SCIENCE
UNIVERSITY OF COPENHAGEN

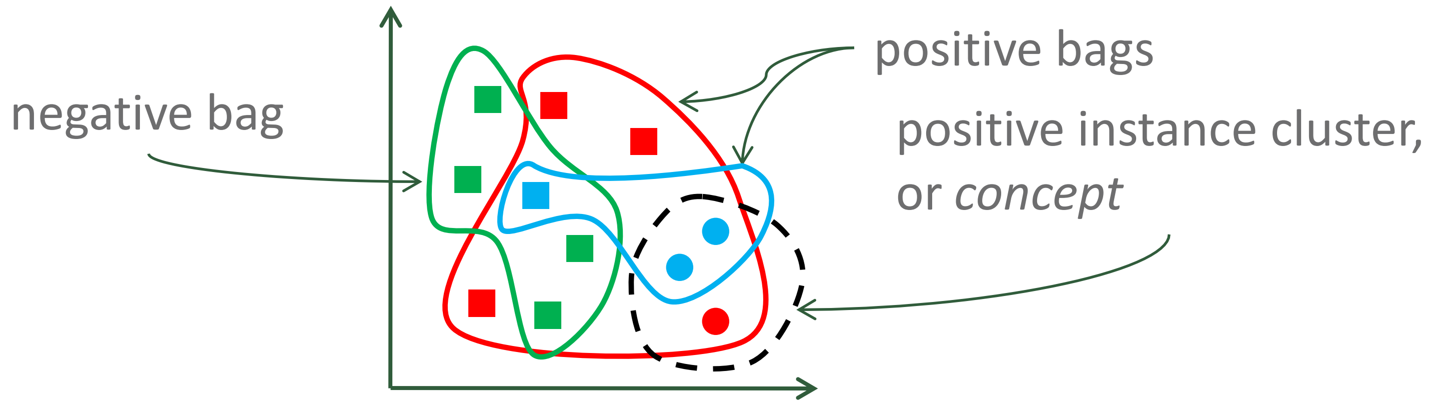
Dissimilarity-Based Multiple Instance Learning

Lauge Sørensen[1], Marco Loog[1,2], David M.J. Tax[2], Wan-Jui Lee[2], Marleen de Bruijne[1,3], and Robert P.W. Duin[2]

1) The Image Group, Dept. of Computer Science, University of Copenhagen, Denmark
2) Pattern Recognition Laboratory, Delft University of Technology, The Netherlands
3) Biomedical Imaging Group Rotterdam, Departments of Radiology & Medical Informatics, Erasmus MC, Rotterdam, The Netherlands¤

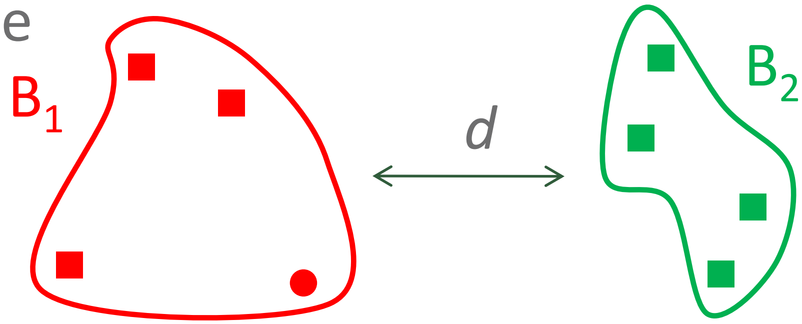
Multiple Instance Learning (MIL)

- In MIL an object is represented by a set (*bag)* of sub-objects (*instances)* with a label (positive/negative) associated with the entire bag
- Unseen bag classified based on the instances it contains:
  - positive = contains at least one positive instance
  - negative = contains no positive instances

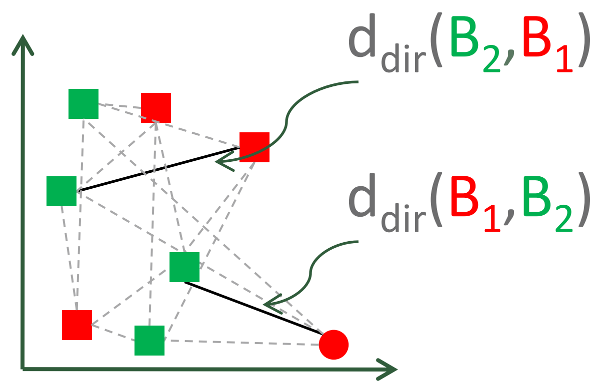


Approach

- Dissimilarity-based classification of bags using a suitable bag dissimilarity measure



- Done before using the *k*NN classifier (and modifications; Citation-*k*NN) with the Hausdorff distance between instances as dissimilarity [1]

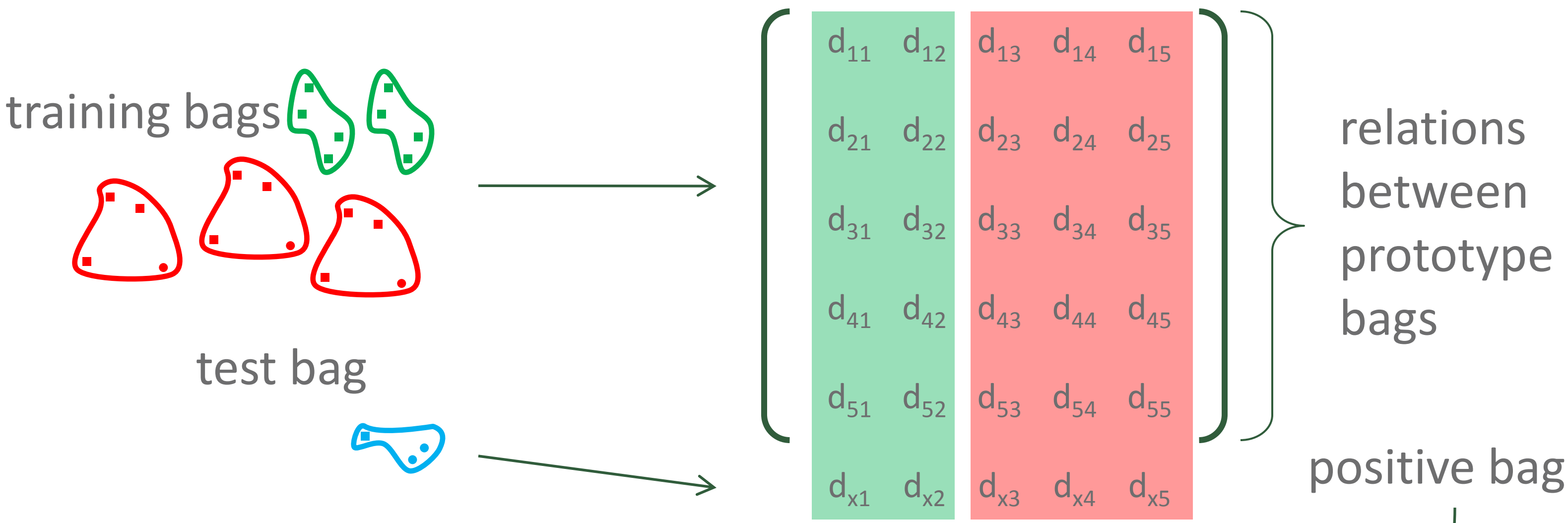


$$d_H = \max\{d_{dir}(B_1,B_2),\ d_{dir}(B_2,B_1)\}$$

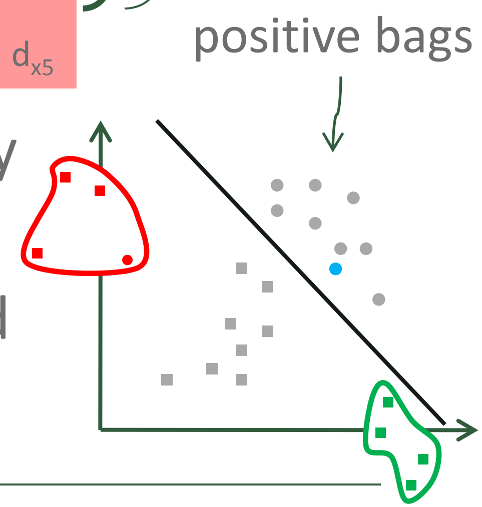
[1] Wang, J. and Zucker, J.D., "Solving the multiple-instance problem: A lazy learning approach," In Proc. ICML, 2000

Approach

- Propose to use the dissimilarity representation approach [2]



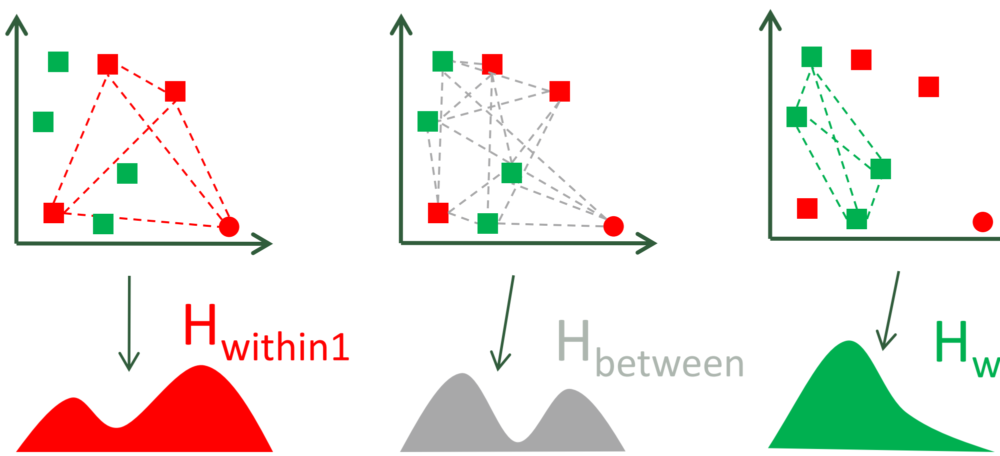
- Dissimilarity space defined by the dissimilarity to prototype bags
- Standard vector space method can be applied

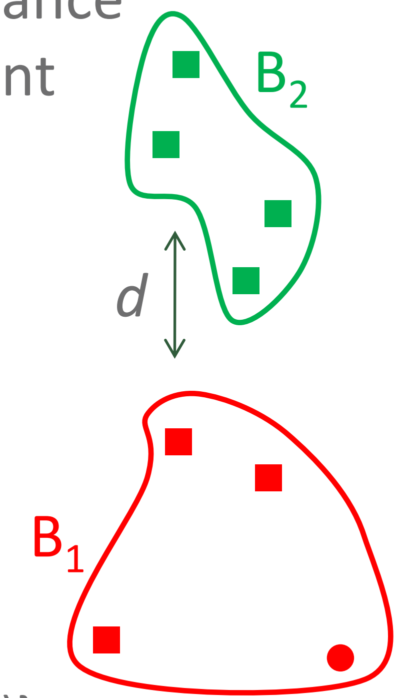


[2] Pekalska, E. and Duin, R.P.W., “Dissimilarity representations allow for building good classifiers,” Pattern Recognition Letters, 2002

Instance Relations in Bag Dissimilarity

- Propose bag dissimilarities that take instance distance distributions between and within bags into account





$$d_{BWmean} = \text{mean}\{\text{EMD}(H_{within1}, H_{between}), \text{EMD}(H_{within2}, H_{between})\}$$

- Assumes positive and negative bags have different distributions of instance distances

Advances

- Better utilization of the training data compared to *k*NN
- Less restrictions on the bag proximity measure compared to, e.g., kernel-based approaches
- Now we can use:
  - L_1
  - Hausdorff (and modifications hereof),
  - Hamming
  - single linkage
  - Fisher criterion
  - d_{BWmean} (proposed bag dissimilarity measures)
  - ...

Advances

- Within-bag instance relations may add discriminative information

uses instance relations within bags

	Elephant	Fox
d_{BWmean} + Fisher	88.5	64.0
MIGraph [3]	85.1	61.2
MI-SVM [4]	81.4	59.4

- Within-bag relations modeled less rigid as opposed to, e.g., using a graph representation + graph kernel [3]

[3] Zhou, Z.H., Sun, Y.Y., and Li, Y.F., "Multi-instance learning by treating instances as non-i.i.d. samples ," In Proc. ICML, 2009
[4] Andrews, S., Tsochantaridis, I., and Hofmann, T., "Support vector machines for multiple-instance learning," In Proc. NIPS, 2002

Challenges

- Designing good bag dissimilarity measures
  - using application specific knowledge
  - using the training set
- Systematic prototype selection for good dissimilarity spaces [5]
- How to account for the number of instances in the bags being compared

[5] Pekalska, E., Duin, R.P.W., and Paclik, P., "Prototype selection for dissimilarity-based classifiers," Pattern Recognition , 2006

Questions

- Can we solve the “strict” MIL problem (bag positive iff. contains at least one positive instance)?
- Which more general MIL problems can we solve?
- Possible to characterize for which MIL problems it is useful to take instance relations (within bags) into account?
- Can we design good bag dissimilarity measures that could otherwise not be used in a kernel-based approach?
- How can one exploit the fact that bags can contain a substantially different number of instances?