

1. INTRODUCTION

La Reconnaissance des formes (RdF) ou Reconnaissance artificielle des formes (par opposition à la reconnaissance naturelle de formes qui est pratiquée par l'Homme) traite de l'apprentissage du processus de reconnaissance et de la prise de décision automatique. Reconnaître une forme c'est donner à une machine la capacité de donner un nom à cette forme à partir des informations descriptives qui lui ont été fournies par l'environnement. C'est dire que ces méthodes et techniques ont vocation à intervenir dans la quasi-totalité des domaines d'intérêt de l'être humain. Les techniques d'automatisation de la production industrielle, les méthodes de contrôle de la qualité, les méthodes de reconnaissance d'objets dans les images numériques, ou la reconnaissance de sons (parole), ou la reconnaissance de gestes, la conduite de robots. Tous ces champs relèvent relativement des méthodes d'apprentissage et de RdF.

La reconnaissance des formes (RdF) est issue de différentes disciplines qui sont les mathématiques (probabilités et statistiques), les sciences de l'ingénieur, l'informatique et l'intelligence artificielle. C'est à partir des années 60 que la reconnaissance de formes est devenue une discipline spécifique. L'extraordinaire développement des ordinateurs ces dernières années a donné un élan à la RdF en permettant des applications temps réel, en particulier dans le domaine des applications visuelles et auditives. Les procédés d'acquisition tels que caméra, scanner sont très accessibles, ainsi que des ordinateurs à la fois puissants et bon marché. Ils permettent le traitement de nombreuses données en un temps raisonnable comme cela est souvent nécessaire en RdF.

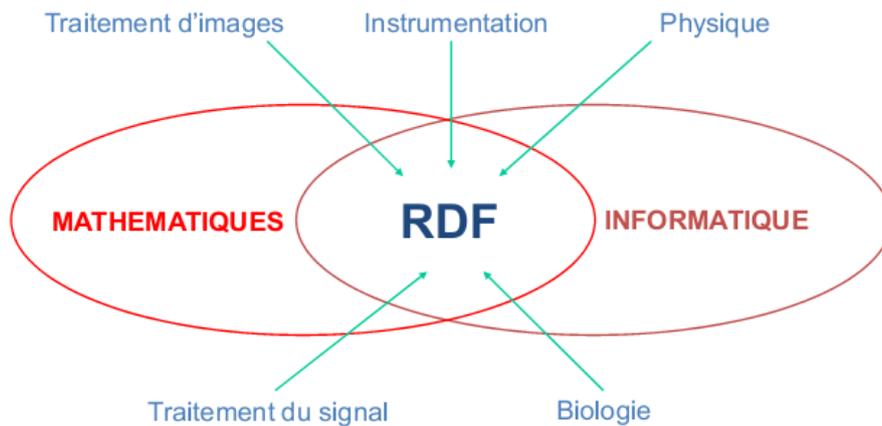
PROBLEMATIQUE

Le problème que cherche à résoudre la reconnaissance des formes est d'associer une étiquette à une donnée qui peut se présenter sous forme d'une image ou d'un signal. Des données différentes peuvent recevoir la même étiquette, ces données sont les réalisations ou les exemplaires de la classe identifiée par l'étiquette.

- Par exemple, le son /a/ prononcé par différents locuteurs conduit à des signaux différents mais ces différences ne sont pas significatives du point de vue de l'identification du son, ces signaux sont des réalisations de la classe /a/.
- De même, l'écriture manuscrite du caractère A varie d'un scripteur à l'autre mais le lecteur identifiera le caractère A pour chacune de ces réalisations.

Des méthodes générales ont été développées en reconnaissance des formes pour extraire automatiquement des informations des données sensibles afin de caractériser les classes de formes (apprentissage) et d'assigner automatiquement des données à ces classes (reconnaissance). La mise en œuvre de ces méthodes générales pour des problèmes particuliers amène à introduire la notion de processus de reconnaissance qui pose la question de l'intégration des méthodes de la reconnaissance de formes dans un système qui a pour but de reconnaître des formes.

Parallèlement aux travaux sur les méthodes de reconnaissance, se développaient le traitement d'image, la vision par ordinateur, et le traitement de la parole. Ces domaines ont focalisé le problème de la reconnaissance sur des données spécifiques, mais par ailleurs ils ont permis de situer la reconnaissance dans un processus plus vaste d'interprétation d'image ou de compréhension de la parole impliquant des niveaux de perception et des connaissances propres au domaine.



La RdF un domaine pluridisciplinaire.

DEFINITION

On désigne par reconnaissance de formes RdF (ou parfois reconnaissance de motifs) un ensemble de techniques et méthodes visant à identifier des motifs à partir de données brutes afin de prendre une décision dépendant de la catégorie attribuée à ce motif.

La RdF est considérée comme une branche de l'intelligence artificielle qui fait largement appel aux techniques d'apprentissage automatique et aux statistiques. Les formes ou motifs à reconnaître peuvent être de nature très variée. Il peut s'agir de contenu visuel (code barre, visage, empreinte digitale...) ou sonore (reconnaissance de parole), d'images médicales (rayon X, EEG, IRM...) ou multi spectrales (images satellitaires) et bien d'autres.

La RdF est défini par divers auteurs comme:

- "Pattern recognition, a field concerned with machine recognition of meaningful regularities in noisy and complex environments" (Duda et Hart 1973)
- "Pattern recognition is a search for structure in data." (Bezdek 1987)

Ces deux définitions laissent apparaître à la fois l'objectif et la difficulté de la tâche à accomplir. La RdF, ainsi définie, est l'apprentissage ou la découverte de structures appelées classes dans un ensemble de données éventuellement perturbées. L'objectif de cet apprentissage est généralement décisionnel suivant qu'il s'agit de déclencher un événement, une action en fonction de l'appartenance de tel ou tel objet à une classe.

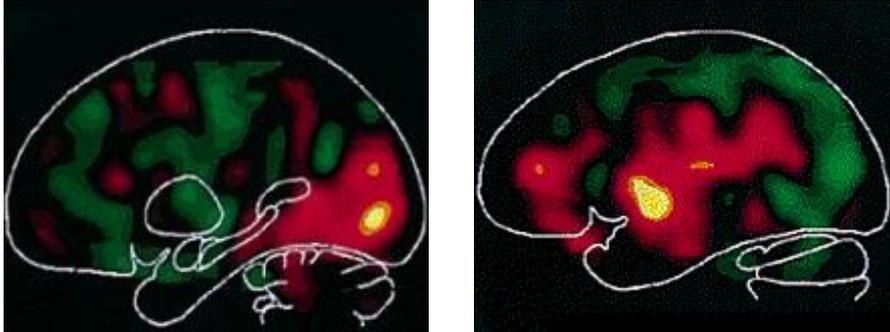
2. EXEMPLES D'APPLICATION DE LA RdF

La reconnaissance de formes est utilisée dans plusieurs domaines d'activités. Parmi ces domaines on peut citer :

2.1 La vision artificielle

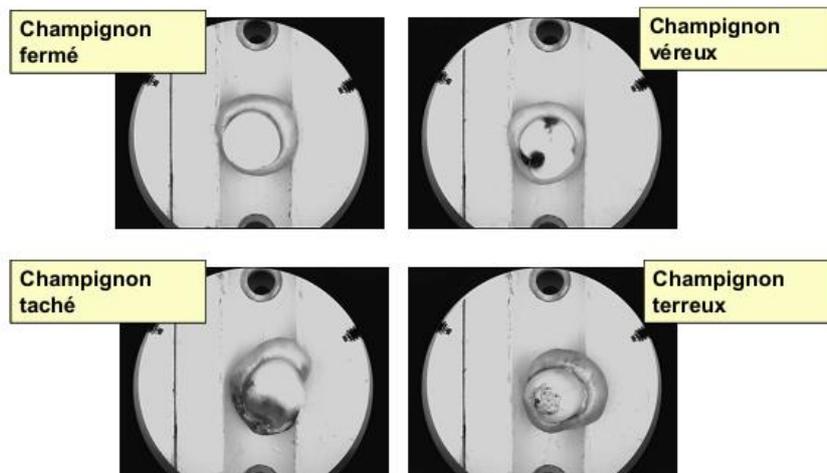
La vision par ordinateur est l'un des principaux champs d'application de la RdF. Actuellement, elle a trouvé une place dominante dans un domaine tel que l'imagerie biomédicale. Par exemple, dans la détection des cellules cancéreuses du cerveau, les images sont acquises en Imagerie par

Résonance Magnétique (IRM). Au cours de l'étape de segmentation, on va rechercher dans les images différentes régions du cerveau : matière blanche, matière grise, liquide cérébro-spinal, etc comme le montre la figure 1. Ensuite, on cherche à détecter d'éventuelles tumeurs bénignes ou malignes à partir de caractéristiques extraites de régions de l'image. Bien évidemment, il ne s'agit là que de fournir une aide au diagnostic validée ultérieurement par le praticien.



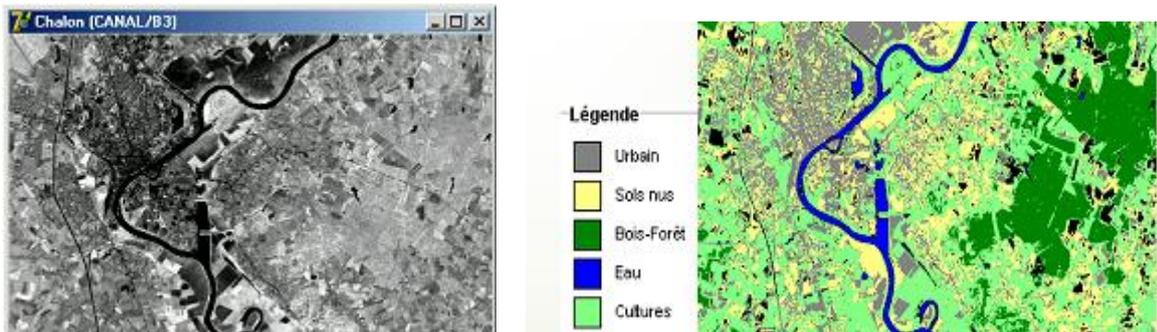
Imagerie médicale.

En industrie, les industriels cherchent à contrôler la qualité de fabrication de leurs produits. La vision artificielle peut être utilisée dans un système de détection de défauts. On peut imaginer diverses pièces industrielles passant sur un tapis roulant et filmées en continu par une caméra. Avec un traitement en temps réel de cette image, on peut, par exemple, vérifier si l'angle formé par deux pièces assemblées est conforme au cahier des charges.



Contrôle de qualité.

En imagerie aérienne, l'interprétation des images aériennes ou de satellites conduit à la surveillance ou aux prévisions agricoles.



Imagerie satellitaire

On peut citer aussi d'autres domaines récents et très actifs comme la reconnaissance des gestes, de l'attitude et même des émotions à partir de séquences vidéos ou audios, la reconnaissance des opinions à partir de tweets.

2.2 La reconnaissance des caractères et des mots

La reconnaissance de l'écriture est une application classique de la RdF. Elle est connue sous le nom d'OCR (Optical Character Recognition) s'il est de l'écriture imprimée. Alors que s'il s'agit de l'écriture manuscrite, la dénomination est ICR (IntelligentCharacter Recognition). Elle vise à transformer un texte manuscrit, sur tout type de support, vers un texte numérique (éditable par un traitement de texte). Le document, sous forme papier par exemple, est préalablement numérisé par un scanner pour en obtenir une représentation sous la forme d'une image. On essaie ensuite d'extraire chacune des lettres présentes sur l'image et d'en déterminer un équivalent "informatique" comme le montre la figure 4 suivante.

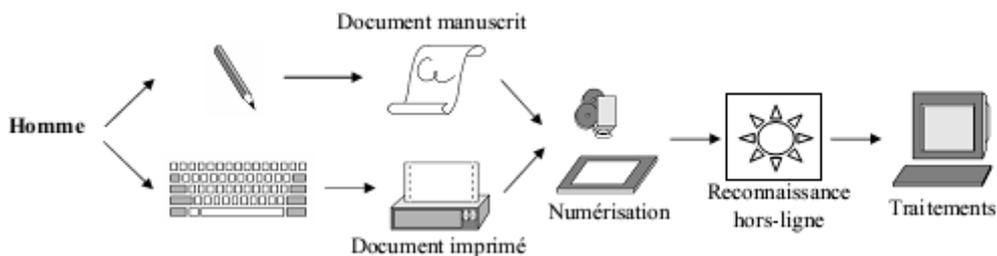
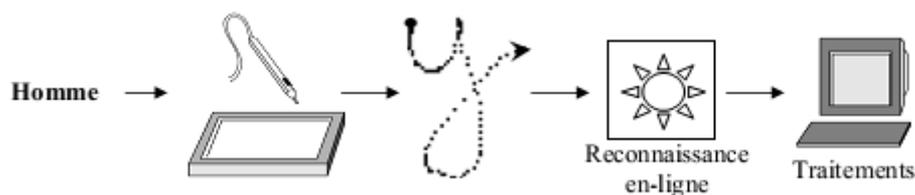


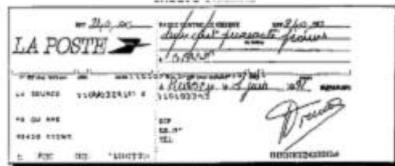
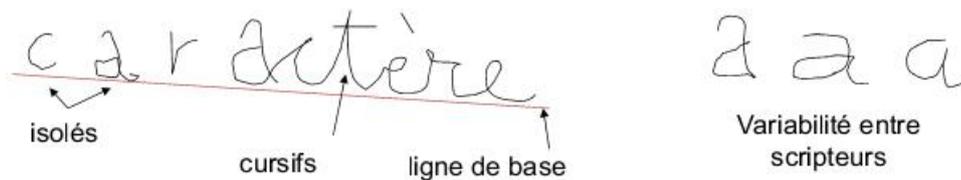
Figure 5. Traitement d'un document papier par numérisation et reconnaissance.



Acquisition d'un signal au moyen d'un stylet et d'une tablette tactile.

L'OCR connaît plusieurs applications pratiques dans différents domaines d'activité parmi lesquels on peut citer :

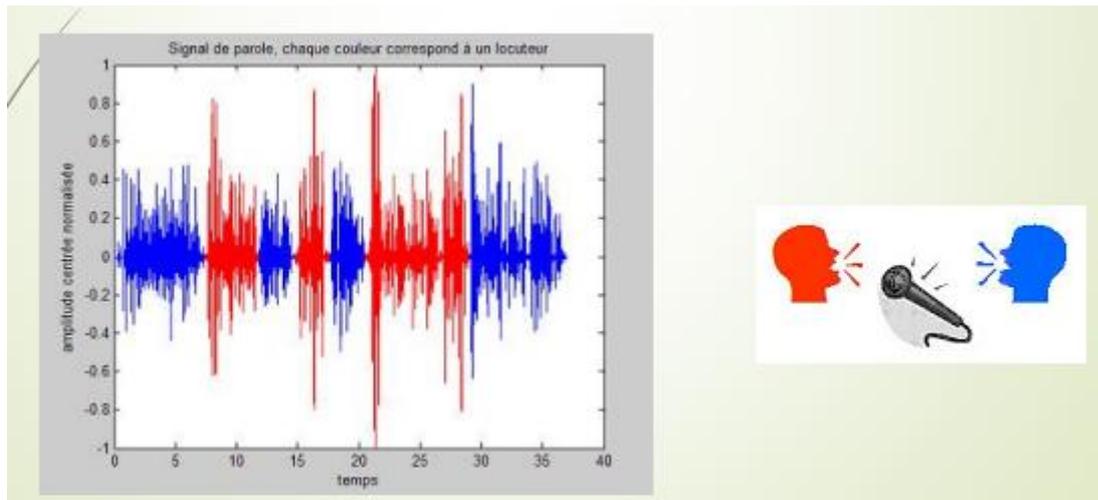
- Les banques et les assurances pour l'authentification de chèques bancaires (correspondance entre montants et libellé d'une part, et correspondance entre l'identité du signataire et sa signature, d'autre part), et la vérification de clauses de contrats pour les assurances.
- La poste pour la lecture des adresses et le tri automatique du courrier.
- Les télécommunications pour l'échange de fichiers informatiques à distance.
 - La police et la sécurité pour la reconnaissance des plaques minéralogiques pour le contrôle routier, l'authentification(et l'identification de manuscrits et l'identification du scripteur (l'identification permet donc de connaître l'identité d'une entité alors que l'authentification permet de vérifier cette identité).
- Les affaires et l'industrie pour la gestion des stocks et la reconnaissance de documents techniques.
- La bureautique pour l'indexation et l'archivage automatique de documents, et pour la publication électronique.
- L'administration pour la reconnaissance de plans cartographiques et la lecture automatique de documents administratifs.



Documents imprimés et manuscrits

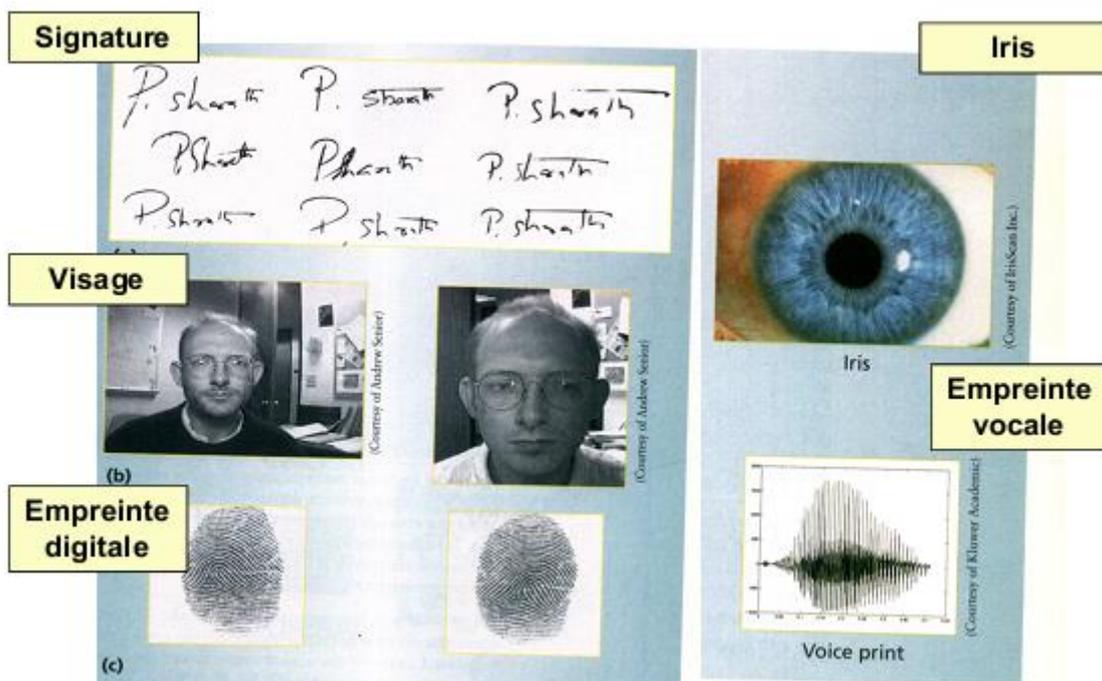
2.3 La reconnaissance vocale

Elle est d'une utilisation relativement récente des techniques de la RdF donnant lieu à de nombreuses recherches. D'une manière générale, l'objectif est de faciliter la communication humaine et Homme/Machine en reconnaissant l'information parlée. La traduction automatique de la parole ou la dictée vocale en sont des applications potentielles au même titre de la commande vocale. Dans ce dernier cas, l'instruction parlée est envoyée deux fois plus rapidement que si elle était dactylographiée. La commande vocale peut également permettre à des personnes handicapées de contrôler des machines en leur parlant.



- (a) Reconnaissance de locuteurs : parmi 10 personnes, qui parle ? et (b)
 Reconnaissance de parole : parmi ces 20 mots, lequel la personne a dit ?

Les systèmes de sécurité utilisent la voix pour identifier ou authentifier des locuteurs. D'autres paramètres biométriques comme les empreintes digitales, la signature, les images (2D et 3D) de la main et du visage, sont également utilisés.



Identification biométrique.

3. LE SCHEMA DE LA RECONNAISSANCE DES FORMES

L'objectif de la RdF est de réaliser des systèmes informatisés qui simulent les activités humaines de perception, de reconnaissance et de compréhension : reconnaissance de l'écrit, de la parole, interprétation de scènes, robotique, reconnaissance des signaux EEG (électroencéphalogramme), ECG (électrocardiogramme). Cela implique aussi une certaine

pluridisciplinarité pour comprendre l'aspect physique des capteurs, les aspects mathématiques de la classification, ceux relatifs à l'informatique.

Le principe de la RdF est de classer les nouvelles formes en utilisant un classifieur qui génère une fonction d'appartenance pour chaque classe. Ainsi la classification d'un nouveau point peut se faire en fonction de la valeur d'appartenance qu'il obtient par rapport à chaque classe.

La RdF est réalisée en deux phases : l'apprentissage à partir des données connues et la classification des nouvelles données. En amont de ces deux phases, une étape de prétraitement est utilisée pour trouver l'ensemble minimal de paramètres informatifs nécessaire à l'établissement de l'espace de représentation. L'ensemble des étapes majeures d'une approche de reconnaissance des formes se décompose comme l'illustre la figure suivante.

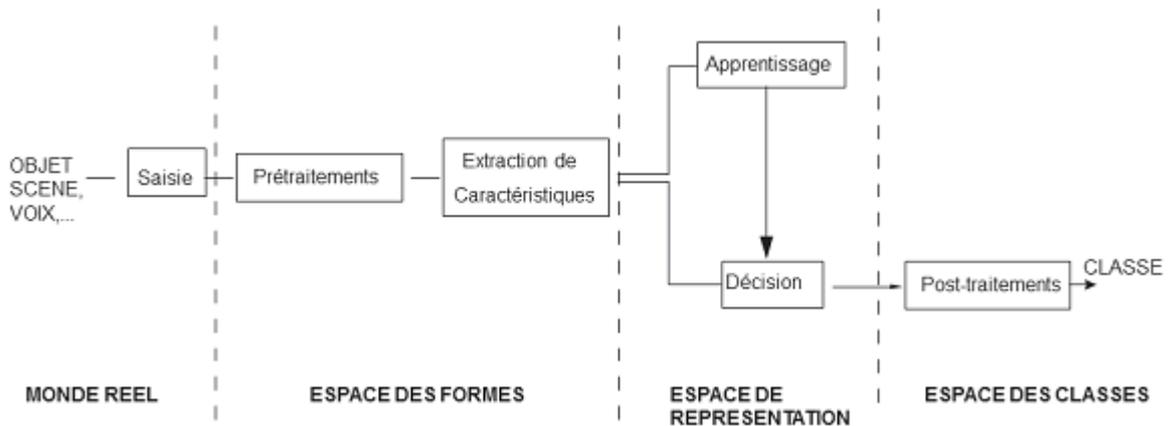


Schéma général d'un système de reconnaissance des formes.

3.1 Saisie (Numérisation)

À partir des informations du monde physique, il s'agit de construire une représentation des données directement manipulable par la machine. Des capteurs (microphone, caméra, instruments de mesure) convertissent les signaux reçus du monde réel en une représentation numérique discrète. L'espace résultant, appelé espace de représentation, a une dimension R très grande lui permettant de disposer du maximum d'informations sur les formes numérisées.

L'échantillonnage est le procédé de discrétisation spatiale d'une image consistant à associer à chaque zone rectangulaire $R(x,y)$ d'une image continue une valeur unique. On parle de sous échantillonnage lorsque l'image est déjà discrétisée et qu'on diminue le nombre d'échantillons.

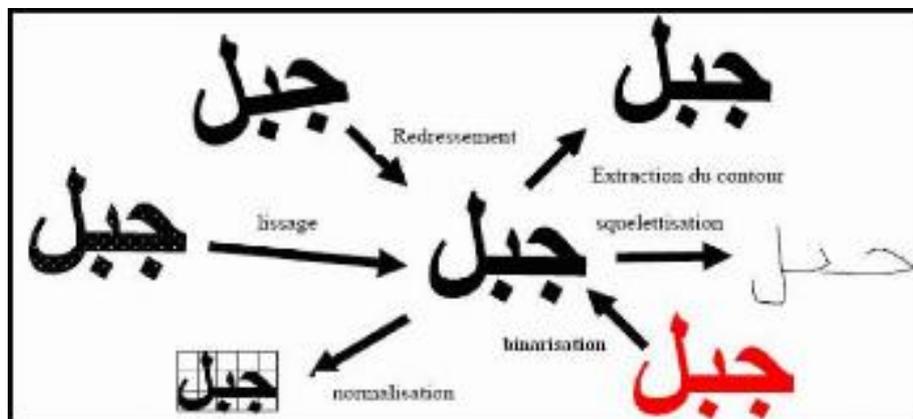
La quantification désigne la limitation du nombre de valeurs différentes que peut prendre $I(x,y)$. Ainsi, une image numérique est une image échantillonnée et quantifiée.

3.2 Prétraitement

Les données brutes issues des capteurs sont les représentations initiales des données à partir desquelles des traitements permettent de construire celles qui seront utilisées pour la reconnaissance. Les données brutes sont bruitées, elles contiennent des informations parasites, elles sont non normalisées et redondantes et elles n'explicitent pas les informations utiles pour la reconnaissance. Ainsi, une étape de prétraitement consiste à sélectionner dans l'espace de représentation l'information nécessaire au domaine d'application. Cette sélection passe souvent par l'élimination du bruit, la normalisation des données, ainsi que par la suppression de la redondance. Le nouvel espace de représentation a une dimension R' très inférieure à R mais demeure un espace de grande dimension et contient des informations encore assez primitives.

Le traitement du signal et le traitement de l'image sont les principales sources pour les méthodes de prétraitement. Ces prétraitements sont utiles pour éliminer des bruits qui peuvent être dus au capteur ou à des interférences avec d'autres sources de signaux (la parole en milieu sonore, l'encre du verso qui traverse le papier et dont la trace est visible sur la feuille du manuscrit, les fonds imagés des chèques, filtrage des bruits hautes fréquences, amélioration des contrastes, séparation la forme du fond, extraction de contours ou de squelettes, modélisation du signal temporel, extraction des primitives, etc).

Par exemple, si la donnée acquise est un caractère, on peut citer comme opérations de prétraitement utilisées: l'extraction des composantes connexes, le redressement de l'écriture, le lissage, la normalisation et la squelettisation (voir figure 11).



Effets de certaines opérations de prétraitement sur un mot arabe.

a. La binarisation

C'est le passage d'une image en couleur ou définie par plusieurs niveaux de gris en image bitonale (composée de deux valeurs 0 et 1) qui permet une classification entre le fond (image du support papier en blanc) et la forme (traits des gravures et des caractères en noir). Pour des images de niveaux de gris, on peut trouver dans [O. D. Trier & T. Taxt 1995] une liste des méthodes de binarisation, proposant des seuils adaptatifs.

b. Extraction de composantes connexes

Une composante connexe est un ensemble de points dans le plan. Elle peut correspondre à un point diacritique, un accent, au corps d'un caractère ou d'une chaîne de caractères. Une fois localisée les composantes connexes sont regroupées pour former les mots. Cette technique est utilisée pour le repérage des points diacritiques dans les images de textes arabes [N. Ben Amara 1999].

c. Redressement de l'écriture

L'un des problèmes rencontrés en OCR est l'inclinaison des lignes du texte, qui introduit des difficultés pour la segmentation. L'inclinaison peut provenir de la saisie, si le document a été placé en biais, ou être intrinsèque au texte. Il convient alors de le redresser afin de retrouver la structure de lignes horizontales d'une image texte.

d. Lissage

L'image des caractères peut être entachée de bruits dus aux artefacts de l'acquisition et à la qualité du document, conduisant soit à une absence de points ou à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales qu'on appelle opérations de bouchage et de nettoyage [Burrow 2004].

L'opération de nettoyage permet de supprimer les petites tâches et les excroissances de la forme. Pour le bouchage il s'agit d'égaliser les contours et de boucher les trous internes à la forme du caractère en lui ajoutant des points noirs.

Plusieurs autres techniques similaires sont utilisées dont la méthode statistique, une méthode basée sur la morphologie mathématique [N. Ben Amara 1999].

e. Normalisation de la taille : La taille d'un caractère peut varier d'une écriture à l'autre, ce qui peut causer une instabilité des paramètres. Une technique naturelle de prétraitement consiste à ramener les caractères à la même taille.

f. Amincissement: Le but de l'amincissement d'un caractère est de simplifier l'image du caractère en une image plus facile à traiter en la réduisant par exemple à une dimension, c'est-à-dire que l'épaisseur du caractère est réduite à un pixel.

3.3 Extraction des caractéristiques (Calcul des représentations)

Il s'agit de la phase finale de la préparation des données. Elle fournit un certain nombre de caractéristiques ou paramètres (attributs) en utilisant des algorithmes de sélection et/ou d'extraction d'attributs. Généralement un grand nombre de paramètres est calculé à partir des données recueillies par les capteurs d'un système. C'est pourquoi, il est nécessaire d'extraire et/ou de sélectionner les paramètres les plus informatifs (Benzécri, 1973), c'est-à-dire ceux qui permettront de discriminer au maximum les classes. Les attributs étant limités en nombre, l'espace des paramètres ainsi obtenu est de dimension P très petite par rapport à R .

Les méthodes de sélection des paramètres choisissent le sous-ensemble de paramètres les plus informatifs. Les méthodes d'extraction des paramètres créent un sous-ensemble de nouveaux paramètres par combinaison des paramètres existants.

Pour réaliser ces phases de sélection et/ou d'extraction, on utilise des techniques d'analyse des données (Benzécri, 1973) telles que l'Analyse en Composantes Principales (ACP) (Jolliffe, 1988), la mesure discriminante de Fisher (Fisher, 1936), etc. Plus les paramètres du système permettent de bien discriminer les formes, plus les résultats de classification sont bons. L'ensemble des paramètres trouvés par ces méthodes représente les attributs qui permettent de caractériser chaque forme. Lorsque les données issues de l'observation du fonctionnement d'un système sont représentées par des paramètres statistiques, elles sont transformées en formes, c'est-à-dire en points, dans l'espace de représentation. Les groupes de formes similaires sont appelés classes. Si les paramètres sont bien déterminés, les classes sont bien discriminées et elles sont situées dans différentes régions de l'espace de représentation. Chaque classe est associée à un mode de fonctionnement (normal ou défaillant). Ces formes, avec leurs assignements à une classe, constituent l'ensemble d'apprentissage. Elles sont représentées par d caractéristiques, ou attributs, ce qui permet de les voir en tant que vecteurs de d dimensions, c'est à dire des points, dans l'espace de représentation. Par exemple dans une image couleur, chaque image contient n pixels et chaque pixel contient d descripteurs RGB. On représente le $i^{\text{ème}}$ pixel ($i = 1, \dots, n$) par un vecteur x_i de dimension d tel que $x_i = (x_{i1}, \dots, x_{ij}, \dots, x_{id})^T \in \{0, 1, \dots, 255\}^d$ avec d ici égale à 3.

3.4 Apprentissage

L'apprentissage ou l'entraînement, est une partie importante du système de reconnaissance. Le classificateur étant généralement une fonction paramétrique, l'apprentissage va permettre d'optimiser les paramètres du classificateur pour le problème à résoudre, en utilisant des données d'entraînement. Lorsque les données d'entraînement sont préalablement classées, l'apprentissage est dit supervisé, sinon il est non supervisé [Theodoridis et Koutroumbas 2003]. Une fois l'espace de représentation établi, il est possible de travailler avec une méthode de RdF. En fonction des informations a priori disponibles sur le système, trois types de méthodes de RdF peuvent être utilisées : les méthodes supervisées, les méthodes non supervisées et les méthodes semi-supervisées.

3.4.1 Apprentissage supervisé

Lorsque les classes des formes sont connues initialement, la RdF est supervisée (Therrien, 1989). Les méthodes supervisées utilisent les formes connues, c'est-à-dire l'ensemble d'apprentissage, pour construire un classifieur qui sépare au mieux les différentes classes connues, dans le but de minimiser l'erreur de classification. Le modèle de chaque classe est alors représenté par une fonction d'appartenance qui détermine la valeur d'appartenance d'une forme à une classe.

3.4.2 Apprentissage non supervisée

Lorsqu'aucune information n'est disponible sur les classes d'un système, la RdF est non supervisée (Bezdek, 1981) (Frigui, 1996) (Frigui, 1997) (Vachkov, 2009). Les méthodes non supervisées sont basées sur des fonctions de similarité. Quand des formes aux caractéristiques similaires apparaissent, elles sont classifiées dans la même classe et à l'inverse quand leurs caractéristiques sont différentes une nouvelle classe est créée par le classifieur. Une fois que le classifieur a estimé les fonctions d'appartenance des classes, les nouvelles formes peuvent être assignées à la classe pour laquelle elles ont la valeur d'appartenance maximale.

3.4.3 Apprentissage semi-supervisé

Les méthodes semi-supervisées (Cozman, 2003) (Gabrys, 2000) (Liu, 00) utilisent les informations connues, c'est-à-dire les formes et classes connues, pour estimer les caractéristiques des classes et leurs fonctions d'appartenances tout en utilisant également l'apprentissage non supervisé pour détecter les nouvelles classes et apprendre leurs fonctions d'appartenance.

3.5 Classification

Cette phase est le noyau de la reconnaissance des formes. En utilisant les modèles (paramètres) obtenus lors de l'apprentissage, le classificateur assigne à chaque forme inconnue sa ou ses formes les plus probables.

3.6 Post-traitement

Cette phase a pour but de corriger les résultats de la classification en utilisant des outils spécifiques au domaine d'application. Par exemple pour un système de reconnaissance de textes manuscrits, le classificateur se charge de classer chaque caractère séparément, alors que le post traitement applique un correcteur orthographique sur tout le texte pour valider et éventuellement corriger le résultat de la classification. Bien que facultative, cette phase permet d'améliorer considérablement la qualité de la reconnaissance.

4. UN EXEMPLE CONCRET : LA RECONNAISSANCE DE CARACTERES MANUSCRITS

Soit un ensemble d'images de caractères isolés ("segmentés")

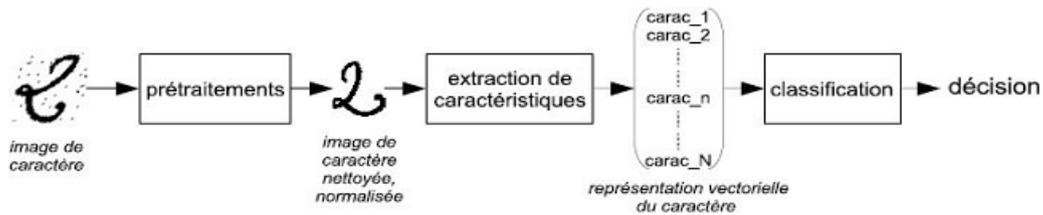
Pour les chiffres : problème à 10 classes

Pour les lettres : 26 classes, + si min/maj



A droite : une image représentant le chiffre 2. A gauche : un ensemble de chiffres manuscrits.

Le but de la reconnaissance de caractères manuscrits est qu'à partir de l'image, un tableau de pixel, on détermine la classe d'appartenance du caractère ("classification") comme l'indique le processus de traitement ci-dessous.



Chaîne de reconnaissance d'une image de caractère.

DIFFICULTES	SOLUTIONS
Présence de bruit	Prétraitements : – Binarisation : niveaux de gris/couleur => noir & blanc, réduction de la quantité d'info – Débruitage : morphologie mathématique – Normalisation => invariance vis-à-vis de la taille – Correction d'inclinaison
Mauvaise numérisation	
32*32 pixels en niveau de gris : – trop d'information – peu pertinente	Extraction de caractéristique : – Augmenter la pertinence de la représentation – réduction de la quantité d'information
La variabilité inter-scripteur : – Taille – Inclinaison – Structure – Epaisseur de trait	Apprentissage Automatique : – Apprendre la diversité et la variabilité des formes grâce en soumettant à la machine de (très) nombreux exemples "machine learning"

Caractéristiques statistiques

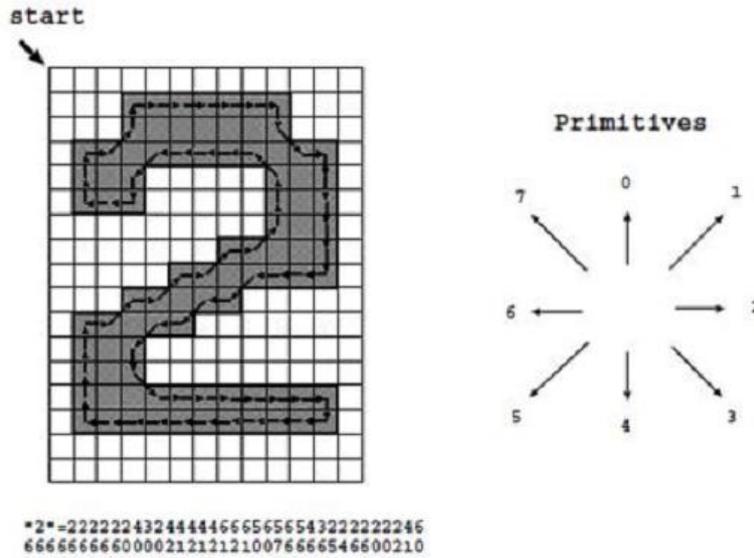
Les primitives statistiques représentent l'information distribuée sur toute la totalité de l'image. Elles sont générées généralement par l'estimation de la densité de pixels appartenant à l'image entière (caractères ou mots) ou dans certaines parties uniquement, en utilisant des mesures statistiques (entropie, moyenne, variance, etc.) associées à des caractéristiques locales qui peuvent être par exemple:

- L'histogramme directionnel qui représente le nombre de pixels sur une ligne dans une direction quelconque de l'image.
- Les moyennes locales de pixels d'une fenêtre dans l'image (principe du "zoning")
- Les projections verticales et horizontales des densités de pixels noir/blanc qui représentent le nombre de pixels sur chaque ligne ou chaque colonne de l'image.
- Les directions des contours dans une fenêtre locale.
- L'histogramme des transitions noir/blanc et vice versa qui, comme son nom l'indique, permet de retenir le nombre des transitions 0-1 et 1-0 entre pixels.

Caractéristiques structurales

Les primitives structurales représentent les propriétés topologiques et géométriques de la forme. Ces caractéristiques sont extraites à partir de la représentation de la forme par le squelette ou le contour. Voici quelques exemples :

- pavage + direction de Freeman

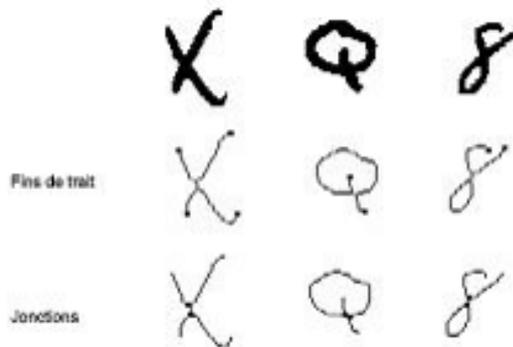


Représentation du chiffre 2 en utilisant la direction de Freeman

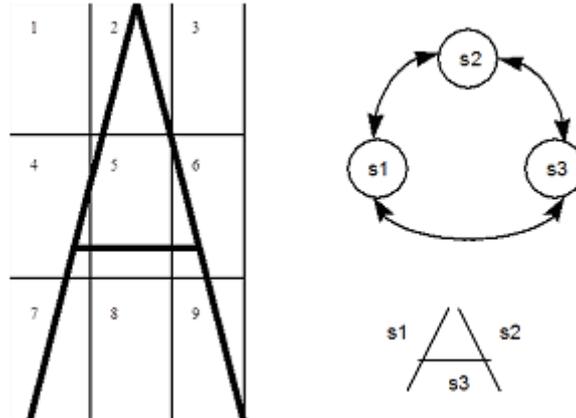
- Squelettisation



- Nombre et position des fins de traits, occlusions, jonctions, etc.



De ce fait, les caractéristiques numériques sont organisées en vecteurs de \mathbb{R}^d , mais elles peuvent être également présentées sous une représentation structurelle (graphe), ou syntaxique.



Exemple simplifié d'extraction de caractéristiques sur une image de caractère. Gauche : l'image est découpée en 9 zones. Sur chacune d'elles, on extrait le pourcentage de pixels noirs. L'ensemble des 9 valeurs forme le vecteur de caractéristiques. Droite : la forme est découpée en segments s1, s2 et s3. L'ensemble de ces segments est agencé sous forme de graphe : s1 est lié à s2, s3 coupe s1 et s2,....

5. EVALUATION DES PERFORMANCES D'UN CLASSIFIEUR

L'objectif est d'obtenir une estimation la plus fidèle possible du comportement du système dans des conditions réelles d'utilisation. Pour cela les critères classiques comme les taux de reconnaissance et taux d'erreurs sont presque systématiquement utilisés. En RdF, il y a 3 catégories de bases :

- Une base de référence ou d'apprentissage utilisée pour apprendre le classificateur.
- Une base de validation pour déterminer les paramètres du classificateur.
- Une base de test constituée d'exemples jamais vus au préalable pour évaluer le classificateur.

En fonction des statistiques sur la base de test, on va pouvoir définir:

Le pourcentage (%) des formes bien classées -> facteur de qualité

Le pourcentage (%) des formes mal classées -> étude des confusions

Le pourcentage (%) des formes non classées -> critères de rejet

Matrice de confusion

Une manière plus complète et très utilisée pour représenter les performances d'un système de classification est de construire une "matrice de confusion". Les intitulés de colonne de cette matrice correspondent aux classes d'appartenance déterminées par le système, tandis que ceux des lignes représentent la vraie classe d'appartenance. Ainsi, la valeur N_{ij} trouvée sur la ligne i et colonne j est le nombre de points appartenant à la classe i qui ont été affectés à la classe j , comme illustré dans le tableau 1.

		Classe prédite			
		C_0	C_1	\dots	$C_{ C }$
Classe réelle	C_0	N_{00}	N_{01}	\dots	$N_{0 C }$
	C_1	N_{10}	N_{11}	\dots	$N_{1 C }$
	\vdots	\ddots	\ddots	\ddots	\ddots
	$C_{ C }$	$N_{ C 0}$	$N_{ C 1}$	\dots	$N_{ C C }$

Tableau 1. Matrice de confusion

Usuellement la matrice de confusion est présentée sous une forme normalisée. Les valeurs de la matrice sont obtenues par la division des valeurs N_{ij} de la matrice du tableau 1.1 par le nombre de points de la ligne i dans l'ensemble de test. Evidemment, pour ce type de matrice de confusion évaluation d'un classifieur parfait sera une matrice de confusion identité, comme représenté dans le tableau 2.

		Classe prédite			
		C_0	C_1	\dots	$C_{ C }$
Classe réelle	C_0	1	0	\dots	0
	C_1	0	1	\dots	0
	\vdots	\ddots	\ddots	\ddots	\ddots
	$C_{ C }$	0	0	\dots	1

Tableau 2. Matrice de confusion normalisée pour un classifieur parfait

Taux de bonne classification sans coûts

Sans rejet

Le taux de bonne classification simplifié (Tb_s) ou sans coût s'agit de l'indicateur le plus naturel et le plus évident permettant d'évaluer les performances d'un système de classification. Cette valeur, simple à calculer, correspond au nombre d'éléments correctement identifiés par le système. La définition du taux de bonne classification sans la prise en compte du rejet est :

$$Tb_s = \frac{Nb \text{ d'exemples bien classés}}{Nb \text{ d'exemples}}$$

On obtient le taux d'erreur par :

$$Te_s = 1 - Tb_s$$

Avec rejet

Lorsque le rejet d'une forme est possible, un troisième taux, dit "taux de rejet" (Tr) est intégré. Il mesure le nombre d'éléments sur lesquels le système n'a pas pris de décision. Nous obtenons ainsi :

$$Tr = \frac{Nb \text{ d'exemples non classés}}{Nb \text{ d'exemples}}$$

On obtient le taux d'erreur par le taux de bonne classification

$$Tb_a = \frac{Nb \text{ d'exemples bien classés}}{Nb \text{ d'exemples}}$$

On obtient le taux d'erreur par :

$$Te_a = 1 - Tb_a - Tr$$

6. LES METHODES DE LA RDF

Les méthodes de RdF peuvent généralement être divisées en deux catégories : les méthodes paramétriques et les non paramétriques.

Les méthodes paramétriques considèrent l'ensemble d'apprentissage comme des données indépendantes, toutes distribuées selon la même loi de probabilité, comme le fait le classifieur Bayésien (Dubuisson, 1990) (Yager, 2006). Cette hypothèse n'est bien souvent pas valide dans le cas d'applications réelles.

Les méthodes non paramétriques génèrent les fonctions d'appartenance des classifieurs (Medasani, 1998) soit :

- En estimant la fonction de densité de probabilité conditionnelle pour chaque classe, comme le fait la méthode des fenêtres de Parzen (Parzen, 1962) (Dubuisson, 2001) et la méthode des K-Plus Proches Voisins (Cover, 1967) (Denoeux, 2001)
- En construisant par apprentissage les régions de décision comme le font les réseaux de neurones (Ripley, 1996) ou les Séparateurs à Vaste Marge (Vapnik, 1995)

Les méthodes de la RdF sont souvent regroupées en grandes classes identifiées par : statistique, syntaxique, structurelle, hybride (une combinaison des autres).

6.1 Méthodes statistiques

Les méthodes statistiques sont fondées sur l'étude statistique des mesures que l'on a effectuées sur les objets à reconnaître. L'étude de leur répartition dans un espace métrique et la caractérisation statistique des classes permet de prendre une décision de reconnaissance du type « plus forte probabilité d'appartenance à une classe ». Ces méthodes s'appuient en général sur des hypothèses concernant la description statistique des familles d'objets analogues dans l'espace de représentation.

6.2 Méthodes syntaxiques

Si l'approche statistique permet de se placer dans un cadre mathématique solide et général, elle présente néanmoins le défaut d'oublier la nature des mesures qui sont faites sur les formes et de les traiter de façon abstraite.

Dans le cas d'une description structurelle, il faudra utiliser des formalismes plus complexes. Les formes sont décomposées en primitives simples (c'est-à-dire en composantes élémentaires –non décomposables- d'une forme. Elles sont mises en évidence à la phase d'extraction. A titre d'exemple : un segment de droite, une boucle... dans une figure géométrique) qui peuvent être des graphèmes ou même les pixels de l'image. Elles sont ensuite représentées par un objet complexe, composé de primitives, comme une chaîne ou un graphe. Le processus de reconnaissance grammatical, stochastique ou graphique est propre à la représentation utilisée. La représentation structurée peut sembler plus naturelle en RdF. Le traitement structuré des formes peut se répartir en deux grandes familles :

- **Syntaxique** : car elle tente de s'appuyer sur les travaux de la linguistique formelle. Les formes sont alors codées par des « mots » utilisant un alphabet dont chaque terme représente un élément de la forme à décrire. Définir une classe de formes revient à établir des règles syntaxiques caractérisant les mots acceptables. Le choix de l'alphabet et des règles syntaxiques est souvent fort délicat, mais des méthodes pour déterminer ces règles ont été proposées.

- **Structurelle** : tout en organisant la représentation des formes à l'aide de primitives de description et de relations entre ces primitives, elle ne s'appuie pas sur les grammaires formelles mais sur des techniques combinatoires (isomorphismes de graphes, arbres d'interprétation, ...) ou numériques (métrique sur les mots, relaxation,...).

6.3 Combinaison des classificateurs

Pour résoudre les problèmes de reconnaissance, on doit parfois combiner plusieurs classificateurs. On le fait pour plusieurs raisons [Jain et al, 00] :

- Le concepteur du système peut avoir accès à plusieurs classificateurs, développés dans des contextes différents pour des représentations/descriptions totalement différentes d'un même problème.
- On peut avoir plusieurs ensembles d'entraînement. Ces ensembles ont été collectés/extraits à différents moments dans différents environnements, et peuvent être représentés par des paramètres différents.
- Des classificateurs différents, entraînés sur les mêmes données, peuvent présenter de grandes différences dans la qualité de la classification. Chaque classificateur peut avoir son propre sous espace de paramètres où il est le plus efficace.
- Certains classificateurs comme les réseaux de neurones montrent des comportements différents selon la phase d'initialisation, ceci est dû à la part aléatoire du processus d'initialisation. Au lieu de ne garder qu'un seul classificateur, on peut combiner les réseaux de neurones obtenus pour bénéficier des résultats de tous les entraînements.