



HAL
open science

Réseaux de neurones en traitement de l'image et du signal

Gilles Burel, Jean-Yves Catros, Isabelle Pottier, Dominique Carel

► **To cite this version:**

Gilles Burel, Jean-Yves Catros, Isabelle Pottier, Dominique Carel. Réseaux de neurones en traitement de l'image et du signal. Flux , 1993, 151-152, pp.18-27. hal-03221743

HAL Id: hal-03221743

<https://hal.univ-brest.fr/hal-03221743>

Submitted on 18 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Réseaux de Neurones en Traitement de l'Image et du Signal

Gilles BUREL (ESE88), Jean-Yves CATROS (ESE76),
Isabelle POTTIER, Dominique CAREL

Thomson CSF, Laboratoires Electroniques de Rennes
Avenue de Belle fontaine, 35510 Cesson-Sévigné

Résumé

Les Laboratoires Electroniques de Rennes de Thomson CSF réalisent depuis plusieurs années des études tant théoriques qu'applicatives dans le domaine des Réseaux de Neurones. A travers divers exemples, cet article montre l'intérêt des Réseaux de Neurones pour la conception de solutions originales à des problèmes de Traitement de l'Image et du Signal.

1 Introduction

Les réseaux de neurones sont des modèles théoriques de traitement de l'information inspirés des observations relatives au fonctionnement des neurones biologiques et du cortex cérébral. Le domaine des réseaux de neurones n'est pas nouveau car il a son origine dans des travaux conduits durant les années 40 (modèle de Hebb pour l'évolution des connexions synaptiques). Ces travaux conduisirent au modèle du perceptron dans les années 60 (modèle qui a principalement été appliqué à la reconnaissance de caractères). Mais ce n'est qu'à partir de 1986 que la recherche dans ce domaine a connu une expansion importante du fait de la publication de modèles de réseaux et d'algorithmes d'apprentissage suffisamment efficaces pour résoudre des problèmes réalistes et complexes.

Il est bien connu que des problèmes aisés à résoudre pour l'être humain (par exemple localiser et reconnaître les divers objets présents sur une image, ou encore comprendre

l'écriture manuscrite) sont très difficiles pour l'ordinateur. On peut donc émettre l'hypothèse que le cerveau humain est plus performant sur ce type de problème parce qu'il utilise un modèle de traitement de l'information très différent de ceux qui ont cours aujourd'hui en informatique. Ainsi, le cerveau traite l'information de façon massivement parallèle. La mémoire y est diffuse et non pas localisée comme dans un ordinateur classique, ce qui rend le système très robuste par rapport aux dégradations (on perd quelques centaines de milliers de neurones par jour sans dégradation significative des performances). De plus, la notion de programmation n'existe pas : l'acquisition des connaissances se fait par apprentissage, c'est à dire par organisation automatique de la structure interne en fonction des données reçues de l'extérieur.

Nous abordons dans cet article certains travaux réalisés aux Laboratoires Electroniques de Rennes (LER) de Thomson CSF, l'un des deux laboratoires centraux de Thomson CSF, dont une des missions est de réaliser des recherches de pointe dans les domaines du traitement du signal et de l'image. Les LER mènent depuis plusieurs années une recherche active dans le domaine des réseaux de neurones. Cette recherche couvre l'aspect algorithmique (recherche de nouveaux modèles et algorithmes d'apprentissage), et l'aspect application dans les domaines de l'image et du signal. Ces travaux ont donné lieu à plusieurs brevets et publications.

Les remarquables performances du système visuel humain laissent entrevoir l'apport que l'on peut espérer de modèles d'inspiration biologique. Pourtant, la plupart des modèles existants sont difficilement exploitables au vu des performances technologiques actuelles, et ne le seront probablement pas avant des dizaines d'années, car trop directement calqués sur la réalité biologique. Notre objectif dans cet article est de montrer la possibilité de résoudre efficacement, par une approche neuronale, des problèmes de traitement de l'image et du signal d'intérêt industriel. Pour cela, l'idée fondamentale de la démarche adoptée consiste à exploiter les connaissances du traitement d'images et les concepts mathématiques développés en traitement du signal, pour modifier, améliorer ou compléter les algorithmes neuronaux existants. En ce sens, nous n'hésitons pas à nous éloigner de la réalité biologique. L'objectif primordial, de notre point de vue, est de mettre au point des algorithmes efficaces au sens des performances, de la robustesse, de la rapidité, et de la facilité de mise en œuvre. Les connaissances de la neurobiologie sont vues avant tout comme une source d'inspiration, mais nullement comme une réalité qui doit être explicitement copiée.

2 Un exemple de modèle neuronal : le perceptron multicouches

On est encore loin de comprendre toutes les subtilités du fonctionnement du neurone biologique. Cependant, les observations réalisées par les neurobiologistes permettent d'élaborer un modèle simplifié qui est représenté figure 1. Un neurone indicé j reçoit des entrées O_i provenant d'autres neurones. Il en effectue une somme pondérée $X_j = \sum_i W_{ij} O_i$, où chaque coefficient de pondération W_{ij} représente la force de la connexion entre le neurone i et le neurone j . Les connexions peuvent être excitatrices (W_{ij} positif) ou inhibitrices (W_{ij} négatif). La valeur de sortie du neurone est obtenue en appliquant une fonction non-linéaire saturante à la somme pondérée.

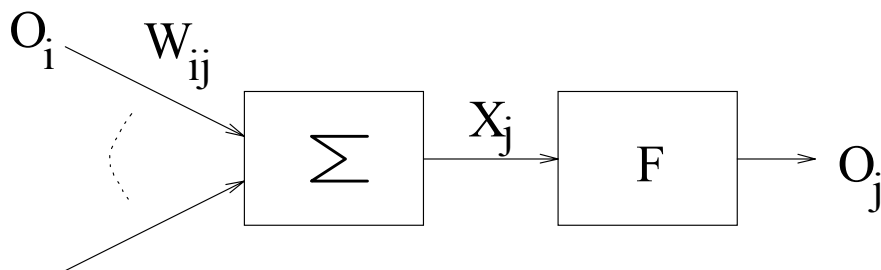


FIG. 1: Le modèle du neurone

Le système nerveux est en première approximation organisé en plusieurs couches (figure 2) : une couche de neurones récepteurs (situés dans la rétine pour la perception visuelle), une couche de neurones moteurs qui activent les muscles et, entre ces deux couches, plusieurs couches dites “cachées” car elles ne sont pas directement en contact avec le monde extérieur. L'apprentissage consiste en une adaptation progressive des connexions synaptiques. Ainsi, en présence de stimulations sensorielles sur sa couche d'entrée, le réseau de neurones fournira une réponse appropriée sur sa couche de sortie.

Les lois régissant la modification des connexions synaptiques dans le cerveau sont très mal connues. Cependant, un algorithme d'apprentissage efficace (mais probablement très éloigné de la réalité biologique) a été proposé en 1986 par Rumelhart et al. [21]. Cet algorithme, connu sous le nom de rétropropagation du gradient, est pour une grande part dans le regain d'intérêt actuel pour les réseaux de neurones. Le principe consiste à présenter successivement des stimulations (ou exemples) sur la couche d'entrée du réseau. Pour chaque exemple, on calcule les sorties fournies par le réseau. La comparaison de ces sorties avec les sorties attendues permet d'obtenir une mesure d'erreur. Il suffit alors de modifier légèrement les coefficients de connexions W_{ij} dans le

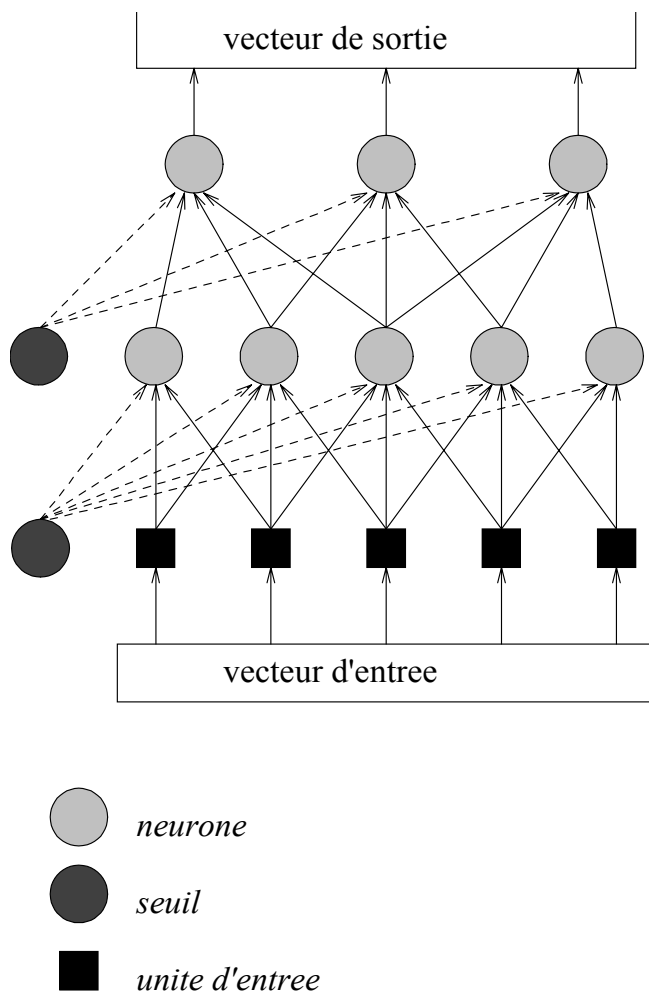


FIG. 2: Le perceptron multicouches (PMC)

sens inverse du gradient de l'erreur pour réduire progressivement l'écart entre les sorties obtenues et les sorties appropriées.

L'un des exemples les plus "médiatiques" d'application de cet algorithme est le système "NET-TALK" conçu en 1987 par T. Sejnowski à l'Université John Hopkins de Baltimore [25]. Il s'agit d'un réseau de neurones qui apprend à lire à haute voix du texte en anglais. Le réseau reçoit en entrée une succession de lettres et doit fournir en sortie la succession de phonèmes et d'intonations appropriée. Le système n'a pas été programmé pour réaliser cette tâche. Il a appris à lire à partir d'exemples fournis par des lecteurs humains.

Les principales applications actuelles des réseaux de neurones sont dans les domaines du traitement du signal, de la parole, et de l'image. Il existe également des applications dans les domaines de la robotique, de la finance, etc. Les premiers résultats semblent très prometteurs, mais il convient bien sûr de ramener les espoirs à de justes proportions. Ainsi, les modèles neuronaux sont généralement simulés sur des ordinateurs classiques (éventuellement équipés de cartes accélératrices) et les réseaux utilisés dépassent rarement le millier de neurones, du fait des limitations technologiques (puissance de calcul et capacité mémoire). On est donc bien loin des quelques 10^{11} neurones du cerveau humain, mais la recherche concernant des réalisations matérielles est actuellement très active.

3 Résultats théoriques et nouveaux algorithmes d'apprentissage

Les modèles neuronaux qui ont donné lieu au plus grand nombre d'applications industrielles sont le perceptron multicouches [21] et le modèle des cartes topologiques [19]. Nous présentons dans cette section des travaux qui visent d'une part à améliorer le modèle du perceptron multicouches (en améliorant son algorithme d'apprentissage, la rétropropagation du gradient, et en généralisant le modèle du neurone grâce à la représentation scalaire distribuée), et d'autre part à mieux comprendre le modèle des cartes topologiques.

3.1 Amélioration de l'algorithme de rétropropagation du gradient

Nous avons dans un premier temps proposé des améliorations de l'algorithme de rétropropagation du gradient, qui visent à permettre une mise en œuvre aisée et efficace sur nos applications de traitement de l'image et du signal [15] [3].

Pour cela, nous avons développé des méthodes de réglage automatique des paramètres d'apprentissage (paramètres d'initialisation des coefficients de pondération, vitesse d'apprentissage, filtrage passe-bas, etc). Nous avons également proposé une fonction d'erreur spécialement adaptée aux problèmes de classification. De plus, afin de limiter l'effet des phénomènes de nature à perturber l'apprentissage, un modèle de neurone (fig 3) qui comprend des contrôleurs de saturation et de perturbation sur les potentiels a été élaboré.

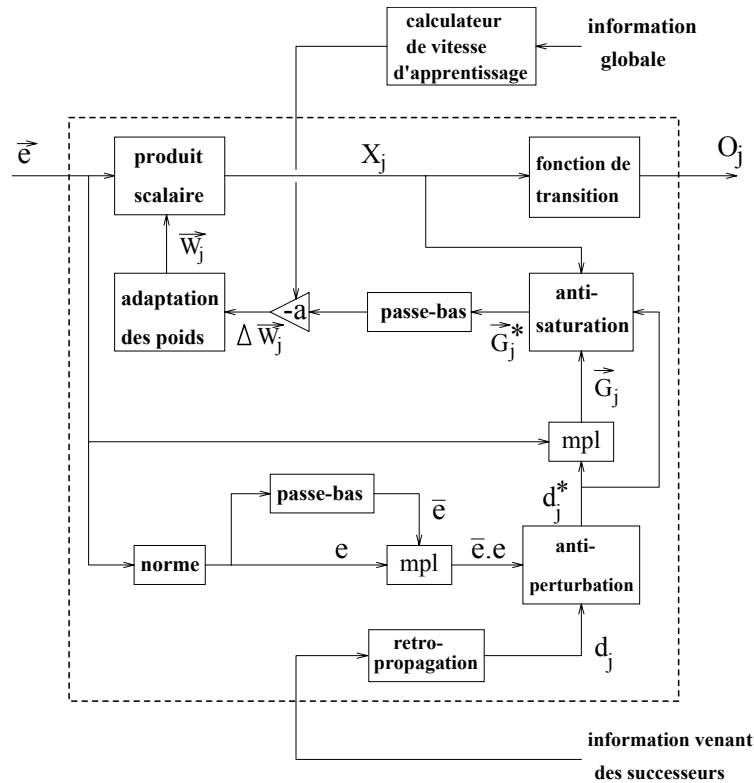


FIG. 3: Nouveau modèle de neurone

3.2 La Représentation Scalaire Distribuée

Une nouvelle approche pour le perceptron multicouches a été proposée [14], en vue d'accroître ses capacités à traiter des problèmes fortement non-linéaires. L'idée consiste à représenter la sortie du neurone non pas par un scalaire, mais par une fonction. Nous désignons par "Représentation Scalaire Distribuée" (RSD) ce passage de l'espace des scalaires à l'espace des fonctions. Nous avons élaboré un algorithme d'apprentissage pour ce nouveau modèle, et nous l'avons validé sur des applications de réduction de dimensionnalité et de prédiction.

A titre d'illustration, considérons un problème de réduction de dimensionnalité. Dans de nombreuses applications de traitement du signal, l'observation est un vecteur appartenant à un espace \mathcal{E} de dimension N . Très souvent, les composantes de ce vecteur ne sont pas indépendantes, et sont en réalité restreintes à une hyper-surface \mathcal{S} de dimension M inférieure à N . La méthode d'analyse en composantes principales (ACP) permet de trouver la meilleure approximation de \mathcal{S} par un hyperplan (meilleure approximation au sens des moindres carrés).

Lorsque l'hyper-surface \mathcal{S} est fortement courbée, une approximation par un hyperplan

n'est pas satisfaisante, et il n'existe actuellement aucune méthode pour traiter ce cas. La figure 4 permet par exemple de comparer les résultats obtenus avec le modèle classique [21] et avec le modèle proposé sur une application de réduction de dimensionnalité. Les données traitées sont des données synthétiques : il s'agit de vecteurs $\vec{v} = [a, b]^T$, aléatoirement choisis sur un quart de cercle (de rayon 1), selon une loi uniforme en fonction de l'angle. Le modèle RSD a conduit à une très bonne approximation de ce quart de cercle, alors que le modèle classique a fourni la même solution que l'ACP, à savoir une approximation par une droite.

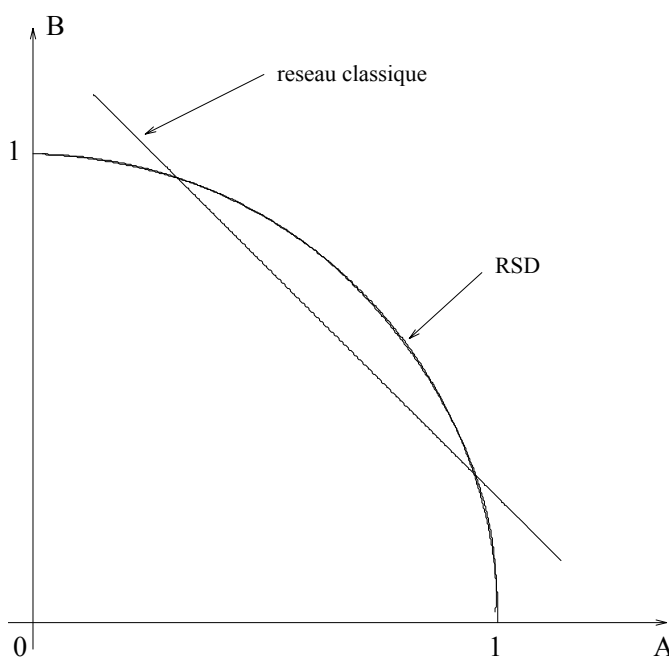


FIG. 4: Réduction de dimensionnalité

3.3 La Quantification Vectorielle avec Voisinage

L'algorithme de Quantification Vectorielle avec Voisinage s'inspire du modèle des Cartes Topologiques [19], ou modèle de Kohonen. Celui-ci trouve son inspiration dans une structure locale particulière existant dans certaines aires du cortex (fig 5). Les neurones sont organisés en couches, et à l'intérieur de chaque couche, chaque neurone émet vers ses voisins les plus proches des connexions excitatrices et vers les neurones plus éloignés des connexions inhibitrices. Tous les neurones reçoivent les mêmes entrées. Kohonen a proposé un algorithme d'apprentissage pour ce type de réseau.

Nous avons mené une étude théorique [9] de l'algorithme de Kohonen, qui nous permet

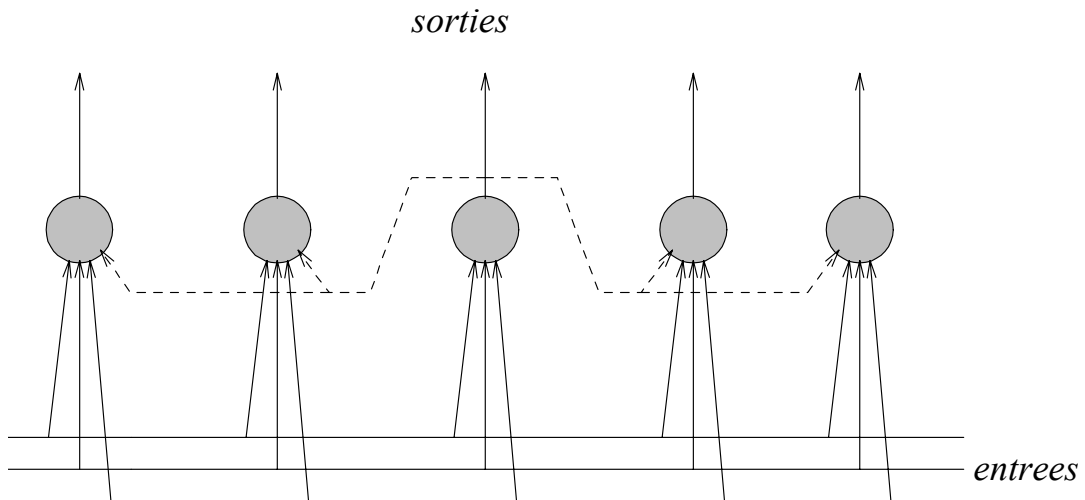


FIG. 5: Modèle de Kohonen (1D)

de trouver l'état de quasi-équilibre de cet algorithme et de démontrer qu'il minimise, sous certaines conditions, une fonction d'énergie. Ces résultats nous conduisirent à proposer un nouvel algorithme de Quantification Vectorielle, l'algorithme VQNf (filtered Vector Quantization with Neighbourhood [3]) qui présente la particularité de conserver la topologie, et possède, par rapport à l'algorithme de Kohonen, l'avantage de ne pas nécessiter de réglage de paramètres, et de minimiser inconditionnellement une fonction d'énergie.

4 Applications de Traitement d'Images

4.1 Idées générales

Bien que l'imagerie numérique semble être le domaine d'application par excellence des réseaux de neurones, ce domaine se caractérise aussi par l'importance et la complexité des données à traiter. Ainsi, une simple image de télévision (noir/blanc) de 720x576 pixels et 256 niveaux de gris est codée sur près d'un demi million d'octets. C'est pourquoi, si chez l'être humain l'information visuelle est directement traitée par un réseau de neurones, une telle approche est impossible avec des neurones formels au vu des performances technologiques actuelles (sauf pour certaines applications très ciblées).

Il est donc indispensable de réaliser sur l'image des traitements préliminaires

permettant de réduire le volume de données à traiter. Il est d'ailleurs intéressant de remarquer que les premières couches de neurones du système visuel humain réalisent effectivement une telle compression de données. De plus, il semblerait que les connexions des neurones correspondant à ces premières couches soient programmées génétiquement et non pas par apprentissage.

Le choix d'un prétraitement permet de fixer l'ordre de grandeur de la taille du réseau. Le deuxième facteur déterminant dans la réalisation d'une application neuronale est la méthode d'apprentissage, qui doit permettre de traiter des données réelles.

Ainsi, pour toute application réaliste de traitement d'image (c'est à dire qui ne se limite pas à des cas d'école), les 2 points suivants sont fondamentaux :

1. Utiliser un algorithme d'apprentissage performant et rapide.
2. Mettre en oeuvre un prétraitement de l'image afin de réduire le volume de données à traiter.

Les résultats présentés dans la section précédente contribuent au premier point. Quant au second point, le problème qui se pose est le choix des bonnes variables caractéristiques à extraire dans le prétraitement. Une approche possible est un choix plus ou moins arbitraire dépendant de la connaissance que l'on a de l'application, ou tout simplement de l'intuition. C'est la démarche qui est généralement adoptée en traitement d'images. Nous avons étudié une autre approche basée sur l'idée que le prétraitement doit réduire le volume de données à traiter, et introduire des invariances par rapport à certaines transformations géométriques, mais ne doit pas introduire de perte d'information. Autrement dit, le prétraitement doit être réversible au sens de l'application à traiter.

Nous avons utilisé une approche neuronale pour diverses applications de traitement d'image. Citons par exemple :

- Classification de textures [5]
- Détection et localisation de visages [10]
- Reconnaissance de chiffres manuscrits [6]
- Reconnaissance de pièces industrielles en deux et trois dimensions [1] [2]
- Vérification de signatures manuscrites [16]
- Classification de véhicules sur images infra-rouges [7]
- Contrôle par la vision [4] du positionnement de composants montés en surface

Nous passons en revue ci-dessous les trois premières applications, afin d'illustrer le type d'approche choisi. Pour une présentation plus détaillée, on se reportera aux références bibliographiques.

4.2 Classification de textures

La texture est une caractéristique importante pour la segmentation de divers types d'images, des images médicales aux images aériennes. De nombreux travaux ont été réalisés dans le domaine de la classification de textures, et de bonnes revues sont fournies par Haralick [22] [23] et Wechsler [24] par exemple. Un point commun aux diverses méthodes proposées est le choix d'un ensemble de caractéristiques considérées comme représentatives de la texture (densités spectrales, paramètres extraits des matrices de co-occurrence, etc).

Une conjecture intéressante a été émise par Gagalowicz [18] : "Deux textures naturelles non structurées possédant les mêmes histogrammes et les mêmes fonctions d'autocovariance ne sont pas discriminables visuellement". La figure 6 illustre cette conjecture. Le cadre de gauche contient quatre textures naturelles. Pour chacune de ces textures, l'histogramme et l'autocovariance ont été calculés et une texture synthétique possédant même autocovariance et histogramme a été générée (cadre de droite). On constate que les textures synthétiques laissent la même impression visuelle que les textures naturelles, ce qui confirme la conjecture.

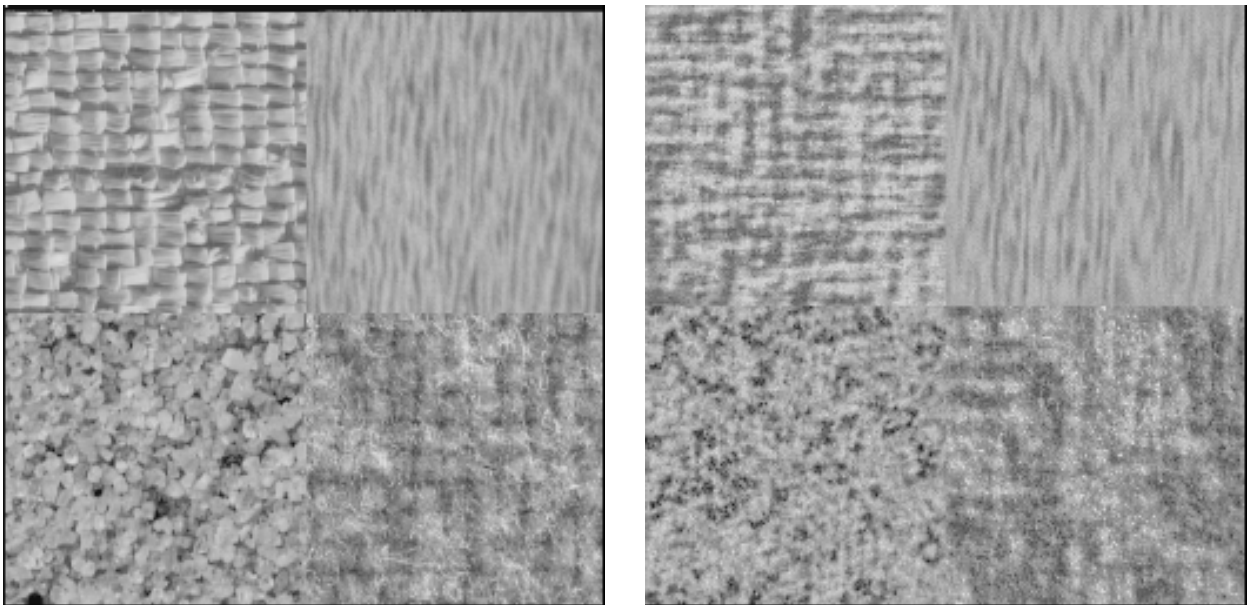


FIG. 6: Textures de Brodatz (à gauche) et leur synthèse (à droite)

On a donc un ensemble de paramètres suffisamment pertinents pour la discrimination visuelle des textures. Cependant, le nombre de ces paramètres est toujours très grand. Une première solution serait de calculer un petit nombre de caractéristiques (entropie de l'histogramme, etc). Mais le risque est alors de perdre de l'information pertinente. Une seconde solution [5] est d'utiliser un classifieur suffisamment puissant pour traiter

directement un tel volume de données. Nos expérimentations ont montré qu'un réseau de neurones est capable d'assurer une telle tâche.

La figure 7 montre, à gauche, deux trames d'un image infra-rouge¹ sur laquelle on distingue trois grands types de fonds : ciel, végétation, et constructions. Nous avons utilisé un perceptron multicouches alimenté directement par l'autocovariance et l'histogramme d'une fenêtre d'observation que l'on déplace sur l'image. En sortie, il y a trois neurones (un pour chaque classe de texture). Après apprentissage, on balaie toute l'image, et la classification obtenue par le réseau est représentée sur l'image de droite de la figure 7.

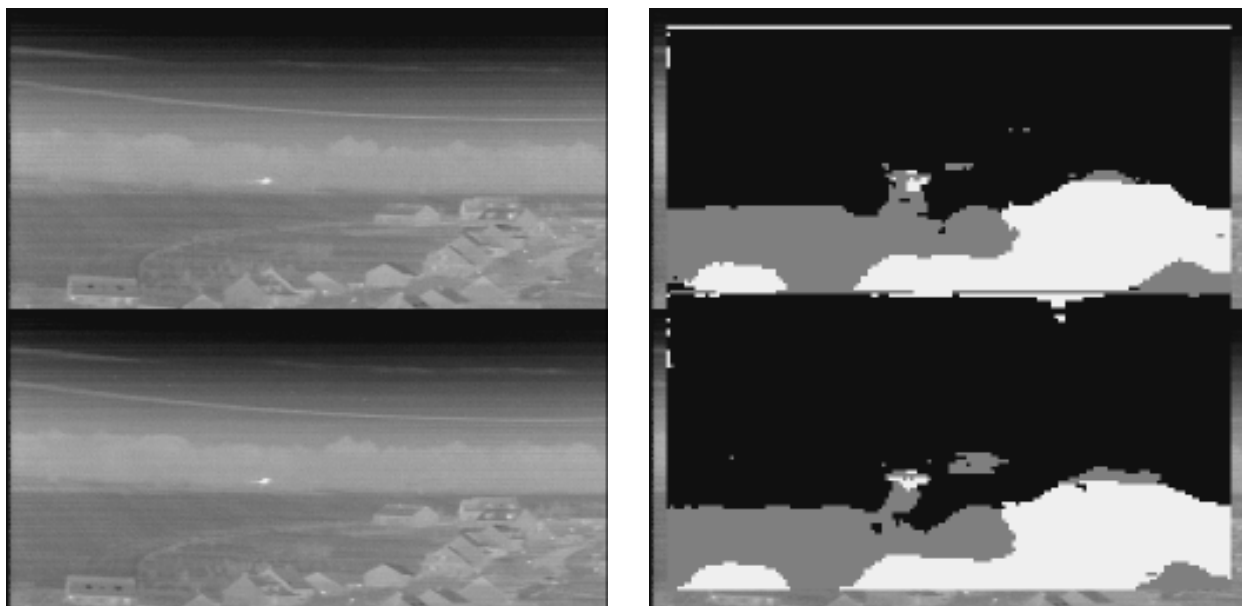


FIG. 7: Classification sur image Infra-Rouge

4.3 Détection et localisation de visages

Un procédé automatique de détection/localisation de visages sur images numériques a été proposé [10]. La détection/localisation de visages consiste, à partir d'une image numérique, à fournir les coordonnées des coins de fenêtres rectangulaires encadrant les visages présents dans l'image. Une indication de confiance peut également être associée à chaque fenêtre.

Il n'existe pas de modèle de visage, du fait d'une part de la forte variabilité d'un

¹Ce travail a été réalisé en partie dans le cadre du programme Thomson CSF "Veille Panoramique Infra-Rouge et Réseaux de Neurones"

individu à l'autre, et d'autre part de la non rigidité d'un visage. On est donc confronté à un problème bien plus difficile que les problèmes classiques de reconnaissance d'objets rencontrés en traitement d'images (reconnaissance de pièces industrielles par exemple), pour lesquels on peut construire des modèles a priori. De plus, dans le cadre d'applications réalistes, on ne maîtrise ni la distance visage-caméra (la taille d'un visage sur l'image peut donc être quelconque), ni les conditions d'éclairage. Enfin, un visage n'est pas nécessairement vu de face. Il faut donc être capable de le détecter même s'il est légèrement de profil.

Pour surmonter ces problèmes, nous avons développé une méthode basée sur les concepts suivants :

- Apprentissage par l'exemple, ce qui permet de surmonter la non-existence de modèle a priori.
- Analyse multi-résolution de l'image, afin de surmonter le fait que l'on ne maîtrise pas la distance visage-caméra. Cette analyse multi-résolution est complétée par un post-traitement visant à supprimer des détections imbriquées.
- Normalisation locale afin d'obtenir une certaine insensibilité par rapport aux conditions d'éclairage.
- Création d'une base d'apprentissage variée, contenant notamment des visages de face et légèrement de profil.

La figure 8 montre un exemple de résultat obtenu.

4.4 Reconnaissance de chiffres manuscrits

La reconnaissance de chiffres manuscrits fait l'objet de nombreuses études à cause notamment de ses applications potentielles dans le tri automatique du courrier. Les méthodes classiques sont généralement basées sur l'extraction de quelques primitives déduites du tracé. Dans le domaine des réseaux de neurones, un grand nombre d'articles a été publié au cours de ces 5 dernières années : l'approche consiste en général à alimenter un réseau de neurones par une image du chiffre à reconnaître.

Nous avons développé une approche différente [6] dans laquelle un prétraitement quasi-réversible est réalisé au préalable sur le chiffre. Ce prétraitement se compose de l'extraction de courbes de profils, de mesures de densité, et de mesures morphologiques. Un perceptron multicouches est alimenté par le résultat de ce prétraitement.

L'approche a été évaluée sur la base de données officielle du Service de Recherche Technique de La Poste. Cette base contient environ 20000 images de chiffres pour l'apprentissage et 12000 pour l'évaluation. La partie servant à l'évaluation est conservée secrète au SRTP afin d'éviter toute contestation sur la validité des performances obtenues. Les images de chiffres proviennent d'enveloppes précasées ou d'enveloppes libres (l'extraction du chiffre est dans ce dernier cas réalisée par un algorithme de



FIG. 8: Détection et localisation automatique de visages

segmentation automatique). Chaque image est étiquetée par 2 opérateurs : le premier a connaissance du contexte, alors que le second n'a accès qu'à l'image du chiffre. Les 12000 chiffres de la base d'évaluation sont reconnus à 87% par notre approche (on utilise comme référence l'étiquetage du premier opérateur, car une partie des chiffres n'est pas correctement reconnue par le second opérateur). Si l'on se restreint aux 9000 chiffres reconnus par le second opérateur, les performances sont de 97%. La figure 9 montre quelques exemples d'erreurs de reconnaissance. Au dessous de chaque chiffre figure la classe de référence (fournie par le premier opérateur), puis la classe et la confiance fournies par notre méthode [12].

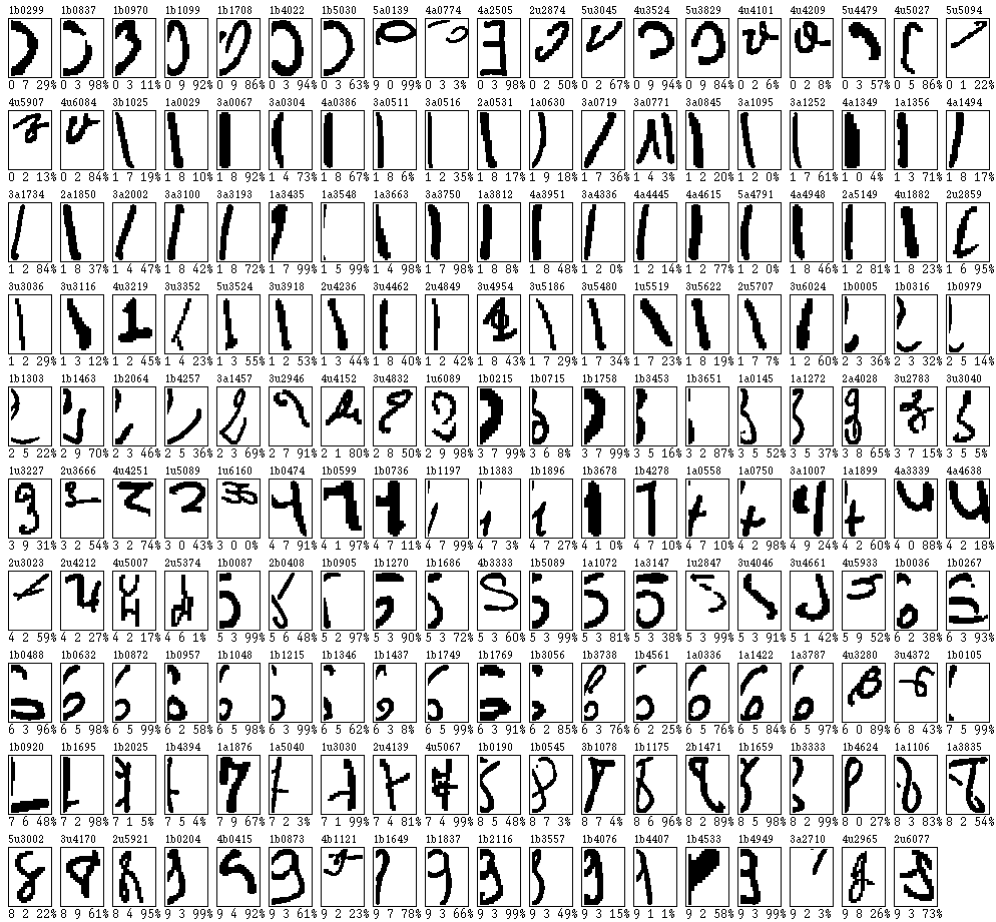


FIG. 9: Erreurs de reconnaissance sur les chiffres manuscrits

5 Applications de traitement du signal

Nous abordons deux applications de traitement du signal : la compression d'images par Quantification Vectorielle [11] et la séparation de mélanges de signaux [8]. D'autres applications telles que l'égalisation de canaux de transmissions numériques [17], et l'estimation d'angles d'arrivée de signaux bande étroite sur une antenne réseau [13] ont également été étudiées.

5.1 Compression d'images par Quantification Vectorielle

Des méthodes de compression d'images par Quantification Vectorielle ont déjà été proposées dans la littérature. L'idée consiste à grouper les pixels de l'image par blocs et à remplacer chaque bloc par un index qui se rapporte à un dictionnaire. Pour que la compression soit efficace, il faut que le dictionnaire soit de bonne qualité, c'est à dire qu'il contienne des blocs bien représentatifs de ce que l'on peut trouver dans une image. Un algorithme efficace pour la génération automatique d'un dictionnaire a été proposé en 1980 : l'algorithme LBG [20].

Nous avons spécifié et évalué une nouvelle méthode de compression d'images [11], qui utilise deux algorithmes neuronaux nouveaux que nous avons développés : l'algorithme VQNf [9] et la Représentation Scalaire Distribuée [14]. L'algorithme VQNf permet de créer un dictionnaire aussi efficace que LBG, tout en présentant le grand avantage de préserver la topologie. Cette préservation de la topologie se traduit par le fait que l'image des index reste cohérente : les figures 10 et 11 permettent de comparer les images des index obtenues, selon l'algorithme qui a créé le dictionnaire. Cette cohérence de l'image des index permet une nouvelle compression : la compression de l'image des index elle-même. Cette compression, est réalisée grâce à un prédictor. La prédiction fournie par un perceptron multicouches (PMC) à Représentation Scalaire Distribuée est nettement meilleure que la prédiction fournie par un PMC classique, ou par un prédictor linéaire classique.



FIG. 10: Image des index (dictionnaire créé par VQNf)

A qualité d'image égale, les gains obtenus sur le taux de compression sont, par rapport

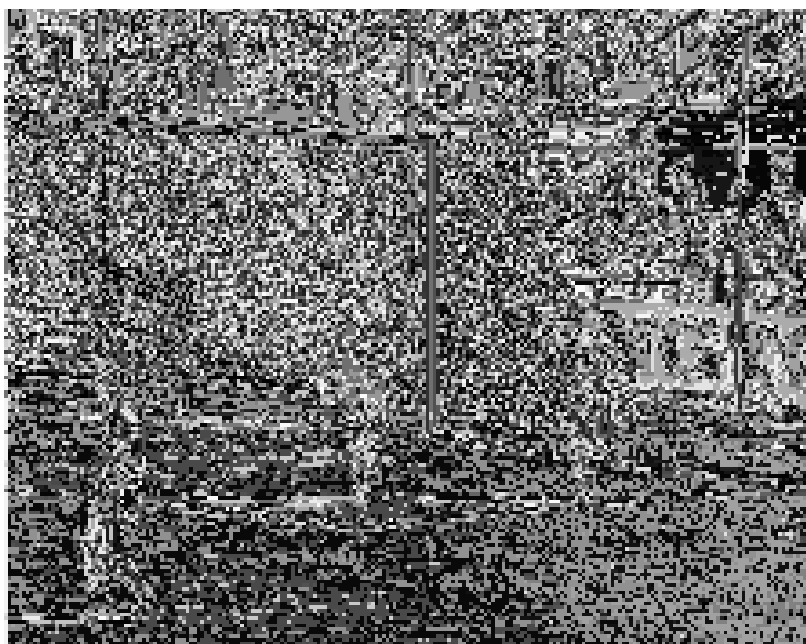


FIG. 11: Image des index (dictionnaire créé par LBG)

à une méthode classique de compression d'images par Quantification Vectorielle, de l'ordre de 15% à 25%. La figure 12 montre, à gauche une image source à 8 bits/pixel, et à droite l'image reconstruite après une compression d'un facteur 18.

5.2 Séparation aveugle de mélanges de signaux

Le problème de la séparation de mélanges de signaux est un problème fondamental en traitement du signal. Par exemple, dans le domaine biomédical, les mesures fournies par les capteurs sont généralement des mélanges de signaux issus de plusieurs sources indépendantes. Le problème consiste à extraire de ces mélanges les signaux utiles. On retrouve le même type de problème en traitement d'antenne pour la séparation de sources sonar ou radar, ou encore en traitement de la parole pour l'amélioration des signaux de diction (élimination des sources sonores parasites : autres locuteurs, bruits de moteurs dans un avion, ...).

Nous avons développé un nouvel algorithme de séparation de sources [8]. Cet algorithme exploite uniquement l'hypothèse d'indépendance des sources, et ne nécessite donc



FIG. 12: Image source et sa reconstruction après compression

aucune hypothèse a priori sur les sources elles-mêmes (séparation aveugle). Il est, de plus, capable de traiter des mélanges non-linéaires. L'idée de base consiste à utiliser un PMC dont la structure est choisie pour correspondre à une forme paramétrique inverse du mélange. Nous avons développé un algorithme d'apprentissage qui permet de minimiser une mesure de dépendance entre les sorties du PMC. A la fin de l'apprentissage, les sorties deviennent statistiquement indépendantes, et sont égales aux sources originales à une permutation et une dilatation près.

La figure 13 montre un exemple de sources originales, et la figure 14 le mélange de ces signaux par une transformation non-linéaire et l'addition d'un bruit. La figure 15 montre les sorties du PMC à l'issue de la convergence : on retrouve bien sur les sorties les sources originales à une permutation et une inversion près. Il est important

de rappeler que l'algorithme a réalisé la séparation sans aucune connaissance a priori sur les sources elles mêmes : seule l'indépendance statistique des sources est exploitée. De plus, aucune hypothèse sur le bruit n'a été utilisée.

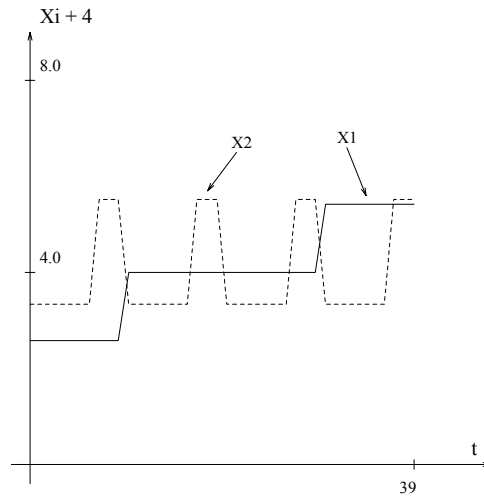


FIG. 13: Les sources

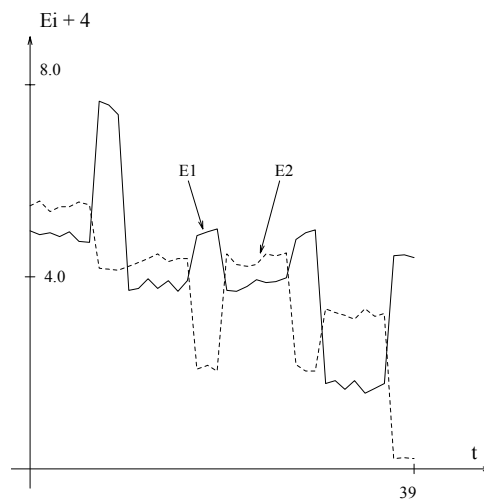


FIG. 14: Mélange bruité des sources

6 Conclusion

Les travaux réalisés aux LER dans le domaine des réseaux de neurones s'étendent suivant deux axes :

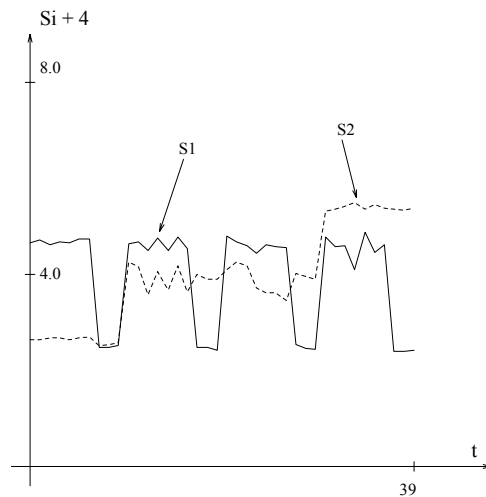


FIG. 15: Signaux obtenus en sortie du Réseau de Neurones

- Un axe théorique, selon lequel nous avons proposé des améliorations des algorithmes existants (réglage automatique des paramètres d'apprentissage), et mis au point de nouveaux algorithmes d'apprentissage.
- Un axe applicatif, selon lequel nous avons traité, par une approche neuronale, différents problèmes de traitement de l'image et du signal (de la segmentation à la séparation de sources, en passant par la reconnaissance de formes et la compression de données).

Cette diversité des problèmes traités permet de montrer dans quelle mesure les réseaux de neurones peuvent être intéressants du point de vue industriel (aussi bien pour la résolution d'applications que pour la mise au point de nouveaux algorithmes). Les travaux relatés permettent de tirer les conclusions suivantes :

1. Les réseaux de neurones sont une puissante source d'inspiration pour la mise au point d'algorithmes nouveaux. Nous avons en particulier proposé les algorithmes suivants :
 - (a) La Représentation Scalaire Distribuée, qui permet la constitution de prédicteurs nettement plus puissants que les prédicteurs classiques. Cette structure, et son algorithme d'apprentissage associé, constituent également un outil de réduction de dimensionnalité qui, contrairement aux méthodes antérieures, n'est pas limité au cas linéaire.
 - (b) La Quantification Vectorielle avec Voisinage a été le fruit d'une étude théorique de l'algorithme de Kohonen. Nous avons montré l'apport de ce nouvel algorithme par rapport aux algorithmes classiques sur une application de compression de données.
2. Dans le domaine du traitement d'images, les réseaux neuronaux permettent de développer rapidement une solution à des problèmes très variés. Pour les différents

problèmes traités en reconnaissance de formes, on retrouve globalement l'approche suivante :

- (a) Réalisation d'un prétraitement réversible, ou quasi-réversible, au sens du problème à traiter. Ceci permet de réduire la complexité du problème, sans risque de perte d'information. La mise en œuvre d'un réseau de neurones autorise un prétraitement simple, ce qui évite toute perte de temps excessive dans la sélection des caractéristiques à extraire.
 - (b) Prise de décision par un réseau de neurones. A ce niveau, les méthodes proposées pour le réglage automatique des paramètres permettent une mise en œuvre aisée. Les réseaux de neurones conduisent généralement à de bons résultats, malgré la simplicité du prétraitement.
3. Dans le domaine du traitement du signal, nous avons montré, sur deux applications, comment les réseaux de neurones, ou les algorithmes dérivés, autorisent la conception de solutions originales à des problèmes classiques :
- En compression de données, nous avons proposé une structure originale, à 2 étages, qui exploite les propriétés topologiques de l'algorithme de Quantification Vectorielle avec Voisinage, et les capacités de prédiction de la Représentation Scalaires Distribuée. Nous avons montré expérimentalement un gain de 15 à 25% par rapport à une structure plus classique.
 - En séparation de mélanges de signaux, nous avons proposé un nouvel algorithme capable de traiter des mélanges non-linéaires. Nous avons montré la robustesse de la méthode par rapport au bruitage des observations.

Du point de vue de l'industriel impliqué dans la recherche, la conclusion générale que l'on peut en tirer, est que les réseaux de neurones constituent un domaine à ne pas négliger, non seulement pour leurs capacités à résoudre rapidement (et donc à moindre coût) des problèmes très variés, mais aussi pour leurs potentialités en tant que source d'inspiration pour le théoricien.

Les travaux relatés dans le présent article ne doivent pas pour autant être considérés comme achevés. Dans le domaine théorique, divers problèmes restent à étudier, notamment en séparation de sources. Dans le domaine applicatif, l'intégration des solutions proposées à des systèmes plus larges est actuellement en cours, notamment en classification de textures (intégration à une chaîne opérationnelle de traitement d'images), et en reconnaissance d'objets dans la cadre de la robotique.

Références

- [1] Gilles BUREL, Dominique CAREL, Jean-Yves CATROS, "A connectionist system for recognition of 2D workpieces", *Revue Technique Thomson CSF*, vol 22, n° 4, décembre 1990
- [2] Gilles BUREL, "Reconnaissance d'objets 3D par réseau d'automates", Congrès AFCET-RFIA, Lyon-Villeurbanne, 25-29 novembre 1991
- [3] Gilles BUREL, "Réseaux de neurones en traitement d'images : des modèles théoriques aux applications industrielles", Thèse de doctorat, Université de Brest, 6 décembre 1991
- [4] Sherif MAKRAM-EBEID (Philips LEP), Gilles BUREL (Thomson CSF-LER), "Neural networks for industrial vision applications", The 5th European Seminar on Neural Networks and Genetic algorithms, Feb. 12-14th, 1992, London
- [5] Gilles BUREL, Jean-Yves CATROS, Hugues HENOCQ, "Caractérisation et classification de textures sur images naturelles", *Traitement du Signal*, vol 9, n°1, mars 1992
- [6] Gilles BUREL, Isabelle POTTIER, Jean-Yves CATROS, "Recognition of handwritten digits by neural network", Congrès IJCNN92, 8-11 juin 1992, Baltimore
- [7] Gilles BUREL, Jean-Yves CATROS, "Vehicle detection and classification by neural networks", CORTEX : Workshop franco-allemand sur les réseaux neuronaux dans la recherche militaire, Saint-Louis, France, 1-2 juillet 1992
- [8] Gilles BUREL, "Blind Separation of Sources : A Nonlinear Neural Algorithm", *Neural Networks*, vol 5, n°6, Nov-Dec 1992, pp. 937-947
- [9] Gilles BUREL, "Nouveaux résultats théoriques concernant les cartes topologiques", *Bulletin d'information des Laboratoires Centraux de Thomson CSF*, décembre 1992
- [10] Gilles BUREL, Dominique CAREL, "Détection et localisation de visages sur image numérique", *Revue Technique Thomson CSF*, vol 25, n°1, mars 1993
- [11] Gilles BUREL, Jean-Yves CATROS, "Image compression using topological maps and MLP", Congrès IEEE-ICNN'93, San Francisco, California, March 28 - April 1, 1993
- [12] Isabelle POTTIER, Gilles BUREL, "Evaluation of a neural system for handwritten digits recognition", Congrès JetPoste'93, Nantes, France, 14-16 juin 1993
- [13] Gilles BUREL, Nadine RONDEL, "Neural networks for array processing : from DOA estimation to blind separation of sources", Invited Paper at IEEE International Conference on Systems, Man, and Cybernetics, Le Touquet, France, October 17-20, 1993
- [14] Gilles BUREL, "Une nouvelle approche pour les réseaux de neurones : la Représentation Scalaire Distribuée", Accepté pour publication dans *Traitement du Signal*
- [15] Gilles BUREL, "Procédé d'apprentissage automatique et réseau connexionniste multi-couches pour la mise en œuvre de ce procédé", Brevet n° 90-13445 déposé le 30 octobre 1990
- [16] Gilles BUREL, Isabelle POTTIER, Thomas VILLARD, "Procédé automatique de reconnaissance de signatures manuscrites par analyse d'images et dispositif destiné à sa mise en œuvre", Brevet n° 91-14814, déposé le 29 novembre 1991
- [17] Jean-Ronan VIGOUROUX, Gilles BUREL, "Procédé de traitement de signaux à composantes complexes, et application à l'égalisation non-linéaire de canaux de transmission numérique", Brevet n° 92-11815, déposé le 6 octobre 1992
- [18] André GAGALOWICZ, "Vers un modèle de textures", Thèse de doctorat d'état ès sciences mathématiques, Université de PARIS VI, 1983
- [19] Teuvo KOHONEN, "Self-Organization and Associative Memory", Springer-Verlag, 1984
- [20] Y.LINDE, A.BUZO, R.M.GRAY, "An algorithm for Vector Quantizer design", *IEEE Trans. on Communications*, vol COM-28, n°1, January 1980
- [21] D.E. RUMELHART, G.E. HINTON, R.J. WILLIAMS, "Learning internal representations by error backpropagation", *Parallel Distributed Processing*, D.E. RUMELHART and J.L. Mc CLELLAND, Chap8, Bradford book, MIT Press, 1986

- [22] R.M. HARALICK, K. SHANMUGAM, I. DINSTEN, "Textural features for image classification", IEEE Trans. Syst.,Man,Cybern., vol SMC-3, pp.610-621, Nov. 1973
- [23] R.M. HARALICK, "Statistical and structural approaches to texture", Proc. IEEE, vol 67, no 5, May 1979
- [24] H. WECHSLER, "Texture Analysis - A survey", Signal Processing 2 (1980) 271-282
- [25] T.J. SEJNOWSKI, C.R. ROSENBERG, "Parallel networks that learn to pronounce English text", COMPLEX SYSTEMS 1, pp145-168, 1987

Table des matières

1	Introduction	1
2	Un exemple de modèle neuronal : le perceptron multicouches	3
3	Résultats théoriques et nouveaux algorithmes d'apprentissage	5
3.1	Amélioration de l'algorithme de rétropropagation du gradient	5
3.2	La Représentation Scalaires Distribuée	6
3.3	La Quantification Vectorielle avec Voisinage	7
4	Applications de Traitement d'Images	8
4.1	Idées générales	8
4.2	Classification de textures	10
4.3	Détection et localisation de visages	11
4.4	Reconnaissance de chiffres manuscrits	12
5	Applications de traitement du signal	14
5.1	Compression d'images par Quantification Vectorielle	15
5.2	Séparation aveugle de mélanges de signaux	16
6	Conclusion	18

Table des figures

1	Le modèle du neurone	3
2	Le perceptron multicouches (PMC)	4
3	Nouveau modèle de neurone	6
4	Réduction de dimensionnalité	7
5	Modèle de Kohonen (1D)	8
6	Textures de Brodatz (à gauche) et leur synthèse (à droite)	10
7	Classification sur image Infra-Rouge	11
8	Détection et localisation automatique de visages	13
9	Erreurs de reconnaissance sur les chiffres manuscrits	14
10	Image des index (dictionnaire créé par VQNF)	15
11	Image des index (dictionnaire créé par LBG)	16
12	Image source et sa reconstruction après compression	17
13	Les sources	18
14	Mélange bruité des sources	18
15	Signaux obtenus en sortie du Réseau de Neurones	19