

## Proposition d'un classificateur en cascade: application à la reconnaissance de polices de caractères rares

Jean Duong, Robert Sabourin, Hubert Emptoz

### ► To cite this version:

Jean Duong, Robert Sabourin, Hubert Emptoz. Proposition d'un classificateur en cascade: application à la reconnaissance de polices de caractères rares. 8ème colloque international francophone sur l'écrit et le document (CIFED'2004), Jun 2004, La Rochelle, France. pp.231-236. sic\_00001172

HAL Id: sic\_00001172

[https://archivesic.ccsd.cnrs.fr/sic\\_00001172](https://archivesic.ccsd.cnrs.fr/sic_00001172)

Submitted on 6 Dec 2004

**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.

# Proposition d'un classificateur en cascade: application à la reconnaissance de polices de caractères rares

Jean Duong<sup>1</sup> – Robert Sabourin<sup>2</sup> – Hubert Emptoz<sup>1</sup>

<sup>1</sup> Laboratoire d'InfoRmatique en Images et Systèmes d'information (LIRIS)  
CNRS FRE 2672 - Institut National des Sciences Appliquées (INSA) de Lyon  
20 avenue Albert Einstein, 69621 Villeurbanne Cedex

<sup>2</sup> Laboratoire d'Imagerie, Vision et Intelligence Artificielle (LIVIA)  
Ecole de Technologie Supérieure (ETS), 1100 rue Notre Dame Ouest  
H3C 1K3 Montréal (Québec, Canada)

[duong,emptoz]@rfv.insa-lyon.fr, robert.sabourin@etsmtl.ca

**Résumé :** *Nous présentons une contribution en reconnaissance de formes motivée par l'étude de polices de caractères rares. Le problème s'est posé à nous lors de l'examen du corpus des inventaires-sommaires des Archives Départementales de Charente Maritime : certaines indications essentielles sont consignées dans un format typographique inusité de nos jours et pour lequel les logiciels de reconnaissance de caractères actuels sont tenus en échec. Nous proposons donc un outil générique de séparation et de classification de formes. Il s'agit d'une architecture de classificateur en cascade construite sur l'analyse de matrices de confusion intermédiaires.*

*Nous testons ce concept sur deux bases de données étiquetées : le corpus académique de chiffres manuscrits MNIST et une collection de chiffres imprimés extraits par nos soins d'images de documents d'archives. Les spécificités de chacune de ces bases sont décrites et les résultats obtenus sont commentés. Nous proposons enfin une série de développements ultérieurs possibles sur notre classificateur en cascade.*

**Mots-clés :** Reconnaissance de caractères, chiffres manuscrits et imprimés, combinaison de classificateurs en cascade, analyse en composantes principales, règle de classification du plus proche voisin, matrice de confusions.

## 1 Introduction

La numérisation des collections de documents anciens représente des enjeux sociétaux et culturels certains, mais le traitement des corpus d'images ainsi constitués pose une multitude de problèmes techniques. C'est dans ce contexte que nous nous sommes intéressés aux prises de vues réalisées sur les inventaires-sommaires des Archives Départementales de Charente Maritime. Ces documents constituent un corpus d'instruments de recherche (c'est à dire des méta-documents) utiles aux archivistes puisqu'ils recensent et commentent de nombreux actes officiels (décisions de justice, actes notariés, décrets civils ou religieux, etc.) remontant jusqu'au seizième siècle. Un ensemble de 209 images capturées forment une base que nous



FIG. 1 – Exemples de chiffres imprimés dans une police de caractères ancienne (extraits des inventaires-sommaires des Archives Départementales de Charente Maritime).

noterons IS-AD17 par la suite et dont l'élaboration s'inscrit dans le cadre d'un projet préliminaire sur la numérisation de la totalité des archives départementales puis nationales.

Lors de l'examen d'IS-AD17 en vue de son indexation automatique nous avons noté la présence d'une police de caractères ancienne et tombée en désuétude à l'heure actuelle. Ce problème est de plus en plus fréquent dans le contexte de la numérisation du patrimoine écrit. Il apparaît dans toutes sortes de collections d'imprimés; des plus anciens à des productions du vingtième siècle. Dans le cas du corpus IS-AD17, le format typographique atypique constaté concerne certains champs d'information comportant principalement des indications chiffrées (exemples en figure 1) et tient en échec les logiciels de reconnaissances de caractères commerciaux. Nous avons donc étudié la séparabilité et la reconnaissance des caractères concernés en faisant appel au paradigme de l'analyse des données multivariées.

Lorsque se pose le problème d'analyser des données multidimensionnelles, la question du choix des méthodes de traitement devient inévitable. Les développements en reconnaissance de formes ont donné lieu à quantité d'outils : systèmes à base de règles, paradigme flou, réseaux neuronaux, machines à vecteurs de support, etc. Il est difficile d'effectuer un choix *a priori* entre toutes ces familles de procédés. Et inversement, il n'est pas envisageable de les essayer toutes.

Nous proposons une architecture de classificateur permettant de revenir aux sources de l'analyse factorielle classique. Cette branche des statistiques présente une maturité et une légitimité appuyées par un cadre mathématique solide ainsi que par une mise en pratique que nous pourrions qualifier d'ancienne (par exemple en sciences humaines ou en sciences de la vie pour ne citer que deux domaines d'application).

Nous présentons l'architecture et le fonctionnement du classificateur proposé (partie 2) [DUO 03]. Puis nous introduisons deux jeux de données issus de la problématique de reconnaissance d'images de documents (partie 3). Nous éprouvons le classificateur sur ces bases et nous exposons les résultats (partie 4). Enfin, nous proposons quelques perspectives de développements ultérieurs (partie 5).

## 2 Classificateur

Considérons  $\mathcal{X}$  un ensemble fini de vecteurs de formes. Notons  $n = |\mathcal{X}|$  et  $d$  le nombre de caractéristiques prises en compte, c'est à dire la dimensionnalité de l'espace de description des formes. Classiquement, les descripteurs seront supposés de type réel. Ce choix est courant en reconnaissance de formes puisqu'il autorise des traitements numériques et analytiques directement empruntés aux mathématiques appliquées. Ainsi nous avons  $\mathcal{X} \subseteq \mathbb{R}^d$ . De plus,  $\mathcal{X}$  est divisé en  $k$  classes.

$$\mathcal{X} = \bigcup_{i=0}^{k-1} \mathcal{X}_i$$

Nous disposons d'une partition de  $\mathcal{X}$  en un ensemble d'apprentissage  $\mathcal{X}_L$  qui servira à entraîner notre classificateur et un ensemble de test  $\mathcal{X}_T$  qui permettra de l'éprouver. Nous exigeons que chacune des  $k$  classes soit représentée dans  $\mathcal{X}_L$  comme dans  $\mathcal{X}_T$ . (L'omission d'une classe lors de l'entraînement pénalisera la précision du classificateur résultant. Et réciproquement, l'absence de représentants pour une ou plusieurs classes peut occulter certains phénomènes de confusion.) Nous scindons encore  $\mathcal{X}_L$  en deux parties,  $\mathcal{L}$  et  $\mathcal{V}$  que nous appelons respectivement **base de construction** et **base de validation**. La subdivision de  $\mathcal{X}_L$  sera telle que chaque classe  $\mathcal{X}_i$  (avec  $i \in \llbracket 0..k-1 \rrbracket$ ) possèdera des représentants dans  $\mathcal{L}$  et  $\mathcal{V}$ . Enfin, nous posons deux seuils  $T_R \in [0, 1]$  et  $T_C \in [0, 1]$  dont nous donnerons la signification dans la suite de cet exposé.

### 2.1 Apprentissage

#### 2.1.1 Etape initiale

Nous effectuons une première analyse en composantes principales (ACP)[SAP 90] sur les individus **préalablement centrés** de  $\mathcal{L}$ . Notons  $\bar{\mathcal{L}}$  l'individu moyen de  $\mathcal{L}$ . Nous définissons ainsi une base  $\mathcal{B}'$  de  $\mathbb{R}^d$ , constituée des vecteurs propres issus de la matrice de covariance de  $\mathcal{L}$ , triés par ordre décroissant des valeurs propres associées. D'où le repère  $\mathcal{R}' = (\bar{\mathcal{L}}, \mathcal{B}')$ . Nous exprimons les formes (centrées) de  $\mathcal{L}$  dans  $\mathcal{R}'$ . Nous translatons les vecteurs de  $\mathcal{V}$  par  $-\bar{\mathcal{L}}$  puis les projetons dans  $\mathcal{R}'$ .

Nous classons ensuite les vecteurs de  $\mathcal{V}$  (exprimés dans  $\mathcal{R}'$ ) en appliquant la règle du plus proche voisin au sens de la distance euclidienne et en utilisant les individus de  $\mathcal{L}$  (également exprimés dans  $\mathcal{R}'$ ) comme prototypes. Nous construisons alors  $Conf(\mathcal{V})$ , la **matrice de confusion** sur  $\mathcal{V}$ . Rappelons [DUD 01] qu'il s'agit d'un tableau carré dans lequel le contenu de la cellule  $(i, j)$  indique le nombre de vecteurs de classe  $i$  affecté d'une étiquette  $j$  par le classificateur. Dans notre cas,  $Conf(\mathcal{V}) \in \mathcal{M}_k(\mathbb{R})$ . Nous normalisons  $Conf(\mathcal{V})$  par rapport à ses colonnes.

#### 2.1.2 Itération

Nous scrutons chaque colonne  $j$  (avec  $j \in \llbracket 0..k-1 \rrbracket$ ) de la matrice normalisée  $Conf(\mathcal{V})$ .  $T_R$  est le **taux de reconnaissance minimum exigé** pour chaque classe : lorsque la cellule  $(j, j)$  réalise  $Conf(j, j) < T_R$ , cela signifie que moins de  $100 \times T_R\%$  des individus étiquetés  $j$  appartiennent réellement à la classe  $j$ .

Soit une classe  $j$  pour laquelle  $T_R$  n'est pas atteint. Nous repérons les positions  $i$  (avec  $i \in \llbracket 0..k-1 \rrbracket$ ) pour lesquelles  $Conf(i, j) > T_C$ .  $T_C$  est le **taux de confusion maximum acceptable**. Notons  $I_j$  l'ensemble des indices  $i$  ainsi trouvés. Lorsque  $I_j \neq \emptyset$ , nous considérons les corpus  $\mathcal{L}_{I_j} = \bigcup_{i \in I_j \cup \{j\}} \mathcal{X}_i \cap \mathcal{L}$  et  $\mathcal{V}_{I_j} = \bigcup_{i \in I_j \cup \{j\}} \mathcal{X}_i \cap \mathcal{V}$ . Il s'agit ici des individus d'origine (avant centrage et projection sur la nouvelle base).

Nous réitérons l'ACP avec  $\mathcal{L}_{I_j}$  et  $\mathcal{V}_{I_j}$  comme ensembles de construction et de validation respectivement, exprimés dans le repère canonique  $\mathcal{R}$ . Cela revient à reconduire la démarche précédente en ne considérant que les classes indicées par  $I_j \cup \{j\}$ . Nous parlerons des **classes adhérentes** à  $j$ . A chaque indice  $j$  est associé une liste  $N_j$  de classes à analyser en particulier lors d'une nouvelle récursion telle que

$$N_j = \begin{cases} I_j \cup \{j\} & \text{si } I_j \neq \emptyset \\ \emptyset & \text{si } I_j = \emptyset \end{cases}$$

Nous réalisons ainsi une **combinaison de classificateurs** simples. Notre démarche est une généralisation de l'approche consistant à examiner systématiquement un nombre fixé *a priori* de classes adhérentes. Ainsi, nous élargissons la procédure de L. Prévost [PRE 03] qui traite les couples de classes adhérentes.

### 2.2 Prédiction

Soit  $X \in \mathcal{X}_T$  un vecteur de la base de test.  $X$  est translaté puis projeté selon l'individu moyen et les composantes principales de l'analyse globale sur l'ensemble des  $k$  classes. Puis  $X$  est étiqueté par rapport aux prototypes de  $\mathcal{L}$  exprimés dans  $\mathcal{R}'$  via la règle du plus proche voisin. Soit  $j$  l'étiquette inférée. Si la liste  $N_j$  associée est vide, alors l'étiquette est acceptée. Sinon,  $X$  est soumis à une nouvelle classification mettant en jeu les classes indiquées dans  $N_j$ .

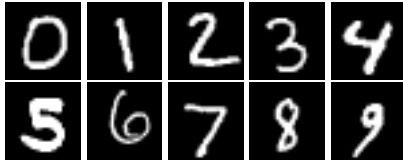


FIG. 2 – Exemples d’images de chiffres manuscrits extraites de MNIST.

### 3 Jeux de données

#### 3.1 Chiffres manuscrits : la collection MNIST

Le corpus *Modified NIST* ou MNIST [LEC 98]<sup>1</sup> est une collection de chiffres manuscrits. Chaque chiffre est associé à une image de taille  $28 \times 28$  en niveaux de gris sur 256 valeurs (exemples en figure 2). Pour notre application, le vecteur de forme correspondant consigne les valeurs de luminance pour chaque pixel. Ainsi, un chiffre est représenté par un vecteur numérique de taille  $28 \times 28 = 784$ . La répartition par classe des bases de MNIST est donnée par le tableau 1.

Classe	Apprentissage	Test
0	5923	980
1	6742	1 135
2	5958	1 032
3	6131	1 010
4	5842	982
5	5421	892
6	5918	958
7	6265	1 028
8	5851	974
9	5949	1 009

TAB. 1 – Distribution des individus par classes dans la collection MNIST.

Enfin MNIST est en réalité la réunion de deux sous-ensembles issus des bases **SD-1** et **SD-3** proposées par le consortium **NIST** (*National Institute of Standards and Technology*). Les échantillons manuscrits de SD-1 et SD-3 ont été collectés respectivement parmi une population de lycéens et un groupe d’employés de bureaux du services des recensements. D’où une différence significative dans la qualité des écritures. Les formes de SD-3 sont réputées plus aisées à reconnaître par classification automatique. La collection MNIST est un réarrangement des ces échantillons. Elle est organisée en une base d’apprentissage et une base de test de 60 000 et 10 000 individus chacune.

#### 3.2 Chiffres imprimés et police de caractères ancienne

Nous avons extrait une base de travail originale en examinant les images d’IS-AD17. Nous nous intéressons aux dates et périodes de référencement utilisées dans les articles de cette

collection. Ces entités sont le plus souvent exprimées sous forme de séquences de chiffres. Elles sont formatées selon une police de caractères non reconnue par les logiciels de reconnaissance de caractères courants. En effet, les formes imprimées (figure 1) sont exagérément épaisses et induisent de fréquentes confusions. Nous avons collecté la totalité des chiffres de la collection IS-AD17 relevant de cette police de caractères. Le tableau 2 montre la répartition des formes recensées.

Classe	Effectif	Classe	Effectif
0	57	5	58
1	356	6	173
2	71	7	251
3	56	8	75
4	54	9	54
		Total	1 205

TAB. 2 – Distribution des chiffres imprimés relevant de la police ancienne dans la collection IS-AD17.

## 4 Résultats expérimentaux

### 4.1 MNIST

#### 4.1.1 Réduction de dimensionnalité

Afin de limiter les coûts calculatoires engendrés par les opérations de prédiction, il est possible de procéder à une **réduction de dimensionnalité** des formes. Il suffit de ne conserver qu’un nombre réduit  $d'$  de composantes principales ( $d' < d$ ) à chaque récursion. Par effet de changement de base, les vecteurs prototypes présenteront donc  $d'$  composantes au lieu des  $d$  initiales. Il en ira de même pour les individus du corpus de validation.

Dans la pratique, nous fixons un seuil minimal de variance exprimée  $\nu \in ]0; 1]$ . Pour chaque sous-problème rencontré, nous arrangeons les vecteurs propres de la matrice de covariance par ordre décroissant des valeurs propres. Notons  $v_k$  (resp.  $V_k$ ) la  $k$ -ième valeur (resp. le  $k$ -ième vecteur) propre ainsi trié. Nous conservons les  $n_\nu$  premiers vecteurs propres pour constituer une matrice de changement d’espace en définissant

$$n_\nu = \min \left\{ n \in \llbracket 1..d \rrbracket \left| \begin{array}{l} \sum_{k=1}^n v_k \\ d \\ \sum_{k=1}^n v_k \end{array} \geq \nu \right. \right\}$$

Autrement dit,  $n_\nu$  est le nombre minimal de composantes principales à conserver afin de restituer  $100\nu\%$  de la variance expliquée sur les données de construction. Il est propre à chaque niveau de récursion et à chaque jeu de données.

Le tableau 3 consigne les résultats des expériences menées sur la base MNIST. Nous avons scindé la base d’apprentissage en deux ensembles de 30 000 individus chacun dédiés à la construction et la validation. Nous avons ensuite fixé un

<sup>1</sup> <http://yann.lecun.com/exdb/mnist/index.html>

seuil de reconnaissance minimal à  $T_R = 0.95$ , et un seuil de confusion tolérable à  $T_C = 0.01$ . Nous examinons l'effet de la réduction de dimensionnalité pour différentes valeurs de  $\nu$  en mesurant le taux de reconnaissance sur la base de test à 10 000 individus.

Variance exprimée ( $\nu$ )	Reconnaissance (%)
0.80	95.78
0.85	95.89
0.90	96.06
0.95	96.03
1.00	96.09

TAB. 3 – Performances réalisées sur MNIST pour différentes parts de variance exprimée.

#### 4.1.2 Liaison des seuils

Afin de faciliter l'étude du classificateur, nous choisissons de nous intéresser à  $T_R$  uniquement.  $T_C$  est déterminé automatiquement par la formule heuristique.

$$T_C = \frac{1-T_R}{k-1} \text{ où } k \text{ est le nombre de classes du problème.}$$

Nous effectuons plusieurs simulations mettant en jeu différentes valeurs de  $T_R$ . Notons qu'en conservant toute la variance exprimée lors de l'ACP, les variations de  $T_R$  sont sans effet sur les performances constatées. Ce propos est illustré par les résultats du tableau 4, qui par ailleurs montre une situation de sur-apprentissage : pour  $T_R = 1.00$ , le taux de reconnaissance exigé en validation est tel que le classificateur construit perd une partie de son pouvoir de généralisation face à la base de test. D'où un taux de reconnaissance en régression.

Seuil de validation ( $T_R$ )	Reconnaissance (%)
0.80	96.09
...	...
0.95	96.09
1.00	94.73

TAB. 4 – Performances réalisées sur MNIST. Le taux de variance exprimée lors de l'analyse en composantes principales est fixé à 1.00

Nous modulons le taux de variance exprimée  $\nu$  et réalisons autant de jeux d'expériences. Les résultats sont reportés dans le tableau 5. Les performances réalisées dépassent celles des classificateurs purement linéaires confrontés à MNIST. Le tableau 6 (détails dans [LEC 98]) donne les performances réalisées par divers classificateurs optimisés sur MNIST.

## 4.2 IS-AD17

Les expériences réalisées sur MNIST donnent des résultats encourageants. Nous avons donc reporté le protocole expérimental sur le corpus des chiffres imprimés de IS-AD17. En effet, MNIST constitue indiscutablement une base de travail reconnue pour la communauté de recherche en

	0.80	0.85	0.90	0.95	1.00
0.80	96.09	96.20	96.17	96.04	96.09
...	...	...	...	...	...
0.95	96.09	96.20	96.17	96.04	96.09
0.97	95.52	95.75	96.04	96.07	96.09

TAB. 5 – Performances réalisées sur MNIST pour différents seuils de reconnaissance exigée (en lignes) et taux de variance exprimée (en colonnes).

Classificateur	Reconnaissance (%)
Réseau neuronal linéaire à une couche	88.0
Réseau neuronal linéaire à une couche avec correction d'inclinaison	91.6
Classificateurs linéaires combinaisons deux à deux	92.4
Plus proches voisins (distance euclidienne)	95.0
Plus proches voisins (distance euclidienne) avec correction d'inclinaison	97.6
Analyse en composantes principales (40 composantes principales) et classificateur quadratique	96.7
Plus proches voisins et distance tangente 16	98.9
SVM à noyau polynômial (degré 4)	98.9
SVM à noyau polynômial (degré 5) (ensemble réduit)	99.0
SVM à noyau polynômial (degré 9) avec base augmentée	99.2

TAB. 6 – Performances réalisées sur la base MNIST par divers classificateurs (linéaires et quadratiques) et machines à vecteurs de support.

reconnaissance de formes et permet de situer notre classificateur en cascade par rapport à d'autres outils. Mais les chiffres imprimés que nous avons isolés dans IS-AD17 présentent un intérêt particulier dans le cadre de la numérisation des collections. La police de caractères utilisée (et aujourd'hui oubliée) est employée dans différents fonds contemporains à IS-AD17. Elle est rencontrée dans les inventaires sommaires d'archives relevant de diverses juridictions, pour ne citer qu'une famille d'exemples. Ainsi, nous avons l'occasion d'éprouver notre classificateur en cascade sur des données "de terrain".

Nous avons choisi de constituer une base d'apprentissage (resp. de validation) comprenant 10 individus de chaque classe. Le reste des vecteurs, soit 1005 formes, est utilisé pour les tests. Les images des chiffres sont normalisées à la taille de  $20 \times 20 = 400$  pixels. Comme pour MNIST, les individus de représentation sont des vecteurs de luminance,

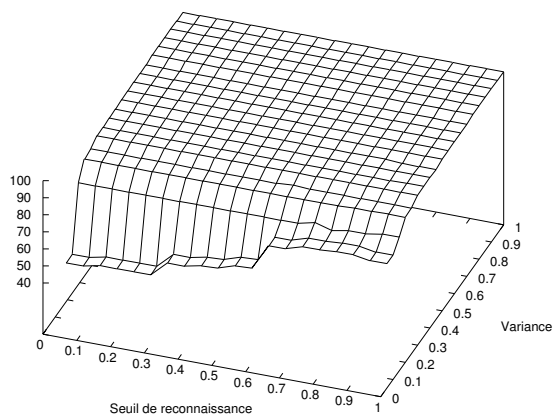


FIG. 3 – Surface de réponse pour les expériences réalisées sur IS-AD17.

c'est à dire ici des vecteurs de 400 composantes. Les performances relevées sont synthétisées graphiquement par la surface de réponse de la figure 3. Deux remarques s'imposent.

Tout d'abord, les taux de reconnaissance sont significativement plus élevés que dans le cas des chiffres manuscrits de MNIST, pour les mêmes valeurs de paramétrage du classificateur. La classification parfaite est même atteinte aisément. Or nous utilisons moins de vecteurs dédiés à la construction et la validation. Cette différence est d'autant plus marquante que la dimensionnalité de l'espace de représentation des individus est plus importante. Signalons encore que notre base de test est largement plus volumineuse relativement à ses homologues servant l'apprentissage. Cette robustesse peut s'expliquer au moins en partie par une certaine régularité morphologique des formes de chaque classe. Cependant, cette stabilité n'est pas absolue car de nombreux caractères sont dégradés. Toutefois, l'impact de ces distorsions est atténué par l'exhubérance typographique propre à la police de caractères : chaque classe de chiffres possède des traits physiques exclusifs.

De plus, la figure 3 montre que pour de faibles parts de variance exprimée et des taux de reconnaissance exigés modestes, les performances restent appréciables. En fait, le seuil de reconnaissance en validation semble jouer un rôle secondaire par rapport au taux d'inertie. Ce constat conforte la conjecture d'un bon pouvoir de généralisation de notre classificateur pour le problème de la séparation des chiffres imprimés de la police de caractères ancienne utilisée dans IS-AD17.

## 5 Conclusion

### 5.1 Mise en perspective

Nous avons proposé une architecture de classificateurs en cascade entièrement conditionnée par les matrices de confu-

sion intermédiaires. Cette initiative constitue notre réponse pratique à un besoin particulier en reconnaissance de formes : la classification de caractères numériques imprimés dans une police ancienne et non reconnue par les logiciels OCR actuels. Elle touche donc à la problématique en plein essor de la numérisation des collections patrimoniales<sup>2</sup>. Nous avons traité un échantillon d'images de chiffres issues de prises de vues sur plusieurs pages du corpus IS-AD17. Lesdits chiffres relèvent d'un format typographique caractéristique des inventaires-sommaires rédigés pour les archives administratives des différents départements de France. Ces documents suscitent un intérêt croissant de par leur valeur sociétale, comme en témoignent de récents travaux [ALL 03]. Ainsi, notre étude représente plus qu'un simple exercice de style académique.

### 5.2 Développements ultérieurs

Notre construction s'est révélée féconde au niveau applicatif, tant sur des bases de données conventionnelles (par exemple les Iris de Fisher [DUO 03] dont nous n'avons pas discuté pour des raisons de contraintes éditoriales) que des compilations plus proches de la vie courante (illustrées par MNIST et IS-AD17) en termes d'enjeux et de difficulté technique. Cependant, cette dernière demeure loin d'être épuisée en matière de possibilités de développement.

En premier lieu, des améliorations dans l'implémentation sont envisageables, tant au niveau des structures de données que de l'algorithmique. Bien qu'il s'agisse d'un aspect technique primordial de toute réalisation logicielle, l'optimisation de code n'est pas une préoccupation majeure de cet exposé. Nous préférons la laisser en suspens et y consacrer ultérieurement des développements spécifiques.

Une autre ouverture peut être faite quant au choix des classificateurs fondamentaux. Nous avons opté pour la règle du plus proche voisin assortie de l'analyse en composantes principales. La motivation sous-jacente était multiple. Il s'agissait d'adopter un paradigme simple afin d'éviter des obstacles conceptuels et des difficultés de mise en œuvre exagérées.

Enfin, les différents paramètres (tels les taux de reconnaissance minimal et de confusion maximal acceptables) peuvent être déterminés par rapport à des heuristiques restant à définir. Ces règles de décision tiendront compte de la taille des échantillons de données, du nombre de classes à discriminer, de la profondeur de récursion par rapport au problème principal, *etc.* Nous envisageons également d'effectuer un examen statistique des effets et interaction de ces facteurs évalués *a priori* par l'utilisateur. Cette réflexion sera fondée sur les paradigmes des plans d'expériences et de l'analyse de la variance [DRO 97, SCH 98, ALE 99]. Des expériences supplémentaires sont en cours afin de collecter suffisamment de données numériques pour mener une telle étude.

<sup>2</sup>Les rencontres sur la numérisation du patrimoine de La Rochelle en sont une illustration.

## Références

- [ALE 99] ALEXIS J., ALEXIS P., *Pratique industrielle des plans d'expériences (La qualité à moindre coût : l'approche Taguchi)*, AFNOR, Paris (France), 1999.
- [ALL 03] ALLIER B., Contribution à la numérisation des collections : apport des contours actifs, PhD thesis, Institut National des Sciences Appliquées (INSA) de Lyon, Villeurbanne (France), 28 November 2003.
- [DRO 97] DROESBEKE J.-J., FINE J., SAPORTA G., *Plans d'expériences (Applications à l'entreprise)*, Technip, Paris (France), 1997.
- [DUD 01] DUDA R. O., HART P. E., STORK D. G., *Pattern Classification*, Wiley Interscience, 2001.
- [DUO 03] DUONG J., Proposition d'un classifieur récursif, rapport n° RR01, July 2003, Laboratoire de Reconnaissance de Formes et Vision (RFV).
- [LEC 98] LECUN Y., BOTTOU L., BENGIO Y., HAFFNER P., Gradient-Based Learning Applied to Document Recognition, *Proceedings of the IEEE*, vol. 86, n° 11, 1998, pp. 2278–2324.
- [PRE 03] PREVOST L., MICHEL-SENDIS C., OUDOT L., MILGRAM M., Combining Model-based and Discriminative Classifiers : application to Handwritten Character Recognition, *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR)*, Edinburgh (Scotland, United Kingdom), 2003, pp. 31–35.
- [SAP 90] SAPORTA G., *Probabilités, Analyse des données et Statistiques*, Technip, Paris (France), 1990.
- [SCH 98] SCHIMMERLING P., SISSON J.-C., ZAÏDI A., *Pratique des plans d'expériences*, Lavoisier Tech & Doc, Paris (France), 1998.