

Reconnaissance des formes

Éléments de base

Houcine Senoussi

29 janvier 2025

- 1 Introduction
- 2 Définitions
- 3 Approches de la RF
- 4 Structure d'un système de RF
- 5 Conclusion
- 6 Références

- Dans ce chapitre, nous introduisons la reconnaissance des formes à travers ses **définitions** les plus courantes, ses **approches** principales et ses '**building blocks**' et ses **concepts** les plus importants.
- Pour ce faire nous allons commencer par discuter la manière dont elle est définie par la plupart des auteurs pour aboutir à une caractérisation qui prend en plus en compte les principaux types de données auxquels elle s'applique, puis nous présenterons et discuterons la structure d'un système de RF que nous illustrerons par un exemple.

De quoi s'agit-il ?

- Il existe dans la littérature plusieurs définitions de la **reconnaissance des formes (pattern recognition)**.
- Pour certains auteurs, l'objet de la reconnaissance des formes est la découverte des **régularités** dans les données, puis l'utilisation de ces régularités pour des tâches telles que la **classification** des données.
- Dans cette définition **générale**, la reconnaissance des formes s'applique à tout type de données (texte, ensemble de points dans un espace multidimensionnel, séries de données, image, signal audio, ...).
- Ici, l'accent est mis sur l'aspect **classification/catégorisation**, souvent supervisée, mais aussi non supervisée (des motifs). Cela a fait dire à un auteur que 'The construction of a classification procedure from a set of data for which the true classes are known has also been variously termed **pattern recognition**, **discrimination**, or **supervised learning**' (et, devrait-on ajouter, **data mining**).

De quoi s'agit-il ?

- Cette définition générale suppose, ou s'appuie sur, une définition aussi générale de ce qu'est un motif (pattern). En effet, on trouve souvent citée cette définition de S. Watanabee : 'un motif est une entité, vaguement définie, à laquelle on peut donner un nom !'.
- Il est parfois précisé, que le pattern peut être observé physiquement (dans une image par exemple) ou défini mathématiquement comme une corrélation entre différentes données (là on retombe sur la proximité de cette définition générale avec celle du data mining).
- Cette définition d'un pattern n'étant manifestement pas suffisante, elle est souvent complétée par des exemples : un pattern peut être une empreinte digitale, un visage humain, (un segment dans) un signal vocal, un mot, une lettre, ou un chiffre manuscrit (dans une image).

De quoi s'agit-il ?

- Les exemples donnés ci-dessus peuvent induire une définition moins générale, plus restrictive de la RF.
- On peut, par exemple, citer la définition selon laquelle l'objet de la RF est principalement la description et la classification d'un **ensemble de mesures** effectuées à partir d'un **processus physique**. Il est précisé à la suite de cette définition qu'il convient de distinguer deux cas : celui où une image est **naturellement présente** (e.g. reconnaître des lettres de l'alphabet inscrites sur un papier) et celui où les données brutes se présentent sous la forme d'un **simple ensemble de mesures** pour chaque object à reconnaître (e.g. concentration de divers produits chimiques dans un échantillon).

■ On peut enfin citer la définition/caractérisation suivante :

- Point de départ : **Reconnaître des objets** est l'un des aspects les plus importants de la **communication** avec le **monde extérieur**, ou encore avec l'**environnement**. Cela vaut pour les humains mais aussi pour les machines.
- Exemples importants : dans une **image**, un ensemble d'images, ou une séquence d'images, nous avons besoin de reconnaître les **directions** dans lesquelles les objets sont orientés, où ils **se situent**, comment ils sont **disposés**, quelle est la **taille** et la **forme** de chacun d'entre eux, et quelles **sortes** d'objets ils sont.
- Ces opérations ne s'appliquent pas seulement aux **images** (objets visuels), mais à toute sorte d'**entités** du monde réel.

Pourquoi la RF est-elle importante ?

- Une citation de Herbert Simon, qui est l'un des pionniers de l'intelligence artificielle explique pourquoi la reconnaissance est essentielle dans la conception de programmes intelligents :
 - 'The more **relevant patterns** at your disposal, the better your **decisions** will be. This is hopeful news to proponents of **artificial intelligence**, since computers can surely be taught to **recognize** patterns. Indeed, successful computer programs that help banks score credit applicants, help doctors diagnose disease and help pilots land airplanes depend in some way on pattern recognition... We need to pay much more explicit attention to teaching pattern recognition.'

De quoi s'agit-il ?

- Les définitions et remarques précédentes peuvent être récapitulées, puis légèrement développées, comme suit :
 - À l'origine de la reconnaissance des formes, il y a le besoin qu'ont les machines 'intelligentes' de connaître le monde extérieur, ou son état à un moment donné. La reconnaissance des formes est donc processus qui commence par l'acquisition des données. Parmi ces données, une classe particulière revêt une importance particulière : ce sont les **données visuelles**, c'est-à-dire les **images**. Au sein de ces données, certains **éléments** (ou objets, ou entités, ou segments) doivent être **trouvés** (on dira aussi détectés, localisés) et **reconnus** (classés, catégorisés). Ces éléments sont les **patterns** (motifs).
- Dans la suite de ce cours, nous considérons quasi-exclusivement les images comme type de données.

- Lorsque l'on considère des images comme type de données, la RF utilise le traitement des images comme **pré-requis**, ou si l'on veut être plus précis, les **premières étapes** d'un système de RF (voir plus bas) utilise le traitement des images.
- Un système de RF met en oeuvre un algorithme de **classification**. De nos jours, la classification se fait par des algorithmes de **machines learning**. Comme le montre Yan Lecun et ses co-auteurs dans un article paru à la fin des années 90s, en s'appuyant sur l'exemple de la reconnaissance des caractères, de meilleurs systèmes de RF sont obtenus en remplaçant les heuristiques 'conçues à la main' (hand-designed) et l'extraction manuelle de variables (handcrafted) peuvent (ont pu) être avantageusement remplacés par du machine learning.

- La reconnaissance des formes est très utilisée dans le domaine médical. Elle permet par exemple de distinguer les cellules saines des cellules malades (première phase d'un diagnostic médical). Elle permet aussi l'analyse de l'électrocardiogramme (ECG).
- Dans le domaine de la sécurité, citons l'authentification à l'aide de la reconnaissance des empreintes digitales, la reconnaissance faciale, ou la reconnaissance vocale.
- Une autre application de la reconnaissance de la parole est ce qu'on appelle le *voice typing* qui consiste à convertir la parole en texte écrit (utilisé notamment par les médecins).
- Dans le domaine de la vie courante, citons la reconnaissance des chiffres et des lettres manuscrites (utilisée pour la lecture des chèques et des lettres), ainsi que celle des plaques d'immatriculation.

- La plupart des auteurs distinguent principalement 3 approches de la reconnaissance des formes : l'approche **syntaxique**, l'approche **statistique**, et l'approche **neuronale**.
- L'approche **syntaxique** est adaptée aux problèmes de RF dans lesquels les patterns sont **complexes** et peuvent être organisés de manière **hiérarchique** : chaque pattern est composé de **sous-patterns**. Les patterns les plus simples sont connus sous le nom de **primitives**. Pour décrire l'ensemble des patterns, une **grammaire** est apprise qui utilise les primitives comme **symboles terminaux** et les autres patterns comme des phrases.
- Pour certains auteurs, l'approche syntaxique est un cas particulier d'une approche plus générale qui est l'approche **structurelle**. Cette dernière, comme son nom l'indique, est adaptée lorsque les différentes classes se distinguent par des propriétés structurelles. Un exemple cité est la reconnaissance des lettres de l'alphabet.

- Dans l'approche **statistique**, chaque pattern est défini par un nombre fixe d de variables numériques, et peut donc être considéré comme un point dans un espace de dimension d . La pertinence de ces d variables peut être évaluée par la facilité avec laquelle les classes de patterns peuvent être distinguées dans cet espace. Le nom *statistique* vient du fait que, historiquement, les frontières des zones de l'espace correspondant à chaque classe de patterns est déterminée par inférence statistique.
- L'approche la plus populaire aujourd'hui est l'approche **neuronale**. Nous verrons dans les chapitres suivants qu'elle change en profondeur la façon dont les différentes composantes d'un système de RF sont agencées.

- La caractérisation de la RF que nous avons donnée plus haut, conduit naturellement à la considérer comme un RF qui commence par l'**acquisition** des données et qui aboutit à la classification du/des patterns concernés.
- Entre l'acquisition et la classification, plusieurs étapes sont indispensables. D'abord, il faut 'préparer' les données (e.g. traiter le bruit), cette phase est appelé le pré-traitement. Ensuite il faut **localiser** le/les patterns. Ensuite il faut le/les **représenter** à l'aide des **variables** adéquates.
- La séquence de ces étapes est récapitulé par le schéma de la page suivante.

Structure d'un système de RF-2

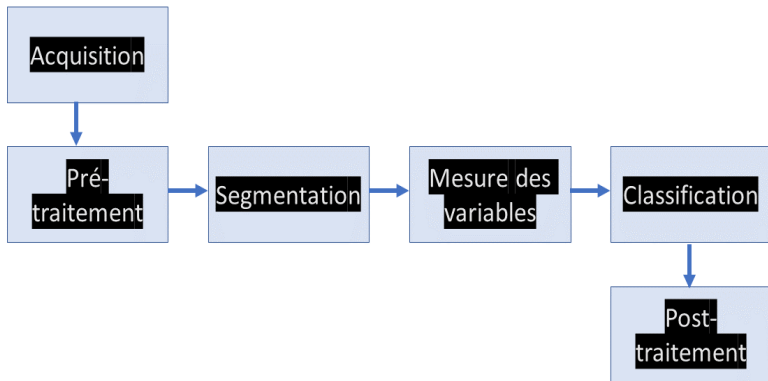


Figure – Etapes d'un système de RF

Structure d'un système de RF-Prétraitement

- Le pré-traitement inclut un ensemble d'opérations, d'ordre divers, dont le but commun est l'**amélioration** des images. Cette amélioration vise essentiellement à corriger les défauts des capteurs (phase acquisition).
- Cette étape inclut par exemple :
 - La suppression du bruit. Plusieurs filtres peuvent être utilisés pour ce faire (e.g. filtre de moyenne, filtre médian).
 - La transformation lorsqu'il y a besoin d'une image en couleur en une image à niveau de gris.
 - Des manipulations de l'histogramme de l'image dont l'objectif est notamment d'améliorer le contraste de l'image (via l'étalement de l'histogramme) et de corriger la répartition des niveaux de gris (via égalisation d'histogramme).

(Voir le cours de traitement d'images pour les détails et plus de précisions)

Structure d'un système de RF-Segmentation

- La segmentation d'une image consiste à la diviser en régions connexes, chacune possédant certaines propriétés qui la différencient de ses voisines. Son objectif est de faire en sorte que les objets se 'détachent' du fond de l'image et des autres objets.
- Il existe plusieurs méthodes de segmentation, qu'on regroupe en méthodes de **segmentation par région** et méthodes de **segmentation par contour**.
- Parmi les méthodes de segmentation par région, citons par exemple :
 - La segmentation par **seuillage**. C'est la méthode la plus simple. Elle consiste à répartir l'image en deux régions selon que la valeur des pixels est plus grande ou plus petite qu'un certain seuil T . Ces deux régions sont supposées correspondre au fond et aux objets. T est en général estimé à partir de l'histogramme de l'image. Cette méthode est efficace lorsque le fond et les objets sont homogènes. Elle peut être étendue à un seuillage multi-niveaux.

■ Segmentation par région (suite) :

- La segmentation par **fusion** (resp. **décomposition**) des régions (merge/split). Cette méthode commence par des petites régions puis les régions voisines sont fusionnées si elles ont les mêmes propriétés. La démarche est inversée dans le cas de la segmentation par décomposition. Il existe aussi une méthode qui combine les deux opérations (split-and-merge).

■ La segmentation par contour construit d'abord les contours (bords des objets) puis en déduit les régions. Il existe plusieurs opérateurs pour détecter les contours (e.g. le **gradient**, qu'on peut calculer par différents filtres, idem pour le **laplacien**).

■ **Remarque** : Il existe aussi des méthodes de segmentation qui utilisent les opérateurs de **morphologie mathématique**.

Structure d'un système de RF-Valeur des variables

- Les variables décrivant les patterns doivent être définies de manière à avoir des valeurs proches pour des objets appartenant à la même catégorie, et des valeurs éloignées pour des objets appartenant à des catégories différentes. Autrement dit, leur principale propriété est d'être **discriminantes**.
- Idéalement, ces variables doivent être invariantes par translation, par rotation, et par d'autres transformations comme le changement d'échelle.
- **Remarque 1** : Dans la littérature en langue anglaise, cette phase est souvent désignée sous le nom de 'feature extraction'. Le terme feature peut être traduit par *caractéristiques* ou *attributs*.
- **Remarque 2** : Dans cette description des variables, et plus généralement dans ce chapitre, nous avons adopté le point de vue (historiquement justifié) où les variables sont définies par 'manuellement' par le concepteur du système. Nous verrons qu'avec le Deep learning, ce point de vue est pour le moins à nuancer.

Structure d'un système de RF-Classification

- Aux débuts de la RF, les classifieurs étaient définis **manuellement** (e.g. sous la forme d'un ensemble de **règles**) par le concepteur du système. Aujourd'hui, ils sont **entraînés** en utilisant des classes **prédéfinies** (apprentissage **supervisé**) ou en utilisant une mesure de **similarité** (apprentissage **supervisé**).
- Un classifieur peut retourner une **classe** ou des **probabilités** d'appartenir à chaque classe.
- Parmi les types de classifieurs que l'on rencontre souvent dans les travaux sur la RF, citons notamment **SVM** (voir chapitre consacré à cet algorithme), le classifieur **bayésien** (pour l'approche probabiliste) et, bien entendu, les **réseaux de neurones** (et plus précisément les réseaux de neurones **convolutifs**, voir suite du cours).

- Plusieurs classifieurs peuvent être **combinés**. Ces classifieurs peuvent utiliser des **descriptions** différentes des données (i.e. des variables différentes), peuvent être **entraînés** sur des ensembles de données différents. Il existe plusieurs schémas selon lesquels les classifieurs sont combinés. Le plus simple est la combinaison **parallèle** (les classifieurs sont invoqués indépendamment, puis leurs décisions sont combinés par un vote **majoritaire** ou en affectant des **poids** différents aux différentes décisions). D'autres schémas invoquent les classifieurs en **cascade** ou les organisent sous la forme d'un **arbre**.
- Parmi les exemples de combinaison des classifieurs que l'on rencontre dans la littérature, certains illustrent la **variété** des types de données utilisés dans la reconnaissance des formes, par exemple celui dans lequel, pour identifier une personne, on combine des classifieurs qui utilisent respectivement leur **voix**, leur **visage** et leur **écriture**.

- Les auteurs qui terminent les étapes d'un système de RF par le postprocessing y incluent le calcul de l'**erreur** du classifieur, l'analyse de sa décision, et éventuellement sa combinaison avec d'autres classifieurs.

Structure d'un système de RF-Un exemple

- Dans cet exemple, nous présentons un résumé d'un système de RF, réalisé par un groupe de chercheurs polonais, et dont les résultats ont été publiés en 2004 (voir référence ci-dessous).
- Dans ce travail, il s'agit de **détecter** les **cellules** dans l'image (voir page suivante) et de les **classer**. Les cellules sont réparties sur 12 classes. L'expérience a porté sur environ 700 cellules, qui ont été réparties entre l'ensemble d'apprentissage et l'ensemble de test.
- Le **prétraitement** a consisté à transformer l'image en couleurs en image à niveau de gris puis en image binaire.
- La **segmentation** a été réalisé à l'aide d'opérations issues de la **morphologie mathématique** et de la transformation dite de **Watershed**. Ces opérations (**dilatation** et **érosion**) supposent la connaissance a priori de la **forme** des patterns (les cellules) à **localiser**. Elles combinent la segmentation par région et celle par contour.

- 3 familles de variables ont été utilisées pour décrire les patterns. Les premières décrivaient la **texture**, les suivants sont des variables **géométriques** (e.g. le rayon, le périmètre, ...), et les derniers sont des variables **statistiques** (e.g. moyenne et variance de l'intensité de chacune des couleurs primaires).
- Le classifieur utilisé est *SVM*.
- Le taux d'erreurs dans la classification des cellules a été assez faible (autour de 3.5%).

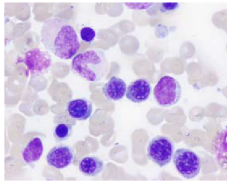


Figure – Cellules à reconnaître

- En prévision de la première étude de cas, consacrée à la reconnaissance des empreintes digitales chargez le dataset SOCOFING à partir de Kaggle.
- Écrire un code permettant de sélectionner une empreinte digitale au hasard, construire son histogramme, puis appliquer l'étalement et l'égalisation de l'histogramme.

- Dans ce chapitre, nous avons introduit les éléments de base de la RF.
- En plus de son côté volontairement **pédagogique** (donc simplificateur), cette présentation a un aspect 'historique' : nous n'avons pas voulu présenter uniquement la RF telle qu'elle se pratique aujourd'hui, avec la domination écrasante du deep learning, mais aussi telle qu'elle se pratiquait avant. C'est précisément cela qui nous a permis de présenter ses principaux concepts.
- Dans les chapitres suivants, nous serons amenés à nuancer certains aspects présentés dans ce chapitre.

■ Livres

- D. Luo, 'Pattern Recognition and Image Processing', 1998.
- R. O. Duda et al., 'Pattern Classification', 2001.

■ Articles

- J.S. Bridle, 'PATTERN RECOGNITION TECHNIQUES FOR SPEECH RECOGNITION', 1980.
- S. Osowski et al., 'FEATURE GENERATION FOR THE CELL IMAGE RECOGNITION OF MYELOGENOUS LEUKEMIA', 2004.
- A. K. Jain, 'Statistical Pattern Recognition : A Review', 2000.

■ Sites internet

- [kaggle.com](https://www.kaggle.com). Plusieurs compétitions.