Sublinear Evolutive Design of Error Correcting Output Codes

Student: Miguel Angel Bautista
Directors: Dr. Sergio Escalera & Dr. Xavier Baró

UNIVERSITAT DE BARCELONA

Centre de Visió per Computador

CCIA 2010
Tretzè Congrés Internacional de l'Associació Catalana d'Intel·ligència Artificial

Outline

- Categorization problems
- Error Correcting Output Codes
- SVMs with Gaussian-RBF kernel
- Genetic optimization
- Experiments & results
- Conclusions

Categorization problems

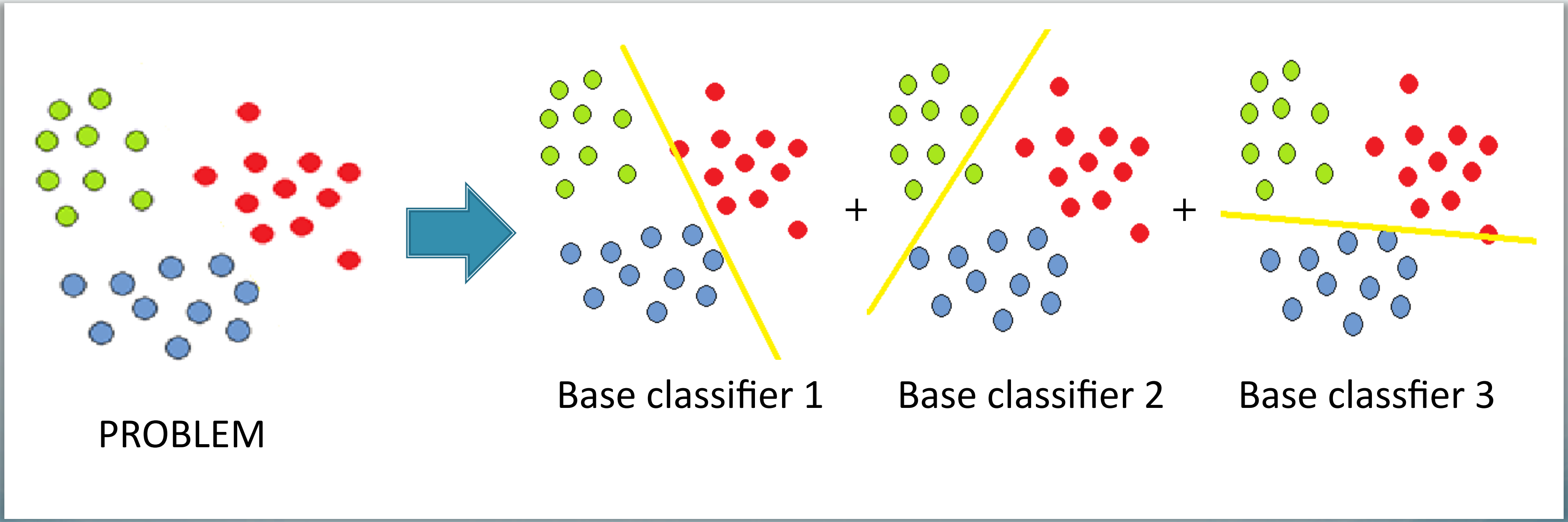
- Humans are involved in classification tasks from their early days.
- Classification is an unavoidable task in intelligent systems.

Multi-class categorization problems

- Real-world problems have more than 2 categories to identify.
- There are several ways to treat multi-class categorization problems.
- Ensemble learning techniques are often used in this type of scenarios.

Error Correcting Output Codes (ECOC)

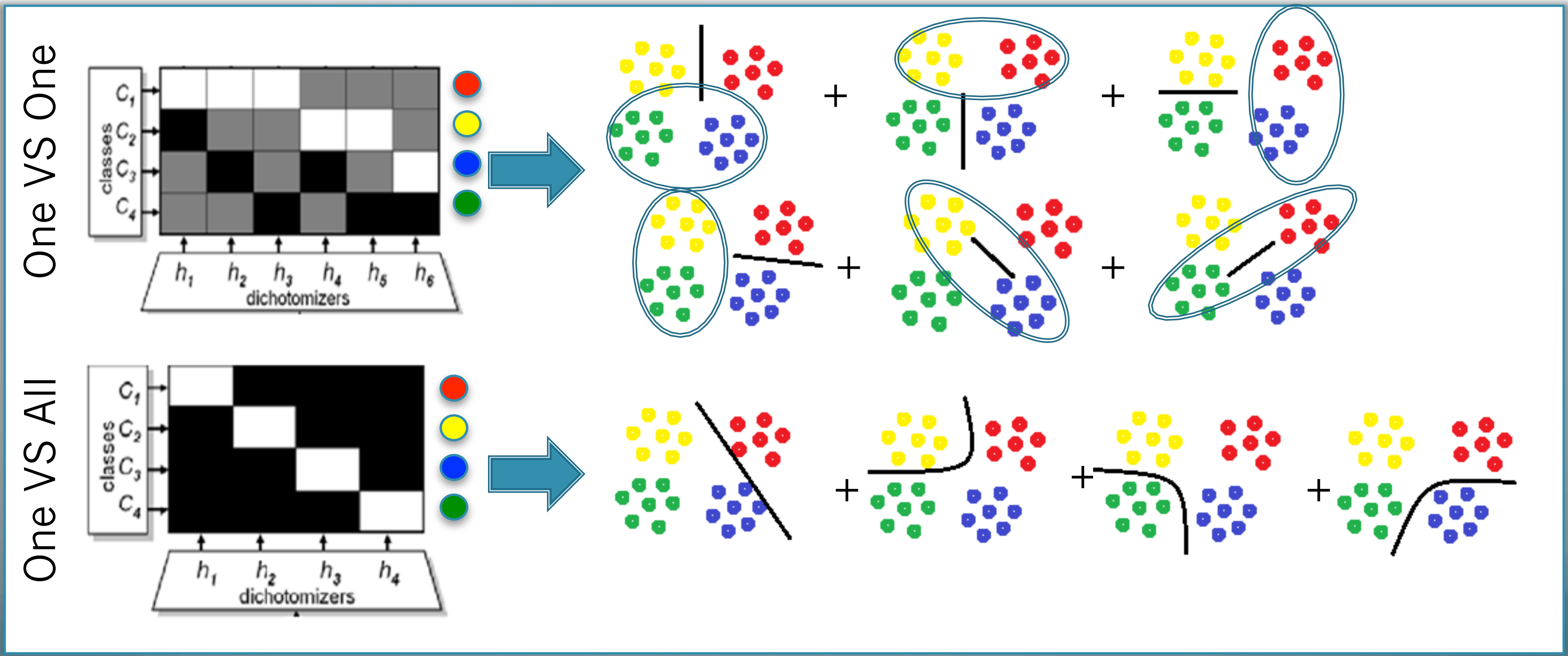
- ECOCs are an ensemble learning methodology which allow to combine dichotomizers (base classifiers) to treat multiclass problems.



ECOC coding

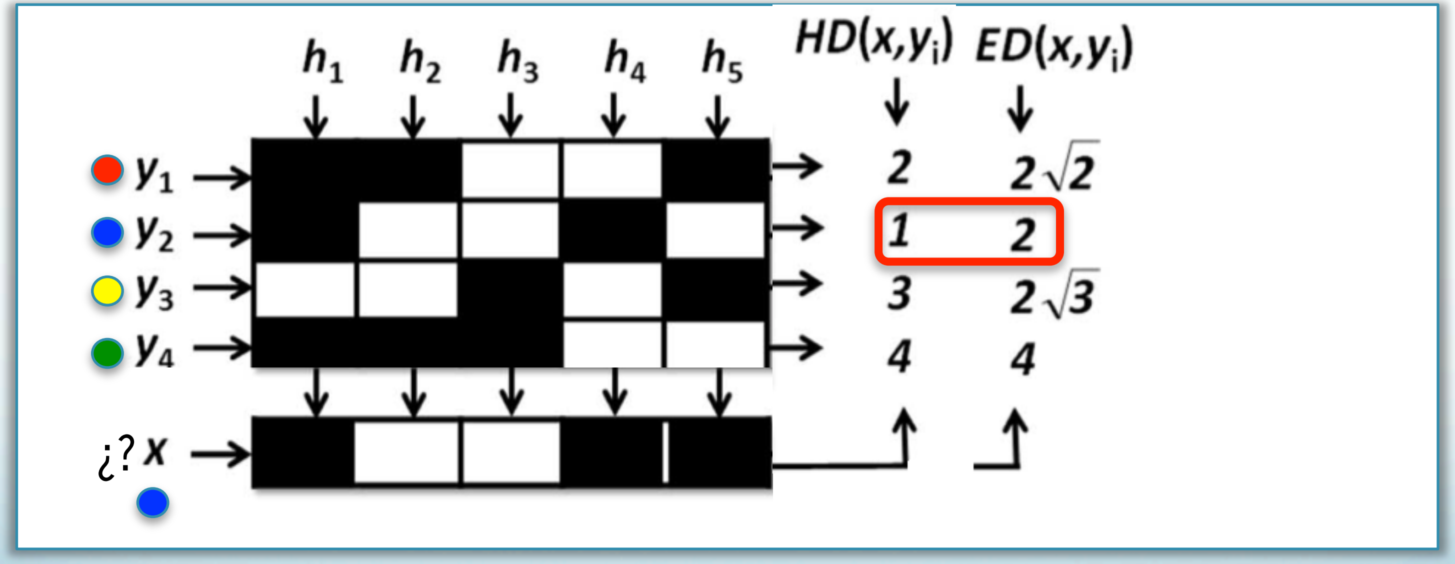
- ECOCs can be represented as matrices, which columns represent the different sub-problems to treat.

- Each column has values that distinguish categories in two groups.



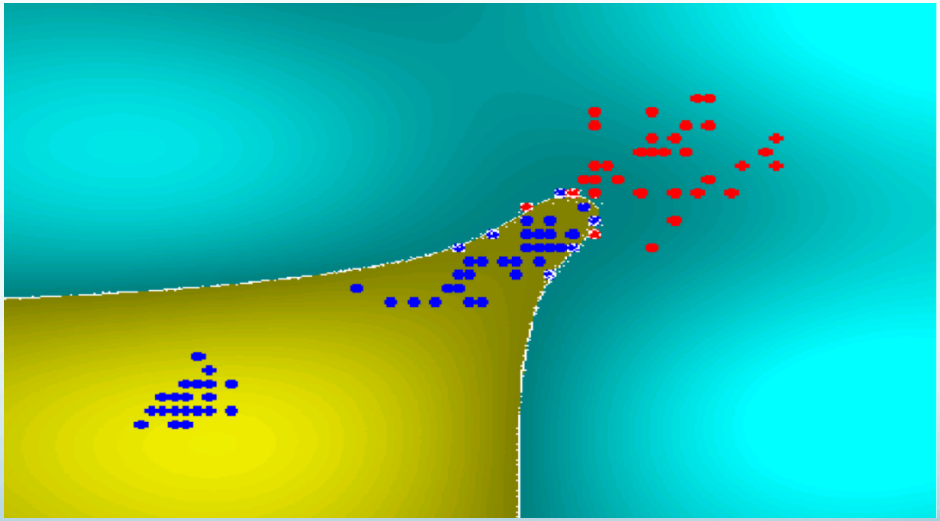
ECOC decoding

- Each sub-problem is trained and the set of predictions are compared to the codewords.
- Various types of decoding based on Euclidean and Hamming distances (only binary codings).



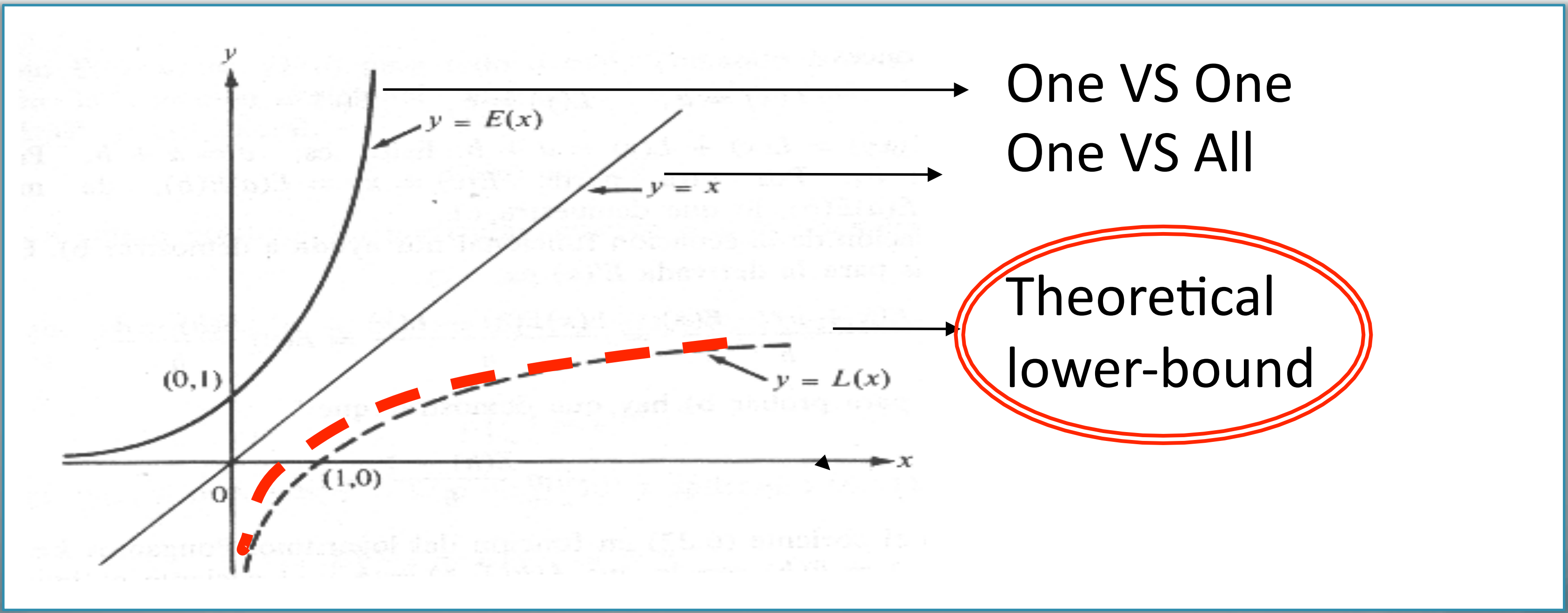
Base classifier: SVM with an RBF kernel

- Each binary problem is learned by a base classifier.
- SVM with RBF kernels have shown a good performance on those kind of problems.
- This type of SVM needs the parameters (C & Gamma) to be optimized.



Motivation: complexity in terms of the number of classifiers used.

- The number of classifiers needed by state-of-the-art approaches becomes inefficient when the number of classes in the problem increases.



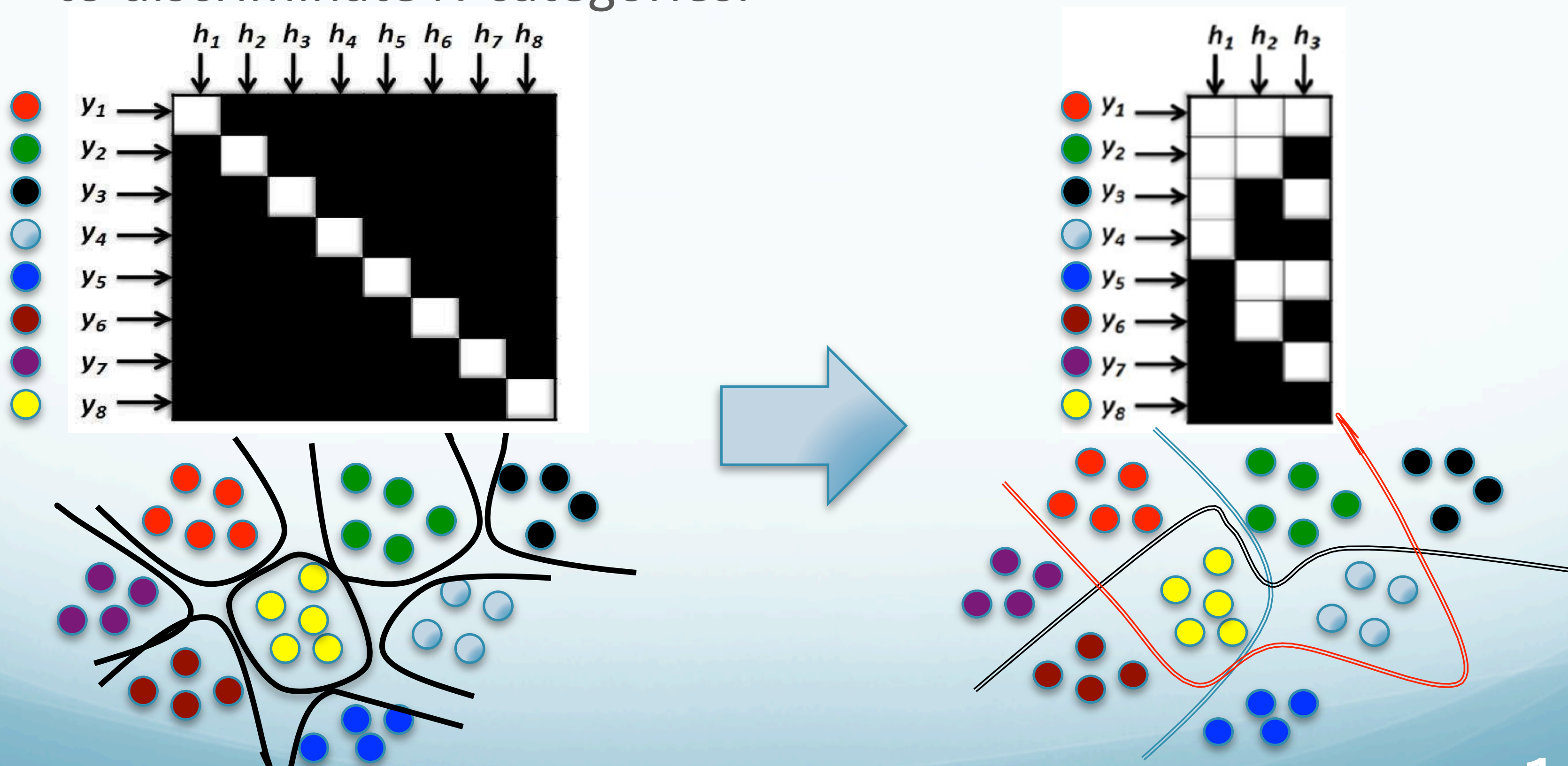
Global overview

Sublinear ECOC coding

Joint Genetic optimization of ECOC & Base classifier

Sublinear coding

- Define the lowest number of base classifiers needed to discriminate *N* categories.
- Taking profit of Information theory only $\log_2 N$ bits are needed to discriminate *N* categories.



Global overview

Sublinear ECOC coding

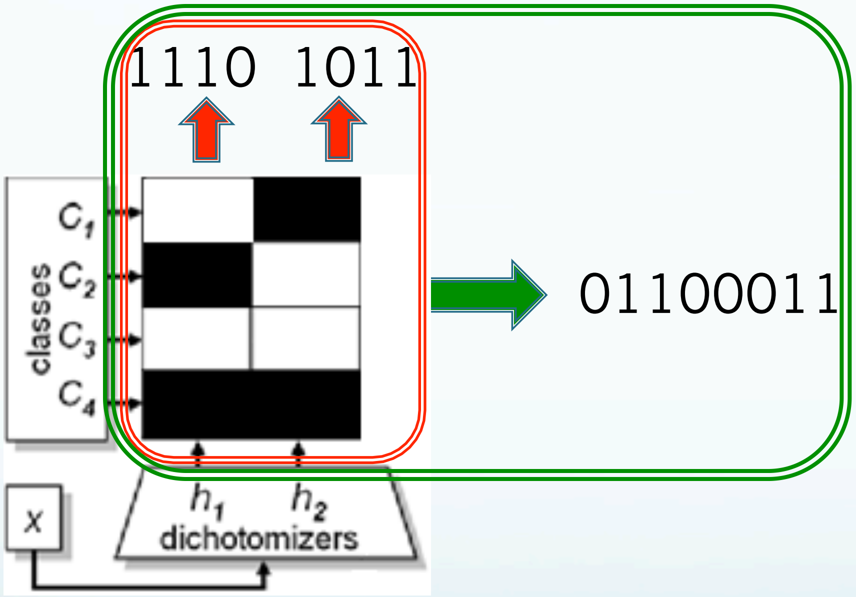
Joint Genetic optimization of ECOC & Base classifier

Genetic algorithms

- Optimization algorithms based on the evolution theory of Darwin.
  - Optimization processes based on evolution of *individuals.*
  - Each possible solution is coded into a *chromosome.*
  - Individuals are evaluated by means of its adaptation to the environment.
- Recommendable method when the space is not continuous neither differentiable.

Evolutionary optimization I

- Each ECOC individual is seen as a binary vector (chromosome) and evaluated by means of its classification error.



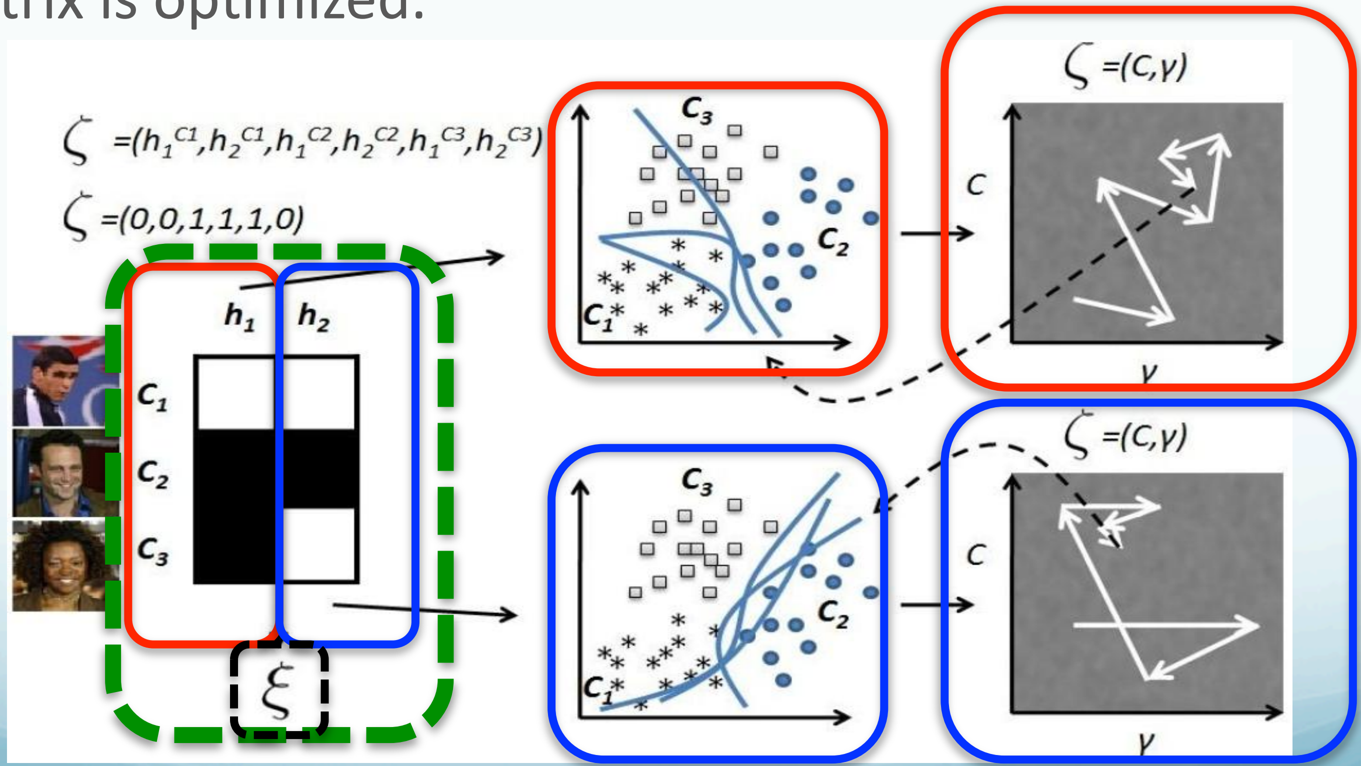
A) Optimize the SVMs looking for suitable parameters.

B) Optimize the coding matrix and return to step A.

- Standard genetic operators are used, scattered crossover and gaussian add unit mutation.

Evolutionary optimization II

- An inner optimization process is carried out to tune the parameters of the SVMs.
- Once each base classifier is optimized, the Sublinear coding matrix is optimized.



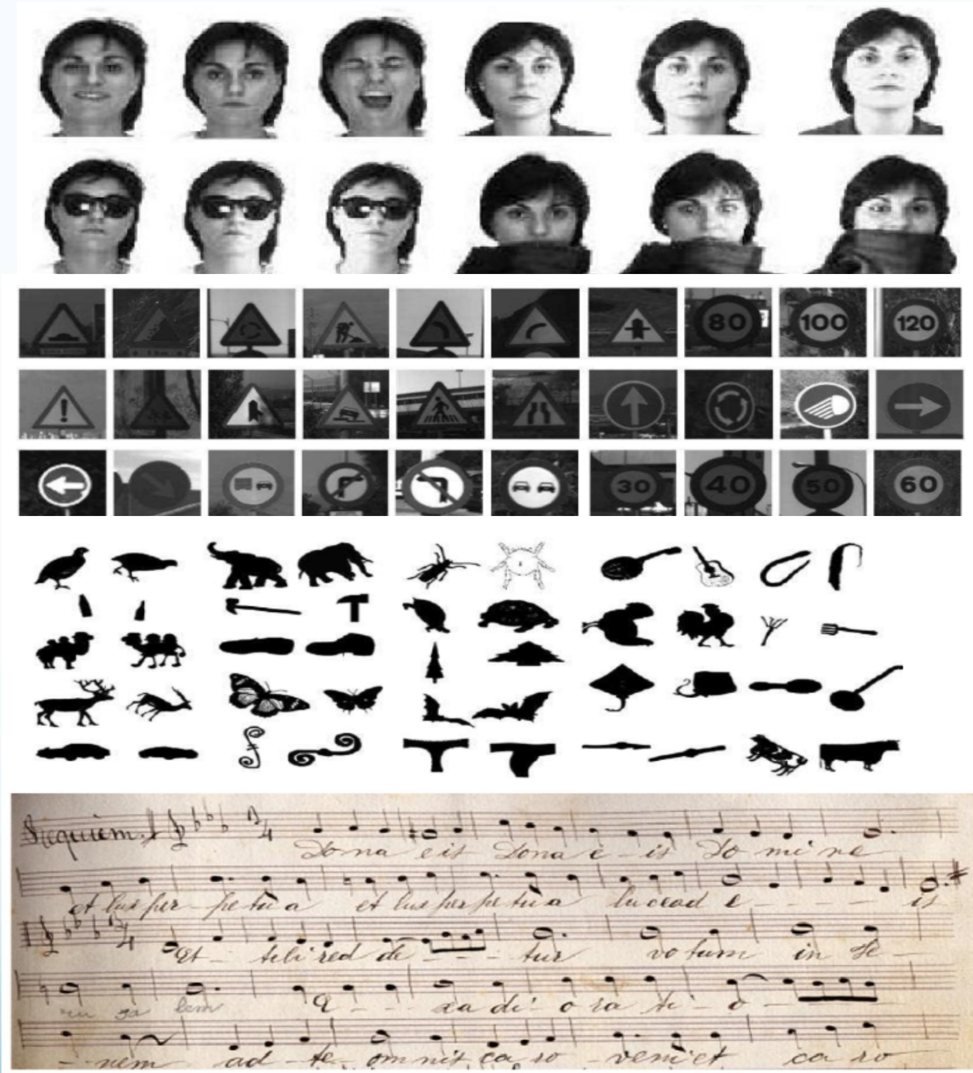
Experiments characteristics

- UCI dataset characteristics.

Problem	#Training samples	#Features	#Classes
Dermathology	366	34	6
Iris	150	4	3
Ecoli	336	8	8
Vehicle	846	18	4
Wine	178	13	3
Segmentation	2310	19	7
Glass	214	9	7
Thyroid	215	5	3
Vowel	990	10	11
Balance	625	4	3
Shuttle	14500	9	7
Yeast	1484	8	10

Experiments Characteristics

- Computer Vision datasets.
- ARFace:520 x 120, 20 classes.
- Traffic: 3481 x 100, 36 classes.
- MPEG: 1400 x 70, 20 classes.
- Cleafs: 4098x65, 7 classes.



Results on UCI problems

- As we can see the evolutive Sublinear performs better than the standard codings.

	ECOC sub. binario		ECOC sub. evol.		ECOC 1vsTodos		ECOC 1vs1	
Base de Datos	Rend.	Clasf.	Rend.	Clasf.	Rend.	Clasf.	Rend.	Clasf.
Derma	96.0±2.9	3	96.3±2.1	3	95.1±3.3	6	94.7±4.3	15
Iris	96.4±6.3	2	98.2±1.9	2	96.9±6.0	3	96.3±3.1	3
Ecoli	80.5±10.9	3	81.4±10.8	3	79.5±12.2	8	79.2±13.8	28
Vehicle	72.5±14.3	2	76.99±12.4	2	74.2±13.4	4	83.6±10.5	6
Wine	95.5±4.3	2	97.2±2.3	2	95.5±4.3	3	97.2±2.4	3
Segment	96.6±2.3	3	96.6±1.5	3	96.1±1.8	7	97.18±1.3	21
Glass	56.7±23.5	3	50.0±29.7	3	53.85±25.8	6	60.5±26.9	15
Thyroid	96.4±5.3	2	93.8±5.1	2	95.6±7.4	3	96.1±5.4	3
Vowel	57.7±29.4	3	81.78±11.1	3	80.7±11.9	8	78.9±14.2	28
Balance	80.9±11.2	2	87.1±9.2	2	89.9±8.4	3	92.8±6.4	3
Shuttle	80.9±29.1	3	83.4±15.9	3	90.6±11.3	7	86.3±18.1	21
Yeast	50.2±18.2	4	54.7±11.8	4	51.1±18.0	10	52.4±20.8	45
Pos. media & # Clasf.	2.9	**2.7**	**2.0**	**2.7**	2.7	5.7	2.2	15.9

Results on Computer Vision problems

- In this experiments we can see how evolutionary approaches outperform standard ECOC codings while decreasing the number of classifiers dramatically.

	ECOC sub. binario		ECOC sub. evol.		ECOC 1vsTodos		ECOC 1vs1	
Base de Datos	Rend.	Clasf	Rend.	Clasf	Rend.	Clasf	Rend.	Clasf
Traffic	90.8±4.1	6	90.6±3.4	6	91.8±4.6	36	90.6±4.1	630
ARFaces	76.0±7.2	5	85.84±5.2	5	84.0±6.3	20	96.0±2.5	190
Clefs	81.2±4.2	3	81.8±9.3	3	80.8±11.2	7	84.2±6.8	21
MPEG7	89.29±5.1	7	90.4±4.5	7	87.8±6.4	6.170	92.8±3.7	2415
Pos, media & # Clasf.	3.0	**5.2**	2.2	**5.2**	3.0	33.2	**1.5**	814.0

Conclusions

- The Sub-linear ECOC represents the lower-bound in terms of number of classifiers.
- The evolutive ECOC optimization obtains comparable results to the standard coding designs (sometimes better) while using far less number of dichotomizers.
- This design is suitable when classifying problems with large number of classes.

Scientific publications associated

- Supervised and Unsupervised ensemble learning and applications 2010, SUEMA-ECML 2010.
- Pattern Recognition Letters Journal (Submitted).
- CVC-RD Workshop, 2010.

Thank you

QUESTIONS?

CCIA 2010

Tretzè Congrés Internacional de
l'Associació Catalana d'Intel·ligència Artificial