

# ASAP : revue de mi-parcours

| Partenaires      | Responsables     |
|------------------|------------------|
| LITIS, U. Rouen  | A. Rakotomamonjy |
| LIP6, UPMC       | M. Cord          |
| INRIA, TAO       | H. Paugam-Moisy  |
| LIF, UP          | L. Ralaivola     |
| Greyc, ENSI Caen | C. Rosenberger   |

<http://sites.google.com/site/anrasap/>

12 septembre 2011

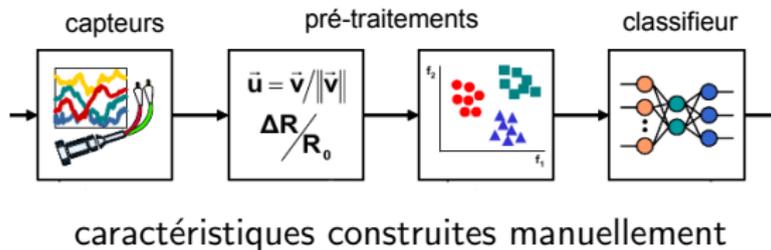
# Plan

1. Mise en contexte et objectifs scientifiques du projet
2. Tâche 1 : Apprentissage de caractéristiques à partir d'une structure simple
3. Tâche 2 : Structure profonde pour l'apprentissage de caractéristiques
  - ▶ Structure d'inspiration biologique
4. Tâche 3 et 4 : Implémentation et Démonstrateurs
  - ▶ Compétition Unsupervised and Transfer Learning Challenge
5. Résultats marquants et impacts
6. Perspectives et conclusions

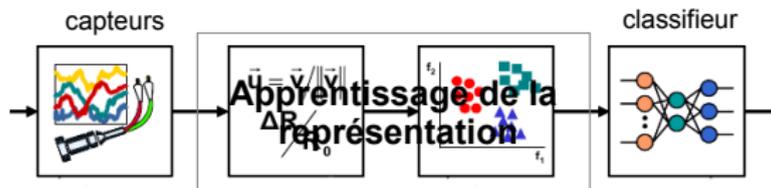
<http://sites.google.com/site/anrasap/>

# ASAP

## Diagramme classique d'un système de reconnaissance de formes



## Rupture envisagée : Apprentissage de la représentation



- ▶ extraction automatique des caractéristiques
- ▶ système *end-to-end*

# ASAP

## Objectif

Etudier et développer des **stratégies d'extraction automatique** de caractéristiques pertinentes pour la reconnaissance de formes à partir

- ▶ structure simple
- ▶ structure hiérarchique

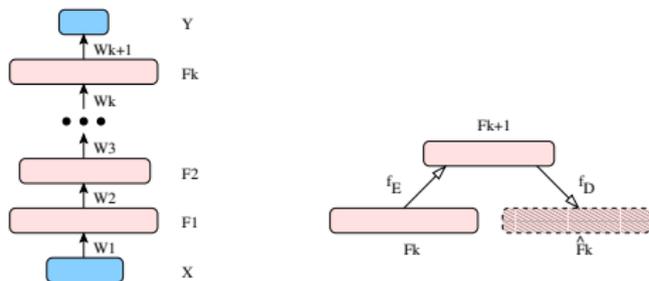
## Principe

On cherche à **transformer** des données brutes (signal ou image) en de **nouvelles représentations**, permettant de faciliter des tâches de classification

- ▶ Superposition de couches
- ▶ Chaque couche apprend une représentation des données qui lui sont présentées
- ▶ Apprentissage **supervisée** ou **non-supervisée**

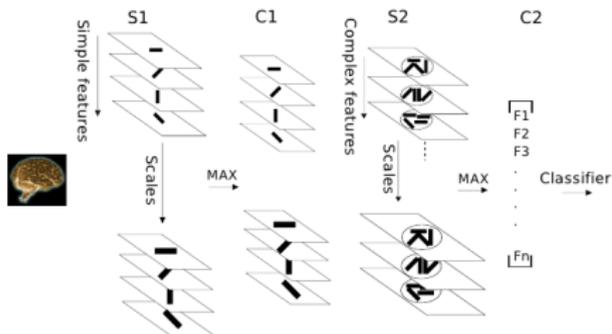
# ASAP : exemples

## Auto-encodeur



- ▶  $X$  données originales
- ▶  $f_E$  et  $f_D$  fonctions de représentations à apprendre tel que  $\hat{F}_k \approx F_k$
- ▶  $Y$  nouvelle représentation des données

## Gabor multi-échelle



- ▶ extraction de caractéristiques de Gabor
- ▶ Organisation des caractéristiques en représentation plus abstraites

# Problématiques scientifiques abordées dans ASAP

- ▶ Apprentissage de structures simples :
  - ▶ quel **modèle** et comment l'**apprendre** ?
  - ▶ propriétés d'**invariance** et **robustesse**
  - ▶ étant donné un ensemble de caractéristiques représentatifs, comment **coder** un nouvel exemple ?
- ▶ Apprentissage de structures profondes
  - ▶ quelles **stratégies** d'empilement de structures simples et quelles propriétés ?
  - ▶ quelles **alternatives** ?
    - ▶ biologiquement inspirée
    - ▶ apprentissage profond et noyaux
    - ▶ représentations profondes par décomposition en ondelettes
- ▶ Mise en Oeuvre et Applications
  - ▶ Image et indexation
  - ▶ Biométrie
  - ▶ Interfaces cerveau-machine

# Plan

1. Mise en contexte et objectifs scientifiques du projet
2. Tâche 1 : Apprentissage de caractéristiques sur une couche
3. Tâche 2 : Structure profonde pour l'apprentissage de caractéristiques
  - ▶ Structure d'inspiration biologique
4. Tâche 3 et 4 : Implémentation et Démonstrateurs
  - ▶ Compétition Unsupervised and Transfer Learning Challenge
5. Résultats marquants et impacts
6. Perspectives et conclusions

# Tâche 1 : apprentissage d'une couche élémentaire

## Etat des lieux

- ▶ Introduction et problématique

## Livrables

- ▶ Livrable 1.1 : Etat de l'art
- ▶ Livrable 1.2 : Approche code et dictionnaire
- ▶ Livrable 1.3 : Optimisation

## Avancement

soldé  
soldé  
80%

<http://sites.google.com/site/anrasap/>

# Apprentissage de la représentation

## Objectif

Transformer des données brutes en une nouvelle représentation à l'aide d'une fonction  $f_{\theta}(\cdot)$ ;

## Cadre

Etant donnée un ensemble de données  $\{\mathbf{x}_i\} \in \mathbb{R}^d$ , on construit une fonction  $f_{\theta}(x) : \mathbb{R}^d \mapsto \mathbb{R}^K$

- ▶  $f_{\theta}(\cdot)$  **génère** un vecteur de caractéristiques
- ▶  $f_{\theta}(\cdot)$  est une fonction apprise à partir des données

# Exemple : Image

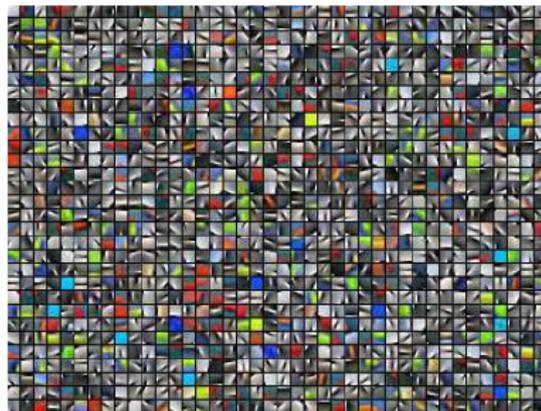
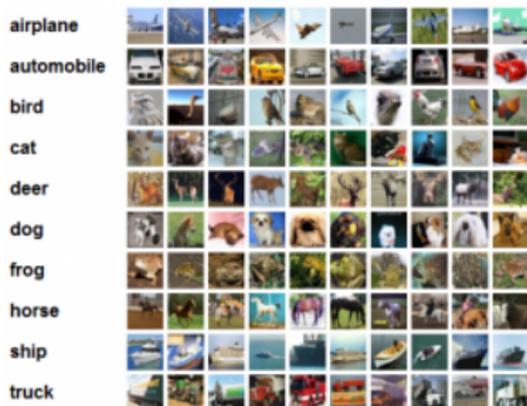
- ▶ Dictionnaire  $\mathbf{D}^*$  appris à partir de **patches** d'images d'apprentissage

$$\mathbf{D}^* = \arg \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{DA}\|_F^2 + \Omega_{\mathbf{A}}(\mathbf{A})$$

- ▶ un nouveau patch  $\mathbf{x}$  est représenté par son code  $a^*$  issu d'un **codage parcimonieux** ou **hard/soft assignment**

$$a^* = \arg \min_a \|\mathbf{x} - \mathbf{D}^* a\|^2 + \Omega(a)$$

- ▶ une nouvelle image est représentée par “l'importance” de chaque atome du dictionnaire dans les codes de tout les patches composant l'image



# Livrable 1.1 : Etat de l'art

## Revue

- ▶ Codage parcimonieux
- ▶ Apprentissage de dictionnaires
- ▶ Auto-encodeur
- ▶ Factorisation de matrices

## Principe

- ▶ Apprentissage de dictionnaire

$$\min_{\mathbf{D}, \mathbf{A}} \sum_i \|\mathbf{x}_i - \mathbf{D}\mathbf{A}_{\cdot,i}\|_2^2 + \Omega(\mathbf{A})$$

$\mathbf{D}$  : le dictionnaire de représentation,  $\mathbf{A}$  : les codes associées à chaque exemple,  $\Omega$  : regularisation

- ▶ Auto-encodeur

$$\min_{h,s} \sum_i \|\mathbf{x}_i - h(s(\mathbf{x}_i))\|_2^2$$

$s$  : fonction d'encodage,  $h$  : fonction de décodage

# Livrable 1.2 : Code et dictionnaire

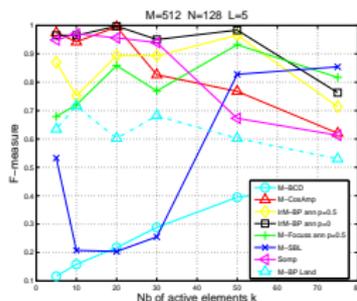
## Code

- Analyse et comparaison des algorithmes permettant de résoudre efficacement [LITIS]

$$\min_{\mathbf{A}} \|\mathbf{X} - \mathbf{DA}\|_F^2 + \Omega(\mathbf{A})$$

où  $\Omega(\mathbf{A})$  regularisation non-differentiable induisant une parcimonie

- méthode gloutonne, itérative, descente par blocs



- version “à noyaux” [LIF]

# Livrable 1.2 : Code et dictionnaire

## Dictionnaire

- ▶ données bruitées : dictionnaire robustes
- ▶ apprentissage de dictionnaires sous contrainte de régularité, de parcimonie et de norme

$$\min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{DA}\|_F^2 + \Omega(\mathbf{A}) + \Omega_D(\mathbf{D})$$

avec

$$\Omega_D(\mathbf{D}) = \sum_j \sum_k |\mathbf{D}_{k+1,j} - \mathbf{D}_{k,j}| + \lambda_1 \|\mathbf{D}_{\cdot,j}\|_1 + \lambda_2 \|\mathbf{D}_{\cdot,j}\|_2$$

- ▶ résolution par méthode des multiplicateurs de Lagrange et opérateur proximal [LITIS]
- ▶ à évaluer sur des tâches de reconnaissances

# Livrable 1.3 : Optimisation

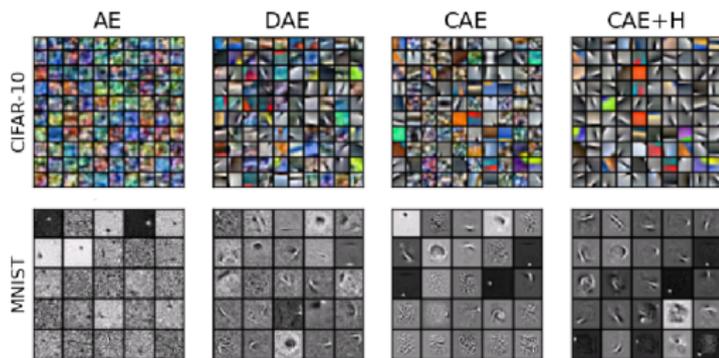
## Amélioration des modèles existants : auto-encodeur robuste aux variations [LITIS + LISA]

- ▶ Le problème d'apprentissage

$$\min_{s,h} \|\mathbf{X} - h(s(\mathbf{X}))\|_F^2 + \Omega(\mathbf{X})$$

Typiquement  $\Omega(\mathbf{X}) = \|J_s(x)\|^2 = \sum_{i,j} \left| \frac{\partial s_j(x)}{\partial x_i} \right|^2$  (norme de la jacobienne de l'encodeur) ou  $\Omega(\mathbf{X}) = \|J_s(x)\|^2 + \|H_s(x)\|^2$

- ▶ Exemple de résultats



# Livrable 1.3 : Optimisation

## Amélioration des modèles existants : Auto-encodeur parcimonieux [INRIA TAO]

### ► Problème

$$\min_{s,h} \|\mathbf{X} - h(s(\mathbf{X}))\|_F^2 + \Omega_p(\mathbf{X})$$

où  $\Omega_p(\mathbf{X})$  impose une contrainte de parcimonie sur  $s(\mathbf{X})$

### ► Résolution

- Apprentissage de l'auto-encodeur
- Sparsification du code par seuillage
- Apprentissage de l'encodeur sur le code épars

## Avantages

- implémentation simple et apprentissage en ligne
- résultats encourageants comparés à un auto-encodeur simple

# Perspectives

- ▶ Evaluation des structures simples sur des tâches de reconnaissance
- ▶ Apprentissage de dictionnaires robustes
- ▶ Utilisation en structure profonde et application à la reconnaissance de formes

# Plan

1. Mise en contexte et objectifs scientifiques du projet
2. Tâche 1 : Apprentissage de caractéristiques sur une couche
3. Tâche 2 : Structure profonde pour l'apprentissage de caractéristiques
  - ▶ Structure d'inspiration biologique
4. Tâche 3 et 4 : Implémentation et Démonstrateurs
  - ▶ Compétition Unsupervised and Transfer Learning Challenge
5. Résultats marquants et impacts
6. Perspectives et conclusions

# Tâche 2

## Livrables

- ▶ Livable 2.1 : Etat de l'art soldé
- ▶ Livable 2.2 : Analyse de la structure en cours
- ▶ Livable 2.3 : Apprentissage et optimisation en cours
- ▶ Livable 2.4 : Multi-tâches débuté

# Livrable 2.1 : Etat de l'art

## Revue sur

- ▶ les différents modèles de “réseaux profonds”
- ▶ apprentissage des réseaux profonds
- ▶ l'architecture des réseaux de neurones multi-couches

## Dissémination

- ▶ Organisation d'une session spéciale à ESANN 2011
- ▶ présentation d'un tutorial

# Livrable 2.2 : Analyse de la structure

## Travaux en cours

- ▶ Livrable prévue à  $T_0+18$  mais tâche en cours à  $T_0+33$
- ▶ Propriétés statistiques théoriques de certaines structures simples (RBM) :  
Approche PAC [LIF]
- ▶ Point de vue “compressed sensing” et “information bottleneck” [TAO]
- ▶ Sélection de modèle [TAO]

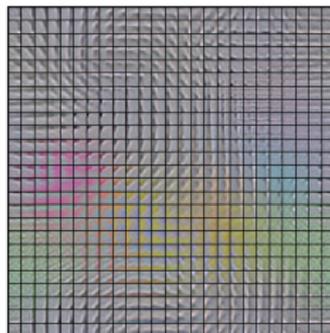
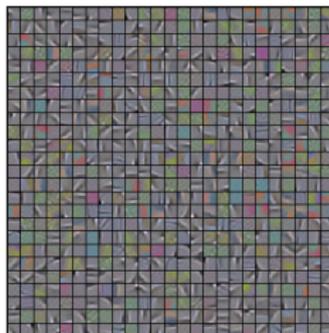
# Livrable 2.3 : Optimisation

## Axe de recherche : étude de solutions “profondes” alternatives

- ▶ Méthodes à noyaux pour l'apprentissage profond [LITIS]

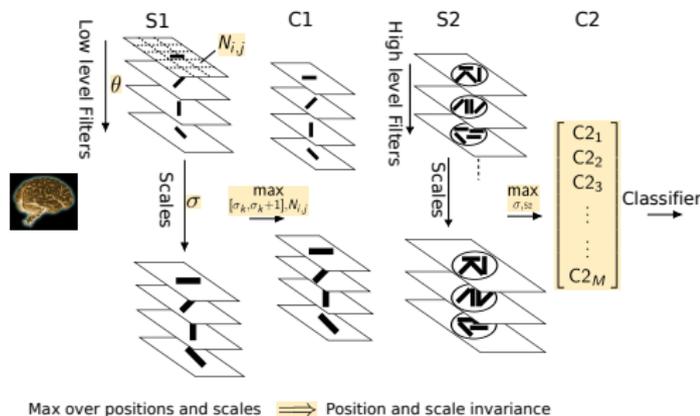
$$K(\mathbf{x}, \mathbf{x}') = \langle \Phi \circ \Phi \circ \dots \circ \Phi(\mathbf{x}), \Phi \circ \Phi \circ \dots \circ \Phi(\mathbf{x}') \rangle$$

- ▶ kernel design
- ▶ apprentissage du modèle
- ▶ Apprentissage de caractéristiques pour des images couleurs [LIP6]
  - ▶ robuste aux changements d'illumination
  - ▶ robuste aux transformations affines
  - ▶ organisé topographiquement



# Livrable 2.3 : Structure d'inspiration biologique

- Cadre : architecture HMAX (Serre et al, 2007, Mutch et al. 2008)



- S1 : filtrage de Gabor
- C1 : invariance locale
- S2 : caractéristiques intermédiaires par matching
- C2 : invariance globale

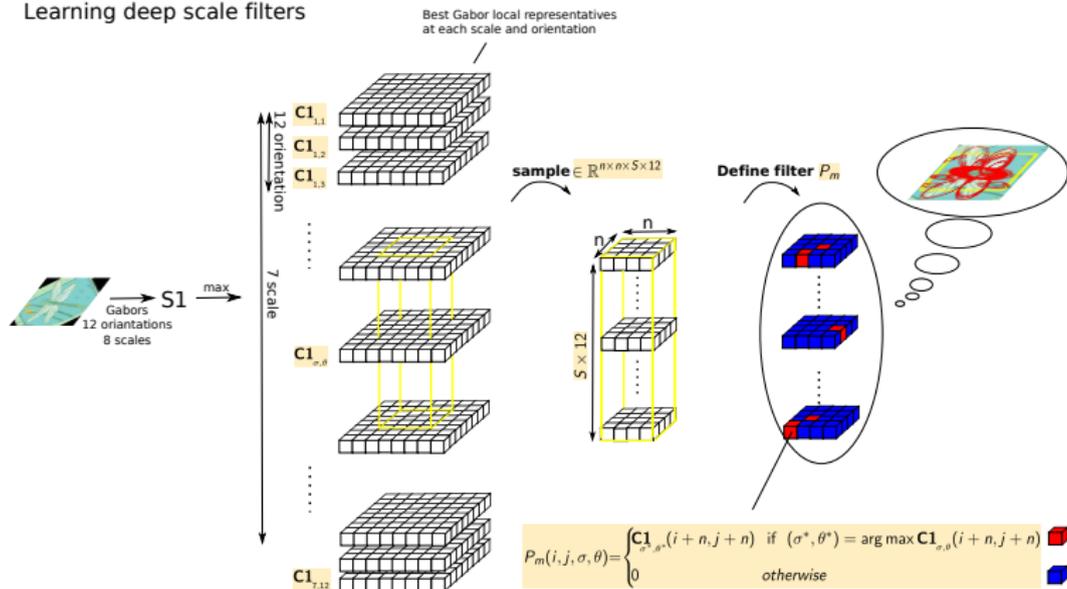
## Contributions [LIP6]

- Prise en compte de l'aspect **multi-échelle** : apprentissage de prototypes visuels adaptés

# Livrable 2.3 : Structure d'inspiration biologique

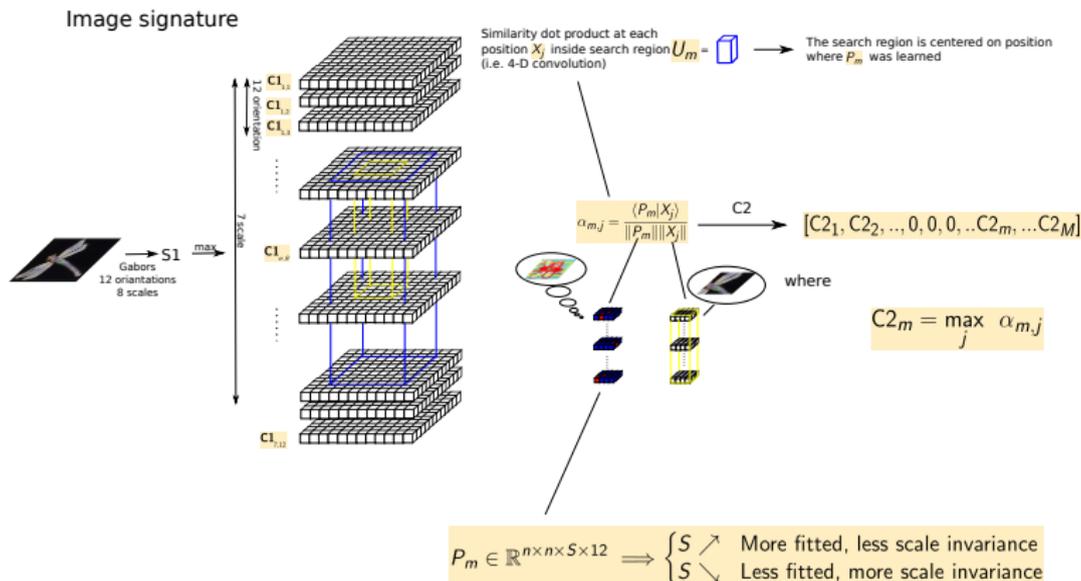
## Apprentissage de prototypes multi-échelles

Learning deep scale filters



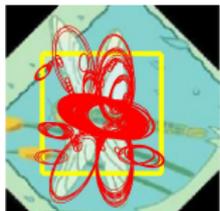
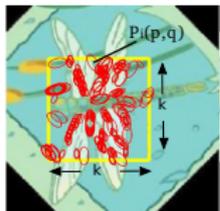
# Livrable 2.3 : Structure d'inspiration biologique

## Caractéristiques multi-échelles

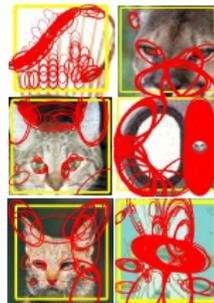


# Livrable 2.3 : Structure d'inspiration biologique

## ► Comparaison



## ► Illustration



| Model     | 15 images           | 30 images          |
|-----------|---------------------|--------------------|
| Our model | <b>59.1 ± 0.2 %</b> | <b>66.9 ± 0.8%</b> |

| Architecture profondes biologiquement inspirées |       |       |
|---|-------|-------|
| Mutch&Lowe                                      | 48    | 54    |
| Zeiler et al                                    | 58.6  | 66.9  |
| Fidler et al                                    | 60.5  | 66.5  |
| Architectures plates                            |       |       |
| Lazebnik et al                                  | 56.4  | 64.6  |
| Zhang et al                                     | 59.1  | 62.2  |
| Wang et al                                      | 64.43 | 73.44 |

# Perspectives

- ▶ Passage à l'échelle des méthodes à noyaux
- ▶ Contrainte de parcimonie dans la couche S2 de HMAX
- ▶ Développement des approches multi-tâches (Livrable 2.4)
- ▶ Etude et analyse des récents travaux de S. Mallat sur une structure profonde à base d'ondelettes

$$||x \star \psi| \star \psi| \cdots \star \psi$$

# Plan

1. Mise en contexte et objectifs scientifiques du projet
2. Tâche 1 : Apprentissage de caractéristiques sur une couche
3. Tâche 2 : Structure profonde pour l'apprentissage de caractéristiques
  - ▶ Structure d'inspiration biologique
4. Tâche 3 et 4 : Implémentation et Démonstrateurs
  - ▶ Compétition Unsupervised and Transfer Learning Challenge
5. Résultats marquants et impacts
6. Perspectives et conclusions

<http://sites.google.com/site/anrasap/>

# Tâche 3 et 4

## Livrables 3

- ▶ Mise à disposition de la communauté de paquets logiciels permettant de reproduire les travaux ASAP

## Livrables 4

- ▶ Livrable 4.1 : Images et indexation en cours
- ▶ Livrable 4.2 : Biométrie en cours
  - ▶ mise en place d'une plate-forme de benchmarking de systèmes biométriques
- ▶ Livrable 4.3 : Interfaces cerveau-machine débutée

**Zoom sur la compétition Unsupervised and Transfer Learning Challenge**

# compétition UTLC

## Objectifs de la compétition

Évaluer **objectivement** les techniques d'apprentissage de la représentation dans un contexte

- ▶ non-supervisé
- ▶ de transfert d'apprentissage

## Principe

- ▶ Etant donné des données de  $\mathbb{R}^d$ , construire une nouvelle représentation dans  $\mathbb{R}^K$
- ▶ non-supervisée : la “qualité” de la nouvelle représentation est évaluée sur une tâche d'apprentissage supervisé (seul les organisateurs connaissent les labels)
- ▶ transfert d'apprentissage : les labels des exemples de certaines classes sont fournis, mais les données de test ne sont pas de même classes

# compétition UTLC

## Données

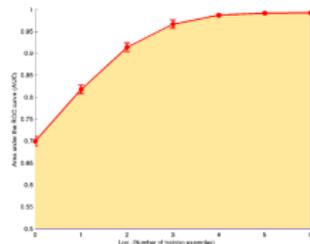
- ▶ Reconnaissance de caractères manuscrites
- ▶ Reconnaissance d'action dans les vidéos
- ▶ Reconnaissance d'objets dans les images
- ▶ Reconnaissance de forêts dans les images satellites
- ▶ Catégorisation de textes

AVICENNA  
HARRY  
RITA  
SYLVESTER  
TERRY



## Evaluation

- ▶ Classifieur linéaire fourni par les organisateurs
- ▶ l'aire sous la learning curve (ALC) en AUC



# compétition UTLC

## Approche [LITIS + LISA]

- ▶ Pré-traitement : normalisation, whitening
- ▶ Apprentissage de la représentation : approche **réseaux profonds**
  - ▶ Empilement de structures simples (au moins 3 niveaux)
  - ▶ Denoising Auto-Encodeurs, Contractive Auto-Encodeurs,...
- ▶ Post-traitement : ACP sur les données de validations ou tests
- ▶ Sélection de modèle

## Résultats

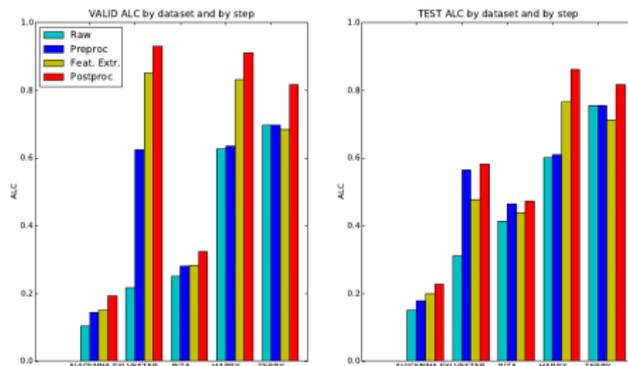
| Phase 1 - Unsupervised Learning |        |            |            |            |            |            |                   |
|---------------------------------|--------|------------|------------|------------|------------|------------|-------------------|
| Rank                            | Team   | Experiment | AVICENNA   | HARRY      | RITA       | SYLVESTER  | TERRY             |
| 1                               | AIO    | AIO        | 0.2183 (1) | 0.7043 (6) | 0.4951 (1) | 0.4569 (6) | <b>0.8465</b> (1) |
| 2                               | 1055A  | exp1       | 0.1906 (6) | 0.7357 (3) | 0.4782 (5) | 0.5828 (1) | 0.8437 (2)        |
| 3                               | Airbus | A3XX       | 0.2174 (2) | 0.7545 (2) | 0.4724 (7) | 0.4949 (4) | 0.8390 (3)        |
| 4                               | LISA   | LISA       | 0.1960 (5) | 0.8062 (1) | 0.4731 (6) | 0.4763 (5) | 0.7959 (6)        |

| Phase 2 - Transfer Learning |       |            |                   |                   |                   |                   |            |
|-----------------------------|-------|------------|-------------------|-------------------|-------------------|-------------------|------------|
| Rank                        | Team  | Experiment | AVICENNA          | HARRY             | RITA              | SYLVESTER         | TERRY      |
| 1                           | LISA  | agartha    | <b>0.2273</b> (1) | <b>0.8619</b> (1) | <b>0.5029</b> (1) | 0.5650 (3)        | 0.8160 (2) |
| 2                           | tkgw  | crush      | 0.1973 (2)        | 0.7533 (2)        | 0.4095 (4)        | <b>0.5933</b> (1) | 0.8118 (3) |
| 3                           | 1055A | phase2exp1 | 0.1511 (4)        | 0.7381 (3)        | 0.4992 (2)        | 0.5873 (2)        | 0.8437 (1) |
| 4                           | FAST  | teaf       | 0.1909 (3)        | 0.3580 (4)        | 0.4275 (3)        | 0.3379 (5)        | 0.6485 (4) |

# compétition UTLC

## Evolution des performances



## Leçons à retenir

- ▶ apprendre des caractéristiques de manière **non-supervisée** peut améliorer les performances de reconnaissances
- ▶ **le transfert d'apprentissage** améliore les performances
- ▶ les approches basées sur les **réseaux profonds** sont compétitives
- ▶ le bon vieux ACP est toujours très utile

# Plan

1. Mise en contexte et objectifs scientifiques du projet
2. Tâche 1 : Apprentissage de caractéristiques sur une couche
3. Tâche 2 : Structure profonde pour l'apprentissage de caractéristiques
  - ▶ Structure d'inspiration biologique
4. Tâche 3 et 4 : Implémentation et Démonstrateurs
  - ▶ Compétition Unsupervised and Transfer Learning Challenge
5. Résultats marquants et impacts
6. Perspectives et conclusions

<http://sites.google.com/site/anrasap/>

# Impacts du projet

- ▶ Publications
  - ▶ 5 journaux (Pattern Recognition, Signal Processing,...)
  - ▶ 13 conférences internationales
  - ▶ 2 conférences nationales
- ▶ co-publications avec ANR Lampada
  - ▶ thèmes : approximation parcimonieux et apprentissage de données fonctionnelles
- ▶ collaborations internationales
  - ▶ U. de Montréal : équipe phare en Apprentissage Profond (auto-encodeur)
    - ▶ thèse en co-tutelle
    - ▶ 2 publications jointes
    - ▶ séjour de recherche (2012)
  - ▶ NEC Labs (Princeton) : équipe reconnue sur le thème
    - ▶ ranking et réseaux de neurones (une publication jointe)
  - ▶ Université du Québec
    - ▶ séjour de recherche en cours
  - ▶ Institute of Infocomm research (Singapour)
    - ▶ thèse en co-tutelle
    - ▶ 2 publication jointes

# Résultats marquants

- ▶ Session spéciale à ESANN 2011
  - ▶ une présentation tutorial
  - ▶ 1 article du consortium
  - ▶ 3 articles hors-consortium
- ▶ Compétition UTLC 2011
  - ▶ Evaluation des méthodes non-supervisées et du transfert d'apprentissage
  - ▶ 4eme/41 (non-supervisée) et 1er/14 place (transfert)
- ▶ Compétition BCI MLSP 2010
  - ▶ Classification de signaux EEG contenant des potentiels évoqués P300
  - ▶ méthodes à noyaux (3eme/35) et approches réseaux profonds (6eme/35)

# Planning et Perspectives

- ▶ Poursuite des autres tâches en accord avec le programme
  - ▶ résultats prometteurs sur plusieurs tâches
    - ▶ invariance multi-échelle
    - ▶ apprentissage profond et noyau
    - ▶ robustesse aux transformations
  - ▶ applications à des problèmes réelles difficiles
    - ▶ Biométrie : intégration et évaluation d'une approche réseau profond sur la plate-forme biométrique
    - ▶ Interface Cerveau-machine
- ▶ nouvelle perspective de recherche **Ondelettes et représentations profondes** :
  - ▶ opportunité
  - ▶ résultats encourageants (ICM)
  - ▶ déroutage des efforts humains LIF - tâche 3 sur cette nouvelle tâche

# Conclusions personnelles

- ▶ première expérience de coordinateur
- ▶ un consortium agréable
- ▶ des équipes qui avancent et atteignent des jalons

# Conclusions personnelles

- ▶ première expérience de coordinateur
- ▶ un consortium agréable
- ▶ des équipes qui avancent et atteignent des jalons

Merci de votre attention