INFO0948
Feature Extraction

Renaud Detry

University of Liège, Belgium

Updated November 27, 2013

These slides are based on Chapter 13 of the book *Robotics, Vision and Control: Fundamental Algorithms in MATLAB* by Peter Corke, published by Springer in 2011.
The Bag-of-feature section is based on a presentation by Cordelia Schmid
`http://www.di.ens.fr/willow/events/cvml2011/materials/CVML2011_Cordelia_bof.pdf`

These slides are based on Chapter 13 of the book *Robotics, Vision and Control: Fundamental Algorithms in MATLAB* by Peter Corke, published by Springer in 2011.
The Bag-of-feature section is based on a presentation by Cordelia Schmid
`http://www.di.ens.fr/willow/events/cvml2011/materials/CVML2011_Cordelia_bof.pdf`

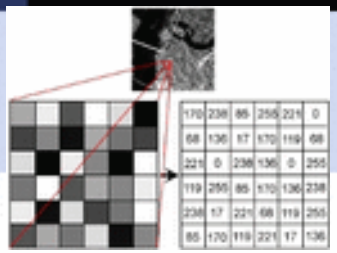
Related topics:
Marc Van Droogenbroeck's "Computer Vision"
and
Louis Wehenkel/Pierre Geurt's "Introduction to Machine Learning"

These slides are based on Chapter 13 of the book *Robotics, Vision and Control: Fundamental Algorithms in MATLAB* by Peter Corke, published by Springer in 2011.
The Bag-of-feature section is based on a presentation by Cordelia Schmid
`http://www.di.ens.fr/willow/events/cvml2011/materials/CVML2011_Cordelia_bof.pdf`

+ More recent topics:
  - End-to-end learning with tree-based methods and deep learning
  - The need for careful data collection for effective computer vision
  - Current developments at ULg and research topics

Related topics:
Marc Van Droogenbroeck's "Computer Vision"
and
Louis Wehenkel/Pierre Geurt's "Introduction to Machine Learning"

Motivation



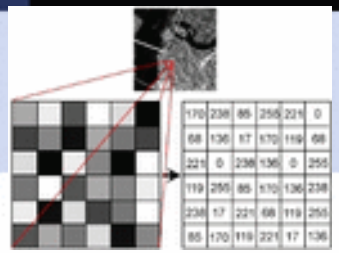
Raw images contain too much data to be of direct practical use for high(er)-level robot vision (object recognition, pose estimation, tracking, ...).

We need to reduce the dimensionality of raw image data, ideally focusing on

- discarding redundant information.
- extracting entities that are invariant to the conditions that typically change while a robot is working (viewpoint, illumination, ...).

Feature extraction is an information concentration step that reduces the data rate from 10^6–10^8 bytes s^{-1} at the output of a camera to something of the order of tens of features (vectors of a few dozen scalars) per frame that can be used as input to a robot's control system.

Motivation



Raw images contain too much data to be of direct practical use for high(er)-level robot vision (object recognition, pose estimation, tracking, ...).

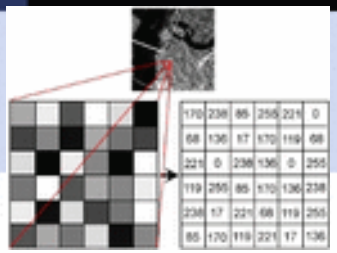
We need to reduce the dimensionality of raw image data, ideally focusing on

??? Deep learning argues it is not necessary anymore ???

- discarding redundant information.
- extracting entities that are invariant to the conditions that typically change while a robot is working (viewpoint, illumination, ...).

Feature extraction is an information concentration step that reduces the data rate from 10^6–10^8 bytes s^{-1} at the output of a camera to something of the order of tens of features (vectors of a few dozen scalars) per frame that can be used as input to a robot's control system.

Motivation



Raw images contain too much data to be of direct practical use for high(er)-level robot vision (object recognition, pose estimation, tracking, ...).

We need to reduce the dimensionality of raw image data, ideally focusing on

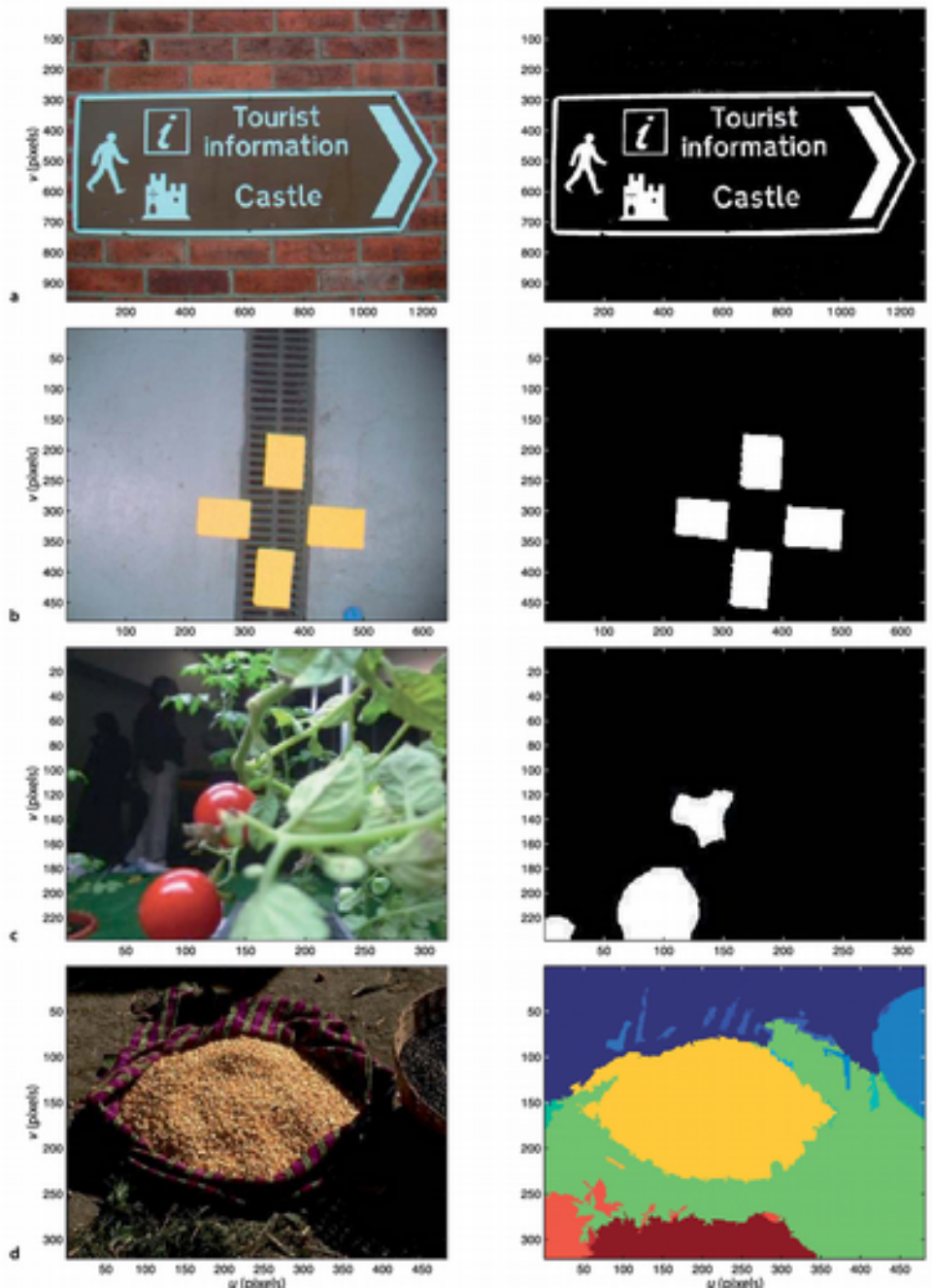
- discarding redundant information.
- extracting entities that are invariant to the conditions that typically change while a robot is working (viewpoint, illumination, ...).

??? Deep learning argues it is not necessary anymore ???
But then you still need impressive computational power

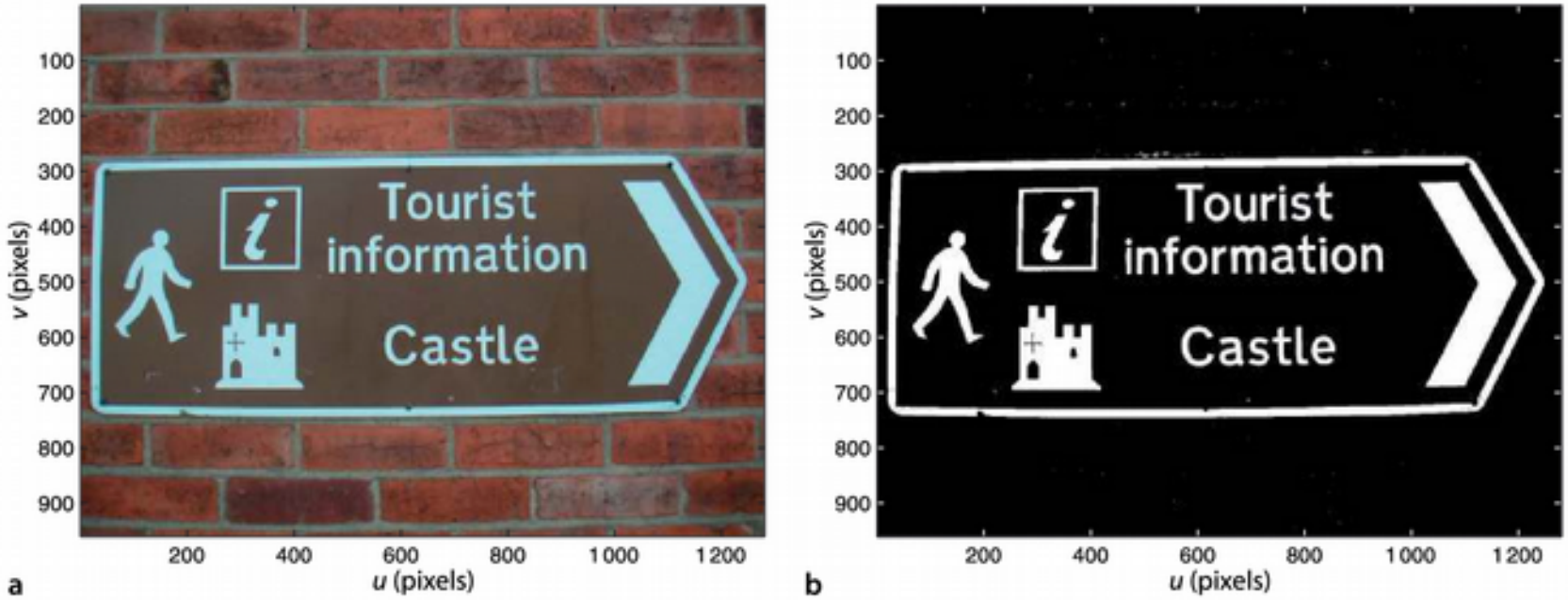
Feature extraction is an information concentration step that reduces the data rate from 10^6–10^8 bytes s^{-1} at the output of a camera to something of the order of tens of features (vectors of a few dozen scalars) per frame that can be used as input to a robot's control system.

Plan

Region Features

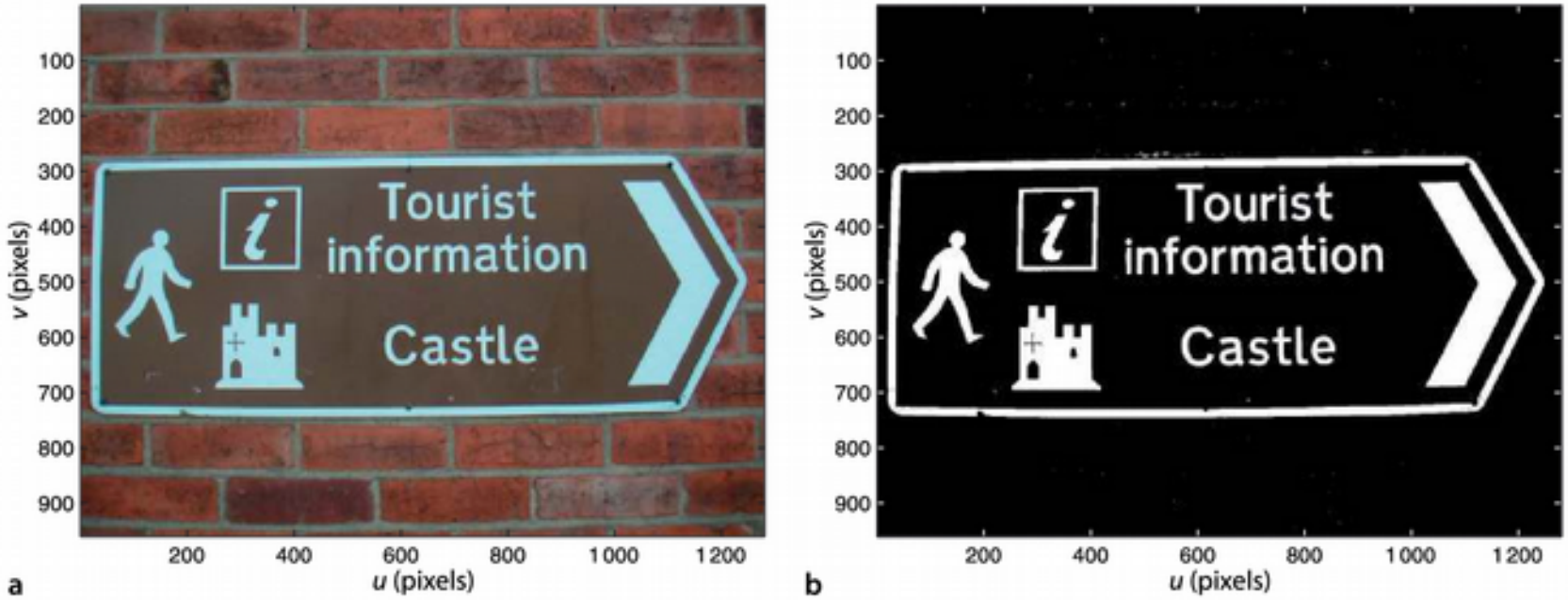


Thresholding

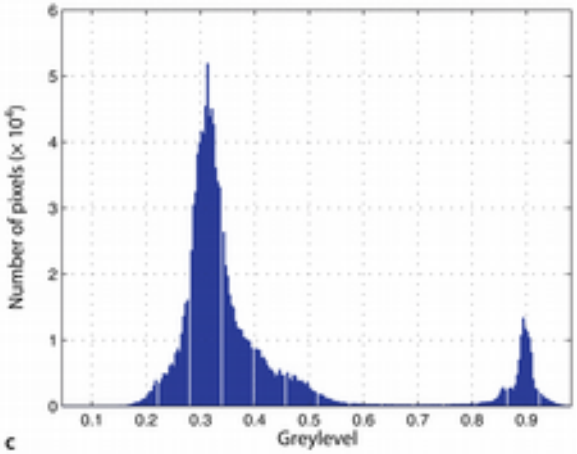


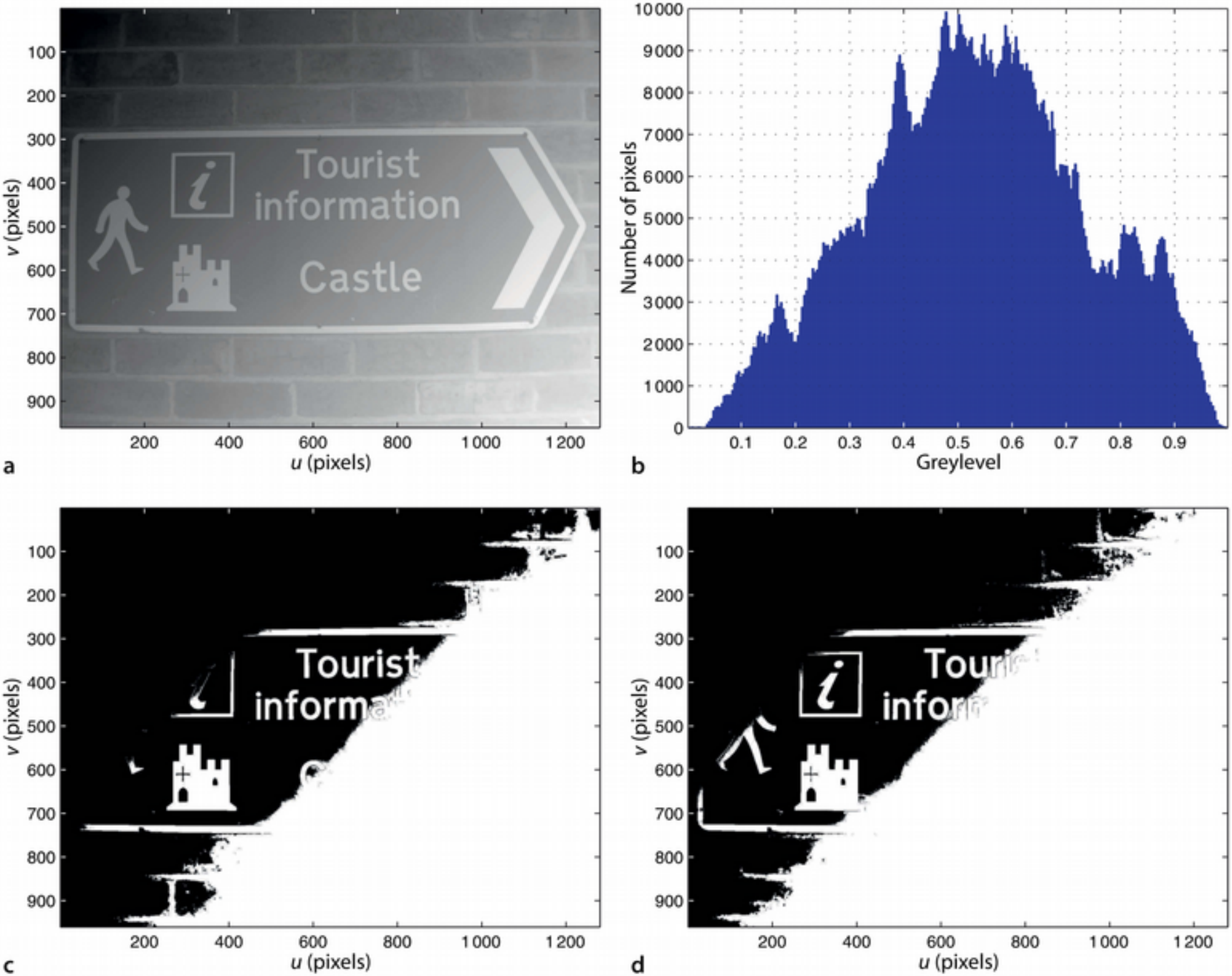
$$c[u,v] = \begin{cases} 0 & I[u,v] < t \\ 1 & I[u,v] \geq t \end{cases} \quad \forall (u,v) \in I$$

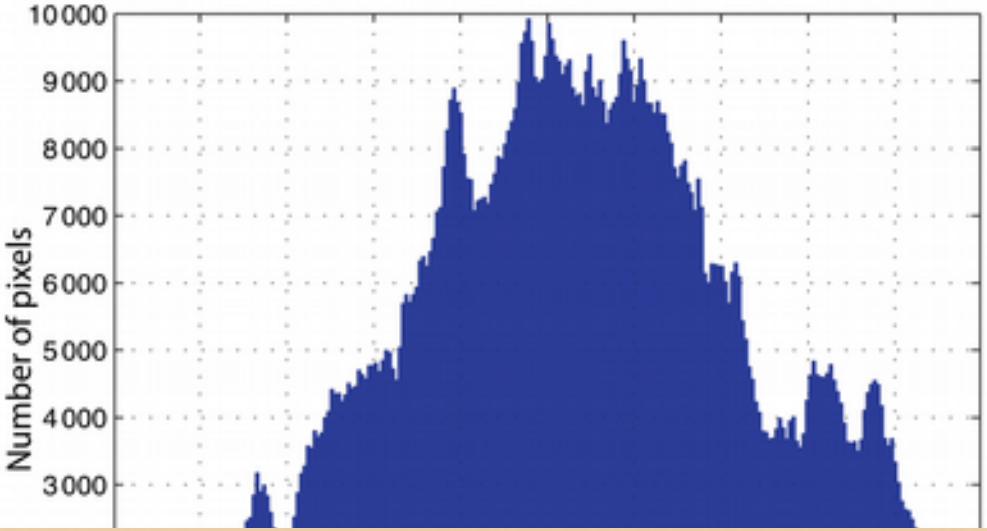
Thresholding



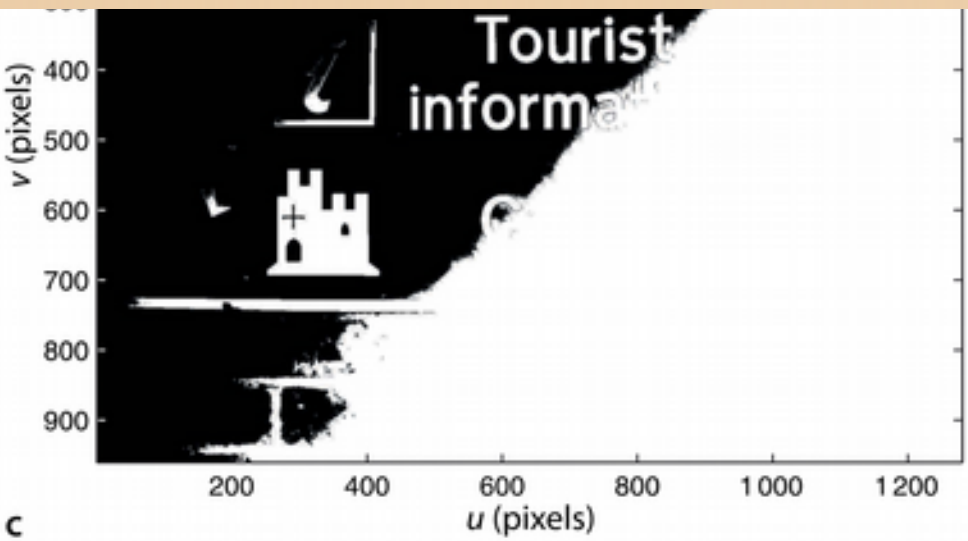
$$c[u,v] = \begin{cases} 0 & I[u,v] < t \\ 1 & I[u,v] \geq t \end{cases} \quad \forall (u,v) \in I$$

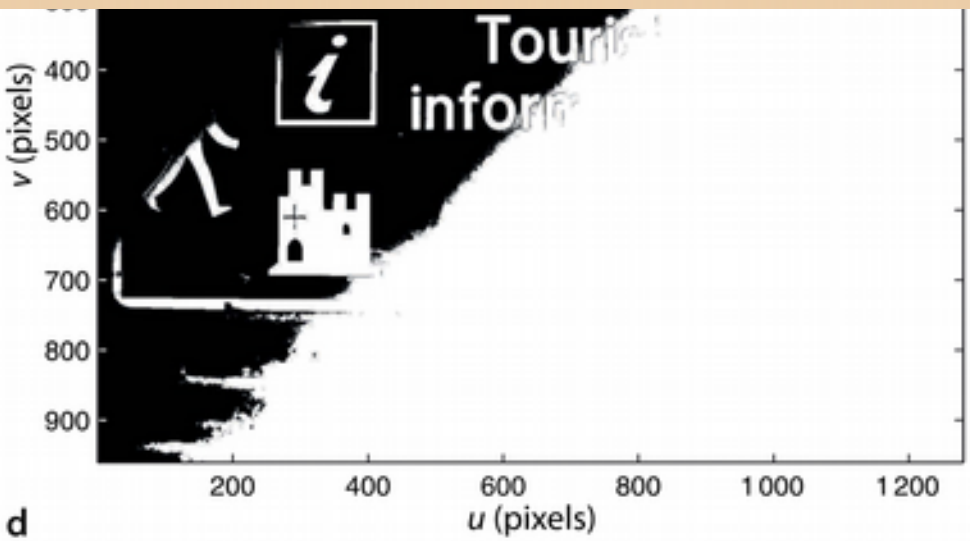



Tourist information
Castle
v (pixels)
u (pixels)
a
Number of pixels
Greylevel
b
c
d



Thresholding-based techniques are notoriously brittle – a slight change in illumination of the scene means that the thresholds we chose would no longer be appropriate. In most real scenes there is no simple mapping from pixel values to particular objects – we cannot for example choose a threshold that would select a motorbike or a duck. Distinguishing an object from the background remains a hard computer vision problem.





Thresholding-based techniques are notoriously brittle – a slight change in illumination of the scene means that the thresholds we chose would no longer be appropriate. In most real scenes there is no simple mapping from pixel values to particular objects – we cannot for example choose a threshold that would select a motorbike or a duck. Distinguishing an object from the background remains a hard computer vision problem.

Many thresholding alternatives:

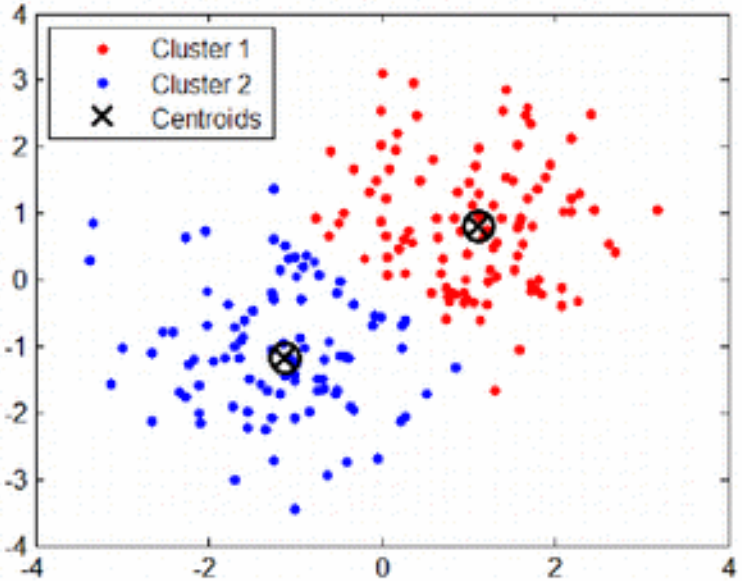
- Local thresholding
- K-means color thresholding
- Maximally stable regions
- …
-

The k-means Algorithm

Given a set of observations $(x_1, x_2, ..., x_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k sets $(k \leq n)$ $S = \{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS)

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2$$

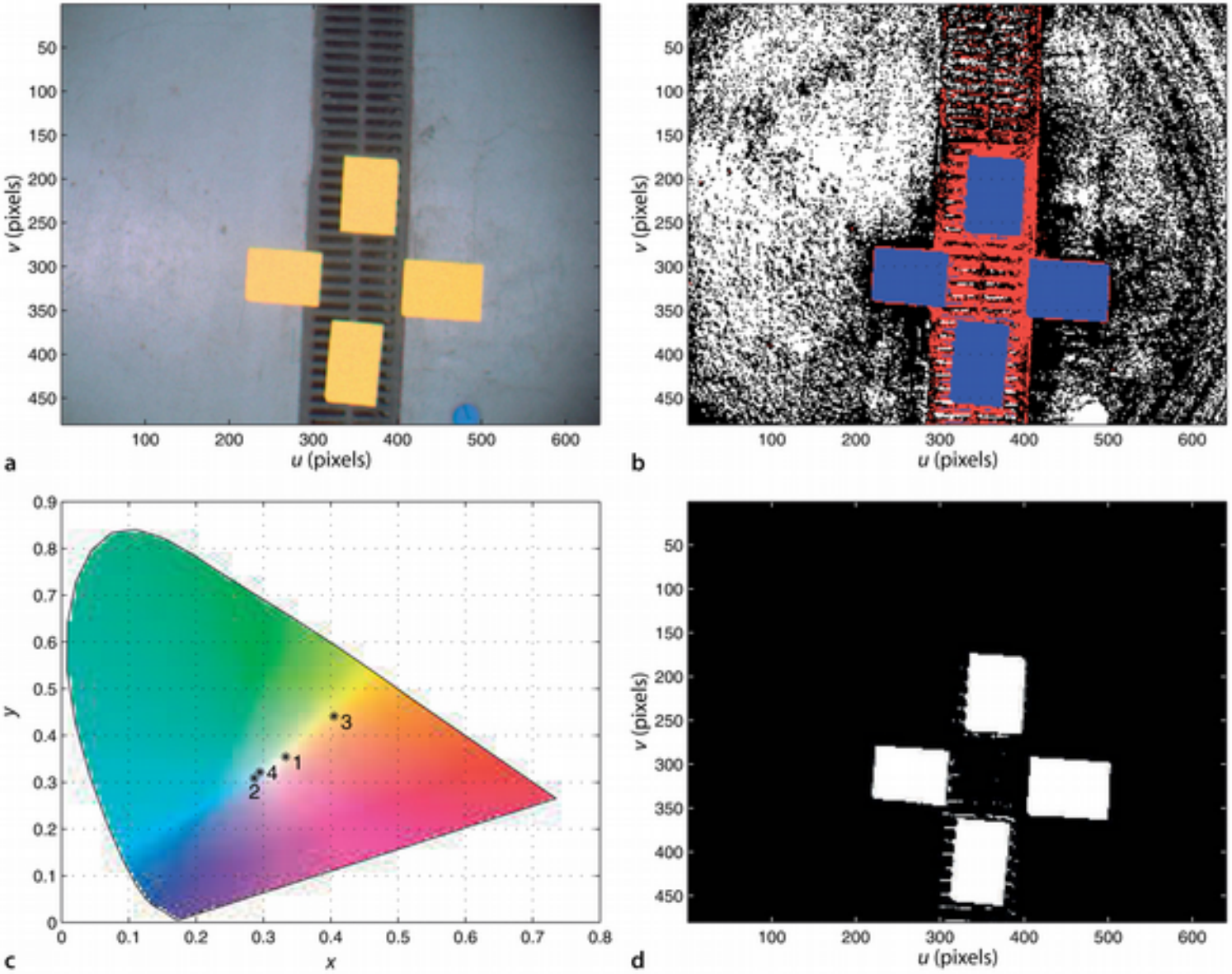
where μ_i is the mean of points in S_i.



(see “Introduction to machine learning” course)

Color Clustering and Classification

```
>> [cls, cxy, resid] = colorkmeans(im_targets, 4);
```

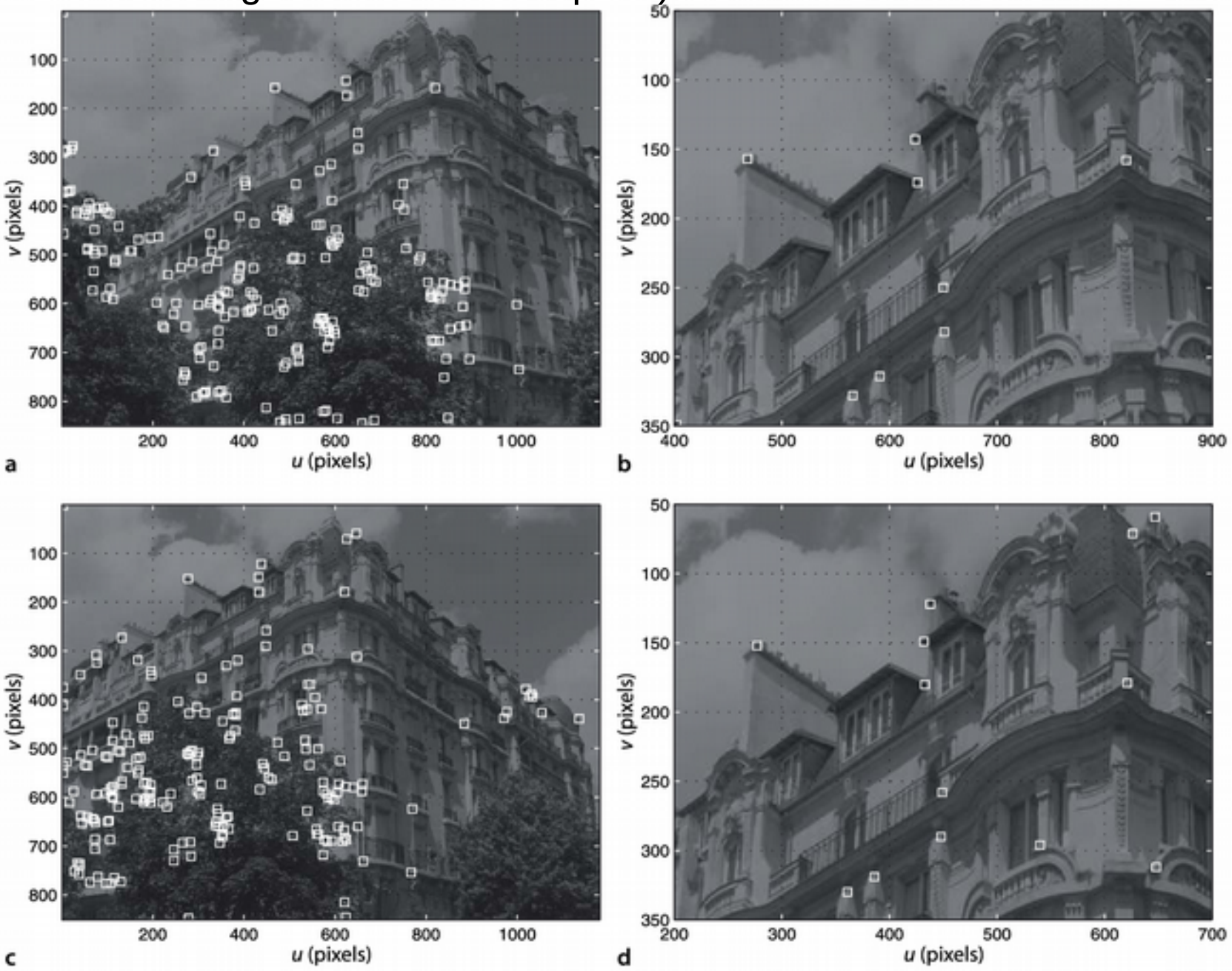


Other features (see chapter 11, Computer vision)

- Corner detectors
  - Harris, …
- Point features
  - SIFT
  - SURF
  - ORB, FREAK, FAST, …
- Line features
  - Hough transform
- Random
- Landmarks

Corner detection

(a point for which there are two dominant and different edge directions in a local neighbourhood of the point)

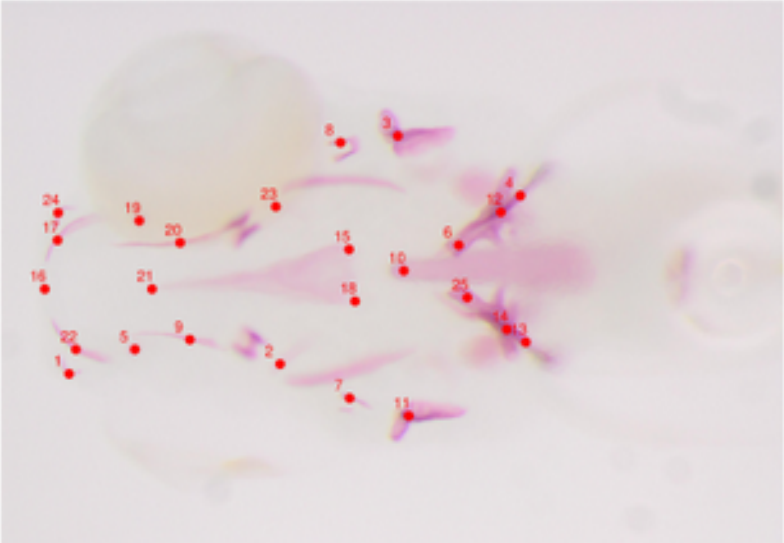


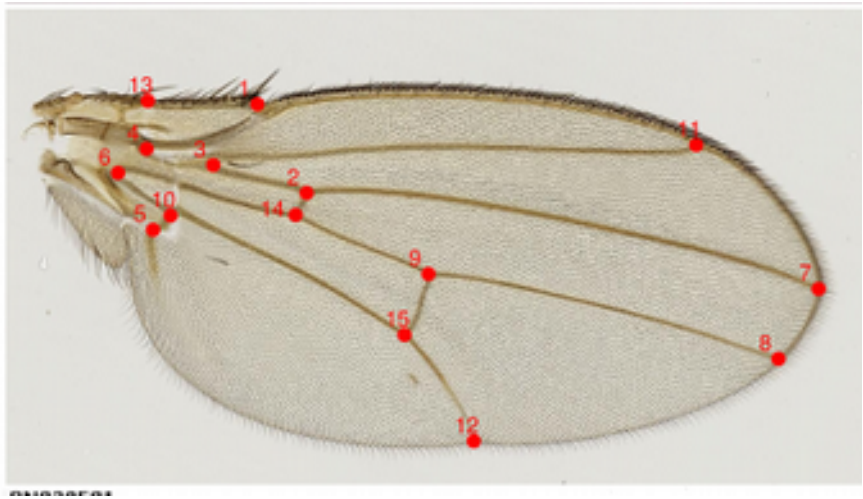
Point detection (FFME, SIFT, ORB, FAST)

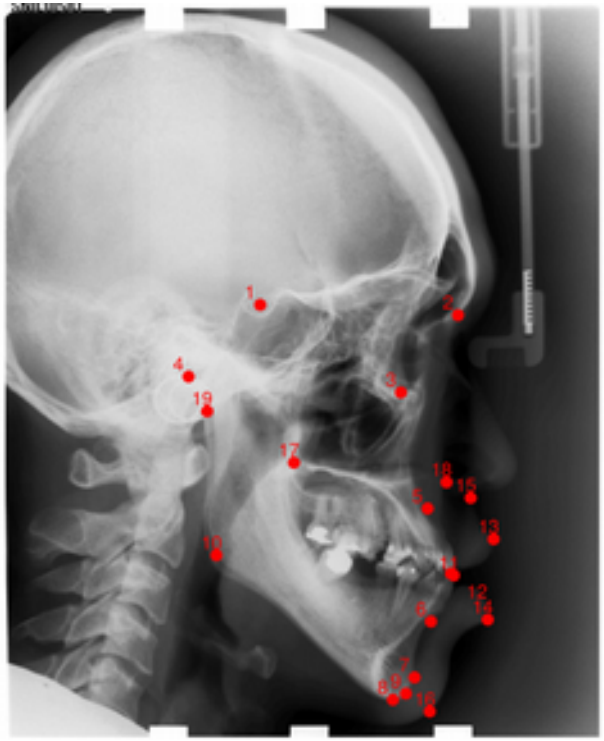
(Anatomical) Landmarks

ALIZARIN RED

- Anguloarticular down
- Anguloarticular up
- Anterior
- Background
- Branchiostegal ray 1 down
- Branchiostegal ray 1 up
- Branchiostegal ray 2 down
- Branchiostegal ray 2 up
- Ceratobranchial down
- Ceratobranchial up
- Ceratohyal down
- Ceratohyal up
- Cleithrum down
- Cleithrum up
- Dentary down
- Dentary up
- Entopterygoid down
- Entopterygoid up
- Eye down
- Eye up
- Hyomandibular down
- Hyomandibular up
- Maxilla down
- Maxilla up
- Notochord
- Opercul down
- Opercul up
- Parasphenoid a
- Parasphenoid b
- Parasphenoid c
- Superposition down
- Superposition up







Object/Category Recognition

Image classification: assigning a class label to the image

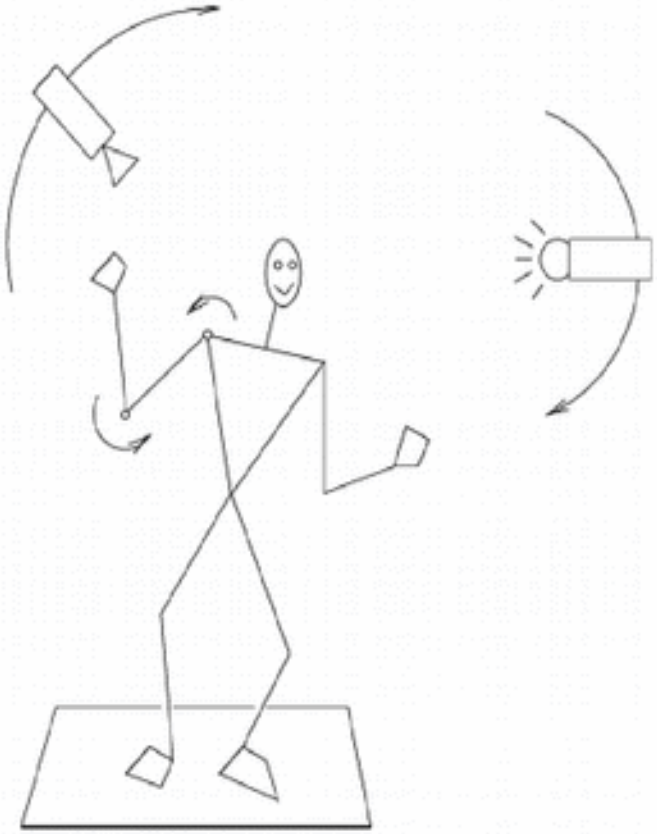
Car: present
Cow: present
Bike: not present
Horse: not present
...

Object localization: define the location and the category

Location

Category

Challenge 1: Intra-instance Variations



Viewpoint, illumination, kinematic configuration, ...

Challenge 2: Intra-class Variations

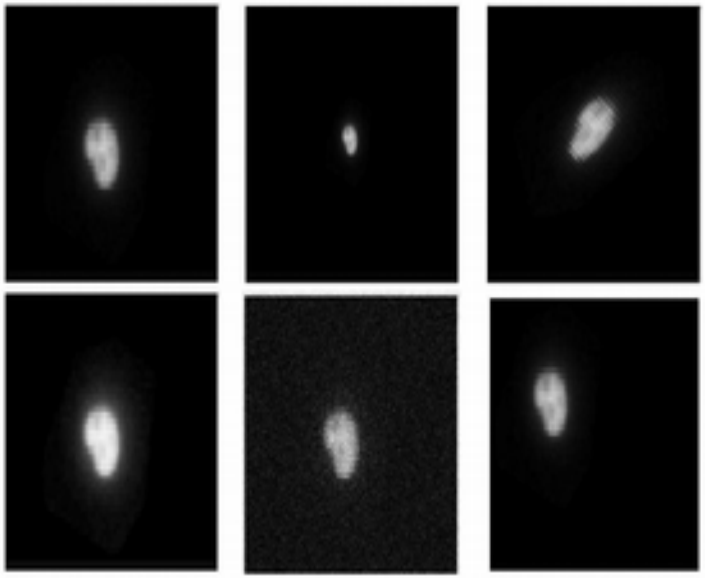





Same challenges for “intelligent microscopes”

Image acquistion conditions can not be fully controlled

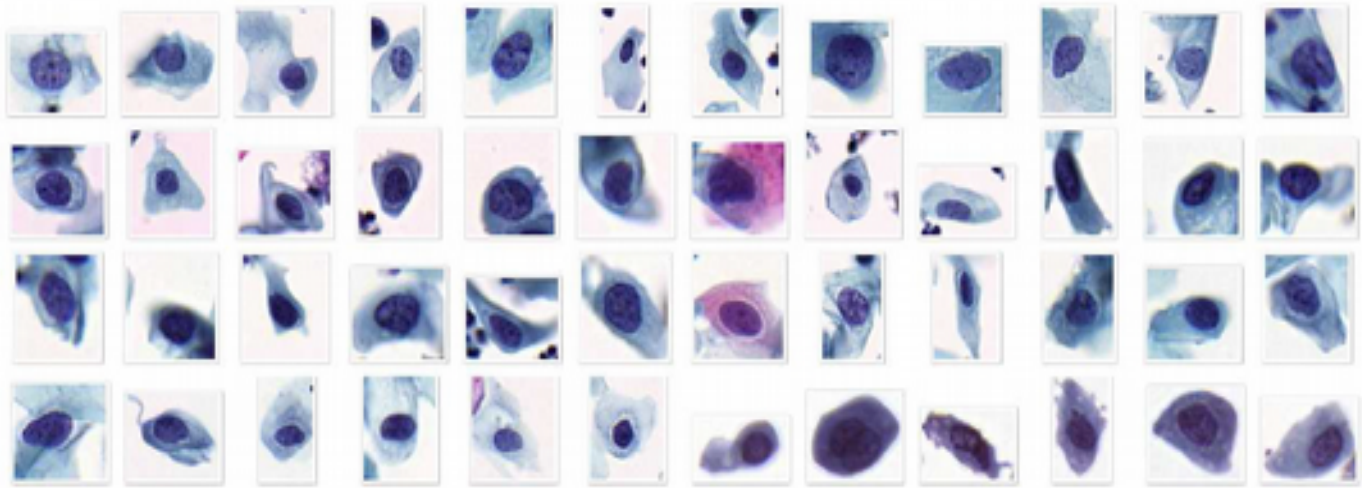
- Illumination, viewpoint, position, scale changes
- Noise, cluttered background, occlusions
- Staining, imaging kits, microscopes, ...



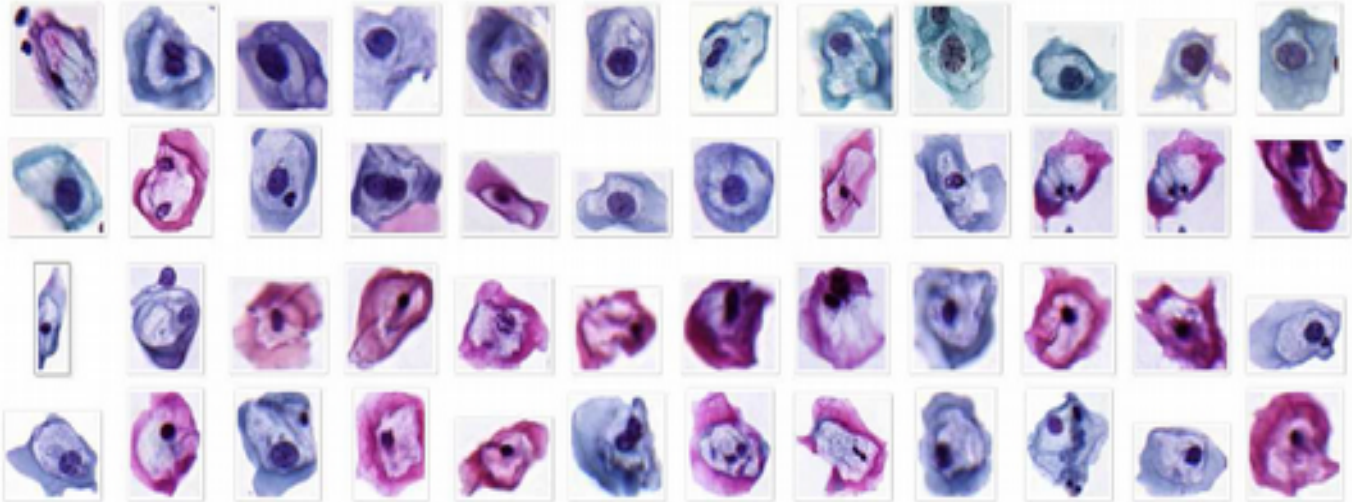
These simple variations yield different image matrices

Same challenges for “intelligent microscopes”

Intra-class / inter-class variations



vs

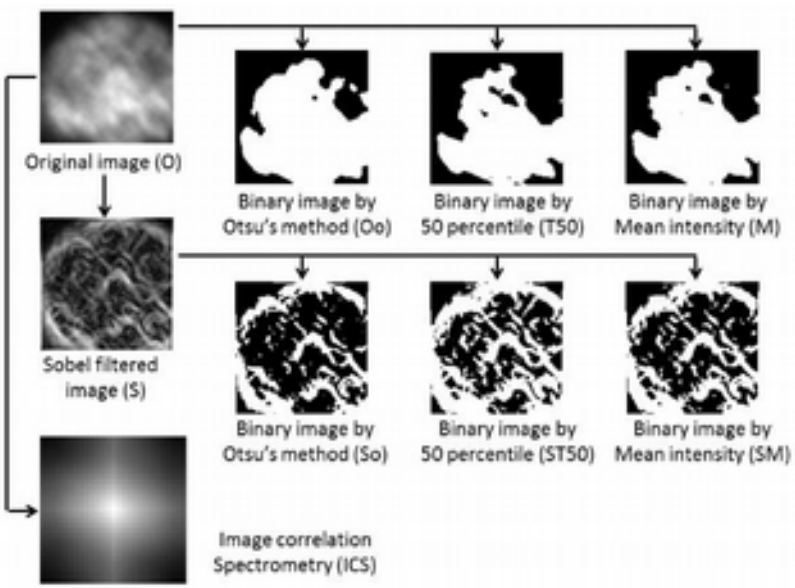


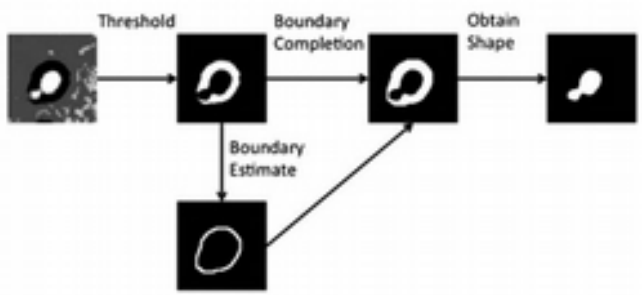
Computer vision approaches

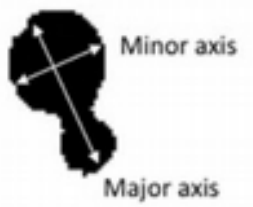
- Traditional : hand-crafted, specific, features +learning

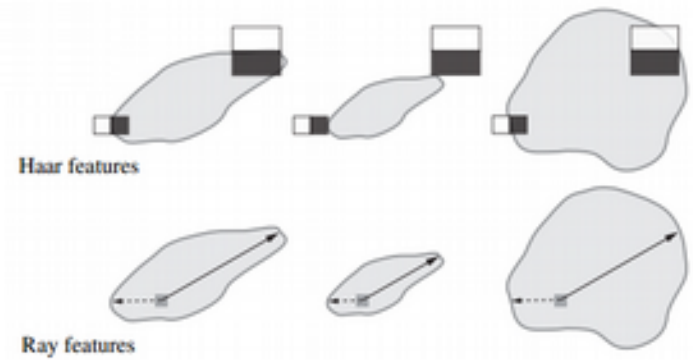
- *Hypothesis : the researcher is very imaginative, and smart*
- *Pros : exploitation of domain knowledge*
- *Cons : need to be adapted when the problem changes*

researchers are indeed imaginative
limited evaluation
} which features to choose ?









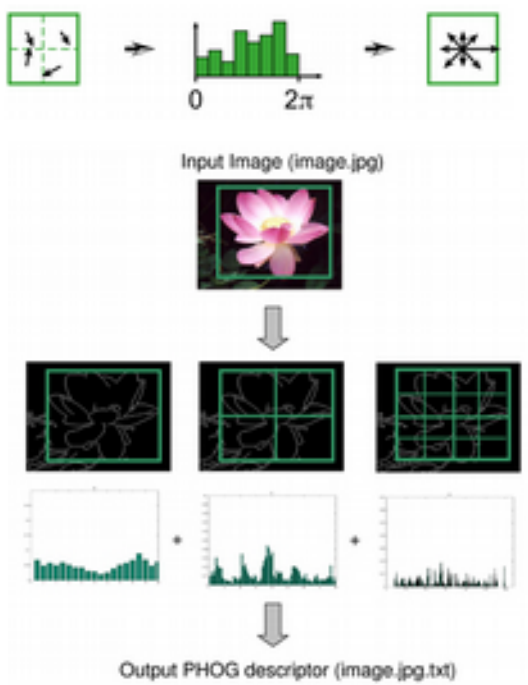
Computer vision approaches

- Traditional : hand-crafted, specific, features +learning

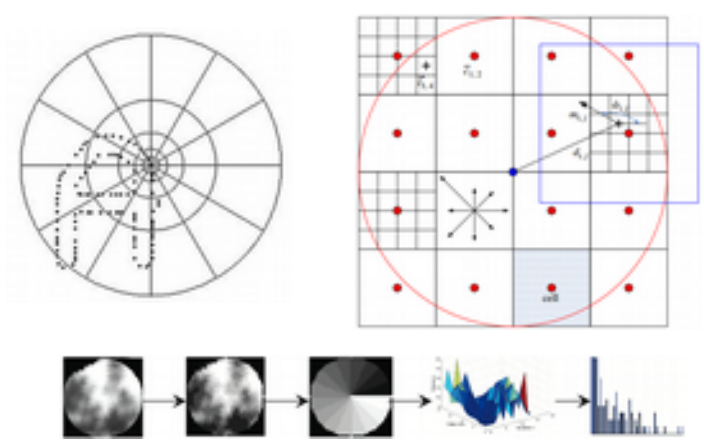
- *Hypothesis : the researcher is very imaginative, and smart*
- *Pros : exploitation of domain knowledge*
- *Cons : need to be adapted when the problem changes*

researchers are indeed imaginative
limited evaluation

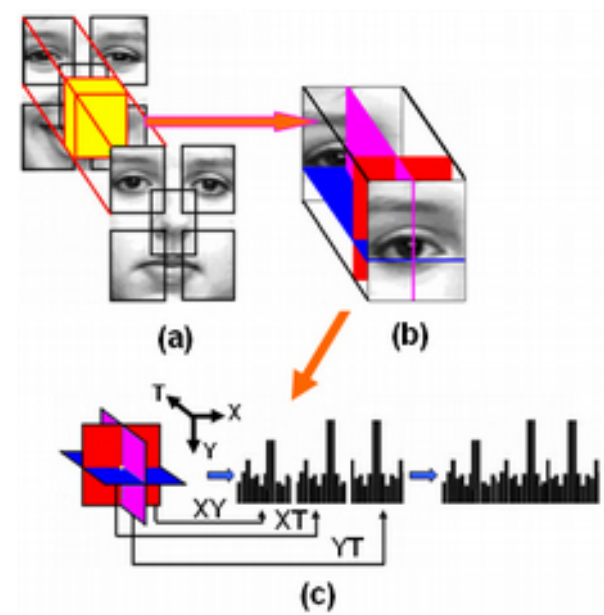
} which features to choose ?



Harris-Affine, Hessian-Affine, EBR, IBR, MSER, SFOP,DAISY, GIST, GLOH, LBP, OSID, PHOG, PHOW, SIFT, RIFT, PCA-SIFT, Spin Image, SURF, VLAD, Shape contexts, Textons, ...



Li & Allison, Neurocomputing 2008

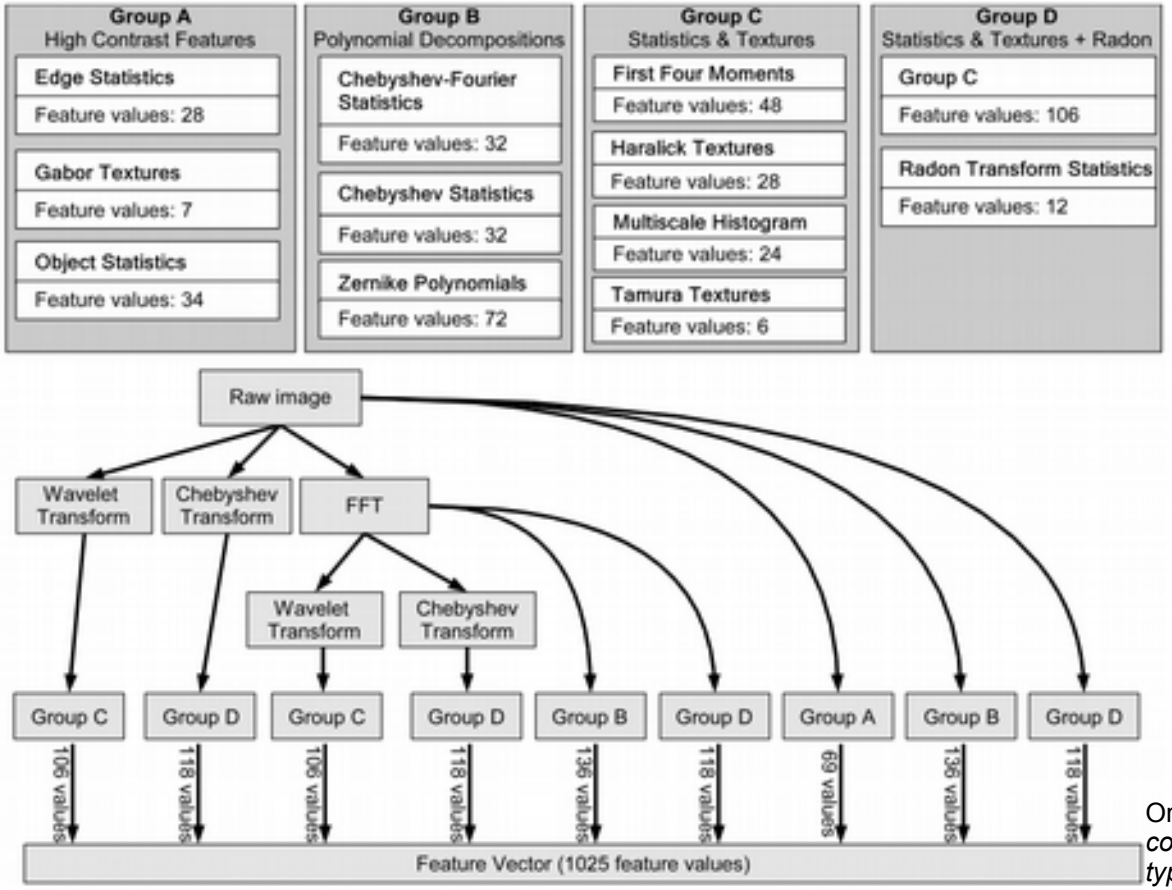


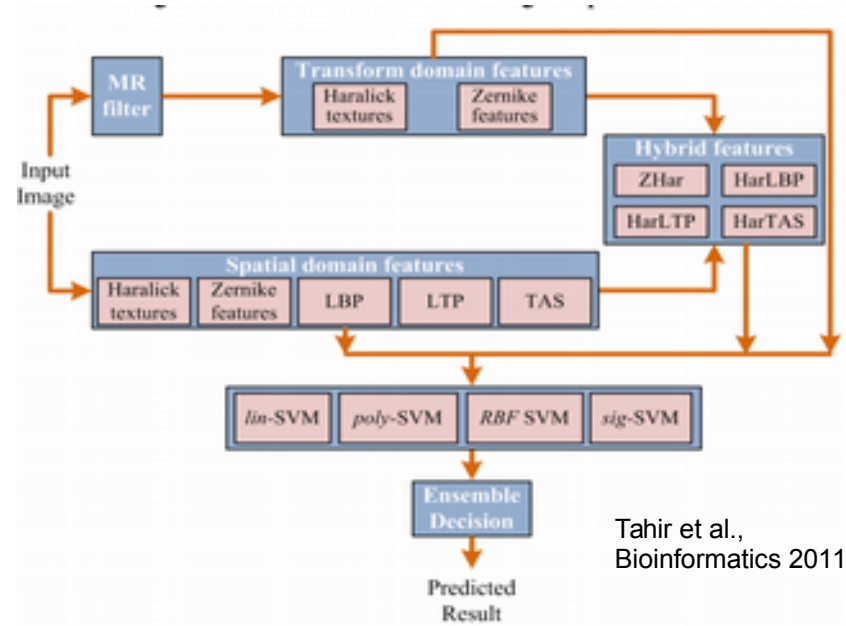
Scholarpedia

Computer vision approaches

- *Recent : Combine many features + learning*

- *Hypothesis : the good features should be among them*
- *Pros : take advantage of previous research efforts*
- *Cons : computationally intensive*





Tahir et al., Bioinformatics 2011

Orlov et al., Pattern Recognition letters, 2008 : « *...poor performance in terms of computational complexity, making this method unsuitable for real-time or other types of applications in which speed is a primary concern.* »

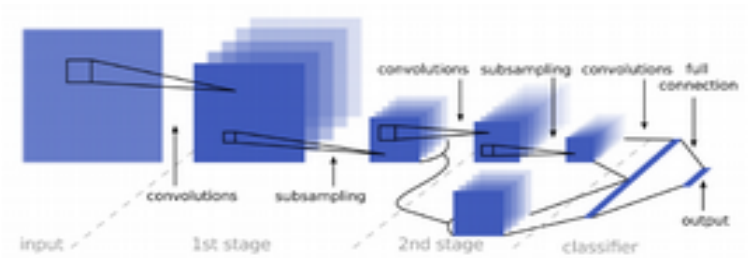
Computer vision approaches

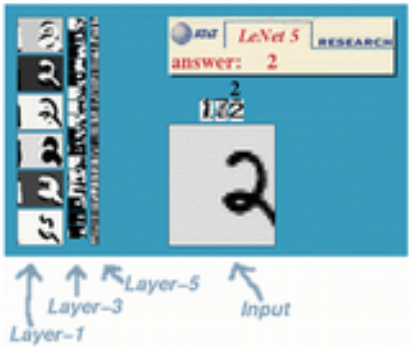
- *Generic : « end-to-end » learning*

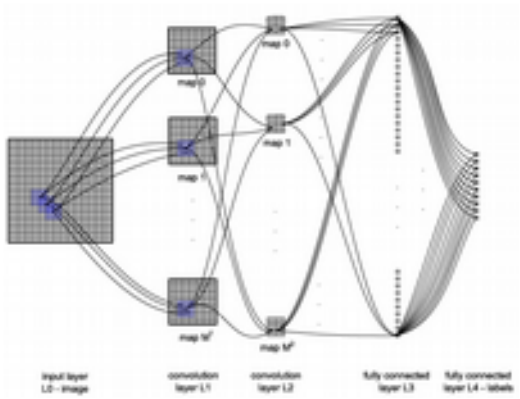
- *Hypothesis : human brain learn from raw data, let's design such an algorithm*
- *Pros : it should work on everything with minimal tuning*
- *Cons : <> architectures*

  many parameters to optimize: need large training data, time-consuming

  does it work ? Is it generic ?







Lecun et al. 1989..., Hinton et al., Ciresan et al. (GPU) 2011

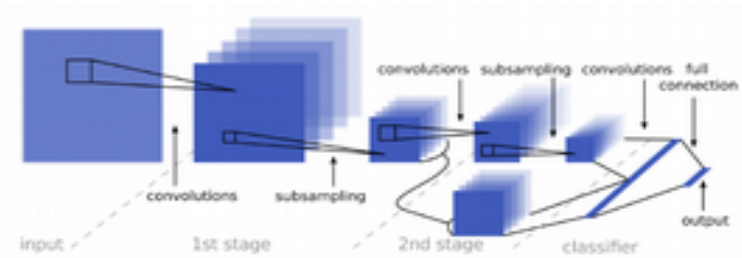
Computer vision approaches

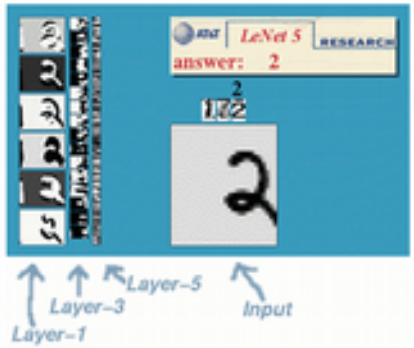
- *Generic : « end-to-end » learning*

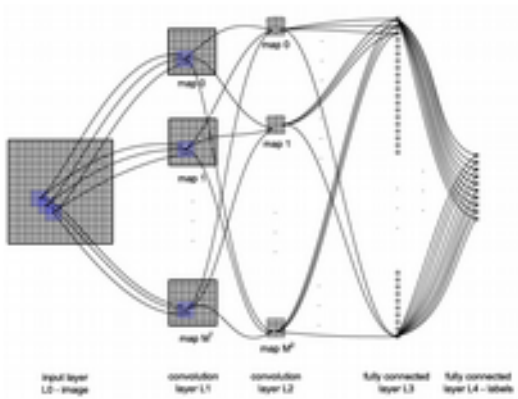
- *Hypothesis : human brain learn from raw data, let's design such an algorithm*
- *Pros : it should work on everything with minimal tuning*
- *Cons : <> architectures*

  many parameters to optimize: need large training data, time-consuming

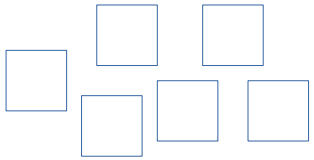
  does it work ? Is it generic ?

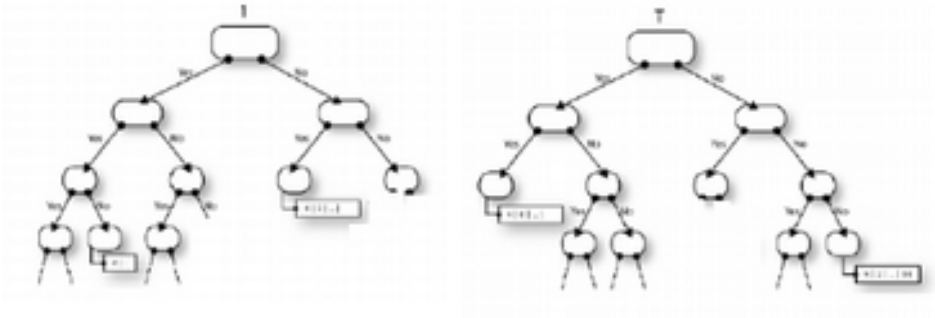






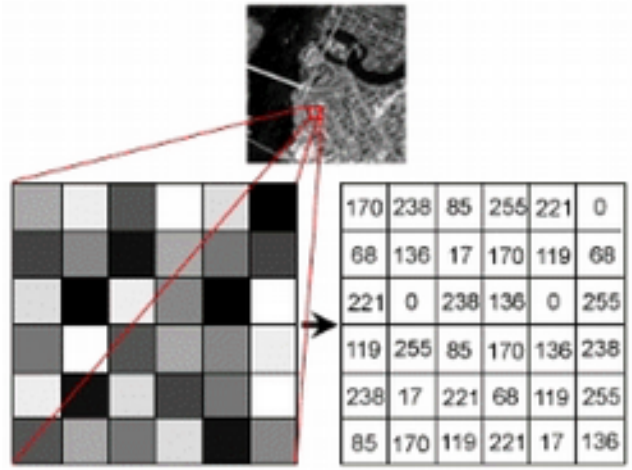
Lecun et al. 1989..., Hinton et al., Ciresan et al. (GPU) 2011





Marée, Geurts, Wehenkel, et al. 2003 ...

Direct application of decision trees on images

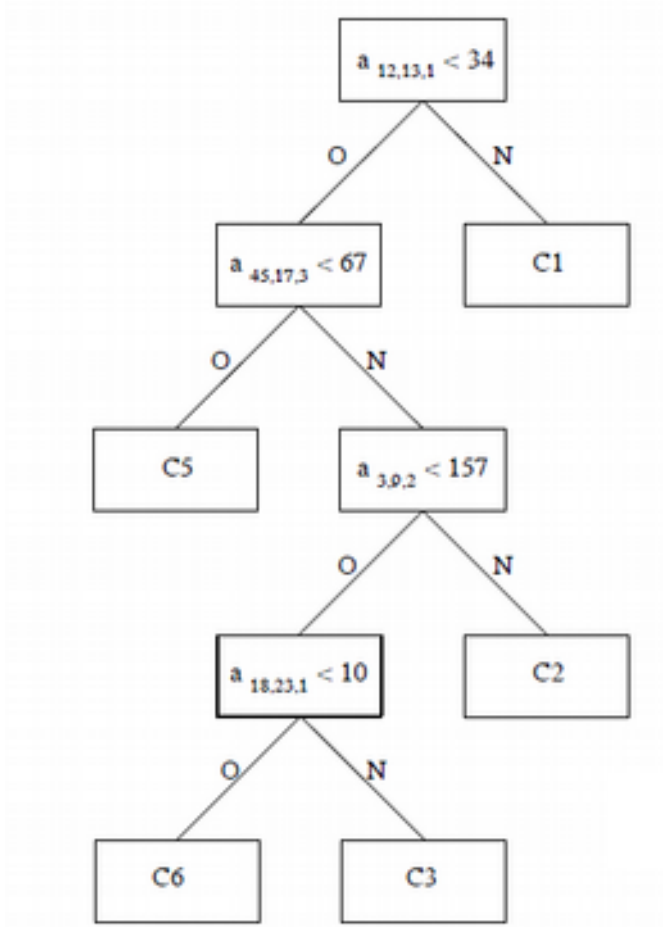


170	238	85	255	221	0
68	136	17	170	119	68
221	0	238	136	0	255
119	255	85	170	136	238
238	17	221	68	119	255
85	170	119	221	17	136

- LEARNING :

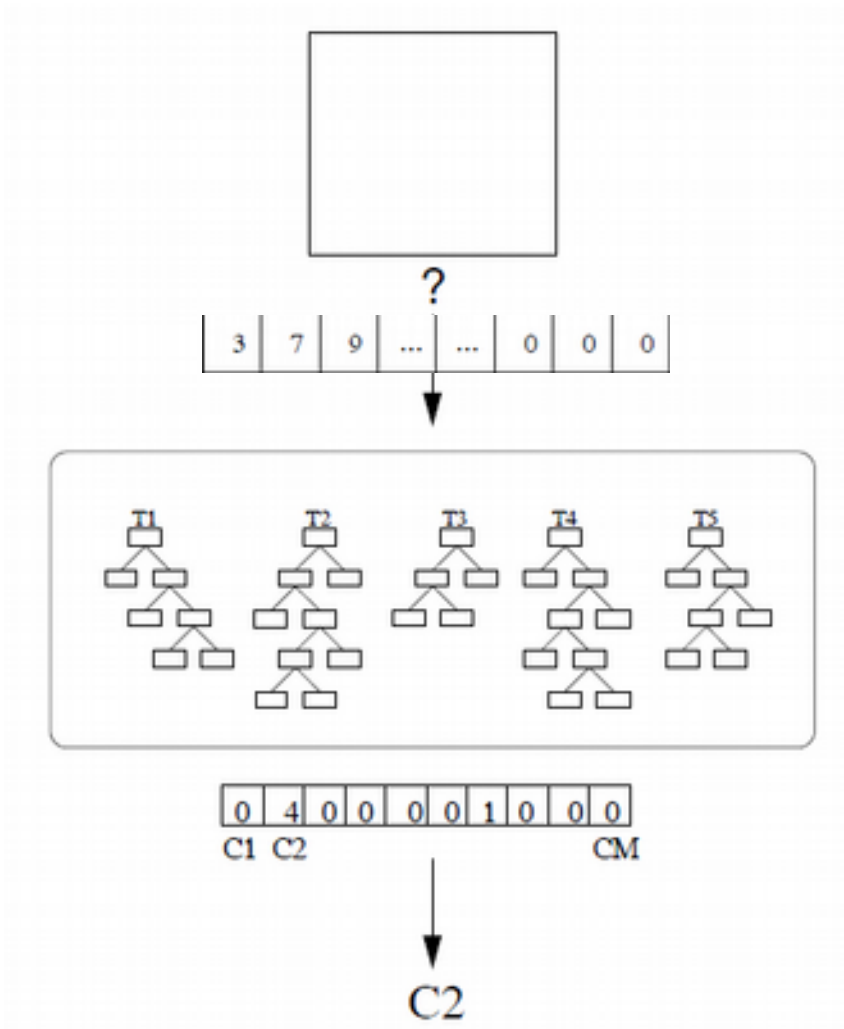
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	...	Classe
Image1	60	19	18	17	0	1	1	...	C1
Image2	60	3	22	23	1	29	11	...	C1
	75	9	2	1	3	77	46	...	C1
	2	10	10	2	234	0	0	...	C2
	3	7	9	18	5	0	0	...	C2
	2	14	5	10	8	10	8	...	C3
	65	3	20	21	2	0	1	...	?

Decision tree learning



Direct application of decision trees on images

- PREDICTION :



Is direct application of ML on structured inputs efficient ?

- Inputs:
  - a grey intensity [0,255] for each pixel
  - each image is represented by a vector of pixel intensities
  - eg.: 32x32=1024 dimensions
- Output:
  - 9 discrete values
  - Y={0,1,2,...,9}

Méthode	Taux d'erreur
Arbre de décision	11.5%
3 Plus proches voisins	5.66%
Bagging ($T = 50$)	4.42%
Arbres aléatoires ($T = 100$)	3.17%
Random Forests ($T = 100$)	3.0%
Boosting ($T = 50$)	2.29%
SVMs (poly2)	1.95%

With 50000 training images
Evaluated on 10000 test images

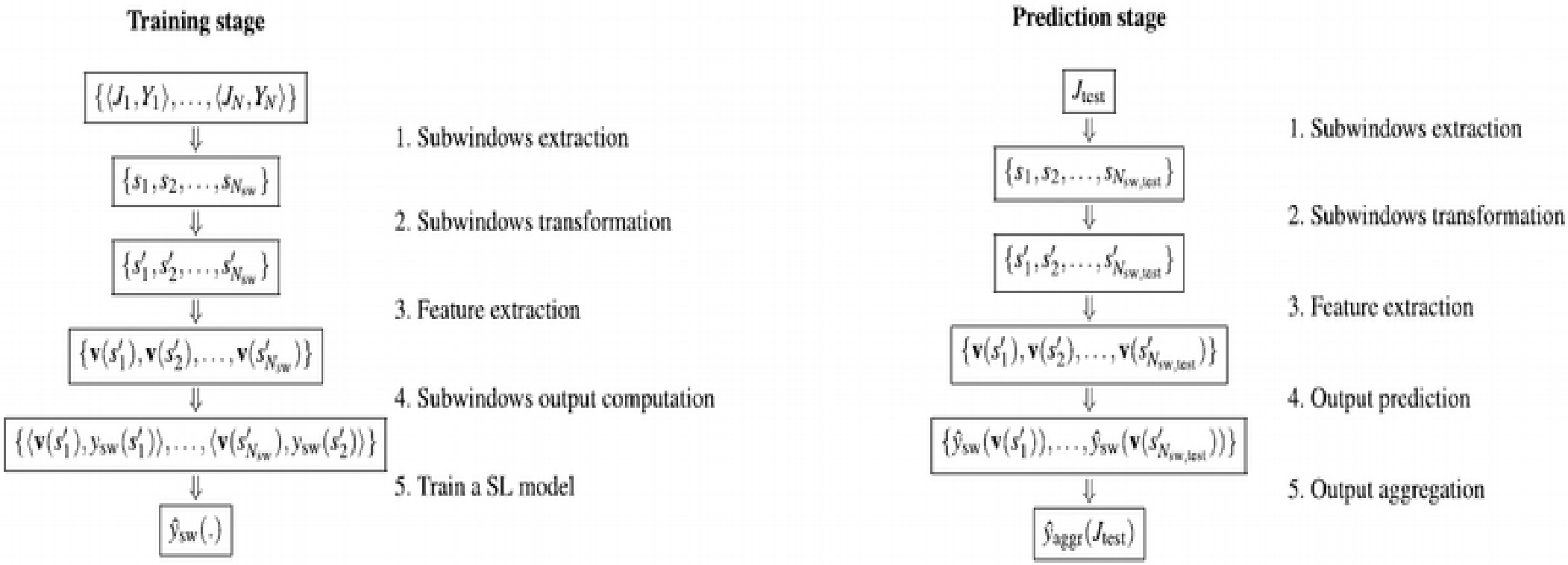
Is direct application of DT on structured inputs efficient ?

e.g. : texture classification

Méthode	Taux d'erreur
Arbre de décision	89.35%
Plus proche voisin	80.79%
Bagging ($T = 50$)	73.15%
SVMs (linéaire)	71.99%
Boosting ($T = 50$)	69.44%
Random Forests ($T = 1000$)	66.90%
Arbres aléatoires ($T = 500$)	65.05%

- Inputs:
  - Rgb color intensities [0,255] for each pixel
  - each image is represented by a vector of pixel intensities
  - eg.: 32x32x3=3072 dimensions
- Output:
  - 40 discrete values
  - Y={0,1,2,...,40}

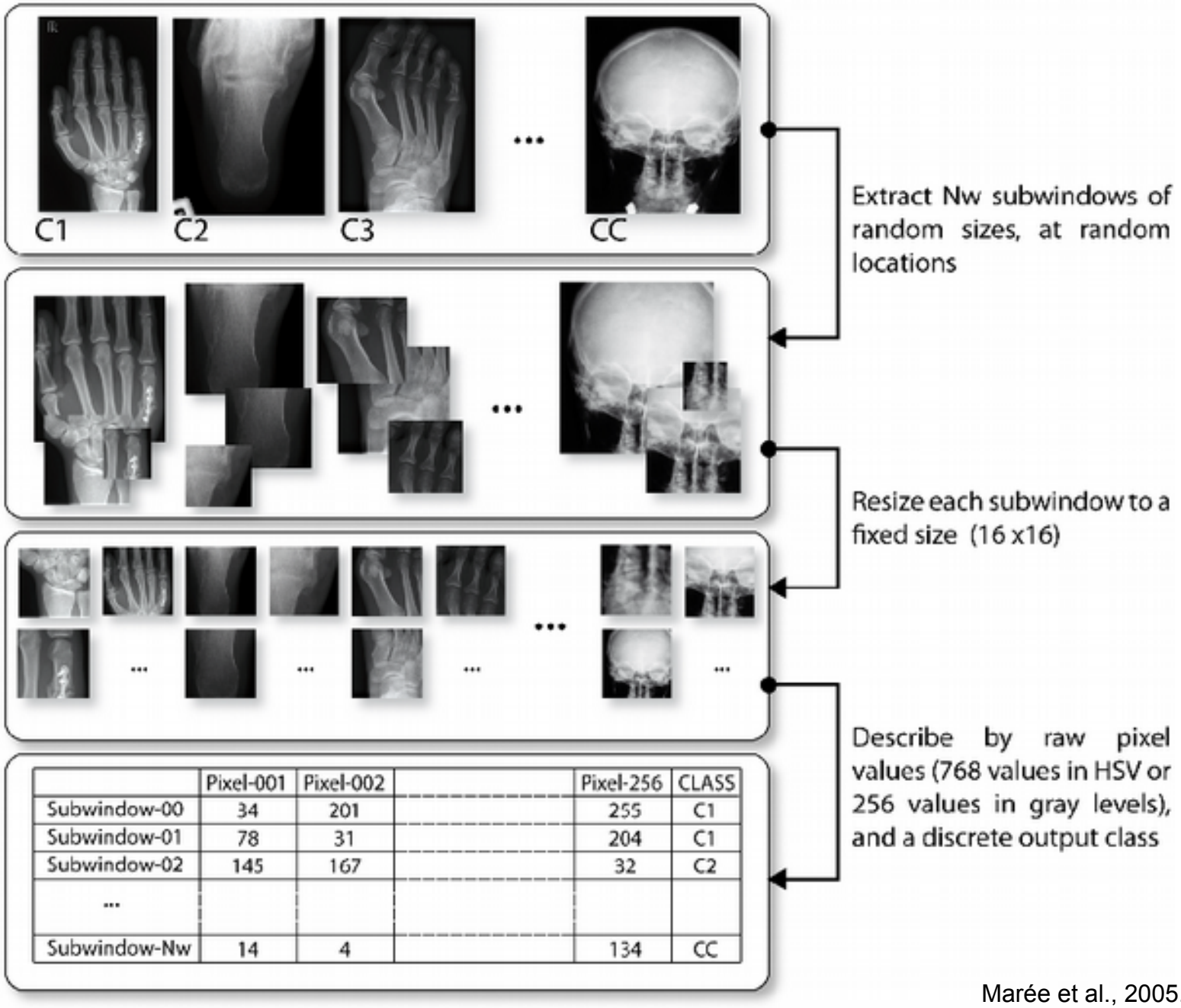
Segment & Combine / Random Subwindows & Extra-Trees : a common framework for classification, segmentation, interest point detection, and retrieval



Chapter 9 (Part II)

Extremely Randomized Trees and Random Subwindows for Image Classification, Annotation, and Retrieval
R. Marée, L. Wehenkel, and P. Geurts

Extraction of Random Subwindows in the whole training set of images

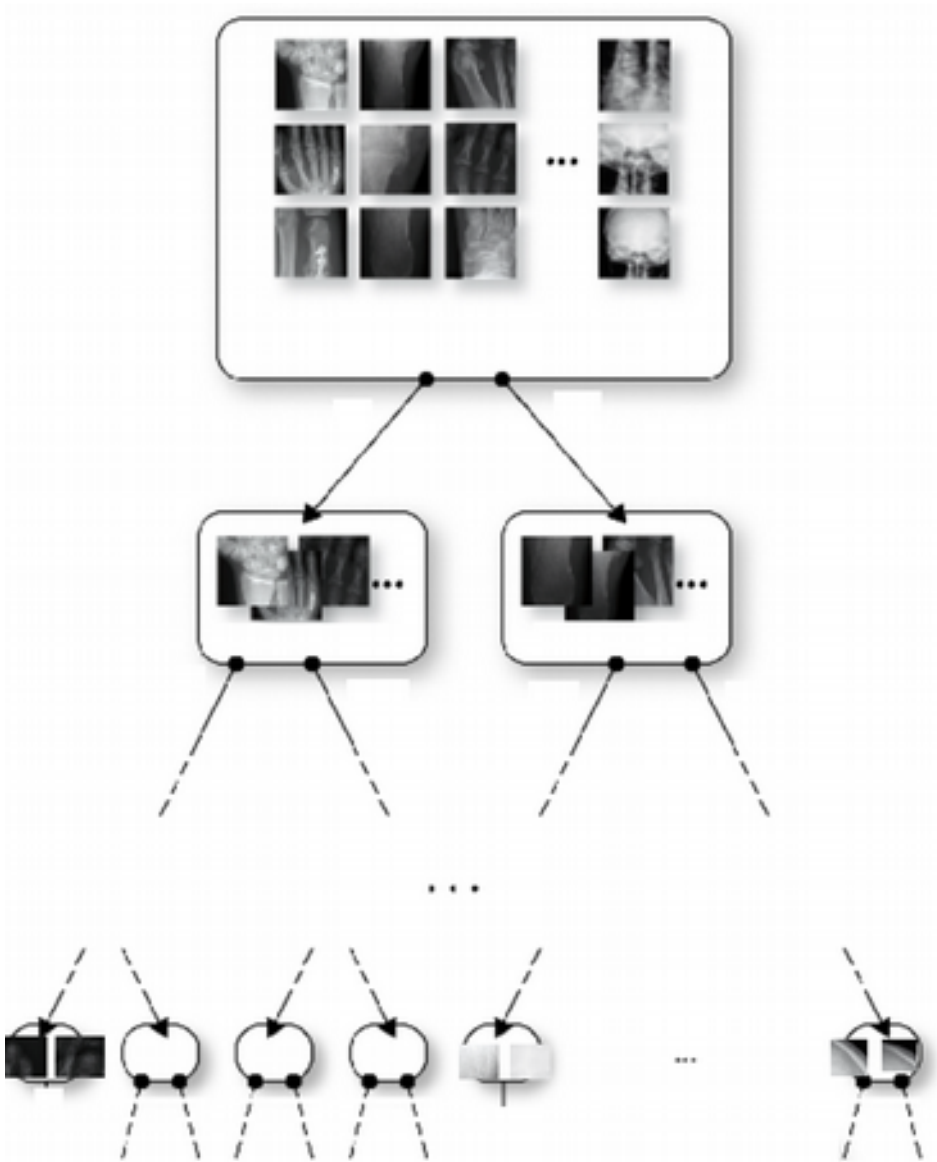


	Pixel-001	Pixel-002		Pixel-256	CLASS
Subwindow-00	34	201		255	C1
Subwindow-01	78	31		204	C1
Subwindow-02	145	167		32	C2
...					
Subwindow-Nw	14	4		134	CC

Parameters :
Nsw = nb subwindows
MinSize = [0%-100%]
MaxSize = [0%-100%]
Resize = 16x16
Colorspace = HSV/GRAY

Marée et al., 2005...

Extra-Trees for Direct Classification : single tree training



Extra-Trees for Direct Classification : single tree training

Top node of the tree with sample S of subwindows (e.g. 1M) extracted from all training images

	Pixel-001	Pixel-002		Pixel-256	CLASS
Subwindow-00	34	201		255	C1
Subwindow-01	78	31		204	C1
Subwindow-02	145	167		32	C2
...					
Subwindow-Nw	14	4		134	CC

- select randomly K attributes $\{a_1, \ldots, a_K\}$ among all non constant (in S) candidate attributes;
- Draw K splits $\{s_1, \ldots, s_K\}$, where $s_i = \textbf{Pick\_a\_random\_split}(S, a_i), \forall i = 1, \ldots, K$
- Return a split s_i such that $\text{Score}(s_i, S) = \max_{j=1,\ldots,K} \text{Score}(s_j, S)$.

K { Pixel-018 > 24
Pixel-123 > 17
Pixel-057 > 213
...
Pixel-202 > 77

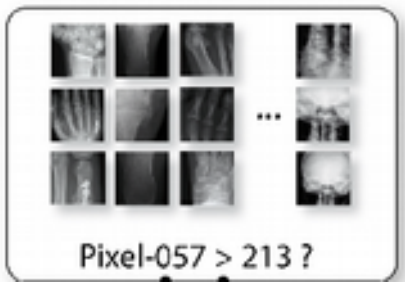
(e.g. logarithmic or Shannon entropy)

Pick_a_random_split(S,a)

Inputs: a subset S and an attribute a

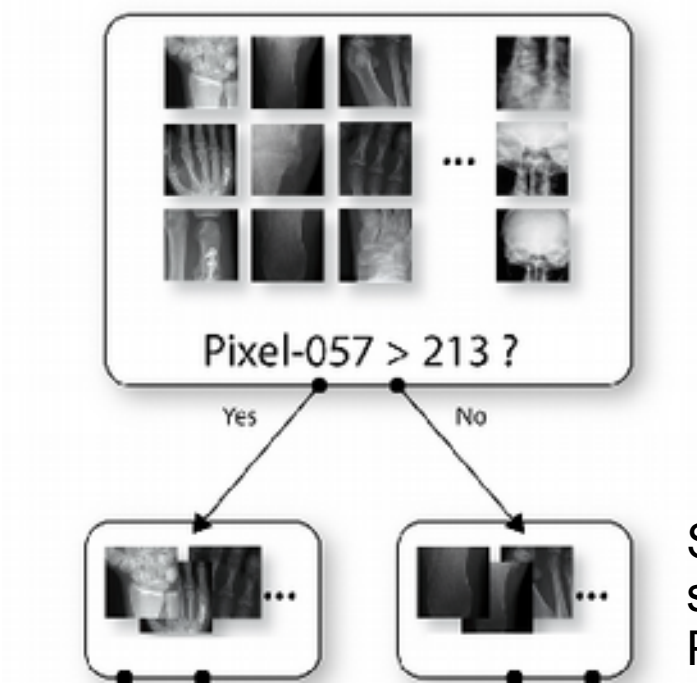
Output: a split

- Let $a^S_{\max}$ and $a^S_{\min}$ denote the maximal and minimal value of a in S;
- Draw a random cut-point a_c uniformly in $]a^S_{\min}, a^S_{\max}]$;
- Return the split $[a < a_c]$.



Marée et al., CVPR 2005...

Extra-Trees for Direct Classification : single tree training



Sample S of subwindows (e.g. 1M) extracted from all training images

Subsample S' of subwindows where Pixel_057 > 213

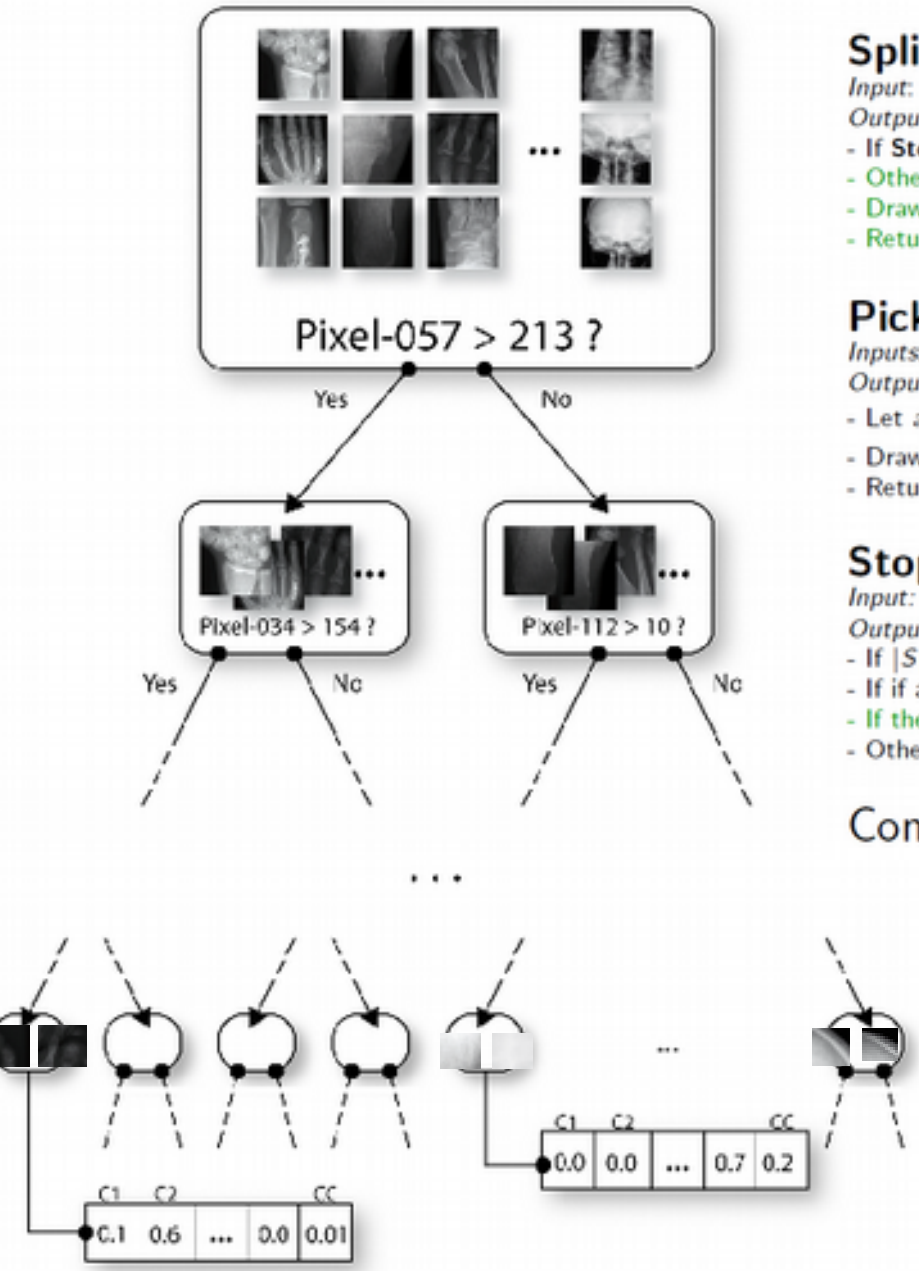
Subsample S'' of subwindows where Pixel_057 <= 213

Stop_split(S)
Input: a subset S
Output: a boolean
- If $|S| < n_{\min}$, then return TRUE;
- If if all attributes are constant in S, then return TRUE;
- If the output is constant in S, then return TRUE;
- Otherwise, return FALSE.

If **Stop_split**(S) is TRUE then attach predictions (ET-DIC)
Otherwise select randomly K attributes .....

Marée et al., CVPR 2005...

Extra-Trees for Direct Classification : single tree training



Split_a_node(S)

Input: the local learning subset S corresponding to the node we want to split
Output: a split $[a < a_c]$ or leaf
- If **Stop_split**(S) is TRUE then attach predictions (ET-DIC) or nothing (ET-BOF).
- Otherwise select randomly K attributes $\{a_1, \ldots, a_K\}$ among all non constant (in S) candidate attributes;
- Draw K splits $\{s_1, \ldots, s_K\}$, where $s_i = \textbf{Pick\_a\_random\_split}(S, a_i), \forall i = 1, \ldots, K$;
- Return a split s_i such that $\text{Score}(s_i, S) = \max_{j=1,\ldots,K} \text{Score}(s_j, S)$.

Pick_a_random_split(S,a)

Inputs: a subset S and an attribute a
Output: a split
- Let $a^S_{\max}$ and $a^S_{\min}$ denote the maximal and minimal value of a in S;
- Draw a random cut-point a_c uniformly in $]a^S_{\min}, a^S_{\max}]$;
- Return the split $[a < a_c]$.

Stop_split(S)

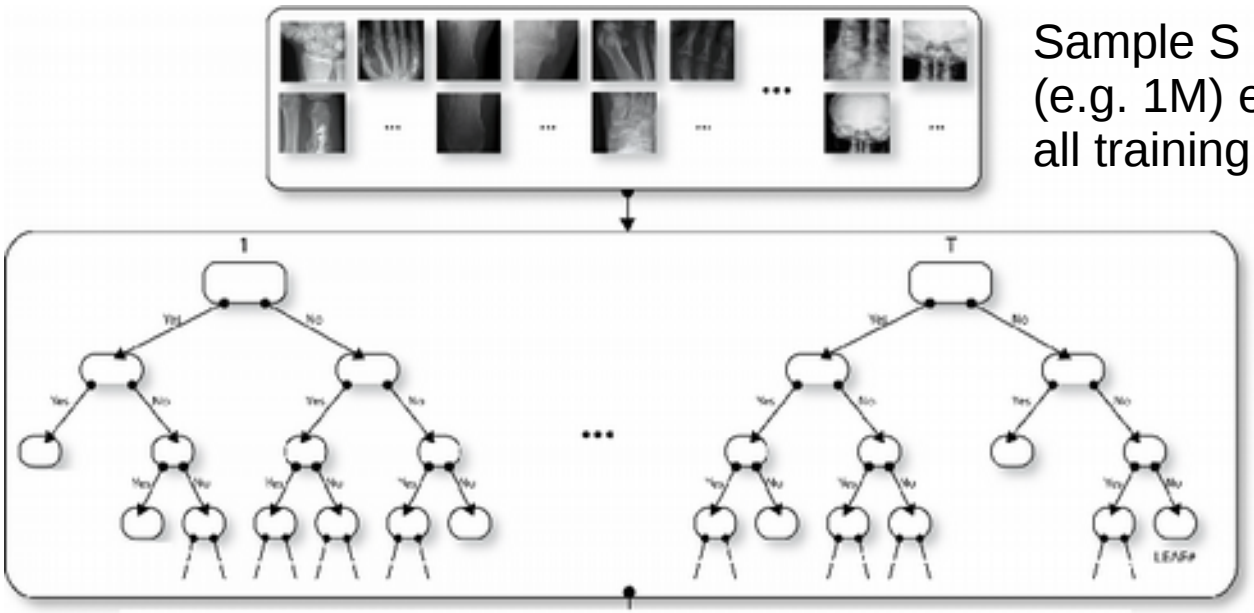
Input: a subset S
Output: a boolean
- If $|S| < n_{\min}$, then return TRUE;
- If if all attributes are constant in S, then return TRUE;
- If the output is constant in S, then return TRUE;
- Otherwise, return FALSE.

Complexity: $O(TKN\log_2(N))$

Parameters :
K = nb random tests
Nmin = minimum node size

Marée et al., CVPR 2005...

Extra-Trees for Direct Classification : ensemble of tree training



Sample S of subwindows (e.g. 1M) extracted from all training images

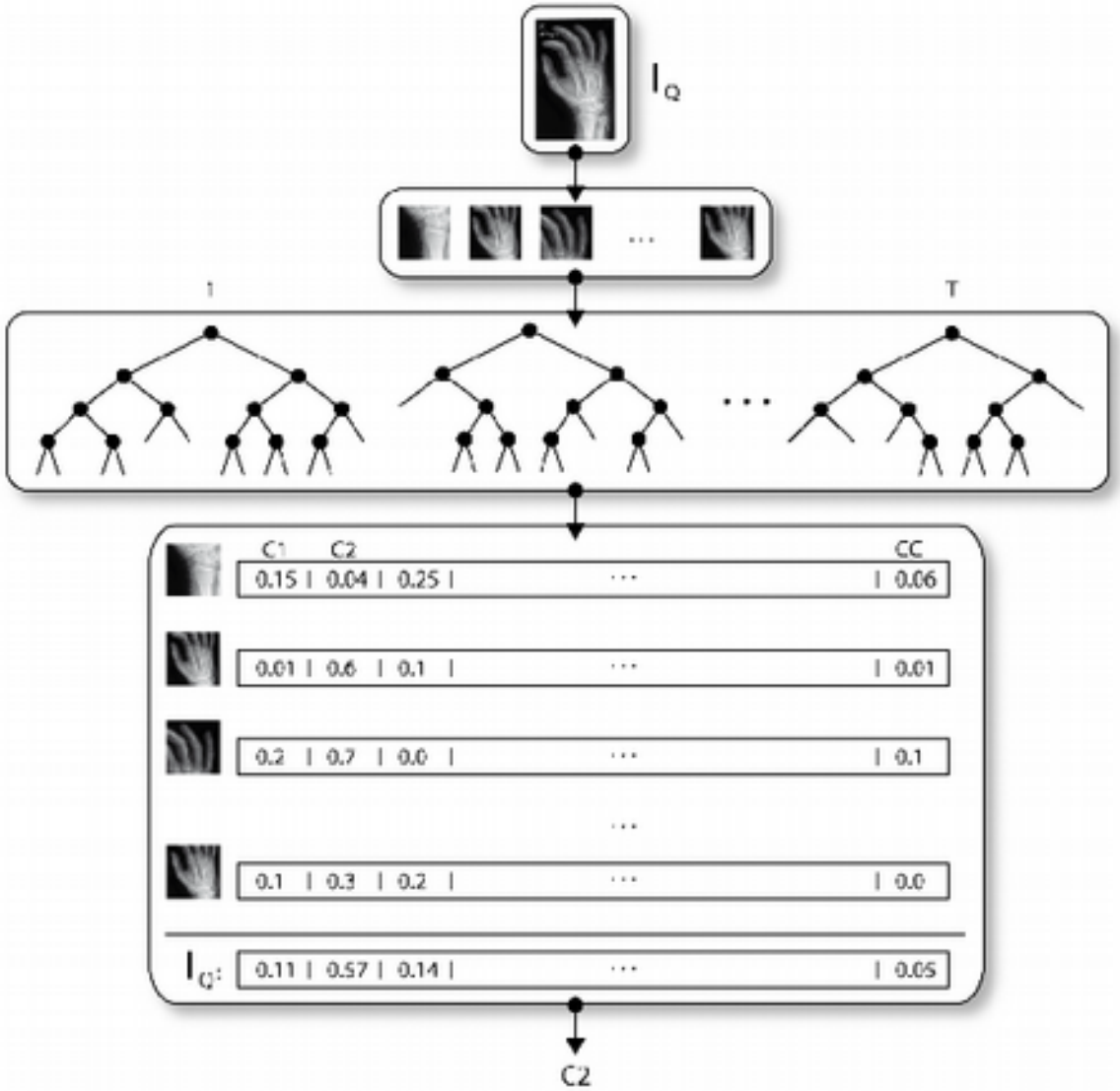
Parameters :
T= nb trees
K = nb random tests
Nmin = minimum node size

Complexity: $O(TKN\log_2(N))$

Marée et al., CVPR 2005...

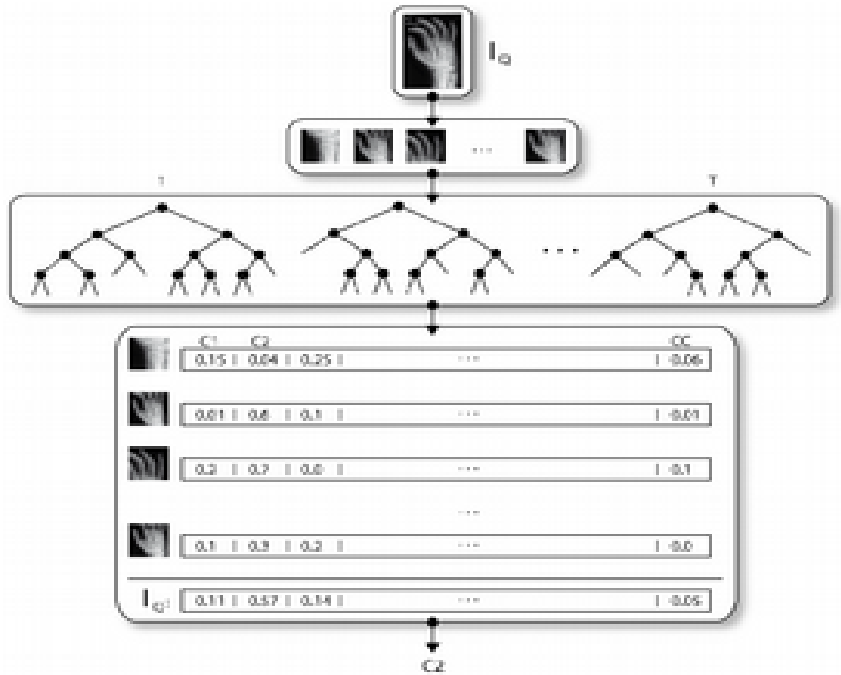
Extra-Trees for Direct Classification : prediction

Parameters :
Nsw = nb subwindows



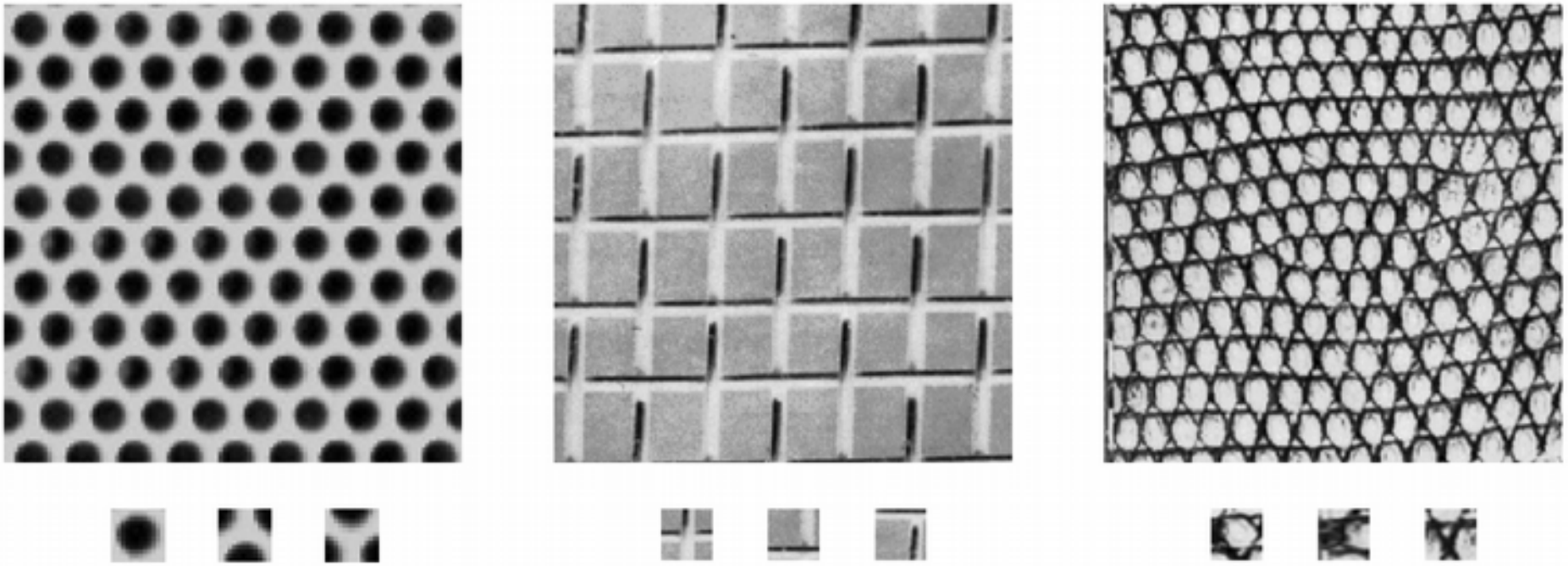
Marée et al., 2005...

Extra-Trees for Direct Classification : prediction



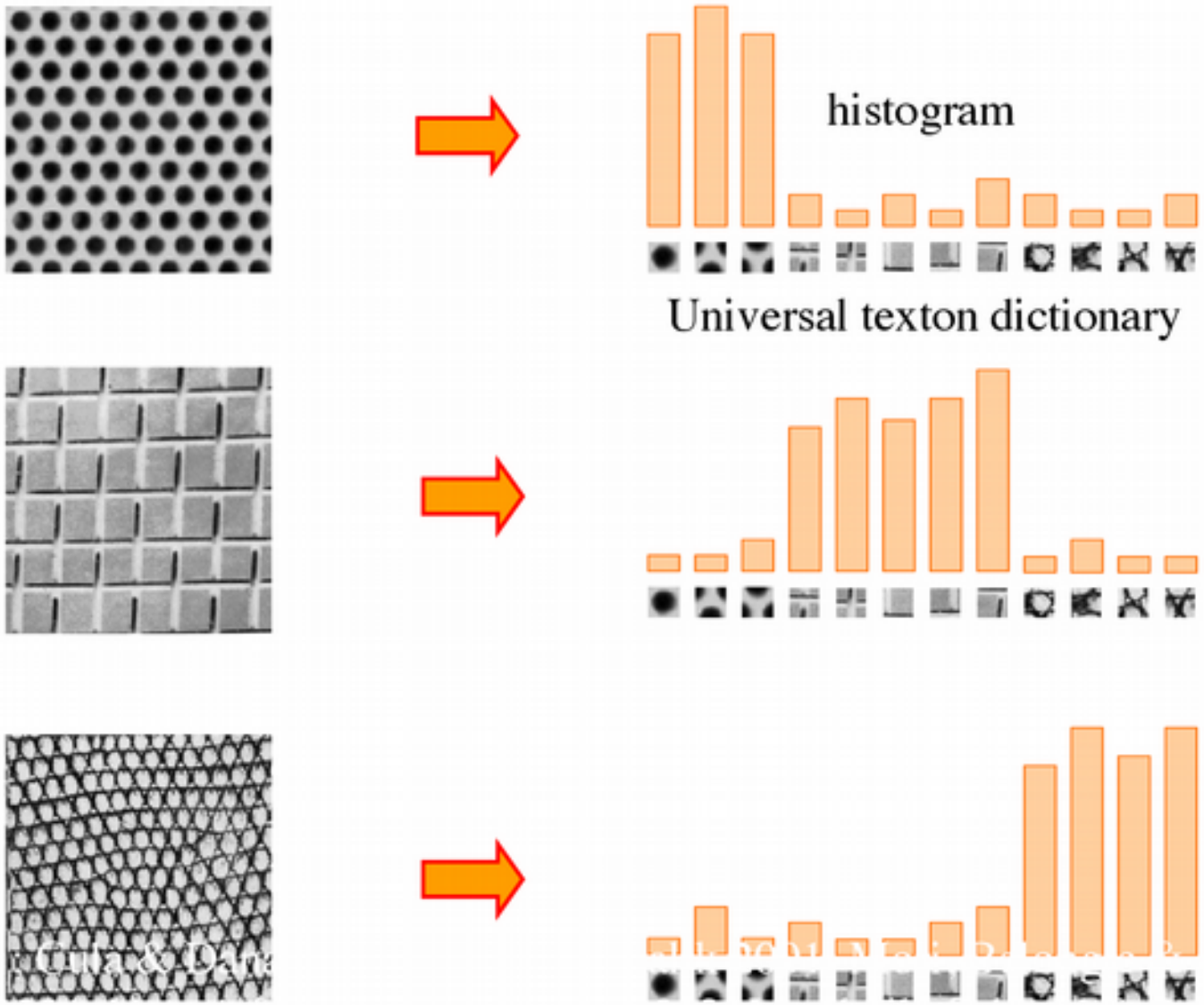
→ From 65% downto 2% error rate (large improvement !)

BoF Origin: Texture Classification



Julesz, 1981; Cula & Dana, 2001; Leung & Malik 2001; Mori, Belongie & Malik, 2001; Schmid 2001; Varma & Zisserman, 2002, 2003; Lazebnik, Schmid & Ponce, 2003

Texture Classification: Histograms over Textons



Bag-of-features for Image Classification

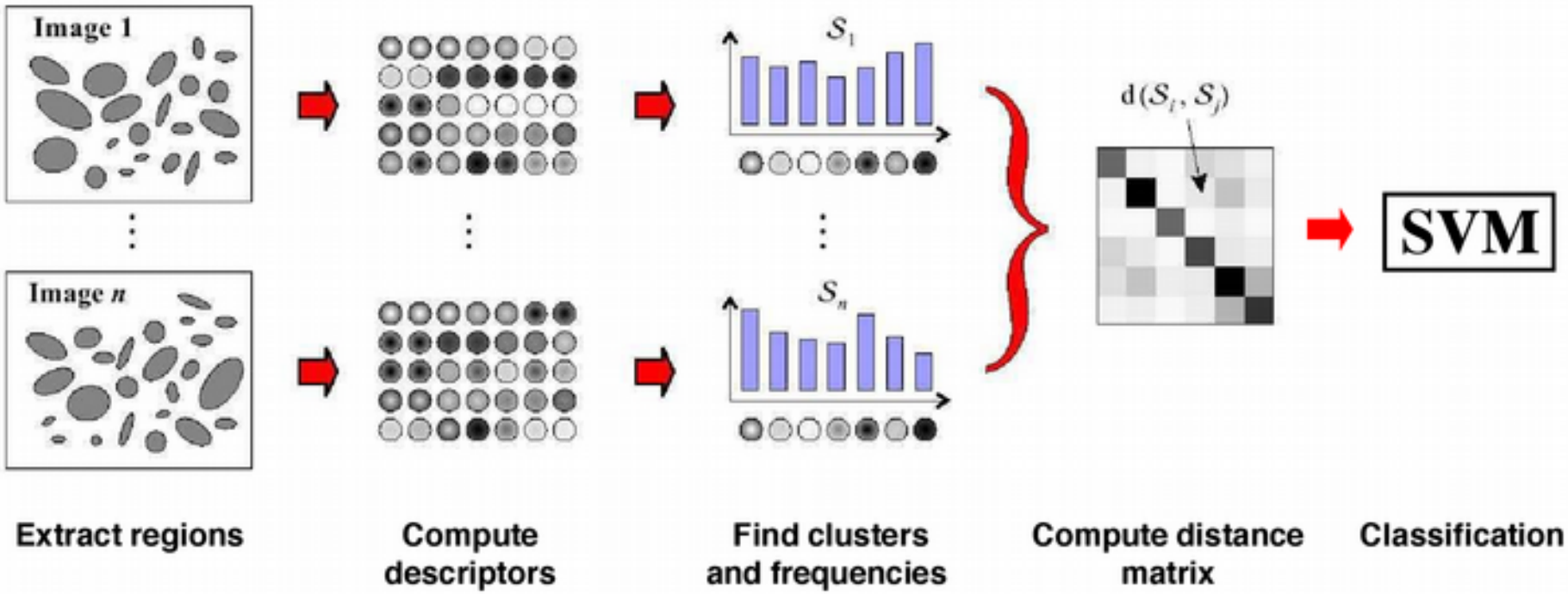
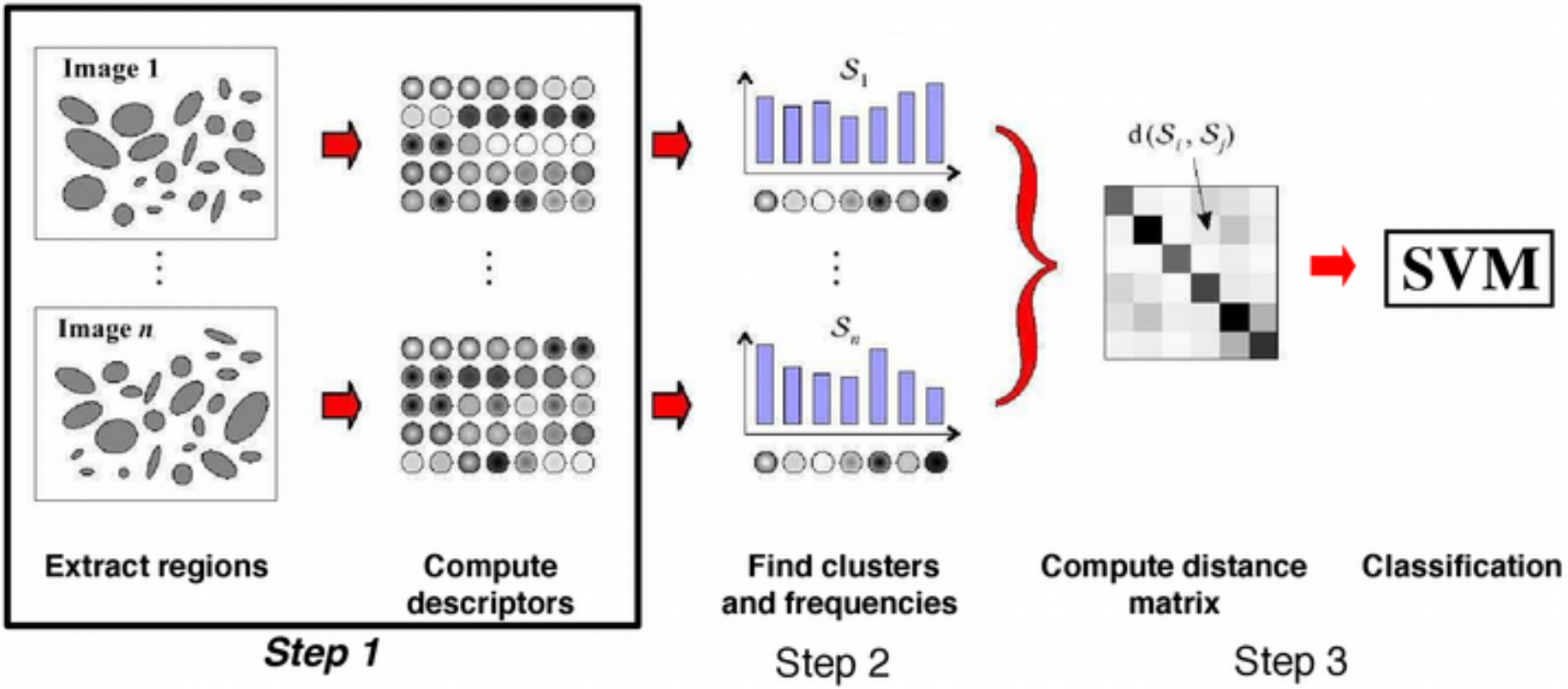


Image 1 contains a bike, image 2 contains a horse, image 3 contains a car, etc...

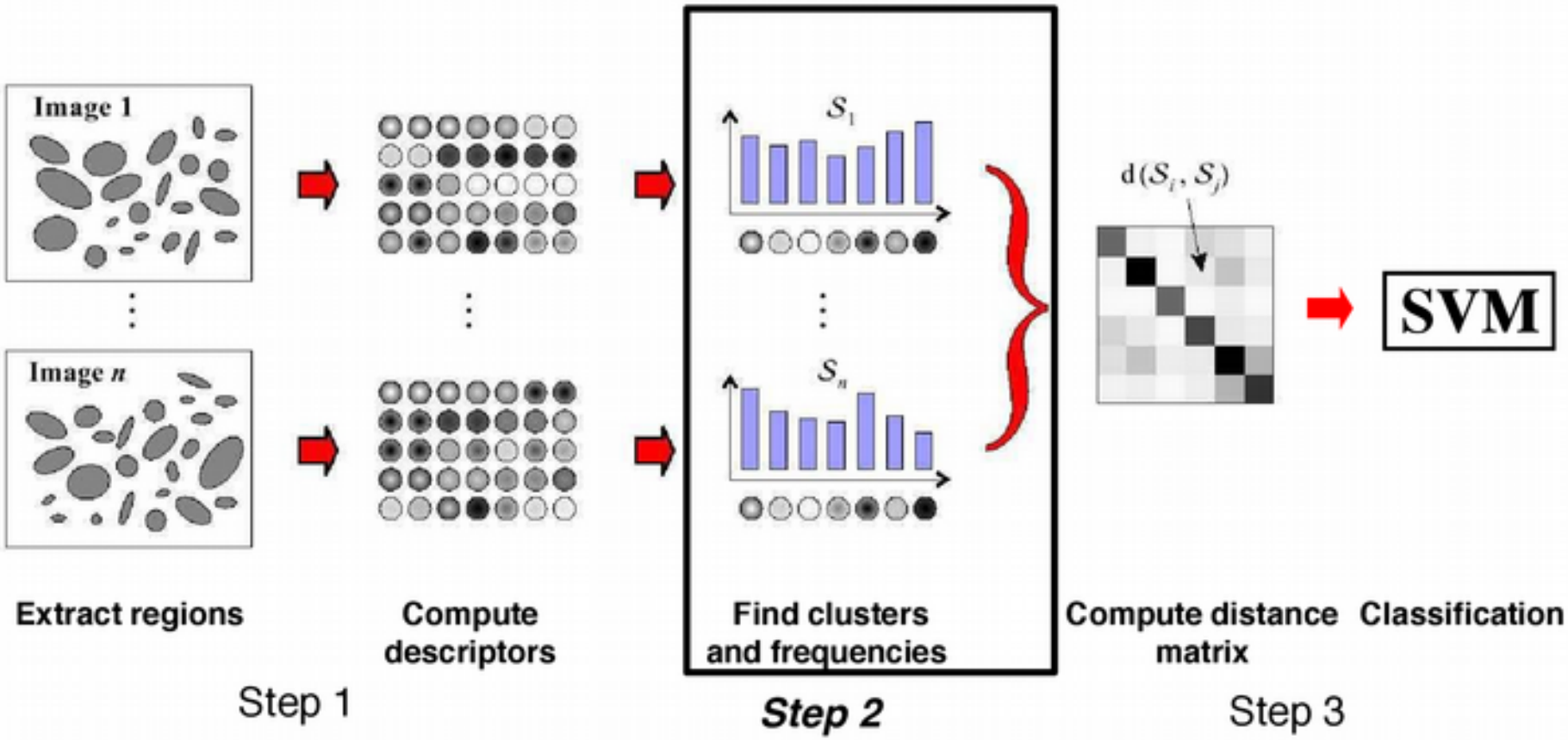
[Csurka et al., ECCV Workshop'04], [Nowak, Jurie, Triggs, ECCV'06], [Zhang, Marszalek, Lazebnik, Schmid, IJCV'07]

Step 1: Extract Features



Corners / Point / Random / ...

Step 2: Cluster Features, Compute Feature Frequencies

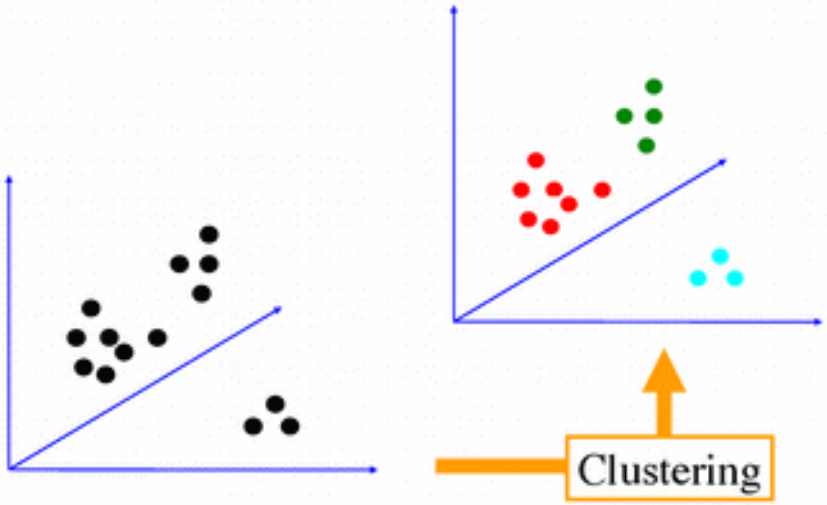


Unsupervised (e.g. k-means)

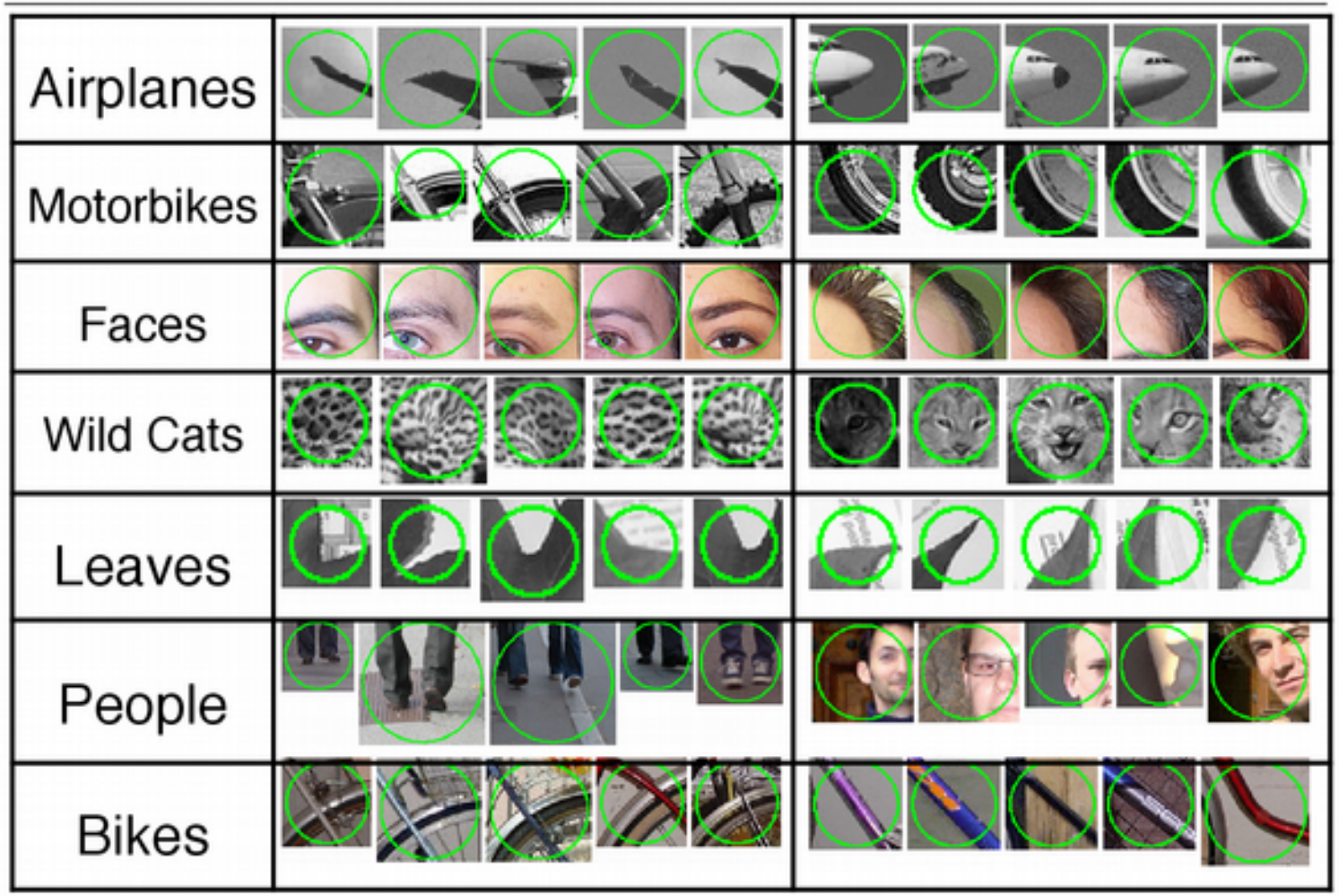
Supervised (e.g. trees)

Clustering Features (k-means)

Because of viewpoint and lighting changes, it is unlikely that two images of even the same object will produce exactly the same features.
This gets even worse when working with different instances of a class (e.g., different cars).
As a result, it is not a good idea to model an object (or object class) with a histogram over all the features that the object produces.
Instead, we cluster *all* the features that come from the training images (of all classes), and keep only the cluster centers. The set of cluster centers is called a codebook, or dictionary. The elements of the codebook are called codewords.



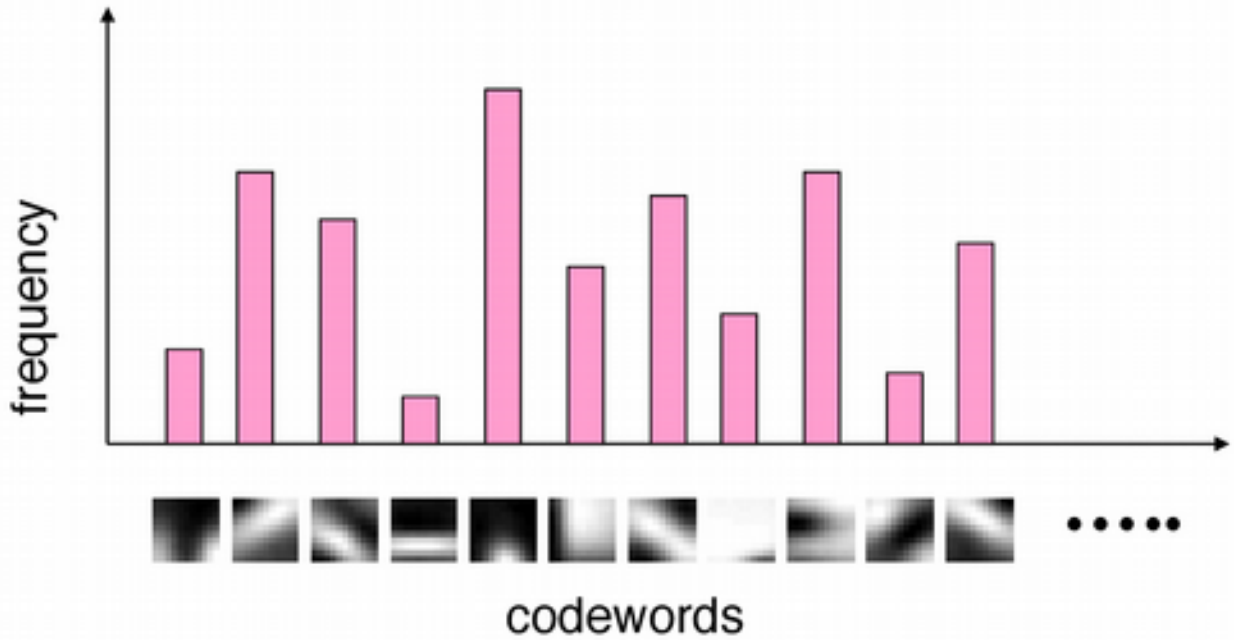
Examples of Clusters of Features



↓ Average of these becomes one codeword

↓ Average of these becomes one codeword

Object/Class Instance Representation: Codeword Frequencies

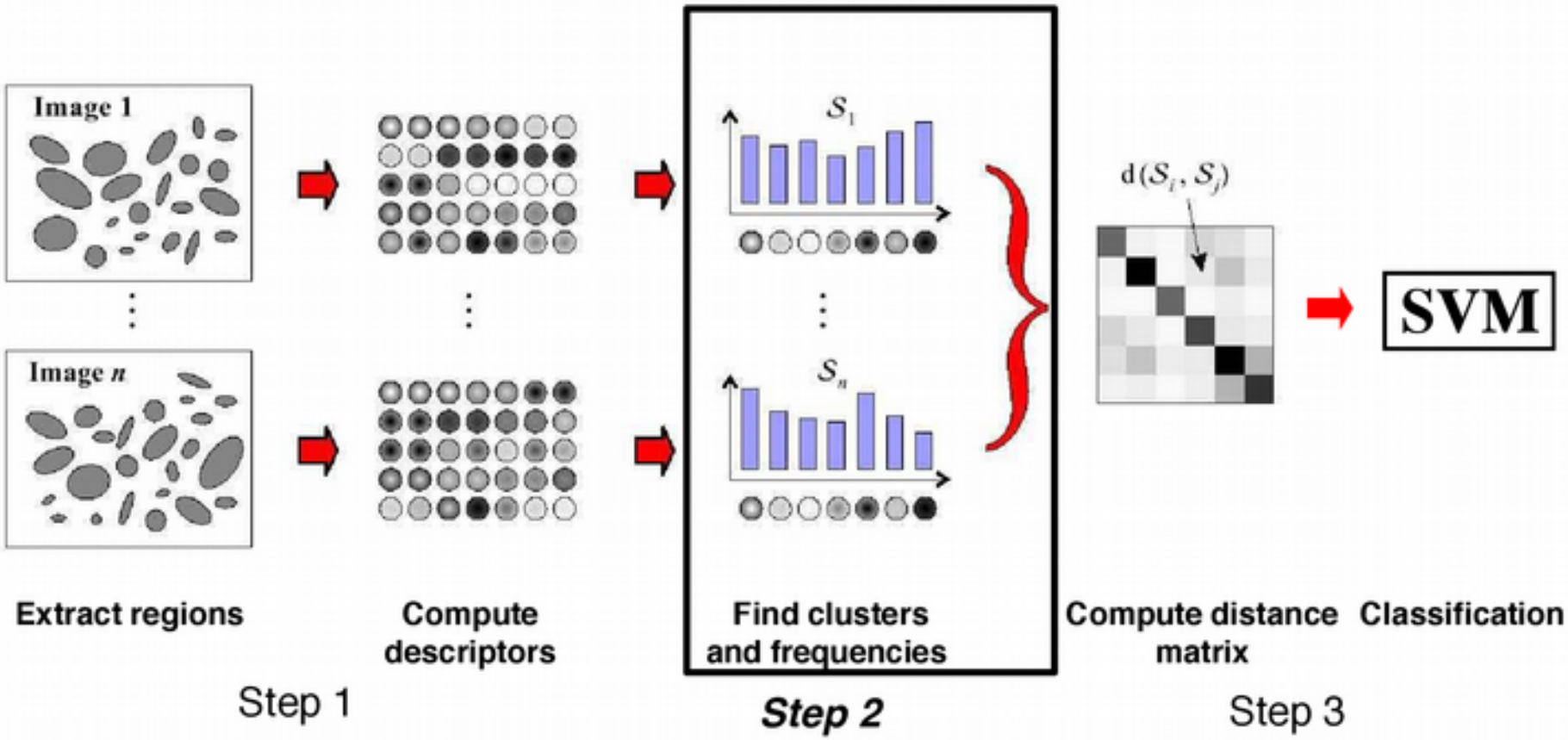


Typically: 1000–4000 codewords:

- More codewords: towards object representation
- Less codewords: towards object class representation

One image of an instance of an object/class is represented with a vector V of frequencies of the codewords. (L1/L2 normalization)

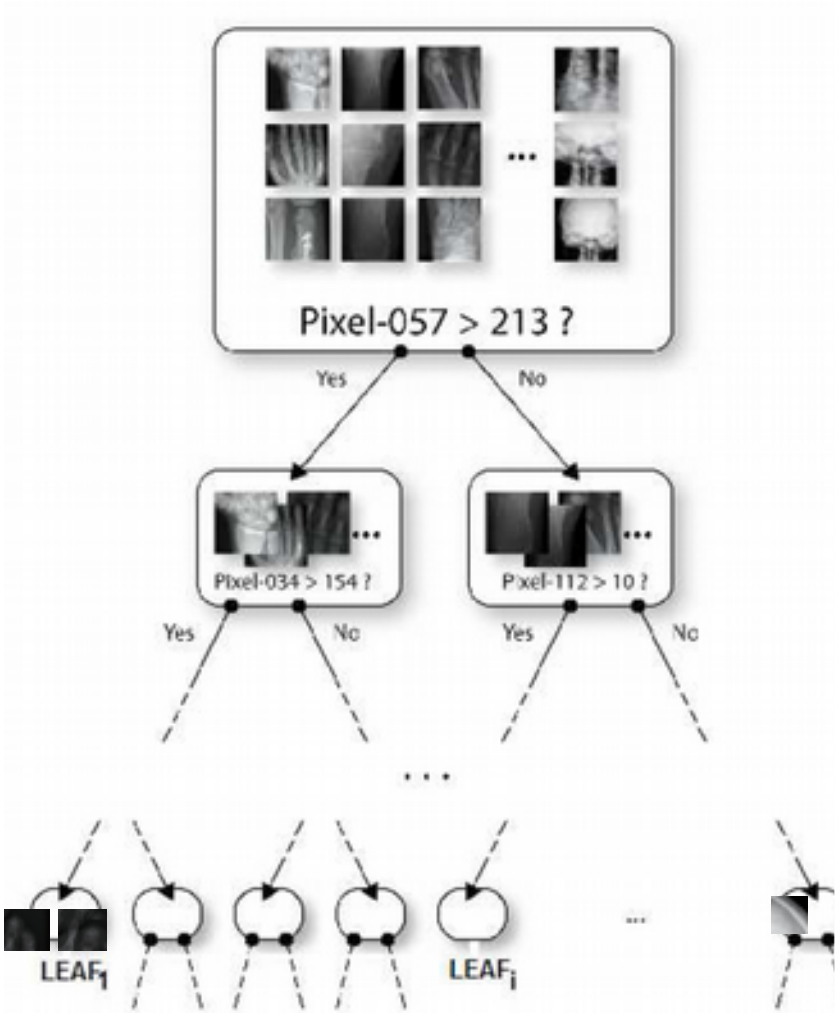
Step 2: Cluster Features, Compute Feature Frequencies



Unsupervised
(e.g. k-means)

Supervised
(e.g. trees)

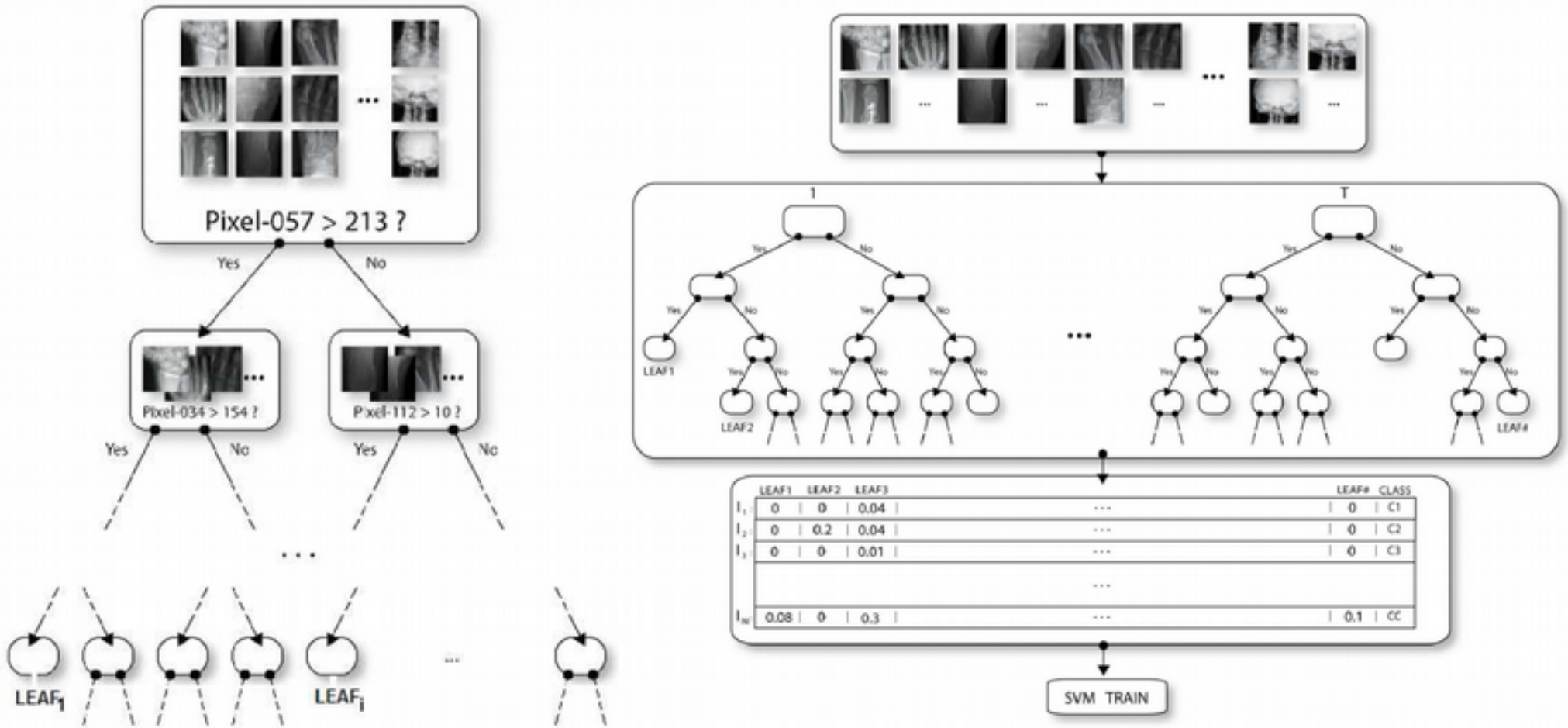
Extra-Trees for Feature Learning : training



Parameters :

K = nb random tests
Nmin = minimum node size

Extra-Trees for Feature Learning : training



Parameters :
T= nb trees
K = nb random tests
Nmin = minimum node size
Coding = binary/frequency
FinalC = liblinear

Step 3: Image Classification

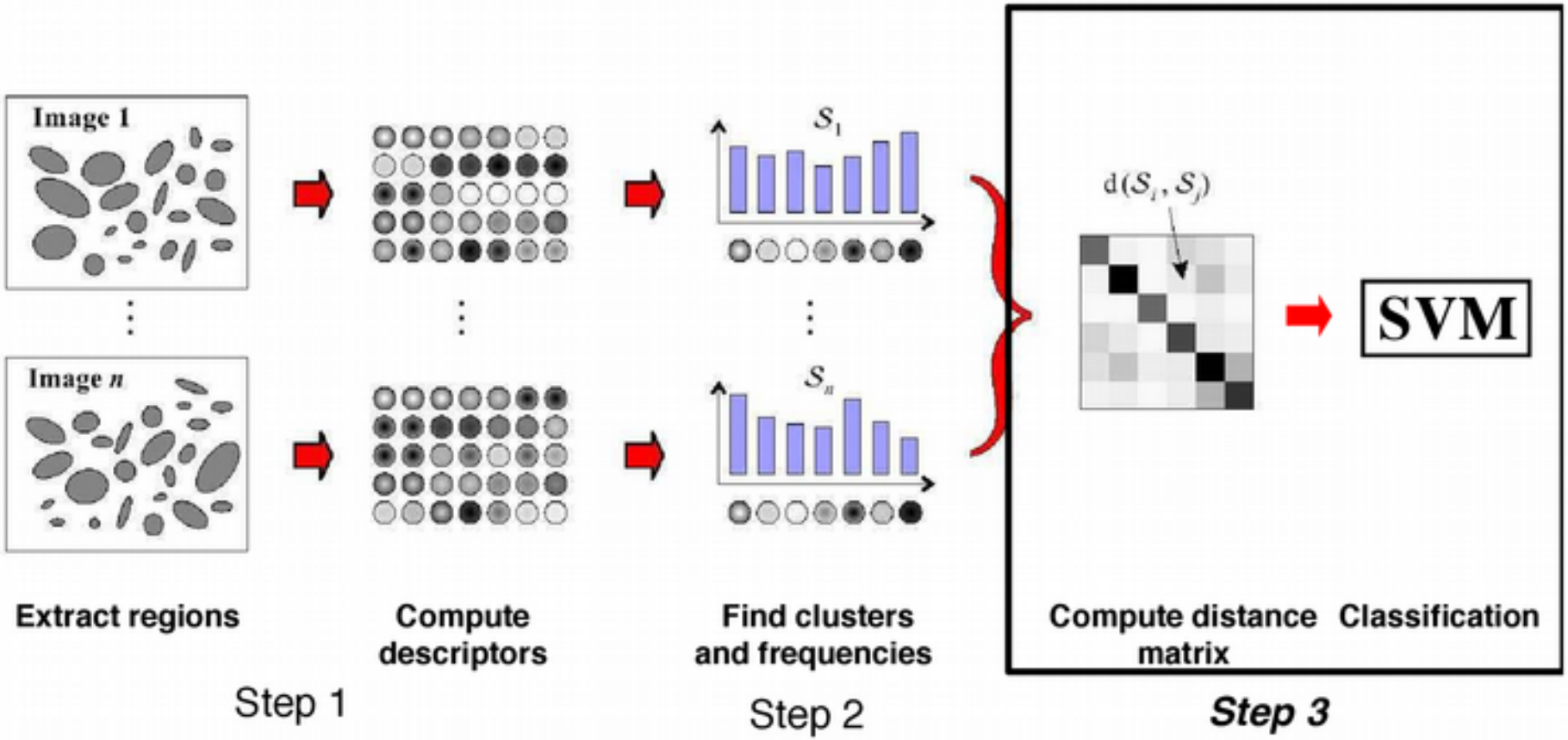
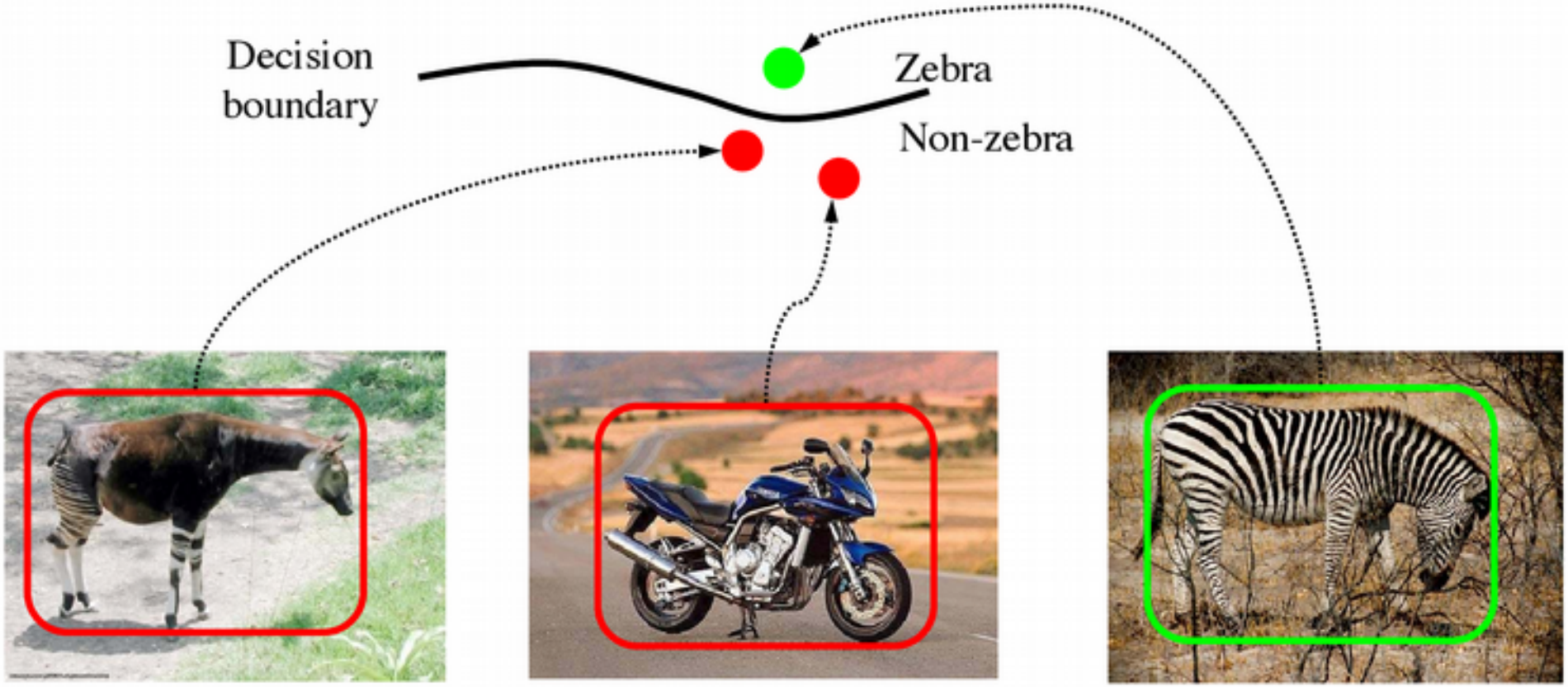
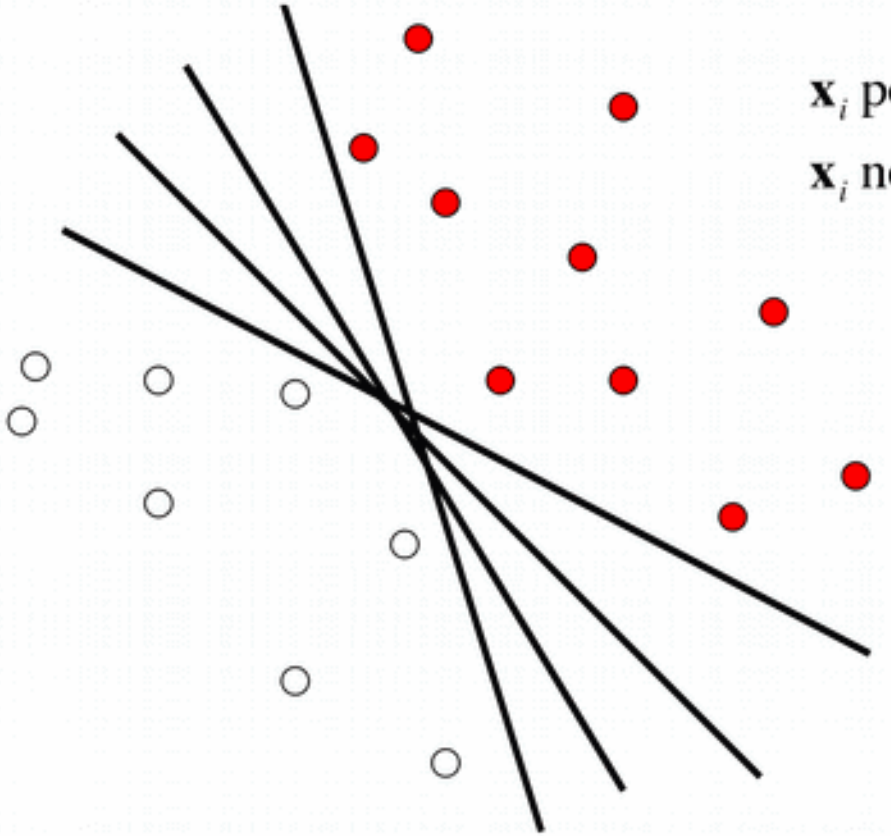


Image Classification

Goal: Learn a decision rule (classifier) to assign V to an object/class.



Linear Classification



$\mathbf{x}_i$ positive: $\mathbf{x}_i \cdot \mathbf{w} + b \geq 0$

$\mathbf{x}_i$ negative: $\mathbf{x}_i \cdot \mathbf{w} + b < 0$

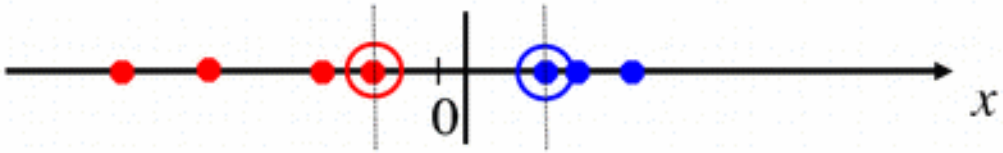
Which hyperplane is best?

For instance: support vector machines (SVM), logistic regression, linear discriminant analysis (LDA), naive Bayes classification, ...

(see “Introduction to Machine Learning”)

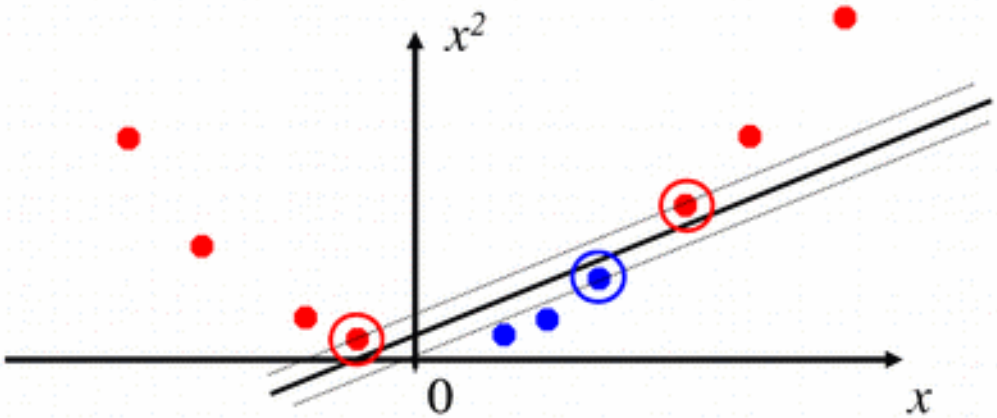
Nonlinear Classification

Datasets that are linearly separable work out great:



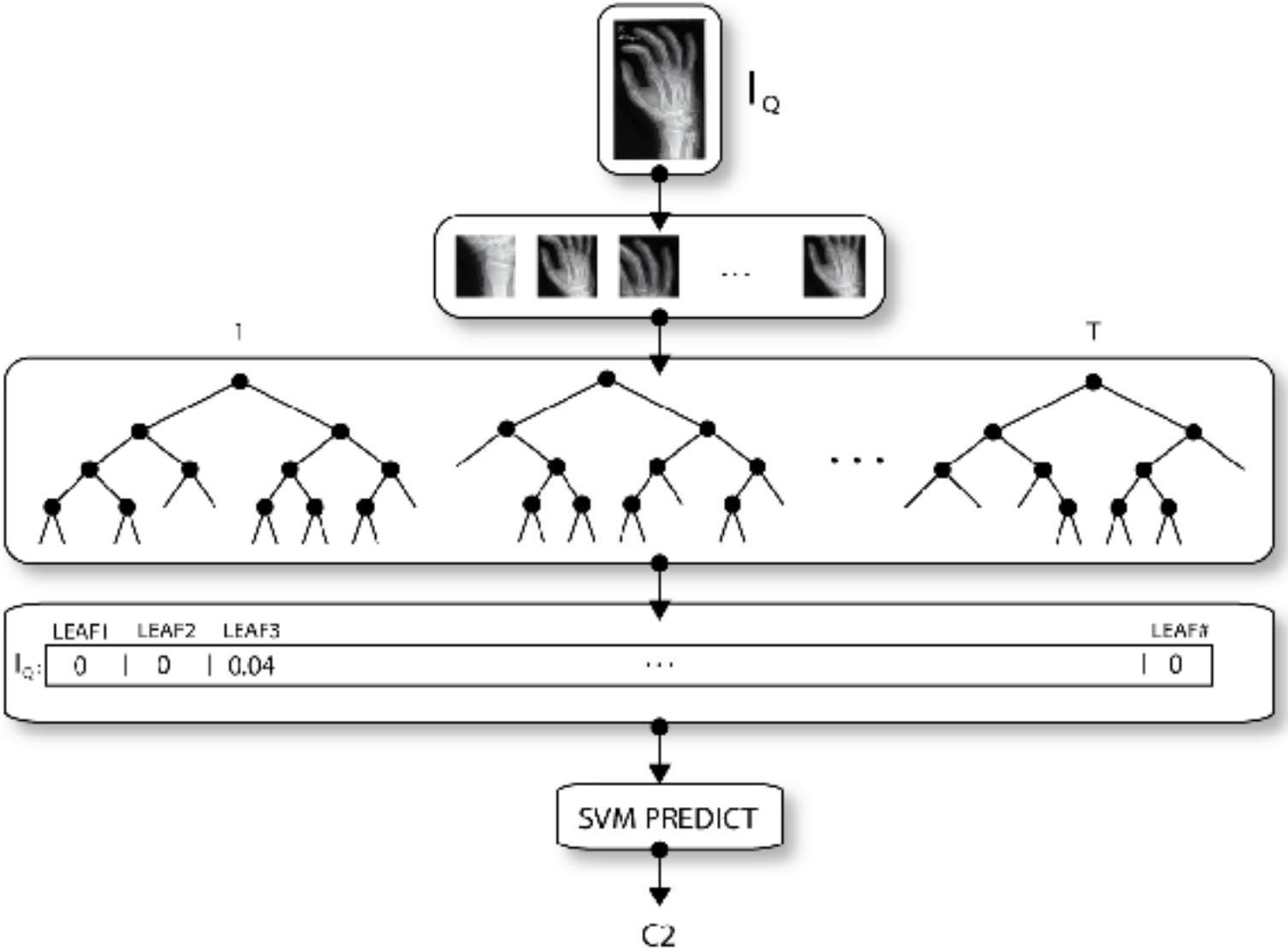
But what if the data set is just too hard?

Map the data to a higher dimensional space where it is linearly separable:



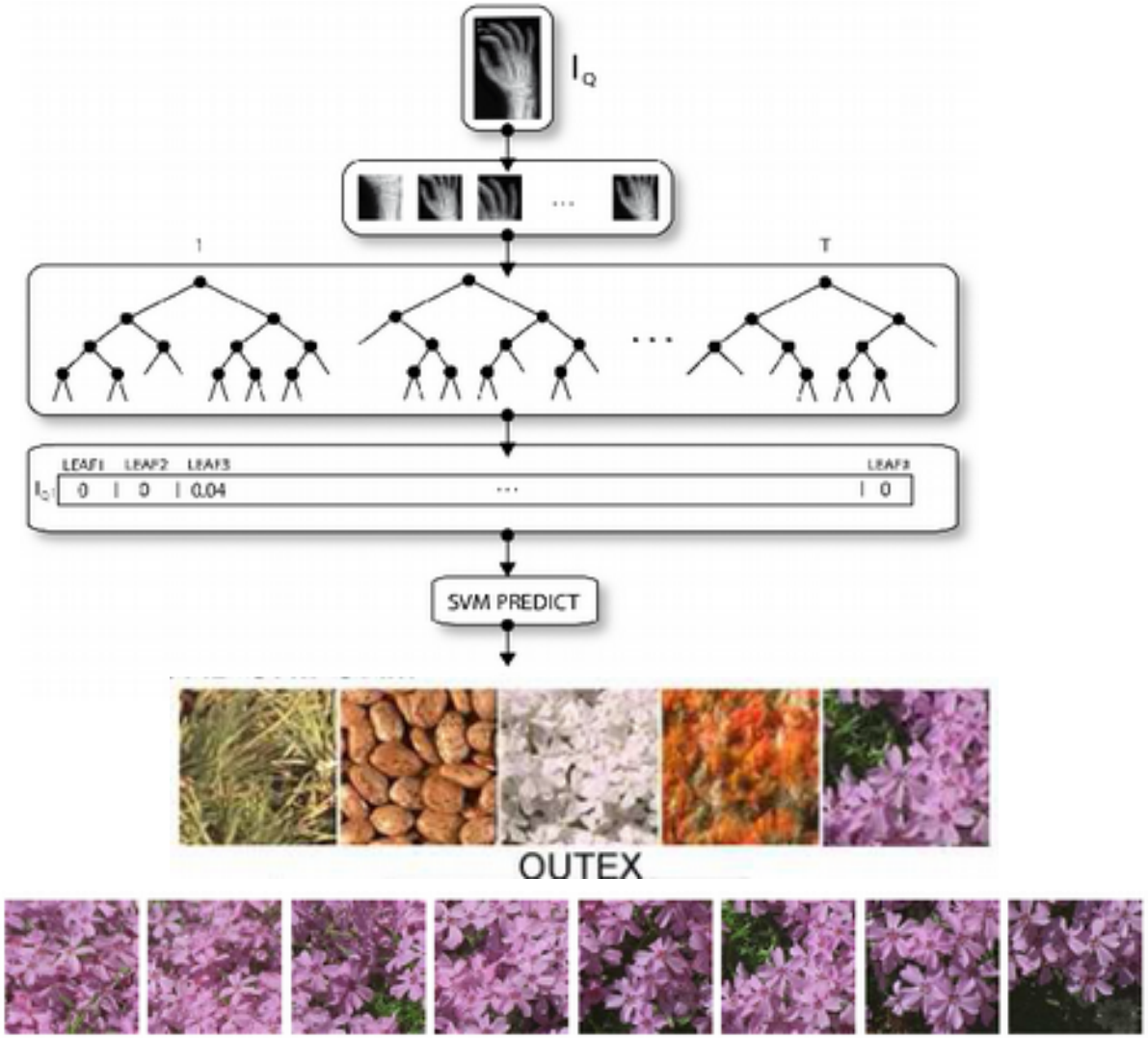
(see “Introduction to Machine Learning”)

Extra-Trees for Feature Learning : prediction



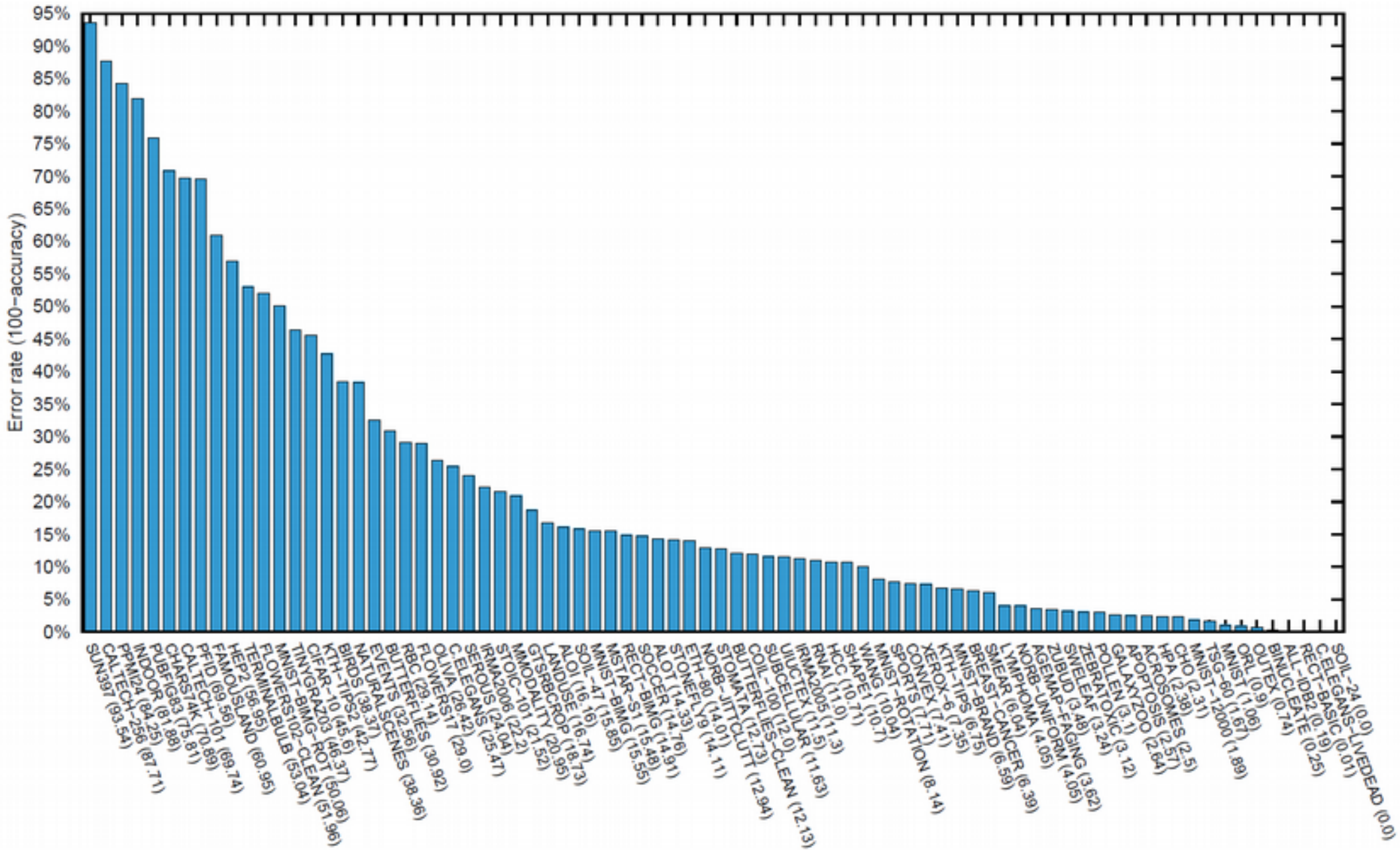
Parameters :
Nsw = nb subwindows

Extra-Trees for Feature Learning : prediction



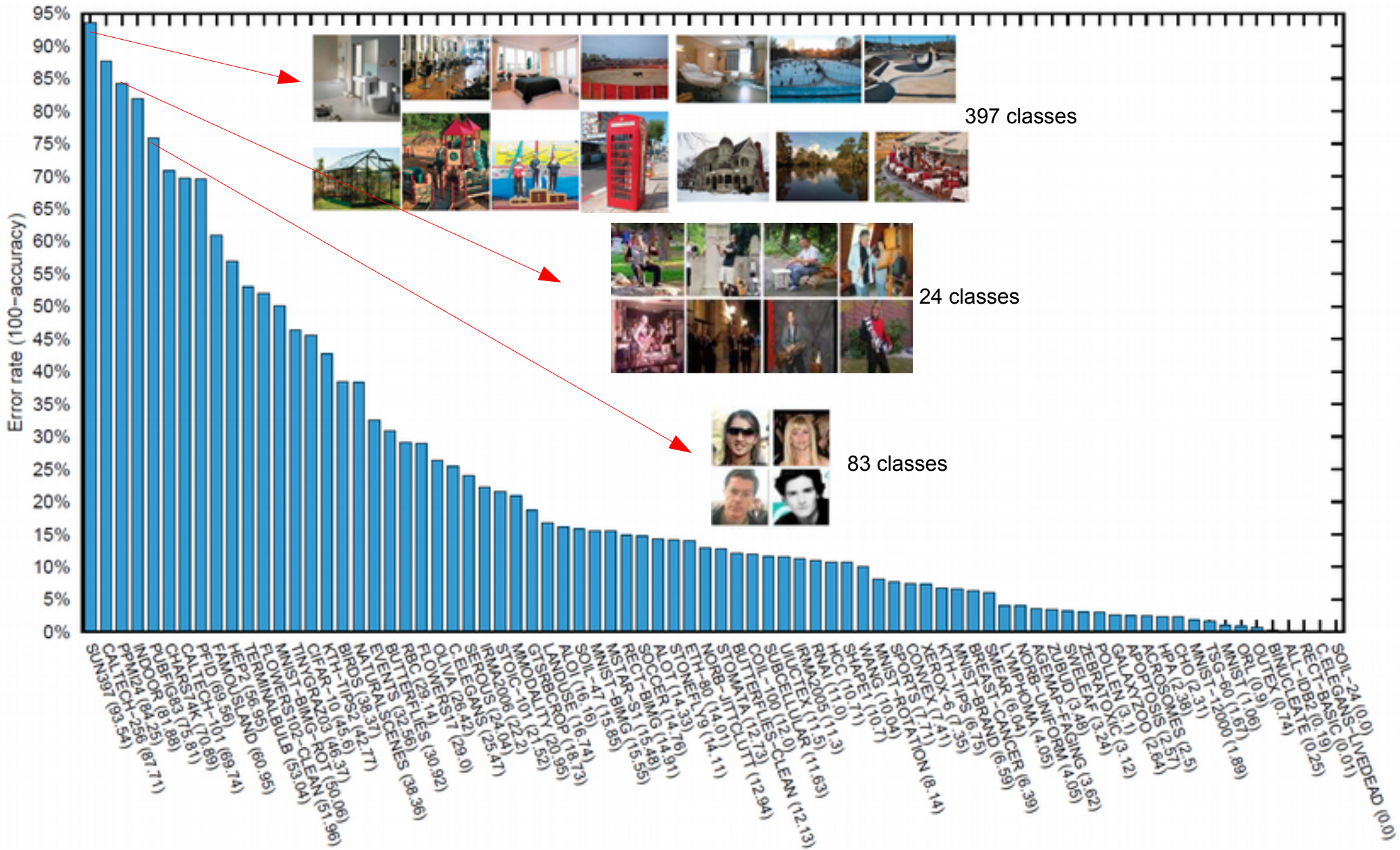
From 2.01% downto 1.04% error rate

Overall results (error rates)



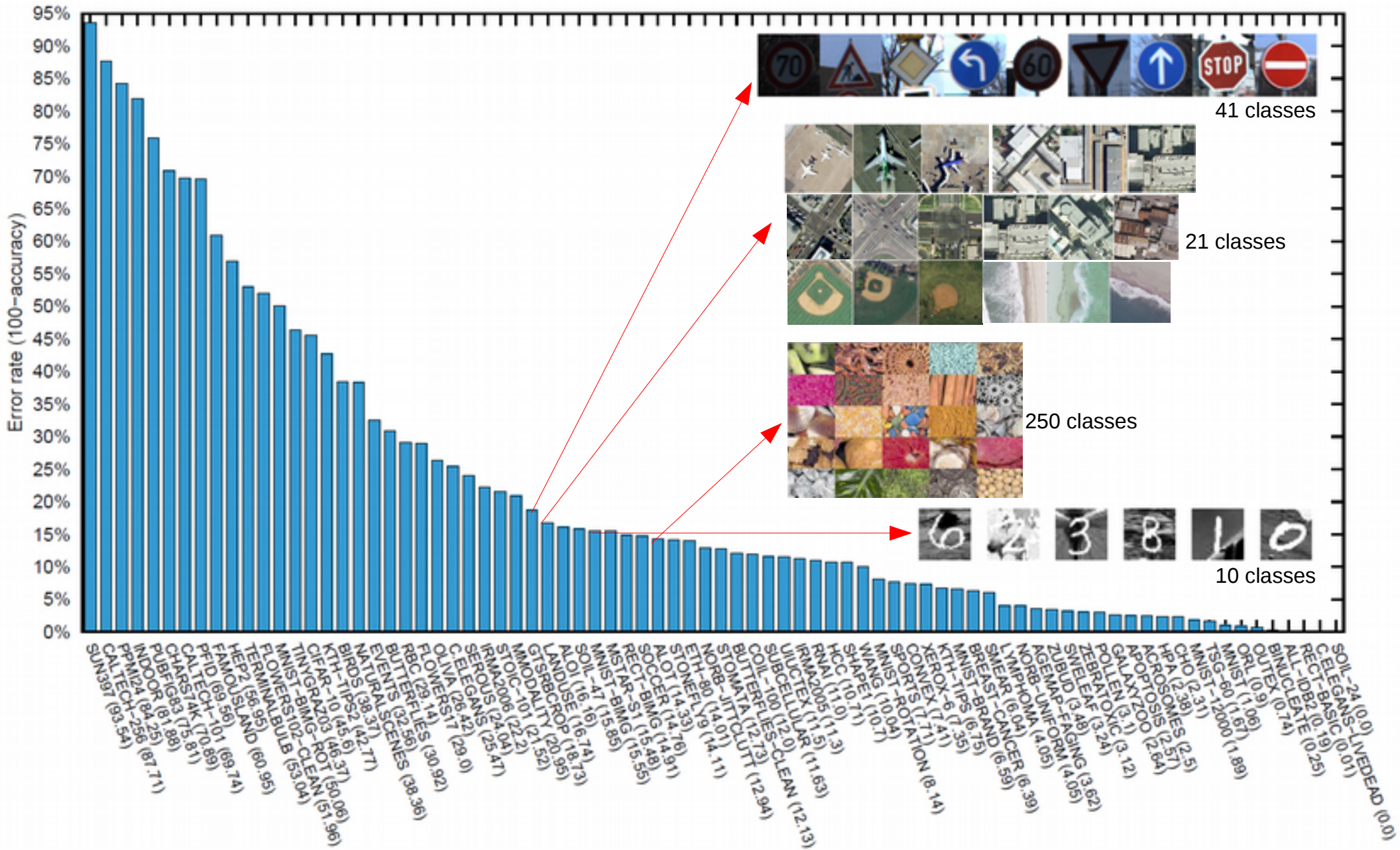
Marée et al., in preparation, 2013

Overall results (error rates)



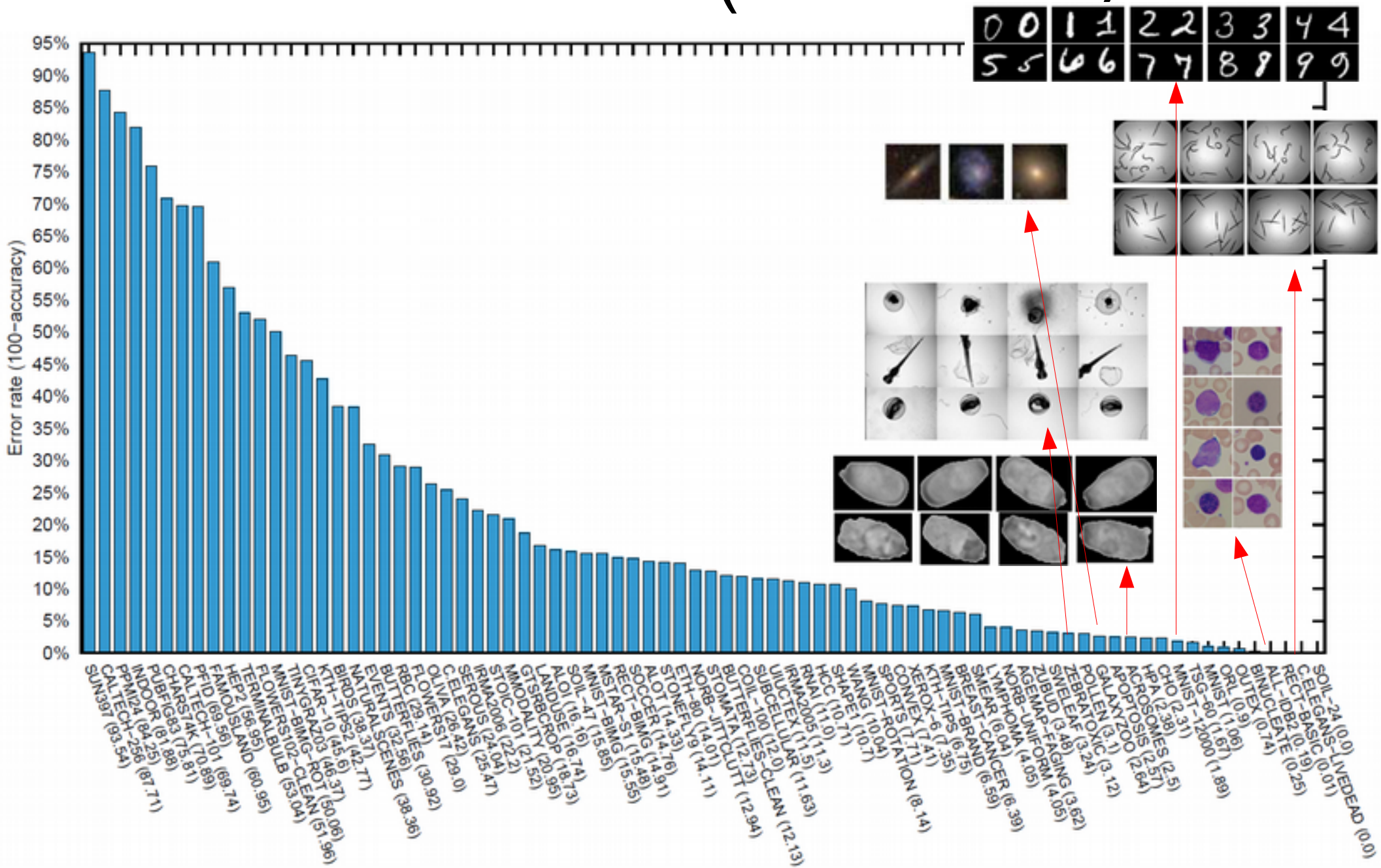
Marée et al., in preparation, 2013

Overall results (error rates)



Marée et al., in preparation, 2013

Overall results (error rates)



Marée et al., in preparation, 2013

Summary

- Many features have been designed to ease vision tasks
- Many learning methods have been designed
- Several (controlled) vision tasks can be solved with end-to-end learning
- But there is still no universal vision method

Pause

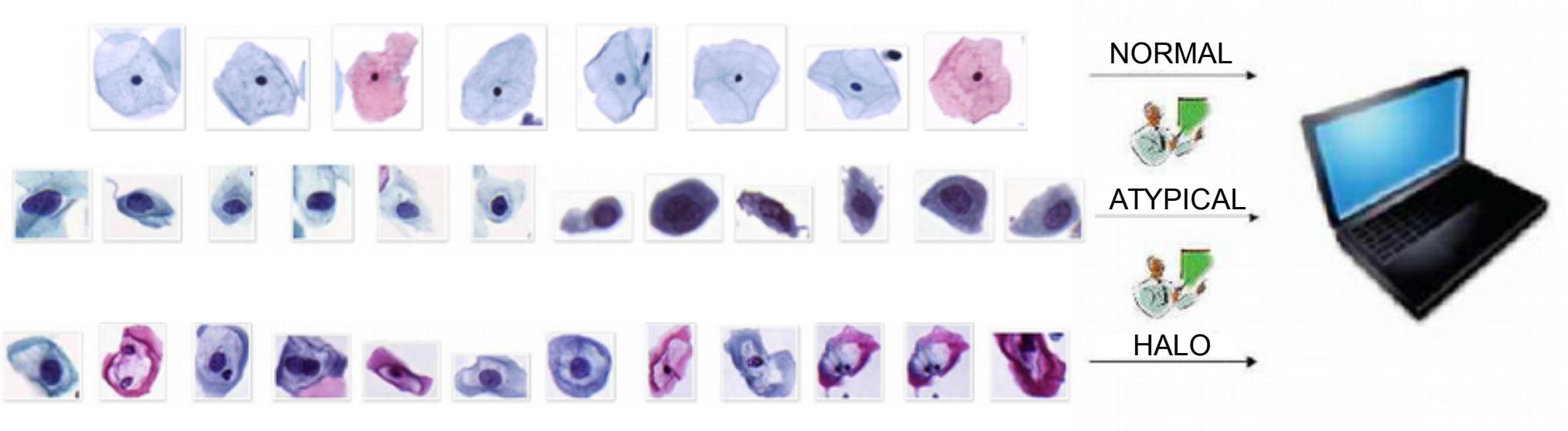
From research to real-world

- **The need for realistic data collection**

- Recent trends
  - Deep learning
  - Multispectral, Multimodal imaging
  - Open hardware/software

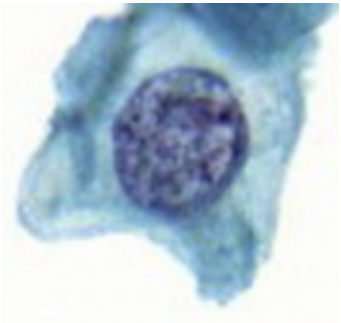
Pattern recognition : training

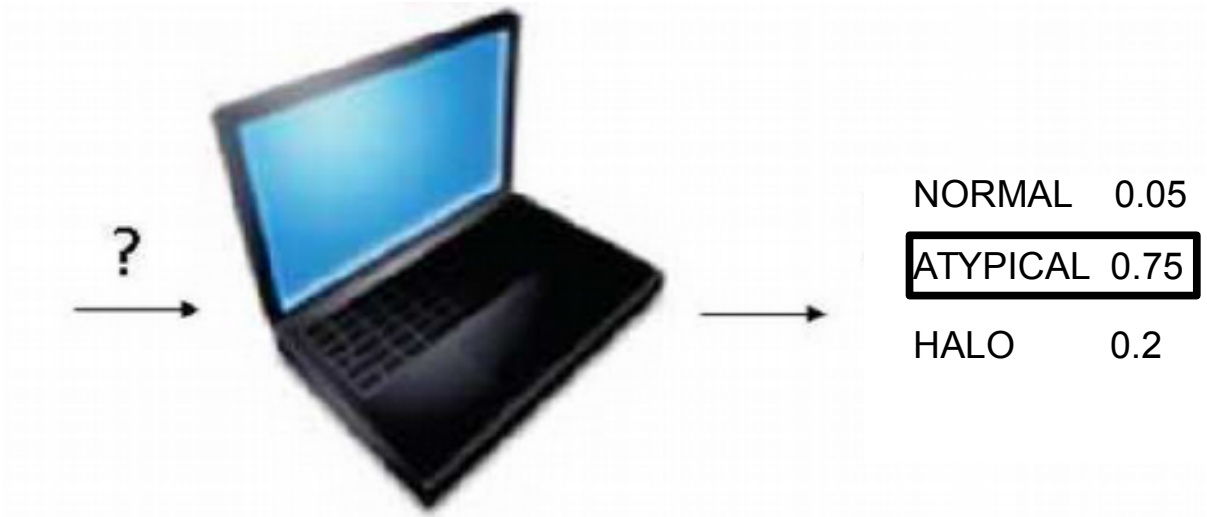
Given a training set of labeled images (one class per image, among a finite number of predefined classes), build a model that will be able to predict accurately the class of new, unseen, objects/images



Pattern recognition : prediction

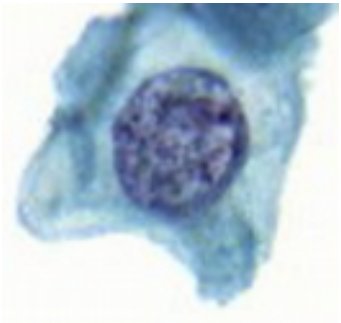
Given a training set of labeled images (one class per image, among a finite number of predefined classes), build a model that will be able to predict accurately the class of new, unseen, objects/images

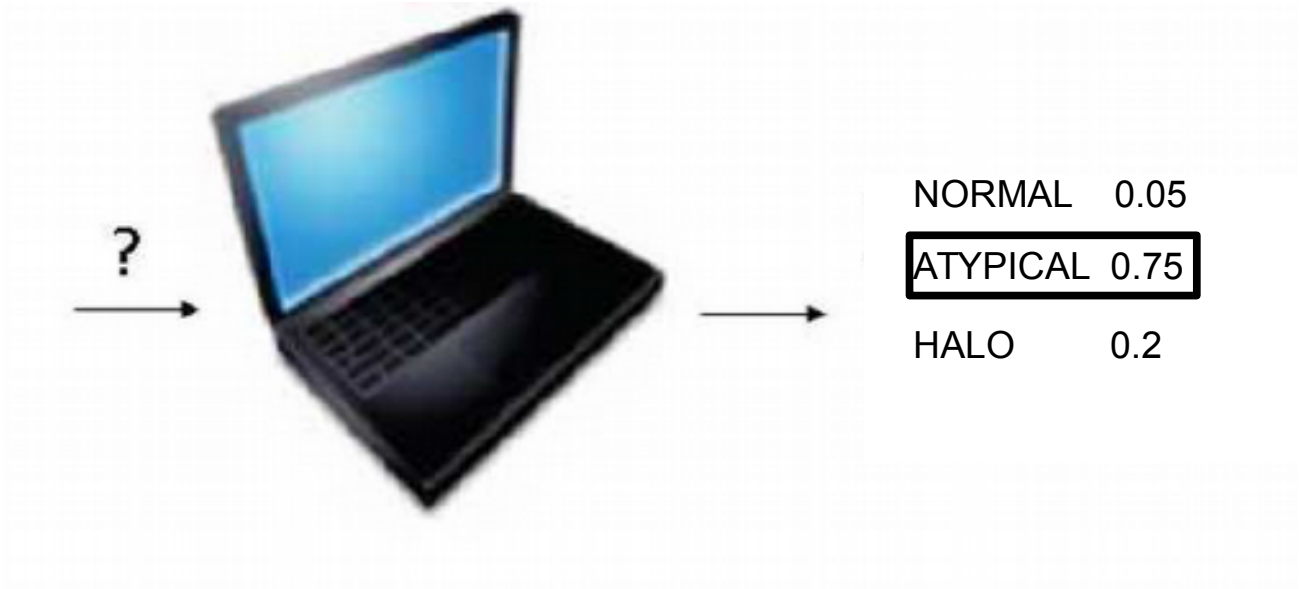




Pattern recognition : prediction

Given a training set of labeled images (one class per image, among a finite number of predefined classes), build a model that will be able to predict accurately the class of new, unseen, objects/images





Pattern recognition algorithms are designed and validated using benchmark datasets

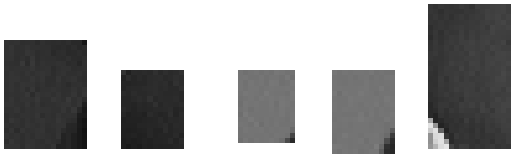
Benchmark dataset quality issues

Int J Comput Vis (2008) 79: 225–230
DOI 10.1007/s11263-008-0143-7

SHORT PAPER

Evaluation of Face Datasets as Tools for Assessing the Performance of Face Recognition Methods

Lior Shamir



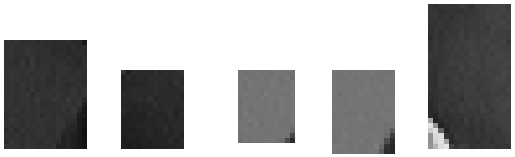
Benchmark dataset quality issues

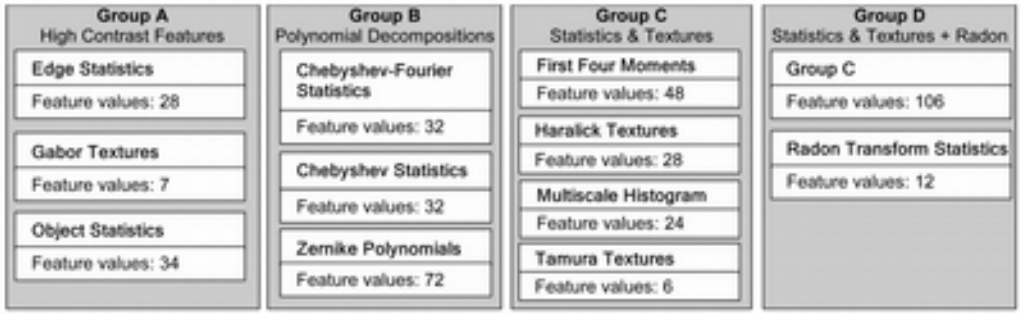
Int J Comput Vis (2008) 79: 225–230
DOI 10.1007/s11263-008-0143-7

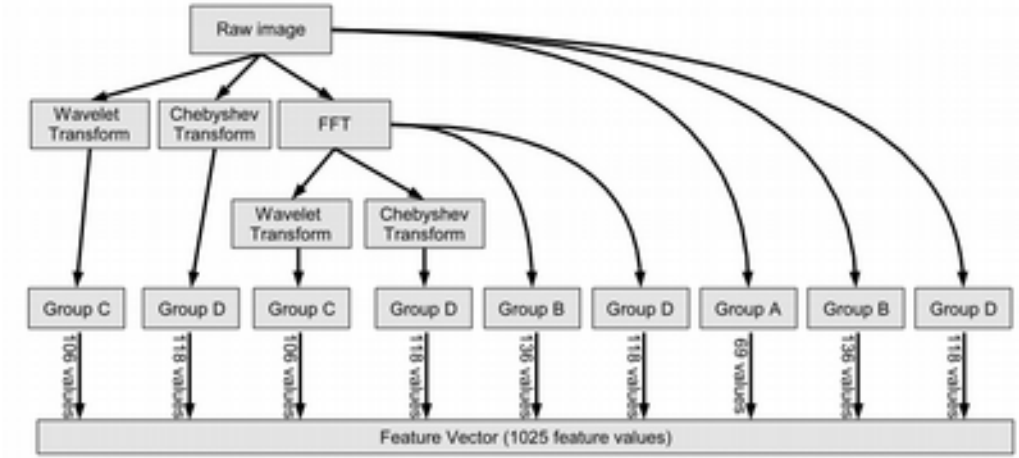
SHORT PAPER

Evaluation of Face Datasets as Tools for Assessing the Performance of Face Recognition Methods

Lior Shamir







Benchmark dataset issues : hidden artefacts

Int J Comput Vis (2008) 79: 225–230
DOI 10.1007/s11263-008-0143-7

SHORT PAPER

Evaluation of Face Datasets as Tools for Assessing the Performance of Face Recognition Methods

Lior Shamir

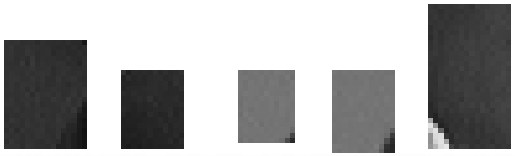


Table 1 Classification accuracy of the face datasets using a small non-facial area

Dataset	Subjects	Images per subject	Original image size	Non-facial area	Random accuracy	Non-facial accuracy
ORL	40	10	92 × 112	20 × 20 (bottom right)	0.025	0.788
JAFFE	10	22	256 × 256	25 × 200 (top left)	0.1	0.94
Indian Face Dataset (Females)	22	11	160 × 120	42 × 80 (top left)	0.045	0.73
Indian Face Dataset (Males)	39	11	160 × 120	42 × 80 (top left)	0.0256	0.58
Essex	100	20	196 × 196	42 × 100 (top left)	0.01	0.97
Yale B	10	576	640 × 480	100 × 300 (top left)	0.1	0.99
Color FERET	994	5	512 × 768	100 × 100 (top left)	~ 0.001	0.135

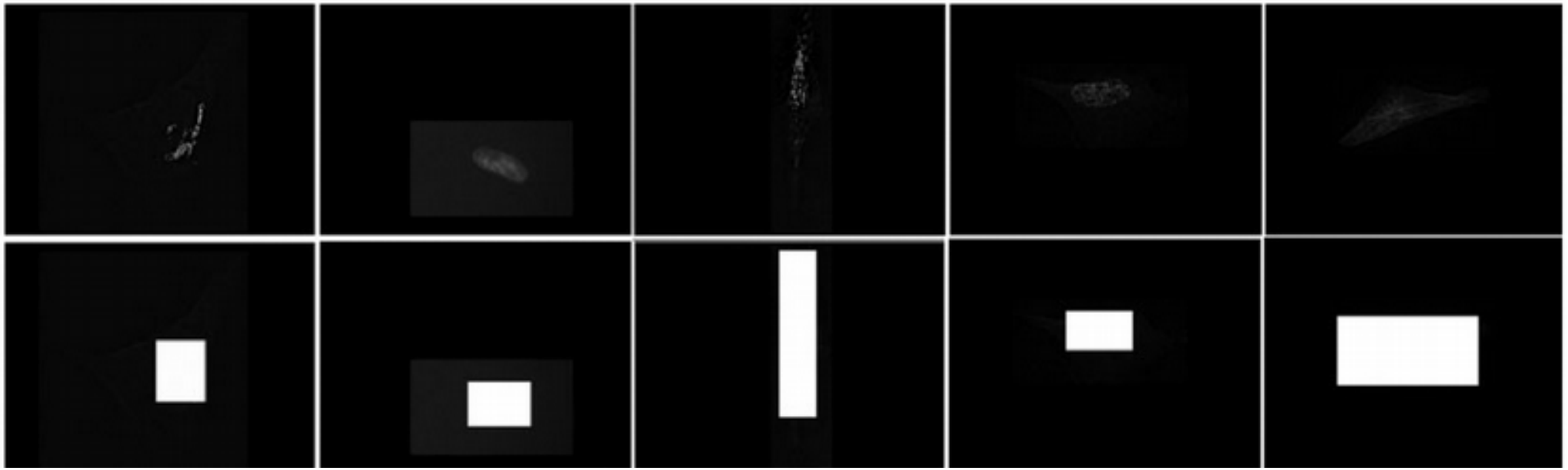
Benchmark dataset issues : hidden artefacts

Journal of Microscopy

Journal of Microscopy, Vol. 243, Pt 3 2011, pp. 284–292
doi: 10.1111/j.1365-2818.2011.03502.x
Received 11 January 2011; accepted 17 March 2011

Assessing the efficacy of low-level image content descriptors for computer-based fluorescence microscopy image analysis

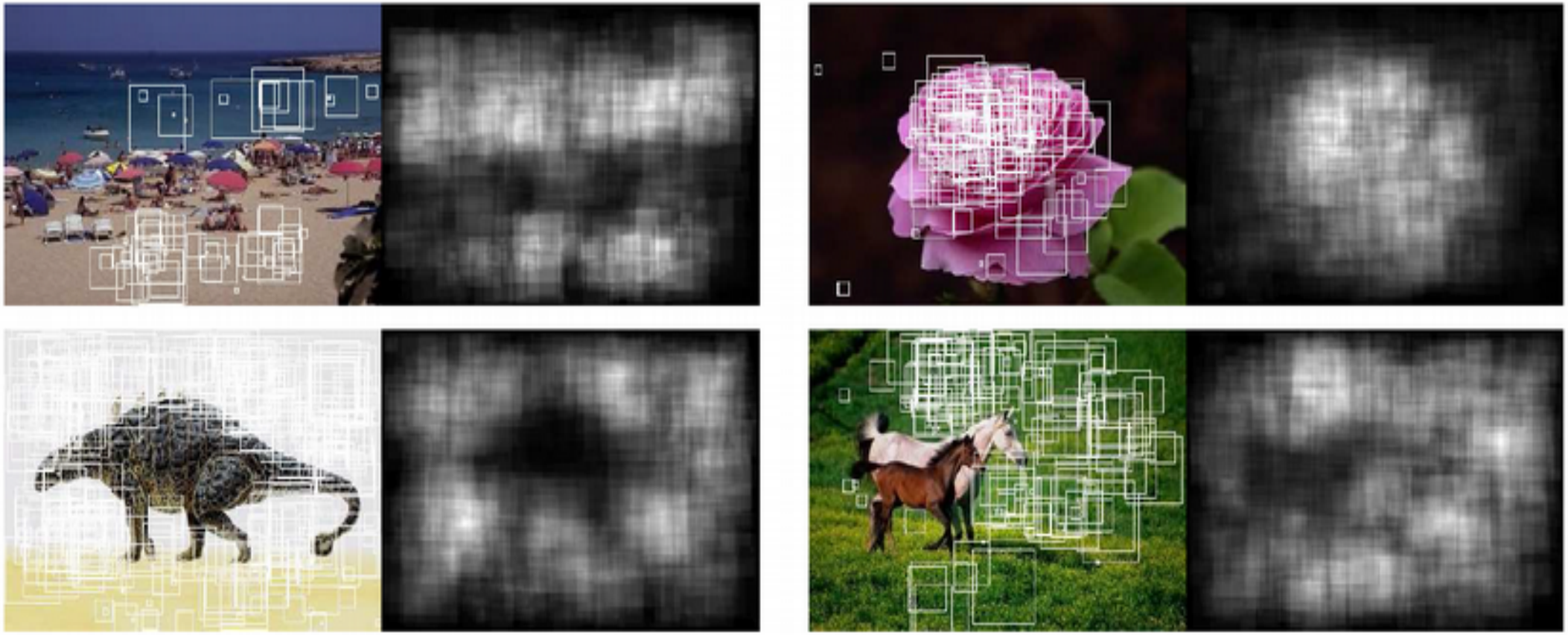
L. SHAMIR
Department of Computer Science, Lawrence Technological University, Southfield, Michigan, U.S.A.



→ 88 % recognition rate using images without protein patterns !

Benchmark dataset issues : hidden artefacts

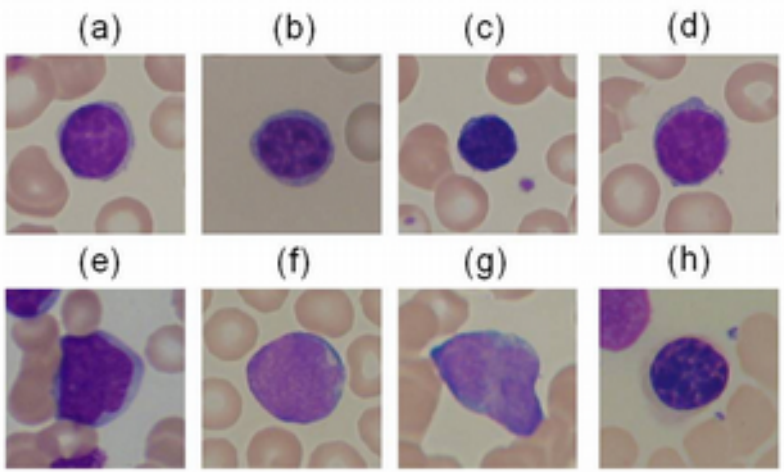
WANG dataset (PAMI, 2001) : 10 categories (beach, dinosaur, flower, horse, food, city, ...)



→ 44 % recognition rate using only 50x50 background data… OK ?

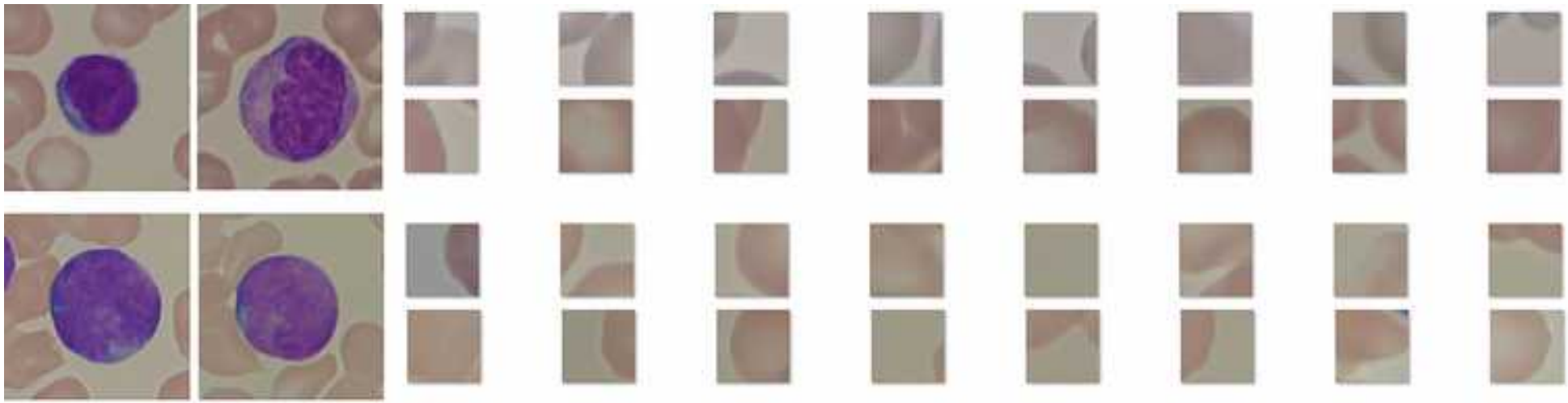
NO ! Two classes (dinosaurs & horses) are almost perfectly recognized using background only !

Benchmark dataset issues : hidden artefacts



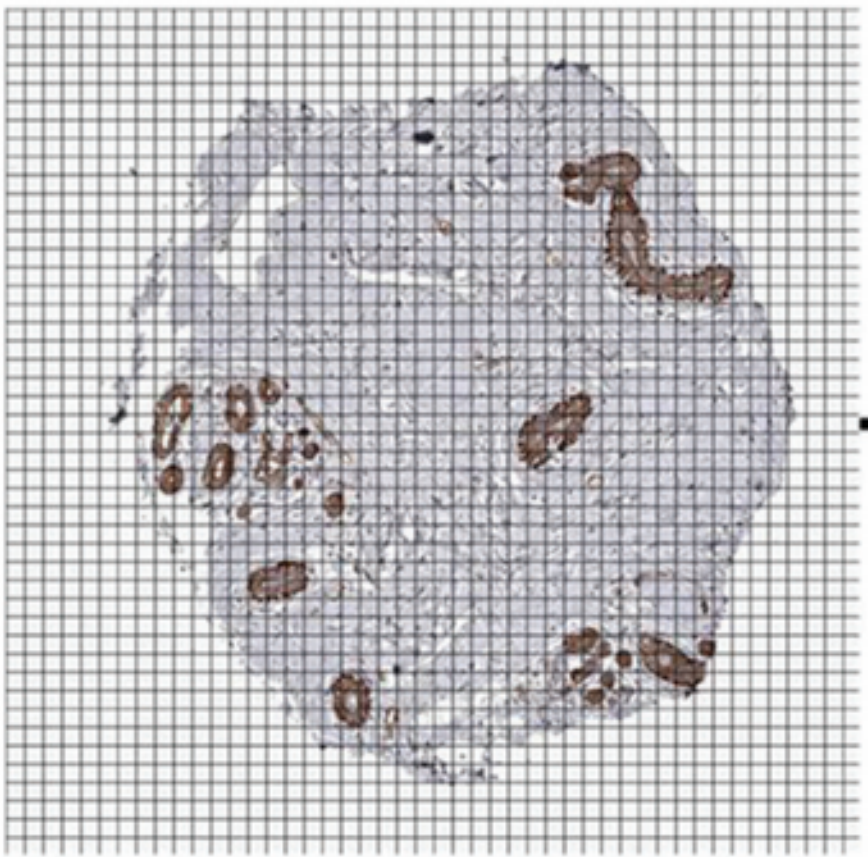
Examples of the images contained in ALL-IDB2: healthy cells from non-ALL patients (a-d), probable lymphoblasts from ALL patients (e-h).

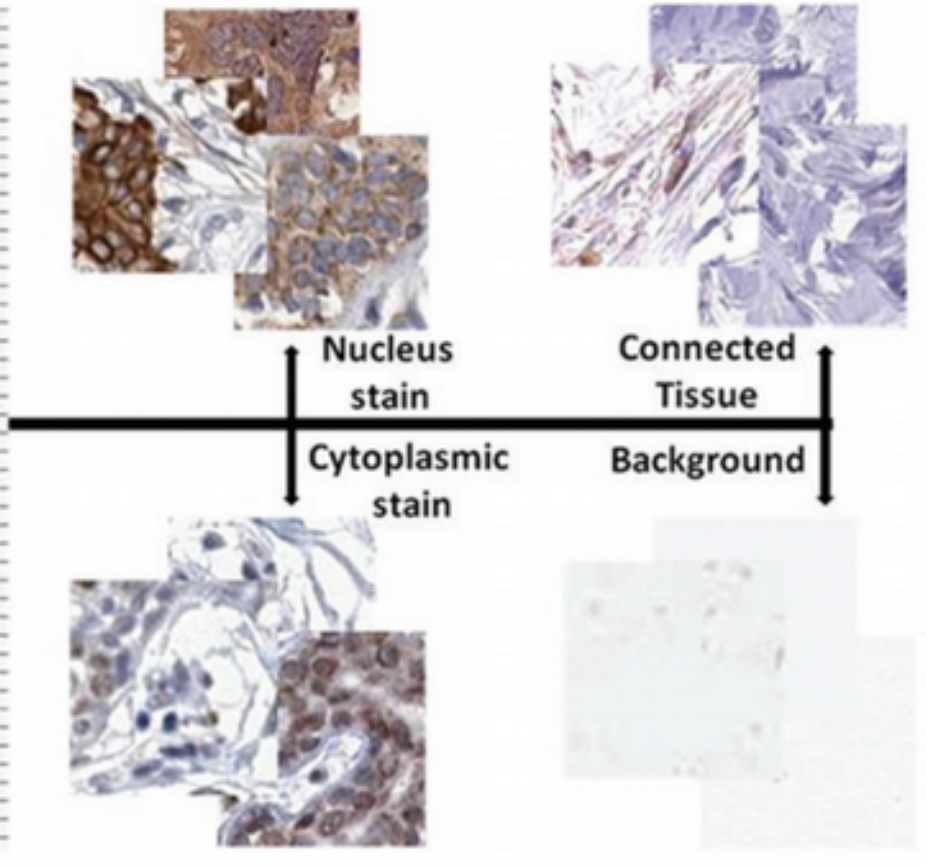
ALL-IDB: the acute lymphoblastic leukemia image database for image processing, Proc. IEEE Int. Conf. on Image Processing (ICIP 2011).



→ 90 % recognition rate using only 50x50 background regions !

Benchmark dataset issues : hidden artefacts



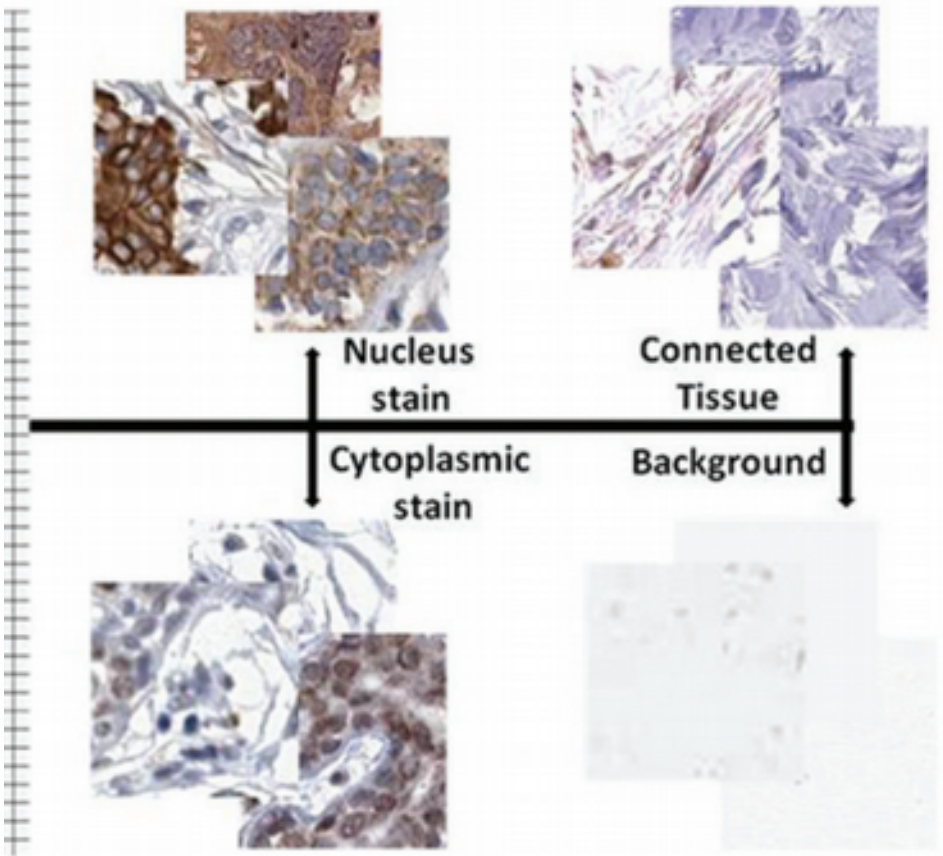


Benchmark dataset issues : hidden artefacts

Table 1: Classification result using different features (where N-Nucleus stains, C-Cytoplasmic stains, T-Connected tissue, B-Background)

Feature-extraction technique (number of features)	Correctly classified patches (%)				Overall accuracy (%)
	N	C	T	B	
GLCM (9)	25,0	77,5	65,8	83,9	66,6
DT-CWT (18)	36,5	87,0	36,6	97,4	69,7
UAADT-CWT (19)	34,0	90,7	89,1	92,8	80,0
GLCM+DT-CWT (27)	68,0	82,9	79,7	97,4	82,2
GLCM+UAADT-CWT (28)	68,5	85,5	93,1	98,5	86,0
PCA-LDA-CHARM (661)	76,0	91,0	100,0	100,0	90,0
Expert 1 (intra-obs. variability)	89,5	95,7	99,0	100,0	95,9
Expert 2 (inter-obs. variability)	94,5	84,9	100,0	100,0	92,3

DT-CWT: Dual-tree complex wavelet transform, UAADT: Undecimated adaptive anisotropic dual-tree complex wavelet transform; GLCM: Gray level co-occurrence matrix; PCA: Principal component analysis; LDA: Linear discriminant analysis; CHARM: Compound hierarchy of algorithms representing morphology; UAADT: Undecimated adaptive anisotropic dual-tree complex wavelet transform



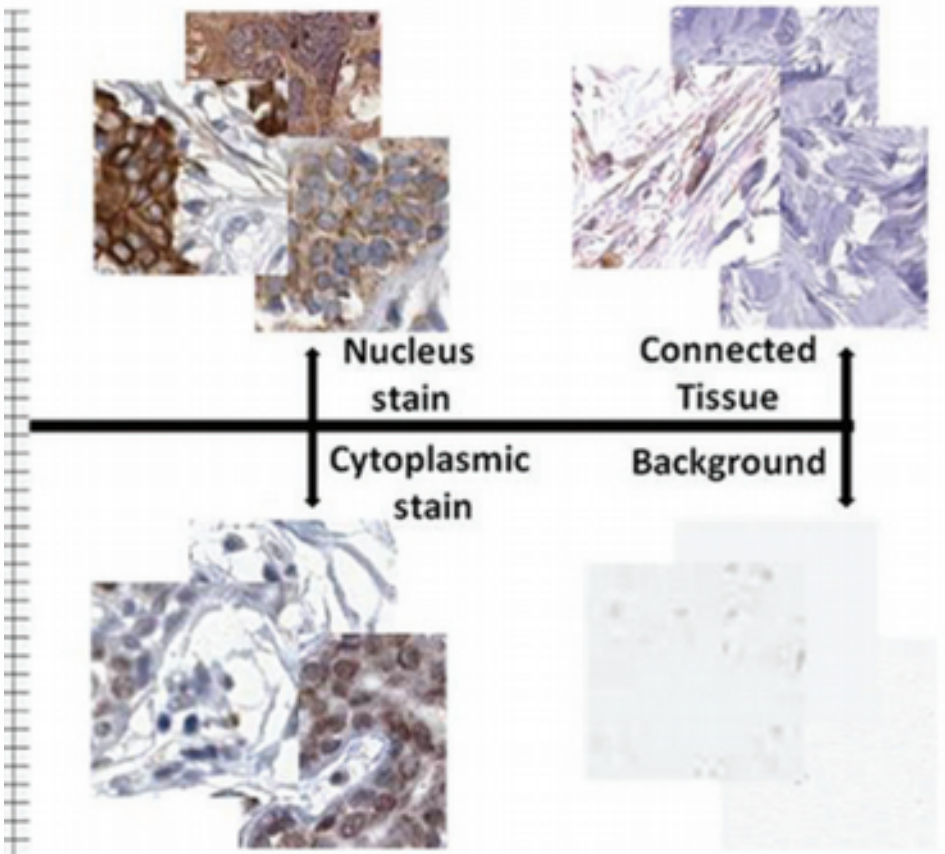
In this paper, we assume that the patches within a core are independent in terms of staining pattern and intensity as well as noise and artifacts. Further experiments are needed to verify that this assumption holds. (Swamidoss et al., 2013)

Benchmark dataset issues : hidden artefacts

Table 1: Classification result using different features (where N-Nucleus stains, C-Cytoplasmic stains, T-Connected tissue, B-Background)

Feature-extraction technique (number of features)	Correctly classified patches (%)				Overall accuracy (%)
	N	C	T	B	
GLCM (9)	25,0	77,5	65,8	83,9	66,6
DT-CWT (18)	36,5	87,0	36,6	97,4	69,7
UAADT-CWT (19)	34,0	90,7	89,1	92,8	80,0
GLCM+DT-CWT (27)	68,0	82,9	79,7	97,4	82,2
GLCM+UAADT-CWT (28)	68,5	85,5	93,1	98,5	86,0
PCA-LDA-CHARM (661)	76,0	91,0	100,0	100,0	90,0
Expert 1 (intra-obs. variability)	89,5	95,7	99,0	100,0	95,9
Expert 2 (inter-obs. variability)	94,5	84,9	100,0	100,0	92,3

DT-CWT: Dual-tree complex wavelet transform, UAADT: Undecimated adaptive anisotropic dual-tree complex wavelet transform; GLCM: Gray level co-occurrence matrix; PCA: Principal component analysis; LDA: Linear discriminant analysis; CHARM: Compound hierarchy of algorithms representing morphology; UAADT: Undecimated adaptive anisotropic dual-tree complex wavelet transform



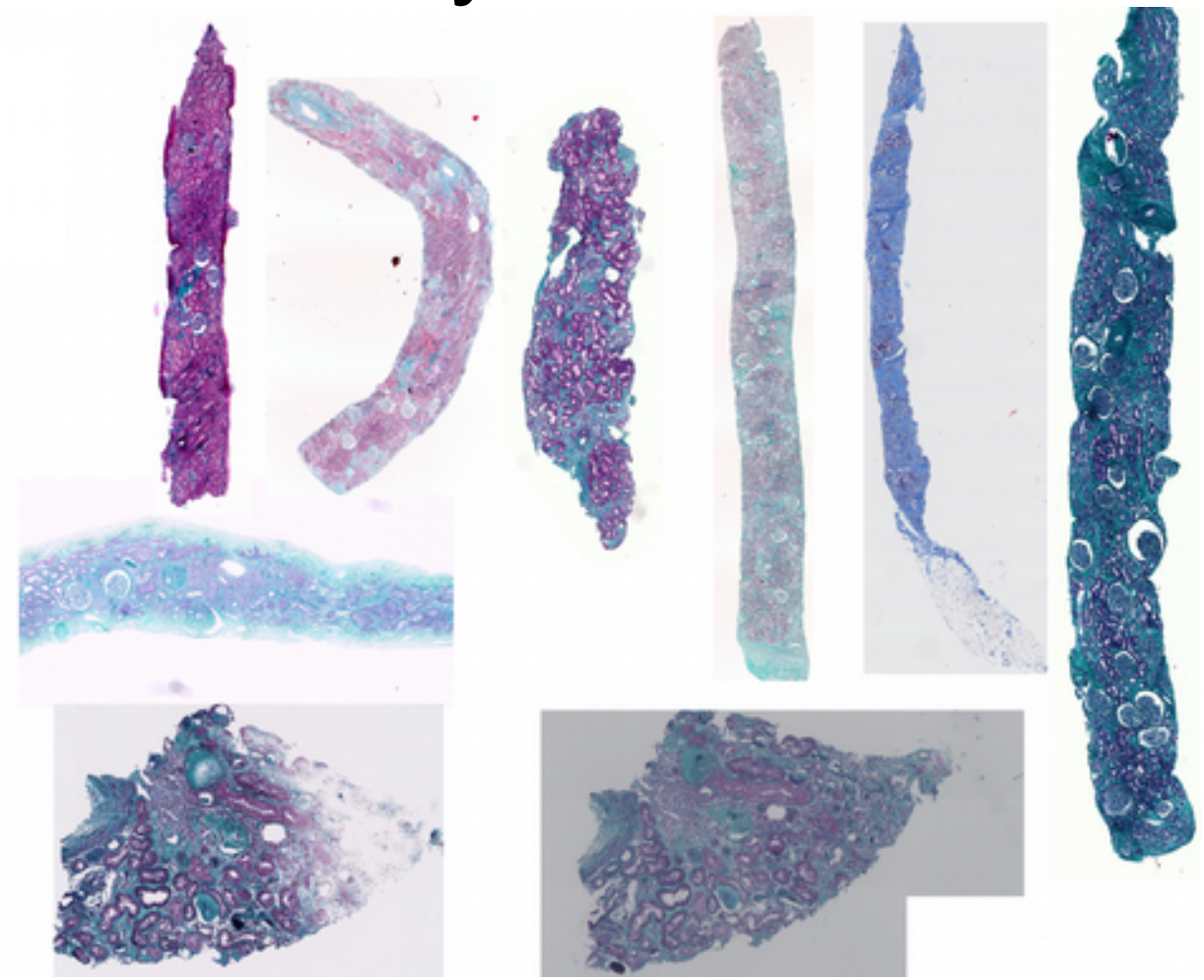
In this paper, we assume that the patches within a core are independent in terms of staining pattern and intensity as well as noise and artifacts. Further experiments are needed to verify that this assumption holds. (Swamidoss et al., 2013)

→ 95 % overall accuracy with « global histograms » learned by Extra-trees built on single pixels.

Potential sources of variability/artefacts in DP

Many sources of variability :

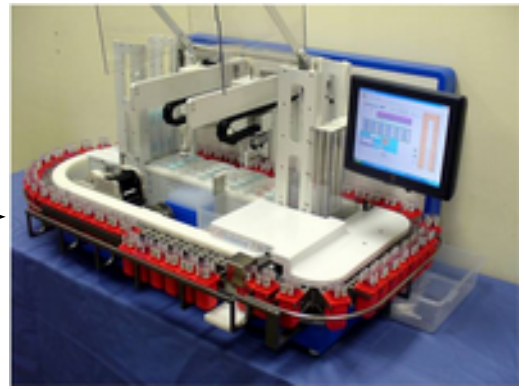
- Tissue (healthy or not, age, …)
- Preparation protocols (staining, sectioning, fixation,…)
- Acquisition setups (slide scanners, acquisition parameters, …)
- Image coding (quality/compression...)



Data collection issues :

- Sample collection bias : e.g. collecting all examples of a given class (e.g. positive cells) from a subset of slides while objects of another class (e.g. negative cells) are collected from another subset of slides.
  → Slide-specific patterns rather than class-specific

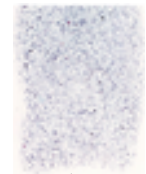
Cervical Cancer screening : hybrid workflow



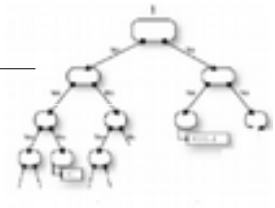
120 slides/hour

CellSolutions®
INNOVATORS IN CYTOLOGY

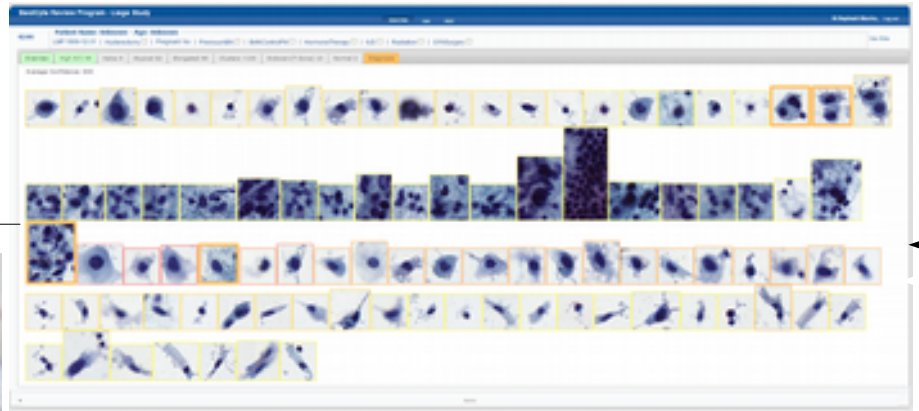
~ 60K cells per slide
~ 65K x 135K pixels



20X



Classify & rank every cells (3min/slide)



Cytotechnologist + pathologist web review of most suspicious cells according to our image recognition algorithms

Final diagnosis

Evaluation of CellSolutions BestPrep Automated Thin-Layer Liquid-Based Cytology Papanicolaou Slide Preparation and BestCyte Cell Sorter Imaging System,
Delga et al. , Acta Cytologica, 2014;58(5):469-77

Data collection guidelines & Quality Control

→ We need ***better*** datasets to train pattern recognition algorithms

Better ? Some suggestions (Marée, Journal of Pathology Informatics 2017):

- Collect examples for each object category from different slide id / sample / scanner / day / lab / pathologist and keep track of provenance to control hidden relationships
- Cover variability of objects (shape, texture, size, color, ...), not only typical ones. Also include an 'others' class (e.g. dust particles, bubbles, various contaminants) to avoid detection of too much false positives ;
- Balance class distributions and follow the experts' annotation process.
- Check dataset with simple approaches (e.g. global histograms ; background regions)
- Realistic evaluation protocols to assess recognition robustness: stratified random sampling + slide hold-out validation

How ? We need **collaborative software** platform for :

- Representative ground-truth creation
- Validation of pattern recognition algorithms on a large-scale

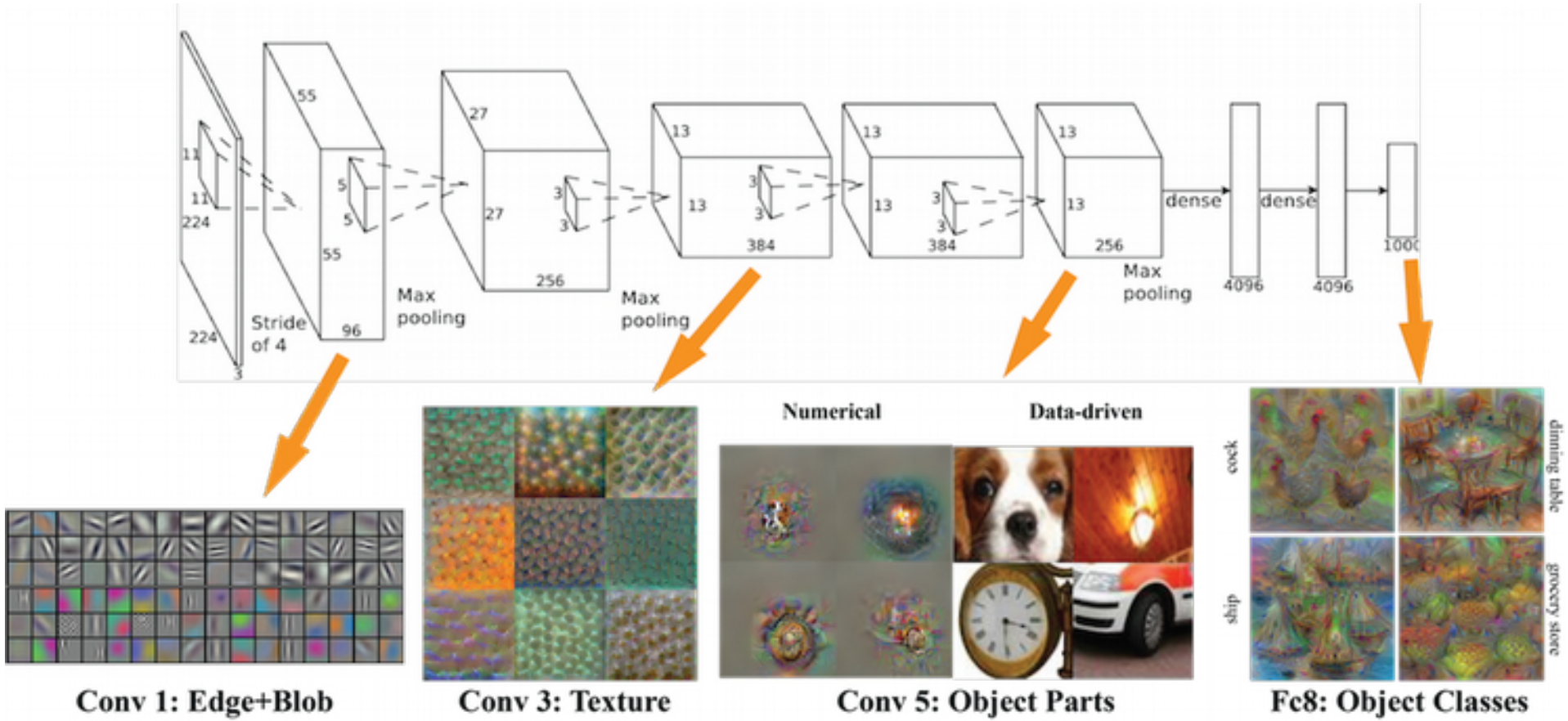
cytomine

From research to real-world

- The need for realistic data collection

- **Recent trends**
  - **Deep learning**
  - Multispectral, Multimodal imaging
  - Open hardware/software

A simplified view of Deep Learning



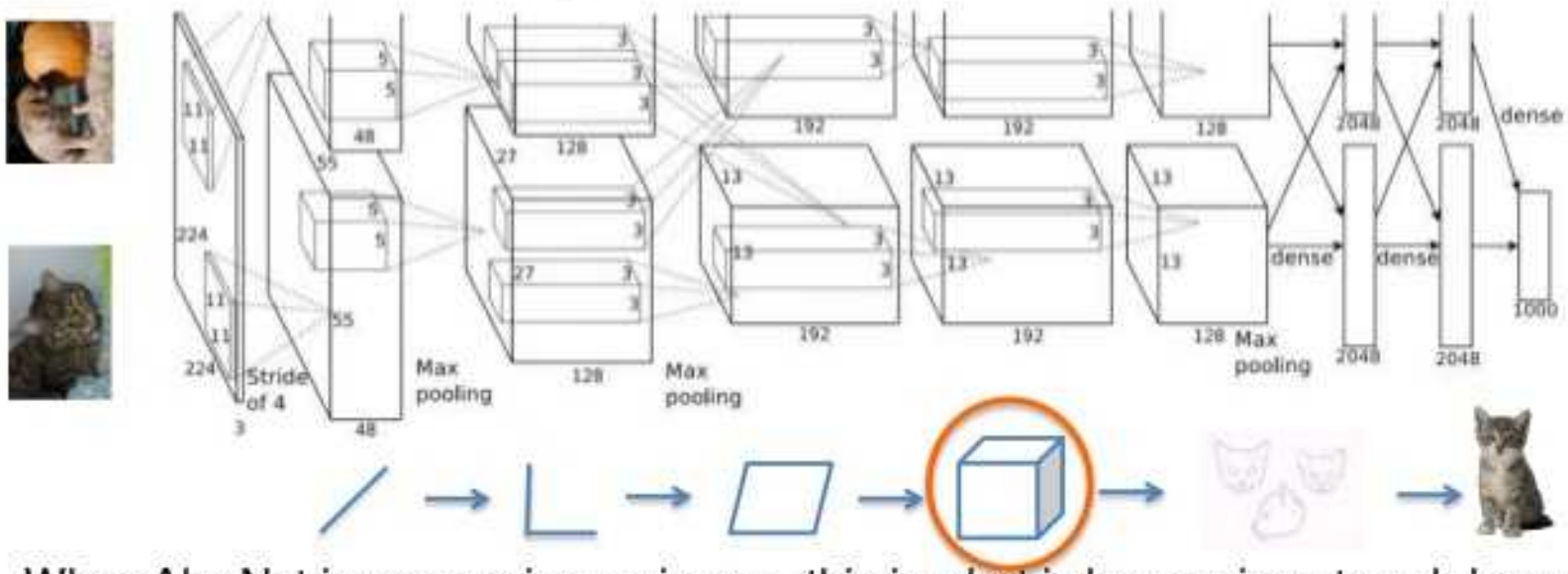
- Convolutions: see Chapter 11, or “Computer vision” lecture
- Pooling
- Backpropagation: see “Machine learning” lecture
- Data

(Farabet et al., PAMI 2013)

Deep learning architectures

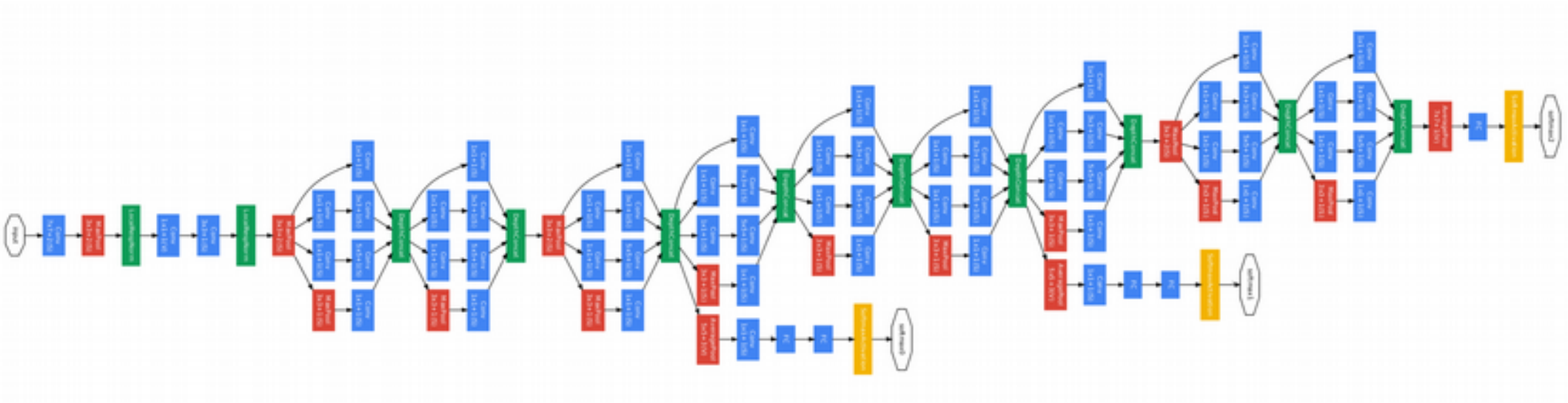
AlexNet (Krizhevsky et al. 2012)

The class with the highest likelihood is the one the DNN selects



When AlexNet is processing an image, this is what is happening at each layer.

Deep learning architectures



(GoogLeNet; Szegedy et al., 2015)

Deep learning issues

Requires very large datasets for training

Requires architecture design and tuning (trial & error)

Not interpretable/user-friendly (black blox)

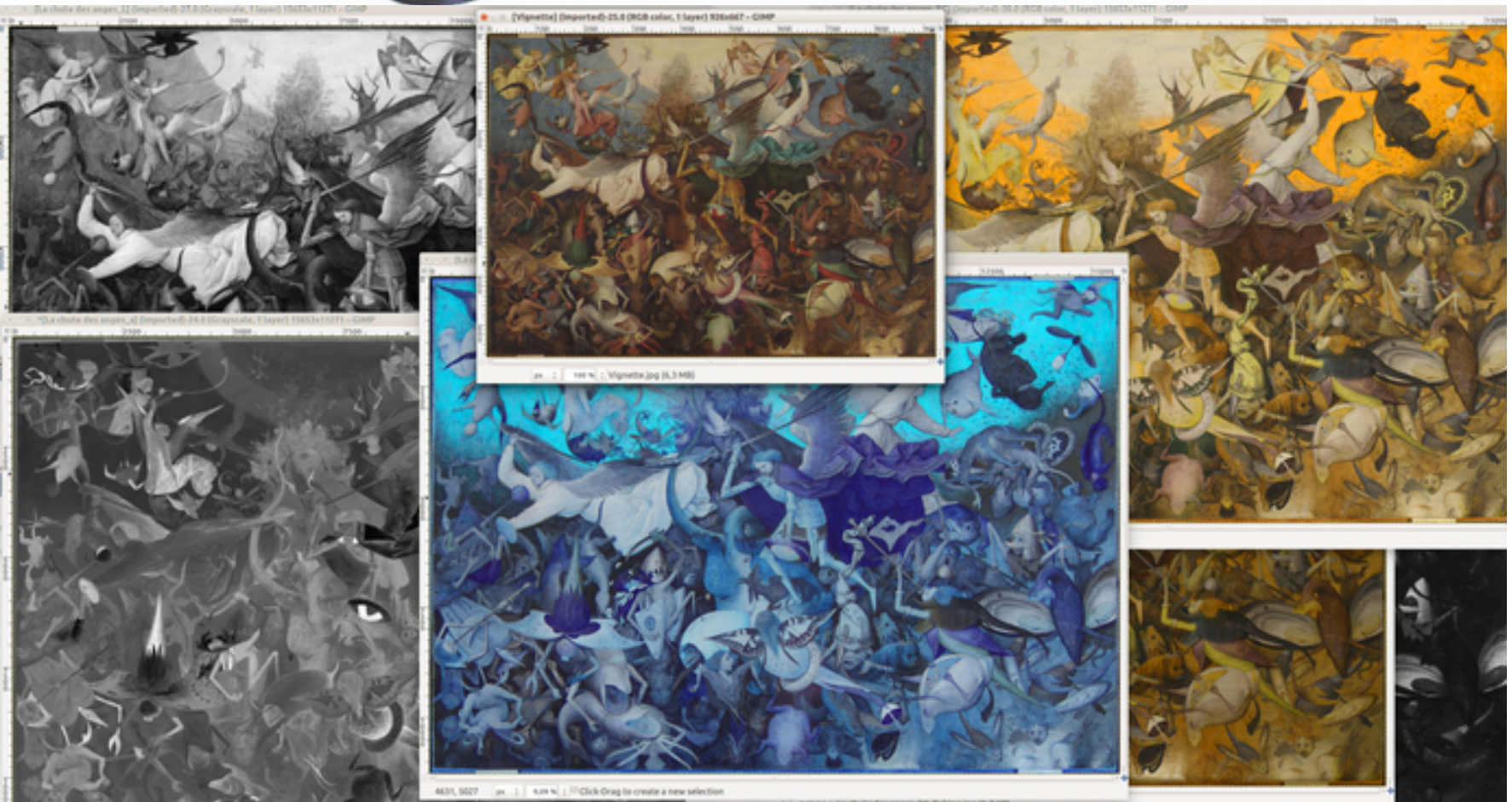
Computationally intensive

From research to real-world

- The need for more realistic data collection

- **Recent trends**
  - Deep learning
  - **Multispectral, Multimodal imaging**
  - Open hardware/software

The **highest resolution** in the World	20 000 x 12 000 pixels (samples)
The **biggest sensor** in the World	78 x 130 mm
The **biggest scan area** in the World	5 x 2 m (6.5 X 16.40 feet)
The **fastest scanner** in the World	280 Mbits/sec. (view a video sample)



20 000

x

12 000

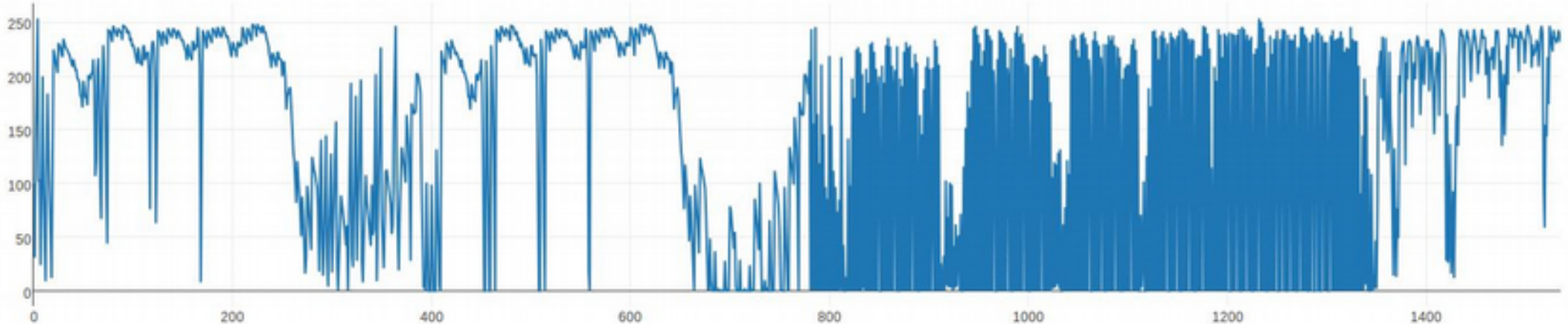
x

1650 bands

The **highest resolution** in the World	20 000 x 12 000 pixels (samples)
The **biggest sensor** in the World	78 x 130 mm
The **biggest scan area** in the World	5 x 2 m (6.5 X 16.40 feet)
The **fastest scanner** in the World	280 Mbits/sec. (view a video sample)

20 000

Spectral distribution



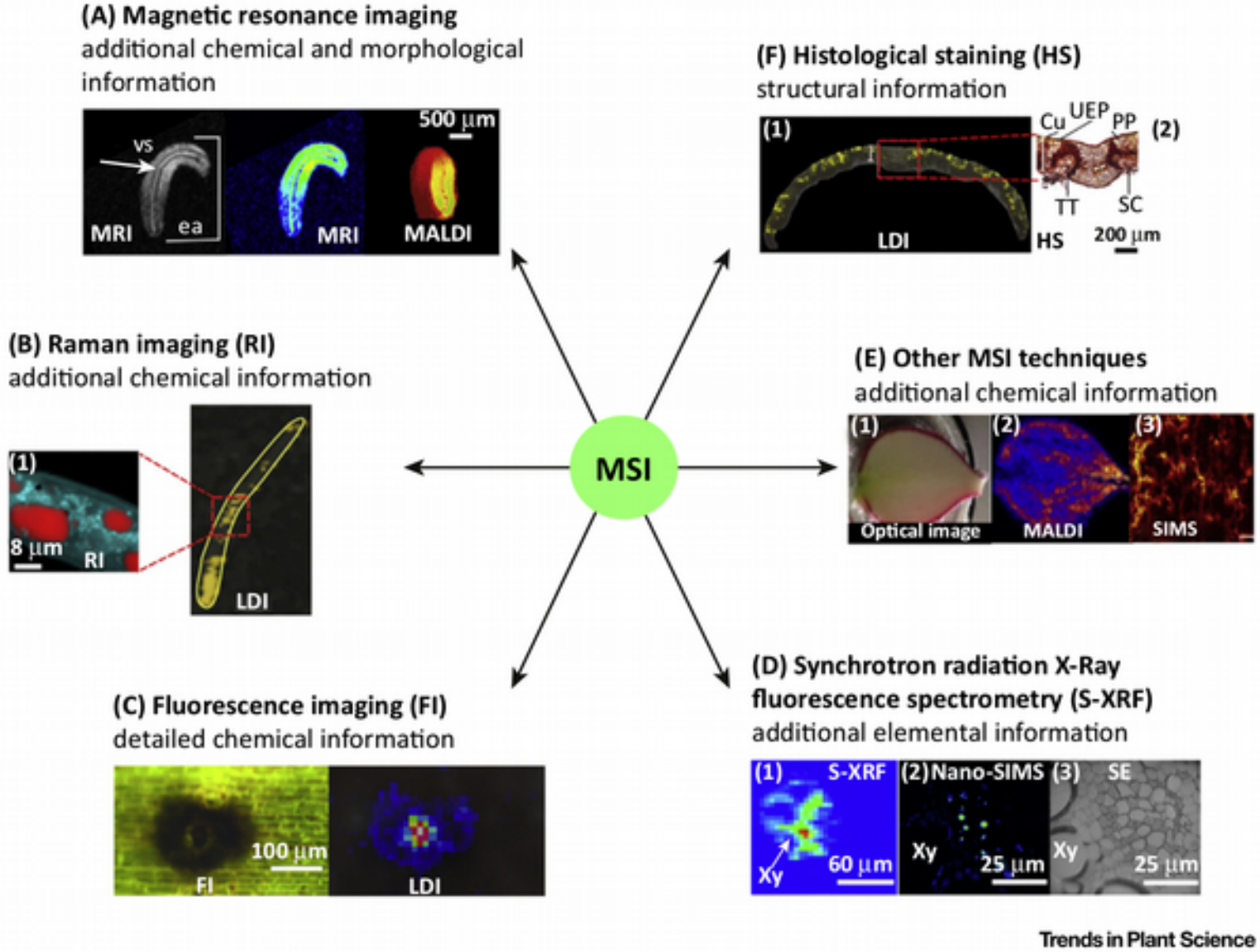


Figure 2. Selected Examples of Multimodal Mass spectrometry imaging (MSI) in Plant Research. (A) Matrix-assisted laser desorption ionization (MALDI) and magnetic resonance imaging (MRI) of lipids in *Camelina sativa* seed. (B) Laser desorption ionization (LDI) and Raman imaging of yellow droplets in nematodes. (C) LDI and fluorescent imaging of hydroxyanigorufone in red stomata and epidermis tissue of *Musa acuminata* ssp. *zebrine* cv. 'Rowe Red'. (D) Nano-secondary ion mass spectrometry (SIMS) and synchrotron X-ray fluorescence imaging of arsenic (As) in a leaf sheath of rice. (E) MALDI and SIMS imaging of choline in in radish bulb. (F) LDI imaging of flavonoid vicenin-2 in the transverse leaf sections of *Leucopogon ericoides*. The section was stained with Sudan for histological analysis. Adapted, with permission, from [59] (A), [60] (B), [97] (C), [98] (D), [58] (E), and [99] (F). Abbreviations: Cu, cuticle; EA, embryonic axis; PP, photosynthetic parenchyma; SC, stomatic crypt; SE, secondary electron image; TT, tector trichome; UEP, upper epidermis; VS, vascular system; Xy, xylem.

Map Compare

search Help

Choose image: Purple tissu

Choose image: Blue tissu

Choose image: Fix tissu

Choose image: Overlay tissu

debug: MapLayer: Overlay tissu

14.23267, 81.55512 zoom=5 number of maps: 1 2 3 4

Multimodal imaging is everywhere

Biomedecine: many molecular probing and imaging techniques allow the detection of single molecules at multiple resolutions. using e.g. multiplexed immunohistochemistry, mass spectrometry imaging, Raman Scattering/vibrational microspectroscopy, immuno/auto-fluorescence, optical coherence tomography, etc.

Geoscience / Remote sensing: various sources of imagery can be used to highlit spectral, spatial and radioactive characteristics (e.g. near infrared, synthetic aperture radar, RGB, LIDAR, …)

Quantitative mineralogy: several techniques at multiple resolutions (e.g. hyperspecral, cross-/plane-polarized light, scanning/transmission electron microscopy, ...) are used to obtain data on structure, composition, texture/fabric, porosity, permeability, and other parameters to perform quantitative mineralogical analysis of samples.

Plant physiology and agronomy: different techniques (bioluminescence, chlorophyll fluorescence imaging, mass spectrometry imaging, ...) to study morphological, chemical, and structural plant parameters

Industrial inspection: non-destructive quality control testing can be done combining e.g. x-rays, thermal infrared, shearography, ...

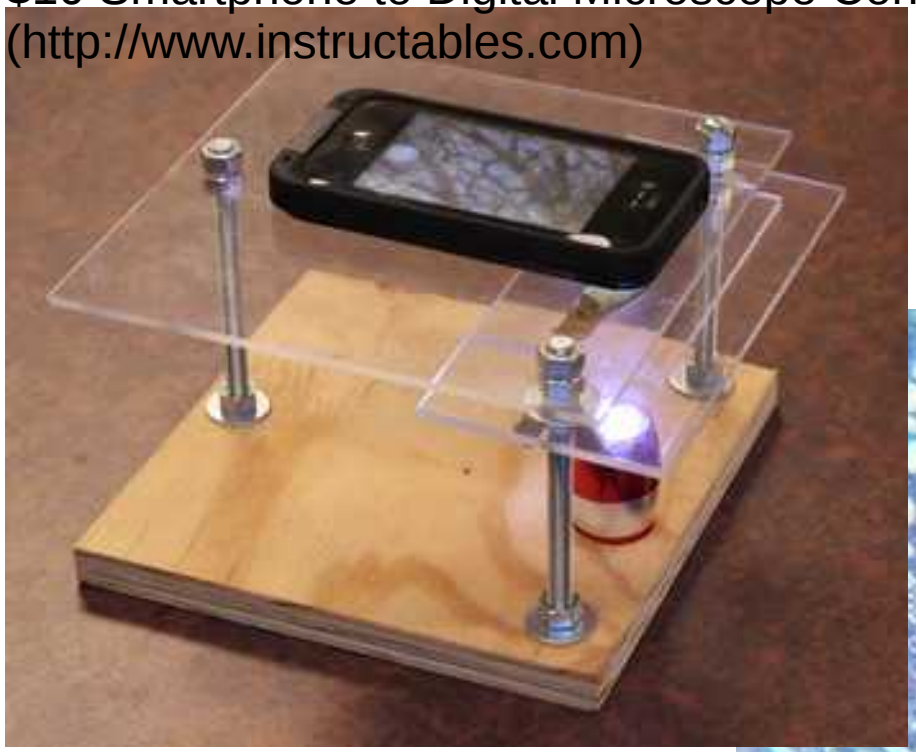
From research to real-world

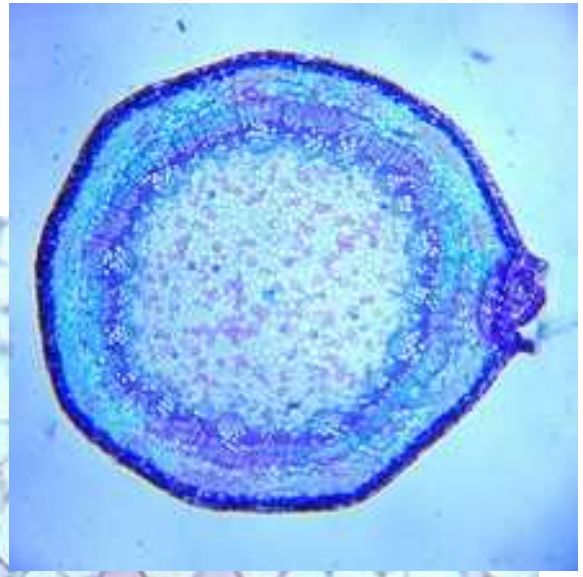
- The need for more realistic data collection

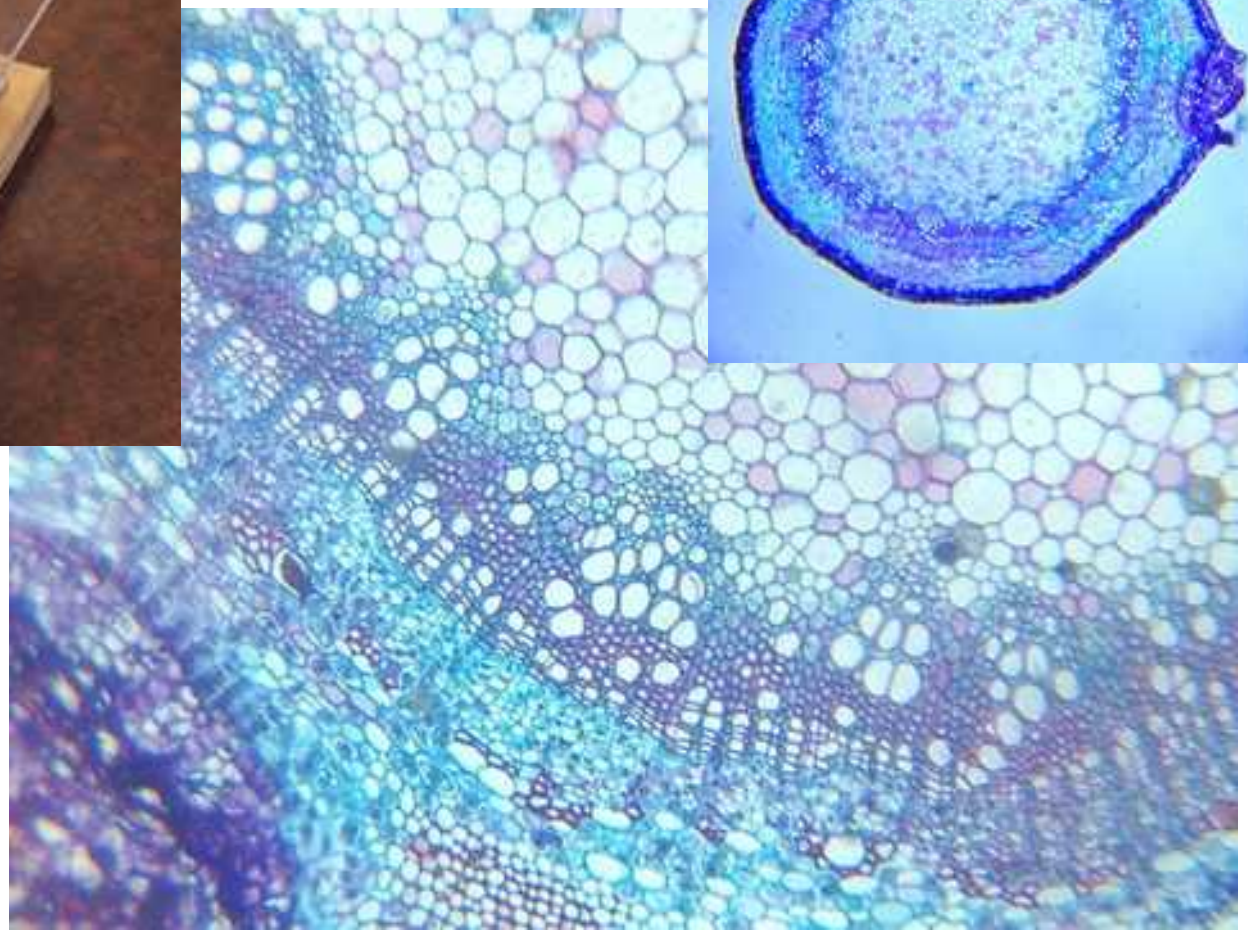
- **Recent trends**
  - Deep learning
  - Multispectral, Multimodal imaging
  - **Open hardware/software**

Towards intelligent microscopes

$10 Smartphone to Digital Microscope Conversion
(http://www.instructables.com)



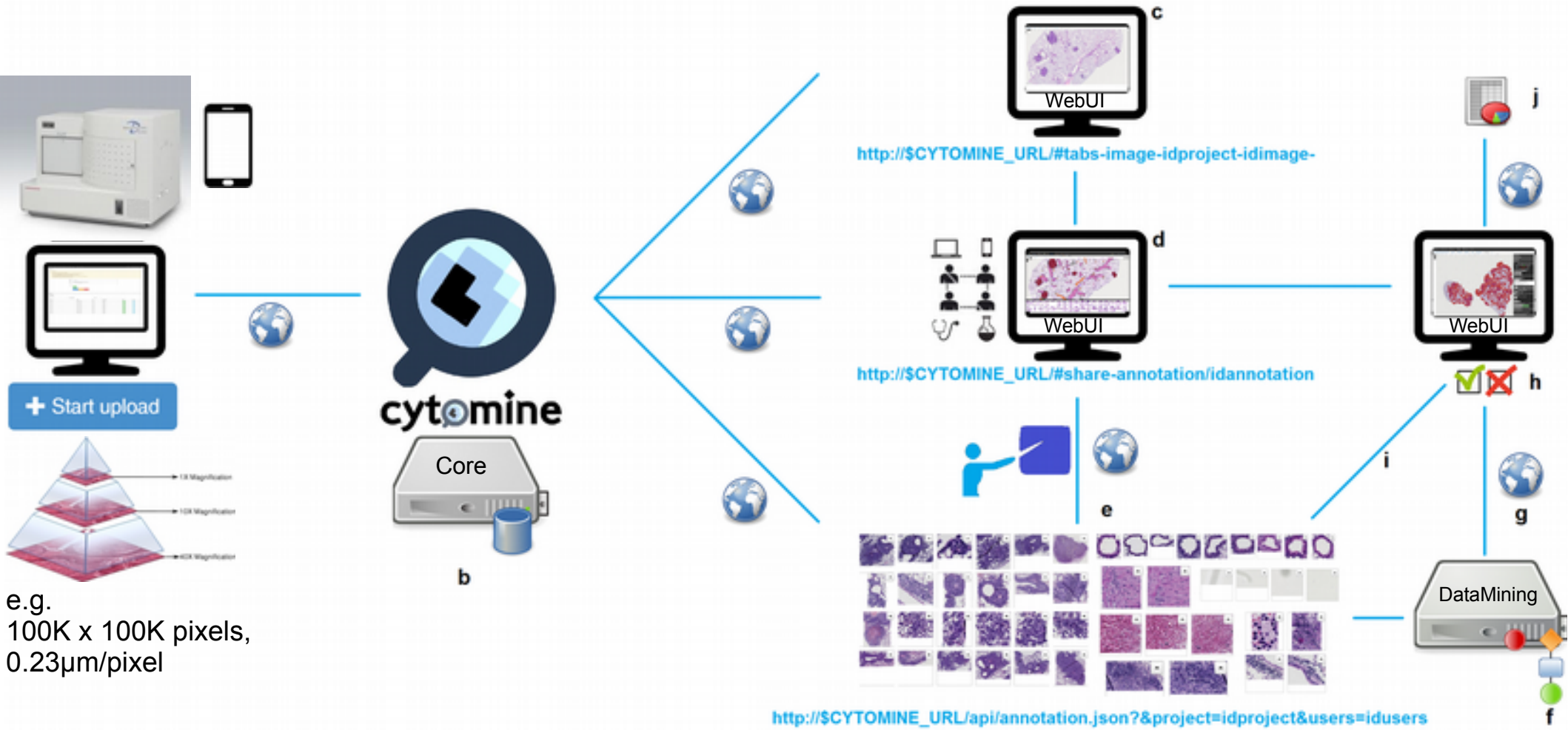




?

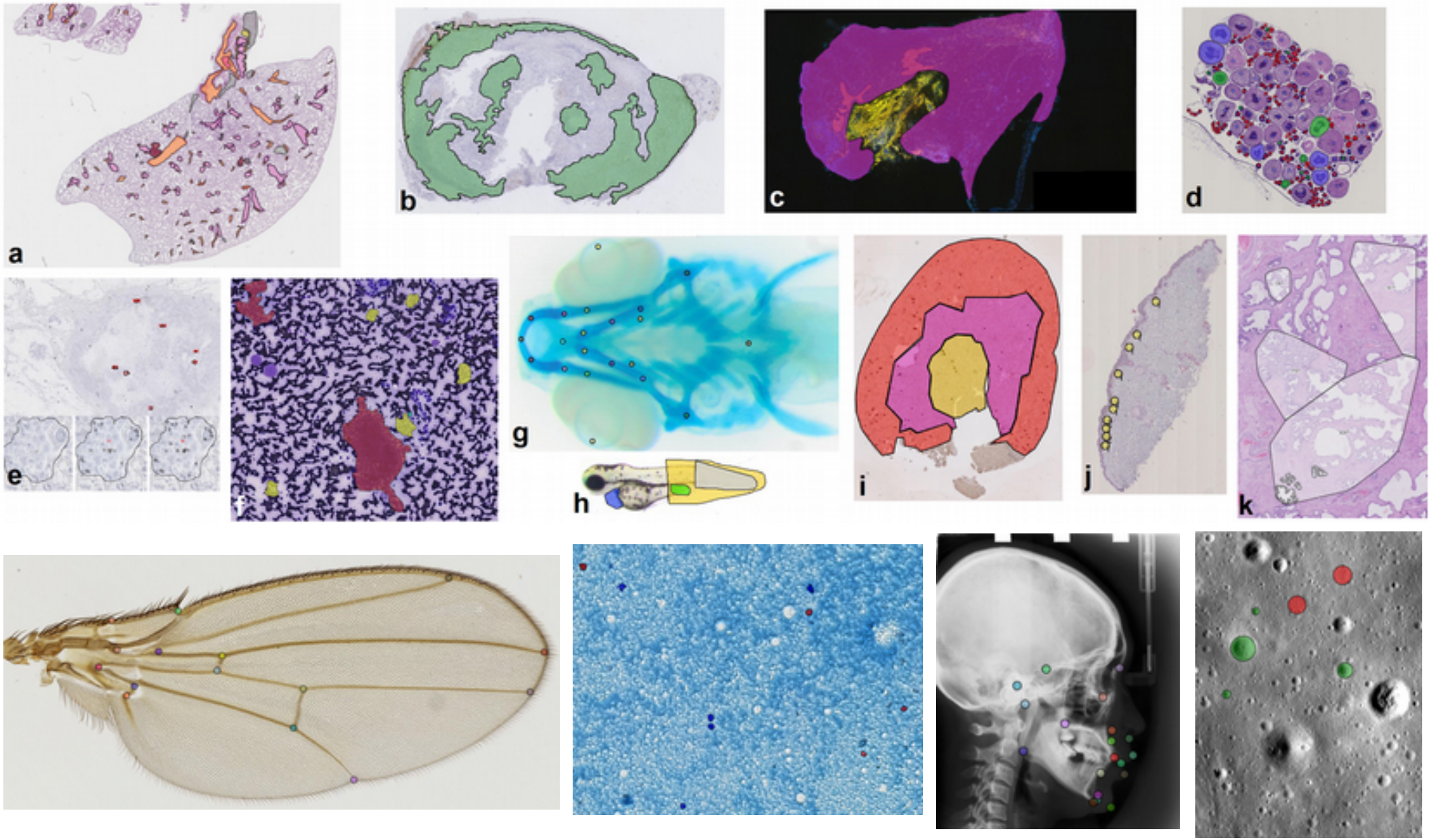
Web services

Towards intelligent microscopes

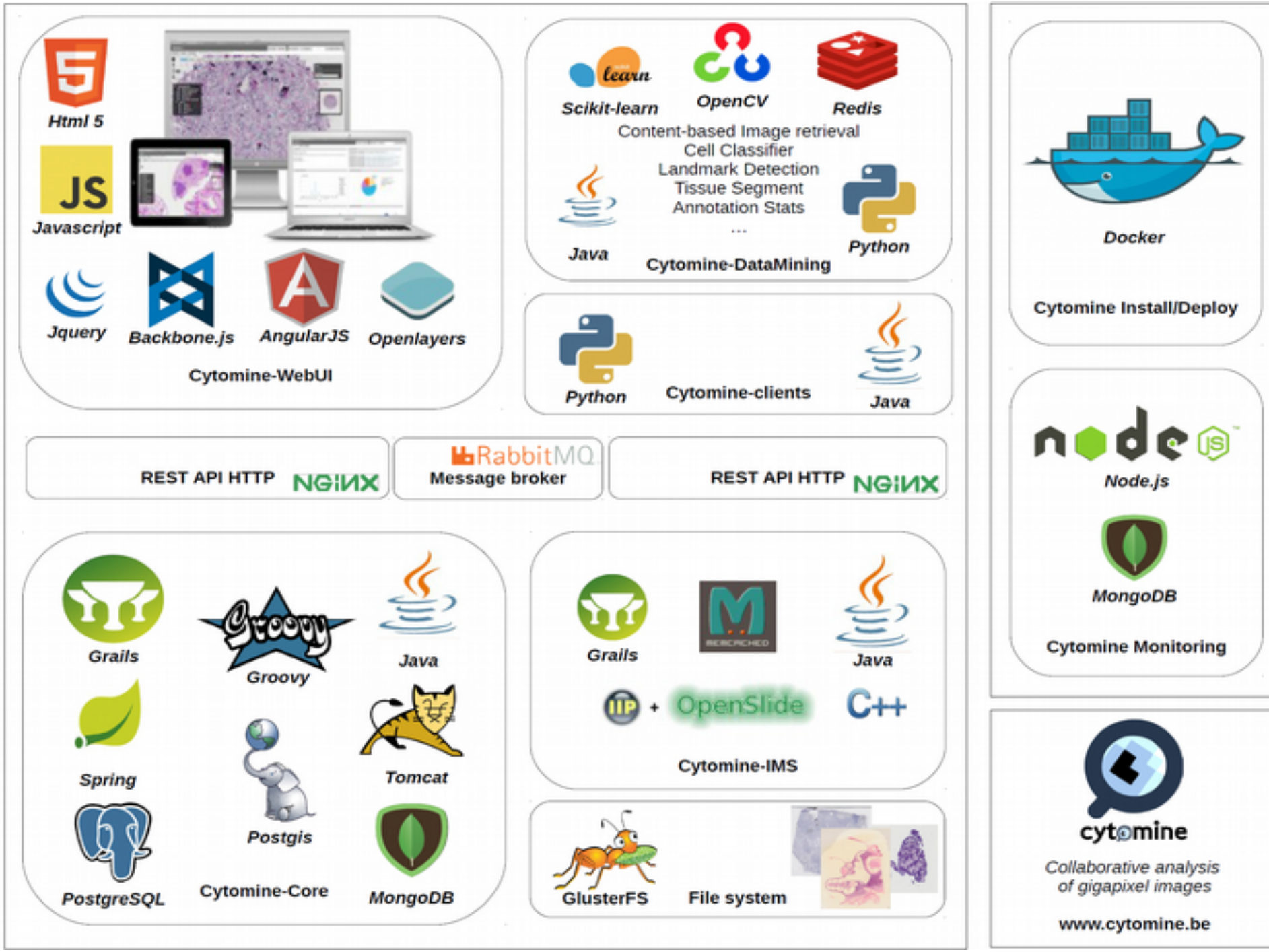


e.g.
100K x 100K pixels,
0.23µm/pixel

Towards intelligent microscopes



(www.cytomine.be/#publications)


Html 5
Javascript
Jquery
Backbone.js
AngularJS
Openlayers
Cytomine-WebUI
Scikit-learn
OpenCV
Redis
Content-based Image retrieval
Cell Classifier
Landmark Detection
Tissue Segment
Annotation Stats
...
Java
Python
Cytomine-DataMining
Python
Cytomine-clients
Java
Docker
Cytomine Install/Deploy
REST API HTTP
NGINX
RabbitMQ
Message broker
REST API HTTP
NGINX
Node.js
MongoDB
Cytomine Monitoring
Grails
Groovy
Java
Spring
Tomcat
Postgis
PostgreSQL
Cytomine-Core
MongoDB
Grails
Java
OpenSlide
C++
Cytomine-IMS
GlusterFS
File system
cytomine
Collaborative analysis
of gigapixel images
www.cytomine.be

From research to real-world

- The need for more realistic data collection
- Usable software
- **Recent trends**
  - Deep learning
  - Multispectral, Multimodal imaging
  - Open hardware/software

Potential Master Thesis (TFE)

Raphael.Maree@ulg.ac.be