Ensemble Learning for Visual Recognition

PhD Defence

Mohammad Ali Bagheri

Supervisor: **Dr. Qigang Gao**
Co-supervisor: **Dr. Sergio Escalera**

DALHOUSIE UNIVERSITY

Huge amount of visual data!
How to deal with this challenge?!!?

OUR GENERAL SOLUTION: ENSEMBLE LEARNING

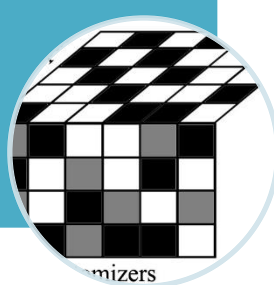
Simple: consult others ! ☺

- **Thesis hypothesise**: possible merits of ensemble learning for visual recognition
  - There is no perfect learner!
  - Better learning in complicated datasets
  - Lowering the risk of an unfortunate selection of a bad learner

Contributions

- Generic subclass Ens.
- GA-SS-ECOC

Novel ensemble classification methods



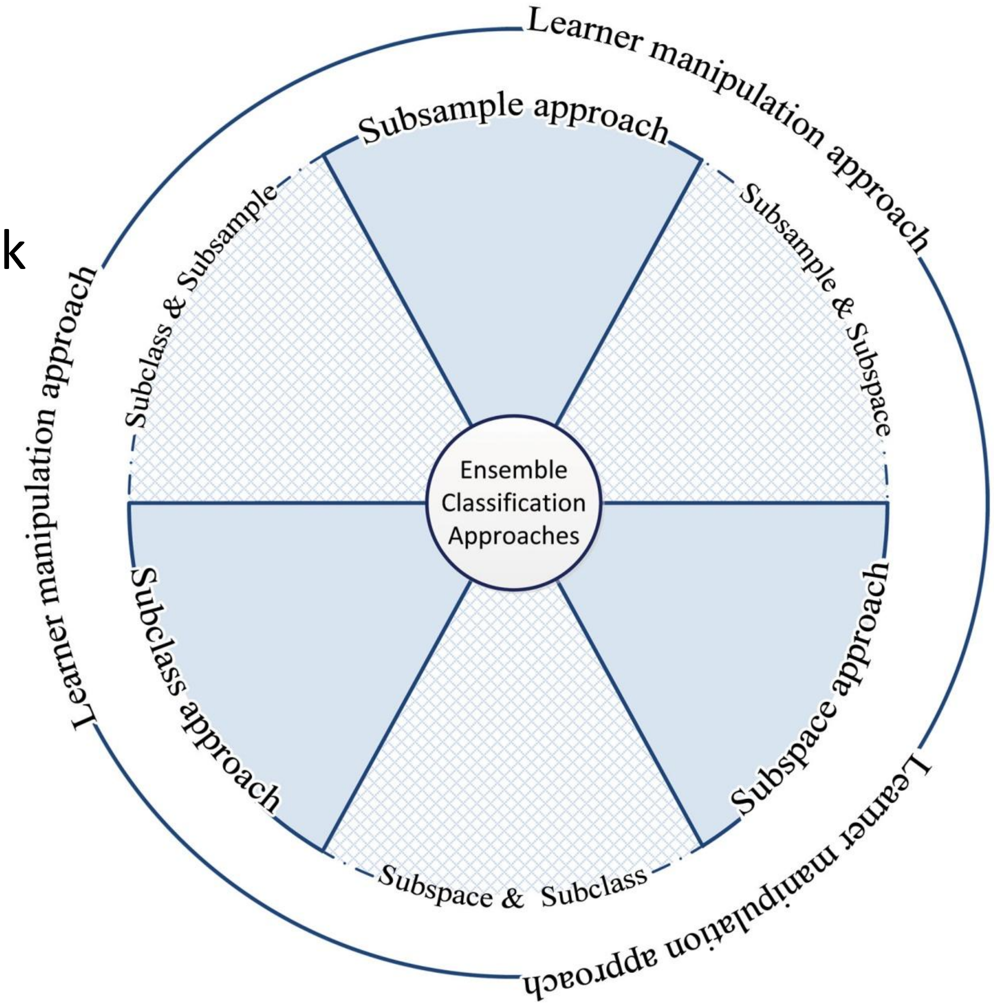
- Shape categorization
  - Application of GA-SS-ECOC
- Action recognition
  - DS-fusion of individual learners
  - An ensemble of ppfSVMs
  - LGSC

Application of ensemble strategy for visual recognition

Ensemble Classification System

- Multiple classifier system
  - Ensemble design
  - Combining classifiers

- A unifying ensemble framework
  1. Subsample
  2. Subspace
  3. Subclass
  4. Learner manipulation

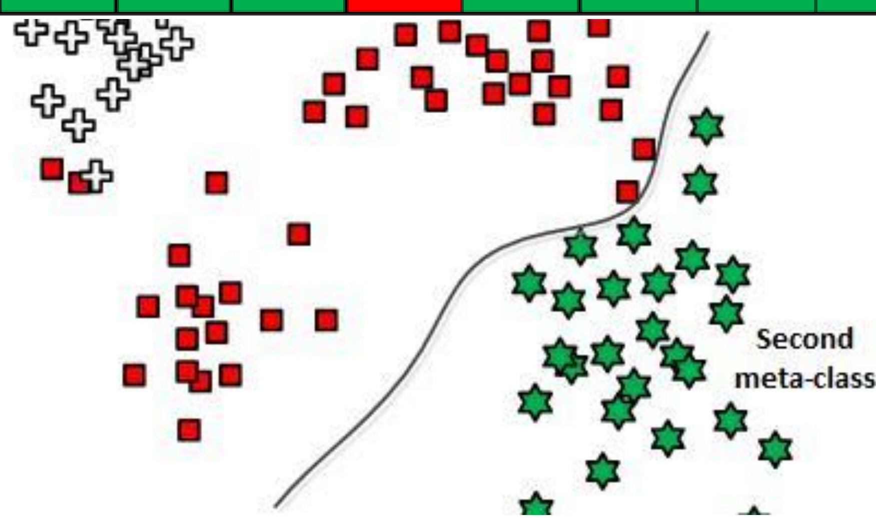


Error Correcting Output Codes

Binary classifiers

	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	
c_1 (✚)	+1	+1	0	+1	0	-1	+1	0	-1	$D_H(c_1) = 5$
c_2 (■)	+1	-1	0	0	-1	0	-1	+1	+1	$D_H(c_2) = 2$
c_3 (★)	-1	-1	-1	-1	+1	-1	0	-1	0	$D_H(c_3) = 4$
c_4 (●)	0	0	1	-1	0	+1	-1	-1	0	$D_H(c_4) = 3$
c_5 (△)	0	+1	0	+1	-1	+1	-1	0	-1	$D_H(c_5) = 3$
Test codeword	-1	-1	+1	+1	+1	+1	-1	+1	+1	

Subspace for h1



Problem Statement

- Goal in designing the ECOC: to improve error correcting capability of the codematrix
  - maximizing a separability criterion between any pair of rows and/or any pair of columns

- Column separation → To have more independent classifiers

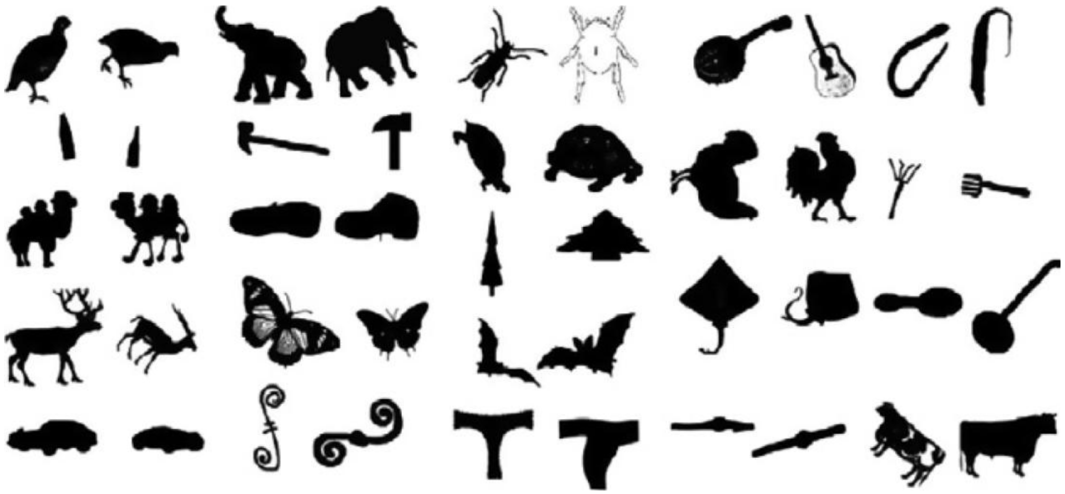
- Conventional strategy: Maximizing Hamming distance

- Problem: **Hamming distance is not an appropriate technique for promoting independency**

OUR SOLUTION/ CONTRIBUTION

- Promoting independency: **SS-ECOC**
  - Employing different features for each classifier
- Enhancing accuracy: **GA-SS-ECOC**
  - Optimization with Genetic Algorithm

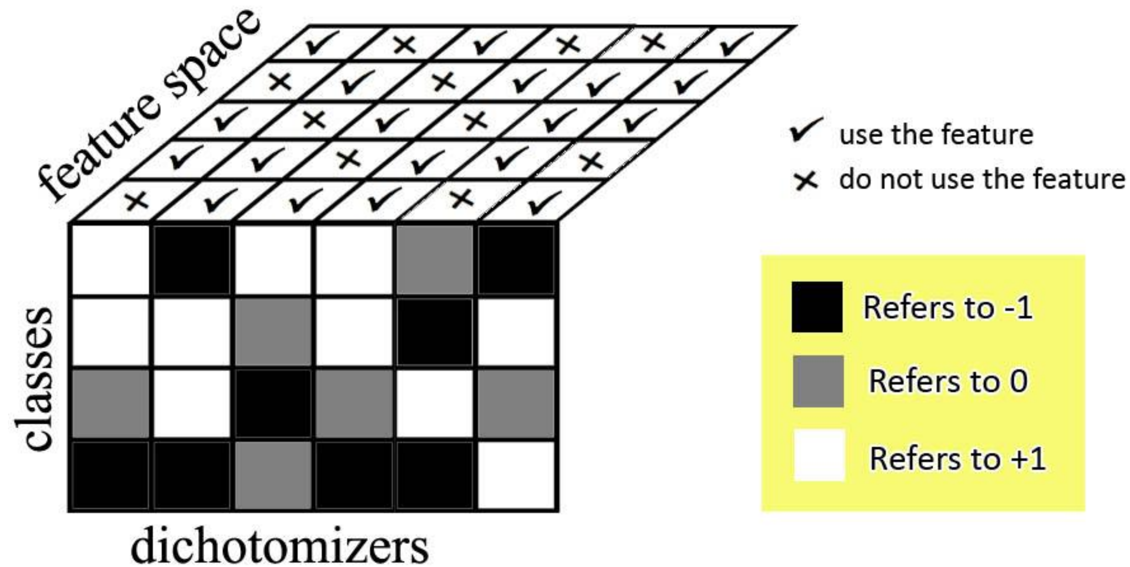
- Applications of the method on two domains
  - Shape categorization
  - Logo recognition



Subspace Error Correcting Output Codes

- Schematic representation

- 3D codematrix

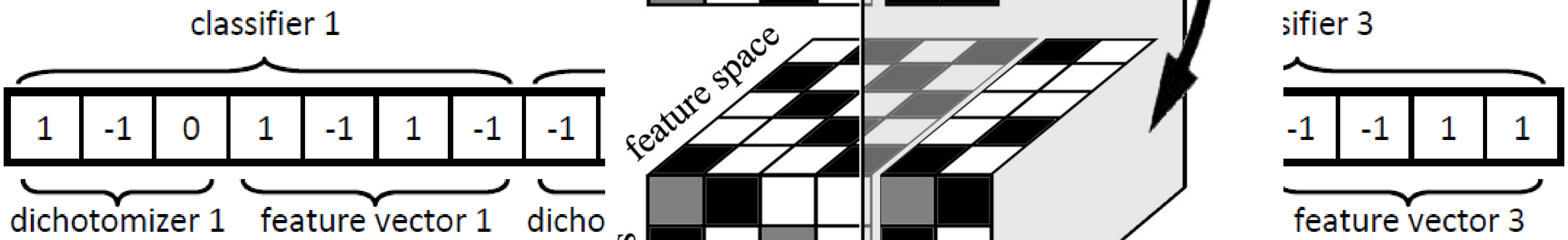


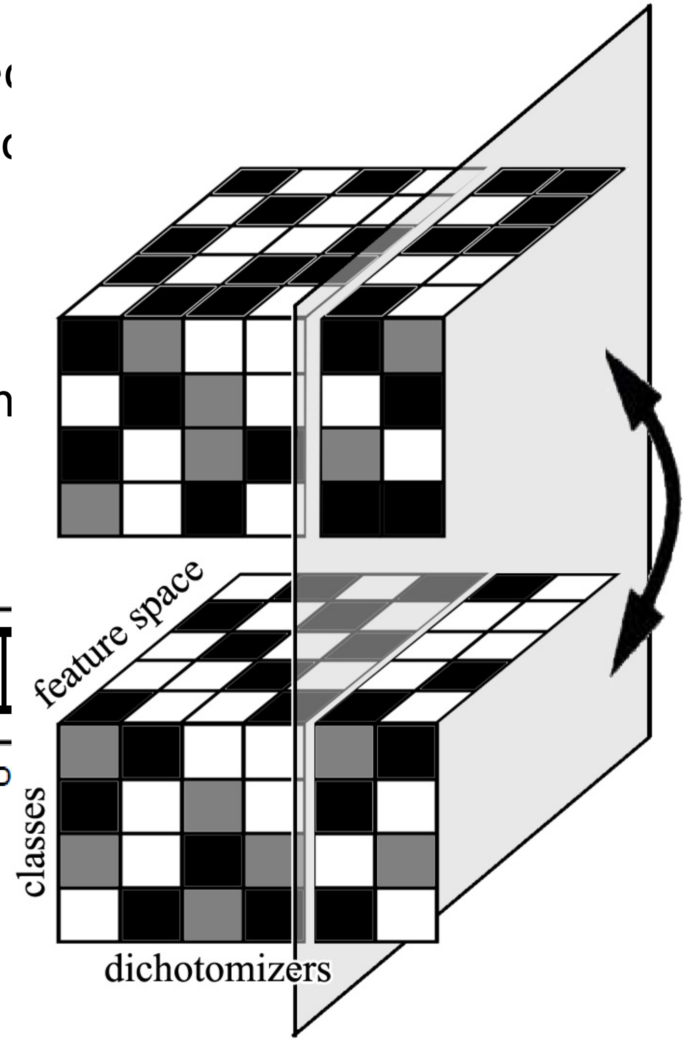
- How to find features for each binary classifier (dichotomizer)?
  - Pseudo-Random *(MLPR Conf. 2012)*
  - Selecting good features for each classifier individually *(ICDM 2012)*
  - **Choosing features of all classifiers simultaneously - > GA-SS-ECOC (PR 2013)**

A Genetic-based Subspace Approach to ECOC

- **Idea:**
  - Optimizing the class de[...] ng together
  - We followed an evolutio

- **Encoding**:
  - Each chromosome of th[...] al codematrix.

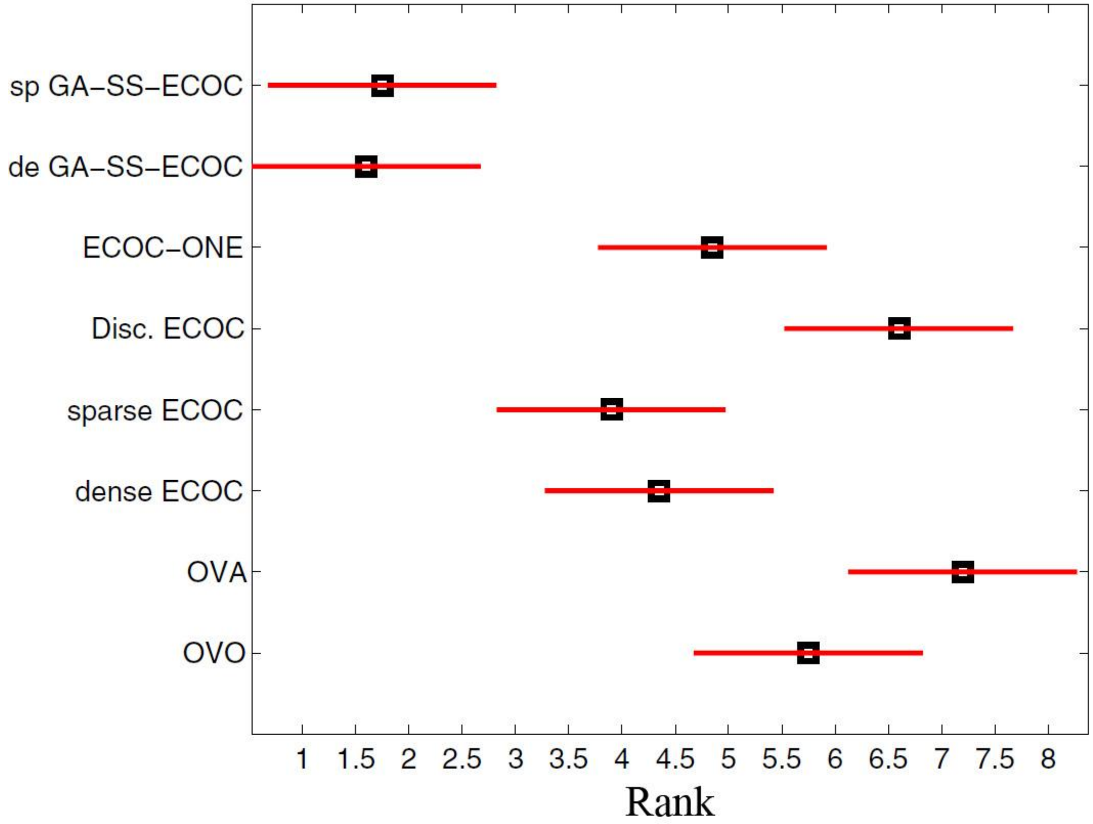




- Cross-over

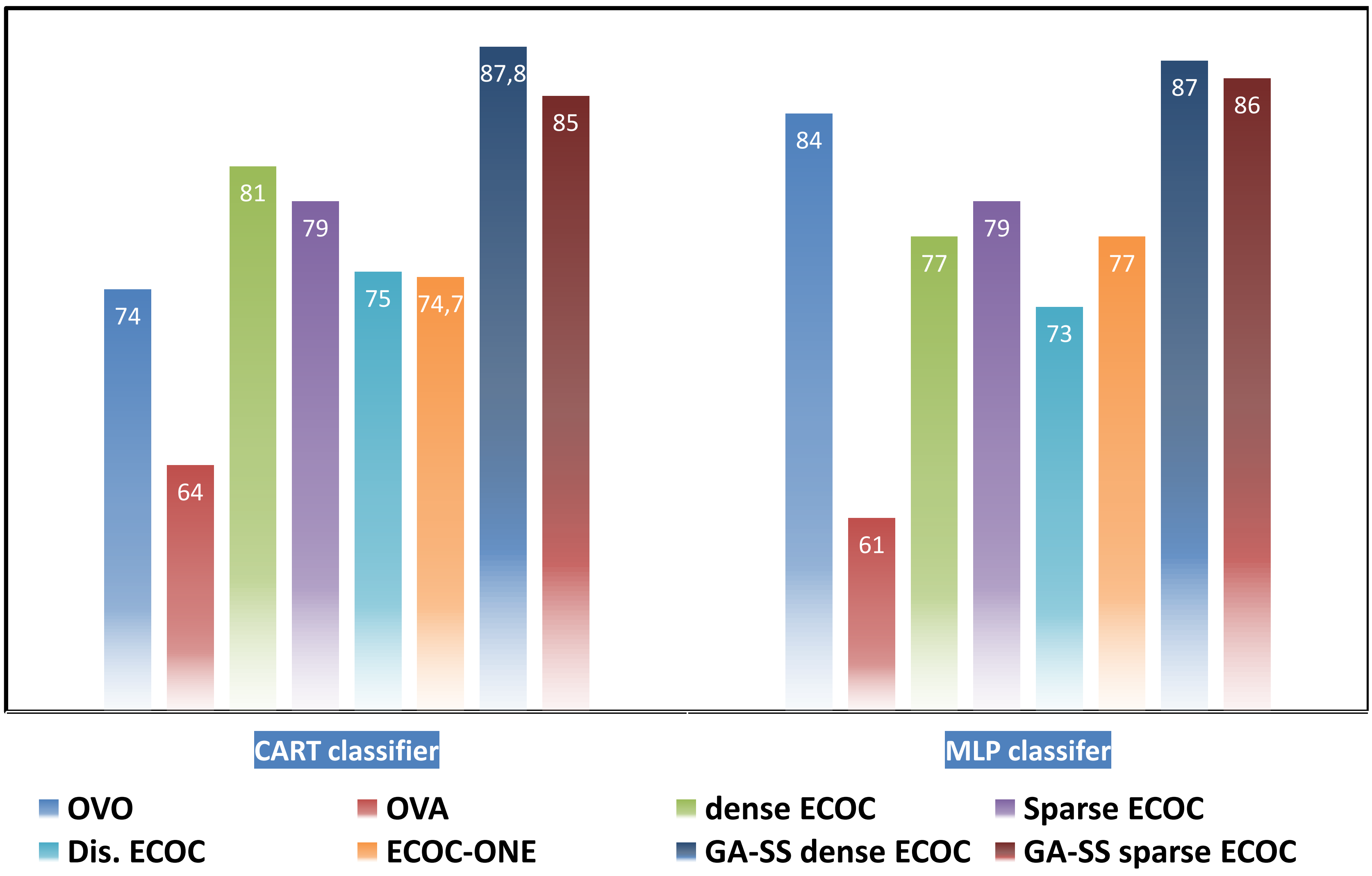
Experimental Results Analysis

- Statistical analysis: Iman-Davenport test - - > Nemenyi



- Disadvantages : more computationally intensive

Application to shape Categorization



Application: Human Action Recognition

- Motivation: the release of Kinect camera and wide range of applications

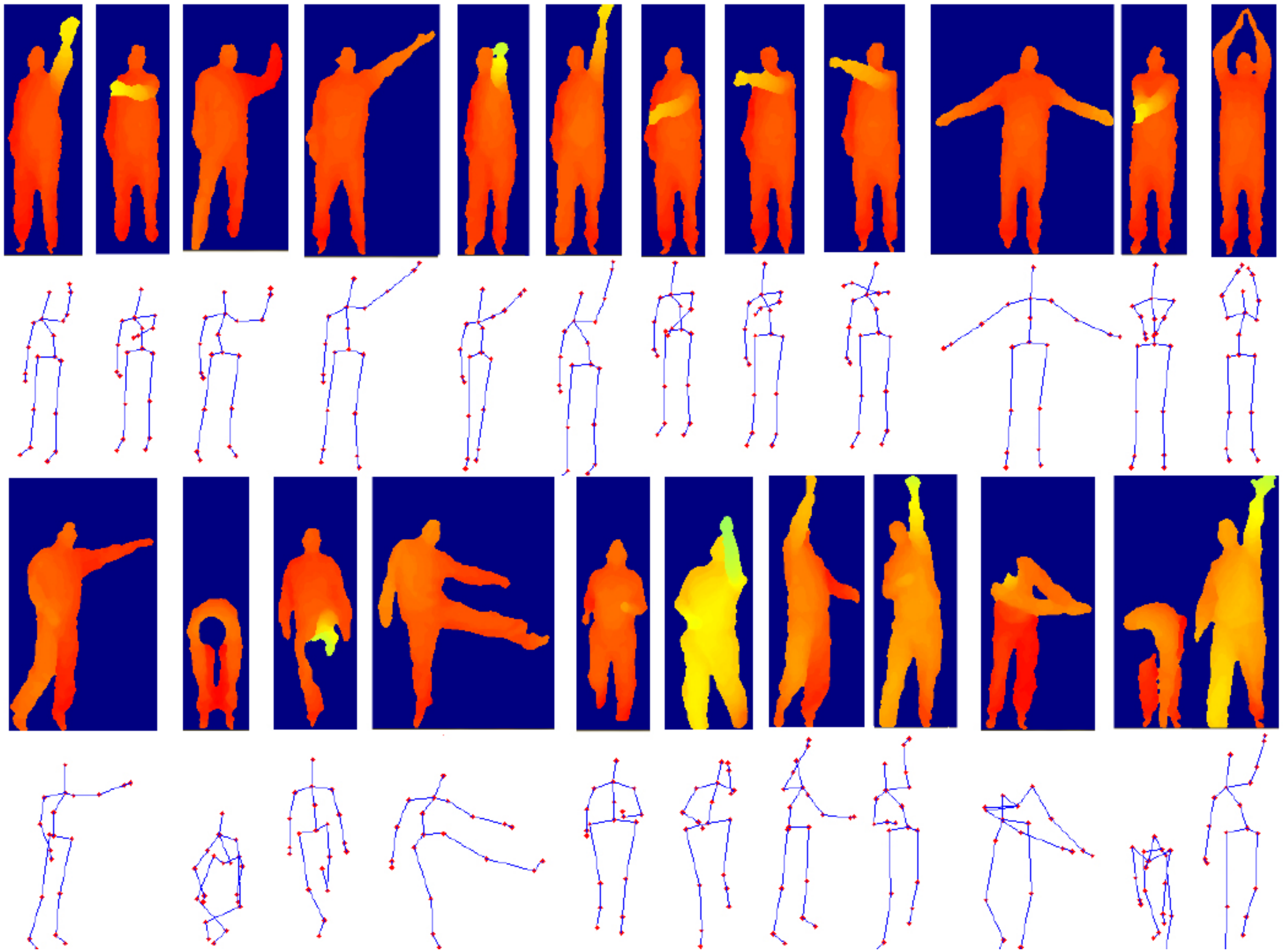
- Three methods:
  1. A Framework of Multi-Classier Fusion for Human Action Recognition
  2. Support Vector Machines with Time Series Distance Kernels for Action Classification
  3. Locality Regularized Group Sparse Coding for Action Recognition (LGSC)

Framework of Multi-Classifier Fusion:
Action Recognition Problems

1) Gesture recognition (Chalearn dataset)

ACTION RECOGNITION PROBLEMS: MSR-ACTION3D

2) Human actions classification

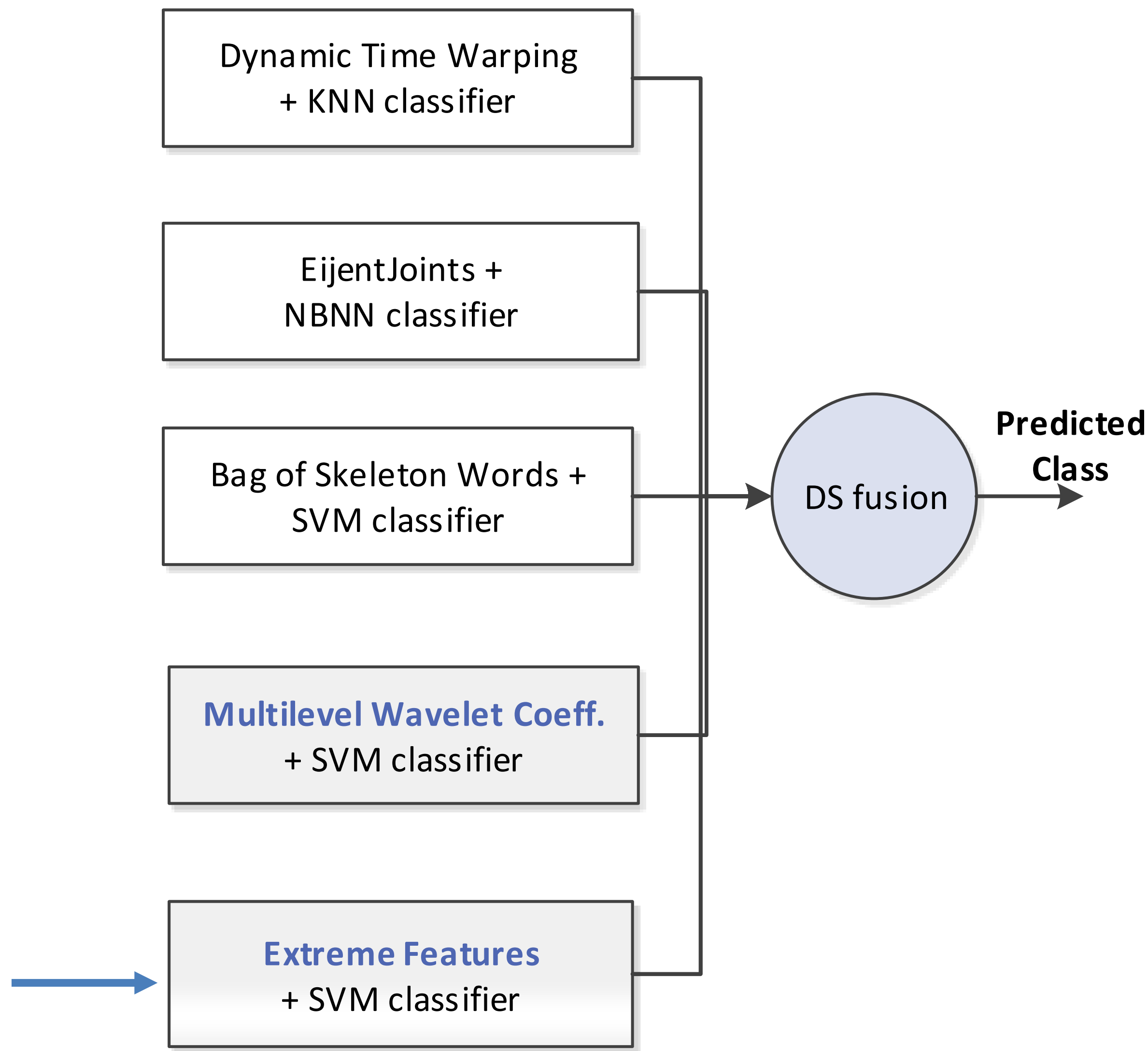


-20 classes

-567 samples

Framework of Multi-Classifier Fusion

- An ensemble of five different learning techniques



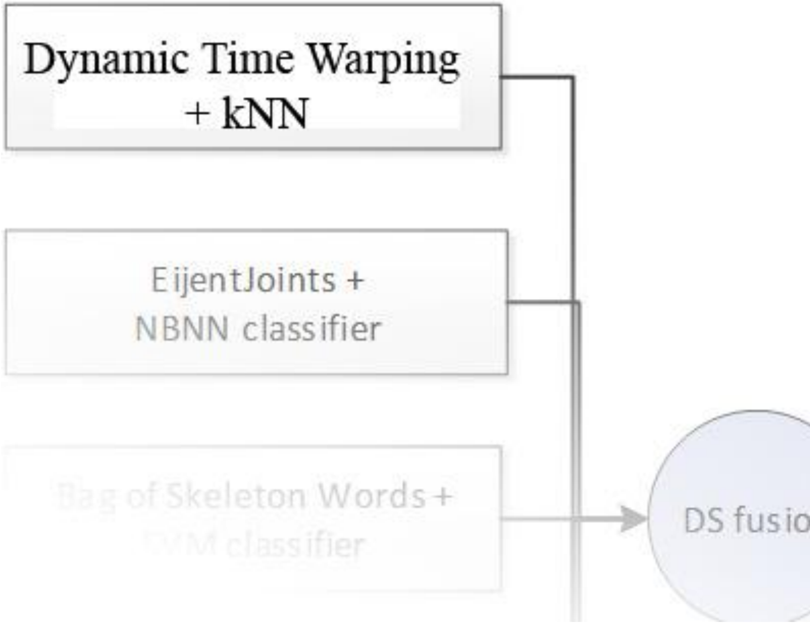
EXTREME FEATURES

- **For many short actions, only a very few salient postures can be a unique representative of the action**
- **The relative position, i.e. the distance, between a set of skeleton joints will reach its extreme value**

Dempster-Shafer Fusion of Individual Learners

- Result of each learning technique **+ Efficient Combination**

Dataset	Single classifier trained only on					DS fusion
	EigenJoints +NBNN	DTW +KNN	BoVW +SVM	Wavelet coefficients +SVM	Extreme features +SVM	
Chalearn	54.30	77.85	73.05	70.40	73.65	**82.60**
MSR-Action3D	47.81	75.76	64.65	58.92	72.39	**80.81**

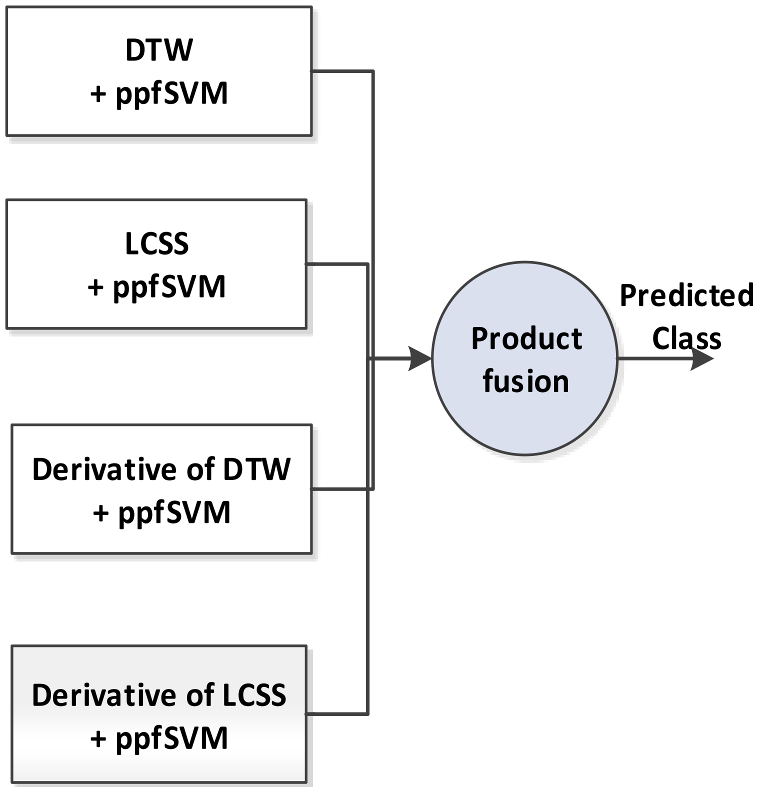


- 3D trajectories of body joints as multi-dimensional time series
  - Traditional recognition technique: Time series distance measure > kNN
- Superiority of SVM

Problem: non-PSD kernels !!

SOLUTION

- Pairwise proximity function SVM (ppfSVM)
  - Each sample is represented with its dissimilarity to all other samples
- An ensemble of four ppfSVMs
  - Two time series distance measure + their derivative, as SVM kernel
  - Classifier fusion



- Results of the ppfSVM and KNN on the MSR-Action3D dataset

	DTW	DDTW	LCSS	DLCSS	Product fusion
Kernel SVM	80.47	83.84	75.76	76.77	**90.57**
kNN	75.42	77.78	72.05	65.66	–

- Results on the CAD-60 dataset

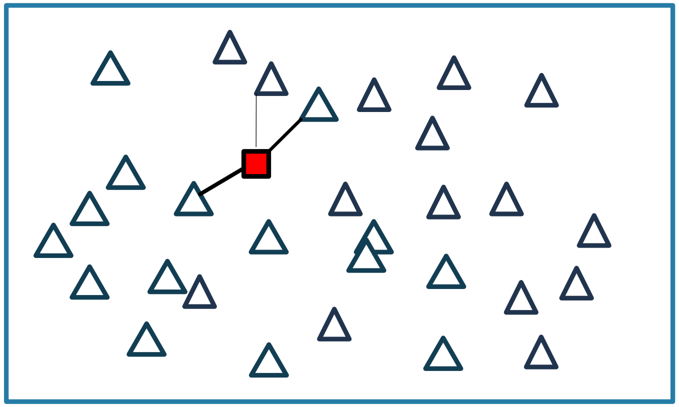
	DTW	DDTW	LCSS	DLCSS	Product fusion
Kernel SVM	73.33	75.00	71.67	70.00	**76.67**
kNN	68.33	68.33	65.00	66.67	–

LOCALITY REGULARIZED GROUP SPARSE CODING

- **Bag of visual words (BoVW)**

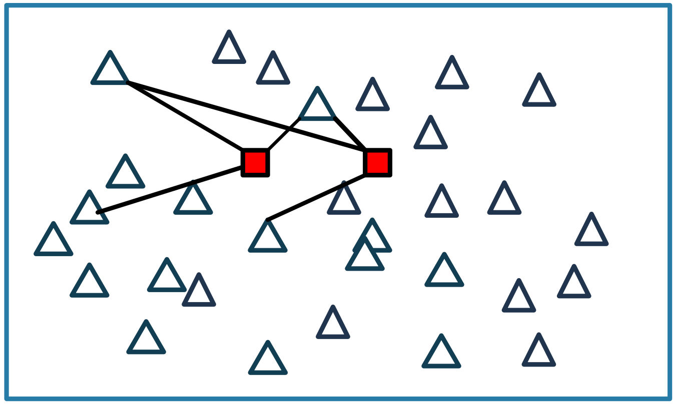
- Core component : **Feature encoding**

- *Locality-constrained Linear Coding*
  - Problems:
    1. Setting a pre-defined number of coeff.
    2. Encoding descriptors of a sample individually.



LLC

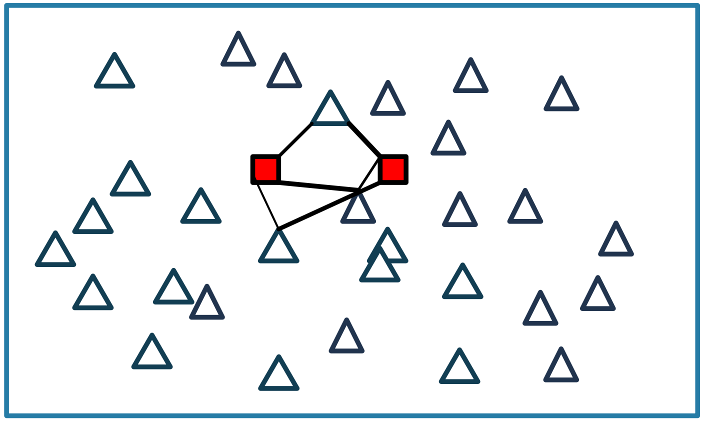
- *Group Sparse Coding*
  - Problems:
    1. Did not consider the locality!



Group sparse coding

SOLUTION: LOCALITY REGULARIZED GROUP SPARSE CODING

- ***Locality Regularized Group Sparse Coding (LGSC)***
  - Utilizes the advantages of locality coding and group coding



- Formulation

$$\mathcal{Q}_c(\mathbf{C}, \mathbf{X}, \mathbf{D}) = \sum_g \mathcal{Q}_c(\mathbf{C}^g; \mathbf{X}^g, \mathbf{D}^g)$$

$$= \sum_g \left(\frac{1}{2} \|\mathbf{X}^g - \mathbf{B}\mathbf{C}^g\|_F^2 + \lambda_1 \mathrm{tr}(\mathbf{D}^{g\top} \mathbf{C}^g) + \lambda_2 \sum_{i=1}^{M} I(\|\mathbf{C}_i^g\|) \right)$$

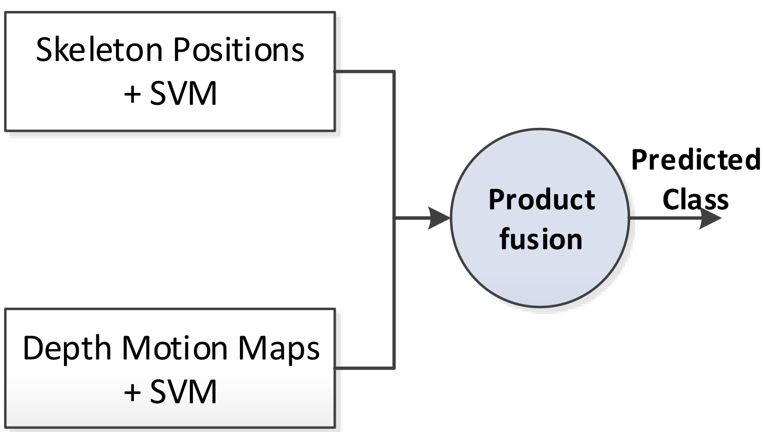
$$s.t. \quad \mathbf{C}^g \geq 0,$$

SOLUTION: LOCALITY REGULARIZED GROUP SPARSE CODING

- Implementation

  Alternating Direction Method of Multipliers (ADMM)

- Classifier fusion



Algorithm 1 : LGSC implementation via ADMM

Initialization: Set $\rho = 10^{-1}$; $\epsilon = 10^{-5}$, maxIter = 10^3.
Initialize $k = 0$; $\mathbf{C}^{(0)} = I$; $\mathbf{Z}^{(0)} = I$; $\Lambda^{(0)} = 0$, error1= 2ϵ, error2= 2ϵ, $\lambda_1 = 0.1$ and $\lambda_2 = 0.15$.

while (error1 $\geq \epsilon$ and error2 $\geq \epsilon$) and ($k <$ maxIter) **do**

Update $\mathbf{C}$ by

$$\mathbf{C}^{k+1} = \arg\min_{\mathbf{C}} \mathcal{L}_\rho(\mathbf{C}, \mathbf{Z}^k, \Lambda^k) \quad (10)$$
$$= \arg\min_{\mathbf{C}} \Big(\frac{1}{2}\|\mathbf{X} - \mathbf{B}\mathbf{C}\|_F^2 + \lambda_1 \mathrm{tr}(\mathbf{D}^T\mathbf{C}) + \frac{\rho}{2}\|\mathbf{C} - \mathbf{Z}^k + \Lambda^k\|_F^2 \Big)$$

Update $\mathbf{Z}$ by

$$\mathbf{Z}^{k+1} = \arg\min_{\mathbf{Z}} \mathcal{L}_\rho(\mathbf{C}^{k+1}, \mathbf{Z}, \Lambda^k) \quad (11)$$
$$= \arg\min_{\mathbf{Z}} (\lambda_2 \|\mathbf{Z}\|_{1,2} + \frac{\rho}{2}\|\mathbf{C}^{k+1} - \mathbf{Z} + \Lambda^k\|_F^2)$$
$$s.t. \quad \mathbf{Z} \geq 0$$

Update the Lagrange multiplier matrix by

$$\Lambda^{k+1} = \Lambda^k + \rho(\mathbf{C}^{k+1} - \mathbf{Z}^{k+1}) \quad (12)$$

Calculate the objective function, Obj^k using Eq. 7

Calculate errors by

$$\mathrm{error1} = \mathrm{Obj}^k - \mathrm{Obj}^{k-1}, k = 2, \ldots$$
$$\mathrm{error2} = \|\mathbf{C}^{k+1} - \mathbf{Z}^{k+1}\|_\infty$$

$k \leftarrow k + 1$

end while

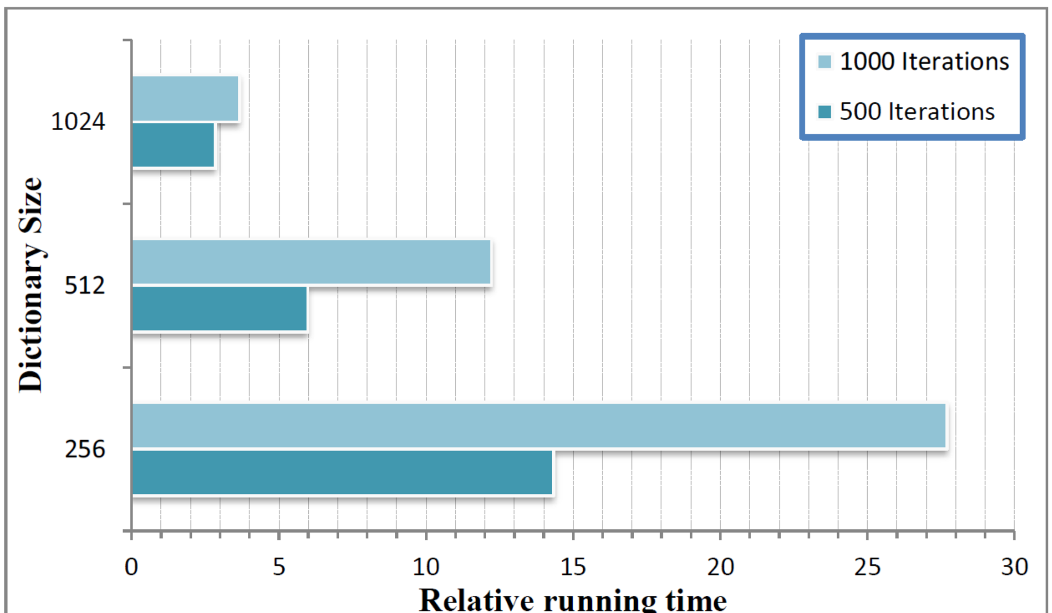
Output: Optimal solution $\mathbf{C}^* = \mathbf{C}^k$

- Classification accuracy
  dataset.

	VQ		LGSC
MSRAction3D	81.2		.45
CAD-60	81.6		.30
Chalearn	72.5		.64

MSRAction3D	
	Accuracy
Studies employed depth data	
Action Graph [89]	74.70
HON4D [108]	85.85
Vieira et al. [150]	78.20
Random Occupancy Patterns [151]	86.50
DMM-HOG [164]	85.52
HOPC [117]	91.64
DMM-LBP-FF [25]	87.90
Studies employed only skeleton data	
	88.20
	78.97
	80.20
	89.48
	93.45

	Accuracy
	86.50
	65.30
a	
	74.70
	51.30
Proposed LGSC Algorithm	**88.33**



Conclusion

1. **A unifying framework** for multiple classifier systems
2. **Generic Subclass Ensemble**: a general approach to ensemble
3. **GA-SS-ECOC:** an evolutionary algorithm-based approach to the design of an application-dependent codematrix in ECOC
4. **Ensemble framework** for action recognition by Dempster-Shafer fusion
5. **ppfSVM**: a new class of SVM that is applicable to trajectory classification, such as action recognition,
6. **LGSC**:

Future Work

- Investigating deep features
- Exploiting the RGB-D data
- Continuous action recognition

PUBLICATIONS (JOURNALS)

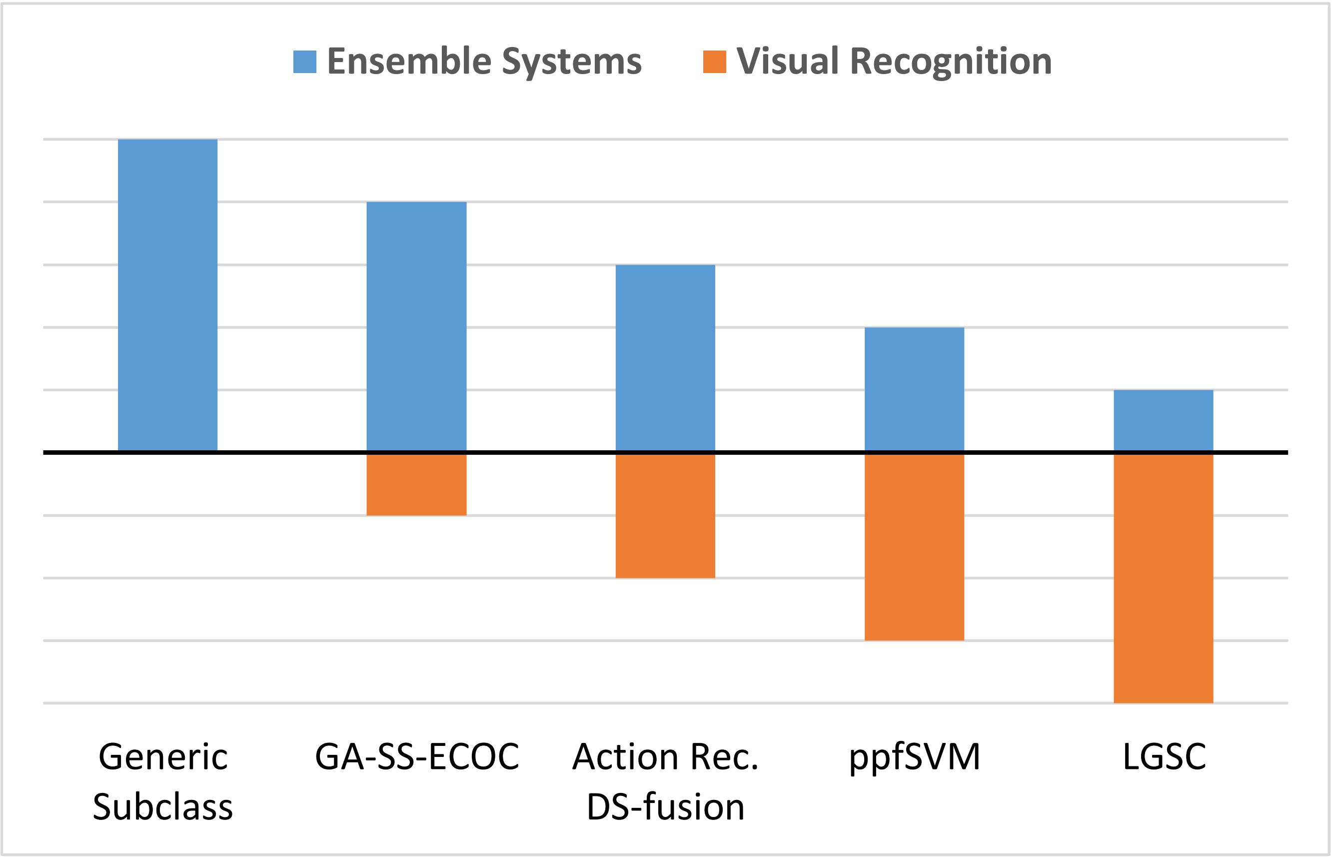
- **M.A. Bagheri**, Q. Gao, and S. Escalera, “A Genetic-based subspace analysis method for improving Error-Correcting Output Coding”, ***Pattern Recognition***, vol. 46, pp.2830–2839, 2013.
- **M.A. Bagheri**, Q. Gao, and S. Escalera, “Combining local and global learners in the pairwise multiclass classification", ***Pattern Analysis and Applications***, pp. 1-16, 2014.
- **M.A. Bagheri**, Q. Gao, and S. Escalera, “Locality Regularized Group Sparse Coding “, ***Computer Vision and Image Understanding.*** **(Under Review)**

Publications (Conferences)

1. **M.A. Bagheri**, Q. Gao, and S. Escalera, "Support Vector with Time Series Distance Kernels for Action Classification", in *Proc. IEEE Winter Conference on Applications of Computer Vision (**WACV**)* , NY, United Sates, 2016.
2. **M.A. Bagheri**, Q. Gao, and S. Escalera, "Action Recognition by Pairwise Proximity Function Support Vector Machines with Dynamic Time Warping Kernels", in *Proc. 29th Canadian conf. on Artificial Intelligence*, BC, Canada, 2016.
3. **M.A. Bagheri**, Q. Gao, and S. Escalera, "Generic Subclass Ensemble: A Novel Approach to Ensemble Classification", in *Proc. 22nd International conf. on Pattern Recognition (**ICPR**)*, Stockholm, Sweden, 2014.
4. **M.A. Bagheri**, G. Hu, Q. Gao, and S. Escalera, "A Framework of Multi-Classifier Fusion for Human Action Recognition", in *Proc. 22nd International conf. on Pattern Recognition (**ICPR**)*, Stockholm, Sweden, 2014.
5. **M.A. Bagheri**, Q. Gao, and S. Escalera, "Logo recognition Based on the Dempster-Shafer Fusion of Multiple Classifiers", in *Proc. 26th Canadian conf. on Artificial Intelligence*, Regina, Canada, 2013. **(Best paper award)**
6. **M.A. Bagheri**, Q. Gao and S. Escalera, "A framework towards the unification of ensemble classification methods.", in *Proc. IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2013.
7. **M.A. Bagheri**, Q. Gao and S. Escalera, "Rough Set Subspace Error Correcting Output Codes", in *Proc. IEEE International Conf on Data Mining (**ICDM**)*, Brussels, Belgium, 2012.
8. **M.A. Bagheri**, Q. Gao, and S. Escalera, "Three-Dimensional Design of Error Correcting Output Codes", in *Proc. International Conference on Machine Learning and Data Mining*, Berlin, Germany, 2012
9. **M.A. Bagheri**, Q. Gao, and S. Escalera, "Efficient pairwise classification using Local Cross Off strategy", in *Proc. 25th Canadian conf. on Artificial Intelligence*, Toronto, Canada, 2012.

Thank You for Your Attention!

Novelty of all proposed methods



- “The important thing with deep learning and machine learning in general is it needs a lot of data to train on, so a computer learns to do a task like recognizing an object in an image or identifying that there is a cancer cell or recognizing which word you’re saying when you’re speaking by looking at millions of examples, and one reason why neural nets didn’t catch on earlier is that we didn’t have that much data in the 90s.”