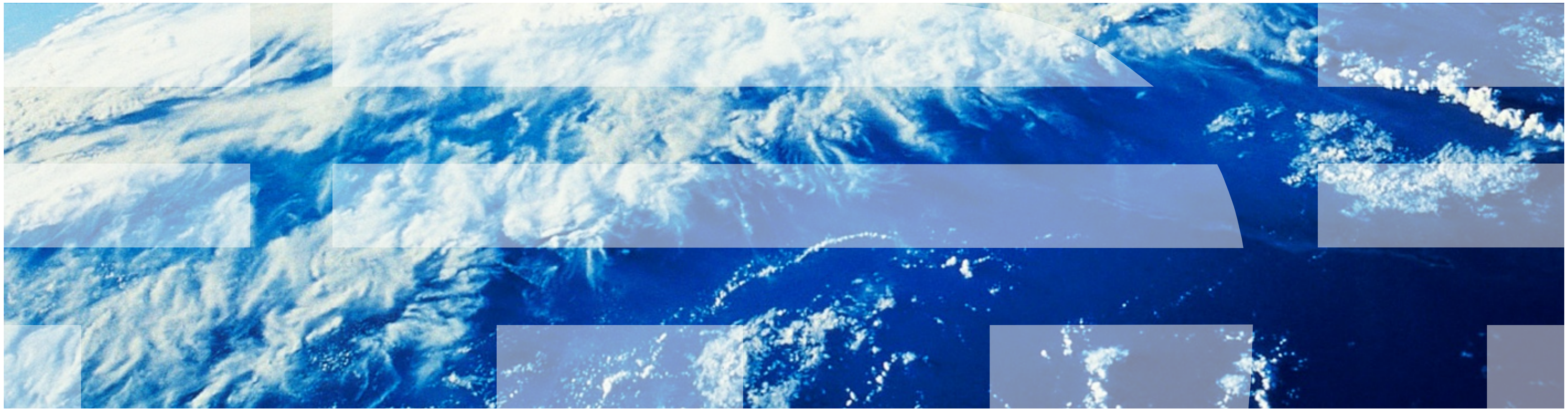
Mani Abedini and Rahil Garnavi - Australia

Amir Geva and Asaf Tzadok - Haifa

Liangliang Cao, Noel Codella, Jonathan H. Connell, **Michele Merler**, Quoc-Bao Nguyen, Sharathchandra U. Pankanti and John R. Smith - TJ Watson

IBM Multimedia Analytics @ ImageCLEF2013



http://www.imageclef.org/2013/medical

Overview

- IBM Multimedia Multi-Lab group @ ImageCLEF 2013
- Modality Classification task
  - Approaches
  - Results
- Case-based retrieval task
- Compound Image Segmentation Task
- Conclusions

IBM Multi-Lab Group @ ImageCLEF 2103

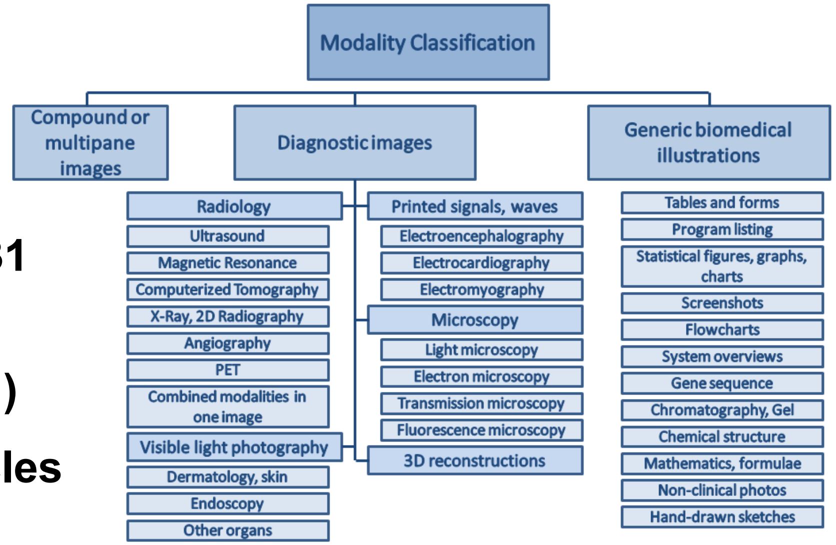
- **In 2013: multi-lab collaboration to solve the tasks**
  - **Australia and TJWatson on Modality Classificationn and Retrieval tasks**
  - **Haifa involved in Compound Figure Segmentation task**

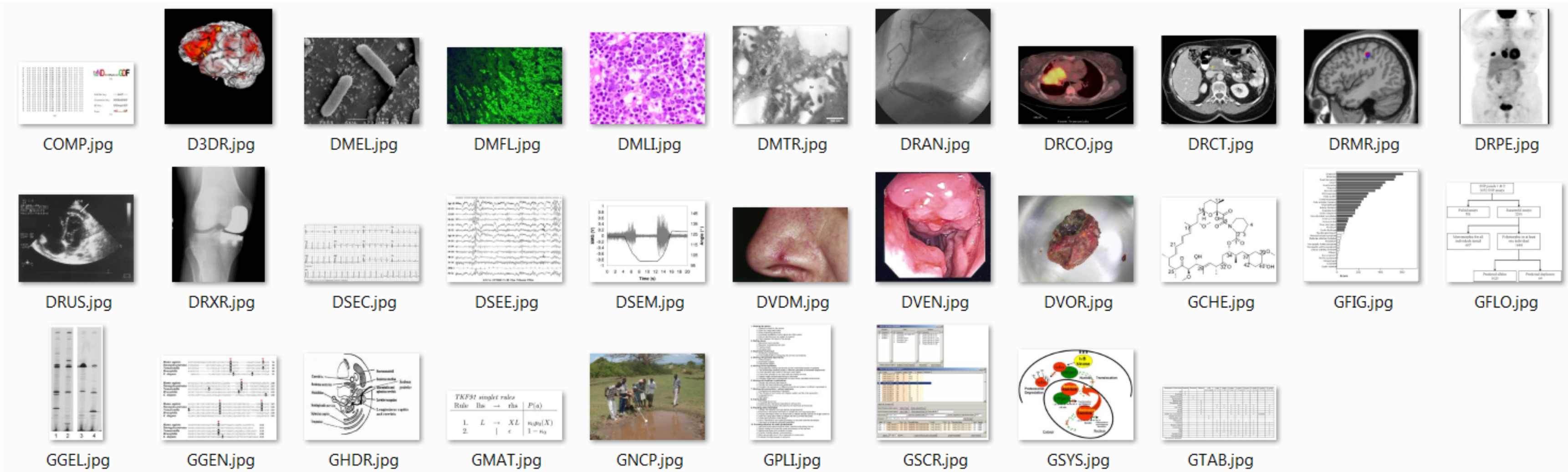
Overview

- IBM Multimedia Multi-Lab group @ ImageCLEF 2013
- Modality Classification task
  - Approaches
  - Results
- Case-based retrieval task
- Compound Image Segmentation Task
- Conclusions

ImageCLEF Medical Imaging Modality Classification Task

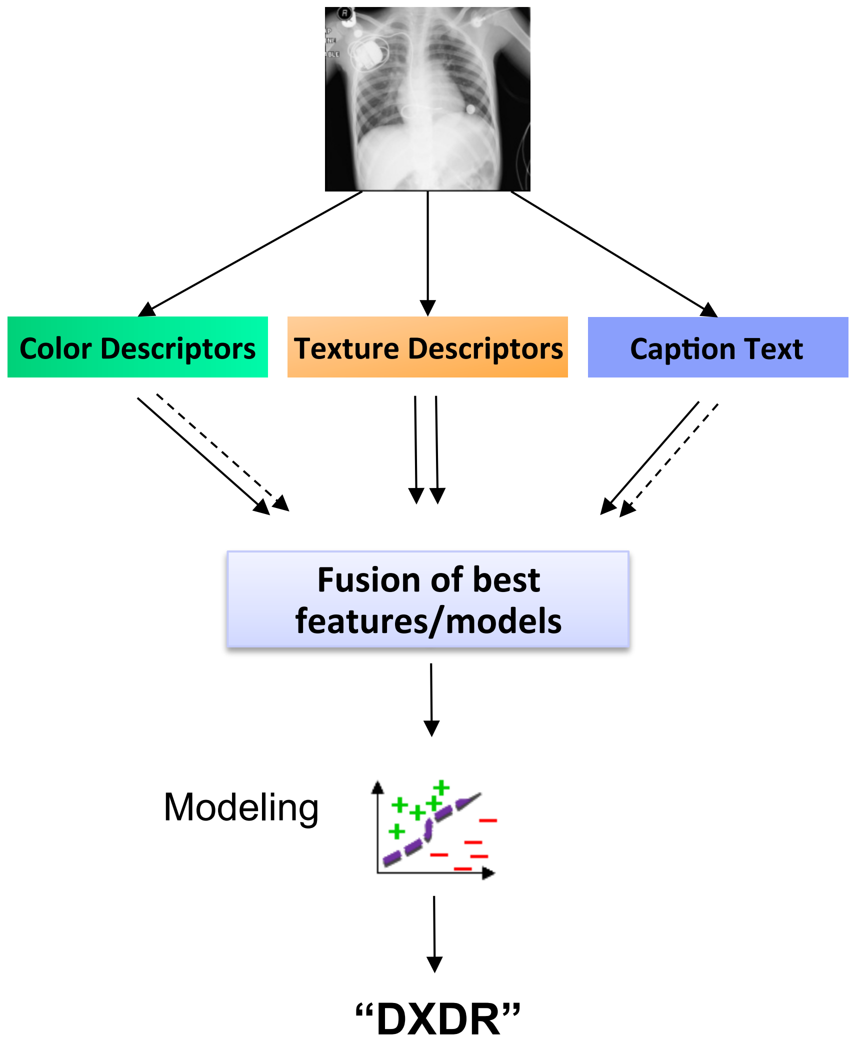
- **In user-studies, clinicians have indicated that modality is one of the most important filters that they would employ for search**
- **TASK: given an image, determine to which out of 31 medical and non-medical modalities it belongs**
  - **31 categories (x-ray, CT scan, ultrasound, etc.)**
  - **Images obtained from 300K real Pubmed articles**
  - **In 2013: 2,845 Training / 2,582 Test images**





Modality Classification Task – General Approach

- Extract **several** descriptors (features)
  - Visual (for texture, color and edges, at multiple granularities)
  - Textual (from captions, articles)
- **Selection** of best features based on held out (validation) set performance
- Learn multi-class image classifier on **fusion** of selected descriptors/ approaches



Modality Classification Task – Visual Descriptors

- Global descriptors
  - Color histogram
  - Color correlogram

  Color

  - Edge histogram

  Edge

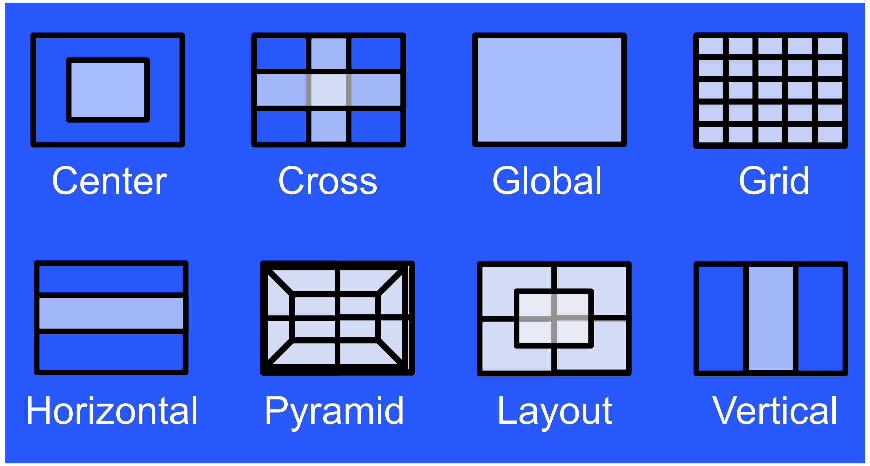
  - GIST
  - Curvelet Texture
  - Fourier Orientation
  - FourierPolarPyramid

  Fourier-texture

  - Thumbnail Vector
  - Image Type, Stats

  Global statistics

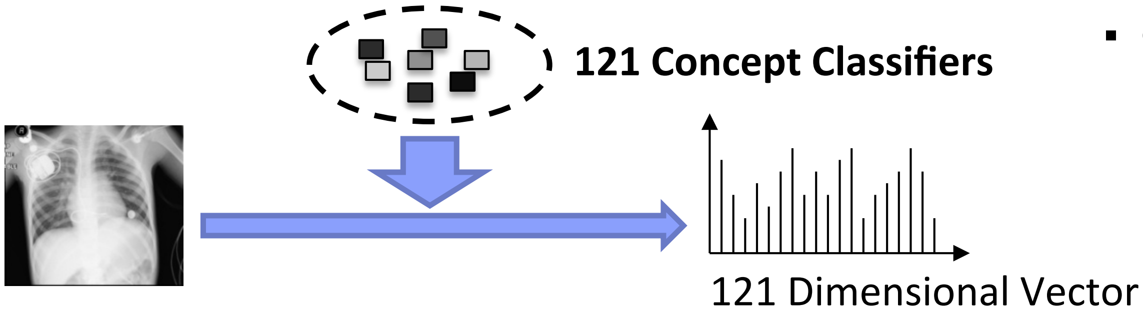
Granularities



- Local descriptors
  - LBP histogram : 58 uniform + 1 non-uniform codes
  - SIFT : different interest point detectors, Bag-of-Words codebooks+ soft assignment
  - Color SIFT (RGB-SIFT, HSV-SIFT, C-SIFT)

Modality Classification Task – Semantic Model Vector

- Set of 121 medical semantic concept classifiers constructed from training data collected from various sources (IRMA, TCIA, JSRT, Web Crawl)
- Classifiers trained using the IMARS learning framework
  - cover a range of radiological modalities, body regions, views, and some instances of disease pathology
- Classifiers responses concatenated into a 121 dimensional vector for each image

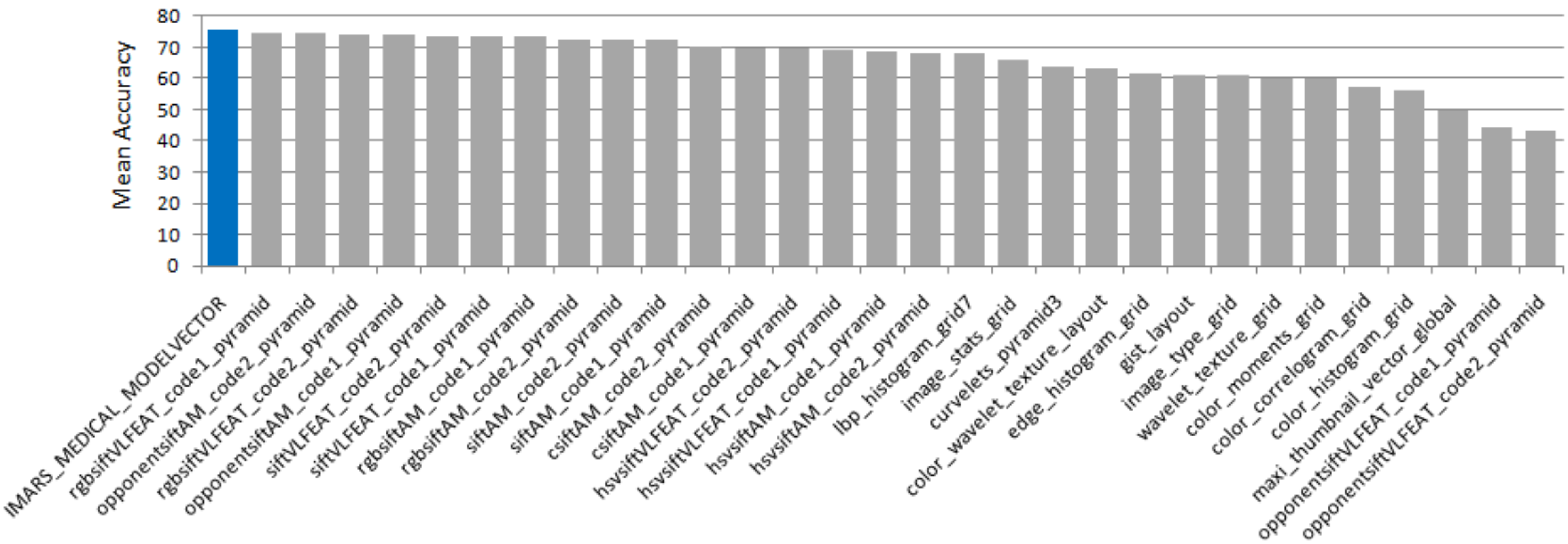


Training Datasets

- IRMA
  - X-Ray, Various Regions
  - 15,000 images
  - 193 categories (Modality, Organ, View)
- TCIA
  - 1,000,000+ images (30+ GB)
  - 17+ Categories (Modality, Body Region, View, Disease)
- JRST
  - X-Ray, Chest
  - 247 images, 154 lung cancer, 93 normal
- Cornell Datasets
  - CT, Chest
  - 25,000 images (11 GB)
- Web Crawl
  - 7,600 images
  - 49 categories (Modality, Organ, View, Disease)
- Cardiac Atlas (TBA)
  - Over 3,000 cases over decades.

Modality Classification Task – Visual Descriptors

- Mean Accuracy measured on official Test Set
- Medical Semantic Model Vector is the Best individual descriptor



Modality Classification Task – Textual Analysis

Modality Tailored Keywords

- Representation
  - Over 400 text patterns (full words, fragments of words, or multi-word phrases)
  - Vocabulary terms hand selected by perusing roughly half of captions in the training set
  - Between 2 and 51 patterns selected for each modality, then combined into one big feature list
  - Related phrases such as *fluorescent*, *immunofluorescence*, and *Alexafuor* merged to variabilized patterns such as *fluor*
  - Asterisks at the front and/or back match an arbitrary number of characters up to the first token delimiter
  - Patterns with all capital letters were only matched to text that was fully capitalized

- Modeling
  - The text-based classier built on top of this representation generates a likelihood score for each modality based on the presence or absence of a number of key words.
  - The number of hits (or an absence of a hit) for each term is weighted by a pseudo-probablistic model derived from the known modalities of the training examples.
  - Conditional probability of seeing a term given a particular modality is divided by that term's background probability.

Fragments of term list

- Pattern syntax
  - Can have variable (*) front and/or back but not middle
  - All capital term must be all capitals in text to match
- Complete list
  - Not segregated by modality (all lumped together)
  - Over 400 terms (best if no repeats)

COMP

```
each
panel*
plots
Images
f
```

DMFL

```
*fluor*
*flour*
immunostain*
spectral confocal micro*
```

DMLI

```
peripheral blood smear
dark field
HE
H&E
H & E
```

DRMR

```
MRI
magnetic resonance
T1*
gadolinium
```

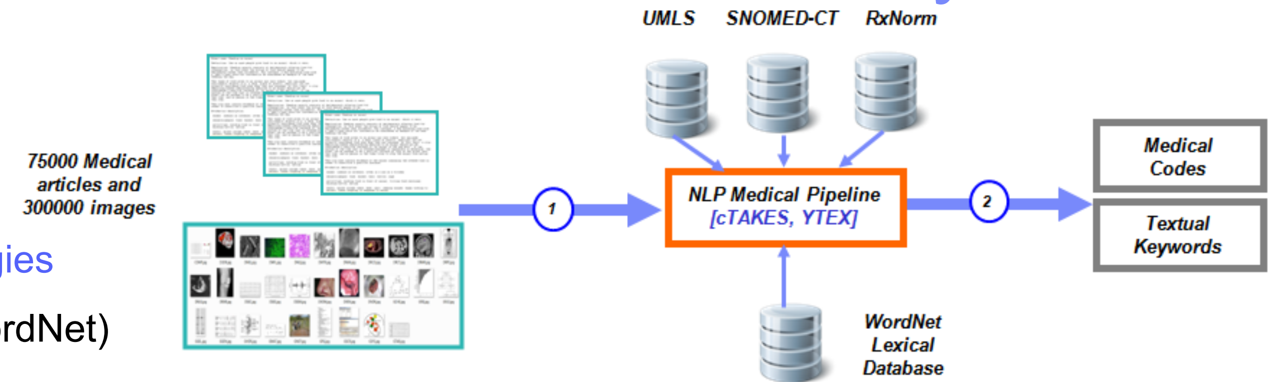
DVDM

```
skin
derm*
psori*
papul*
melanoma*
```

GGEN

```
*sequence*
align*
amino-acid*
*codon*
```

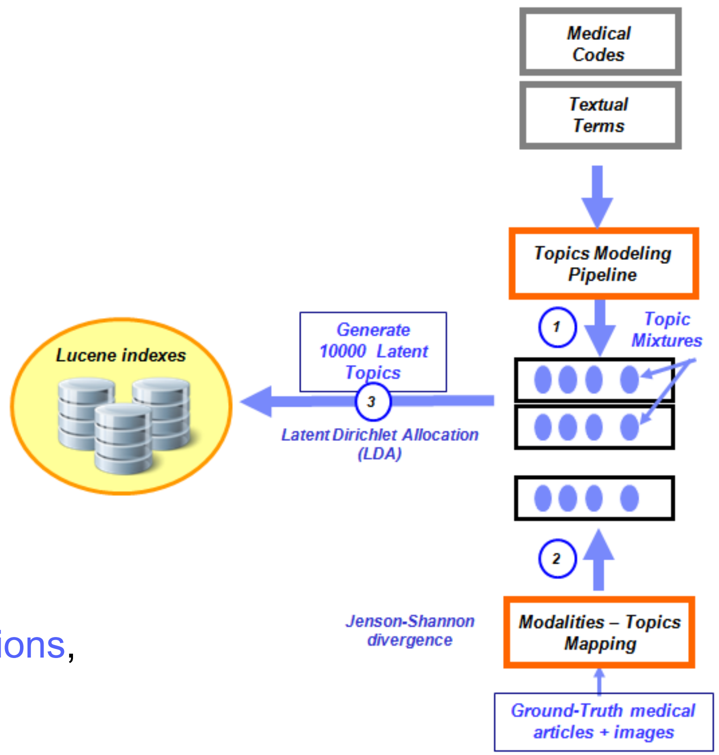
Modality Classification Task – Textual Analysis



Ontology Based Vocabulary

- Representation
  - Terms from two types of Ontologies
    - General lexical ontology (WordNet)
    - Medical specific domains medical knowledge-bases

- Modeling
  - NLP pipeline that consist of
    - WordNet lexical relations
    - Clinical Text Analysis and Knowledge Extraction System (cTAKES) and theYale cTAKES
  - Word-sense disambiguation and sliding window based part-of-speech to identify
    - relationships among words in the medical context
    - types of clinical named entities such as drugs, diseases,…
  - Lucene indexing on Articles Titles, Abstracts and Image Captions, TF-IDF weight
  - Modality classification based on modality *search*



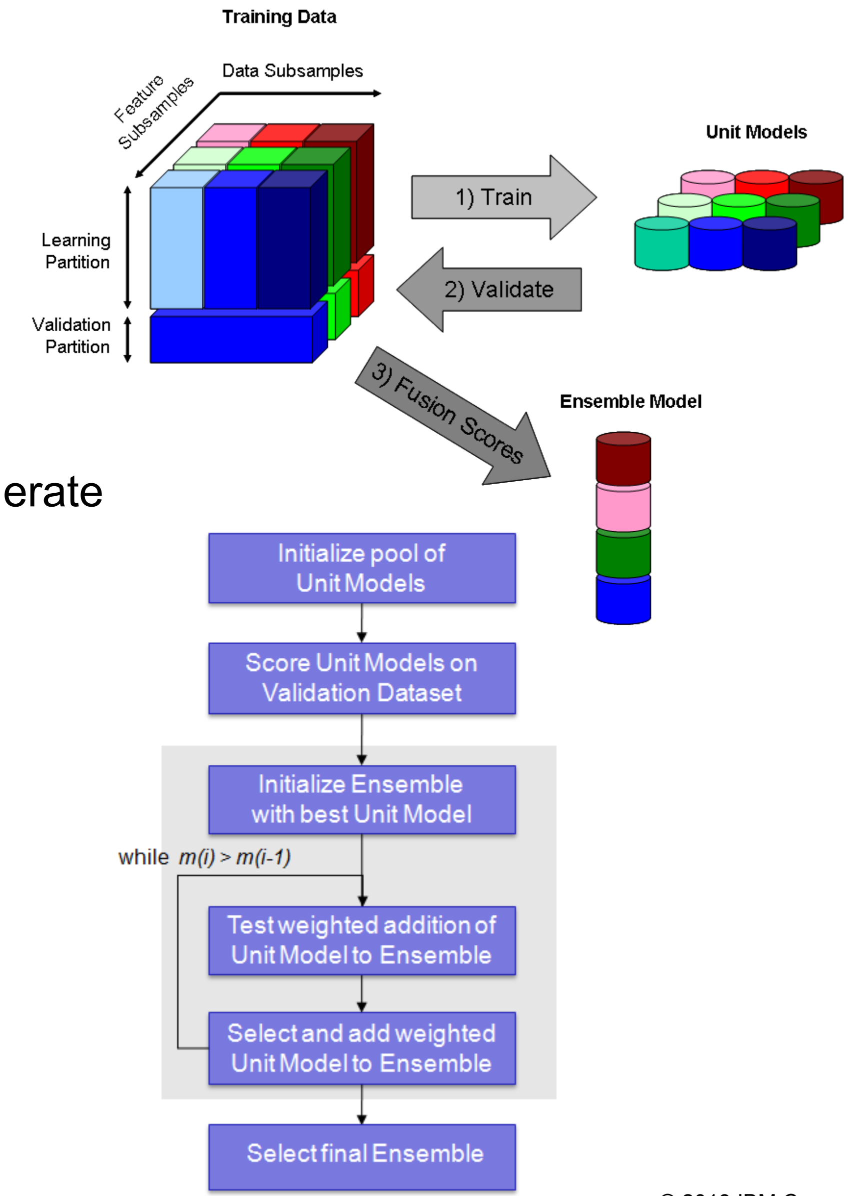
Modality Classification Task - Modeling and Fusion Strategies

- IMARS MODELING
- Two level SVM + Kernel Approximation
- Meta Classifiers
- Early (Kernel) and Late Fusion

Modality Classification Task – IMARS Modeling

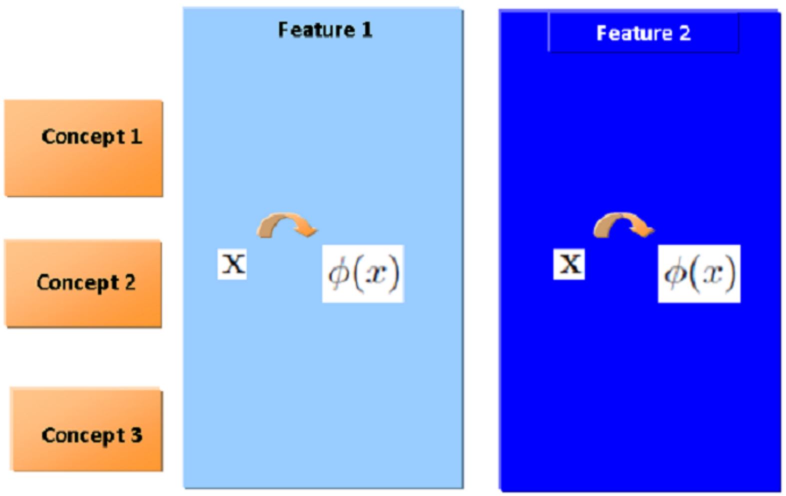
- Train collection of Unit Models on various subsets of data, image granularities, and features
- Each Unit Model on its own is "weak"
  - highly under-sampled entity
- Collection of Unit Models can be "strong"
  - cover most of the data/feature space
- Forward model selection Fusion strategy to generate strong Ensemble Classifier

- 1 Vs All classifiers learned for each class
- Max pooling used for multiclass classification

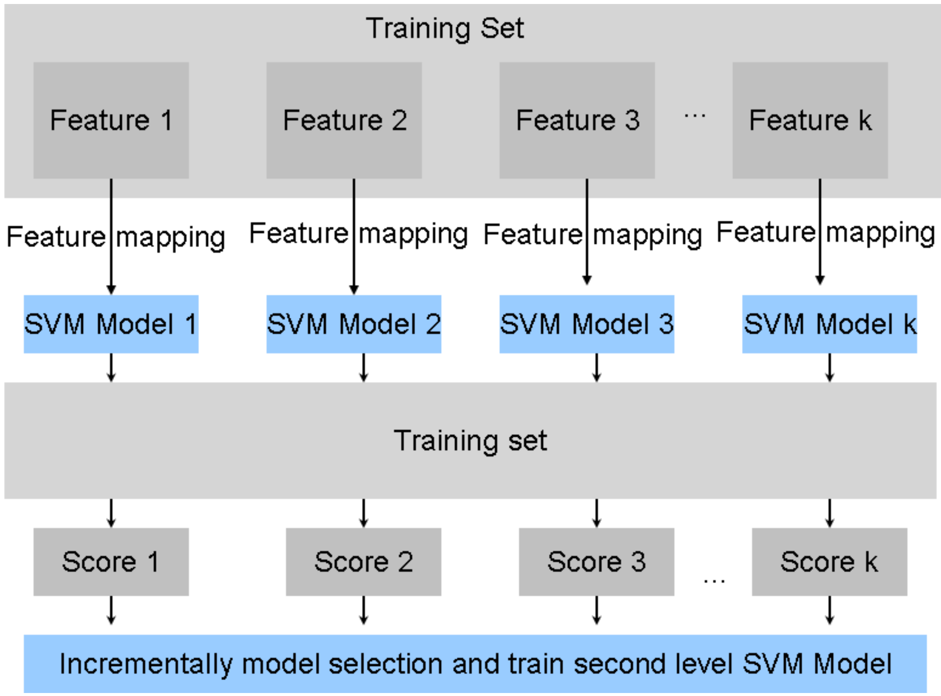


Modality Classification Task – Two level SVM + Kernel Approximation

- Motivated by the success of “deep-learning”, we make traditional SVM one layer deeper
- Traditional nonlinear kernel evaluation is very expensive, so we use kernel approximation to speed up the process
- 100% training accuracy and 81.05% (12 features) and 81.23% (24 features) for validation accuracy



Over all model: $K(x_i, x_j) = \sum_m \alpha_m K_m(x_i, x_j) = \sum_m \alpha_m \phi(x_i)\phi(x_j)$

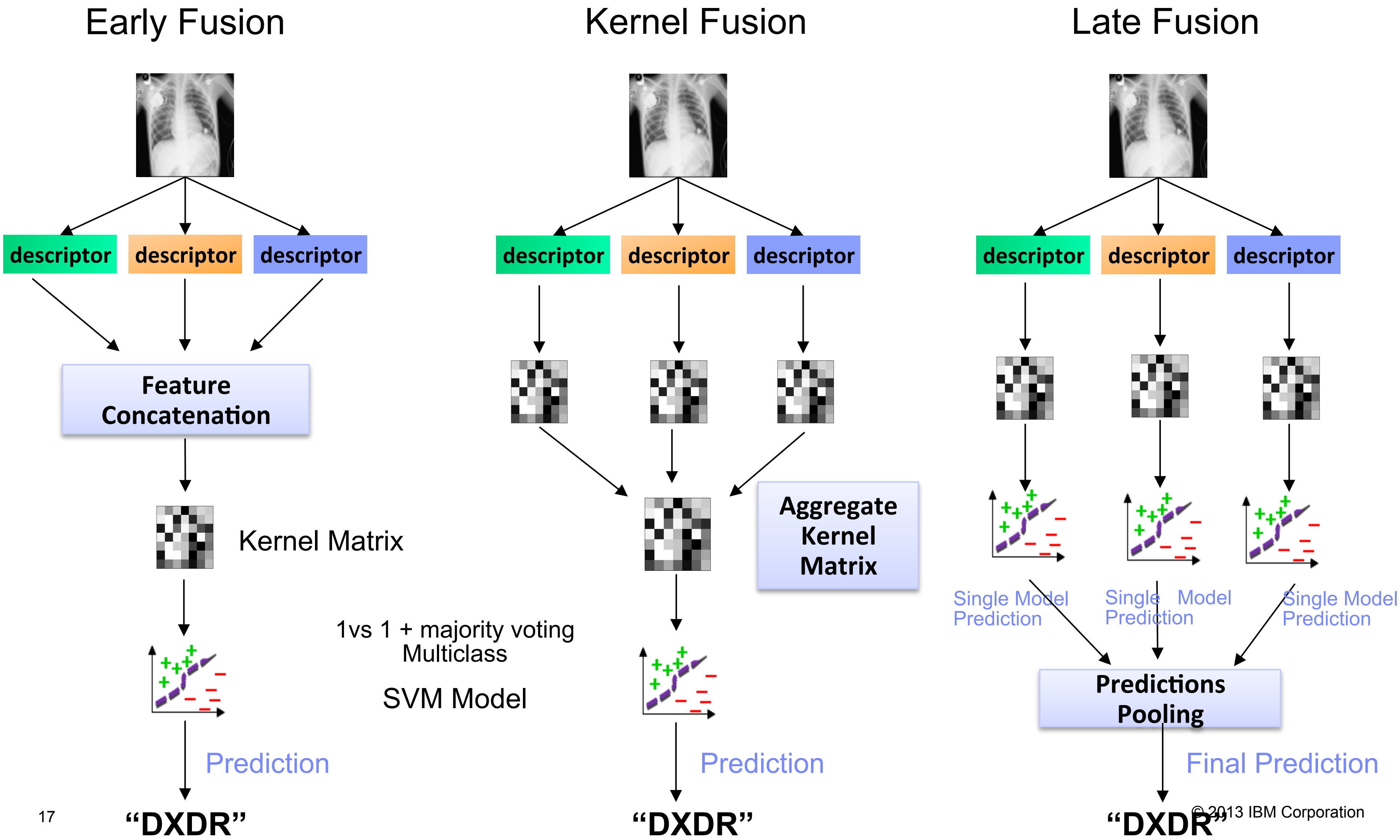


Modality Classification Task – Meta Classifiers

- Meta-learning[1] is a strategy to learn from learned knowledge
- Another level of supervised learning for combining the results of existing fusion models
- Collaboration model to combine the fusion models predictions
- **INPUT**: vector of different IMARS Ensemble models scores on top of visual and textual descriptions
- Learning algorithms tested:
  – Decision Tree
  – SVM (RBF Kernel, Poly kernel, Normalized Poly kernel and Puk kernel)
  – Random Forest
  – Logistic Model Tree (LMT)
  – Naive Bayesian

1. Kumari, D.M.U.R.G.P.: A study of meta-learning in ensemble based classier. Engineering Science and Technology: An International Journal (ESTIJ) 2(1) (February 2012) , pages 36-41

Modality Classification Task – Early/Kernel/Late Fusion



Modality Classification – Official Results

IBM submission runs: **10 Runs** →

- Top Textual
- Top Visual
- Top Mixed

Overall Best Performance for every submission type

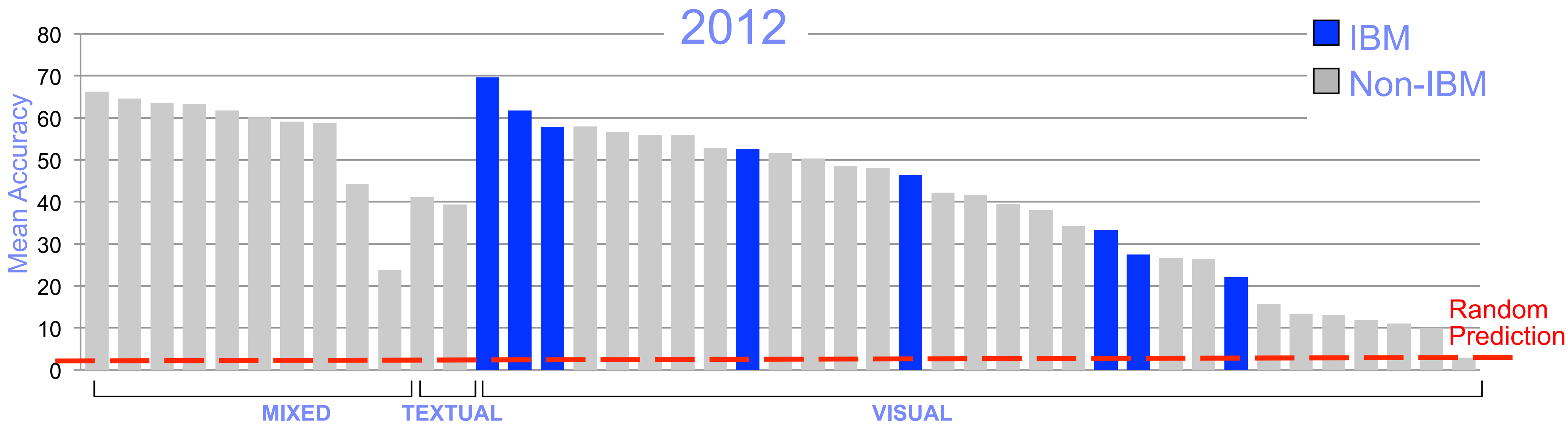
Modality Tailored Keywords

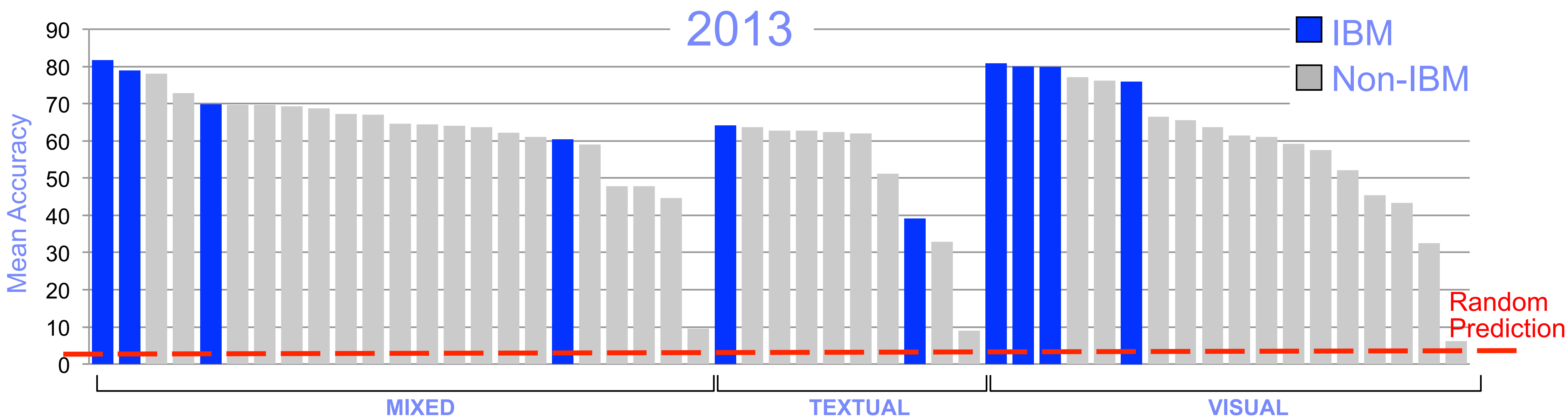
TYPE	NO EXTERNAL DATA	EXTERNAL DATA
Textual	IBM_modality_run1	IBM_modality_run2
Visual	IBM_modality_run3	IBM_modality_run5
Visual	IBM_modality_run4	IBM_modality_run6
Mixed	IBM_modality_run7	IBM_modality_run9
Mixed	IBM_modality_run8	IBM_modality_run10

Late fusion of all visual features and classification strategies

Late Fusion of Run1 and Run4

Modality Classification - Results



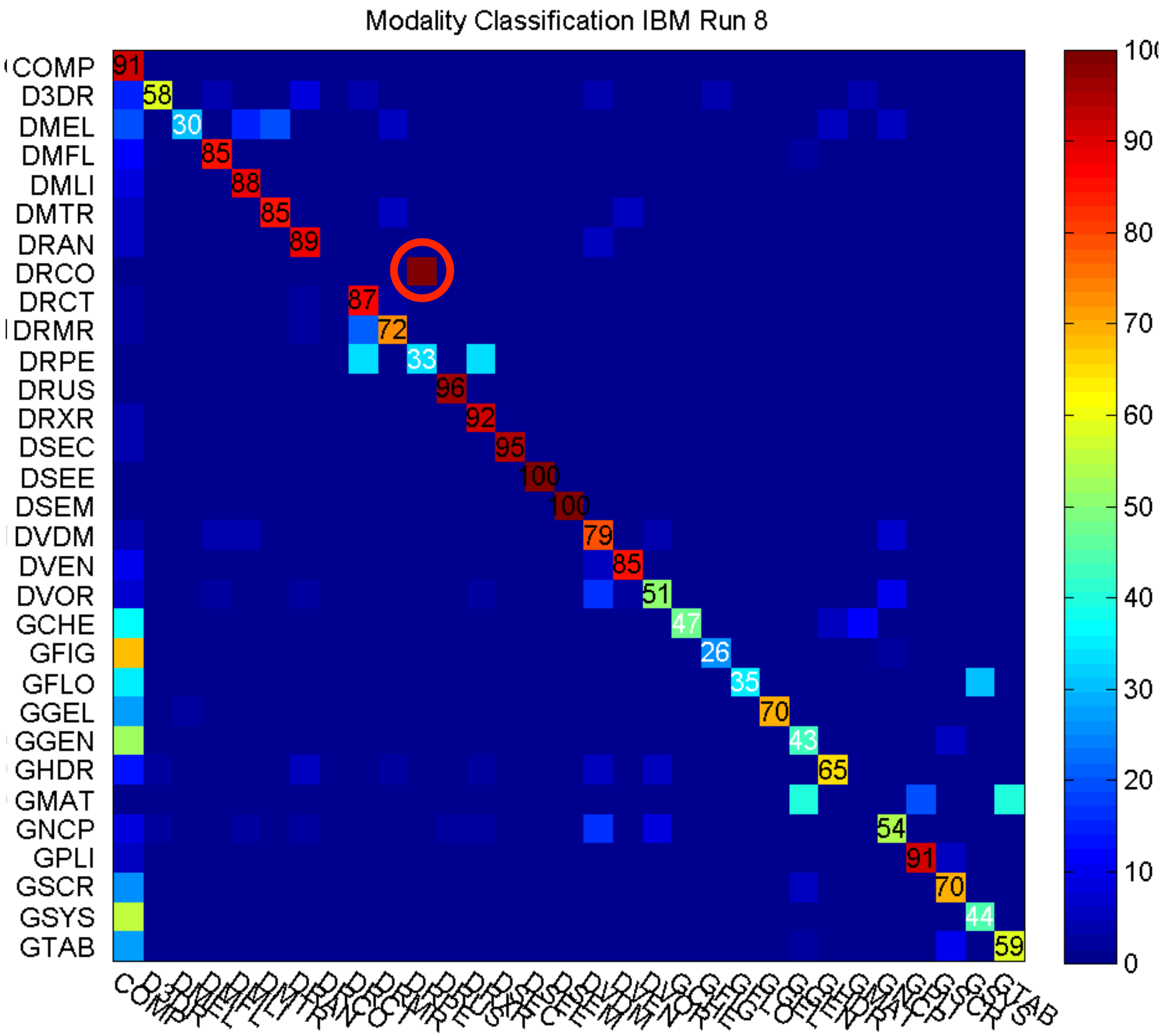


Modality Classification - Results

- Textual
- Visual
- Mixed

DRCO – Combined Radiology modalities in one image

Confused with DRPE (PET)

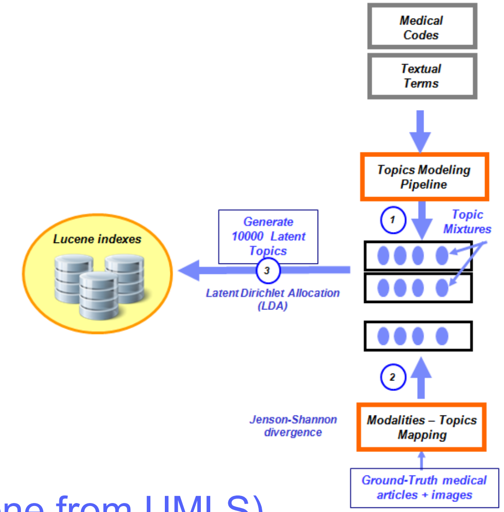


Overview

- IBM Multimedia Multi-Lab group @ ImageCLEF 2013
- Modality Classification task
  - Approaches
  - Results
- Case-based retrieval task
- Compound Image Segmentation Task
- Conclusions

Case Based Retrieval

- 35 query cases
- Dataset: 300K Pubmed Articles
- GOAL: return list of 1000 most relevant articles, given a query



APPROACH

- Based on textual Ontology Based Vocabulary (one vocabulary from WordNet, one from UMLS)
- Topic modeling approach to identify meaningful patterns from the medical documents
- LDA to detect the probability distribution P(w|z) over words given topic *z*
- Each medical document defined as a mixture of latent topics characterized by a multinomial distribution over words.
- Number of topics ranging from 100 to 10,000 topics. Gibbs sampling and Bayesian estimation to assign the multinomial distributions over a set of words to each latent topic
- Separated the topics that are defined for titles, abstracts and captions and grouped the medical documents that share the same topics
- Lucene index with TF-IDF

Results

	Runid	Retrieval type	MAP	P10	P30
	SNUMedinfo9	Textual	0.2429	0.2657	0.1981
WordNet	IBM_run_1	Textual	0.1573	0.1571	0.1057
Fusion	IBM_run_3	Textual	0.1573	0.1943	0.1276
UMLS	IBM_run_2	Textual	0.1476	0.2086	0.1295

Overview

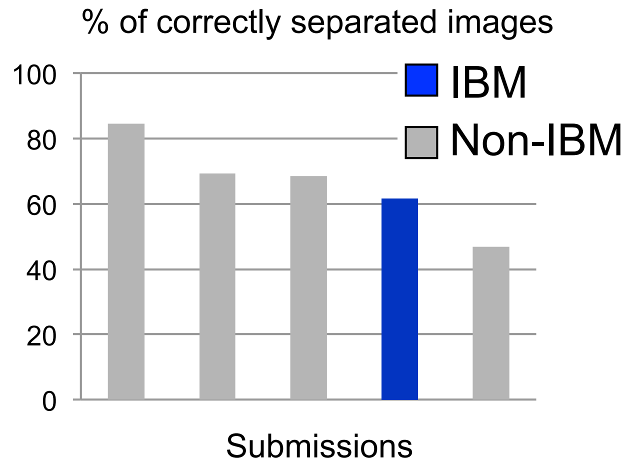
- IBM Multimedia Multi-Lab group @ ImageCLEF 2013
- Modality Classification task
  - Approaches
  - Results
- Case-based retrieval task
- Compound Image Segmentation Task
- Conclusions

Compound Image Segmentation Task

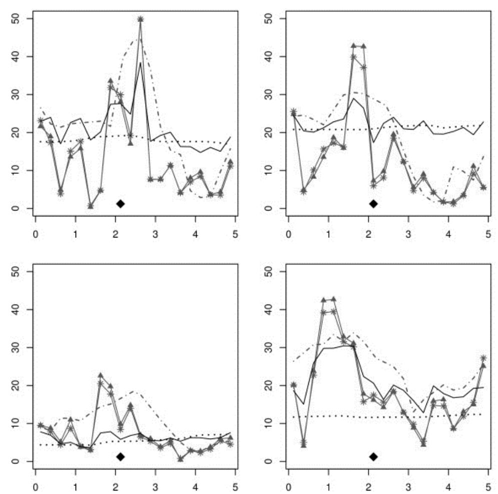
Combination of two approaches

- Analysis of connected components in a binarized image
  - Grayscale conversion
  - Binarization
  - Connected Components analysis
  - Geometric based filtering (size, proximity)
- Use of common notation of subfigures using text
  - OCR to recognize isolated components as letters (A, B , C)
  - Analysis of geometric layout of letters

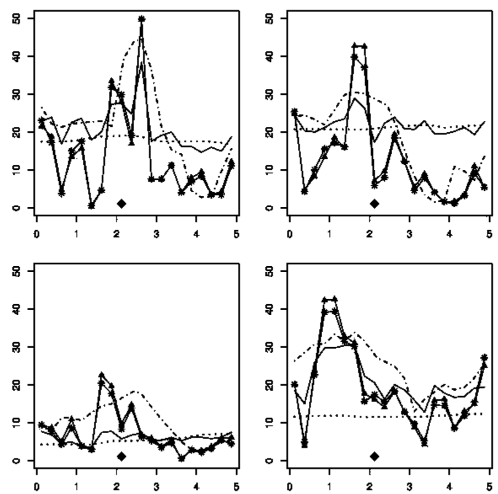
Results



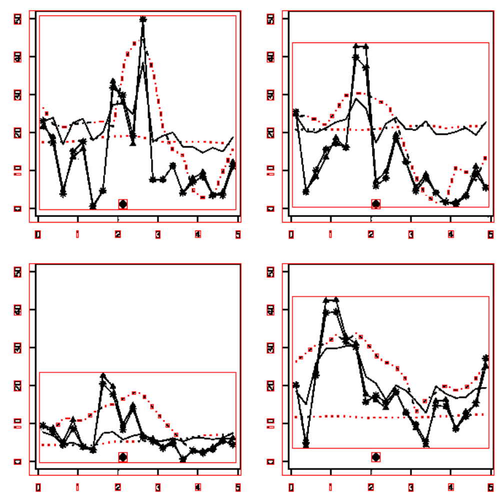
Input Image



Binarization Result



Connected Components



Overview

- IBM Multimedia Multi-Lab group @ ImageCLEF 2013
- Modality Classification task
  - Approaches
  - Results
- Case-based retrieval task
- Compound Image Segmentation Task
- Conclusions

Conclusions

- Semantic Model Vector best single performing feature

- Combination/fusion of different visual and textual based representations, as well as learning frameworks

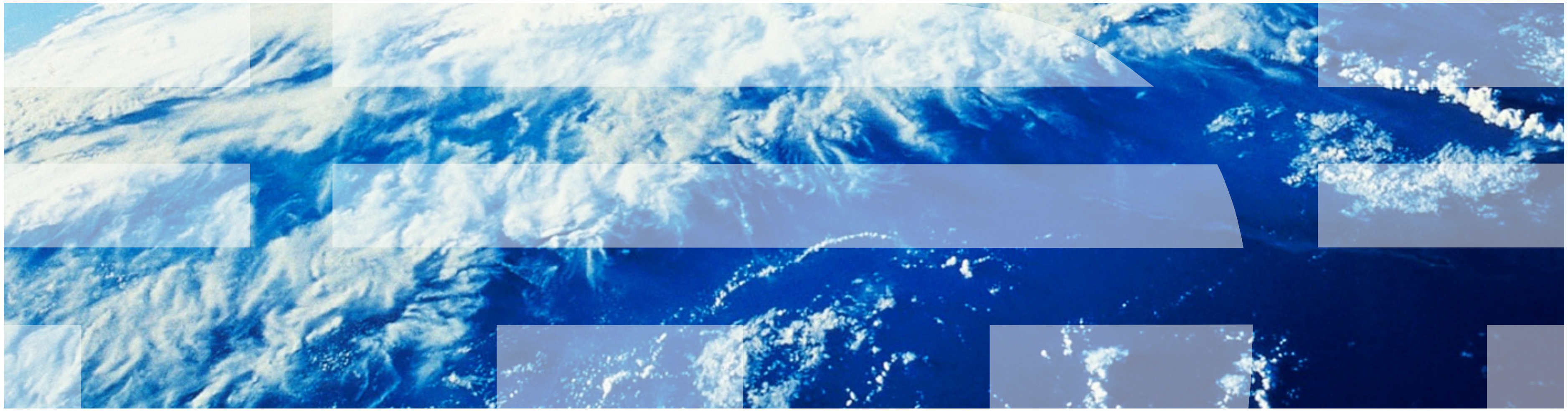
- Leverage combination of different sources for textual search/classification
  - Modality tailored extracted lexicon
  - General lexical ontology (WordNet) and
  - Medical specific domains medical knowledge-bases

- Future directions
  - Improve combination of complementary information from Visual and Textual domains

Mani Abedini and Rahil Garnavi - Australia
Amir Geva and Asaf Tzadok - Haifa
Liangliang Cao, Noel Codella, Jonathan H. Connell, Michele Merler, Quoc-Bao Nguyen, Sharathchandra U. Pankanti and John R. Smith - TJ Watson

Thank you!

Questions?



Modality Classification – Results 2012

- **Confusion Matrix**

  Better Performance

  Red Diagonal

- **Limited Training Data**
- **Extended Training Data**
  - **Reduced confusion**
  - **Still confused categories:**

System diagram vs. Flowchart