

## Une Méthode de Reconnaissance de Mots Manuscrits Arabes par Réseau de Neurones Transparent

Samia Snoussi Maddouri et Hamid Amiri

Laboratoire de Systèmes et Traitement du Signal  
Ecole National d'Ingenieurs de Tunis  
B.P 37 Belvédère 1002, Tunisie  
Tél : 874 700 Fax : 872 729

E-mail: samia.maddouri@enit.nru.tn  
hamid.amiri@enit.nru.tn

**Résumé :** la reconnaissance de l'écriture arabe manuscrite constitue un problème de recherche crucial dû à la complexité du script arabe. Dans cet article, nous proposons une approche par réseau de neurones transparents qui se base sur les caractéristiques structurelles des lettres arabes. L'avantage de ce système est qu'il ne nécessite pas une phase d'apprentissage, ni une large base d'exemples. Une étude expérimentale sur des mots isolés est aussi présentée.

**Mots clés :** reconnaissance de l'écriture, écriture arabe manuscrite, réseau de neurones transparent, caractéristiques structurelles.

### 1.Introduction

Les technologies informatiques connaissent une évolution rapide. Bien que la machine ait une capacité de stockage, de calcul et de gestion de données importantes et complexes, elle présente encore des limites. Toutefois, l'acquisition de données, surtout celles de nature textuelle, n'a depuis longtemps que très peu évolué [Chen96]. Il s'avère que l'utilisation d'un clavier est restée un point de passage obligé pour la plupart des applications. Ce procédé est particulièrement peu rentable lorsque les informations à saisir existent déjà sur papier. Rendre la machine capable de lire permettrait de saisir les informations de manière plus aisée et de traiter les documents de façon plus rapide [Nedeljkovité92]. Si la reconnaissance de l'imprimé [Amara99, Srihari93, Bose92], a fait une grande avancée vers ces buts, une solution complète pour la lecture automatique de l'écriture manuscrite n'est pas encore trouvée [Mahjoub99, Miled98, Avila96, Belaïd96, Saon96].

Dans cet article nous présentons une approche de reconnaissance des mots isolés hors ligne qui se base sur un réseau de neurones transparents. Dans la première section on effectue une étude des caractéristiques de l'écriture arabe manuscrite inspirée du comportement et de la lecture de l'homme [Côté98]. Une méthode de projection est utilisée dans cette section pour l'extraction de ces caractéristiques. La deuxième section est réservée pour la méthode de reconnaissance proprement dite.

La méthode proposée utilise un réseau de neurones avec une représentation locale des connaissances, un traitement parallèle de l'information; et une propagation progressive des activations entre les niveaux adjacents des cellules entraînant certains processus ascendants et descendants [Côté98]. Une description de l'architecture de notre système est effectuée en première partie. Le cycle perceptuel du système se compose de deux processus. Le processus ascendant correspond à la phase de propagation de l'information. Le processus descendant se traduit par rétropropagation de l'information afin de générer de

nouvelles hypothèses pour un nouveau cycle de propagation/rétropropagation. La saturation du cycle est la condition d'arrêt pour notre système de reconnaissance. Les différents modules décrits sont présentés dans la figure 1.

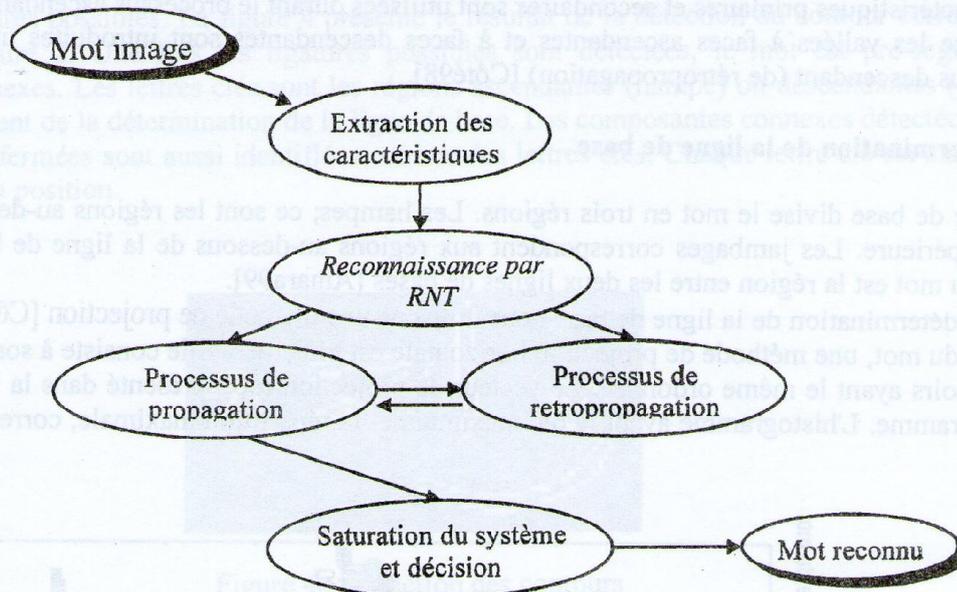


Figure 1: Système de reconnaissance

## 2. Caractéristiques de l'écriture arabe manuscrite

Etant donnée la variabilité de l'écriture manuscrite, nous incluons dans notre système de reconnaissance certaines caractéristiques spécifiques à l'écriture cursive telle que : Hampe, Jambage, Position relative des lettres, Technique de correspondance entre les lettres de l'image et ceux du lexique, Analyse contextuelle [Tang92, Lam88, Miclet87]. Le tableau 1 résume les caractéristiques des lettres arabes que nous essayons de déterminer à partir de l'image scannée.

Type	Primaire		Secondaire		
Caractéristiques	Hampe	Jambage	Hampe/ Jambage	Boucle	Semi boucle
Lettres arabes	أ ل ك ط ظ	ح خ غ ع س ش ص ض ق ر ز و م ن	ل	ف ق و م ة ه ه ق ص ض ط ظ	غ ع ن

Tableau 1: lettres arabes et leurs caractéristiques

L'extraction des caractéristiques est basée sur trois étapes: le prétraitement, la détermination de la ligne de base, et la détection des primitives. Etant donné une image scannée, le prétraitement consiste à la détection des contours, l'extraction des boucles, et l'identification d'un minimum locale du contour supérieur afin de détecter les ligatures potentielles. Suite à ce prétraitement, l'identification de la ligne de base est nécessaire pour la dernière étape d'extraction des caractéristiques.

Trois types de caractéristiques sont à détecter : les caractéristiques primaires (correspondant aux lettres isolées), les caractéristiques secondaires, les vallées à face ascendante et les vallées à face descendante. Les caractéristiques primaires et secondaires sont utilisées durant le processus ascendant (de propagation); alors que les vallées à faces ascendantes et à faces descendantes sont introduites uniquement dans le processus descendant (de rétropropagation) [Côté98].

## 2.1 Détermination de la ligne de base

La ligne de base divise le mot en trois régions. Les hampes; ce sont les régions au-dessus de la ligne de base supérieure. Les jambages correspondent aux régions au-dessous de la ligne de base inférieure. Le corps du mot est la région entre les deux lignes de bases [Amara99].

Pour la détermination de la ligne de base nous utilisons une méthode de projection [Côté98]. Etant donné l'image du mot, une méthode de projection horizontale est effectuée. Elle consiste à sommer le nombre de pixels noirs ayant le même ordonnée. Le vecteur de projection est représenté dans la figure 2 sous forme d'histogramme. L'histogramme ayant le pic maximum et la répartition maximale, correspond à la ligne de base.

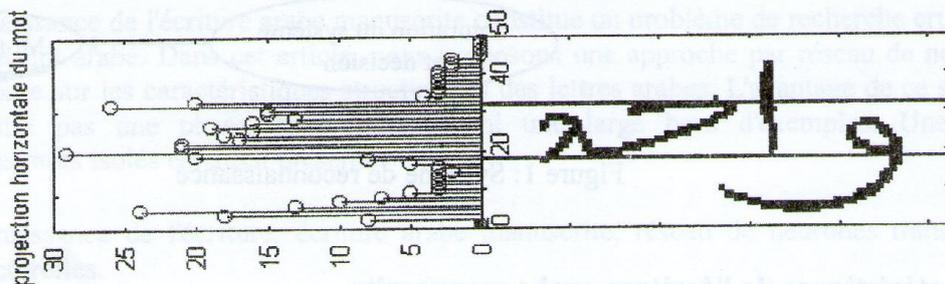


Figure 2 : Illustration de l'extraction de la ligne de base

## 2.2 Extraction des Hampes et des Jambages

L'idée de la détection des caractéristiques primaires consiste à ne pas segmenter le mot en caractères mais à délimiter uniquement les lettres clés (contenant les hampes et les jambages). Cette étape utilise la projection verticale. La figure 3 illustre cette méthode. Une première estimation du nombre et des positions des caractéristiques primaires telles que les hampes et les jambages peut être déduite à partir des pics maximaux.

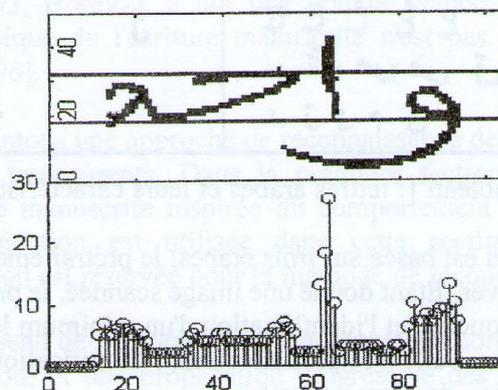


Figure 3 : extraction des caractéristiques primaires

### 2.3 Détermination des boucles

La détermination des boucles se base sur la détection du minimum locale du contour supérieur afin de localiser les ligatures possibles. La figure 4 présente le résultat de la détection du contour correspondant au mot de la figure 2. Une fois les ligatures possibles sont détectées, le mot est pré-segmenté en composantes connexes. Les lettres clés sont les régions ascendantes (hampe) ou descendantes (jambage) détectées au moment de la détermination de la ligne de base. Les composantes connexes détectées comme étant des boucles fermées sont aussi identifiées comme des lettres clés. Chaque lettre clé est caractérisée par sa largeur et sa position.

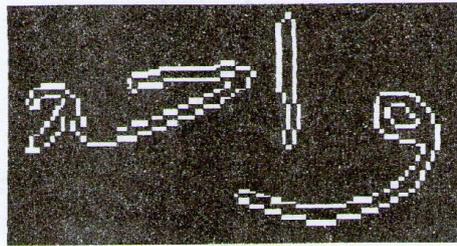


Figure 4 : extraction des contours

### 2.4 Extraction des caractéristiques secondaires

Les caractéristiques secondaires sont détectées pour améliorer l'information sur les caractéristiques primaires. Par exemple, pour la détection d'une boucle, nous essayons d'identifier à quelle lettre correspond cette boucle. Si c'est une boucle sans jambage, ceci implique qu'il s'agit d'un "Fa ف". Si elle est avec jambage, elle peut être un "Qaf ق", "Mim م" ou un "waw و". Si le point de ligature est à gauche de la boucle il s'agit d'un "Mim م" s'il est à sa droite il s'agit d'un "waw و". Des caractéristiques secondaires plus détaillées sur la forme de la boucle permettent de séparer la lettre "Mim م" de la lettre "Ha ه".

### 2.5 Détermination des vallées

Ici, l'écriture et l'arrière plan de l'image sont tous les deux pris en considération. Les deux composantes connexes de l'arrière plan, à savoir la vallée ascendante et la vallée descendante sont détectées entre la partie haute et la partie basse du contour. La figure 5 illustre la détection des deux types de vallées.

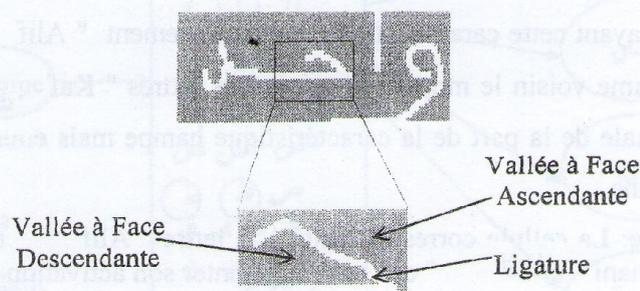


Figure 5 : Vallée à face ascendante/descendante

### 3 Reconnaissance des mots manuscrits arabes

La méthode de reconnaissance se base sur un réseau de neurones à trois couches qui identifie un mot à partir des caractéristiques qui le décrivent à travers deux cycles. Un premier cycle est dit processus ascendant ou de propagation et un deuxième cycle dit descendant ou de retropropagation.

Une correspondance entre les zones de lettres dans l'image d'entrée du mot manuscrit et les lettres dans chaque mot imprimé du lexique en se basant sur des informations contextuelles est effectuée à la fin de chaque cycle. La sortie est une liste de mots candidats qui seront classés selon l'ordre décroissant de leurs activation afin de décider sur le résultat de la reconnaissance.

Le système est basé sur trois couches de cellules organisées de manière hiérarchique (voire figure6). Le réseau de neurones de la figure6 contient une cellule pour chaque mot du lexique, chaque lettre de l'alphabet arabe (28 lettres arabes + les lettres à l'intérieure des mots).

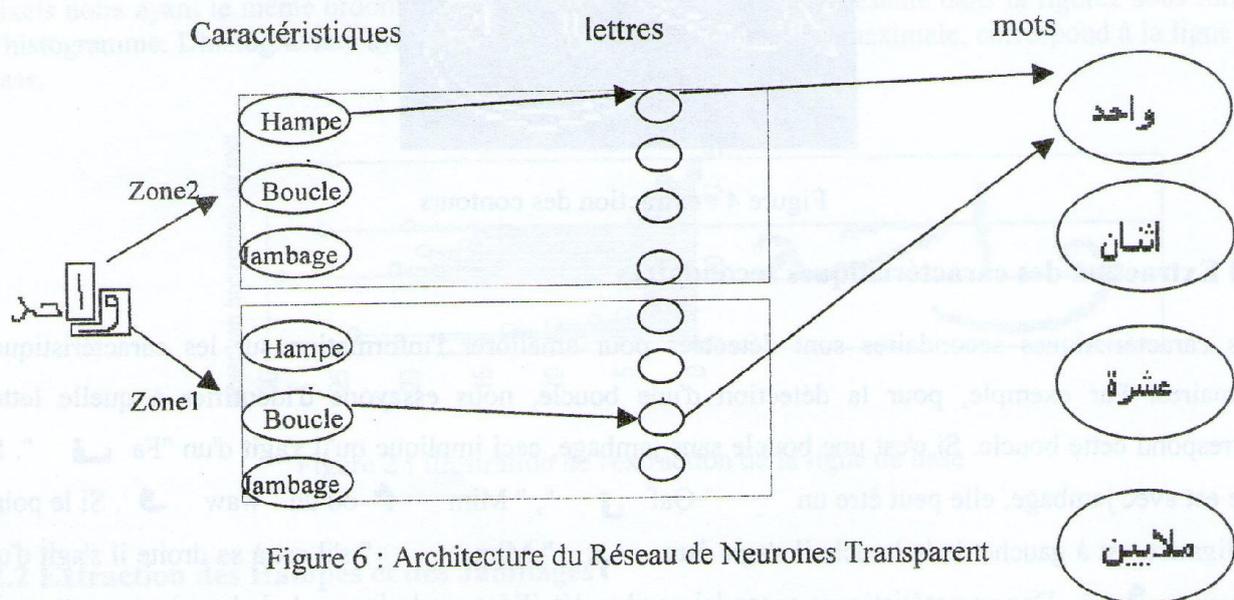


Figure 6 : Architecture du Réseau de Neurones Transparent

#### 3.1 Etat des cellules

Une cellule peut être soit active soit passive, la valeur de son énergie interne varie entre 0 et 1. Lorsqu'une cellule est stimulée, son degré d'activation croît et peut influencer ses voisins (ce sont les cellules du niveau supérieur auxquelles elle est connectée). Donc l'activation d'une cellule dépend non seulement de son énergie interne, mais aussi de l'activation de ses voisins.

- **La cellule désactivée :** c'est un noeud qui a reçu une activation maximale, mais celle-ci ne préserve pas cette activation puisqu'elle n'a pas de voisins. Par exemple, le mot " Wahidon واحد " va activer toutes les lettres ayant cette caractéristique mais uniquement " Alif أ " va garder son activation puisqu'elle a comme voisin le mot en question. Les lettres " Kaf ك " et " Lem ل " ont reçu une activation maximale de la part de la caractéristique hampe mais elles vont être désactivées car elles n'ont pas de voisins.
- **La cellule activée:** La cellule correspondante à la lettre " Alif أ " a deux voisins " Wahidon " et " Ethnani اثان " qui vont augmenter son activation jusqu'à atteindre un maximum.
- **La cellule inactivée:** La caractéristique hampe ne peut pas activer la cellule de la lettre " Bae ب " donc elle reste inactive et elle ne peut par la suite activer son voisin " Arbaaton أربعة " .

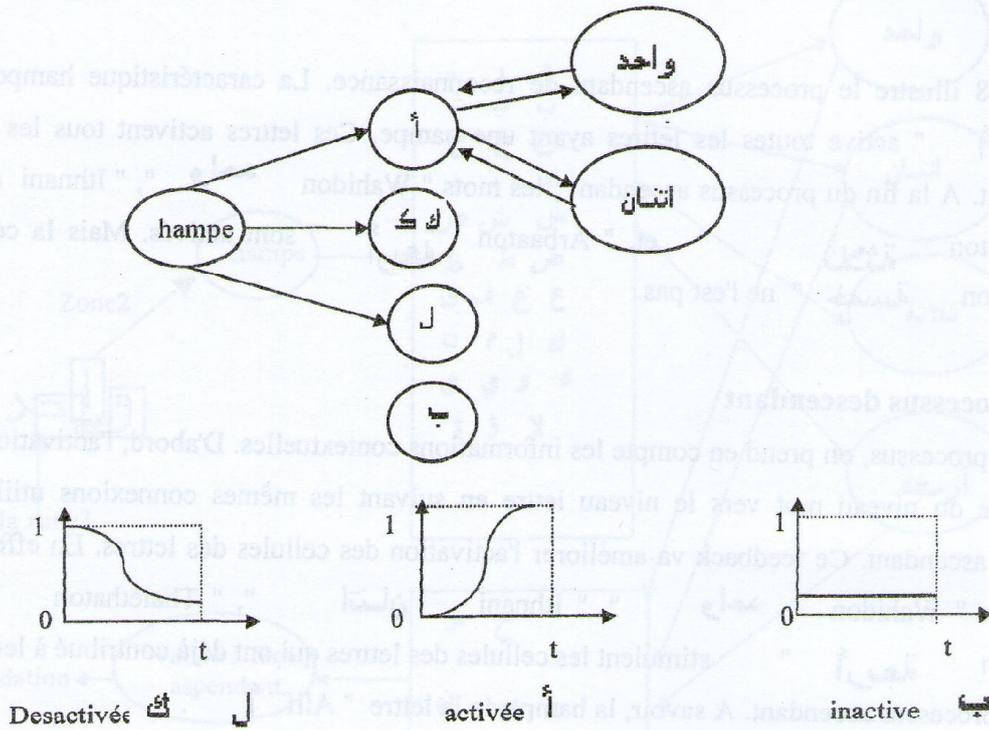


Figure 7 : Les états d'activation des cellules

### 3.2 Le processus ascendant

Les caractéristiques correspondantes aux lettres clés telles que " hampe, jambage, boucle ", sont traitées en priorités puisqu'elles constituent les parties stables de l'image.

L'étape d'extraction des caractéristiques fournit :

- Les descriptions des lettres clés,
- Leurs ordres d'apparition,
- Leurs positions relatives (et non absolues), cette position est décrite par une zone qui peut contenir plus d'une seule lettre.

L'initialisation consiste à activer les cellules correspondantes aux caractéristiques des zones construites pour chaque lettre clé détectée. L'activation des cellules relatives à la couche des caractéristiques initialise la propagation de l'activation vers les cellules de la couche supérieure (des lettres). De même, l'activation des cellules des lettres initialise la propagation de l'activation vers les cellules des mots. A la fin du processus ascendant certaines cellules des mots parmi lesquelles il y a la solution seront actives.

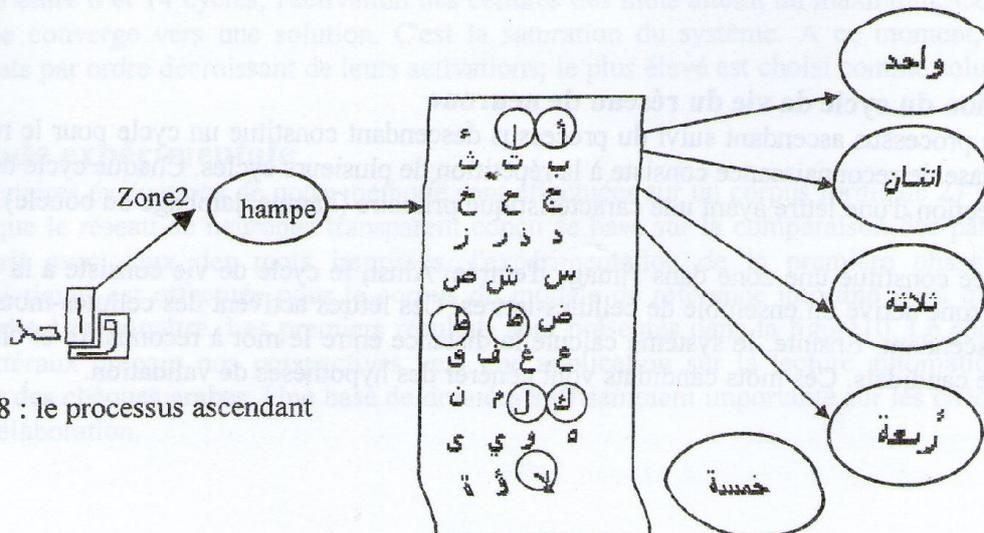


Figure 8 : le processus ascendant

La figure 8 illustre le processus ascendant de reconnaissance. La caractéristique hampe de la lettre " Alif  $\text{أ}$  " active toutes les lettres ayant une hampe. Ces lettres activent tous les mots qui les contiennent. A la fin du processus ascendant ; les mots " Wahidon  $\text{واحد}$  ", " Ithnani  $\text{اثنان}$  ", " Thalèthaton  $\text{ثلاثة}$  " et, " Arbaaton  $\text{أربعة}$  " sont activés. Mais la cellule du mot " Khamsaton  $\text{خمسة}$  " ne l'est pas.

### 3.3 Le processus descendant

Durant ce processus, on prend en compte les informations contextuelles. D'abord, l'activation des cellules se propage du niveau mot vers le niveau lettre en suivant les mêmes connexions utilisées dans le processus ascendant. Ce feedback va améliorer l'activation des cellules des lettres. En effet, les cellules des mots : " Wahidon  $\text{واحد}$  ", " Ithnani  $\text{اثنان}$  ", " Thalèthaton  $\text{ثلاثة}$  ", " Arbaaton  $\text{أربعة}$  " stimulent les cellules des lettres qui ont déjà contribué à leurs activations durant le processus ascendant. A savoir, la hampe de la lettre " Alif  $\text{أ}$  ".

Ensuite, nous utilisons l'information contextuelle donnée par le lexique afin d'améliorer les chances de reconnaissance du mot présenté au système. Les cellules actives des mots génèrent des hypothèses devant être présentes dans l'image. Si les caractéristiques correspondant aux lettres du mot du lexique sont perceptibles dans l'image du mot à reconnaître, les hypothèses sont validées et les cellules correspondantes sont activées; si non, elles seront rejetées.

Dans l'exemple de la figure 9, la cellule du mot " wahidon  $\text{واحد}$  " propose comme lettre " Ha  $\text{ح}$  ", le mot " Ithnani  $\text{اثنان}$  " propose la lettre " Noun  $\text{ن}$  " et le mot " Arbaaton  $\text{أربعة}$  " propose la lettre " Ra  $\text{ر}$  ". Les hypothèses sont donc faites sur les lettres " Ha  $\text{ح}$  ", " Noun  $\text{ن}$  " et " Ra  $\text{ر}$  " à gauche de la lettre " Alif " précédemment reconnue. Mais dans l'image à gauche de cette zone, on détecte une vallée à face ascendante. Cette caractéristique valide la présence de la lettre " Ha  $\text{ح}$  ", et rejette les lettres " Rae  $\text{ر}$  " et " Nour  $\text{ن}$  " qui ont en plus la caractéristique "Jambage".

### 3.4 La saturation du cycle de vie du réseau de neurone

L'application du processus ascendant suivi du processus descendant constitue un cycle pour le réseau de neurones. La phase de reconnaissance consiste à la répétition de plusieurs cycles. Chaque cycle de vie, est initié par la détection d'une lettre ayant une caractéristique primaire (hampe, jambage ou boucle).

La lettre détectée constitue une zone dans l'image d'entrée. Ainsi, le cycle de vie consiste à la détection d'une zone. La zone active un ensemble de cellules-lettres. Ces lettres activent des cellules-mots. Il s'agit du processus ascendant. Ensuite, le système calcule la distance entre le mot à reconnaître et chacun des mots du lexique candidats. Ces mots candidats vont générer des hypothèses de validation.

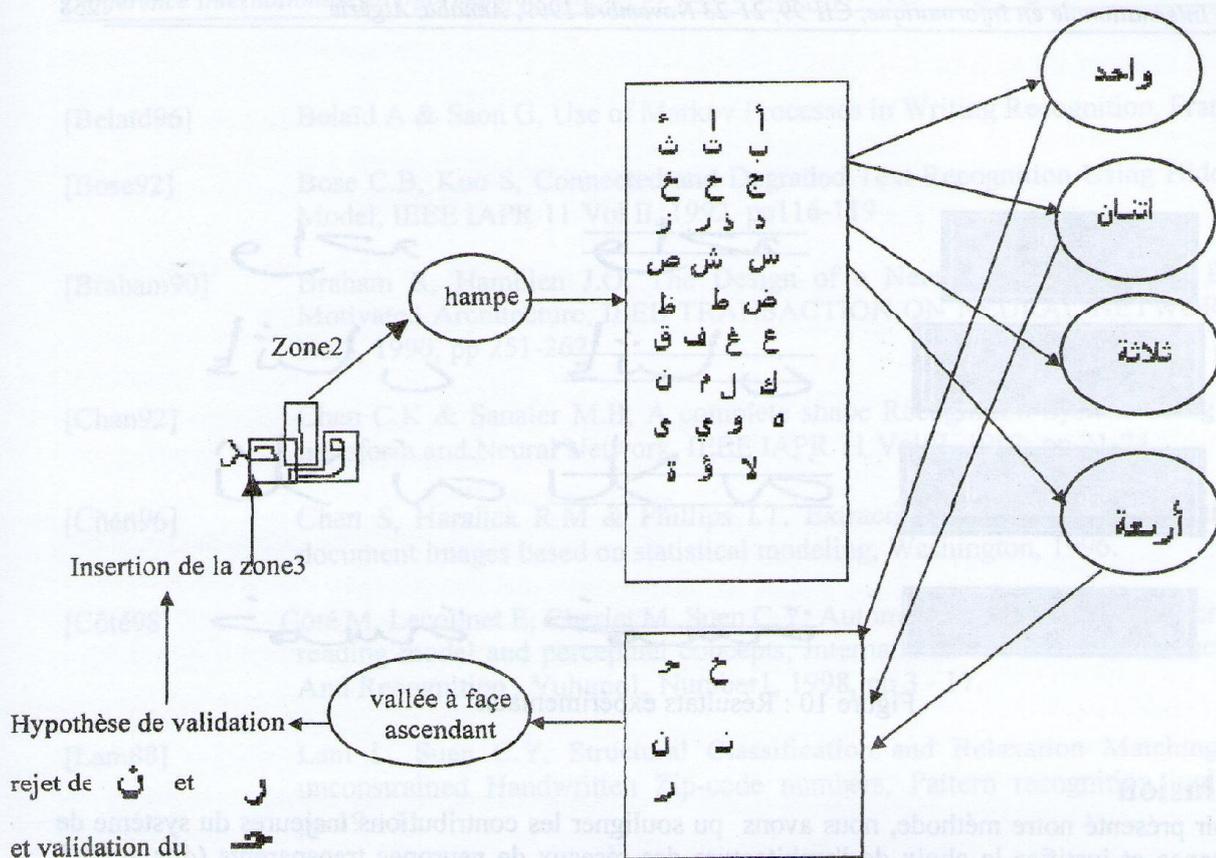


Figure 9 : processus descendant

Pour chaque mot possible, on essaye de valider les hypothèses qu'il génère sur l'image d'entrée. Une lettre dans le mot candidat sera validée si ses caractéristiques existent dans l'image. Quand une lettre est validée le score associé au mot correspondant sera incrémenté. Les mots ayant les meilleurs scores vont participer dans le cycle de vie suivant. Les autres mots vont être ignorés. Pour chaque mot candidat et pour chaque lettre validée de ce mot, une nouvelle zone de l'image va être détectée. Les cellules des caractéristiques associées à cette zone sont activées. Ainsi, on commence un nouveau cycle de vie, durant lequel ces caractéristiques vont contribuer à la propagation et l'activation des différentes couches du système.

Après un certain nombre de processus ascendant (propagation) et descendant (rétropropagation) en général entre 6 et 14 cycles; l'activation des cellules des mots atteint un maximum. Ceci implique que le système converge vers une solution. C'est la saturation du système. A ce moment, on liste les mots candidats par ordre décroissant de leurs activations; le plus élevé est choisi comme solution.

#### 4. Etude expérimentale

Les premières évaluations de notre méthode sont effectuées sur un corpus formé de 27 mots arabes. Etant donné que le réseau de neurones transparent conçu se base sur la comparaison des paramètres d'un mot manuscrit avec ceux des mots imprimés, l'expérimentation de la première phase d'extraction des caractéristiques est effectuée pour le corpus du modèle de référence imprimé ainsi que celui des mots manuscrits à reconnaître. Les premiers résultats sont présentés dans la figure 10. Le choix du corpus des mots littéraux prépare nos perspectives vers une application sur la lecture automatique des montants littéraux des chèques arabes. Une base de données suffisamment importante sur les chèques arabes est en cours d'élaboration.

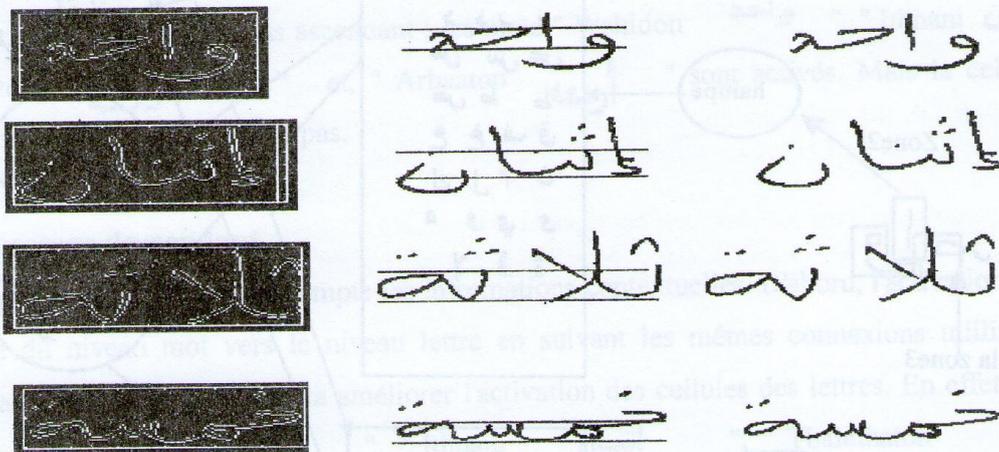


Figure 10 : Résultats expérimentaux

## 5. Conclusion

Après avoir présenté notre méthode, nous avons pu souligner les contributions majeures du système de reconnaissance et justifier le choix de l'architecture des réseaux de neurones transparents (dits aussi à représentation locale de l'information). En effet, les réseaux de neurones à représentation distribuée de l'information (ou à boîte noire) nécessitent une durée d'apprentissage élevée ainsi qu'une large base de données [Sankar96, Chan92, Alpayding92, Braham90]. L'explication du comportement du réseau de neurones et l'analyse de la cause de l'erreur sont tous les deux des tâches difficiles pour cette gamme de réseaux. Dans les réseaux de neurones à représentation locale de l'information, chaque neurone correspond à un concept. Ce codage est utilisé dans les réseaux de neurones sans phase d'apprentissage car il est relativement facile d'établir les connexions entre neurones en tenant compte de leur interprétation. Ce sont des réseaux de neurones déjà appris. Ils modélisent directement les connaissances de l'expert.

Les premiers tests de notre système sont faits sur un ensemble de 27 mots manuscrits pour un monoscripteur et 27 mots imprimés. Ces mots isolés forment le premier pas vers la reconnaissance des montants littéraux arabes. Afin d'effectuer une évaluation robuste du système, une base de données de 10000 chèques arabes multiscripteurs (100 scripteurs) est en cours d'élaboration.

## Bibliographies

- [Alpayding92] Alpayding E, Multiple Neural Networks and Weighted Voting, IEEE IAPR 11 Vol II, , 1992, pp 29-32.
- [Amara99] Ben Amara N.E, Utilisation des modèles de Markov Cachés Planaires en reconnaissance de l'écriture Arabe Imprimée; thèse de doctorat en Génie Electrique, Ecole National d'Ingénieurs de Tunis, 1999.
- [Avila96] Avila M, Optimisation de Modèles Markovienne pour la reconnaissance de l'écrit, thèse de doctorat en sciences appliquées spécialité informatique, automatique et systèmes, Université de Rouen, 1996.

- [Belaïd96] Belaïd A & Saon G, Use of Markov Processes in Writing Recognition, France, 1996.
- [Bose92] Bose C.B, Kuo S, Connected and Degraded Text Recognition Using Hidden Markov Model, IEEE IAPR 11 Vol II, 1992, pp116-119
- [Braham90] Braham R, Hamblen J.O, The Design of a Neural Network with a Biologically Motivated Architecture, IEEE TRANSACTION ON NEURAL NETWORKS, Vol.1, No.3, 1990, pp 251-262.
- [Chan92] Chan C.K & Sanaler M.B, A complete shape Recognition System Using the Hough transform and Neural Network, IEEE IAPR 11 Vol II, 1992, pp 21-24.
- [Chen96] Chen S, Haralick R.M & Phillips I.T, Extraction of text lines and text blocks on document images based on statistical modeling, Washington, 1996.
- [Côté98] Côté M, Lecolinet E, Cheriet M, Suen C.Y, Automatic reading of cursive scripts using a reading model and perceptual concepts, International Journal on Document Analysis And Recognition , Volume1, Number1, 1998, pp 3 - 17.
- [Lam88] Lam L, Suen C.Y, Structural Classification and Relaxation Matching of totally unconstrained Handwritten Zip-code numbers, Pattern recognition, vol 21, 1988, pp 19 - 31.
- [Mahjoub99] Mahjoub M.A, Application des Modèles de Markov Cachés stationnaire et non stationnaire à la reconnaissance en ligne de l'écriture arabe, thèse de doctorat en Génie Electrique, Ecole National d'Ingénieurs de Tunis, 1999.
- [Miclet84] Miclet L, Méthodes structurelles pour la reconnaissance des formes, Germain, 1984.
- [Miled98] Miled H, Stratégie de résolution en reconnaissance de l'écriture semi-cursive: Application aux manuscrits arabes, thèse de doctorat en sciences appliquées, Université de Rouen, 1998.
- [Nedeljković92] Nedeljković V & Milosavljević M, On the influence of the training Set Data Preprocessing on Neural Network Training, IEEE IAPR 11 Vol II, 1992, pp 33-36.
- [Sankar96] Sankar K.Pal, Pradip K.Sriamani, Neurocomputing Motivation, Models, and Hybridization, Neural Computing, IEEE Computational Science and Engineering, 1996, pp 24 - 44.
- [Saon96] Saon G & Belaïd A, Cursive Word Recognition Using a Random Field Based Hidden Markov Model, France, 1998.
- [Srihari93] Srihari S.N, Recognition of Handwritten and machine printed text for postal address interpretation, Pattern Recognition Letters14, Nort-Holland, 1993, pp219-302.
- [Tang92] Tang T.LY.Y, Suen S.C, L.Y.Fang and A.J.Jennings, A Structural Adaptive Neural Tree for the Recognition of large Character Set, IEEE IAPR 11 Vol II, 1992, pp187-190.