

# *Apprentissage par Réseaux de Neurones Profonds*



COLLÈGE  
DE FRANCE  
—1530—

---

## *Cours 1*

### *Introduction aux réseaux de Neurones Profonds*

*Stéphane Mallat*

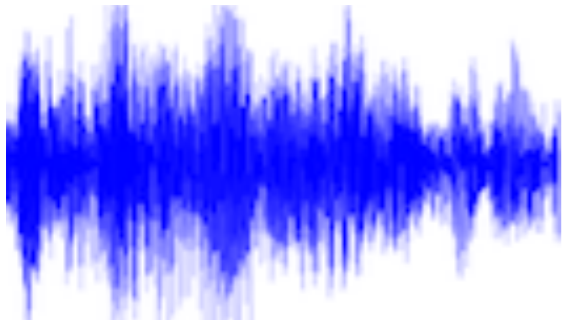


# Réseaux de Neurones Profonds

- Domaine algorithmique et empirique avec des applications spectaculaires et soulève des questions fondamentales.
- But du cours: essayer de comprendre *comment* et *pourquoi* cela marche.
- Au-delà de l'algorithmique:
  - Besoin de nouveaux outils mathématiques: statistiques, probabilités, optimisation, analyse harmonique, systèmes dynamiques, géométrie...
  - L'intelligence artificielle
  - « L'architecture de la complexité » en sciences et sciences humaines.

Données diverses mais avec un grand nombre  $d$  de variables

$$d \sim 10^6 / mn$$



Audio

$$d \sim 10^6$$

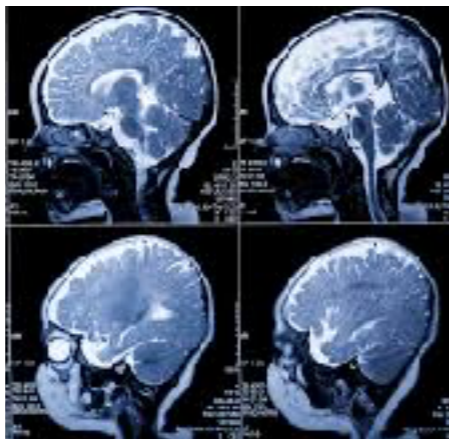


Image 2D/3D

$$d \sim 10^6$$



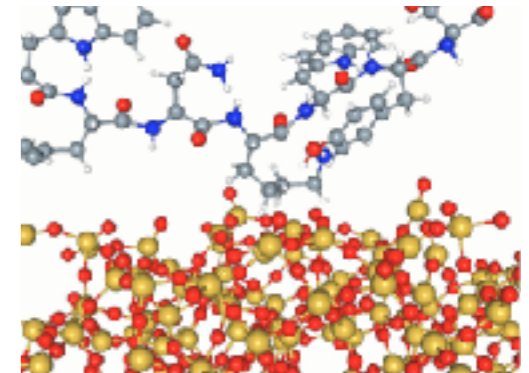
Texte

$$d \sim 10^9$$



Reseau social

$$d \sim 10^{24}$$



mole de matière

- *Modéliser*: capturer la nature et la variabilité des données
  - compression, restauration, synthèse
- *Prédire*: estimer la réponse d'une question à partir de données
  - Reconnaissance d'images ou de sons, diagnostics médicaux, analyse de textes ou traductions, prédire la physique...

Apprentissage statistique  $\longrightarrow$  Intelligence artificielle



- High-dimensional  $x = (x(1), \dots, x(d)) \in \mathbb{R}^d$ :
- **Classification:** estimate a class label  $f(x)$   
given  $n$  sample values  $\{x_i, y_i = f(x_i)\}_{i \leq n}$

Image Classification  $d = 10^6$

Anchor



Joshua Tree



Beaver



Lotus



Water Lily



Huge variability  
inside classes

Find invariants



# Prediction: Supervised Learning

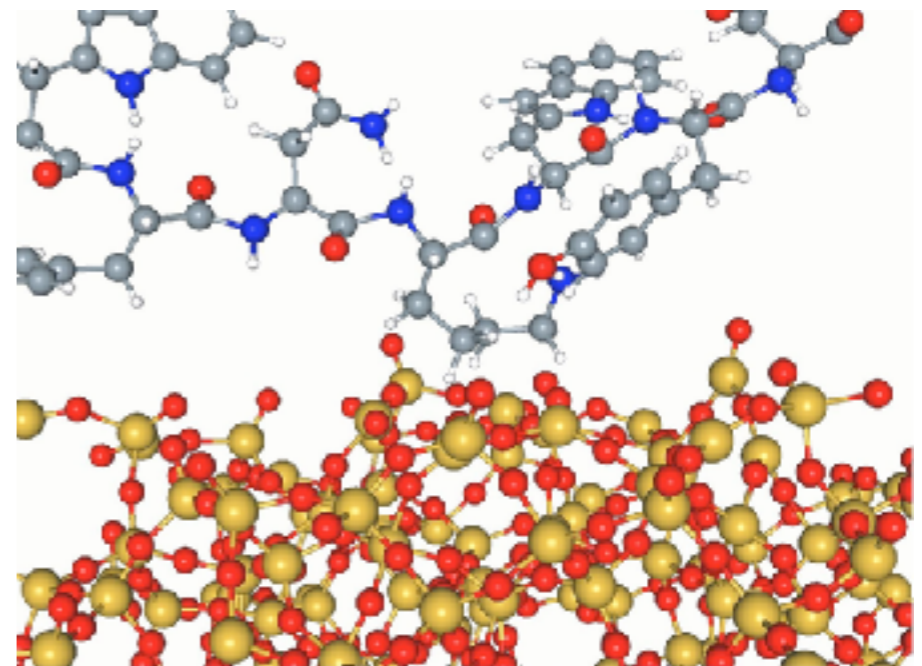
- High-dimensional  $x = (x(1), \dots, x(d)) \in \mathbb{R}^d$ :
- **Regression:** approximate a *functional*  $f(x)$   
given  $n$  sample values  $\{x_i, y_i = f(x_i) \in \mathbb{R}\}_{i \leq n}$

Physics: energy  $f(x)$  of a state vector  $x$

Astronomy



Quantum Chemistry



Importance of symmetries.



# Algorithmes d'Apprentissage

Estimer la réponse  $y$  à partir des données  $x$



Quel est cet animal ?  $y = \text{chien}$



## Supervisé

**Apprentissage:** optimise les paramètres pour faire peu d'erreur sur les exemples  $\tilde{y}_i \approx y_i$

**Généralisation** si peu d'erreur pour des  $x$  inconnus:  $\tilde{y} \approx y$

*Mathématiques:* comprendre les conditions de généralisation  
**régularité**

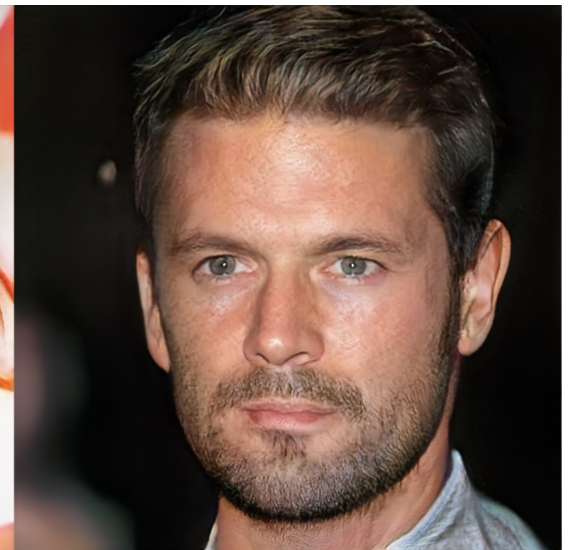
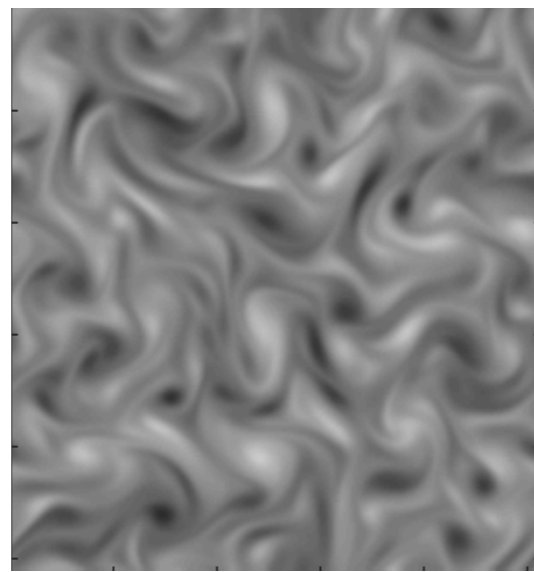
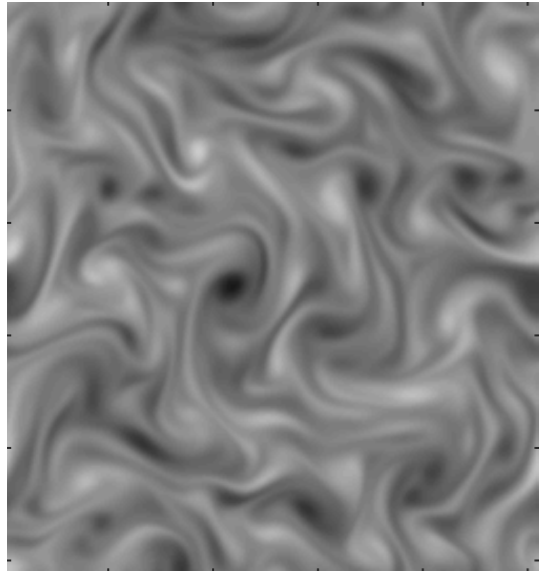


*Informatique:* développer des algorithmes efficaces

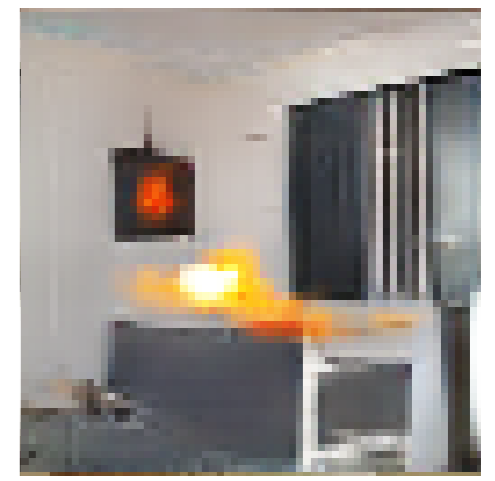
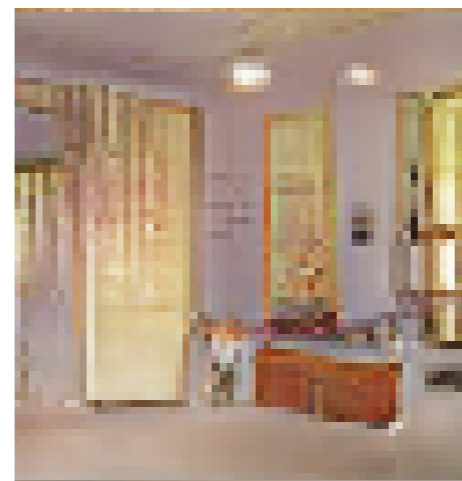
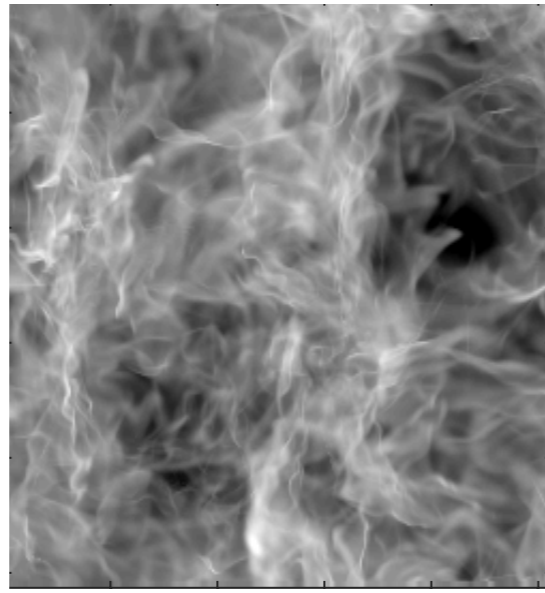


- Estimate the probability density  $p(x)$  of  $X(u)$  from few realisations  $\{x_i(u)\}_i$

## Turbulence de fluide



gaz interstellaire (hydrogène, hélium)



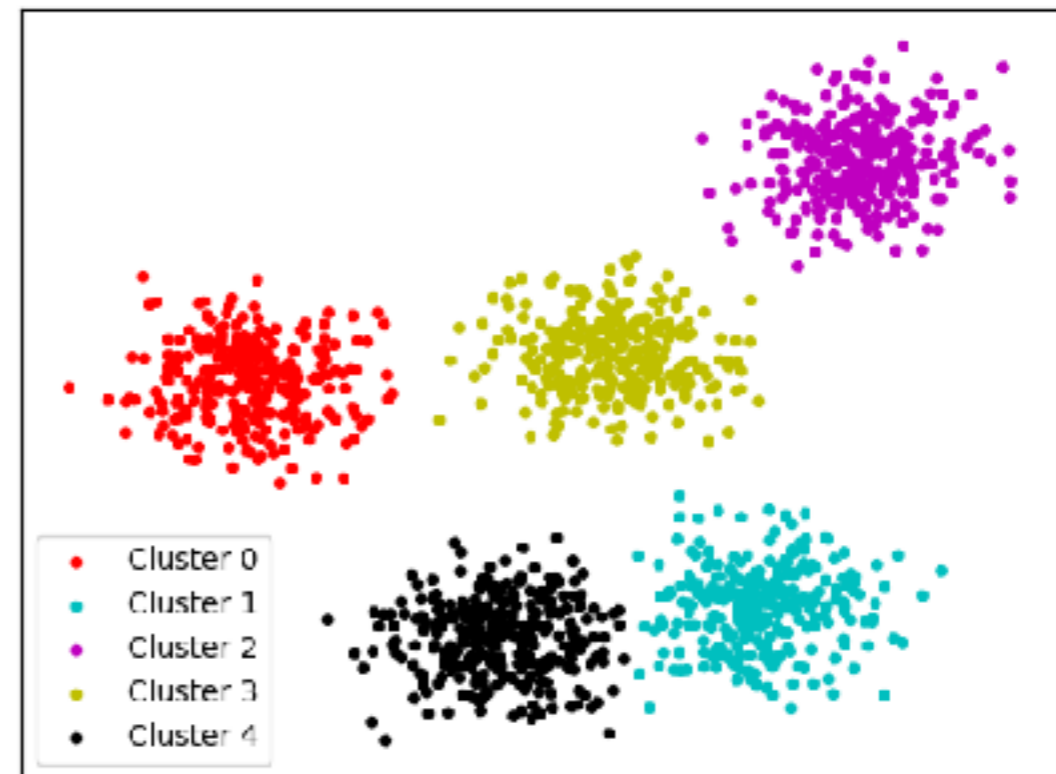




# Unsupervised Clustering

- Unsupervised learning for classification

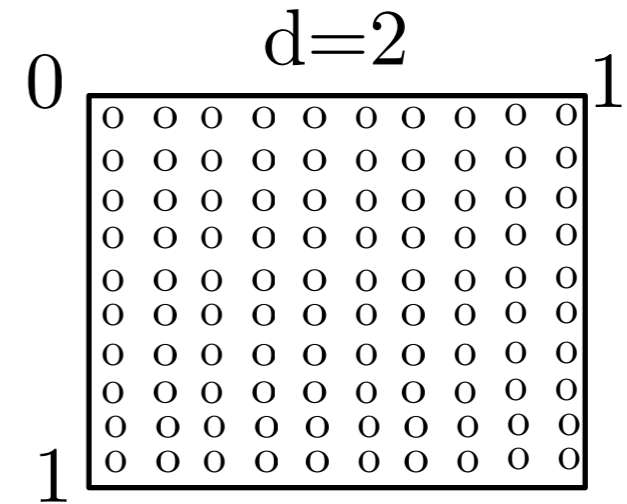
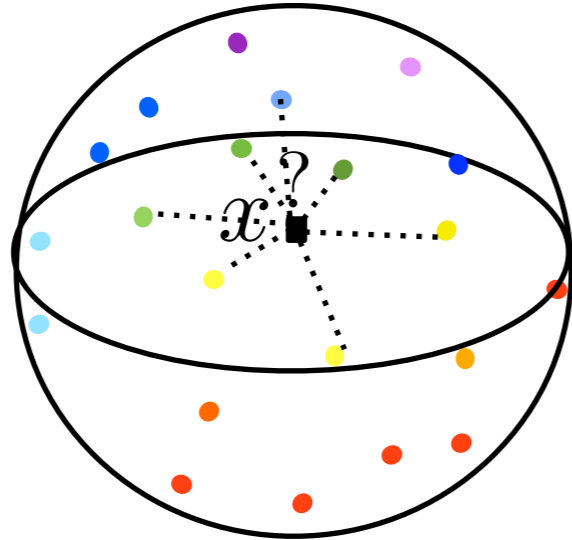
3 6 8 1 7 9 6 6 9 1  
6 7 5 7 8 6 3 4 8 5  
2 1 7 9 7 1 2 8 4 5  
4 8 1 9 0 1 8 8 9 4



- Does not work in high dimension

# Curse of Dimensionality

- $f(x)$  can be approximated from examples  $\{x_i, f(x_i)\}_i$  by local interpolation if  $f$  is regular and there are close examples:



- To cover  $[0, 1]^d$  at a distance  $10^{-1}$  we need  $10^d$  points

Problem:  $\|x - x_i\|$  is always large



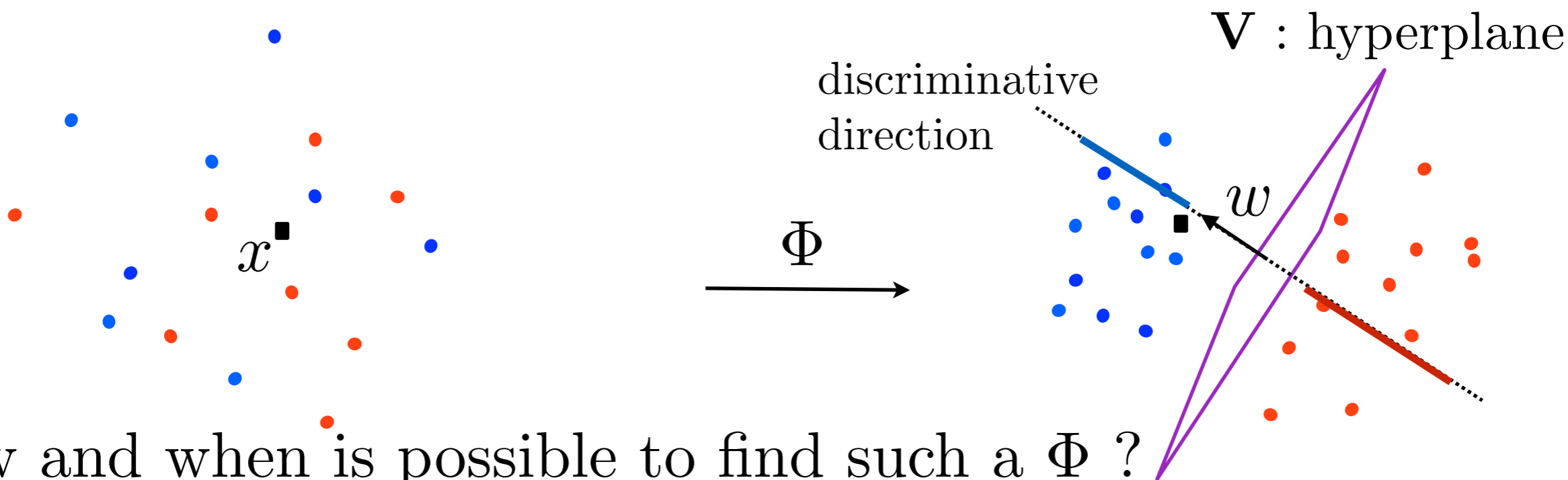


# Kernel Classifiers: Linearize

Change of variable  $\Phi(x) = \{\phi_k(x)\}_{k \leq d'}$  to nearly linearize class boundaries, and approximate  $f(x)$  by:

$$\tilde{f}(x) = \text{sign}(\langle w, \Phi(x) \rangle + b) = \text{sign}\left(\sum_k w_k v'_k + b\right)$$

$$x = (v_1, \dots, v_d) \longrightarrow \Phi(x) = (v'_1, \dots, v'_d)$$



- How and when is possible to find such a  $\Phi$  ?  
A priori/learned ?



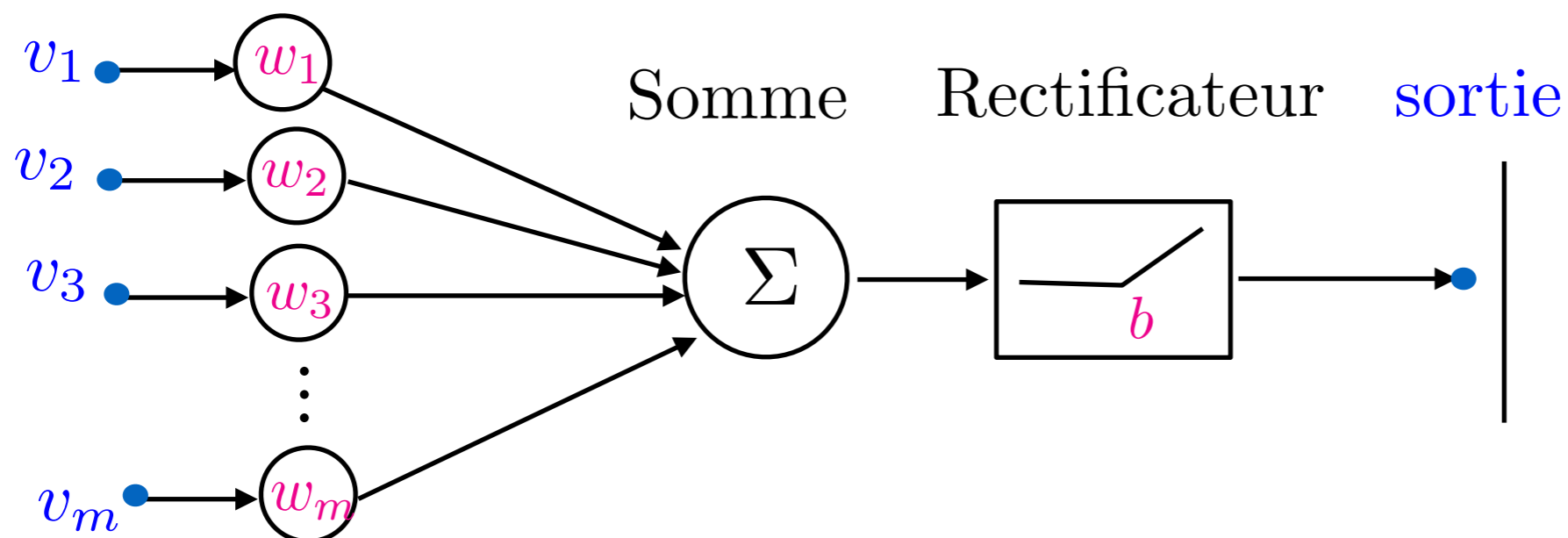
camions de pompiers versus voitures



# Neurone Informatique

- Début dans les années 1950 : inspiration biologique  
*McCulloch & Pitts* : modèle de neurone de grenouilles

entrées



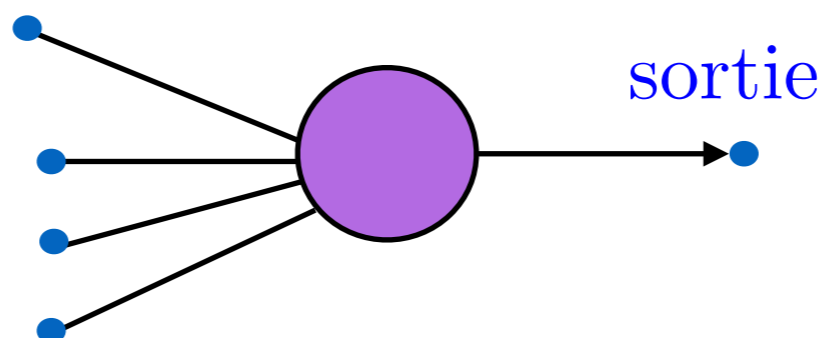
*Parcimonie*

$$0 \text{ si } \sum_{k=1}^m v_k w_k < b$$

$$\sum_{k=1}^m v_k w_k - b \text{ sinon}$$

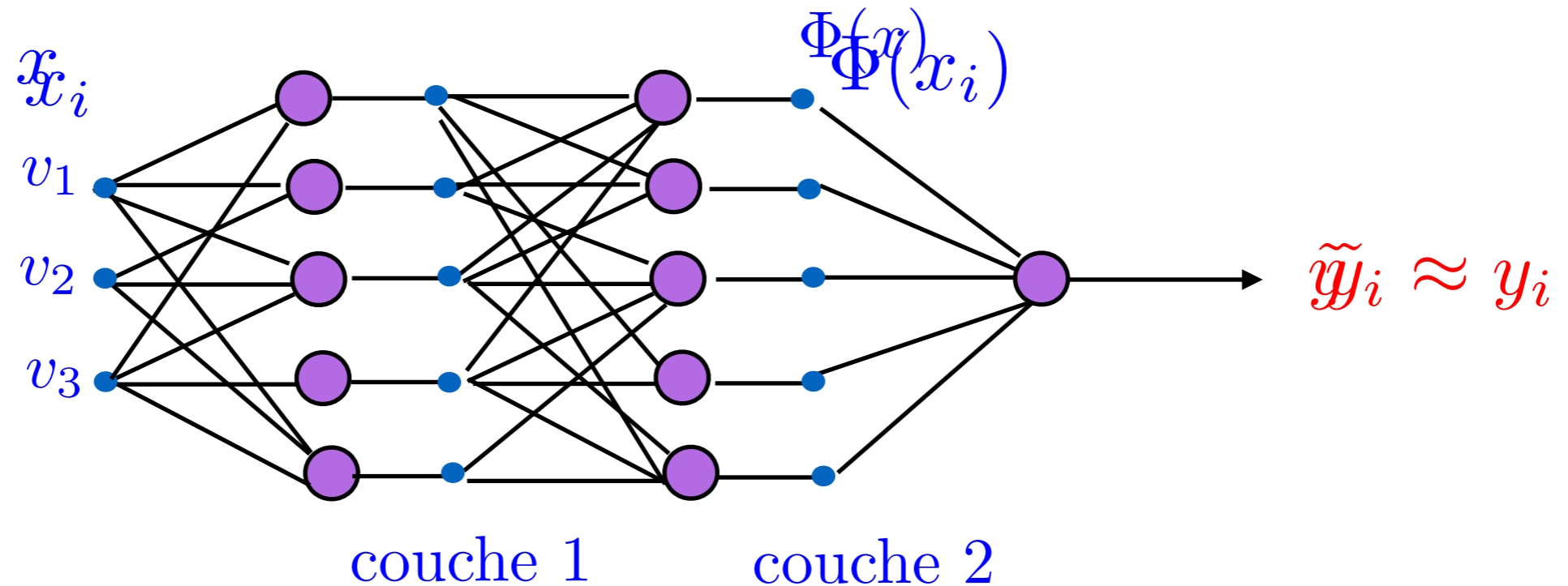
Un neurone est paramétré par les  $w_k$  et le seuil  $b$

entrées





# Réseaux de Neurones

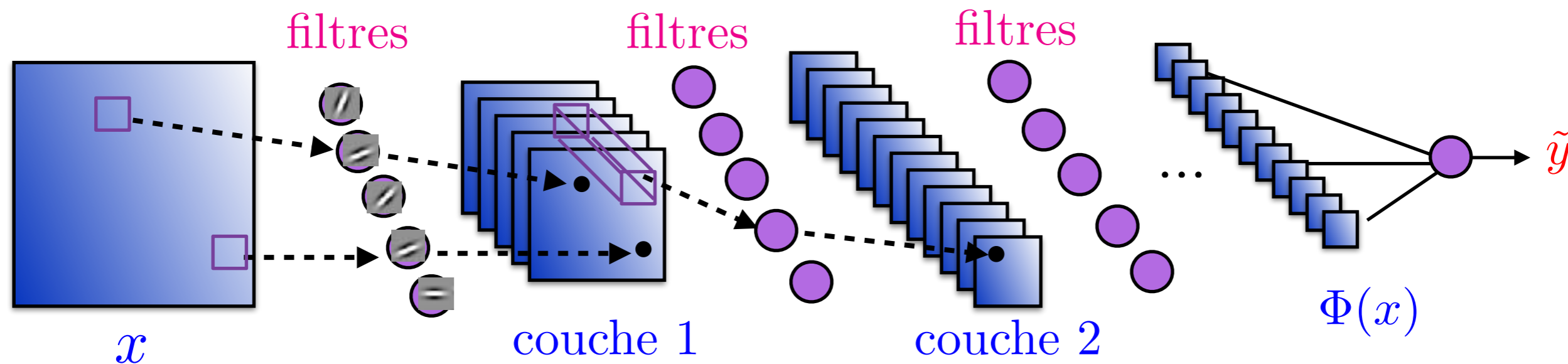


**Apprentissage:** optimise les **paramètres** pour minimiser les erreurs sur les exemples. Problèmes difficiles d'optimisation.

**Information à priori:** elle est dans l'architecture du réseau.

- Les paramètres sont invariants par translations: **filtres**

*Y. LeCun*



Centaines de millions de paramètres

Prédictions exceptionnelles et *génériques*: pour l'image, parole, diagnostiques, traduction, en physique...

Le  $\Phi(x)$  a aplati les frontières. Comment ?

Importance de l'architecture





# Plan du Cours

- Applications spectaculaires et questions mathématiques
- L'origine: la Cybernétique: *Wiener 1947, Herbert Simons 1960*  
le Perceptron: *Franck Roseblatt 1957*
- Réseaux à deux couches:
  - Théorème d'approximation universelle: *Cybenko 1988*
  - Théorie de l'approximation
- Réseaux multicouches:
  - Performances d'approximations ?
  - Optimisation par gradient stochastique: *Robbins, Monro 1951*
  - Rétro-propagation: *Rumelhart, Hinton, Geoffrey 1986*
- Réseaux convolutifs: *Y. LeCun 1990, Kriev. Hint. .Sus. 2012*  
*Un monde algorithmique, mathématique et d'applications*



# Challenges de Données

challengedata.ens.fr

- Calculer une prédiction  $\tilde{y}$  de  $y$  à partir de données  $x$
- Exemples supervisés  $\{x_i, y_i\}_{i \leq n}$
- Données de test  $\{x_i^t\}_{i \leq n_t}$
- On doit calculer une prédiction  $\tilde{y}_i^t$  de chaque  $y_i^t$  inconnu
- Erreur évaluée par une fonction de risque  $r(\tilde{y}, y)$

$$\text{Score} = n_t^{-1} \sum_{i=1}^{n_t} r(\tilde{y}_i^t, y_i^t)$$

- Maximum 2 soumissions par jour
- Résultat intermédiaire en Juin et final en Décembre 2019.
- Oral en début Avril pour étudiants du MVA



- 23 Janvier: Présentation de 7 challenges de données
- 30 Janvier: Présentation de 7 challenges de données
- 6 Février: Présentation des gagnants des challenges 2018
- 13 Février: Josef Sivics, ENS, sur la classification d'images
- 20 Février: Pitor Bojanowski , Facebook, sur la langage naturel
- 27 Février : Cordelia Schmid, INRIA sur la classification vidéo
- 6 Mars: Pas de cours et de séminaire.
- 13 Mars : Yann Olivier, sur l'apprentissage par renforcement
- 20 Mars: Chercheur de Owkin, sur l'imagerie médicale





# Applications: La Vision

- Reconnaissance d'animaux: 150ms

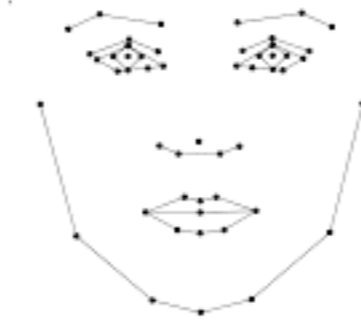
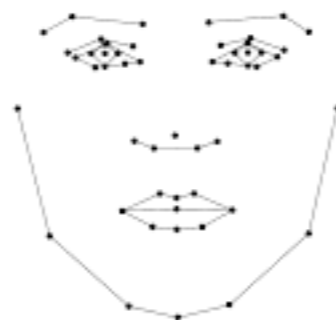
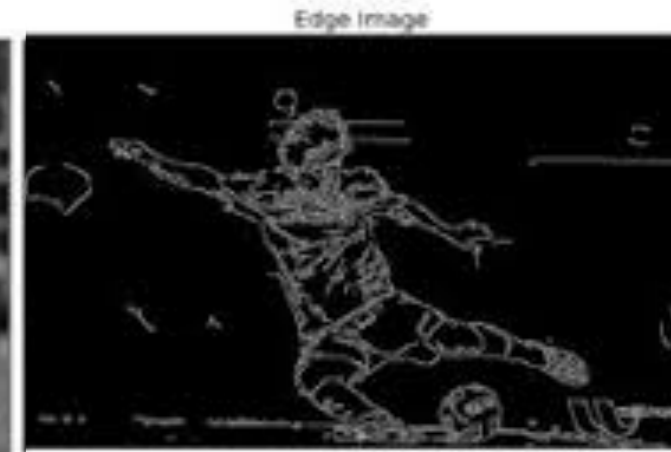






# La vision par ordinateur

- Quel sont les structures élémentaires ? paternes ?
- Detection de contours, filtrage, détections de coins
- Comment sont elles agencées ?
- Qu'est-ce qui caractérise une forme, un objet ?
- Représentation symbolique de la connaissance
- Une approche influencée par le langage: la grammaire des images





# ImageNet Data Basis

- Data basis with 1 million images and 2000 classes



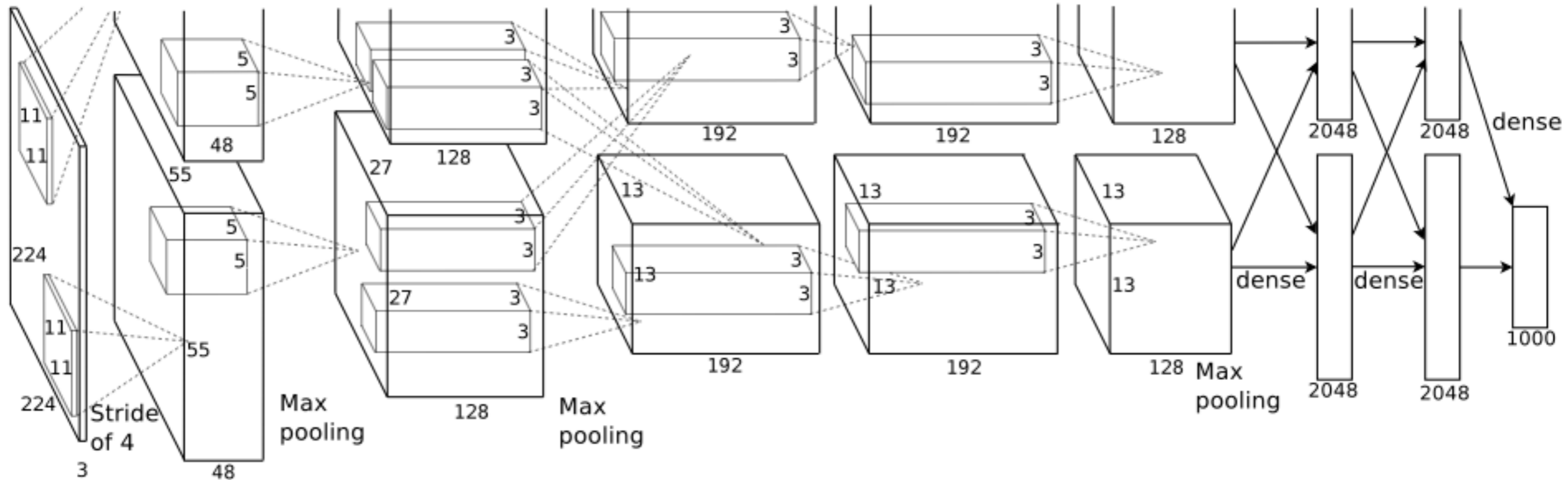




# Alex Deep Convolution Network

*A. Krizhevsky, Sutssever, Hinton*

- Imagenet supervised training:  $1.2 \cdot 10^6$  examples,  $10^3$  classes  
15.3% testing error in 2012



New networks with 3% errors.

Up to 150 layers!

Pourquoi cela marche ?



# Image Classification



**mite**

**container ship**

**motor scooter**

**leopard**

█	mite
█	black widow
█	cockroach
█	tick
█	starfish

█	container ship
█	lifeboat
█	amphibian
█	fireboat
█	drilling platform

█	motor scooter
█	go-kart
█	moped
█	bumper car
█	golfcart

█	leopard
█	jaguar
█	cheetah
█	snow leopard
█	Egyptian cat



**grille**



**mushroom**



**cherry**



**Madagascar cat**

█	convertible
█	grille
█	pickup
█	beach wagon
█	fire engine

█	agaric
█	mushroom
█	jelly fungus
█	gill fungus
█	dead-man's-fingers

█	dalmatian
█	grape
█	elderberry
█	ffordshire bullterrier
█	currant

█	squirrel monkey
█	spider monkey
█	titi
█	indri
█	howler monkey

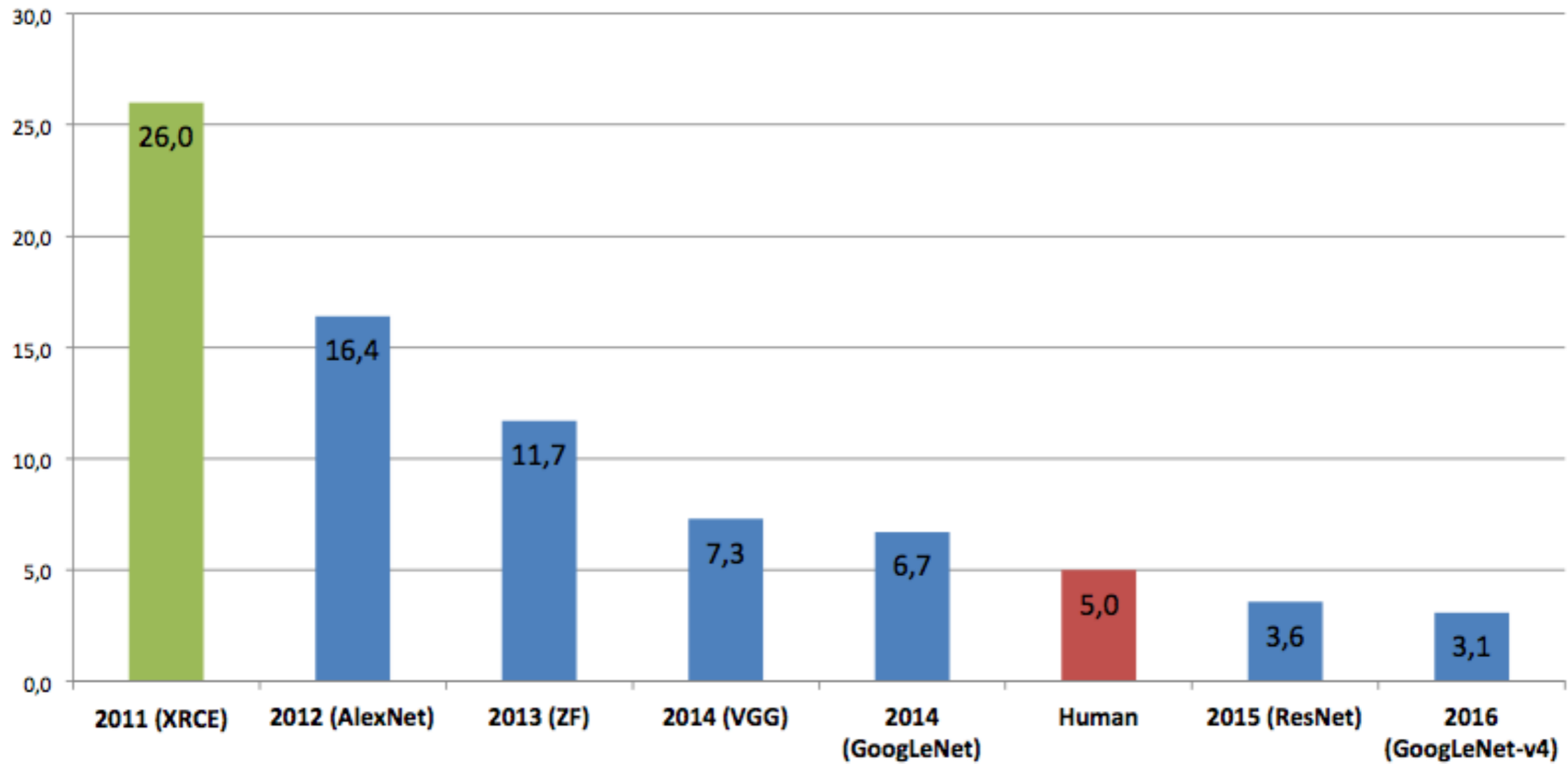
Quels types d'invariants (connaissance) ont été appris ?





# Erreurs sur ImageNet

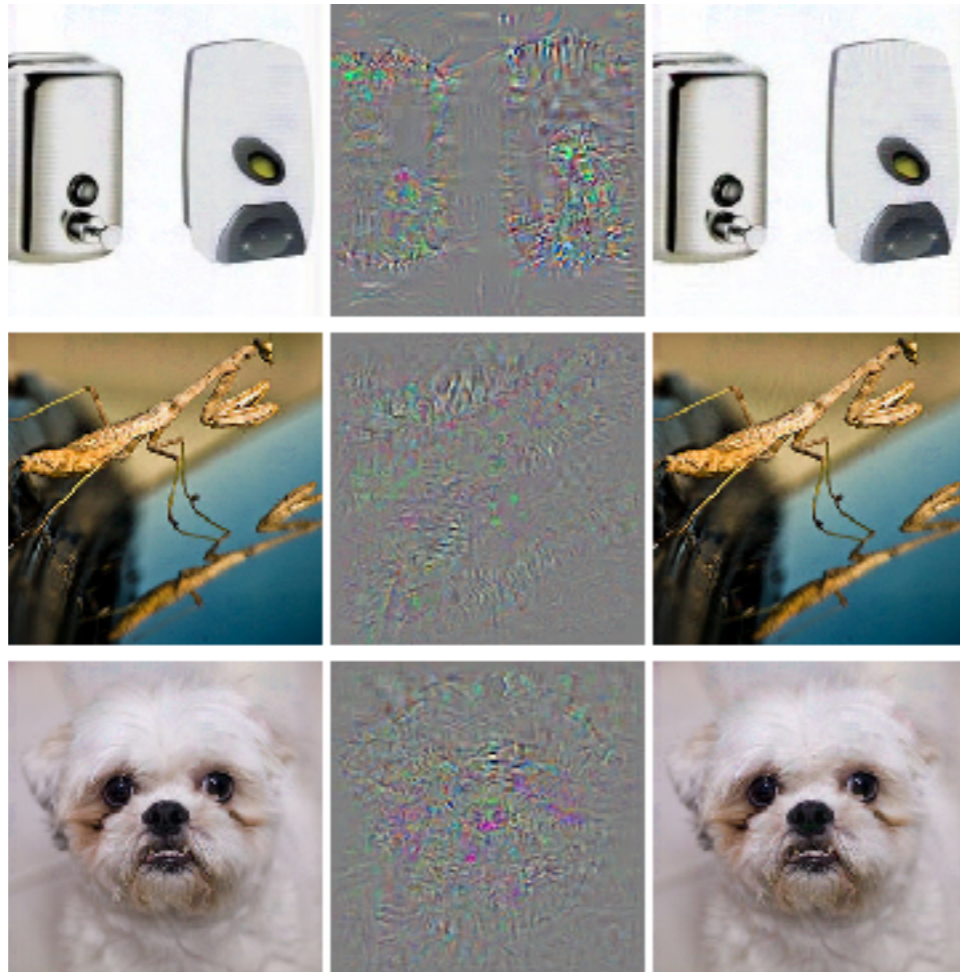
ImageNet Classification Error (Top 5)



# Why Understanding ?

Szegedy, Zaremba, Sutskever, Bruna, Erhan, Goodfellow, Fergus

$$x + \epsilon = \tilde{x} \quad \text{with} \quad \|\epsilon\| < 10^{-2} \|x\|$$



correctly  
classified

classified as  
ostrich

- Trial and error testing can not guarantee reliability.



# Sports Video Classification

Large-scale Video Classification with Convolutional Neural Networks  
CVPR 2014

- Trouver le chien:



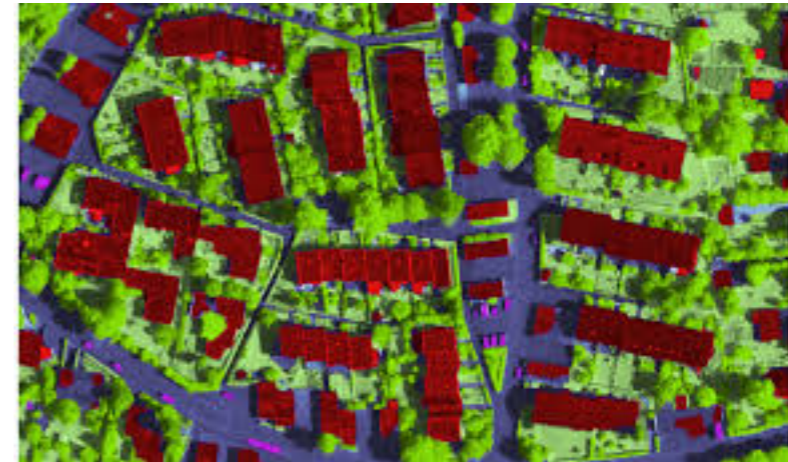
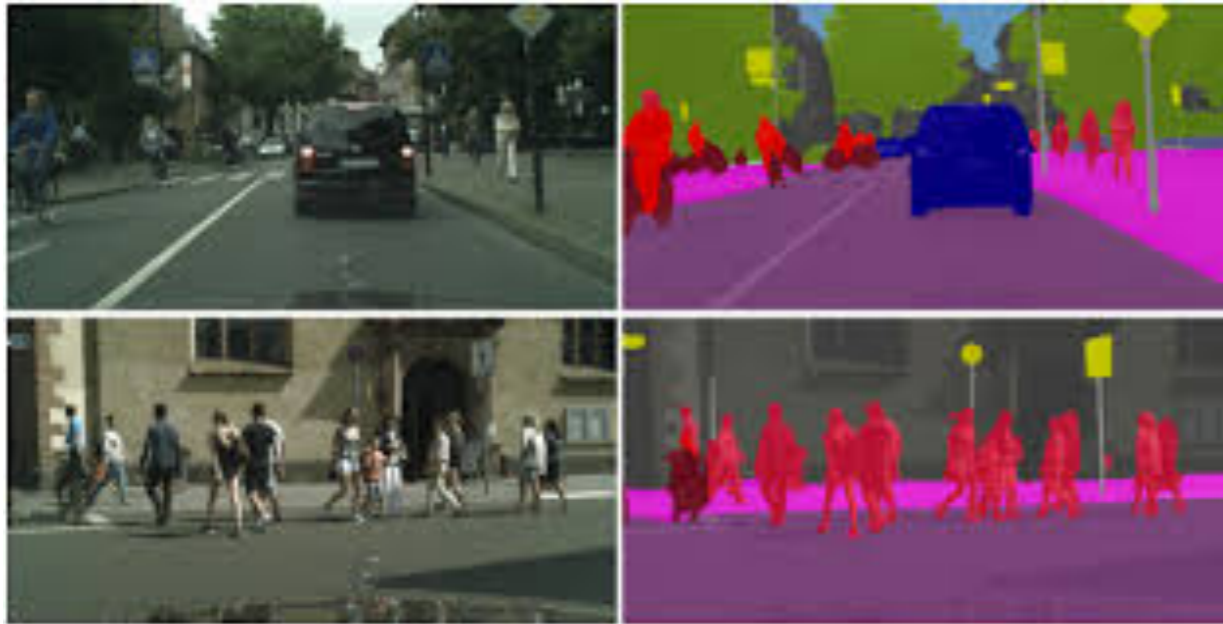
- Séparer les composantes:







# Segmentation d'Images





- Sky
- Building
- Pole
- Road Marking
- Road
- Pavement
- Tree
- Sign Symbol
- Fence
- Vehicle
- Pedestrian
- Bike



