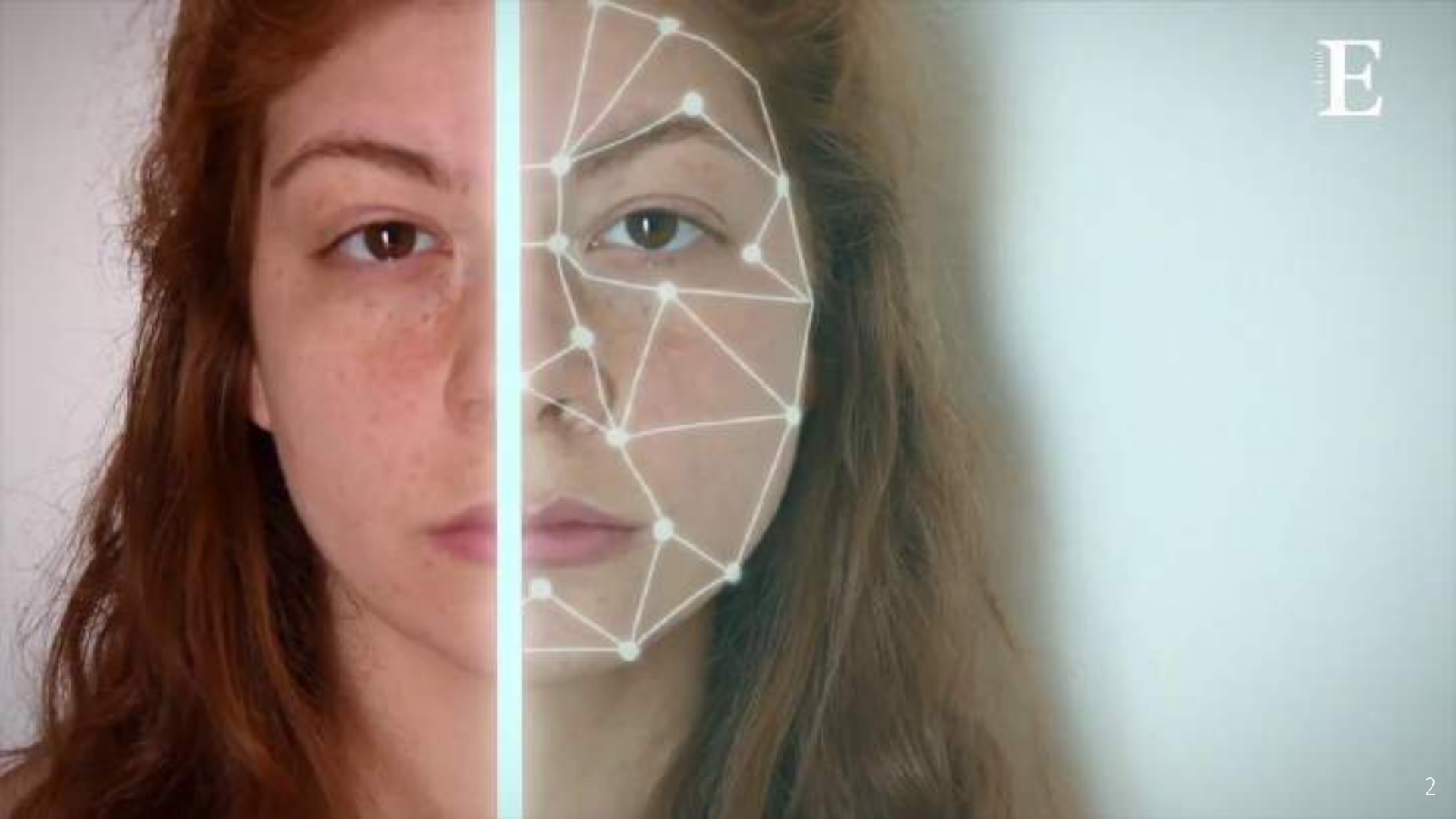


PIR 11 : Reconnaissance de visages

Encadrant : Khalid IDRISSE, LIRIS



Mokhtar EL BOURAQUI, Zaineb HARRIZI , Noëline MIGEON, Elhadji NIANG, Pénélope ROULLOT



E

DÉTECTION VS. RECONNAISSANCE



https://en.wikipedia.org/wiki/Face_detection



face scan by Franck Boston/www.shutterstock.com

DIFFÉRENTES APPROCHES

DESCRIPTION



SIFT & LBP

APPRENTISSAGE



Réseaux de
neurones



Eigenfaces &
Fisherfaces

01 | **SIFT**
Scale-Invariant Feature Transform

02 | **LBP**
Local Binary Patterns



03 | **EIGENFACES**

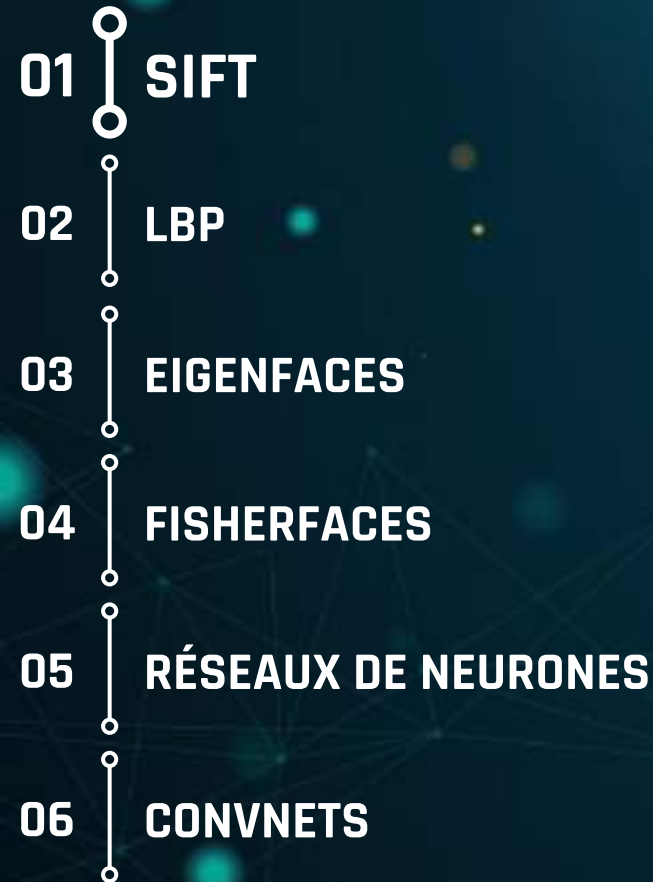
04 | **FISHERFACES**



05 | **RÉSEAUX DE NEURONES**
Théorie, AlexNet, transfer learning

06 | **CONVNETS**
Convolutional Neural Networks (CNN)



- 
- 01 SIFT
 - 02 LBP
 - 03 EIGENFACES
 - 04 FISHERFACES
 - 05 RÉSEAUX DE NEURONES
 - 06 CONVNETS

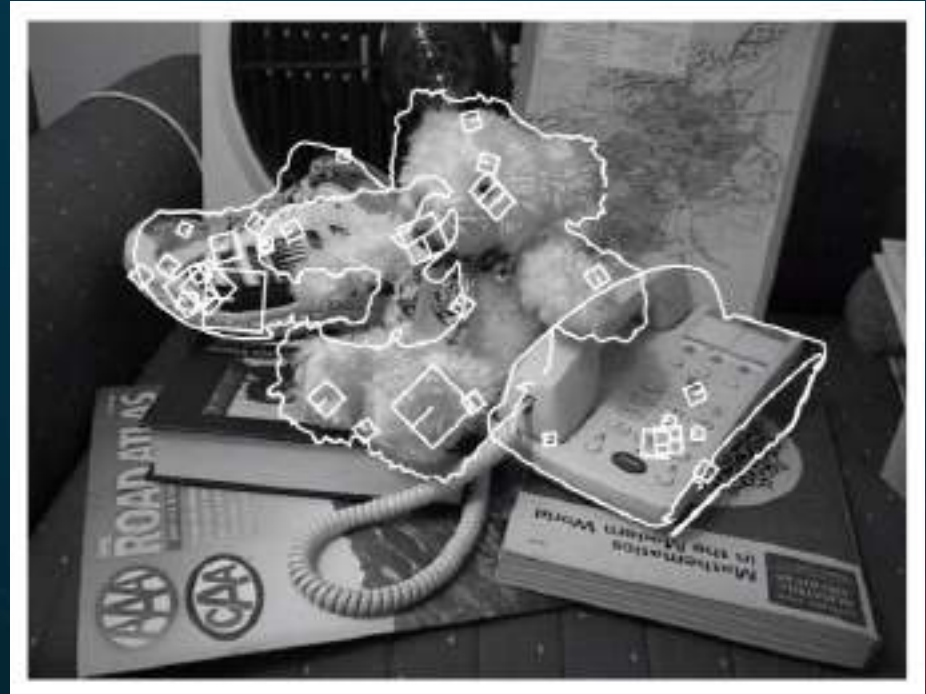
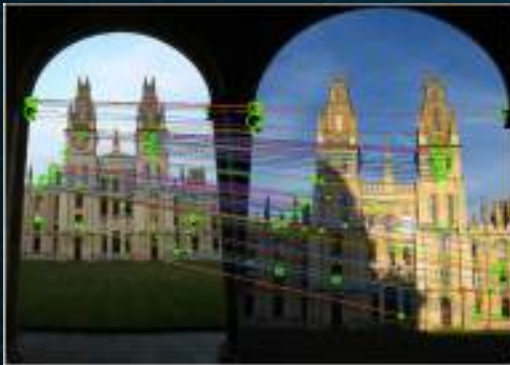
SIFT

Scale-Invariant Feature Transform

Noëline Migeon

SIFT

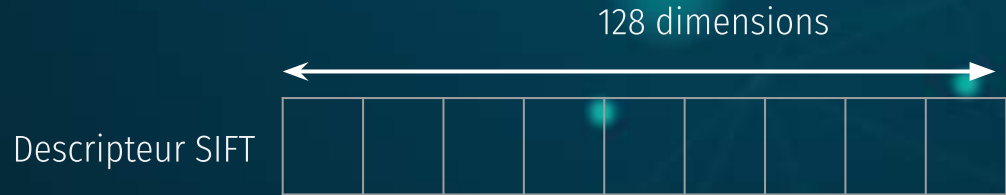
- David Lowe, 1999
- Reconnaissance d'objets
- Approche par descripteurs
- Approche locale : extraction de caractéristiques **stables**



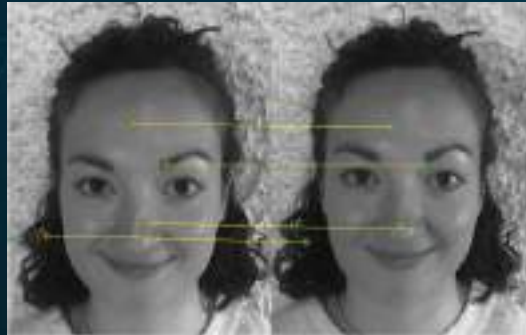
1. Détection



2. Description



3. Matching



SIFT - Principe

Image
d'entrée



1 point d'intérêt
=
1 descripteur SIFT
=
1 vecteur de 128 dimensions

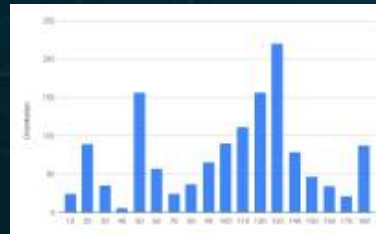
Détection

Détection des
points d'intérêts

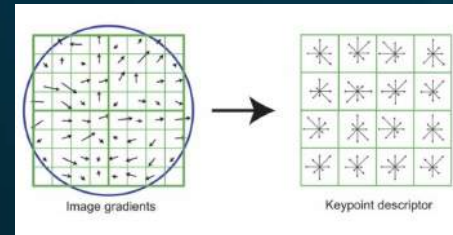


Description

Attribution d'une
orientation à
chaque point
d'intérêt

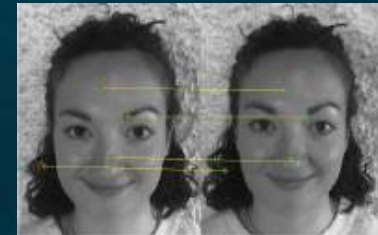


Calcul du
descripteur pour
chaque point
d'intérêt

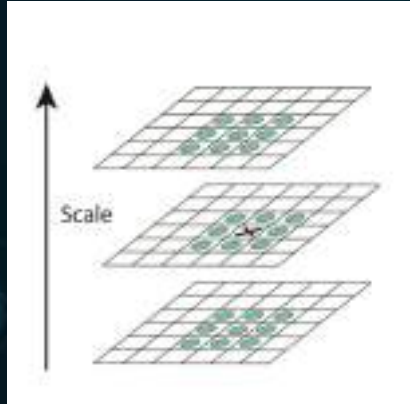


Matching

Mise en
correspondance
avec d'autres
images

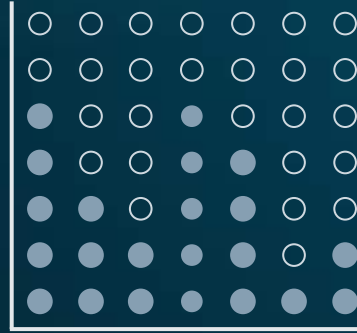


Détection des points d'intérêt



Recherche d'extrema locaux à
plusieurs échelles

→
Sélection des variations
les plus significatives



→
Obtention des points d'intérêts

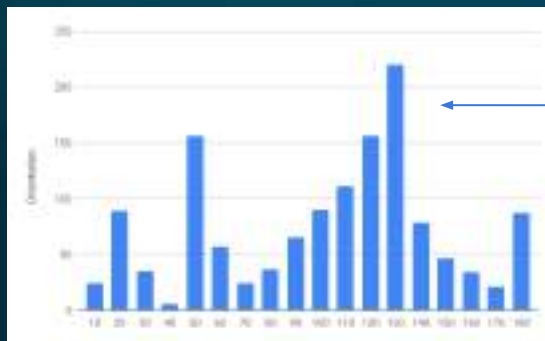


→
Visualisation

Description

1. Assignment d'orientation

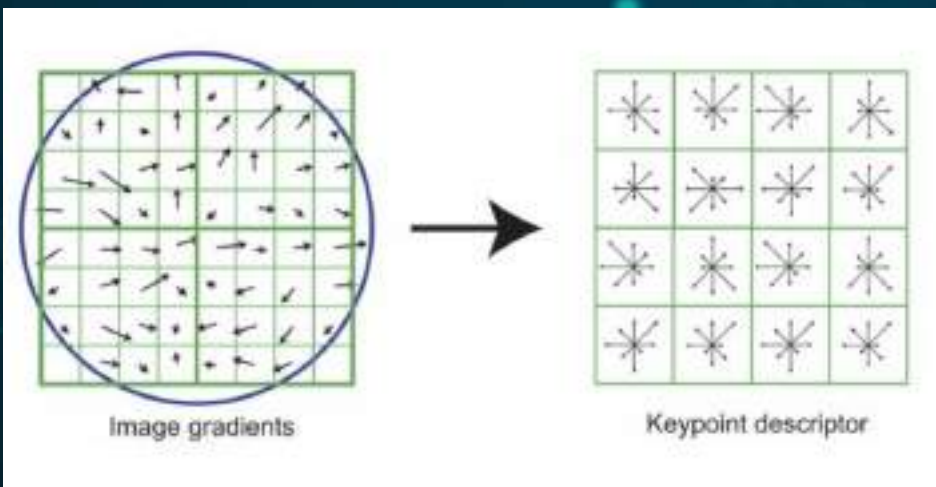
Orientation de chaque pixel
=
direction du gradient



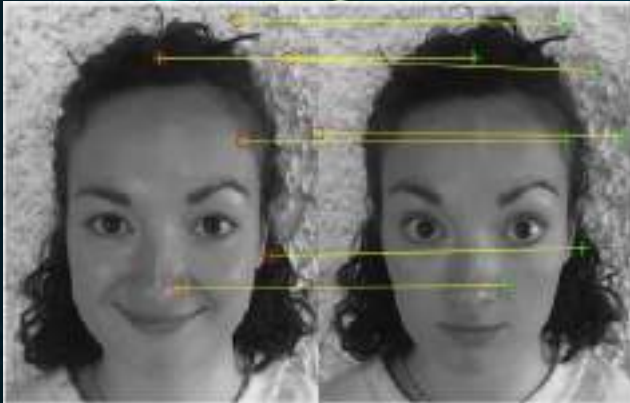
“Pic” de l’histogramme :
Orientation principale du
point d’intérêt

2. Calcul du descripteur

16 sous-régions * 8 orientations =
vecteur de 128 dimensions



Mise en correspondance



Mise en correspondance des descripteurs
“proches” (distance euclidienne)



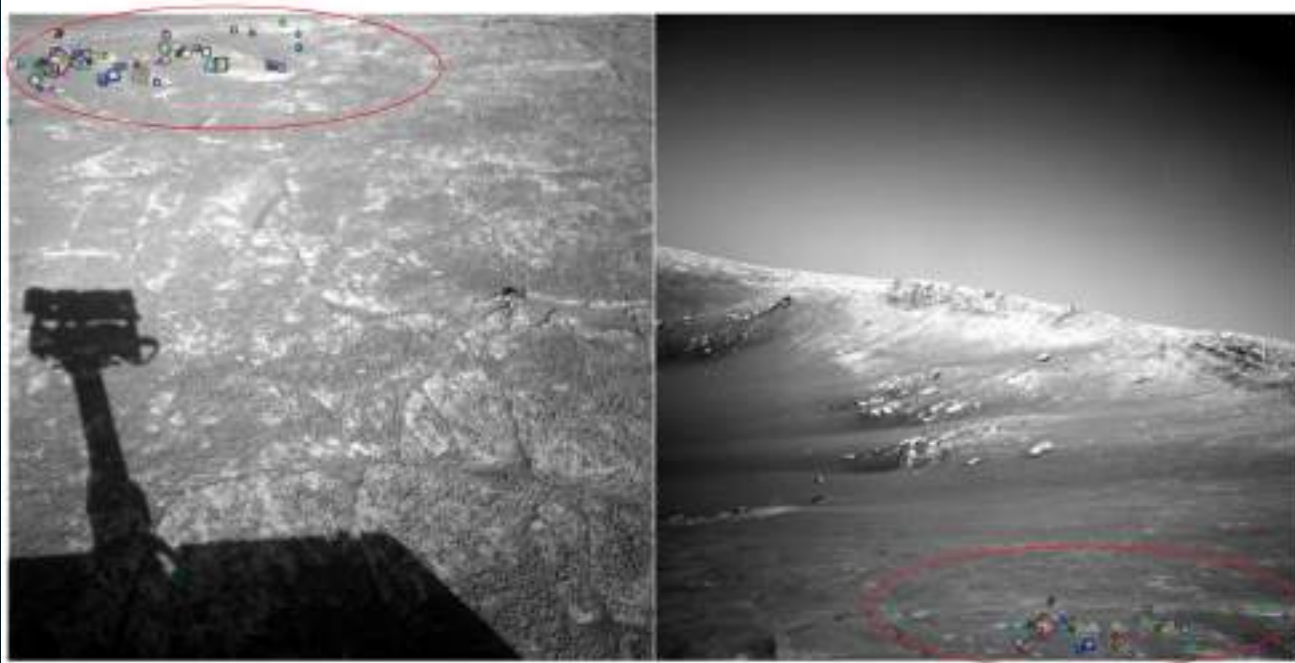
Mise en correspondance de points
d'intérêt après rotation de l'image

Résultats SURF

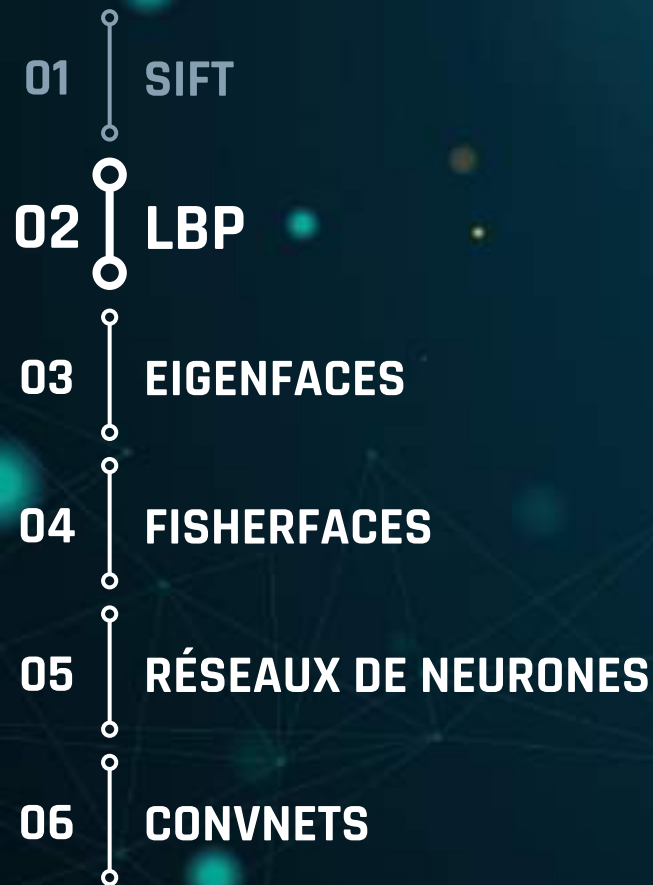


Détection d'objets

Utilisations



NASA Mars Rover images
with SIFT feature matches
Figure by Noah Snavely

- 
- 01 SIFT
 - 02 **LBP**
 - 03 EIGENFACES
 - 04 FISHERFACES
 - 05 RÉSEAUX DE NEURONES
 - 06 CONVNETS

LBP

Local Binary Pattern

Pénélope Roullot

Le descripteur LBP

Opérateur
LBP



67	153	102
0	67	183
183	102	153

seuillage

1	1	1
0	1	1
1	1	1

$$(01111111)_2 = (127)_{10}$$

Calcul du
descripteur

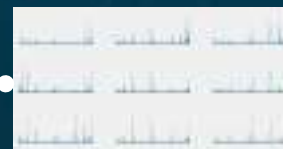


Histogramme spatial

Détection de visages



Parcours de l'image à différentes échelles



Modèle de visage



Modèle de non-visage



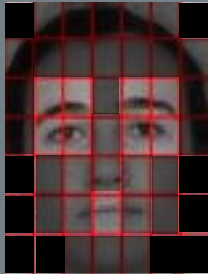
Comparaison des histogrammes

Calcul du descripteur LBP pour chaque imagerie



Améliorations apportées

- D'après Ahonen et al. [2] : **pondération des régions** en fonction des points d'intérêts **(a)**
- **Opérateur LBP** : Inversement de l'ordre de concaténation des pixels
- Comparaison des histogrammes : ajout d'un **critère plus discriminant** (pourcentage de différence)
- **Regroupement des** multiples **rectangles** à différentes échelles encadrant un seul visage (moyennage des coordonnées, avec une pondération en fonction de leur importance et de l'échelle) **(b)**
- **Détection des yeux** dans chaque visage détecté et **recadrage** du visage en fonction de la position des yeux, pour améliorer la reconnaissance **(c)**
- Amélioration de la **vitesse d'exécution** du code



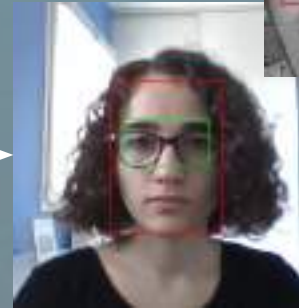
(a)



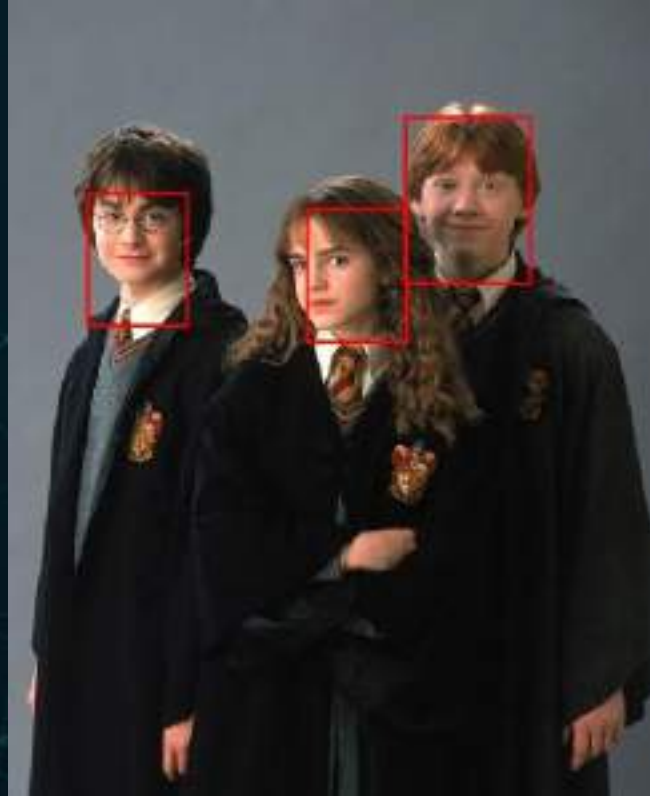
(b)



(c)



Résultats



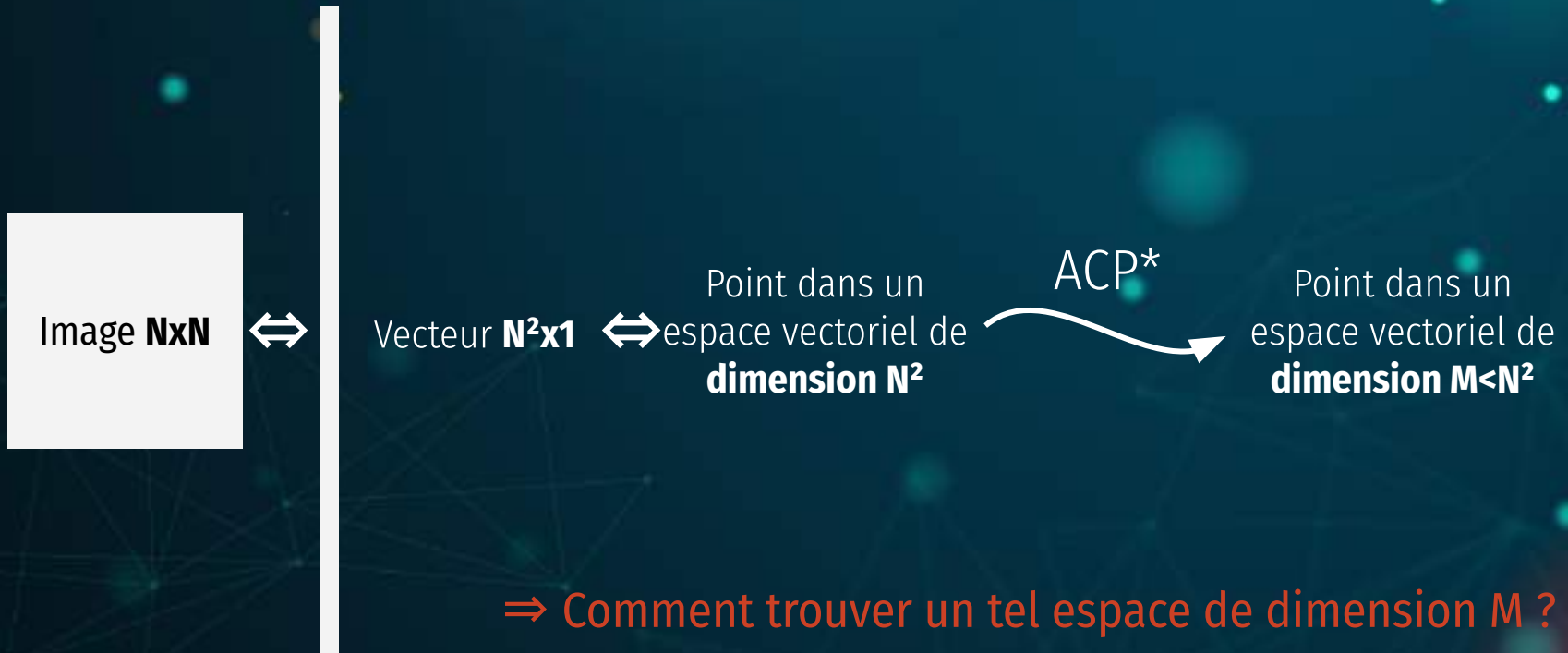
- 01 SIFT
- 02 LBP
- 03 **EIGENFACES**
- 04 FISHERFACES
- 05 RÉSEAUX DE NEURONES
- 06 CONVNETS



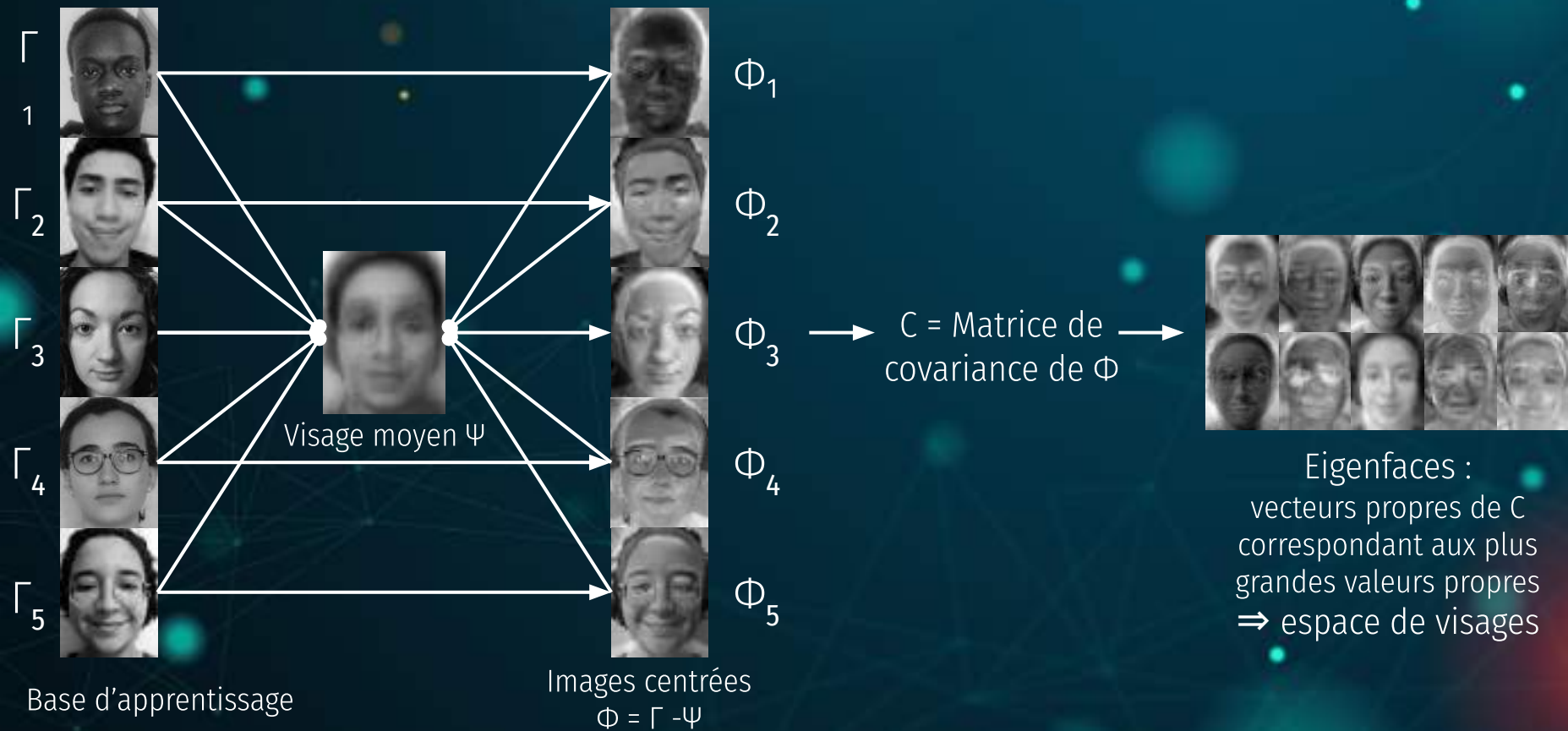
Eigenfaces

Pénélope Roullot

ACP* - Eigenfaces



Calcul des eigenfaces

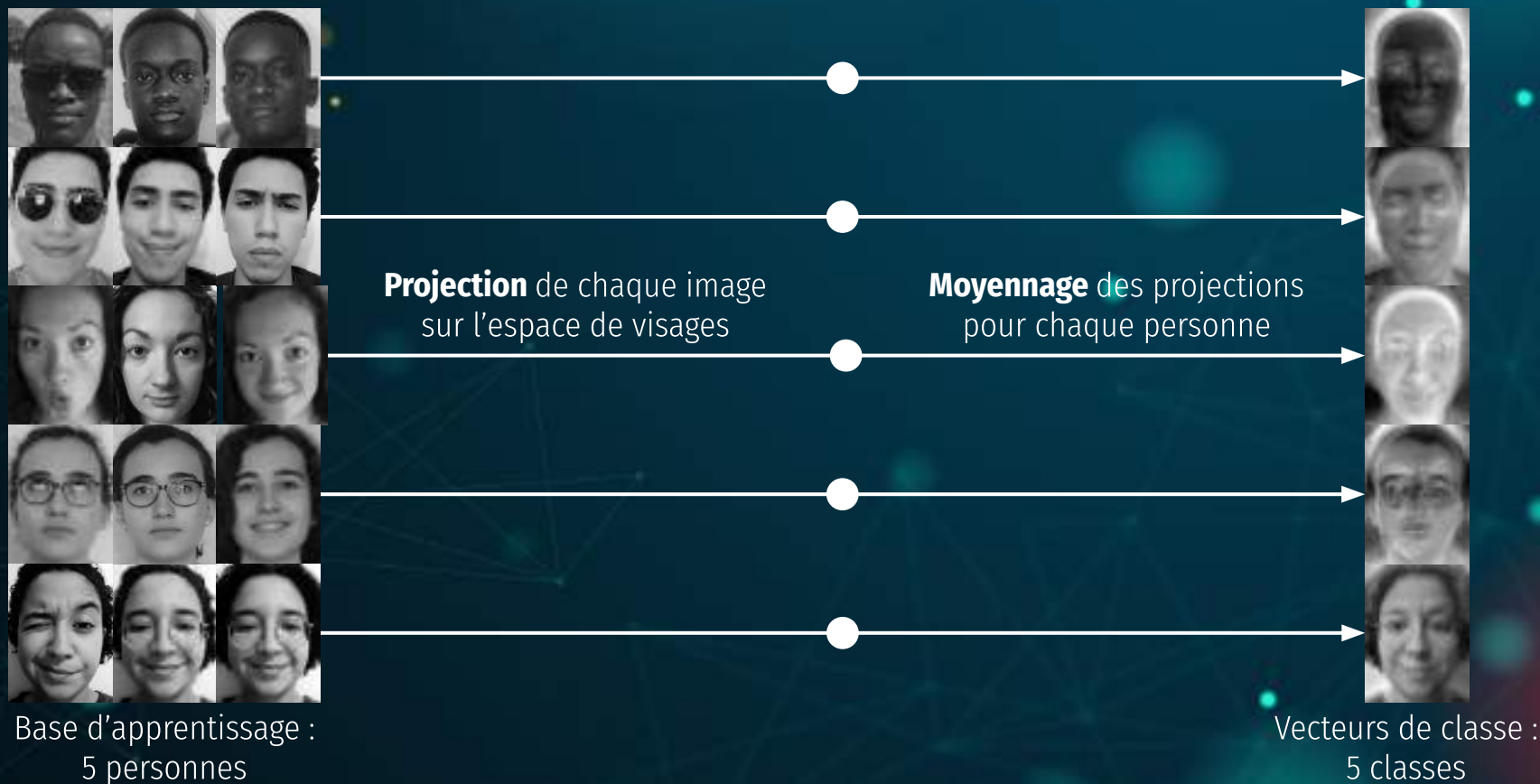


Projection d'un visage sur l'espace de visages


$$\text{Face} = \omega_1^* \text{Face}_1 + \omega_2^* \text{Face}_2 + \omega_3^* \text{Face}_3 + \dots + \omega_M^* \text{Face}_M$$

$$\Omega = [\omega_1 \ \omega_2 \ \omega_3 \ \dots \ \omega_M]$$

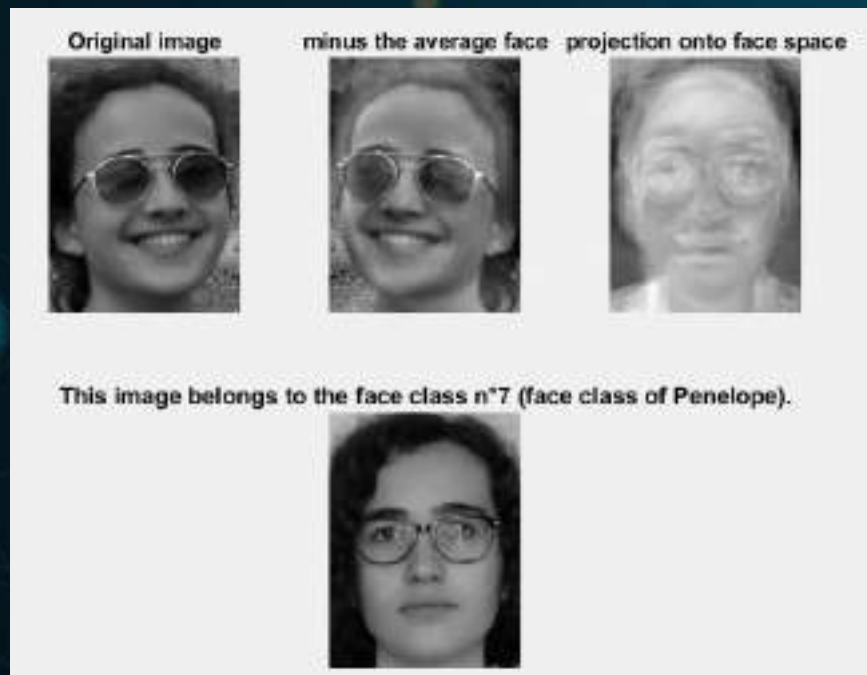
Apprentissage : Construction des classes



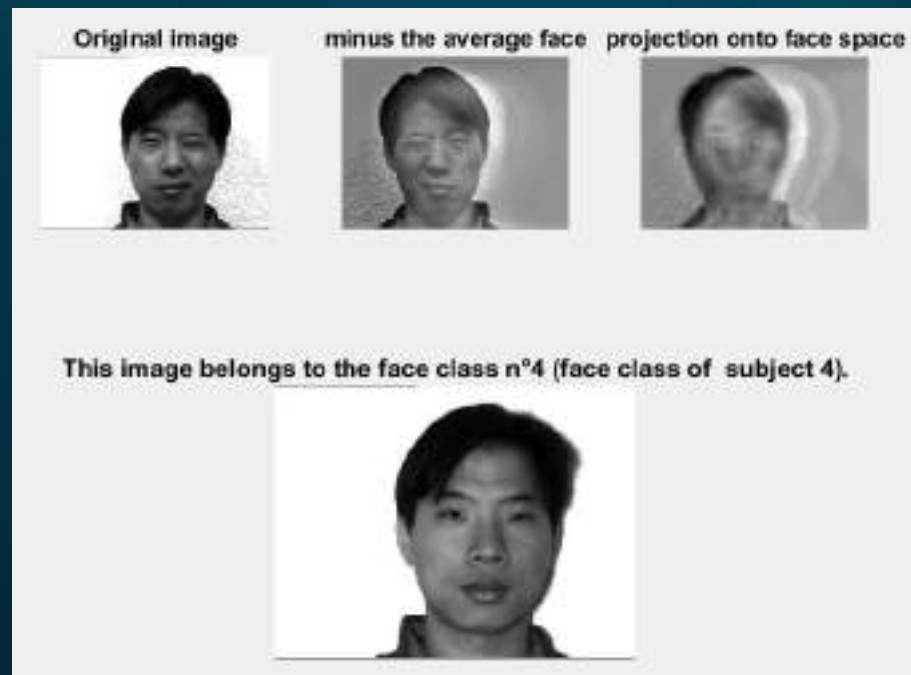
Reconnaissance de visage



Résultats



Base personnelle (20 images de 10 personnes)



Yale Database (11 images de 15 personnes)
(<http://vision.ucsd.edu/content/vale-face-database>)

Résultats



Base d'apprentissage : 250 images

Base de test : 250 images

-

Pourcentage de reconnaissance : 80%

Reconnaissances erronées : 12,8%

Non-reconnaissances : 7,2%

- 01 SIFT
- 02 LBP
- 03 EIGENFACES
- 04 FISHERFACES**
- 05 RÉSEAUX DE NEURONES
- 06 CONVNETS

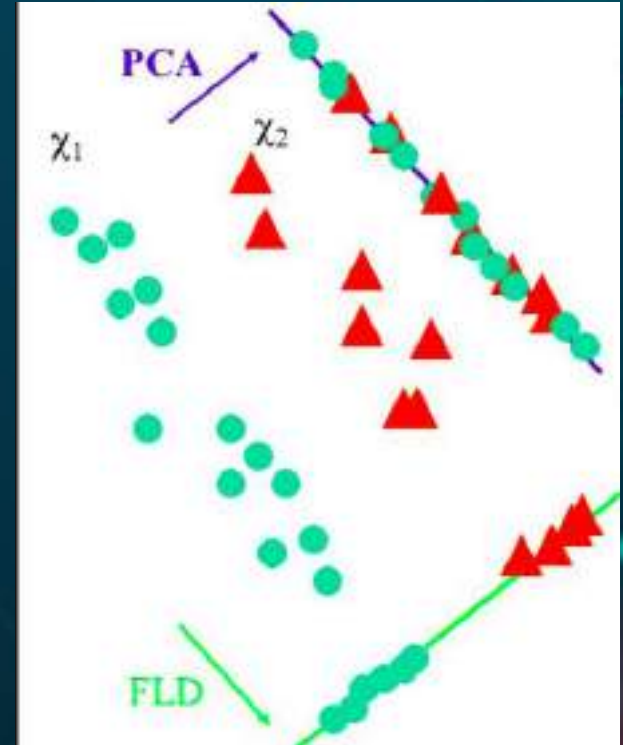


Fisherfaces

Mokhtar El Bouraqui

ACP vs ADL*

- ACP maximise la variance totale des données
- ADL assure la séparation entre les classes



Calculer les Fisherfaces

- Matrice de variance **interclasse** :

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T$$

- Matrice de variance **intra-classe** :

$$S_i = \sum_{x_k \in \gamma_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad S_W = \sum_{i=1}^c S_i$$

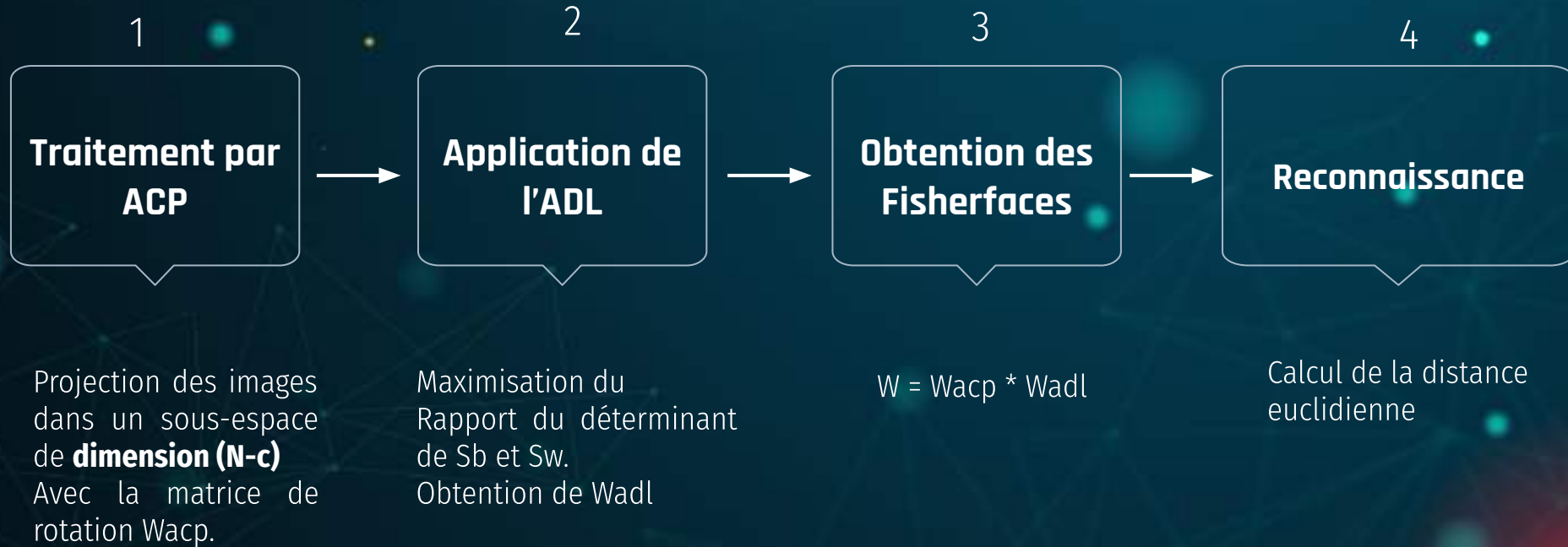
- Maximiser :

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|}$$

Problème

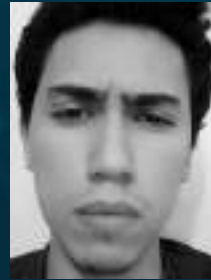
S_W est **toujours** singulière

Solution des Fisherfaces



Base de données (Personnelle)

- Input : 100 images de 5 personnes
- Train : 50 images
- Test : 50 images



Base de données (Yale)

- Input : 150 images de 15 personnes
- Train : 149 images
- Test : 1 image



Avec
lunettes

Sans
lunettes

3 conditions d'éclairage

5 expressions faciales

Résultats




Résultats



Base d'apprentissage : Yale Database
Pourcentage de reconnaissance: 93%

-

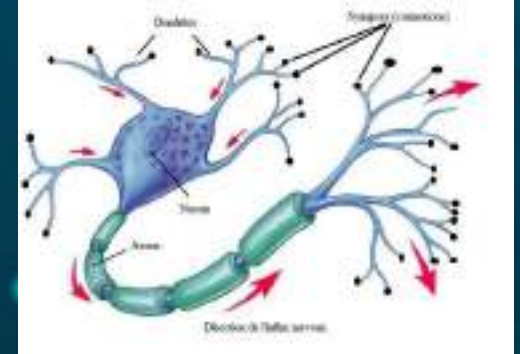
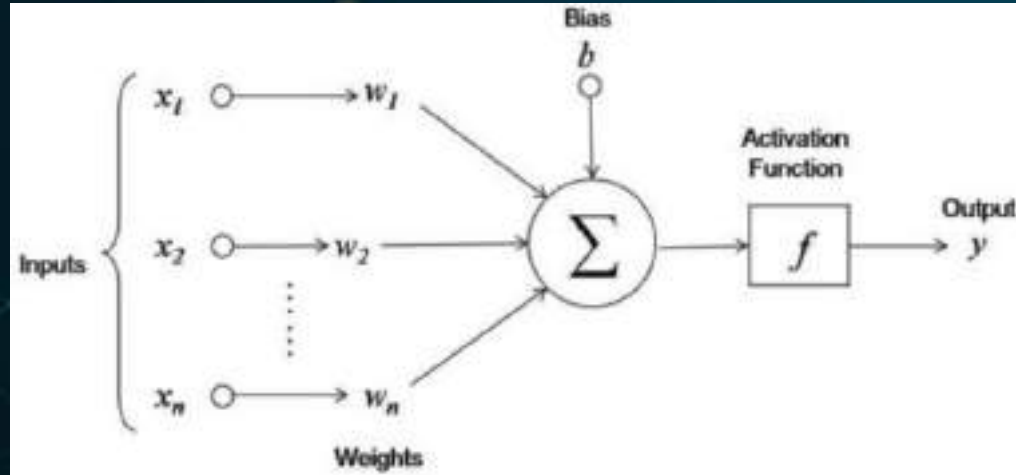
Base d'apprentissage : Base personnelle
Pourcentage de reconnaissance : 60%

- 
- 01 SIFT
 - 02 LBP
 - 03 EIGENFACES
 - 04 FISHERFACES
 - 05 RÉSEAUX DE NEURONES**
 - 06 CONVNETS

Réseaux de neurones

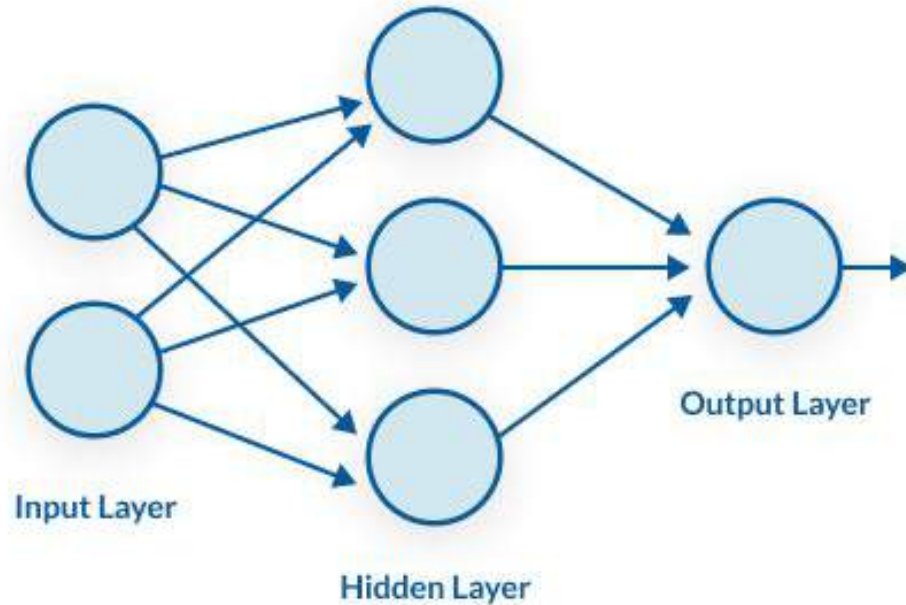
Zaineb Harrizi

Modélisation mathématique d'un neurone

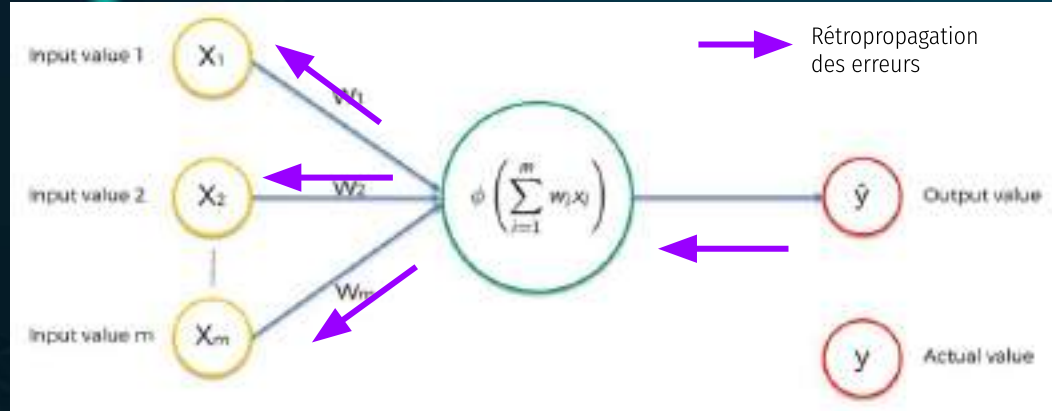


- x_i : les données
- w_i : les poids associés à chaque entrée
- b : biais
- f : fonction d'activation du neurone
- y : la sortie

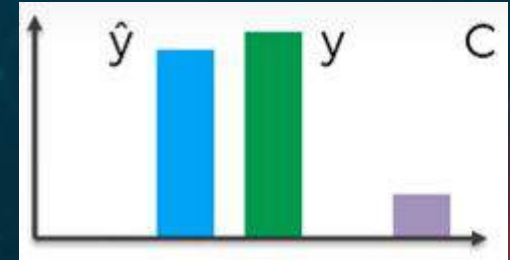
Perceptron multicouche



Phase d'apprentissage

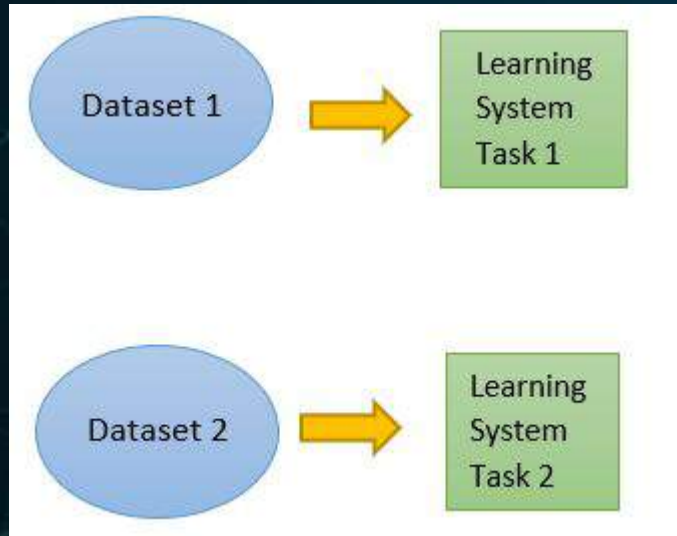


$$C = \frac{1}{2} (\hat{y} - y)^2$$

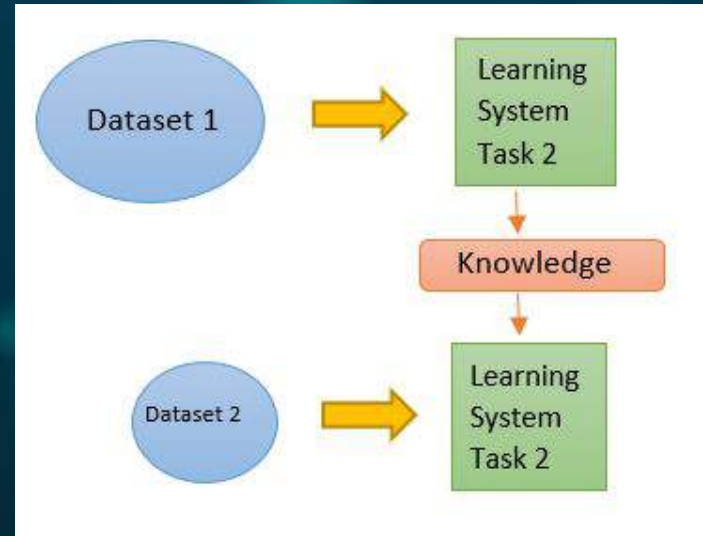


Transfer Learning

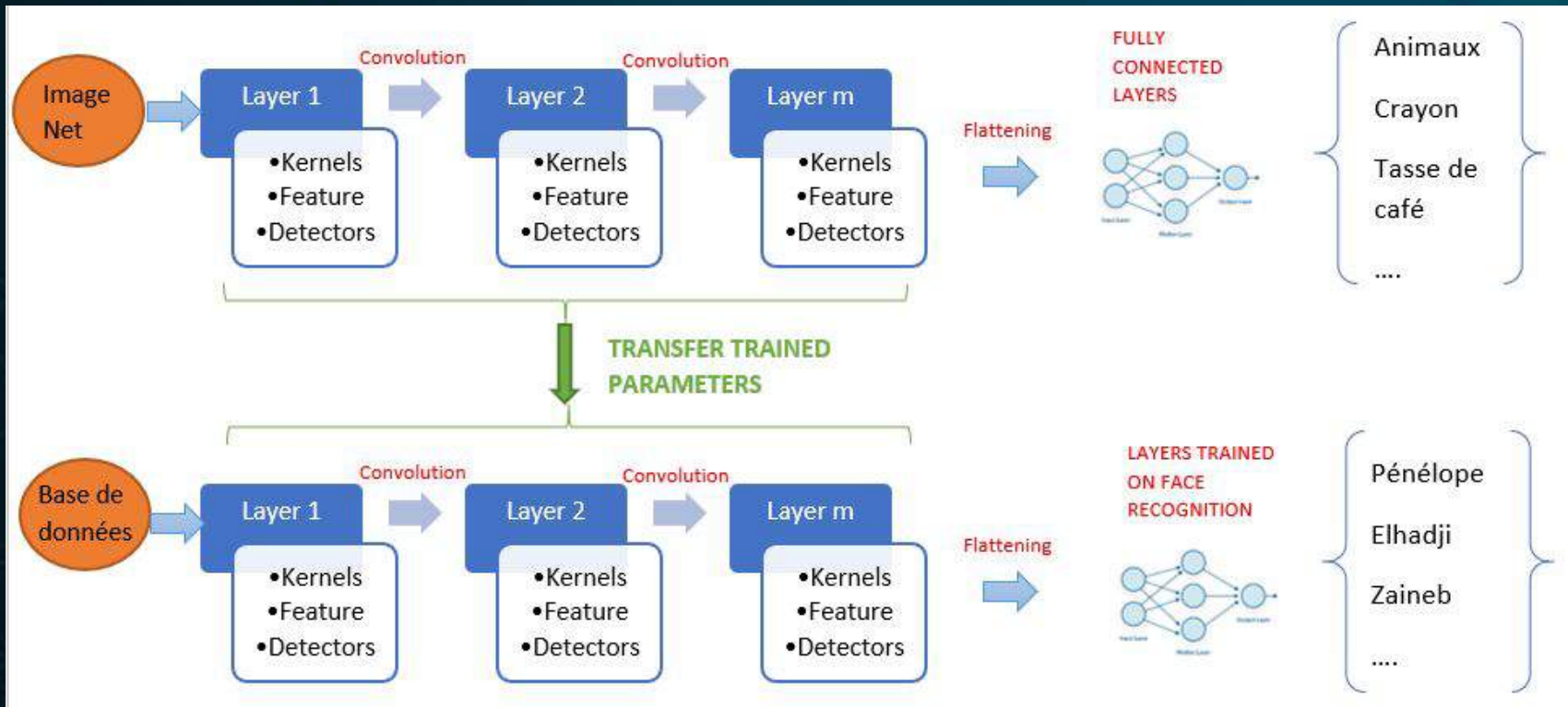
Traditional Learning



Transfer Learning



AlexNet



Base de données

- 9 classes
- 123 images
- Training set : 86 images
- Validation set : 37 images



Résultats



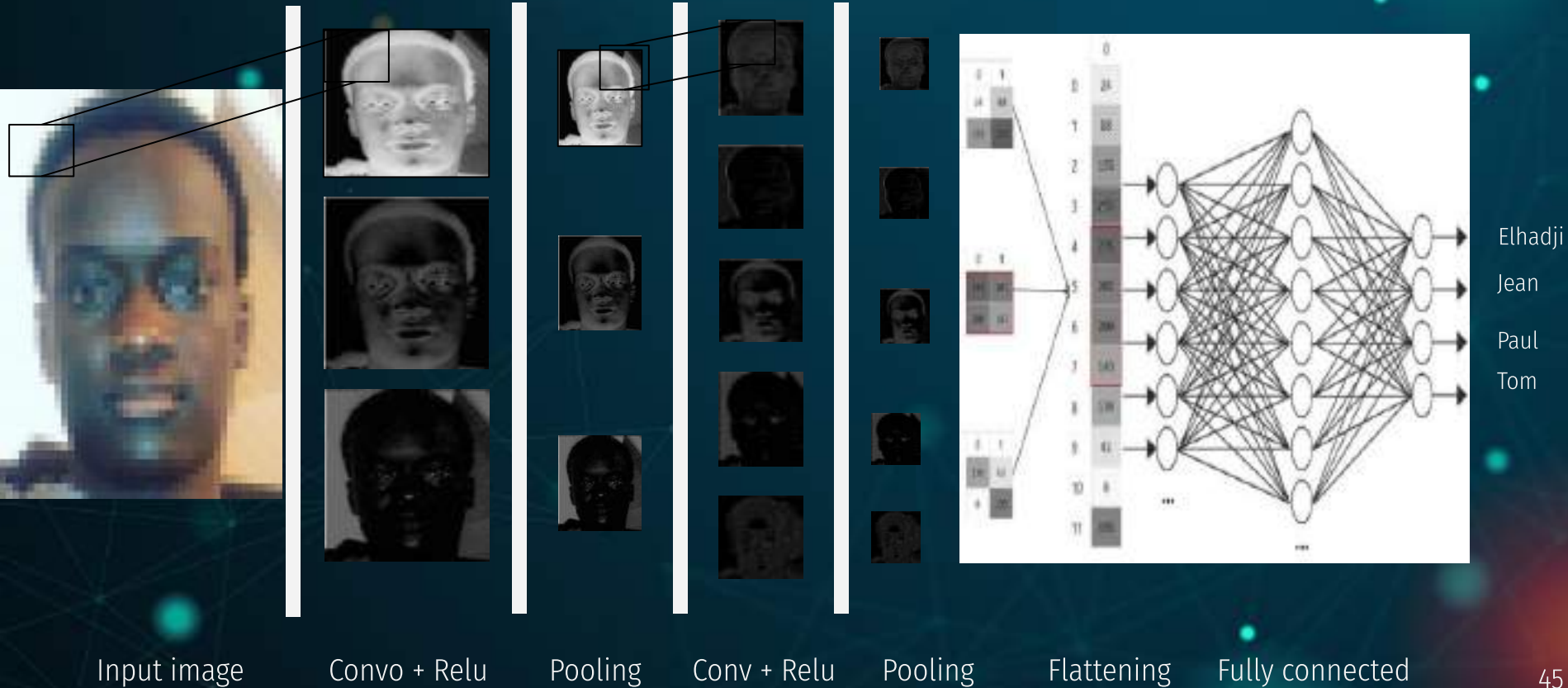
Pourcentage de reconnaissance :
87,65%

- 
- 01 SIFT
 - 02 LBP
 - 03 EIGENFACES
 - 04 FISHERFACES
 - 05 RÉSEAUX DE NEURONES
 - 06 **CONVNETS**

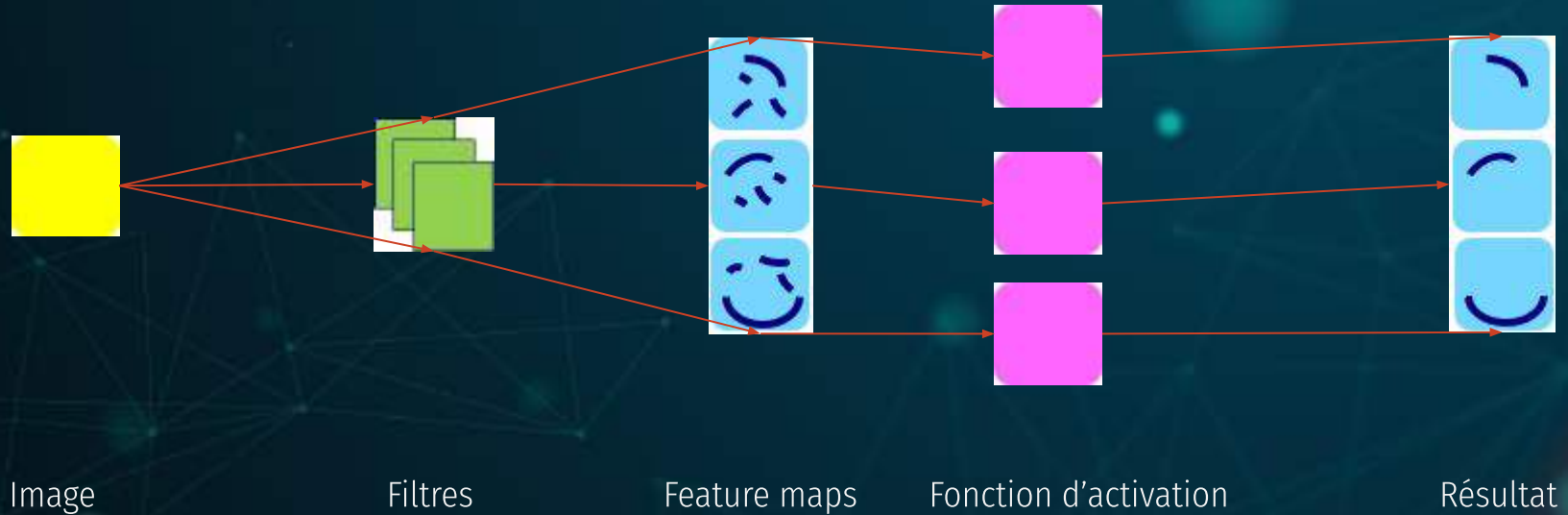
Convnets

Elhadji Niang

Architecture des Convnets



Extraction de caractéristiques

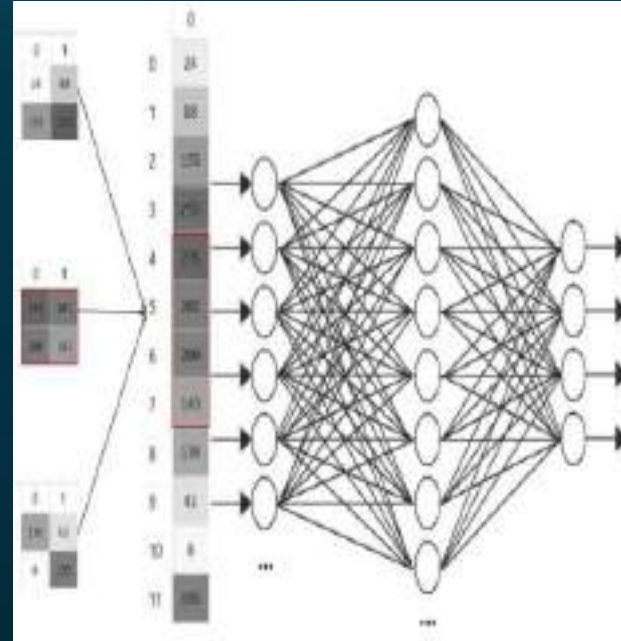


Apprentissage et Prédiction de la classe de l'image

Activité d'apprentissage



Partie Convolution



0.32 = Elhadji

0.53 = Jean

0.08 = Paul

0.07 = Tom

Rétropropagation des erreurs

Bases de données d'images utilisées

- Face96

2520 images , 125 classes. 20 images par classes.
1764 images pour l'apprentissage et 756 images pour le test



- Images_PIR

81 images, 5 classes. 57 images pour l'apprentissage et 24 images pour le test



Résultats

- Pooling

Max pooling

94	30
98	36

VS

94	20	30	22
8	12	2	0
34	54	24	8
70	98	12	36

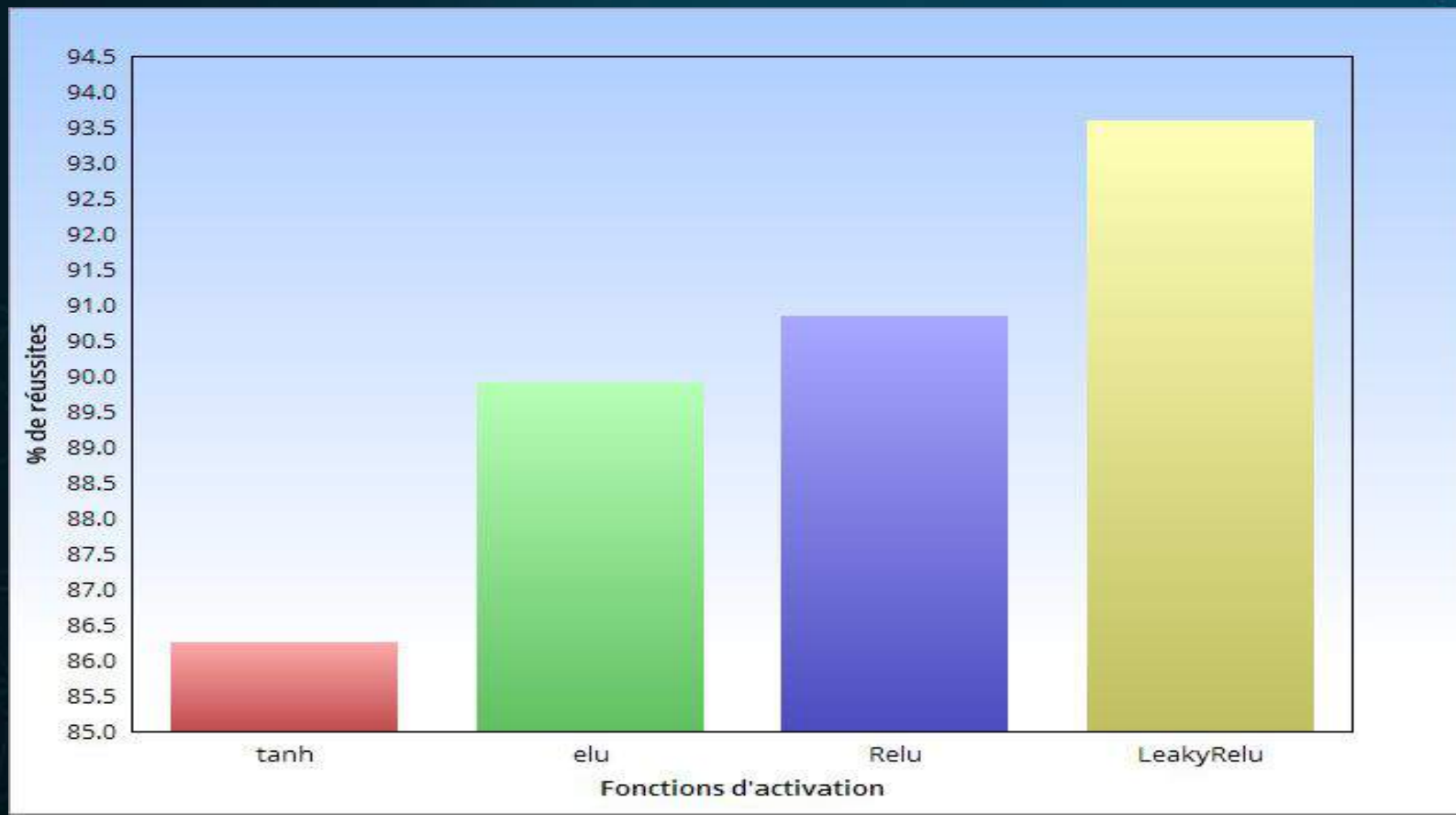
Average pooling

33.5	13.5
64	20



- Fonctions d'activations

4 fonctions d'activations testées



- **Pourcentage de réussites**

Avec la base face96

86.73 % de précision pendant l'apprentissage

13.27 % d'erreur

Avec la base Image_PIR

83.2 % de précision pendant l'apprentissage

16.8% d'erreur



Conclusion

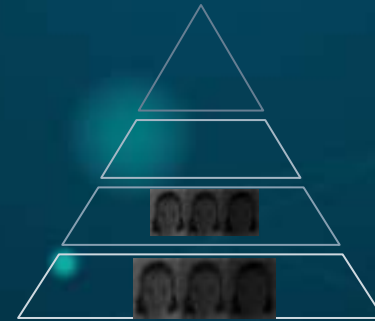
CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Bibliographie

- [1] D. G. Lowe, « Object recognition from local scale-invariant features », in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999, p. 1150-1157 vol.2, doi: 10.1109/ICCV.1999.790410.
- [2] T. Ahonen, A. Hadid, et M. Pietikainen, « Face Description with Local Binary Patterns: Application to Face Recognition », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, n° 12, p. 2037-2041, déc. 2006, doi: 10.1109/TPAMI.2006.244.
- [3] L. S. López, « Local Binary Patterns applied to Face Detection and Recognition », Universitat Politecnica de Catalunya, Barcelona, 2010.
- [4] M. Turk et A. Pentland, « Eigenfaces for Recognition », *Journal of Cognitive Neuroscience*, vol. 3, n° 1, p. 71-86, janv. 1991, doi: 10.1162/jocn.1991.3.1.71.
- [5] P. N. Belhumeur, J. P. Hespanha, et D. J. Kriegman, « Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection », *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 19, n° 7, p. 10, 1997.
- [7] Y. LeCun, « Convolutional networks for images, speech, and time series », *The handbook of brain theory and neural networks*, Consulté le: juin 17, 2020. [En ligne]. Disponible sur: https://www.academia.edu/2813453/Convolutional_networks_for_images_speech_and_time_series.
- [8] F. Mamalet, S. Roux, et C. Garcia, « Embedded facial image processing with Convolutional Neural Networks », in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, Paris, France, mai 2010, p. 261-264, doi: 10.1109/ISCAS.2010.5537897.
- [9] AMIDI Afshine, AMIDI Shervine. *Pense-bête de réseaux de neurones convolutionnels*. Stanford University [en ligne]. Disponible sur : stanford.edu/~shervine/l/fr/teaching/cs-230/pense-bete-reseaux-neurones-convolutionnels
- [10] SPACEK, Libror. *Face Recognition Data*. University of Essex [en ligne]. Disponible sur : cswww.essex.ac.uk/mv/allfaces/
- [11] Moodle INSA Lyon : Cours : Télécommunications/TC-4/Traitement de signaux numériques : Parole et Image. Disponible sur : moodle.insa-lyon.fr/course/view.php?id=1222

Annexes

Construction de la pyramide "différence de gaussiennes"



Pyramide DoG

Soustraction

