Hierarchical Annotation of Medical Images

Ivica Dimitrovski[1], *Dragi Kocev*[2], *Suzana Loškovska*[1], *Sašo Džeroski*[2]

[1]Department of Computer Science, Faculty of Electrical Engineering and Information Technologies, Skopje, Macedonia
[2]Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia

Overview

- Introduction and problem definition
- Feature extraction
  - Edge Histogram Descriptor (EHD)
- Classifier
  - PCTs for HMLC
- Experiments and results
- Conclusions and future work

Introduction

- The amount of medical images is constantly growing
- The cost of manually annotating these images is very high
  - automatic image annotation algorithms to perform the task reliably
- Feature extraction from images
- Classifier to distinguish between different classes
- Application: Multilingual image annotations and DICOM standard header corrections

IRMA code

- **IRMA** coding system: Four axes marked with {0, …, 9, a, …, z}
  - T (Technical): image modality
  - D (Directional): body orientation
  - A (Anatomical): body region
  - B (Biological): biological system
- IRMA code: **TTTT – DDD – AAA – BBB**
- The code is strictly hierarchical
  - Example:

2	cardiovascular system
21	cardiovascular system; heart
216	cardiovascular system; heart; aortic valve

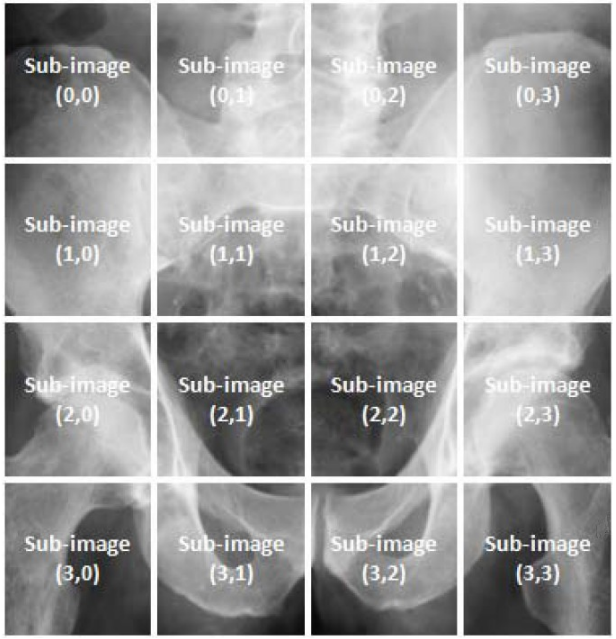
IRMA code - example

- IRMA code: 1123-211-520-3a0
  - 1123 (x-ray, projection radiography, analog, high energy)
  - 211 (sagittal, left lateral descubitus, inspiration)
  - 520 (chest, lung)
  - 3a0 (respiratory system, lung)

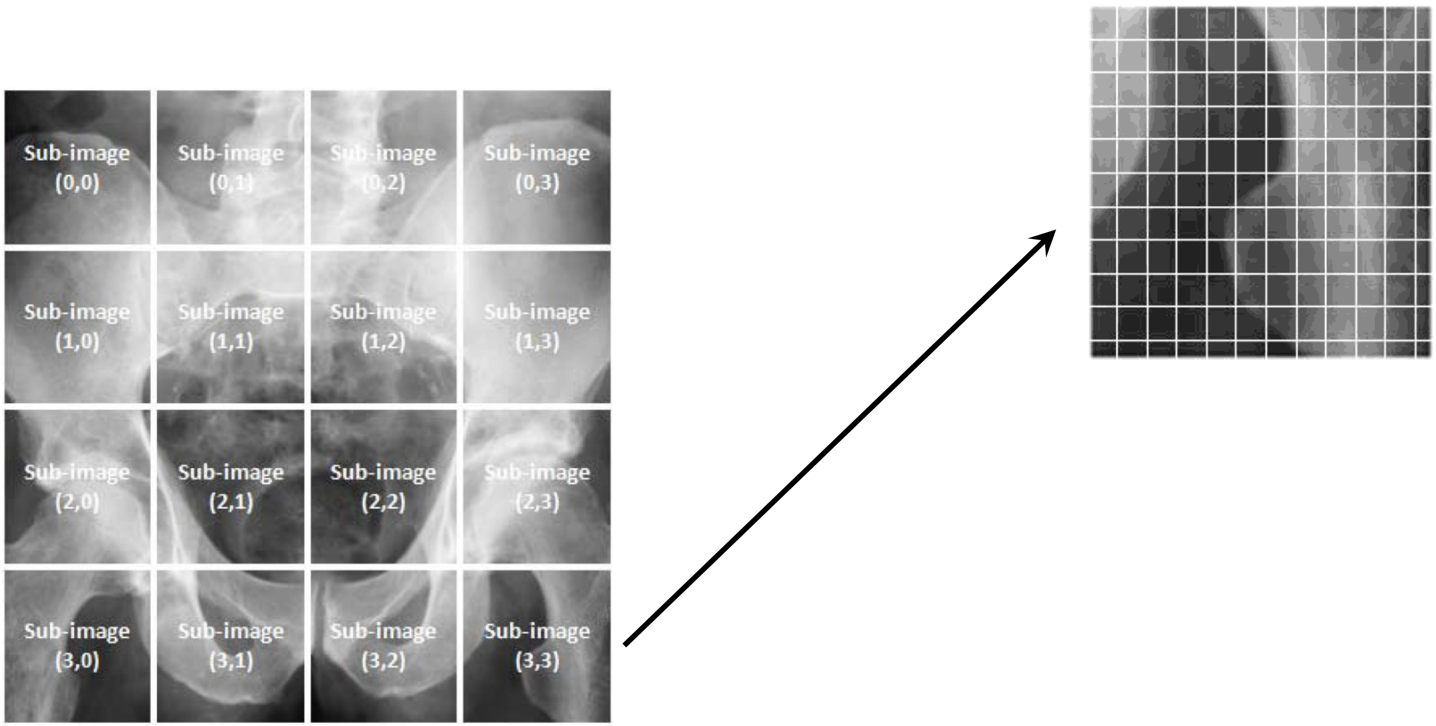
Feature extraction

- Obtain features that describe the visual content of an image
- Histogram of local edges
  - Mark the points in a digital image at which the luminous intensity changes sharply
  - Reduction of the amount of data to be processed, while retaining important information about the shapes of objects in the image
  - Frequency and the directionality of the brightness changes in the image

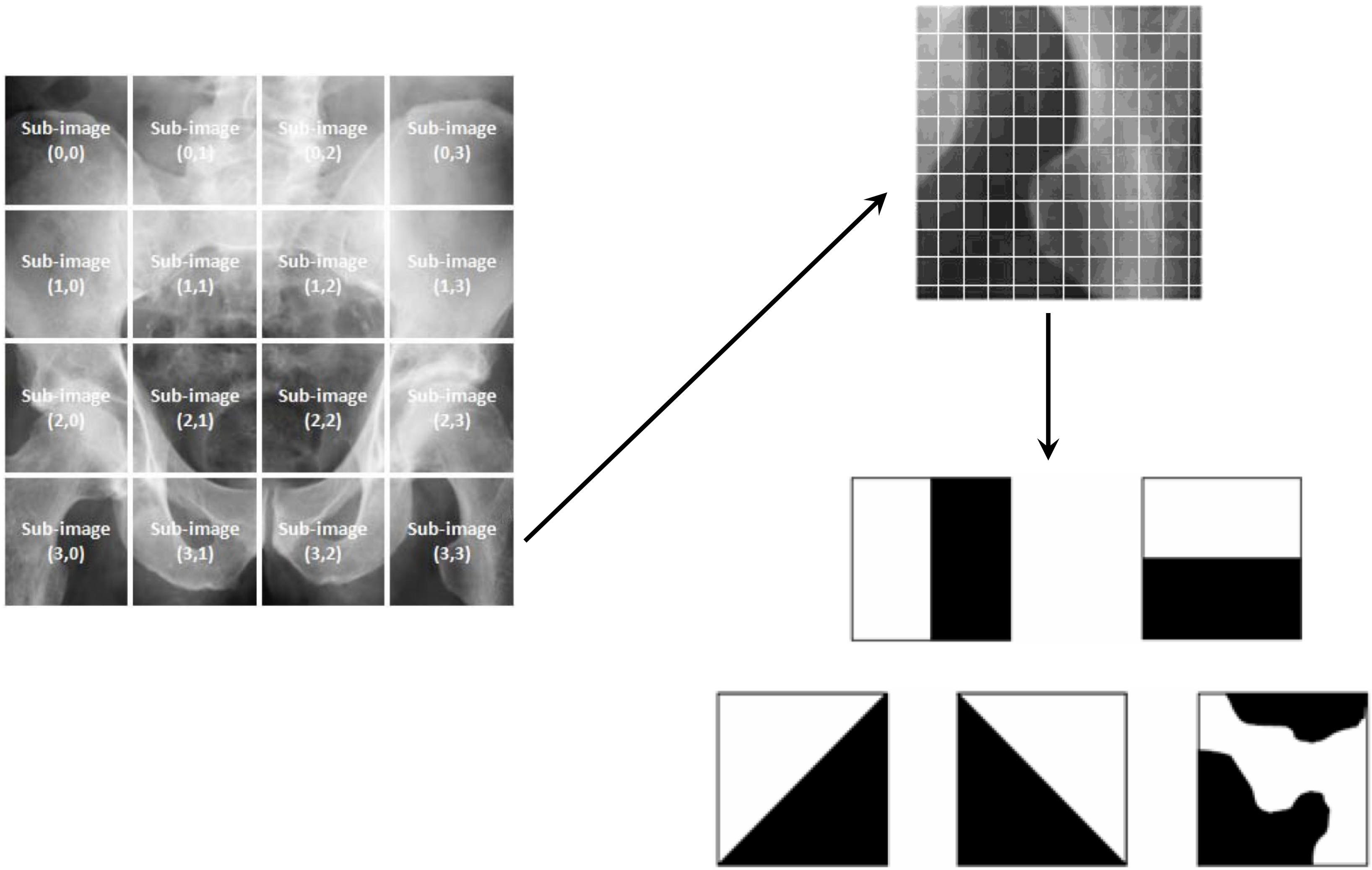
Feature extraction



Feature extraction



Feature extraction



Classification Methodology

- Predictive Clustering Trees framework (PCTs)
  (Blockeel et al. Top-down induction of clustering trees. In Proc. of the 15th ICML, p.55-63, 1998)

- Ensemble methods to improve the predictive performance
  - Bagging (L. Breiman. Bagging predictors, Machine Learning Journal, vol. 24 Issue 2, p. 123-140, 1996)
  - Random Forests (L. Breiman. Random Forests, Machine Learning Journal, vol. 45, p.5-32, 2001)

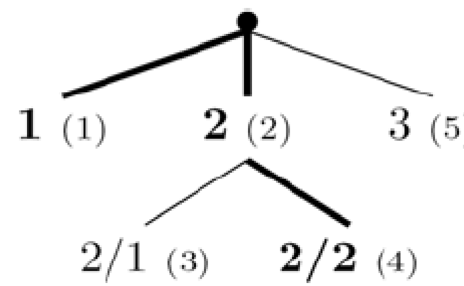
Predictive Clustering Trees (PCTs)

- A tree is a hierarchy of clusters
- Standard top-down induction of decision trees (TDIDT) algorithm
- best acceptable attribute-value test that can be put in a node
- The heuristic for selecting the tests is the reduction in variance in the induces subsets
  - Maximizes cluster homogeneity and improves predictive performance
- PCTs can handle different types of target concepts: multiple targets, time series, hierarchy
  - Instantiation of the variance and prototype function

Hierarchical Multi-Label Classification

- HMLC: an example can be labeled with multiple labels that are organized in a hierarchy

{ 1, 2, 2.2 } ⟶

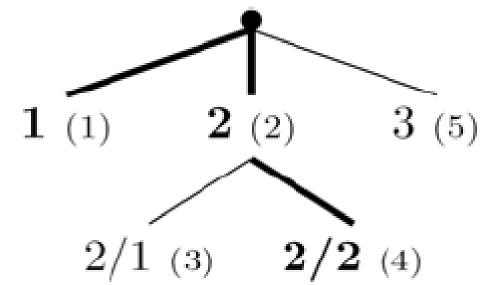


$$v_i = \overset{(1)\,(2)\,(3)\,(4)\,(5)}{[1,\, 1,\, 0,\, 1,\, 0]}$$

Hierarchical Multi-Label Classification

- HMLC: an example can be labeled with multiple labels that are organized in a hierarchy

$\{ 1, 2, 2.2 \}$ ⟶



$$v_i = \overset{(1)}{[1}, \overset{(2)}{1}, \overset{(3)}{0}, \overset{(4)}{1}, \overset{(5)}{0]}$$

- Variance instantiation:
    - average squared distance between each example's label and the set's mean label
    - the arithmetic mean of a set of such vectors contains as i'th component the proportion of examples of the set belonging to class ci

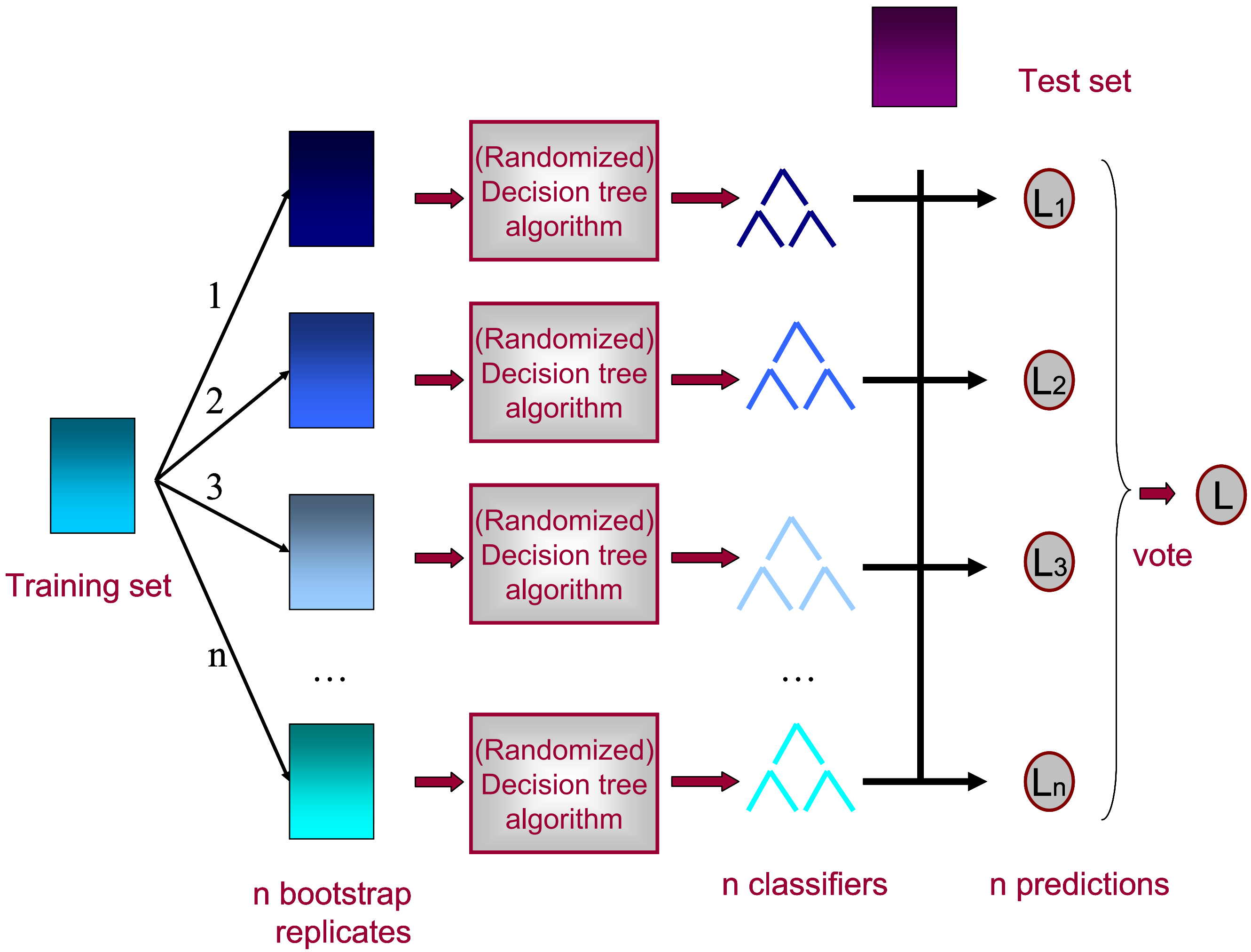
$$Var(S) = \frac{\sum_i d(v_i, \bar{v})^2}{|S|} \qquad d(v_1, v_2) = \sqrt{\sum_i w(c_i) \cdot (v_{1,i} - v_{2,i})^2}$$

Ensemble methods

- Ensemble - Set of classifiers
- Classification of new example by combination of the predictions of each classifier from the ensemble
  - Regression: Average
  - Classification: Majority Vote

- Bagging
- Random Forests

Ensemble methods



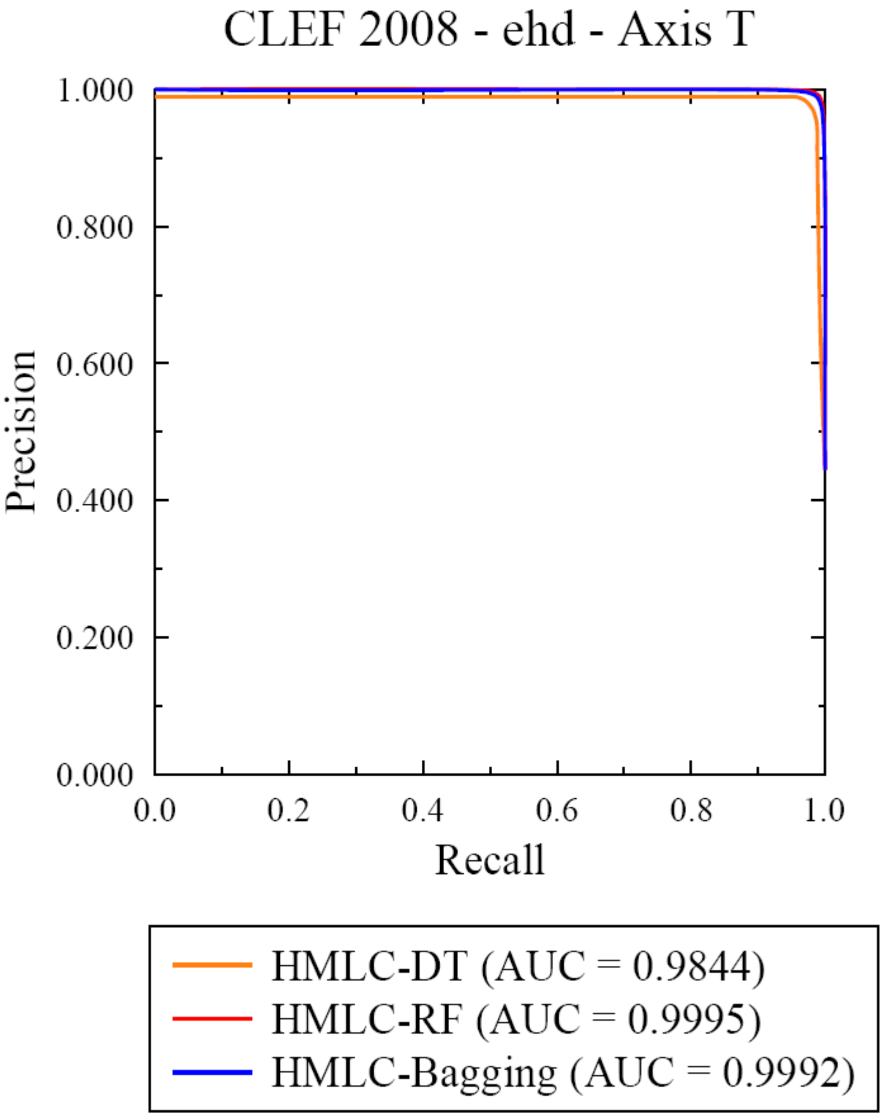
Experimental design

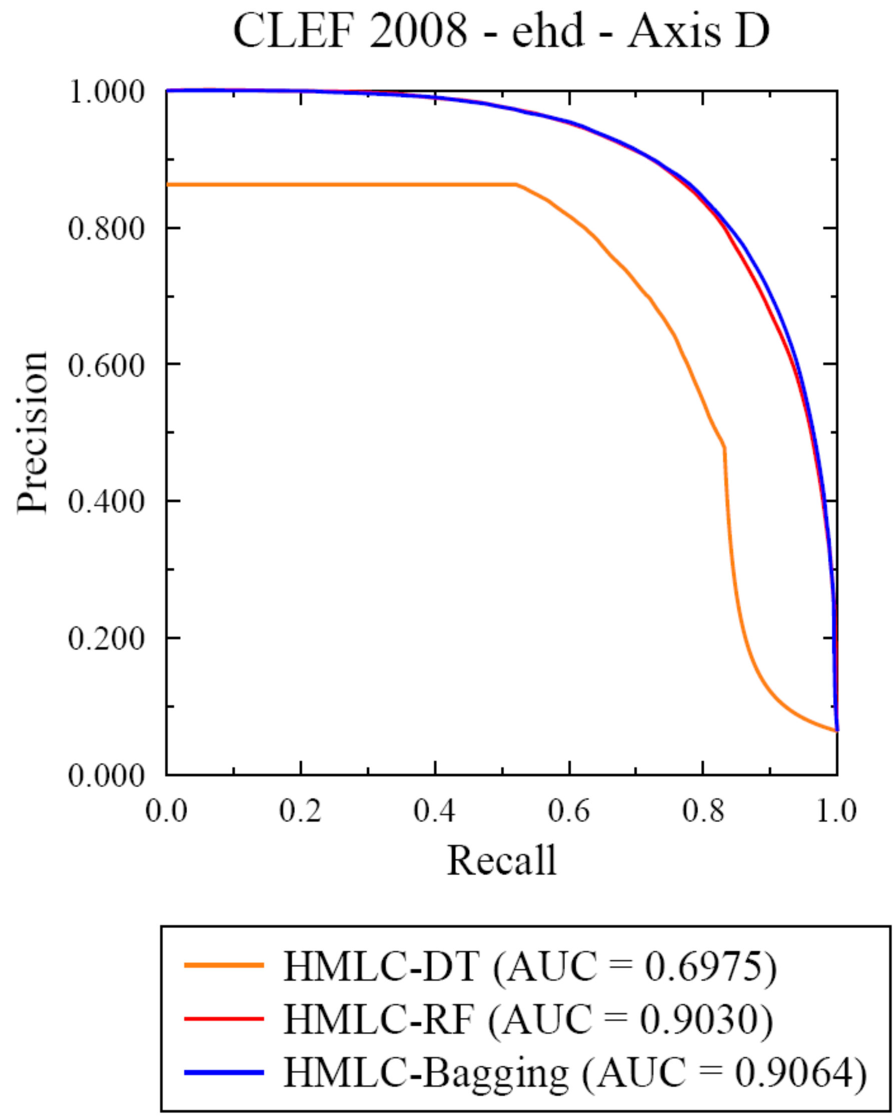
- Goal: provide the IRMA code for an image
- Data
  - ImageCLEF 2008
  - 12.076 images sorted in 197 classes
  - 82 classes have less than 10 elements (129 images)
  - Each image is described with 80 features
- Feature extraction
  - Contrast enhancement
  - Histogram equalization

Experimental design

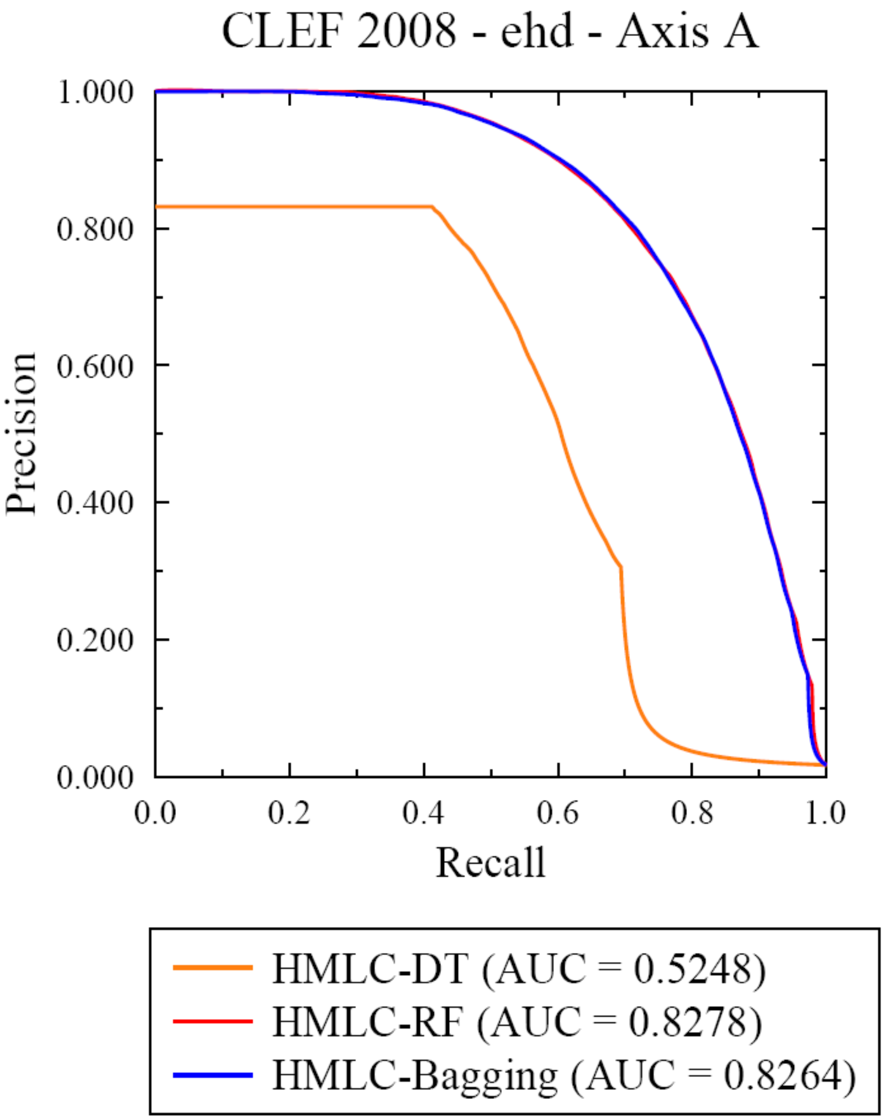
- Classifier
  - Number of classifiers: 100 un-pruned trees
  - Random Forests Feature Subset Size: 7 (log)
- Comparison of the performance of a single tree and an ensemble
  - Precision-Recall (PR) curves - “area under the PR curve” (AUPRC)
  - 10 fold cross-validation
- Two scenarios
  1) Each axis is an dataset (4 in total)
  2) Single dataset for all axes

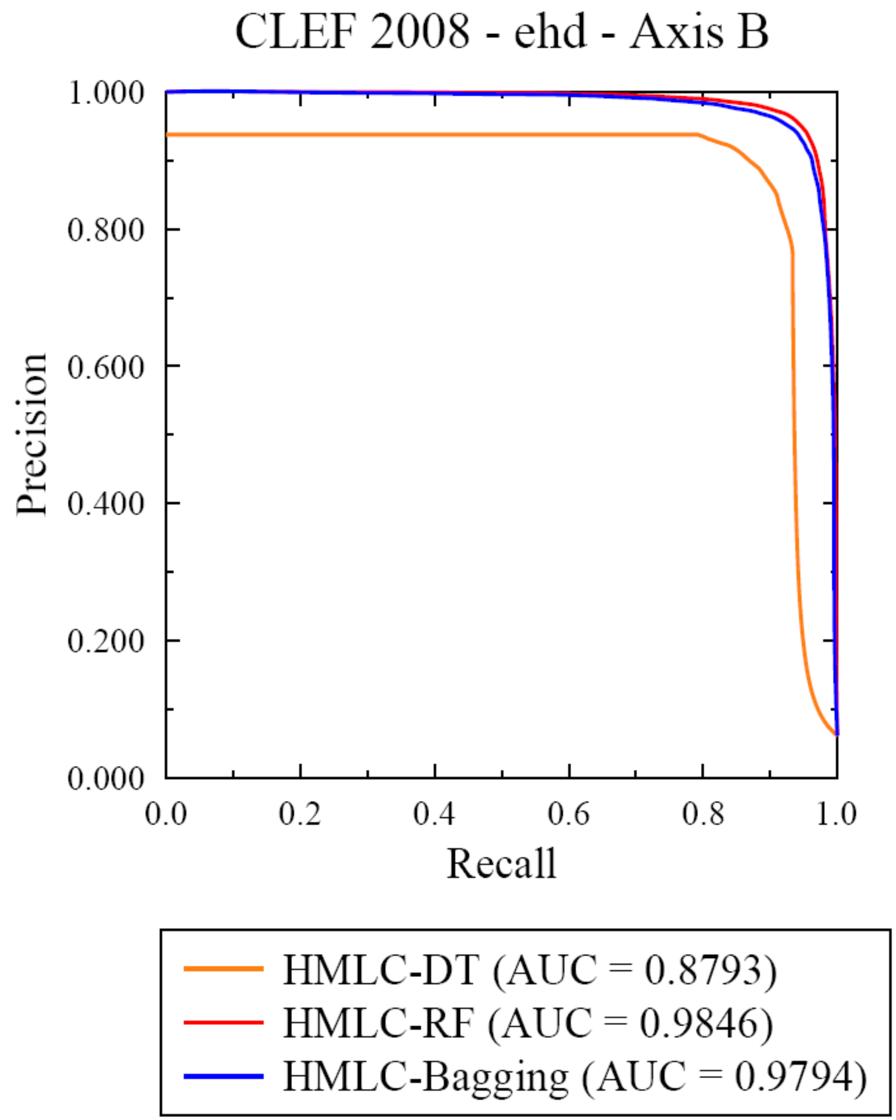
Results per axis



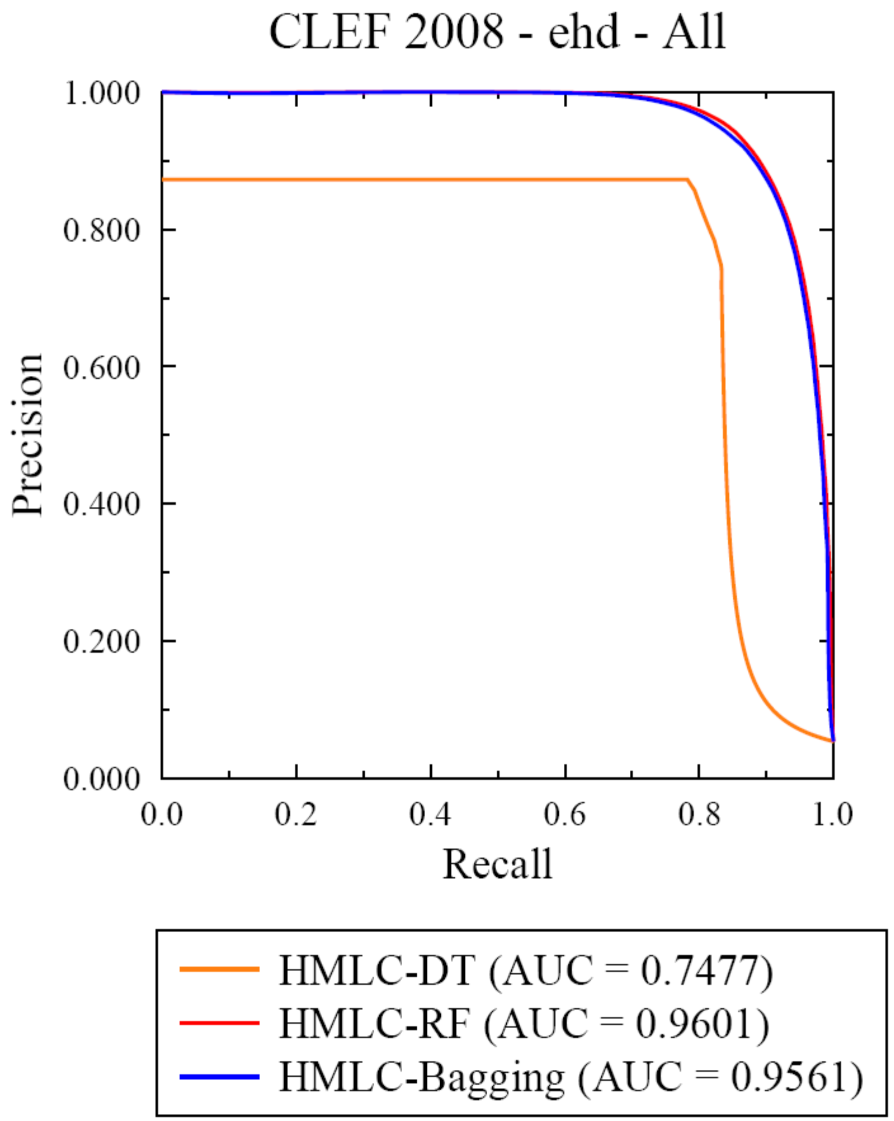


Results per axis





Results for all axes



Discussion

- Increase of the predictive performance with ensembles compared to a single tree
- Excellent performance for axes T and B (AUPRC of 0.9994 and 0.9862)
  - The hierarchies for axes T and B contain only few nodes (9 and 27, respectively)
- The classifiers for axes A and D have high predictive performance (AUPRC of 0.8264 and 0.9064)
  - The hierarchies for axes A and D contain 110 and 36 nodes, respectively
- Predicting the complete hierarchy at-once yields improvements

Summary

- Medical image annotation using Hierarchical Multi-Label Classification (HMLC)
- Local Edge Histogram Descriptor (EHD) to represent gray-scale radiological (X-Ray) images
- Images annotated with IRMA code
- Ensembles of PCTs for HMLC as classifier

Future work

- Other algorithms for feature extraction:
  - SIFT, TAMURA, Scale, Color Histogram…
- Combination of the features obtained from different techniques:
  - Each technique captures different aspects of an image
- Extension of the classification algorithm:
  - Distance measures for hierarchies
  - Learning under covariate shift