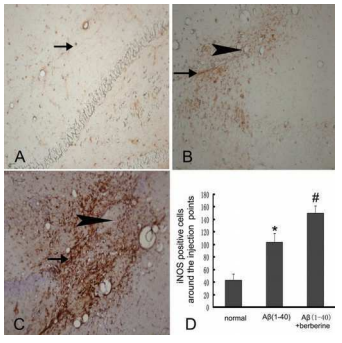
- Dataset:
  - Subset of the of ImageCLEF 2013 dataset, part of PubMed Central







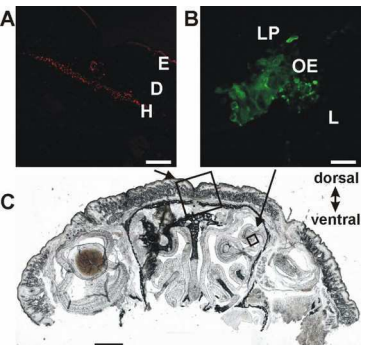


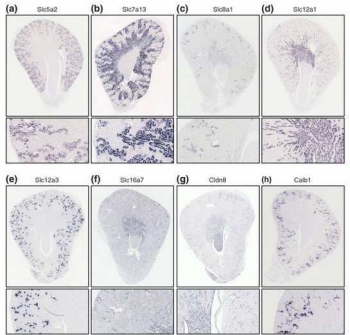


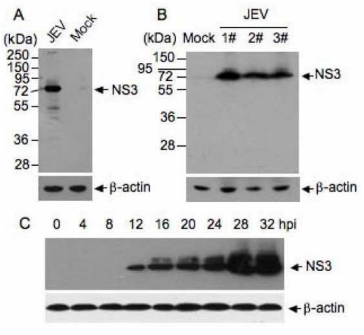
General information

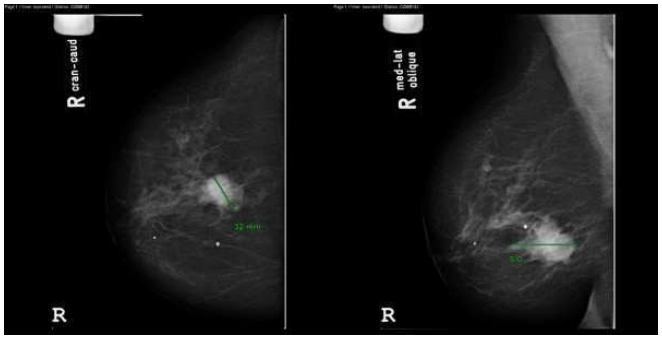
- Past editions:
  - Current format since 2015
  - Compound figure separation subtask in 2013
  - Figure classification subtask since 2010

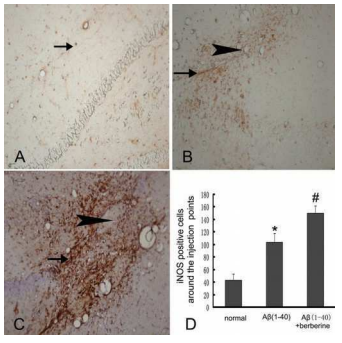
- Participation:
  - Eight groups participated, 69 system runs in total
  - No participants in caption prediction task





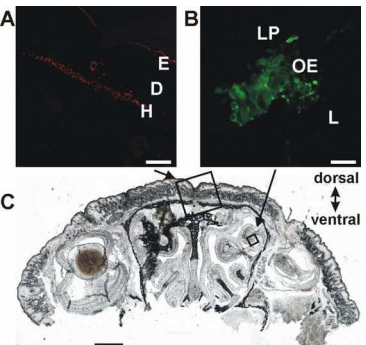


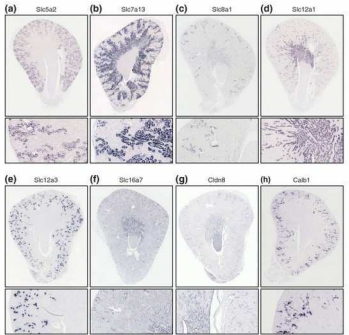


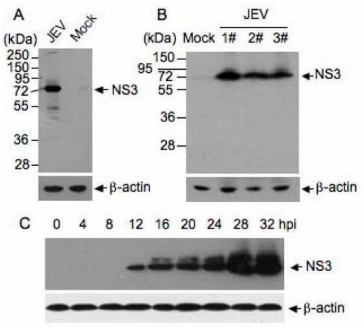


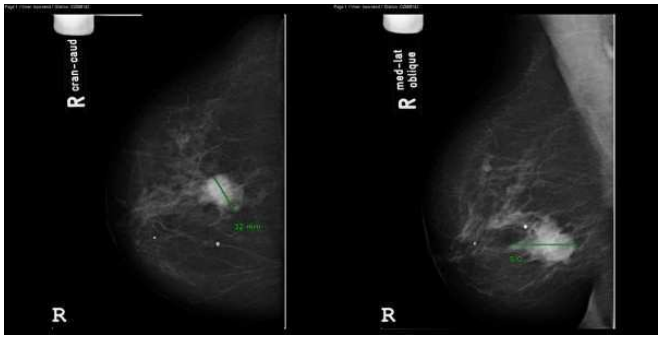
Results and observations

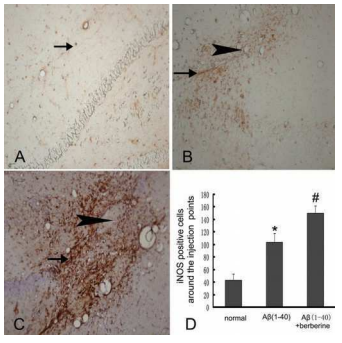
- Around 90% accuracy in:
  - Compound figure detection
  - Subfigure classification
- 84% accuracy in compound figure separation
- 0.0131 Hamming loss for multi–label classification
- Prominent use of CNNs
- Connected component analysis for figure separation











Scalable Concept Image Annotation Task

Concept Annotation, Localization, Sentence Generation and Text Illustration

(5th edition)

Motivation and aim

- Motivation:
  - Users struggle with the ever-increasing quantity of data available to them
  - Large number of images cheaply found and gathered from the Internet
  - More valuable is mixed modality data, web pages of both images and text
- Aim:
  - To develop techniques to allow computers to reliably describe images, localise the different concepts depicted in the images, generate a description of the scene and retrieve relevant images, using noisy mixed modality data

Data and task

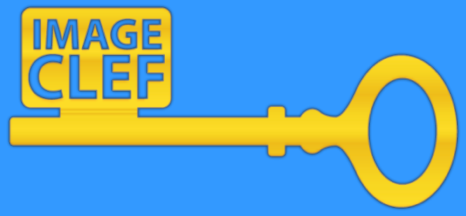
- 4 subtasks: single dataset of 510k webpages, images + text
  - Test set ⊂ training set (unknown to participants)
- Subtask 1: Image localisation/detection
  - For each of the 510k images, annotate+localise with 251 concepts
- Subtask 2: Natural language caption generation
  - Generate textual descriptions for all 510k images
- Subtask 3: Content selection
  - Given labelled bounding box instances for 450 test images as input, select the instances most likely to be mentioned by humans in a description

Data and task

- Teaser task: Text Illustration (new for 2016)
  - Given a piece of text, select the most relevant image to illustrate it (out of 200k)

- 510k dataset split into 310k for training and 200k for testing

Results and observations



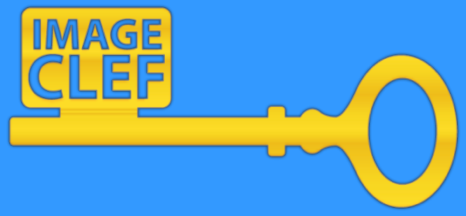
Subtask 1

- All approaches using exist CNN structure VGG
- Impressive performance
- Not using the provided web data much – too noisy?
- All image based features

Subtask 2

- Varied approaches
  - Deep learning
  - Template based

Results and observations



Subtask 3

- Teams focused on obtaining high recall
- Deep learning not used directly

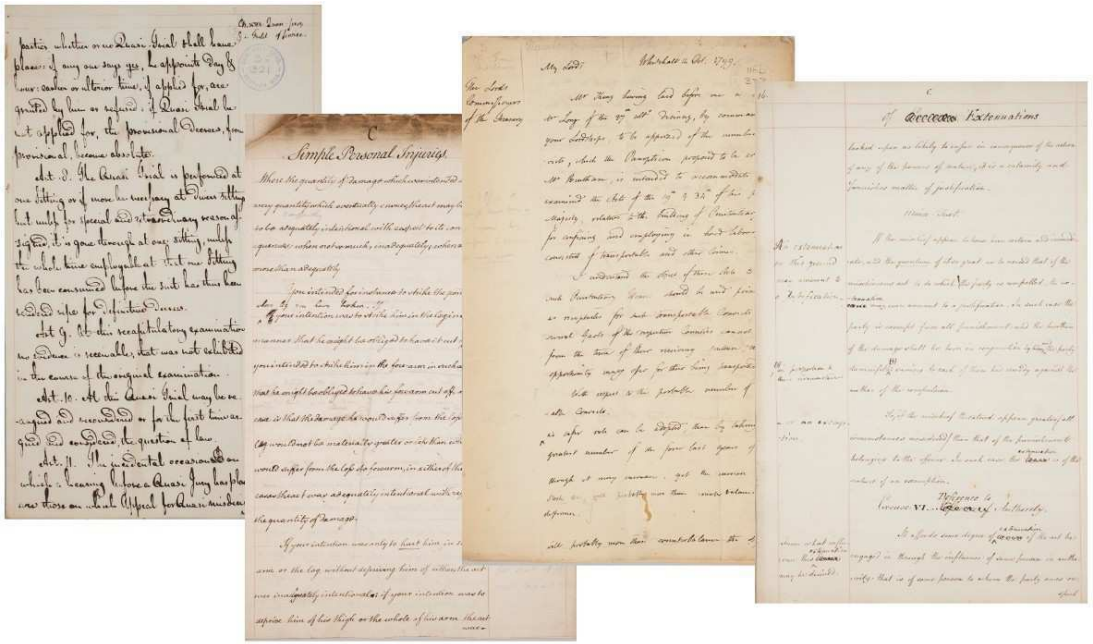
Teaser task

- IR queries
- Mapping onto common multimodal space

Handwritten Scanned Document Retrieval Task

Retrieval of image segments for multi-word queries

(1st edition)



Motivation and aim

- Motivation:
  - Currently there is a boom in digitization (scanning) of books, a large percentage being handwritten
  - Information access tools required to ease searching these handwritten collections
  - Automatic recognition of scanned handwriting is not as mature as for printed text
- Aim:
  - Develop indexing and retrieval techniques for handwritten documents that deal with its particular challenges

Data and task

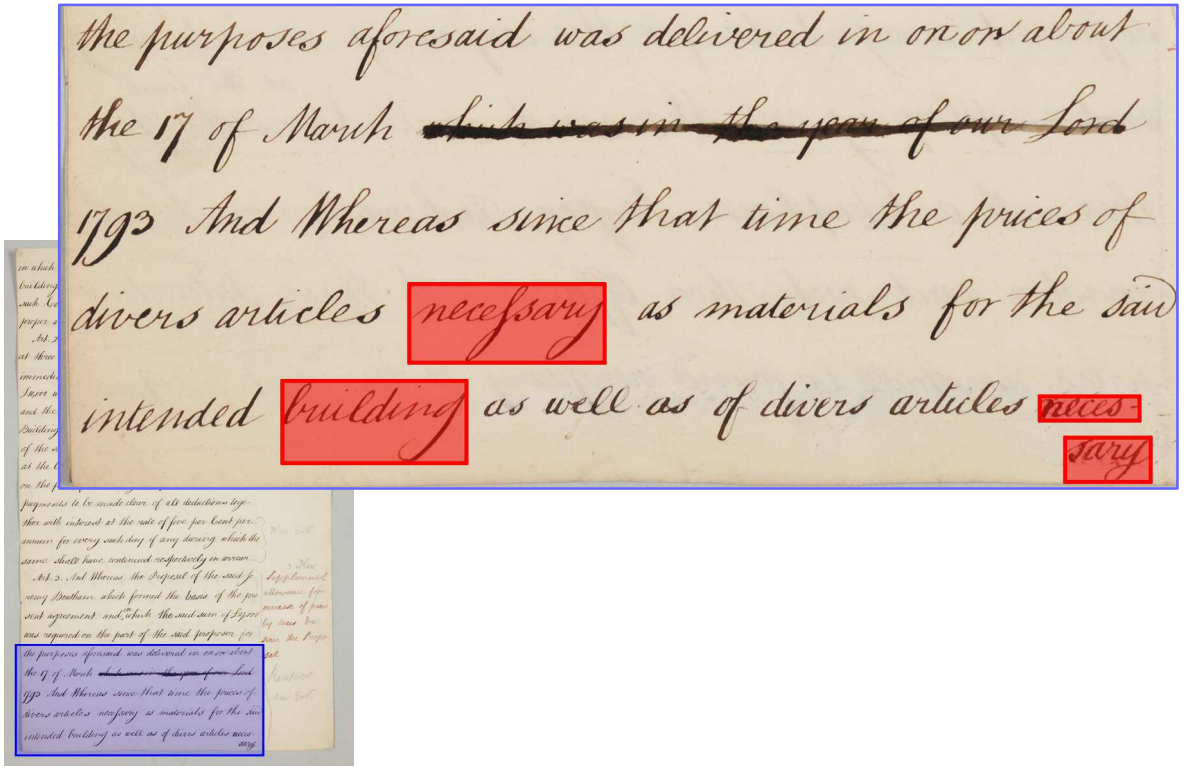
Dataset:

- Jeremy Bentham manuscripts (996 pages: 363 training, 433 development, 200 test)
- Provided page images, pre-extracted line images and 100-best recognitions using baseline system

Task:

- Retrieve 6-line segments for multi-word queries
  - Segments traverse pages
  - Broken words
  - Words unseen in training
  - Queries with zero results
  - Queries with repeated word

Query: building necessary



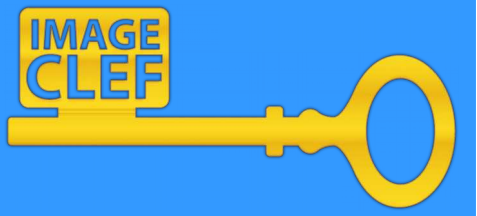
Results and observations

- Participation:
  - Four groups participated, 37 system runs in total
- Good performance obtained, in particular:
  - Words unseen in training
  - Words broken between lines
- Handling of broken words only based on the detection of hyphenation symbols

CLEF 2016 programme

Wednesday September 7	
13:30 – 14:50	Task overview presentations
14:50 – 15:30	Participant oral session 1
15:30 – 16:15	Participant poster session
16:15 – 17:35	Participant oral session 2
17:35 – 18:00	General discussion, feedback and future
Thursday September 8	
10:30 – 11:20	Best of Labs 2015: ImageCLEF

Organising Committee



Overall coordination:
Mauricio Villegas <mauvilsa@upv.es>
Henning Müller <henning.mueller@hevs.ch>

Technical support:
Ivan Eggel

Image Annotation:
Andrew Gilbert
Luca Piras
Josiah Wang
Fei Yan
Arnau Ramisa
Emmanuel Dellandrea
Robert Gaizauskas
Mauricio Villegas
Krystian Mikolajczyk

Medical:
Alba García Seco de Herrera
Roger Schaer
Stefano Bromuri
Henning Müller

Handwritten Retrieval:
Mauricio Villegas
Joan Puigcerver
Alejandro H. Toselli
Joan Andreu Sánchez
Enrique Vidal