

# La reconnaissance des visages : une comparaison entre les réseaux des neurones compétitifs et les réseaux des neurones à spike

## A Comparison between Artificial and Spike Neural Network for Face Recognition

**Adjoudj Réda<sup>1</sup>, Gafour Abdel-Kader<sup>1</sup>, Boukelif Aoued<sup>2</sup>,  
Lehireche Ahmed<sup>1</sup>**

<sup>1</sup>Laboratoire des Systèmes Distribués & de l'ingénierie évolutive, EEDIS, Département Informatique, Université de Sidi Bel Abbès, Algérie  
AdjReda@yahoo.fr, Gafour@yahoo.com, elhir@yahoo.com

<sup>2</sup>Laboratoire des Télécommunications et du Traitement Numérique de Signal, Équipe de recherche des techniques vidéo,  
Département Electronique, Université de Sidi Bel Abbès, Algérie  
aboukelif@yahoo.fr

Manuscrit reçu le 30 juillet 2007

### Résumé et mots clés

Il est souvent utile d'avoir une machine qui effectue la reconnaissance. En particulier, les machines qui peuvent reconnaître des images de visage sont très coûteuses. Effectivement, une machine qui contient un système qui peut traiter ou reconnaître plus d'image de visage qu'un opérateur humain. Ce type de système fait gagner du temps et de l'argent, et élimine les tâches répétitives pour l'opérateur humain. Cet article montre comment un système de reconnaissance des visages peut être réalisé par un réseau de neurones artificiel de type perception multicouche et par un réseau de neurones à spike. Le système à base de spike est développé pour acquérir les importantes caractéristiques du visage, pour simuler le système de la vision humaine et pour optimiser le temps de calcul, ce dernier objectif c'est la principale force derrière le développement des systèmes à base des réseaux de neurones à spike. Avec ce derniers système, les meilleurs résultats de reconnaissance ont été obtenus et un gain du temps considérable d'apprentissage a été enregistré. Dans ce cas-ci, le système à base de réseau neurones à spike a enregistré un taux de reconnaissance de 92.10 %, 95.00 % et de 93,10 % avec un temps maximum d'apprentissage de 150 ms (milliseconde) pour les trois bases de données utilisées lors des tests, comparé au premier système, où le taux de reconnaissance est de 71.05 %, 77.50 % et de 73.89 %, avec un temps maximum d'apprentissage de 2 heures 17min 45s. Cela fait, que le système de reconnaissance à base des spikes est meilleur. A noter que l'apprentissage du réseau sur différents ensembles d'images lui force à apprendre comment il se comporte vis-à-vis la variété des visages, un problème commun dans le monde réel...

**Biométrie, reconnaissance des formes, réseau de neurones artificiel, réseau de neurones à spike, traitement d'image.**

### Abstract and key words

Face recognition is the process of the automatic recognition of the person's identity based on individual informations that are included in face image. This technique makes the face recognition possible, to use the person's image, to verify their identity and control access to services such as e-commerce, internet access, physical access control, cellular phones, border control, passport control, database access services, information services, security control for confidential

information areas, and remote access to computers. This document demonstrates how a face recognition system can be designed by a conventional artificial neural network and by another more recent neural network, which is called Spike neural network. The latter is developed to capture the important characteristics of the face, to simulate the human visual system and to optimise the computational time, this last characteristic has been one of the driving forces behind the development of spike neural networks. This kind of neural networks has scored a perfect recognition rate of 92.10 %, 95.00 % and 93.10 %, with a maximum training time of 150ms for the three databases, compared to the case when the artificial neural network is applied, where the rate was 71.05 %, 77.50 % and 73.89 %, with a maximum training time of 2 hours 17 min 45 s. Note that the training process of the two networks, on different sets of noisy images, forced the two networks to learn how to deal. In this case, a common problem in the real world.

Biometric, Pattern Recognition, Artificial Neural Network, Spike Neural Network, Image processing.

# 1. Introduction

**S** La plupart des contrôles d'accès automatiques actuels utilisent un code secret ou une clé pour valider l'identité de la personne désirant accéder à un lieu, un compte, un ordinateur, etc. On utilise un numéro d'identification personnel pour valider l'identité au guichet automatique. On utilise une clé pour entrer chez soi. Ces deux techniques de contrôle d'accès partagent les mêmes failles : le test de contrôle est duplicable et indépendant de son utilisateur. Quelque malfrat possédant votre numéro d'identification personnel et votre carte de guichet automatique peut vider vos comptes bancaires à sa discrétion. De même, une serrure n'est qu'une invitation pour un cambrioleur qui en possède la clé.

Ces problèmes ne surviennent pas lorsque le contrôle d'accès est lié directement à l'identité d'une personne. Les systèmes d'identification basés sur des caractéristiques personnelles, aussi appelés biométriques, utilisent des aspects propres à chaque individu pour en valider l'identité. Ces caractéristiques biométriques peuvent être les empreintes digitales, l'identification vocale, l'empreinte rétinienne, identification faciale ou encore l'empreinte génétique...

Nous avons choisi la reconnaissance des visages comme moyen d'identification, par rapport aux autres méthodes, pour différentes raisons :

- Cette identification est naturellement utilisée par une personne normale ; on n'a généralement qu'à voir quelqu'un pour savoir qui la personne est.
- Cette méthode ne demande pas d'équipement d'acquisition très complexe ; une simple caméra peut suffire.
- Ce type de reconnaissance peut éventuellement s'appliquer à la localisation d'un individu dans une foule, contrairement aux autres méthodes.
- Cette approche recevant une information de nature optique, les techniques de traitement optique de l'information présentent un intérêt.

Nous prendrons le temps ici, de bien établir le contexte d'utilisation du système de reconnaissance de visages que nous voulons utiliser, car le contexte établira les limites de notre travail. Le système de reconnaissance se veut être un contrôle d'accès pour un site ou une information privilégiée. On considère, par exemple, une personne se présentant à un guichet automatique ou à un contrôle de sécurité d'un local. Dans ce contexte, on peut faire certaines suppositions sur le visage de la personne qui se présente au système de reconnaissance :

- La personne sera toujours sensiblement placée à la même distance de la caméra prenant l'image du visage. On peut donc présumer que l'image du visage aura toujours la même taille. Le système de reconnaissance n'aura donc pas à être robuste aux changements d'échelle.
- La personne devrait être placée sensiblement au même endroit dans l'image prise par la caméra. Le système peut donc se contenter d'une invariance sous translation limitée.
- On peut présumer également que la personne fera face à la caméra. Nous nous restreindrons donc à produire un système supportant des rotations de la tête inférieures à 15 degrés de chaque côté d'une vue de face dans un plan horizontal et 5 degrés, dans un plan vertical.
- Le système devra avoir une certaine robustesse à des modifications du visage comme le port de lunettes, un sourire, les yeux fermés, etc.

Différentes méthodes de reconnaissance de visage existent et plusieurs logiciels commerciaux ont vu le jour pour effectuer ce type de reconnaissance. Les compagnies Visionics Corporation, Miro, Siemens Nixdorf, Visionsphere et Polaroid Corp produisent de tels systèmes.

Dans les logiciels commerciaux actuels, deux méthodes de reconnaissance sont principalement utilisées : l'analyse des composantes principales (l'ACP) et l'analyse des caractéristiques locales. On trouve aussi dans la littérature d'autres méthodes d'identification de visages utilisant des réseaux de neurones ainsi que la corrélation optique comme mesures de similitude. Une approche que nous a particulièrement intéressé

pour l'identification de visages [1] : c'est à base d'un corrélateur à transformées conjointes non-linéaire et un réseau de neurones pour effectuer l'authentification des visages.

Les travaux qui sont présentés dans cet article forment une vérification, un approfondissement et une synthèse des méthodes de reconnaissance des visages proposées et développées par nous même dans notre labo<sup>1</sup> depuis 2002, et qui ont fait l'objet de contributions.

Nous verrons premièrement le principe de fonctionnement du réseau de neurones, qui est le coeur du système de reconnaissance présenté. Nous étudierons par la suite un nouveau concept du réseau de neurones de type Spike ou à potentiel d'action asynchrone, qui forme l'outil d'optimisation des performances du réseau de neurones à perceptron [10, 9, 1].

Nous examinerons par la suite les deux systèmes de reconnaissance obtenus par implémentation de ces deux technologies. Nous parlerons ensuite de la base de données des visages collectée pour la réalisation de cette méthode et des méthodes d'acquisition prises. Nous serons à ce moment rendus à présenter les résultats obtenus grâce au système de reconnaissance décrit. Une fois les résultats présentés et analysés, un résumé des travaux effectués et des résultats obtenus conclura ce document.

### 1.1. Comment la Reconnaissance par le Visage Fonctionne

La reconnaissance de visage peut être utilisée en mode identification (un-à-plusieurs) ou en mode vérification (un-à-un). L'identification de visage est le processus qui détermine si la personne présente et qui a fournit une expression donnée est déjà enregistrée ou non. Par contre, la vérification de visage est le processus d'accepter ou de rejeter l'identité proclamée d'une personne [21, 17, 10, 18].

Les technologies de la reconnaissance de visage, soit l'identification et ou la vérification, chacune a ses propres avantages et inconvénients et elles exigent différents traitements, prétraitement et techniques de transformation. Le choix de quelle technologie à employer est spécifique à l'application [9, 26, 10, 15].

### 1.2. Description

Au niveau le plus élevé, tous les systèmes de reconnaissance des visages contiennent deux principaux modules [21, 24, 29, 17, 22] : extraction des caractéristiques et comparaison des caractéristiques. L'extraction des caractéristiques est le processus qui extrait les informations utiles à partir de l'image capturée du visage, ces informations peut être ensuite employées pour représenter chaque personne. La comparaison des caractéristiques est le procédé réel qui identifié la personne inconnue en comparant ses caractéristiques extraites de sa photo d'input à celle de l'en-

semble de personnes connues. Nous discuterons chaque module en détail dans les sections suivantes. Bien que la reconnaissance des visages semble être une méthode d'authentification facile à mettre en application et en utilisation, mais l'efficacité de cette technologie est fortement influencée par des facteurs environnementaux, et spécialement par les conditions d'éclairage [23]. Les variations de l'appareil photo et la position du visage, l'expression et les caractéristiques faciales (coup du cheveux, lunettes de soleil, barbes) tous affectent la performance du système de reconnaissance. L'alignement précis de l'image sur les positions des yeux est nécessaire pour le bon déroulement des algorithmes de reconnaissances des visages. En conséquence, la technologie courante de la reconnaissance faciale est la plus efficace lorsque l'éclairage utilisé est conforme et les sujets se montrent coopératifs en respectant la position adéquate, où des chapeaux et les lunettes de soleil sont enlevés et chacun regarde directement vers l'appareil photo, en même temps.

Comme essai, nous avons téléchargé des images faciales de référence, et nous avons traité ces photos de façon numérique, ainsi la reconnaissance/comparaison est faite avec succès, pour le cas de notre base, nous avons accordé aux personnes inscrites de déplacer leur tête légèrement à gauche et à droite pour voir si le système sera dupé ou non, le même principe a été utilisé dans les travaux de [24, 12, 10, 23].

## 2. L'approche proposée et développée

La reconnaissance des visages fait partie du domaine de la reconnaissance des formes. Le but de la reconnaissance des formes est de classifier des objets d'intérêt dans un certain nombre de catégories ou de classes. Les objets d'intérêt sont appelés généralement les modèles ou patterns et dans notre cas ils sont des vecteurs de caractéristiques appelés les matrices de code, ces derniers sont extraits à partir des images de visage d'input en utilisant les techniques décrites dans les sections suivantes. Les classes ici représentent les différentes personnes. Puisque la procédure de classification dans notre cas sera appliquée sur des vecteurs de caractéristiques, elle peut être donc désigné par le nom de la comparaison de caractéristiques [11]. Notre approche démontre comment un système de reconnaissance des visages peut être conçu par un réseau de neurones artificiel conventionnel et par un autre réseau de neurones plus récent, qui est appelé un réseau de neurones à spike ou un réseau de neurones à potentiel d'action asynchrone, les deux réseaux de neurones sont employés en tant que processus de comparaison ou de reconnaissance. La figure 1 résume comment le système de reconnaissance proposé fonctionne.

1. Laboratoire des Systèmes Distribués & de l'ingénierie évolutive, EEDIS, Département Informatique, Université de Sidi Bel Abbès.

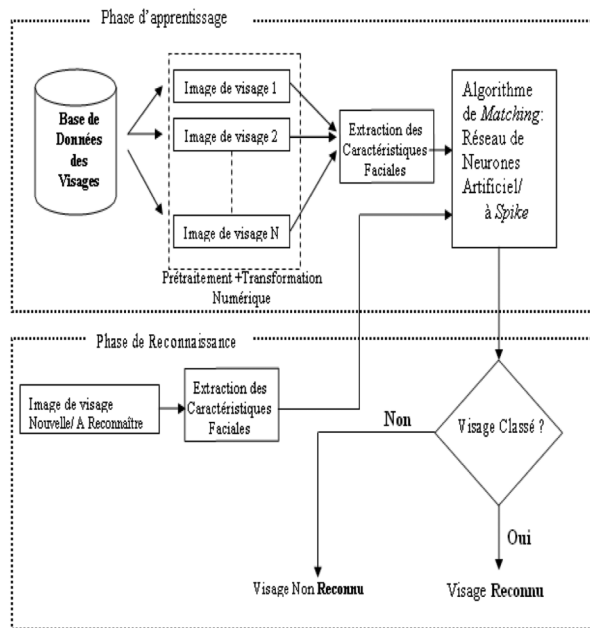


Figure 1. Système de Reconnaissance des visages basé sur le Réseau de Neurones Artificiel/à Spike.



### 3. Le problème posé

Un réseau de neurones artificiel et un réseau de neurones à spike sont désignés et réalisés pour reconnaître les personnes à partir d'une base de données. Ces deux présents systèmes doivent convertir chaque image de visage en un vecteur code. Le but de cette première phase est d'obtenir un vecteur de 960 éléments. (premièrement, la taille de l'image du visage initiale est de  $\sim 32 \times 30$  pixels) Figure 2. L'obtention d'un vecteur de 960 éléments est reflété par l'extraction des caractéristiques à partir des petites régions locales de l'image, cette étape est détaillée dans la section suivante.

### 4. La phase de reconnaissance des visages

La reconnaissance est le noyau de ce système, et c'est la comparaison du vecteur code du visage en input (établi dans l'étape précédente) avec ceux de la base de données, et partons du fait que nous voulons modéliser une fonction du cerveau humain et qu'on a un problème de classification, nous avons choisi d'implémenter un RN (Réseau des Neurones) qui est un simulateur du réseau de neurones biologique.

L'utilisation des RN en reconnaissance automatique des personnes est relativement récente, et reste en évolution continue vu les taux de reconnaissance obtenus qui non pas encore satisfait les chercheurs, ces taux sont souvent calculés sur des ensembles de données fermés en vu de la difficulté d'avoir des résultats satisfaisants sur des ensembles ouverts (en mode on-line). Plusieurs types de RN sont utilisés dans le domaine (par exemple, RBF, MLP, TDNN, LVQ,...). Cependant dans la phase de reconnaissance de notre système, nous avons utilisé deux type de réseau de neurones, le premier est un réseau de neurones conventionnel multicouches de type MLP (Multi Layer Perceptron), et le deuxièmes est un réseau de neurones à potentiel d'action asynchrone/à Spike, ce dernier est une implémentation du principe décrit par les biologistes Délorne *et al.* [12]. Dans les deux approches, le système décide que la personne à identifier correspond à la personne de la base qui a le maximum de points similaires. Donc le rôle de la reconnaissance, c'est de confirmer qu'il s'agit de telle ou telle personne, ou de dire que cette personne n'appartient pas à la base de données (voir la phase de reconnaissance de la figure 1).

#### 4.1. Réseau de neurones artificiel, « Système 1 »

Le réseau MLP utilisé dans la première approche recevra les 960 valeurs réelles qui correspondent aux 960-pixels de l'image d'input (la taille de l'image  $\sim 32 \times 30$ , la Figure 2). Il sera alors exigé pour reconnaître la personne de répondre en activant un seul neurone parmi les N-éléments de la couche de sortie ou le vecteur d'output. Chaque élément des N-éléments du vecteur d'output représente une seule personne/classe (Figure 2).

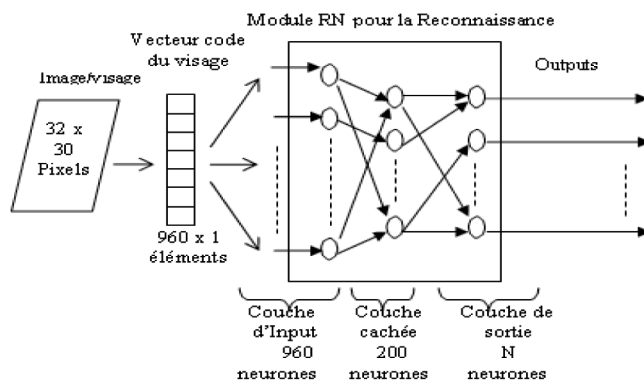


Figure 2. Architecture du réseau de neurones.

Pour que le réseau fonctionne correctement, il devrait forcer la valeur par un 1 du neurone qui correspond au rang de la classe de la personne présente au réseau. Toutes les autres valeurs des neurones dans le vecteur d'output devraient être égale à 0.

**Remarque :** Le nombre de neurones dans les couches du réseau dépend de la base de données utilisée. Dans la couche d'entrée le nombre de neurones correspond au nombre des éléments du

vecteur d'input (960 éléments). Dans la couche cachée le nombre de neurones est désigné par expérience et selon le principe utilisé dans les articles suivants [5, 8]. Si le RN a des troubles pendant l'apprentissage, on doit ajouter des neurones. Enfin le nombre de neurones (N) dans la couche de sortie correspond au nombre de personnes (classes) de chaque base [8].

L'Apprentissage (le processus et les paramètres)

L'apprentissage du réseau est supervisé, pour cela nous avons utilisé l'algorithme de rétro-propagation (Backpropagation) pour régler ses poids synaptiques. Il doit respecter un ensemble de paramètres tel que : le but de convergence ou le critère d'arrêt (goal), et le nombre d'époques maximum (epochs). En plus, il doit faire l'apprentissage sur différents ensembles pour satisfaire au problème posé. Initialement, le réseau de neurones fait l'apprentissage pour un maximum de 5 000 époques ou jusqu'à atteindre une valeur de la somme de l'erreur quadratique (SSE) inférieure ou égale à 0.001. Figure 3 montre la convergence du processus d'apprentissage du réseau neurones proposé, et le critère d'arrêt (goal) est fixé à  $10^{-3}$  ou à 0.001.

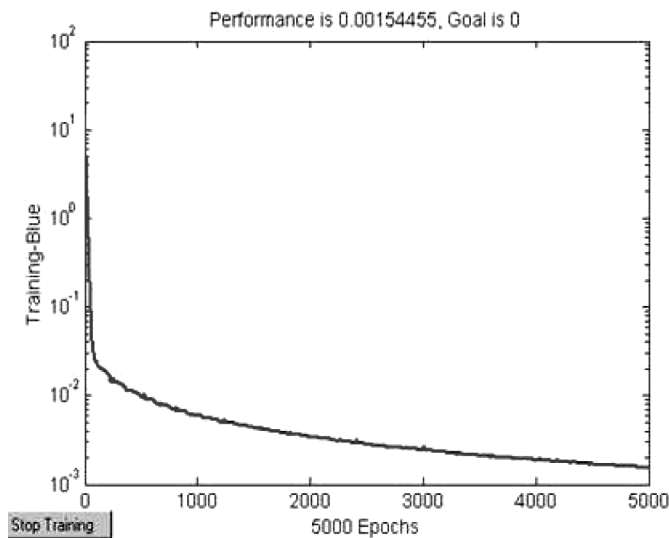


Figure 3. Processus d'apprentissage du RN « système 1 ».

#### 4.2. Réseau de neurones à potentiel d'action asynchrone/ à Spike, « Système 2 »

Maintenant, nous reprenons la même architecture du système de reconnaissance précédent, mais nous implémentons un nouveau réseau de neurones, la différence est que ce deuxième réseau fait l'apprentissage selon le principe décrit par Délorne *et al.* [12]. Dans les tâches comme la reconnaissance des formes, la performance des réseaux de neurones artificiels est toujours médiocre par rapport à celle de l'être humain. En effet, si on utilise une tâche bien précise de catégorisation/classification appelée go no-go, dans laquelle des personnes activent un bouton

lorsque ils détectent un animal par exemple dans une seine de photos normales flashées selon une fréquence de 20 ms (milli-seconde), où Thorpe *et al.* (1996) a prouvé que le traitement exigé dans une telle tâche pourrait être effectué dans moins de 150 ms. Afin d'atteindre des secteurs corticaux de l'oeil d'ordre supérieur, l'information rétinienne doit passer par au moins 10 étapes de traitement, et des contraintes bien connues des neurones [27].

Les propriétés de ce type de catégorisation visuelle ultra-rapide (ultra-rapid visual categorization : URVC) ont depuis fait l'objet de nombreuses études, et sont ainsi mieux connues (voir par exemple la thèse de Denis Fize, 2000) [27]. Tout d'abord, cette capacité ne dépend pas directement d'une allocation de l'attention spatiale : la performance n'est pas affectée lorsque le stimulus peut apparaître aléatoirement à plusieurs positions, ce qui requiert pourtant une répartition des ressources attentionnelles [28]. L'information de couleur dans les stimuli ne semble pas non plus être un élément déterminant de la performance : en l'absence de couleurs, la performance des sujets rapides n'est pas affectée [12], et la différence de potentiels évoqués entre cibles et distracteurs survient à la même latence (*i.e.* 150 ms). Ce type de tâche de catégorisation ultra-rapide peut également être réalisé en vision périphérique extrême, avec toutefois une diminution de la performance, mais dans des conditions où les sujets sont souvent incapables de reporter verbalement le stimulus présenté [27].

Enfin, cette catégorisation ultra-rapide ne peut être encore accélérée par une connaissance préalable des images présentées, même après 3 semaines de pratique intensive [28], ce qui suggère que cette durée de traitement minimale de 150 ms serait en quelque sorte « incompressible ».

De fait, de récentes études où des électrodes sont implantées sous le crâne chez des sujets humains épileptiques<sup>2</sup>, montrent que les régions occipito-temporales ventrales (correspondant au cortex inféro-temporal du singe) s'activent spécifiquement pour des visages humains ou d'autres objets, avec des latences de réponse entre 150 et 200 ms [30].

Thorpe et Délorne ont proposé que ce traitement soit essentiellement basé sur une décharge continue de l'information, dans n'importe quelle couche donnée du réseau de neurones, où chaque neurone produirait un spike ou une décharge d'action. Avec un spike par neurone, des codes basés sur ce dernier sont remis en cause.

Comme une alternative aux codes neuronaux, il y a le codage temporel basé sur l'ordre relatif dans lequel les spikes sont produites dans les premiers secteurs du cortex visuel. Le réseau neurones à spike peut alors employer la latence ou la latence relative des décharges neuronales, il signifie que leur ordre d'arrivée atteint toujours des performances élevés surtout dans le cas des applications de traitement d'image avec seulement un spike par neurone. En outre, le codage d'ordre est à base de

2. Personne qui souffre de troubles nerveux chroniques caractérisés par des crises soudaines et brutales.

contraste et il peut être analysé d'un point de vue mathématique. Ce codage est exploité dans notre approche, où les neurones peuvent agir en tant que convertisseurs analogiques de latence, avec les neurones qui ont un potentiel le plus fort, ils s'activent les premiers. Par exemple, comme il est montré dans la figure 4, on peut également employer l'ordre de la mise à feu suivant ( $B > A > H > C > G > D > F > E$ ) comme code. Avec 8 neurones, on aura  $(8!) i.e.$  40 320 ordres possibles différents.

Le réseau de neurones à spike est conçu pour simuler des réseaux qui contiennent un nombre très grands des neurones à décharge asynchrone. Ces neurones sont simulés avec un nombre limité de paramètres qui inclut les propriétés classiques comme le potentiel post-synaptique et le seuil de décharge, et également des nouvelles caractéristiques telles que la sensibilité dendritique. Le réseau de neurones à spike peut être employé pour simuler des réseaux avec des millions de neurones et des centaines de millions de poids synaptiques.

L'optimisation du temps et l'exécution en temps réel du calcul sont la force dérivée derrière le développement du réseau neurones à spike [12]. Ce réseau neurones à spike a certaines caractéristiques [30, 12]:

- Effectue le traitement d'image en utilisant le principe du réseau de neurones biologique.
- Simule des millions de neurones qui sont organisés dans la carte retino-optical en les activants par une décharge.
- Relie ces cartes neuronales en utilisant des projections, et regroupe les poids synaptiques communs pour optimiser la sauvegarde en mémoire, dans le but de déclarer et de manipuler plusieurs centaines de milliards de connexions synaptiques.
- Convertit des images en niveau de gris en listes de spike (également le réseau neurones à spike peut exécuter une variété de pré-traitement sur les images d'input)
- Met en application un mécanisme complexe pour la projection des cartes neuronales de différentes tailles.
- Utilise un apprentissage supervisé.
- Utilise un schéma de codification efficace à base d'ordre de décharge neuronal ou Rank- Order-Coding scheme.

Pour ces caractéristiques, nous réalisons un nouveau système de reconnaissance de visage, qui fait l'apprentissage par un réseau neurones à base de spike. Lorsque nous avons examiné rigoureusement sa performance, on a obtenu approximativement 20 millions de connexions par seconde sur un ordinateur standard de bureau (PC avec micro-processeur de 1.7 GHz). Ce qui est suffisant pour modéliser un réseau de 400 000 neurones en temps réel, en utilisant 1ms pour chaque étape (affectant 40 connexions par neurone, et un taux moyen de mise à feu de 1 spike ou une décharge par seconde, une valeur qui est une évaluation raisonnable pour le taux moyen d'une mise à feu des neurones corticaux). Notez que, si l'augmentation de cette résolution par le temps n'a aucun effet sur le temps global de calcul, par contre l'addition d'un paramètre aux neurones engendre une augmentation approximative de 30 à 40 % du temps de calcul. Vers la fin, les résultats des deux systèmes de reconnaissance des visages (« système 1 » & « système 2 ») sont comparés, (pour plus de détails voir la section 5 et 6).

## L'Architecture du réseau de neurones à spike « Système 2 »

Pour illustrer comment un réseau neurones à spike peut être utilisé, nous décrivons un réseau multi-échelles de reconnaissance des visages qui une extension du modèle de la localisation des visages décrit par Van Rullen *et al.* (1998) et Delorme *et al.* (2001), où ils utilisent une architecture inspirée de l'organisation du système visuel des primats. Des images d'input sont d'abord analysées par des vecteurs de cellules centrales On-Center et Off-Center dans la « rétine » à trois résolutions spatiales différentes ou niveaux. Ces cellules envoient des spikes aux neurones de la prochaine couche qui contient des neurones qui correspondent aux 8 différentes orientations à chaque niveau spatial. Des interactions latérales entre les cellules dans cette couche ont été employées pour améliorer la sélectivité, et sont semblables à ceux décrites par Zhaoping Li (1998). Une inhibition de faible intensité a été introduite pour rendre les neurones sensibles à l'ordre de l'activation de leurs inputs. Une couche intermédiaire dans le réseau contient des neurones sélectifs spécifiques aux visages du troisième niveau spatial. Les connexions, entre les cartes d'orientation du niveau 2 et ces unités sélectives du niveau 3, ont fait l'apprentissage sur environ un ensemble de 200 photos des visages, cet apprentissage est supervisé, et relie les poids élevés qui correspondent aux premiers inputs qui mettent le feu ou activent la décharge, et des poids progressivement plus petits à la mise à feu aux derniers inputs. Finalement, la dernière couche contient les neurones qui intègrent l'information des trois différents niveaux spatiaux dans la couche précédente. La simulation est réalisée dans la carte de la couche finale, le feu ou la décharge des neurones est activé si un visage, à n'importe quel niveau, est présent dans l'image d'input. Le modèle n'est clairement pas très réaliste. Par exemple, aucune tentative n'a été faite de modéliser le changement de la résolution avec l'excentricité rétinienne, mais l'architecture démontre comment le réseau neurones à spike peut être utilisé pour réaliser des architectures multicouche tout à fait complexes, et impliquant un grand nombre d'unités, et il montre comment différentes hypothèses pourraient être évaluées et intégrées facilement dans un réseau de neurones convainquant qui effectue la reconnaissance des visages. La figure 4 expose les différentes couches de l'architecture du réseau neurones à spike.

## 5. Expérience et discussion

Pour évaluer la performance des deux méthodes proposées, nous avons collecté un grand nombre d'images de visage à différents moments en utilisant un appareil d'acquisition (ex: Caméra) pour former notre propre base de données 'own Database'. La base de données inclut 240 images de visage ou prises de vue de 40 sujets différents ( $N = 40$ ). Les images sont

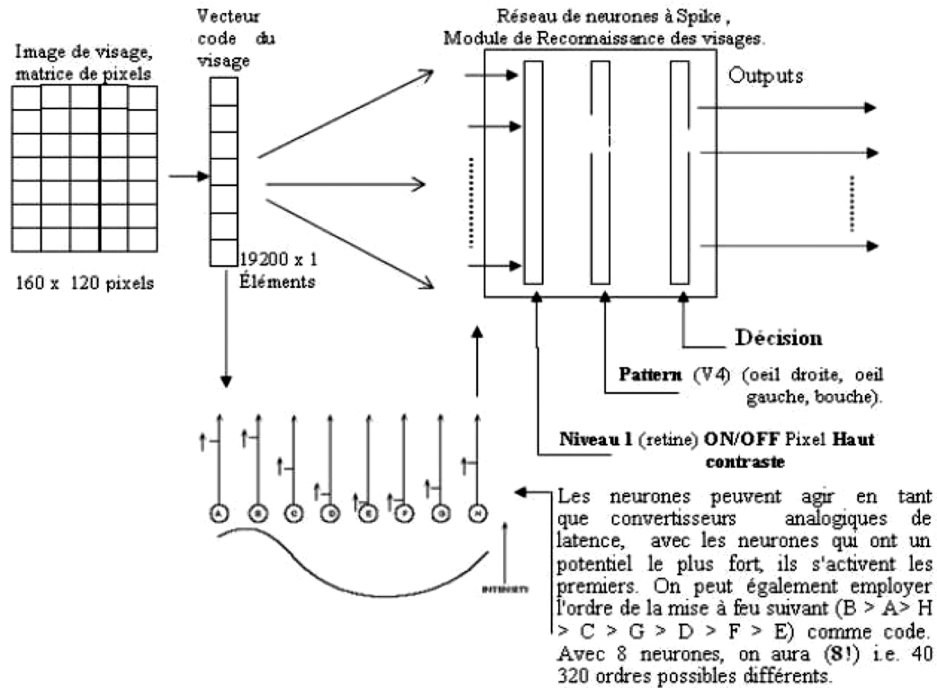


Figure 4. L'Architecture du « system 2 », réseau à spike.

acquises pendant différentes sessions et avec différents types de bruit, ceci pour défier nos deux systèmes (« système 1 » & « système 2 ») et voir leurs performances respectives. Pour tester les deux systèmes, un visage avec du bruit peut être pris et présenté aux deux réseaux de neurones, ainsi une autre base de données des images de visage est téléchargée et présentée aux deux systèmes comme une data-set, cette dernière est nommée CMU Face Database. Elle contient des images des visages de 203 membres de l'université CMU de Chicago (USA). Une troisième base de données est appelée MIT Face Database [10], avec 38 personnes de l'institut de technologie : Massachusetts Institute of Technology (MIT), (dans ce cas  $N = 38$ ). Donc, ces trois bases de données (own Database, CMU Face Database, et

MIT Face Database) peuvent être présentées aux deux systèmes comme des data-sets ou des ensembles de test.

À noter que, les images ont été acquises dans différentes conditions et avec différents types de bruit (c-à-d. différents bruits synthétiques, surtout le bruit gaussien, que nous l'avons employé dans toutes nos expériences). Les images originales ont été collectées comme images en niveau de gris. Le tableau 1 montre le taux de reconnaissance (RR) et l'efficacité des deux systèmes proposés, pour le deuxième système, qui utilise les Spikes, les meilleurs résultats de reconnaissance ont été obtenus et un gain du temps considérable d'apprentissage a été enregistré. Dans ce cas-ci, le système a enregistré un RR de 92.10 %, 95.00 % et de 93,10 % avec un temps maximum d'apprentis-

Table 1. Le résultat de reconnaissance pour chaque base de données, est réalisé sur un PC 1.7 GHz de CPU. RR : le taux de reconnaissance. N : Le nombre de classes distinguées de personnes dans chaque base de données.

Système de reconnaissance des visages	Algorithme de Matching : « Pattern recognition Module »	Temps d'apprentissage maximum	Taux d'apprentissage (RR)					
			MIT Database [10] « N = 38 »		Our own database [7] « N = 40 »		CMU Data-set [21] « N = 203 »	
			Train Set	Test Set	Train Set	Test Set	Train Set	Test Set
« Système 1 »	Réseaux de neurones artificiel	2 heures 17 min 45 sec	94.73 %	71.05 %	95.00 %	77.50 %	88.66 %	73.89 %
« Système 2 »	Réseaux de neurones artificiel à spike	150 ~ 200 ms (milli-seconde)	100 %	92.10 %	100 %	95.00 %	100 %	93.10 %

sage de 150 ms pour les trois bases de données, comparé au premier système, où le taux (RR) est de 71.05 %, 77.50 % et de 73.89 %, avec un temps maximum d'apprentissage de 2 heures 17 min 45 s. Cela fait, que le second système est meilleur que le premier.

## 6. La Performance du système

La fiabilité du système de reconnaissance est mesurée par test, ce test est réalisé par l'introduction d'une centaine d'images des visages en input avec des quantités variables de bruit, *i.e.* bruit synthétique. Effectivement, dans ce cas, nous testons les deux systèmes « Système 1 » & « Système 2 » par des niveaux divers de bruit, puis nous représentons graphiquement le pourcentage des erreurs de reconnaissance de chaque système en fonction du niveau du bruit appliqué.

Un bruit moyen de 0, puis une déviation de 0 à 0.5 sont ajoutés aux images d'input. À chaque niveau de bruit, la sortie de chaque système est calculée après introduction d'une centaine (100) de différentes versions d'image de visage bruitées. La sortie, ou l'output des deux réseaux des deux systèmes « Système 1 » & « Système 2 », est alors calculée par la fonction de transfert compétitive de sorte que seulement une des sorties de N, représentant la classe de la personne de la base de données ait une valeur de 1.

Le nombre d'erreurs de classification/reconnaissance est alors ajouté et des pourcentages sont obtenus (voir la Figure 5). La ligne discontinue sur le graphe (de couleur bleue) montre la fiabilité du premier système « Système 1 » (qui utilise un réseau de neurone conventionnel comme un module de reconnaissance).

La performance du deuxième système « Système 2 » est représentée dans le graphe de la Figure 5 par une ligne continue (de couleur verte), ce dernier utilise un réseau de neurone à spike dans son module de reconnaissance.

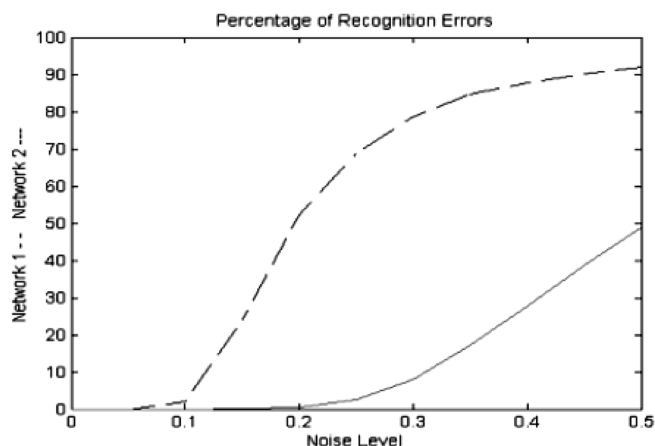


Figure 5. La performance des deux systèmes.

Ainsi, l'apprentissage des deux réseaux des deux systèmes sur des images des visages bruitées a considérablement réduit leurs erreurs lors de la phase de reconnaissance. Les deux réseaux n'ont fait aucune erreur pour des images de visage introduites avec un bruit moyen de 0.00 ou 0.05. Quand un bruit moyen plus de 0.05 est ajouté aux images, le premier réseau du 'système 1' commence à faire des erreurs mais le réseau du deuxième système résiste encore aux bruits, et il commencera à faire des erreurs à partir de 0.20. Si une fiabilité plus élevée est nécessaire, les deux réseaux pourraient être entraînés pendant une plus longue période ou être ré-entraînés avec plus de neurones dans leurs couches cachées, respectivement [5]. En outre, la résolution des images des visages de l'input peut être augmentée jusqu'à avoir une matrice, de 640 x 480 pixels. Et finalement, les deux réseaux des deux systèmes pourraient être entraînés sur des images d'input avec des quantités plus élevées de bruit si une plus grande efficacité était nécessaire.

## 7. Conclusion

Nous avons démontré dans cet article comment un système de reconnaissance des visages peut être conçu par, premièrement, un réseau de neurones artificiel en utilisant l'image du visage comme input, et deuxièmement, par un réseau de neurones à spike, dans le but de capturer les caractéristiques importantes du visage, et de simuler le système visuel humain. Bien que, le réseau de neurones à spike soit conçu comme outil de modélisation des réseaux de neurones biologiques. Le niveau de la performance obtenu est au moins meilleur que le modèle neuronal conventionnel de reconnaissance des formes, sachant que, les niveaux de performance réalisés par le système visuel humain sont, dans l'ordre de grandeur, meilleur que même les systèmes de vision artificiels les plus sophistiqués.

En élucidant les principes informatiques qui rendent ce niveau de la performance possible, il est bien possible non seulement de démontrer la puissance des réseaux de neurones informatiques comme paradigme pour la simulation biologique de la compréhension, mais d'indiquer également le potentiel de la discipline dans des secteurs divers comme la vision et l'intelligence artificielle.

En fin de compte, on peut dire que la reconnaissance des visages est un problème difficile mais motivant, et il reste encore des travaux de recherche qui devraient être effectués dans ce secteur. Au cours de ces dix dernières années, la reconnaissance des visages a suscité l'attention des chercheurs en biométrie, en reconnaissance des formes, en vision par ordinateur, et la communauté des psycho-cognitivistes. Cet intérêt commun, pour la technologie de reconnaissance des visages, est motivé, premièrement par la capacité remarquable de la reconnaissance des personnes, et deuxièmement par l'attention accrue consacrée aux applications de sécurité.



# Références

- [1] R. ADJOUDJ, « Détection & Reconnaissance des Visages En utilisant les Réseaux de Neurones Artificiels », Magister Thesis of artificial Intelligence, Computer Science Department, University of Sidi Bel-Abbès, 2002
- [2] R. ADJOUDJ, A. BOUKELIF, « Artificial Neural Network & Multilevel 2-D Wavelet Decomposition Code-Based Iris Recognition », the International Conference on Computing, Communications and Control Technologies, Paper No. T645XC, CCCT 2004, Austin (Texas), USA, on August 14-17, 2004.
- [3] R. ADJOUDJ, A. BOUKELIF, « Détection et reconnaissance des visages basée sur les réseaux de neurones artificiels », Communication Science & Technologie Journal, COST Journal, Vol. 03, 2005, pp. 78-92.
- [4] R. ADJOUDJ, A. BOUKELIF, « Artificial neural network-based face recognition », International Symposium on Control, Communications and Signal Processing, Paper No 1107, IEEE-ISCCSP 2004, Hammamet, Tunisia, March 2004.
- [5] R. ADJOUDJ, A. BOUKELIF, « Detection and recognition of the faces based on the artificial neural networks », RIST Journal, Vol. 13 n°02, 2003, pp. 93-108.
- [6] R. ADJOUDJ, A. BOUKELIF, « Artificial neural network & mel-frequency cepstrum coefficients-based speaker recognition », 3rd international conference : sciences of electronic, technologies of information and telecommunications, IEEE-SETIT 2005, pp. 174-179, Tunisia, march 27-31, 2005.
- [7] R. ADJOUDJ, « Reconnaissance & Identification Des Personnes Par Authentification Automatique Dans Un Système De Haute Sécurité », PhD thesis, Computer Science Department, University of Sidi Bel-Abbès, 2006.
- [8] Souad BELHIA, Mohamed BENYATTOU, Ahmed LEHIRECHE, Réda ADJOUDJ, « Vérification et reconnaissance des signatures manuscrite hors-ligne en utilisant les SVMs multi-classes et les réseaux de neurones à Impulsion », International conference on Web and information technologies, Sidi Bel Abbes, Algeria, June, 2008.
- [9] D.J. BEYMER, « Pose-Invariant Face Recognition Using Real & Virtual Views », PhD thesis at the Massachusetts Institute of Technology, MIT, USA, 1996.
- [10] C.M. CARRILLO, « Continuous Biometric Authentication For Authorized Aircraft Personnel », master thesis, Naval Postgraduate School Monterey, California, June 2003.
- [11] J. DAUGMAN, « Face and Gesture Recognition: Overview », IEEE PAMI, vol.19, no.7, July 1997.
- [12] A. DELORME, R. VAN RULLEN & S. THORPE, « SpikeNet: a simulator for modeling large networks of integrate and fire neurons », Neurocomputing Journal, Elsevier Ed., Vol.38-40(1-4), 2001, pp. 539-545.
- [13] H. DEMUTH, M. BEALE, Neural Network Toolbox User's Guide For Use with MATLABR, The Math Works Ed., USA, 1998.
- [14] R.P.W. DUIN, P. JUSZCZAK, P. PACLIK, E. PEKALSKA, D. De RIDDER, D.M.J. TAX, PRTools4, A Matlab Toolbox for Pattern Recognition, Delft University of Technology, 2004.
- [15] H. GAMBOA, A. L. N. FRED and A. K. JAIN, « Webbiometrics : User Verification Via Web Interaction ». Biometric Symposium, BCC, Baltimore, September, 2007.
- [16] S. HAYKIN, Neural Networks: A Comprehensive Foundation, Prentice Hall Ed., NJ, USA, 1999.
- [17] R.L. HSU, « Face Detection and Modeling for Recognition », PhD thesis, Department of Computer Science & Engineering, Michigan State University, USA, 2002.
- [18] A. K. JAIN, P. FLYNN, A. ROSS, Handbook of Biometrics, Springer Ed., 2007,
- [19] A. K. JAIN, « Biometric recognition: Q&A », Nature, Vol. 449, pp. 38-40, 6th September 2007.
- [20] D. MALTONI, D. MAIO, A.K. JAIN, S. PRABHAKAR, Handbook of Fingerprint Recognition, Springer Ed., New York, 2003.
- [21] T. MITCHELL, « Neural Net & Face Images », PhD thesis, Carnegie Mellon University, Pittsburgh, USA, October 1997.
- [22] K. NANDAKUMAR, A. K. JAIN and S. PANKANTI, « Fingerprint-based Fuzzy Vault: Implementation and Performance », IEEE Trans. on Information Forensics and Security, 2007.
- [23] K. NANDAKUMAR, « Multibiometric Systems: Fusion Strategies and Template Security », PhD thesis, Department of Computer Science and Engineering, Michigan State University, 2008.
- [24] J.S. NORRIS, « Face Detection and Recognition in Office Environments », Master thesis of Computer Science at the Massachusetts Institute of Technology, M.I.T, USA, May 21, 1999.
- [25] X. LU and A. K. JAIN, « Deformation Modeling for Robust 3D Face Matching » IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007.
- [26] H.A. ROWLEY, « Neural Network-based face detection », Phd Thesis, Carnegie Mellon University, Pittsburgh, USA, May 1999.
- [27] S. THORPE, D. FIZE, C. MARLOT, « Speed of processing in the human visual system », Nature journal, vol.381, 1996, pp. 520-522.
- [28] S. J. THORPE, J. GAUTRAIS, « Rank Order Coding: A new coding scheme for rapid processing in neural networks », Computational Neuroscience: Trends in Research, J. Bower Ed. New York, 1998, pp. 113-118.
- [29] Y.L. TIAN & T. KANADE, « Recognizing Action Units For Facial Expression Analysis », Carnegie Mellon University, Pittsburgh, USA, 1999.
- [30] R. VAN RULLEN, J. GAUTRAIS, A. DELORME, S.J. THORPE, « Face detection using one spike per neurone », Biosystems journal, vol. 22, 1998.
- [31] L. ZHAOPING, « A neural model of contour integration in the primary visual cortex », Neural Computation journal, Elsevier Ed., vol.10, 1998, pp. 903-40.





Réda **Adjoudj**

Ingénieur d'État en informatique, option 'système informatique', de l'université de Sidi Bel Abbès en 2000, Magister en informatique, option 'intelligence artificielle', de l'université de Sidi Bel-Abbès en 2002, titulaire d'un doctorat en informatique, option 'intelligence artificielle' de l'université Djillali Liabes de Sidi bel Abbès 2006. Les domaines de recherche sont la biométrie, la reconnaissance des formes, le traitement d'image, les paradigmes de l'intelligence artificielle, membre du laboratoire des Systèmes distribués & de l'ingénierie évolutive, EEDIS, département Informatique, université de Sidi Bel-Abbès.



Abdel-Kader **Gafour**

Ingénieur d'État en informatique du Centre de Recherche en Informatique (Ex-INI, actuellement ENSI) d'Alger en 1982, Magister en Informatique de l'université de Sidi bel-Abbès en 1997, titulaire d'un doctorat en informatique de l'Université Djillali Liabes de Sidi bel Abbès 2005. Les domaines de compétence sont l'architecture des ordinateurs, le traitement et synthèse d'images, la compression fractale d'images, les réseaux de neurones, les algorithmes génétiques, le test de primalité des nombres premiers.



Aoued **Boukelif**

Ingénieur diplômé de l'université de Pittsburgh (USA), titulaire d'un doctorat d'état en électronique, et d'un Master UTICEF « Sciences de l'Education », spécialité professionnelle de l'université Louis Pasteur (ULP). Les domaines de compétence sont le traitement numérique des images, la télévision digitale, les TIC, les communications mobiles par satellite, la cryptographie, les multimédias et l'interactivité. Il dirige actuellement l'équipe de recherche « NTIC » du laboratoire « Réseaux de Communications, Architectures et Multimédia » de l'université de Sidi Bel-Abbès.



Ahmed **Lehireche**

Ingénieur diplômé du Centre de Recherche en Informatique (Ex-INI, actuellement ENSI) d'Alger en 1981, Magister de l'université des sciences et de technologie d'Oran en 1993, titulaire d'un doctorat d'état de l'université Djillali Liabes de Sidi bel Abbès 2005 et maître de conférence Classe A. Les domaines de compétence sont les outils formels ou heuristiques dédié à L'IA. Actuellement les ontologies dans les SBC. Il dirige actuellement l'équipe d'IA du laboratoire EEDIS de l'université UDL de Sidi Bel-Abbès.