MULTIPLE INSTANCE CLASSIFICATION IN THE IMAGE DOMAIN

Ilaria Bartolini, Pietro Pascarella, Marco Patella

University of Bologna, Italy

SISAP 2019, Newark, October 4th

Two tigers on the grass

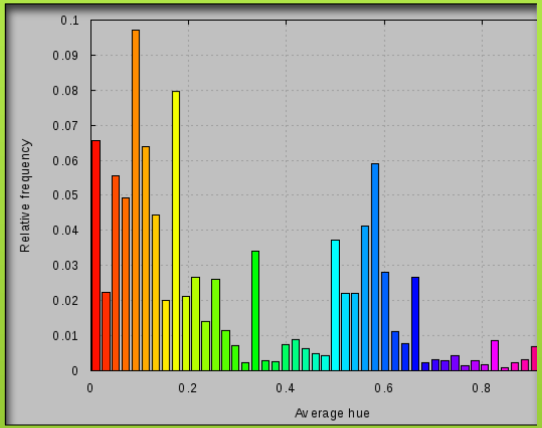
- Searching for images of interest based on their visual content
  - Retrieve the images most visually similar to a given query image
- Different wrt concept-based image retrieval
  - Uses text-based techniques

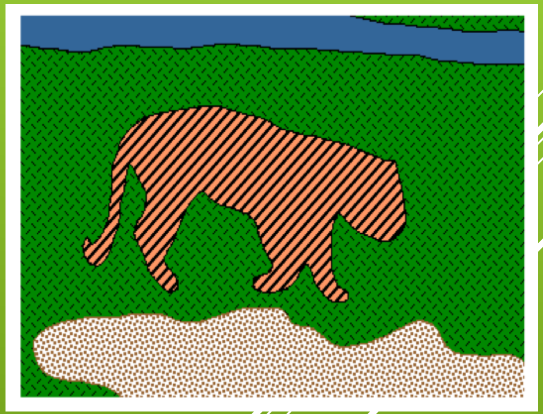
CONTENT-BASED IMAGE RETRIEVAL

- Evaluating the similarity between two images entails:
  - Automatically extract *features/descriptors* from the images
  - Compare such features to assess a *similarity score*
  - The higher the score, the more similar the images' contents are

SIMILARITY-BASED CBIR

- Global features describe the visual content of an image as a whole
  - Color histograms
- Local features describe the visual characteristics of a (small) set of image pixels
  - Region descriptors
  - Salient point descriptors
- Nowadays, the latter overshadow the former





A BIT OF HISTORY

- Typically by classifying a set of query images (whose class is known)over a knowledge base of images

"A feature that performs well for the task of classification on a certain data set, it will most probably be a good choice for retrieval of images from that data set, too."

Desalaers, Keysers, Ney. Inf. Retr., 2008

HOW DO YOU ASSESS EFFECTIVENESS OF A CBIR SYSTEM/TECHNIQUE?

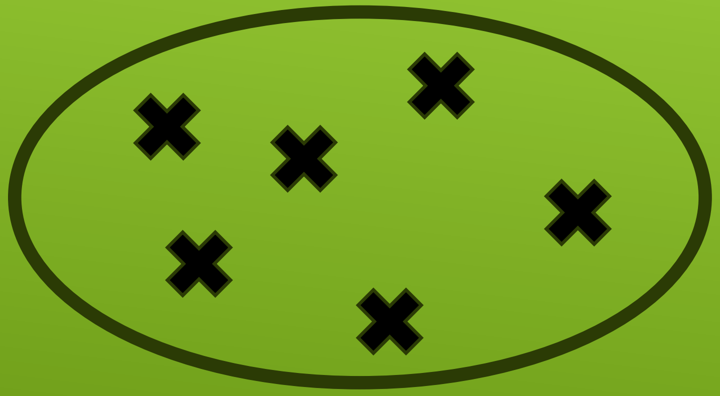
- Most of the proposed techniques lack emphasis on classification techniques
- And what about efficiency?
  - Indexing, anyone?

HOWEVER…

- A number of classification techniques drawn from the realm of Machine Learning
- Local features perfectly fit the scenario of applicability of MIC
- It could help researchers working in CBIR to evaluate their proposed features and/or indexing techniques in a more structured way
  - What are the alternatives at hand?
- Each technique also suggests a retrieval method

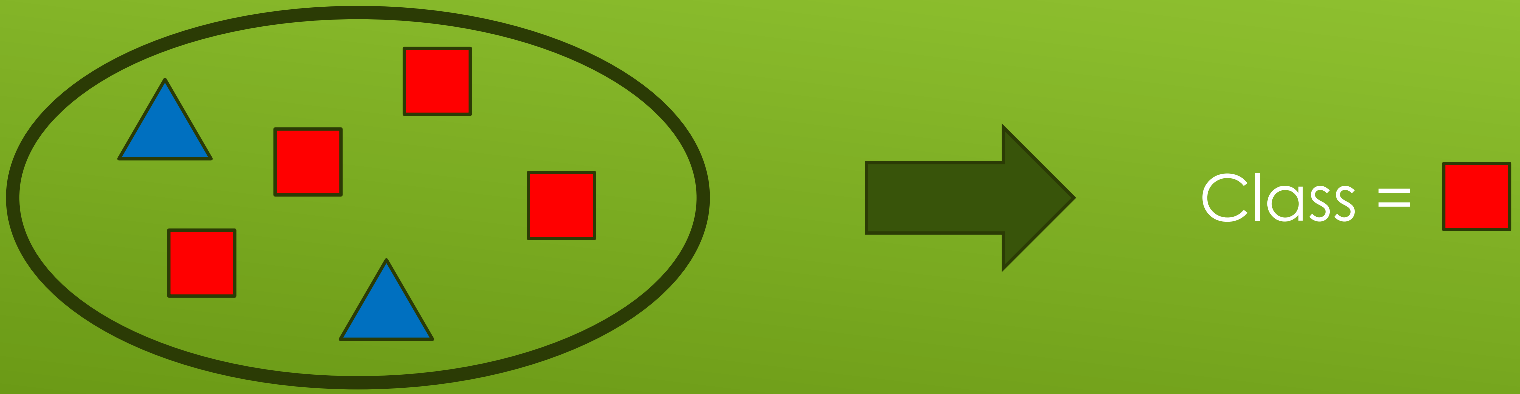
INTRODUCING MULTIPLE INSTANCE CLASSIFICATION (MIC)

- Each object (*image*) is seen as a *bag* of individual *instances* (*features/descriptors*)
- The *class* of each bag can be transferred to all (or some) of its instances



THE MIC MODEL

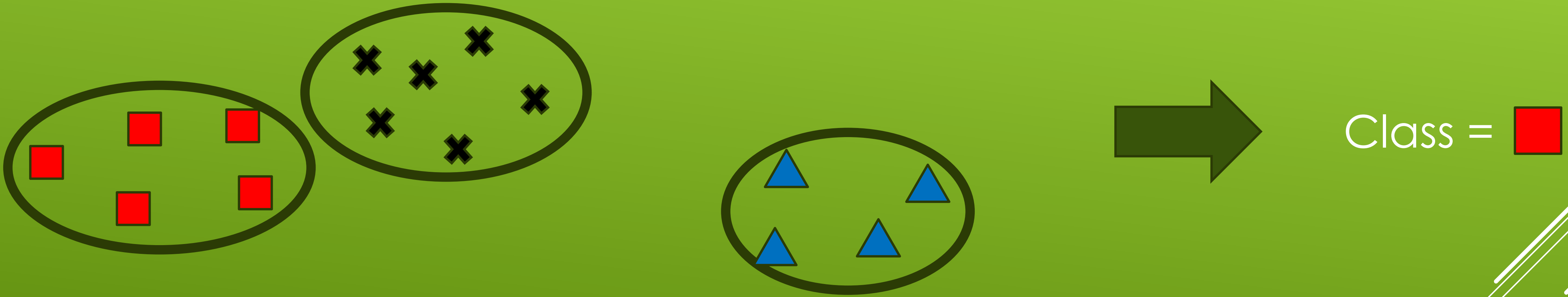
- The discriminative information lies at the instance level
  - Classification is performed on instances
    - For example, using a distance between descriptors
  - The overall classification is performed by *aggregating* classifications obtained at the instance level
- Retrieval (and indexing) is based on instances (local features)



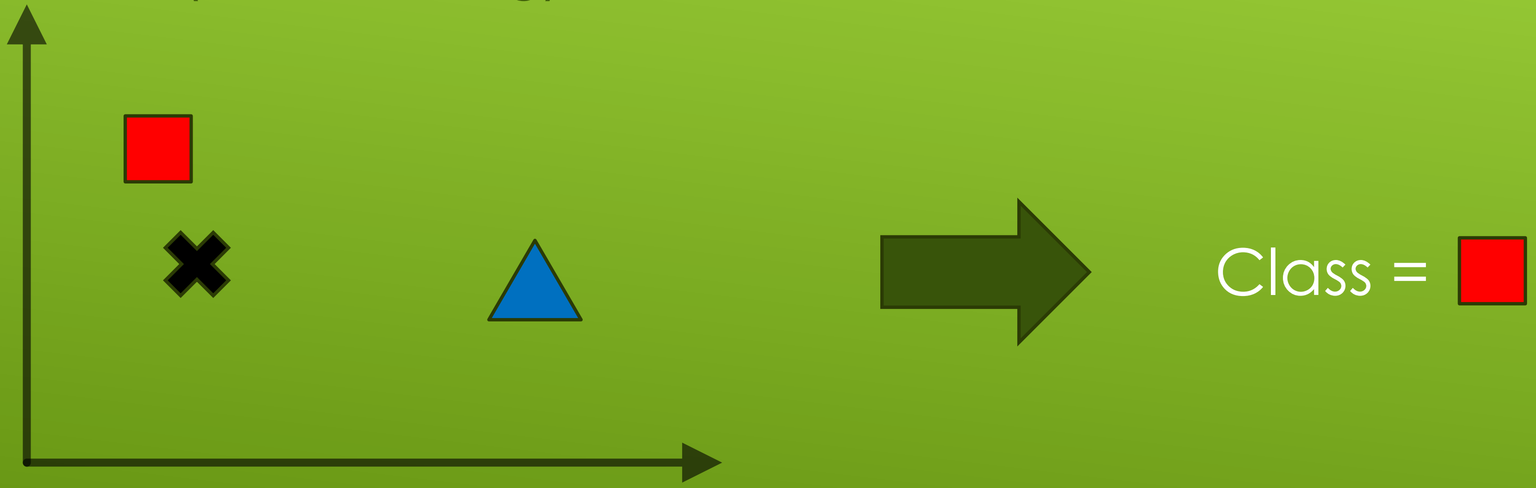
INSTANCE SPACE PARADIGM

BAG SPACE PARADIGM

- The discriminative information lies at the bag level
  - This cannot be distributed to instances
  - The overall classification is performed by *aggregating* distances at the instance level
    - For example, using an overall *distance* function between images
    - Commonly used aggregators: EMD, Hausdorff, etc.
- Retrieval (and indexing) is based on images



- Each bag is mapped to a single feature vector
  - A vector-based classifier is exploited
  - Only makes sense when the number of instances in a bag is very high
    - Example: Bag-Of-Visual-Words
- Retrieval (and indexing) is based on vectors



EMBEDDED SPACE PARADIGM

- Comparing performance of two local features
  - WINDSURF region-based features
    - On average, 4-5 regions per image (ES not applicable)
  - SIFT salient point descriptors
    - On average, thousands of keypoints per image
- Alternatives implemented on top of the WINDSURF framework
  - Provides algorithms and indexing structures
    for efficient query processing based on local features
  - All three paradigms are implemented
  - M-tree indexing instances (descriptors), bags (images), and vectors

EXAMPLE OF USE

features	WINDSURF	SIFT
Instance Space	Accuracy: **good** Efficiency: **very good**	Accuracy: **very good** Efficiency: **very bad**
Bag Space	Accuracy: **bad** Efficiency: **very good**	Accuracy: **bad** Efficiency: **bad**
Embedded Space	Not applicable	Accuracy: **good** Efficiency: **very good**

COMPARISON OF ALTERNATIVES

- MIC can be a handy tool for researchers in CBIR
- When proposing features,
  always consider *all* available alternatives
- And never forget efficiency!

LESSONS LEARNED