Overview of ImageCLEF 2015

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, 2013 and 2014
  - 4 tasks, 11 and 13 working notes respectively

- Current edition, 2015
  - 4 tasks: 3 medical related and 1 with web images
  - 148 registrations, 72 signed EUA
  - 31 participants, 25 working notes papers

ImageCLEF objectives

- Automatic image annotation 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

Registration and submission system

CLEF 2015

logged in as: mauvilsa@upv.es My Account Logout

Home
Subtracks
Collections
Runs

Users
Pending Signatures

ImageCLEF-Website
CLEF-Website

New subtrack

All available subtracks

Subtrack name	Start date	End date		
ImageCLEF2015:image-annotation	Apr 25, 2015	May 5, 2015	Detail	Update
ImageCLEF2015:liver-ct-annotation	Apr 1, 2015	May 14, 2015	Detail	Update
ImageCLEF2015:medical-classification	Mar 15, 2015	May 1, 2015	Detail	Update
ImageCLEF2015:medical-clustering	Apr 20, 2015	May 5, 2015	Detail	Update
LifeCLEF2015:bird	Oct 1, 2015	May 17, 2015	Detail	Update
LifeCLEF2015:fish	Oct 1, 2015	May 23, 2015	Detail	Update
LifeCLEF2015:plant	Oct 1, 2015	May 16, 2015	Detail	Update

Lab Advertising

- Web:
  - ImageCLEF website: http://imageclef.org/2015
  - Twitter and facebook page

- Other Conferences:
  - Presentation at CVPR Language and Vision Workshop (Scalable Annotation task)

- Calls for participation:
  - We had 3 rounds: November, January and March
  - There have been general ImageCLEF calls and more targeted task specific calls
  - Printed versions of the calls have also been distributed at a few conferences

Tasks organised in 2015

- Scalable Concept Image Annotation (4th edition)
  - Concept detection and localization in web images
  - Generation of sentence descriptions of the content

- Medical classification (6th edition)
  - Identification, multi-label classification and separation of compound figures from biomedical literature

- Clustering of body part x-rays (1st edition)
  - Hierarchical clustering of x-rays containing mostly bone fractures from all over the body

- Liver CT Annotation (2nd edition)
  - Prediction of radiological annotations of liver CT images

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 and generate a description of the scene, using noisy mixed modality data

Data and task

- Single dataset of 500k webpages, images + text
  - Test set $\subset$ training set (unknown to participants)

- Subtask 1: Image localisation/detection
  - For each of the 500k images, annotate+localise with 251 concepts

- Subtasks 2 and 3
  - Noisy track: Generate sentence for all 500k images
  - Clean track: Given bounding boxes (with concept labels) for 450 test images as input, generate sentences

Results and observations

Subtask 1

- Nearly all approaches using CNN features

- Impressive performance

- Not using the provided web data much – too noisy?

- Generally mostly image based features

Subtask 2,3

- Varied approaches
  - Deep Learning
  - Joint Image-Text Retrieval
  - Template based

- Promising results across different approaches

Medical Classification task

Compound figure separation and multi-label classification

(6th edition)

Motivation and aim

- Motivation:
  - About 40% of figures in PubMed Central are compound figures
  - In the digital articles often these figures are available as a single block without information of its parts
  - The compound figures need to be handled by image indexing and retrieval systems

- Aim:
  - Develop techniques to reliably detect, separate and classify compound figures from the biomedical literature

Data and task

- ## Dataset:
  - 20,867 figures, subset of the of ImageCLEF 2013 dataset, part of PubMed Central

- ## Tasks:
  - Compound figure detection
  - Compound figure separation
  - Multi-label classification
  - Subfigure classification
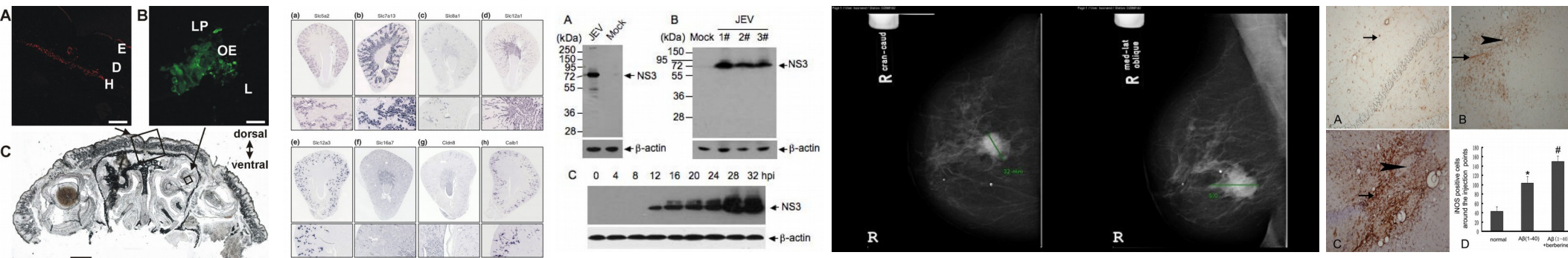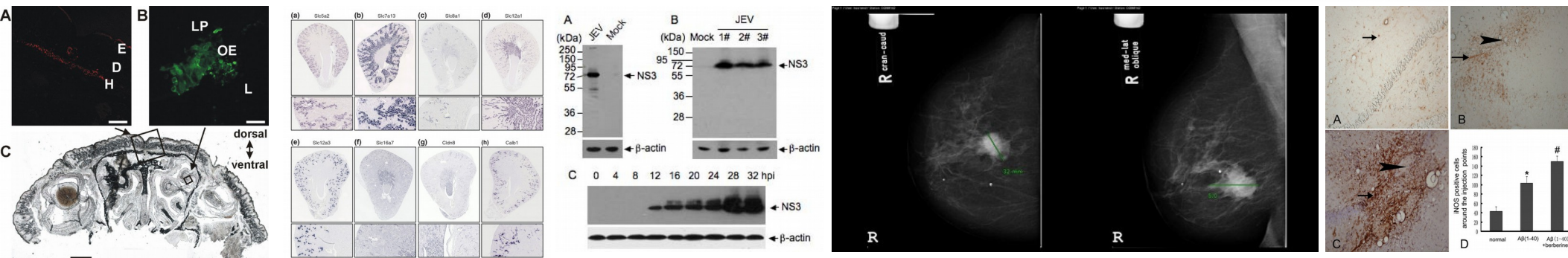
- More than 80% accuracy in:

  - Compound figure detection

  - Compound figure separation

- New approaches proposed for multi–label classification

- Mix approach achieves best results for subfigure classification
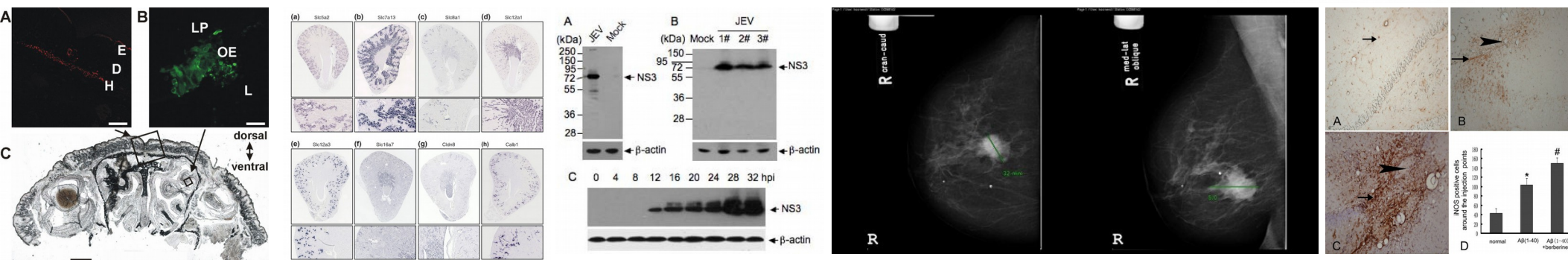
Medical Clustering task

Clustering of body part x-ray

(1st edition)

Motivation and aim

- Motivation:
  - Hospitals have large collections of x-rays available for teaching and research, but normally not annotated or searchable
  - Non-supervised approaches such as clustering can help to make sense out of such large collections

- Aim:
  - Develop tools to automatically cluster collections of x-ray images to make them searchable or browsable in several ways

Data and task

- Dataset:
  - 500 training and 250 development images in off-line DICOM format
  - X-ray images divided in 5 groups: Body, Head-Neck, Lower Limb, Upper Limb and True Negative
- Task:
  - Classify x-ray images into the 4 body part categories (multi-label classification)

Results and observations

- Since it was a novel task, only four categories were considered

- Very good participation (8 groups, 6 working notes), even though in its first edition

- Due to availability of general tools for feature extraction and classification, diverse approaches were presented

- Good performance, although the aim (for future editions) is to have a more fine grained clustering of the data

Motivation and aim

- ## Motivation:

  - Searchable collective medical knowledgebases enable experience sharing among the community for clinical and educational purposes

    - Empower comparative diagnosis
    - Assist medical students

- ## Aim:

  - Develop Content Based Case Retrieval systems focused on liver cases which allow querying with incomplete representation

- ## Dataset:

  - 50 training samples and 10 test samples (with incomplete radiology reports)

  - A cropped 3D CT image of the liver

  - ROI, which defines lesion area in the image

  - A set of 73 UsE features annotated using LiCO ontology

- ## Task:

  - Given a cropped CT of the liver and LiCO ontology, fill in a standardized radiology report composed of UsE features

Cluster size: 1
Segment: SegmentIII, SegmentIV
Lobe: Left lobe
Width: 60, Height: 64
Is gallbladder adjacent? False
Is peripheral localized: True
Is sub-capsular localized: False
Is central localized: False
Margin type: Lobular
Shape: Round
 Is contrasted: True   ...

Results and observations

- Unfortunately low participation, possibly due to inadequate advertising

- Nevertheless, the task is interesting and good results obtained

- Performance measured using completeness and accuracy

Year	Grp	Comp.	Acc.	Score	Method
2015	CREDOM	0.99	0.83	0.910	Image Retrieval
2014	BMET	0.98	0.91	0.947	Image Retrieval
2014	CASMIP	0.95	0.91	0.93	LDA + KNN
2014	piLabVaVlab	0.51	0.89	0.677	GCTF

CLEF 2015 programme

Wednesday September 9	
08:30 – 10:30	Tasks overview presentations
11:00 – 13:00	Participants oral session 1
16:00 – 16:30	Participants poster session
Thursday September 10	
13:30 – 14:30	Joint ImageCLEF / eHealth / BioASQ session Panel "Challenges and synergies in the evaluation of health IR/IE"
14:30 – 15:00	Participants oral session 2
15:00 – 15:30	Closing: Feedback and discussion on the future of 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
  Emmanuel Dellandrea
  Robert Gaizauskas
  Mauricio Villegas
  Krystian Mikolajczyk

Medical Classification:
  Alba García Seco de Herrera
  Stefano Bromuri
  Henning Müller

Liver CT Annotation:
  Burak Acar
  Suzan Uskudarli
  José F. Aldana
  María del Mar Roldán García

Medical Clustering:
  M. Ashraful Amin
  Mahmood Kazi Mohammed