

## 1. Introduction

La reconnaissance biométrique est nouvelle tâche dans le domaine sécurité et qui est devenu une solution alternative se basée sur les traits physiques ou morphologique d'un humain afin d'accéder system protégé ayant une technologie biométrie. Parmi ces techniques biométriques, la reconnaissance de l'ECG est l'une des approches les plus prometteuses en raison de sa grande fiabilité pour l'identification humain. Les premières travaux [36] montrent que le signal ECG contient des caractéristiques uniques appropriés à chaque personne.

Dans ce chapitre, nous présenterons les différentes méthodes d'identification biométrique basée sur l'ECG.

## 2. Historique

Dans notre société sait que les technologie biométries, le visage, iris, l'empreinte digital soient le plus utilise dans nos vies quotidiennes. L'utilisation de l'ECG en biométrie est relativement nouvelle. En fait, il existe plusieurs méthodes biométriques basées sur l'ECG proposées précédemment. En 2001, Biel et al ont tenté de développer un system biométrique basé sur les caractéristiques temporelles d'un signal ECG [36]. Dans cet optique, Wang et al. ont proposé une nouvelle méthode d'identification biométrique basé sur les variation morphologiques d'un signal ECG codés en des composants fréquentielle. Au cours ces derniers années plusieurs approches ont étéproposé avec succès qui exploitent les information temporelle, spatial et fréquentielle d'un signal ECG. Dans plusieurs travaux, les approches neuronales ont été aussi bien exploité dans cette application. Dans ce travail ont proposé un system hybride qui exploite l'information temporelle d'un signal ECG et l'efficacité de la discrimination d'une approche neuronal laquelle va identifier une identité humaine selon la variation de son signal ECG.

## 3. La biométrie par l'ECG :

Dans la section suivante, nous donnerons un rappel anatomique du système cardiovasculaire.

### 3.1 Le système cardiovasculaire

Le cœur est une pompe à quatre chambres du système circulatoire. La fonction principale depompage est assurée par les ventricules ; les oreillettes stockent le sang avant que les ventricules pompent. Le cœur est composé de différents types de tissus (tissu nodal SA et AV, tissu auriculaire, Purkinje, et tissu ventriculaire). Les cellules représentatives de chaque type

de tissu différent anatomiquement d'une manière considérable. Elles sont électriquement excitables et chaque type génère son propre potentiel d'action ; dans ce cas, le cœur est assimilé à un générateur électrique. Ces potentiels sont mesurés à la surface de la peau et constituent l'électrocardiogramme ou ECG. [17]

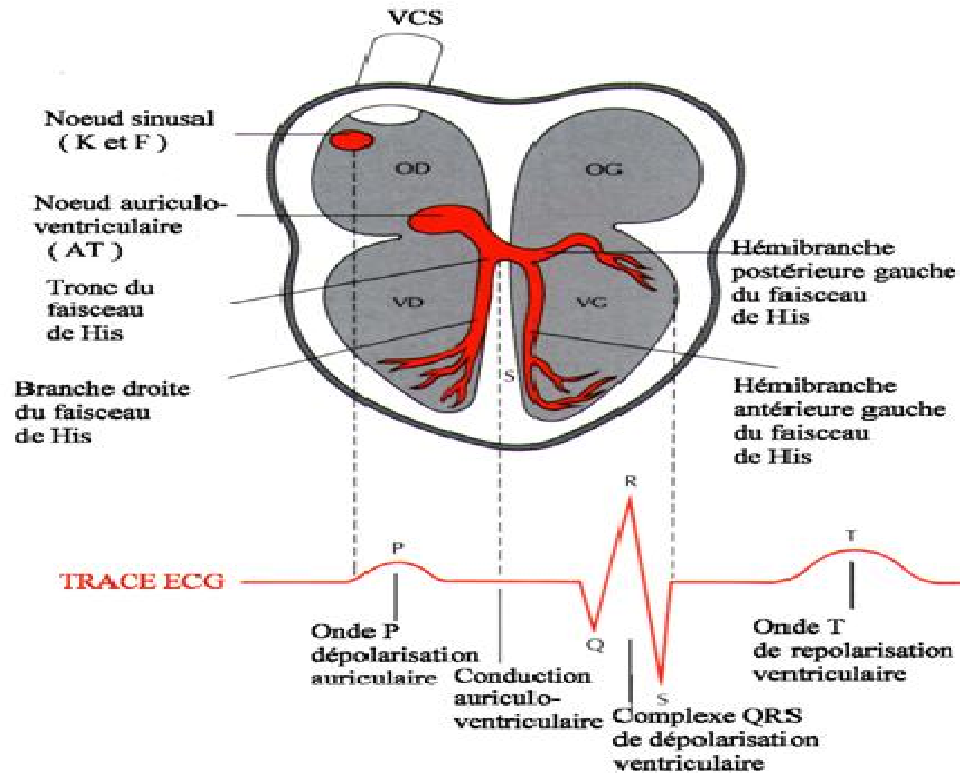


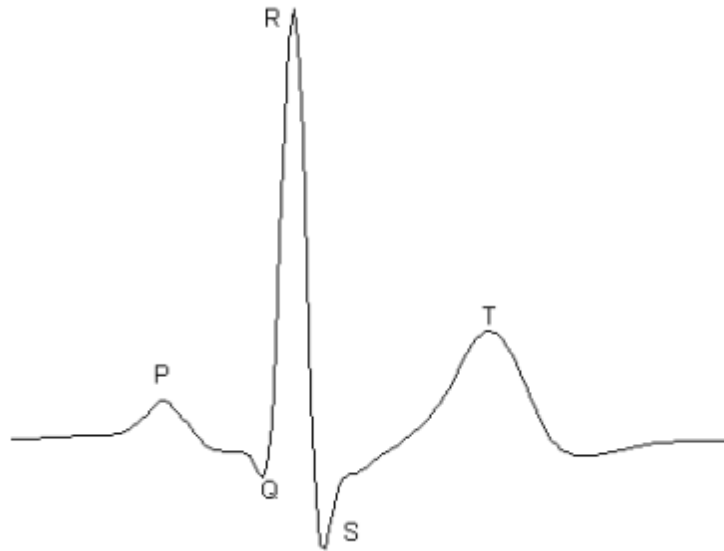
Figure 2.1 : Propagation de l'influx à l'intérieur du cœur.

### 3.2 Analyse de l'ECG

L'ECG est l'activité électrique du cœur souvent enregistrée au niveau de la poitrine. Pendant son activité, le myocarde - le muscle cardiaque - se comporte comme une série de dipôles électriques connectés d'une manière unique appelée «syncytium» fonctionne. L'activité électrique du cœur est communément décrite en utilisant un vecteur électromagnétique individuel variable dans le temps, dont les projections peuvent être enregistrées sur la surface du corps. [19]

Un ECG est considéré comme normal quand les différents phénomènes électriques répondent à des critères de forme et de temps, et quand il existe un rythme sinusal régulier de fréquence normale. La figure 2.2 présente un tracé ECG avec les ondes P et T et le complexe QRS. Ces différentes vagues formant l'ECG représentent les séquences de dépolarisation et de repolarisation des oreillettes et des ventricules (figure 2.1). Le complexe QRS est la forme

d'onde la plus saisissante de l'ECG puisqu'il reflète l'activité électrique du cœur pendant la contraction ventriculaire. En raison de sa forme caractéristique, il sert de base à la détermination automatisée de la fréquence cardiaque. La détection du QRS fournit les principes fondamentaux pour presque tous les algorithmes automatisés d'analyse de l'ECG. [17]



**Figure 2.2:** Battement d'un ECG.

Le signal ECG se compose d'ondes, d'intervalles, de segments et d'un complexe appelé QRS, comme indiqué dans la figure 2.2 :

- Onde: Une déflexion positive ou négative par rapport à la ligne de base indique un phénomène électrique précis. L'onde d'un signal ECG comprend l'onde P, l'onde Q, l'onde R, l'onde S, les ondes T et U.
- Intervalle : temps entre deux événements ECG spécifiques. Les intervalles qui sont couramment mesuré sur un signal ECG comprennent l'intervalle PR (intervalle PQ), l'intervalle QT et RR.
- Segment : La longueur entre deux points spécifiques sur le signal ECG qui sont supposés être à l'amplitude de base (pas négatif ou positif). Le segment ST a segmenté le TP segment.
- Complexe : La combinaison de plusieurs vagues regroupées, le seul complexe principal sur l'ECG est le complexe QRS.
- Point : Il n'y a qu'un seul point sur l'ECG connu sous le point J qui est la jonction entre la fin du QRS et le début du segment ST.[18]

✓ **L'onde P**

L'onde P correspond à la dépolarisation auriculaire et à la diffusion du signal électrique à travers les oreillettes droite et gauche, cette onde s'achève au niveau où le tracé rejoint la ligne de base. sa durée est de 0.08 à 0.1 secondes, son amplitude inférieure à 2 mm, elle est monophasique, arrondie, et toujours positive.

✓ **Le complexe QRS**

Le complexe QRS est formé de 3 ondes Q, R et S et correspond à la diffusion de l'influx électrique à travers les ventricules (dépolarisation ventriculaire). Il représente la dépolarisation ventriculaire et dure normalement entre 0.06 et 0.1 seconde.

Cette durée courte indique que la dépolarisation ventriculaire est rapide. L'onde Q est la première déflexion négative du complexe QRS et l'onde R est la première déflexion positive. Une déflexion négative suit l'onde R, elle est appelée onde S. Ce complexe est d'une grande amplitude (signal de quelques millivolts).

✓ **L'onde T**

L'onde T représente la repolarisation ventriculaire. Elle est plus longue en durée que la dépolarisation (la vitesse de conduction de l'onde de repolarisation est plus faible que celle de l'onde de dépolarisation). Elle survient pendant la diastole, et ne témoigne d'aucun événement mécanique. C'est un phénomène purement électrique, pendant lequel les ventricules redeviennent stimulables. [17]

#### **4. Méthodes d'extraction des caractéristiques**

Les approches de reconnaissance basées sur l'ECG sont nombreuses et très variées. Les attributs (caractéristiques) d'un ECG sont extraits pour classer un signal ECG d'une personne 'X' par minimiser la variation intra-signal et maximiser la variation inter-signal. En général, les caractéristiques sont basées sur la morphologie du rythme cardiaque, à un moment particulier les intervalles dérivés des ondes ECG ou des caractéristiques extraites de manière spécifique. Le choix du mode d'identification est généralement déterminé par la complexité du système de reconnaissance, le besoin d'identification en temps réel et le dispositif d'enregistrement spécifique, etc.

Dans ce mémoire, l'analyse ECG est souvent effectuée sur des bases de données off-ligne afin de faciliter la tâche de la reconnaissance.

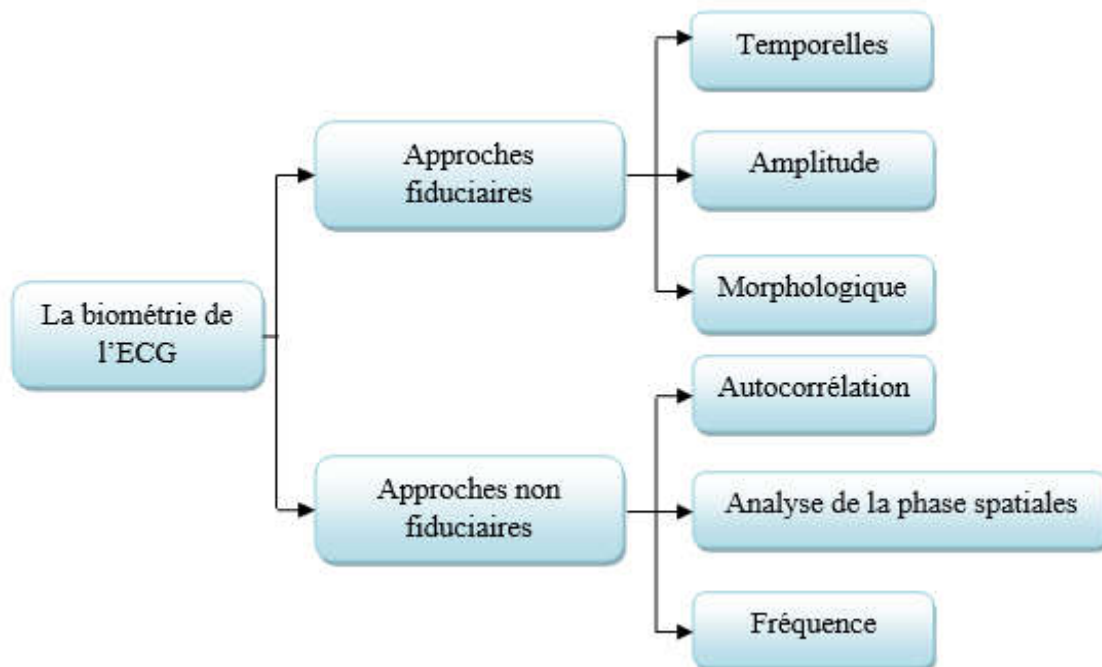
Ici, nous avons récapitulé les approches existantes qui ont été regroupées en deux grandes catégories principales: fiduciaire et non fiduciaire.

Approches fiduciaire: sont les méthodes le plus utilisées pour faire l'extraction des attributs d'un signal ECG (Les pics, les intervalles, les pentes ou d'autres mesures d'une onde), qui sont simplement basés sur l'information temporelle et morphologique lesquelles apparue au fur et à mesure d'un rythme cardiaque ECG.

Dans ce mémoire, nous suivons une approche de seuillage pour localiser les pic 'R' (voir la figure 2.4) dans chaque enveloppe complexe QRS, puis on calcul les distances mesurées entre  $R_i$  et  $R_{i+1}$  (Baptiser l'intervalle RR), qui sont localisées successivement dans un signal ECG. Une fois les distances sont calculées et enregistré dans un vecteur de taille N que sera représenté les attributs d'un signal ECG.

Approches non fiduciaire: les caractéristiques naturel d'un signal ECG ont fait un grand effet sur localisation d'une information invariante dans le temps, celle-ci peut être considéré un challenge dans le traitement des signaux non stationnaire, l'information fréquentielle a montré un atout dans plusieurs domaine, la transforme Fourier, l'ondelette, ...etc. [19] sont des outils pour visualiser cette information laquelle considère comme une moyenne pour coder un signal ECG.

Chaque catégorie peut être divisée et représenté sur la figure 2.3 :



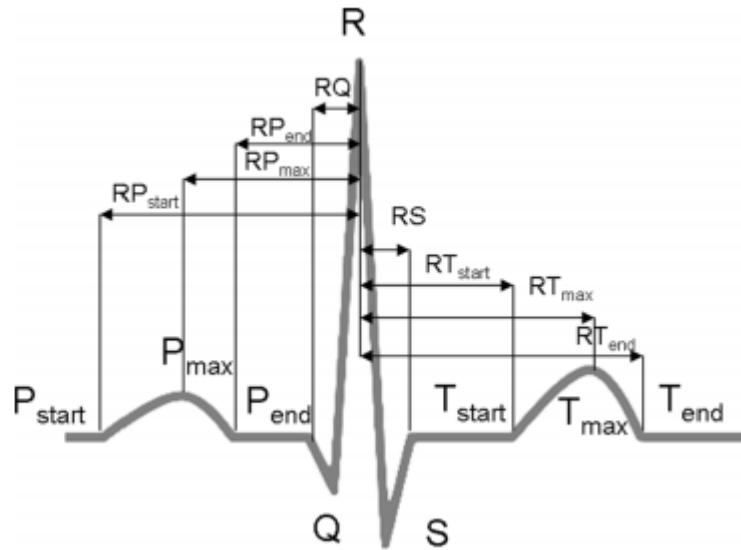
**Figure2.3 :** Taxonomie de l'analyse biométrique basée sur l'ECG.

## 4.1 Approches fiduciaires

### 4.1.1 Caractéristiques temporelles

Les caractéristiques temporelles se sont défini par les relations temporelles entre les différentes ondes ECG (P, QRS, et T) reflètent les variations de battement qui peuvent être utilisées comme une signature biométrique.

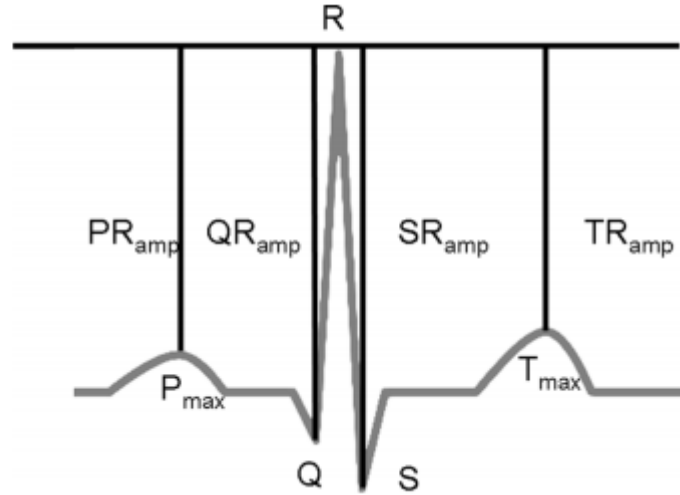
Dans la figure 2.4, la localisation des points spécifiques permet de distinguer de plusieurs intervalles temporels. Les caractéristiques temporelles utilisées sont les ondes de pulsations cardiaques durée (c'est-à-dire P, QRS, T) et intervalles mesurés (PQ, RS, ST, etc.). [19]



**Figure 2.4 :** Caractéristiques temporelles: sont obtenus en localisant des fiduciaires durant un intervalle du temps d'un rythme cardiaque.

### 4.1.2 Caractéristiques d'amplitude

La variabilité inter-classe d'amplitude du rythme cardiaque localisée dans des ondes peuvent être facilement identifier chez les individus. Les caractéristiques d'amplitude sont mesurées par la hauteur d'un pic R. Nous pouvons aussi distingues d'autres d'amplitudes mesurées dans l'enveloppe complexe PQRST comme ils sont montres dans la figure 2.5. [19]



**Figure2.5 :**Caractéristiques d'amplitude: les amplitudes relatives peuvent être obtenues par rapport au pic R.

#### 4.1.3 Caractéristiques morphologiques

Les attributs morphologiques sont ceux qui portent des informations liées à la forme de l'ECG, par mesure du mécanisme de la variation géométrique située dans un PQRST (P-QRS-T).

Une façon simple, nous pourrions extraire les caractéristiques morphologiques d'un battement de cœur en effectuant un alignement des segments QRS répétés mesurant dans la durée d'un signal ECG, cet alignement s'est fait par rapport au pic R. En revanche, un alignement centré sur un pic R, s'explique moins d'inter-variation, néanmoins, un désalignement s'exprime de plus d'inter-variation. [19]

Cependant, l'étude de la morphologie de l'ECG par utilisation des pentes des segments ST, RS, et les angles dérivés par les ondes Q, R et S a montré que les attributs extraits sont très sensibles aux facteurs liés au enregistrement d'un signal ECG et les anomalies du cœur. Les Palaniappan et Krishnan [28] ont introduit un attribut particulier calculé à partir de la morphologie du complexe QRS. Le facteur de forme QRS (FF) est défini comme dans l'équation (2.1) :

$$FF = \sqrt{\frac{\text{var}(x'') / \text{var}(x')}{\text{var}(x') / \text{var}(x)}} \quad (2.1)$$

où  $x$  est la forme d'onde complexe QRS,  $x'$  est la première dérivée du QRS et  $x''$  est sa deuxième dérivée seconde.

Toutes les caractéristiques décrites ci-dessus présentent un inconvénient principal: l'onde cardiaque entière change avec la fréquence cardiaque (HR). Ceci, à son tour, induit des fluctuations dans les relations temporelles, d'amplitude et morphologiques entre les vagues pendant l'activité physique, la prise de médicaments ou les fortes émotions.

De nombreux auteurs ont donc tenté de normaliser les caractéristiques en ce qui concerne les changements de RH. [19]

Israël et al [29] ont supposé l'existence d'une relation linéaire entre la fréquence cardiaque et les caractéristiques (temporelles) calculées par rapport au pic R. Pour éviter le choix critique de la fréquence cardiaque, ils utilisent une simple normalisation des caractéristiques à la longueur de l'ensemble du complexe P-QRS-T [19].

Shen [30] et Singh[31] ont utilisé la formule de Bazett pour normaliser l'intervalle QT, Tawfik et Sansone ont appliqué l'approche par régression proposée par Sagie [19].

Fatemian et Hatzinakos [37] ont proposé une nouvelle normalisation de l'ECG de fréquence cardiaque en rééchantillonnant l'onde T pour une durée de 120 ms. Après le rééchantillonnage des ondes T, ils ont combiné le segment obtenu avec les ondes P et QRS afin de composer l'ensemble du rythme cardiaque. Enfin, le rythme cardiaque entier est rééchantillonné à 850 ms. [32]

### **4.2 Approches non fiduciaires**

Les caractéristiques temporelles, d'amplitude et morphologiques sont influées par l'efficacité de l'algorithme d'appliquer et les variations intrinsèques et extrinsèques d'acquisition d'un signal ECG. Pour surmonter le problème, les approches fréquentielles, qui ne demandent pas des informations à post-priori sur les informations fiduciaires d'un signal ECG.

Une de ces approches, l'approche de Fourier qui est basée sur l'hypothèse de la périodicité et quasi périodicité d'un signal ECG afin d'extraire une composante fréquentielle invariante dans le temps. Nous pouvons classer ces approches en trois grandes classes : autocorrélation, la phase spatiale et la fréquence. [19]

#### **4.2.1 Caractéristiques basées sur l'autocorrélation**

La première approche non fiduciaire a été proposée par Plataniotis et Hatzinakos[33]. Afin d'extraire des attributs pertinents, ils sélectionnent au hasard 5 segments  $s$  à partir d'un enregistrement ECG et estiment l'autocorrélation normalisée (AC) sur une fenêtre de décalages (voir l'équation 2). AC incorpore des informations sur les caractéristiques particulières de l'ECG: il est invariant par décalage et approprié au modèle aléatoire. Le complexe QRS, en



particulier, conserve une forte invariance de la forme et de la largeur du temps. Avec cette approche, les échantillons qui devraient être influencés par la détection des repères sont combinés en une séquence de sommes de produits comme indiqué dans Eq (2.2) mentionnée ci-dessous:

$$r[m] = \frac{1}{r[0]} \sum_i s[i] s[i+m] \quad (2.2)$$

Dans l'équation,  $r[m]$  est l'AC,  $s[i]$  est le signal à l'instant  $i$  et  $m$  est choisi plus grand que la durée moyenne du QRS (en échantillons).

#### 4.2.2 Analyse de la phase spatiales

Comme pour l'approche basée sur AC, le signal ECG peut être caractérisé dans un espace bidimensionnel ou même tridimensionnel en utilisant la technique de time-delay.

L'analyse de la trajectoire de l'espace de phase peut en effet mettre en évidence des particularités inexplorées de l'activité cardiaque.

Fang [34] a représenté les données dans un nouvel espace afin bien discriminer les données ECG dans un espace trois dimensions ( $s(t)$ ,  $s(t+dt)$ ,  $s(t+2dt)$ )

Coutinho [35] a analysé les similitudes des ECG en utilisant un algorithme de matching. Chaque enregistrement est converti en appliquant une quantification uniforme de 8 bits, qui produit une séquence de symboles (chaînes) à partir d'un alphabet de  $2^8$  symboles. Ensuite, un algorithme d'analyse de Ziv-Merhav (ZMCP) a été utilisé pour évaluer la complexité croisée entre les différentes chaînes. L'idée principale de l'utilisation de ZMCP est que la complexité croisée devient plus faible lorsque les deux séquences sont similaires. [19]

#### 4.2.3 Caractéristiques basées sur la fréquence

D'autres auteurs se penchent sur l'analyse des caractéristiques d'ECG par des techniques fréquentielles.

Loong et al [38] ont utilisé une technique prédictive linéaire (codage prédictif linéaire ou LPC) pour modéliser le contenu fréquentiel des signaux ECG. Plus loin, le modèle de spectre pour chaque sujet est obtenu en considérant les quarante premiers points du reconstruction du spectre ECG en utilisant Eq (2.3):

$$\hat{x}[n] = -\sum_{i=1}^p a_i x[n-i] \quad (2.3)$$

Où les coefficients  $a_i$  sont évalués en minimisant l'erreur  $e[n]$  (voir l'équation 2.4) en utilisant la récursivité de Levinson-Durbin :

$$e[n] = x[n] - \hat{x}[n] \quad (2.4)$$

où  $x[n]$  représente la valeur réelle.

Kouchaki et Zhao [39] ont utilisé des procédures similaires à la transformation de Hilbert-Huang pour obtenir des données ECG dans un espace fréquentielle. Les signaux originaux ont été réduits par un processus adaptatif appelé Décomposition Empirique en Mode (EMD). [19]

## 5. Méthodes de Classification

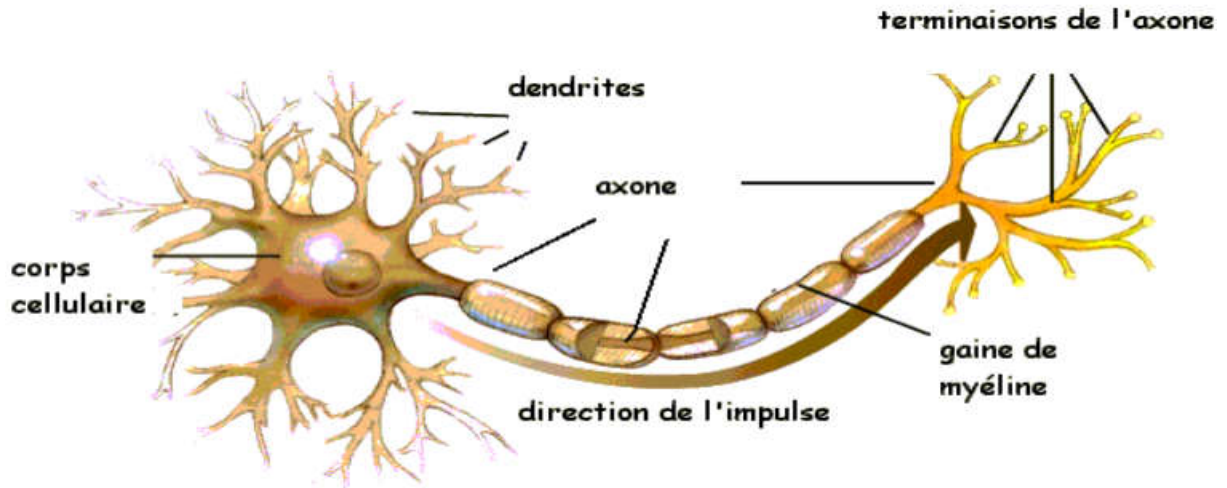
### 5.1 Réseaux de neurones

Les réseaux de neurones artificiels (RNA) sont des réseaux fortement connectés de processeurs une sortie unique sur la base des informations qu'il reçoit. Ils infèrent des propriétés émergentes permettant de solutionner les problèmes jadis qualifiés de complexe. Dans un réseau, chaque sous-groupe fait un traitement indépendant des autres et transmet le résultat de son analyse au sous-groupe suivant. L'information donnée au réseau va donc se propager de la couche d'entrée à la couche de sortie, en passant soit par aucune, une ou plusieurs couches intermédiaires. Il est à noter qu'en fonction de l'algorithme d'apprentissage. [20]

#### 5.1.1 Neurone biologique

Le neurone est une cellule composée d'un corps cellulaire et d'un noyau. Le corps cellulaire se ramifie pour former ce que l'on nomme les dendrites. Celles-ci sont parfois si nombreuses que l'on parle alors de chevelure dendritique ou d'arborisation dendritique. C'est par les dendrites que l'information est acheminée de l'extérieur vers le soma, corps du neurone.

L'information traitée par le neurone chemine ensuite le long de l'axone (unique) pour être transmise aux autres neurones. La transmission entre deux neurones n'est pas directe. En fait, il existe un espace intercellulaire de quelques dizaines d'Angstrom ( $10^{-9}\text{m}$ ) entre l'axone du neurone afférent et les dendrites (on dit une dendrite) du neurone efférent. [21]



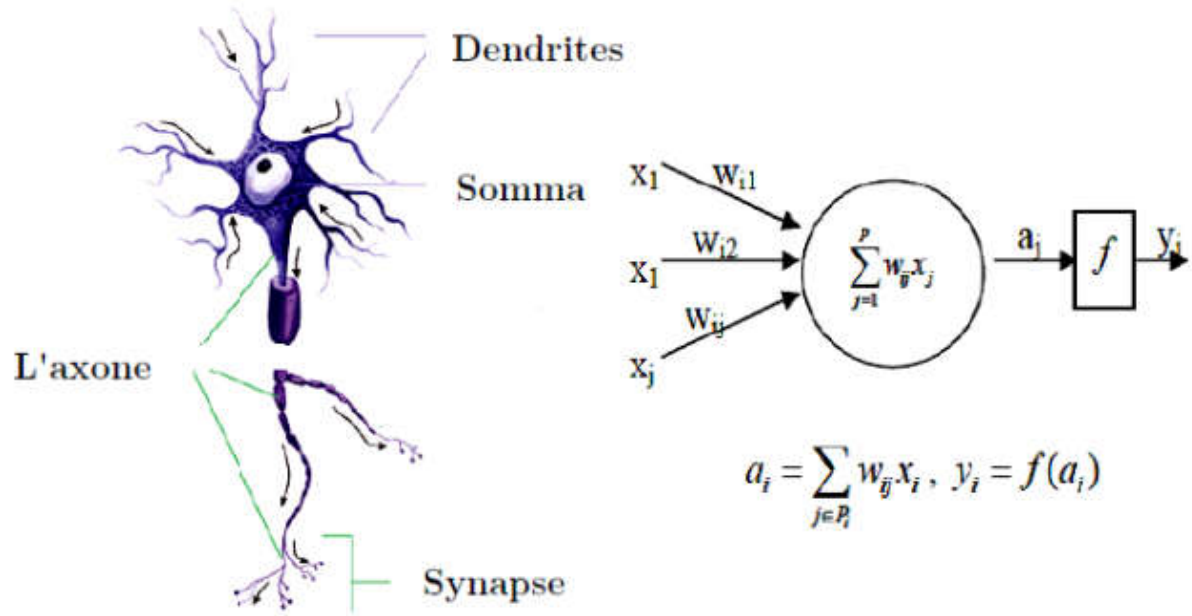
**Figure 2.6 :** Neurone biologique.

Ainsi les axones des neurones (sorties) sont connectés aux dendrites d'autres neurones par les synapses, ce qui fait de notre cerveau un réseau extrêmement complexe de neurones, compte tenu que nous possédons plusieurs milliards de neurones et que chaque neurone peut avoir plusieurs milliers de dendrites. Les synapses jouent un rôle important car elles sont le lien entre la sortie d'un neurone et l'entrée d'un autre neurone. L'influx nerveux passe par ces synapses, qui peuvent inhiber ou laisser passer le signal nerveux. Si l'amplitude du signal arrivant à la synapse est assez élevée, alors elle laissera passer le signal, sinon il sera inhibé. La synapse est capable de s'adapter pour laisser passer le signal ou empêcher son passage : ce mécanisme d'adaptation est à la base de l'apprentissage de notre cerveau. [21]

### 5.1.2 Neurone formel

Un "neurone formel" (ou simplement "neurone") est une fonction algébrique non linéaire et bornée, dont la valeur dépend de paramètres appelés coefficients ou poids. Les variables de cette fonction sont habituellement appelées "entrées" du neurone, et la valeur de la fonction est appelée sa "sortie".

Un neurone est donc avant tout un opérateur mathématique, dont on peut calculer la valeur numérique par quelques lignes de logiciel. On a pris l'habitude de représenter graphiquement un neurone comme indiqué sur la Figure 2.3 [22]



**Figure 2.7 :** Neurone biologique et neurone formel.

Avec :  $x_i$  : paramètres d'entrées.

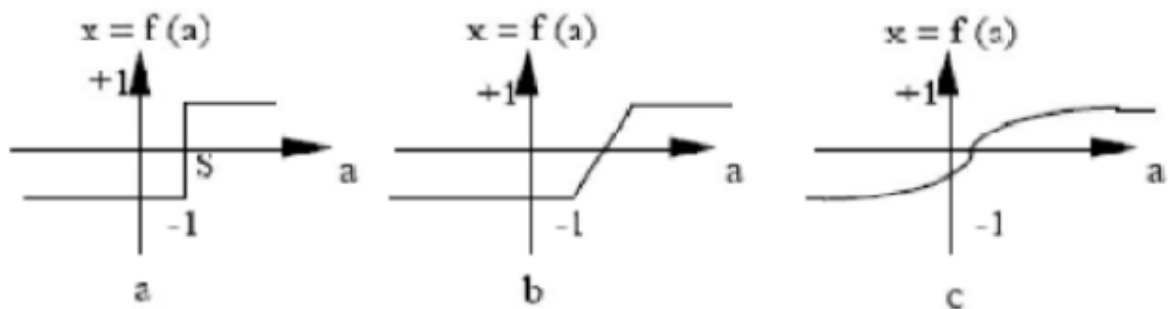
$w_{ij}$  : poids synaptiques.

$a_i$  : potentiel neuronal.

$y_i$  : sortie du neurone  $i$ .

$f$  : la fonction d'activation.

Il existe de nombreuses formes possibles pour la fonction d'activation. Les plus courantes sont présentées sur la (Figure 2.8). On remarquera qu'à la différence des neurones biologiques dont l'état est binaire, la plupart des fonctions d'activations sont continues, offrant une infinité de valeurs possibles comprises dans l'intervalle  $[0, +1]$  ou  $[-1, +1]$ .



**Figure 2.8 :** Différents types de fonctions d'activation pour le neurone formel,

a : fonction à seuil (S : valeur du seuil), b : linéaire par morceaux, c : sigmoïde.

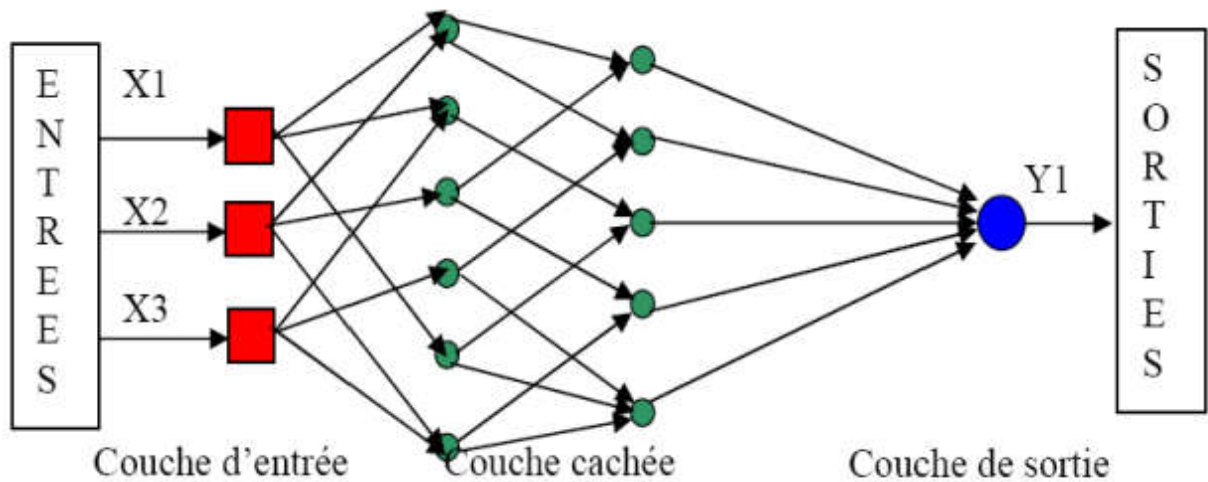
Une collection de neurones formels est appelée réseaux de neurones. Ces réseaux sont souvent constitués de couches de neurones interconnectés.

### 5.1.3 Architecture des réseaux de neurones formels

On distingue deux grands types d'architectures de réseaux de neurones : les réseaux de neurones non bouclés et les réseaux de neurones bouclés.

#### 5.1.3.1 Réseaux non bouclée (Statique)

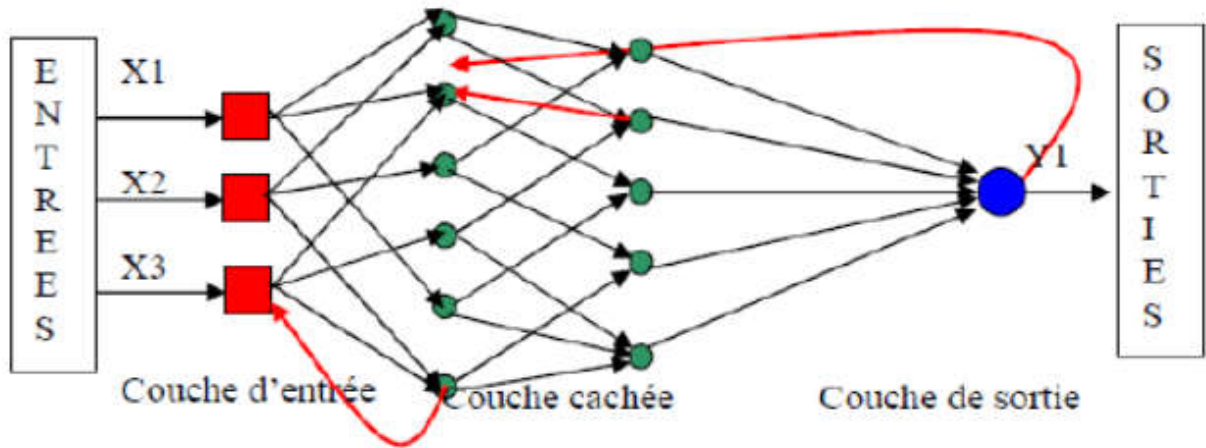
Réseaux d'avance permettent aux signaux de voyager d'une manière unique ; De l'entrée à la sortie. Il n'y a pas de commentaires (boucles), c'est-à-dire que la sortie de n'importe quelle couche n'affecte pas cette même couche. Les réseaux d'avance ont tendance à être des réseaux directs qui associent les entrées aux résultats. Ils sont largement utilisés dans la reconnaissance de motifs. Ce type d'organisation est également appelé Bas - En haut ou Haut - vers le bas. [23]



**Figure 2.9 :** Réseau de neurone non bouclé.

#### 5.1.3.2 Réseaux bouclée (Dynamique)

Les réseaux de rétroaction peuvent transmettre des signaux dans les deux sens en introduisant des boucles dans le réseau. Les réseaux de rétroaction sont très puissants et peuvent devenir extrêmement compliqués. Les réseaux de rétroaction sont dynamiques; Leur «état» change continuellement jusqu'à ce qu'ils atteignent un point d'équilibre. Ils restent au point d'équilibre jusqu'à ce que l'entrée change et qu'un nouvel équilibre soit nécessaire. Les architectures de rétroaction sont également appelées interactives ou récurrentes, bien que ce dernier terme soit souvent utilisé pour désigner des connexions de rétroaction dans des organisations à une seule couche. [23]



**Figure 2.10 :** Réseau de neurone bouclé.

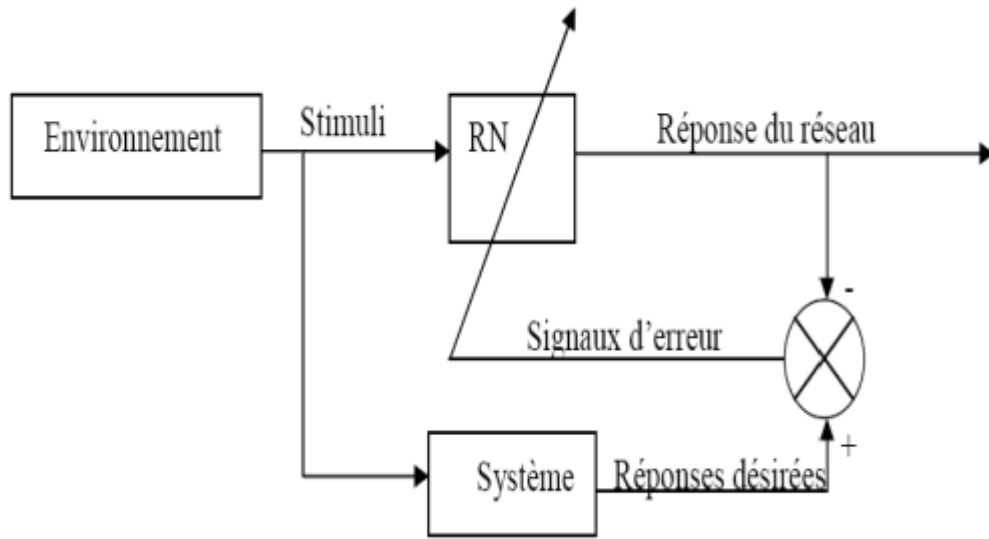
#### 5.1.4 L'apprentissage des réseaux de neurones

On appelle « apprentissage » des réseaux de neurones la procédure qui consiste à estimer les paramètres des neurones du réseau, afin que celui-ci remplisse au mieux la tâche qui lui est affectée.

On peut distinguer deux types d'apprentissages : l'apprentissage « supervisé » et l'apprentissage « non supervisé »

##### 5.1.4.1 L'apprentissage supervisé

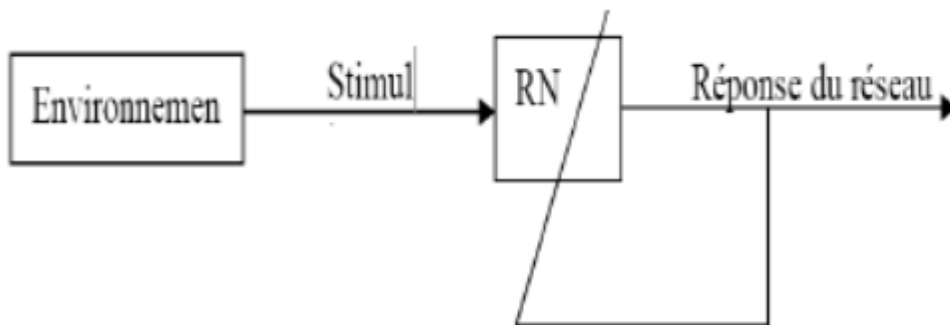
L'apprentissage supervisé consiste à présenter au réseau une série d'exemples sur ses entrées, ainsi le réseau calcule sur ses sorties un vecteur d'état qui est comparé au vecteur des états désirés. Une erreur est calculée entre le vecteur des sorties désirées et le vecteur des sorties calculées par le réseau. L'algorithme d'apprentissage adapte les poids du réseau de façon à corriger son comportement comme le montre la figure 2.11. [24]



**Figure 2.11 :** Apprentissage supervisé.

#### 5.1.4.2 Apprentissage non supervisé

L'apprentissage est qualifié de non supervisé lorsque, seules les valeurs d'entrée sont disponibles. Il repose sur un critère interne de conformité du comportement du réseau par rapport à des spécifications générale et non sur des observations externes. Dans ce cas, l'apprentissage est basé sur des probabilités. Le réseau va se modifier en fonction des régularités statiques de l'entrée. L'algorithme d'apprentissage exploite pour se faire une mesure prédéterminée de la qualité de représentation de la connaissance afin d'ajuster les paramètres du réseau. Ce type d'apprentissage est représenté schématiquement en Figure 2.12. [24]



**Figure 2.12 :** apprentissage non supervisé.

### 5.1.5 Le réseau de neurones de fonctions à base radial (RBF)

La structure typique d'un réseau de neurones RBF (Figure 2.13) se compose de trois couches à savoir: une couche d'entrée, une couche cachée et une couche de rendement. Chaque couche se compose d'un certain nombre de neurones (nœuds). Dans la couche d'entrée, les nœuds sont employés pour faire passer seulement les données d'entrée à la couche cachée. Aucun calcul n'est exécuté dans les nœuds de la couche d'entrée. La couche cachée contient les  $k$  nœuds, qui appliquent une transformation non linéaire sur les variables d'entrées. [25]

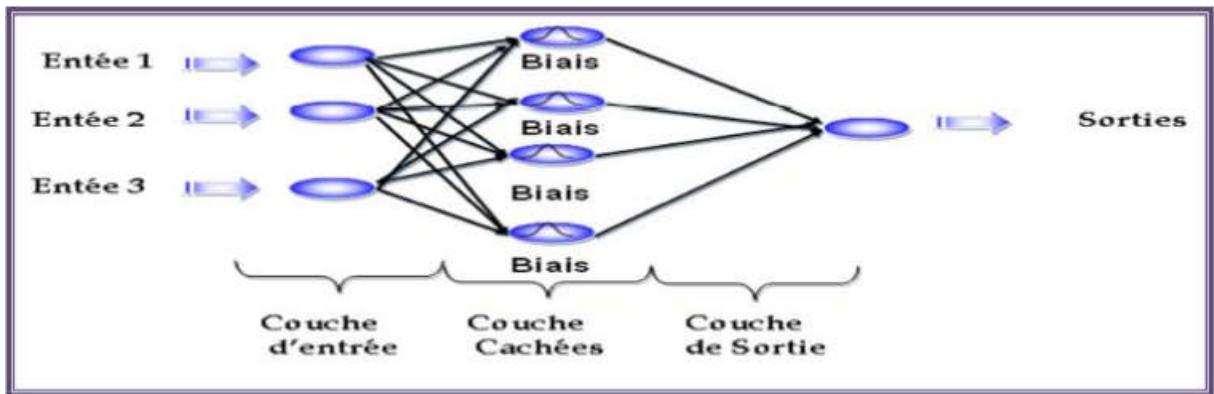


Figure 2.13 : Architecture d'un réseau de neurone RBF.

### 5.1.6 Le réseau de neurones probabiliste (PNN)

Le réseau de neurones probabiliste PNN (Probabilistic Neural Network) introduit par Donald Specht en 1988, est un réseau feed-forward à 3 couches utilisé pour la classification des données. À l'inverse des autres réseaux de neurones qui sont basés sur la rétropropagation, le réseau de neurones probabiliste est basé sur des principes statistiques dérivés de la stratégie de décision de Bayes et des fonctions estimateurs de la densité de probabilité.

Le PNN emploie des fonctions gaussiennes à base radiale et sphériques centrées à chaque vecteur d'apprentissage. [26]

Pour l'apprentissage d'un réseau probabiliste, il faut d'abord que chaque vecteur  $x$  en entrée soit normalisé de façon que :  $\sum_{i=1}^n x_i^2 = 1$ . Le nombre de neurones d'entrée est égal au nombre des paramètres (variables) de la forme à classifier.

Pour chaque échantillon de la base d'apprentissage, on crée un neurone dans la couche cachée (couche de motifs) avec les connections correspondantes avec les neurones d'entrée de façon que  $w_k = x_k$  pour  $k = 1, 2, \dots, n$ .



Ensuite une seule connexion est créée vers l'neurone de la couche de classification correspondant à la classe de l'échantillon. Ce processus est illustré dans l'algorithme suivant : [27]

**Début**

**Initialisation**

$j=0$  ;  $n$  = premier échantillon d'apprentissage ;

**Faire**

$j \leftarrow j+1$  ;

**Normalisation :**

$$x_{jk} \leftarrow \frac{x_{jk}}{\sqrt{\sum_i^d x_{jk}^2}}$$

**Apprentissage des poids :**  $w_{jk} \leftarrow x_{jk}$  ;

**Si**  $x \in w_j$  **alors**  $a_{ic} \leftarrow 1$  ;

**Jusqu'à**  $j = n$  ;

## 5.2 Le classificateur du K plus proches voisins (KNN)

C'est un classificateur simple basé sur le calcul de distance entre les exemples d'apprentissage et les exemples de tests, généralement la norme euclidienne est souvent employée comme mesure de distance, dans chaque étape de l'apprentissage, l'algorithme mémorise les k meilleurs exemples de l'ensemble d'apprentissage ( $kppv(x)$ ) qui sont proches à l'exemple de test  $x$ . Cet algorithme est souvent performant s'il y a suffisamment d'exemples d'apprentissage, mais demande un temps de prédiction très long pour passer tous les exemples afin de trouver les K meilleures solutions. [26].

***L'algorithme***

Soit  $L = \{(x', c) | x' \in R_d, c \in C\}$  est un ensemble d'apprentissage.

( $x'$  est un exemple d'apprentissage et  $c$  sa classe qui lui est associée.

Soit  $x$  l'exemple de test dont on souhaite déterminer sa classe

**Début**

**Pour chaque** (*exemple*  $(x', c) \in L$ ) **faire**

Calculer la distance  $D(x, x')$

**Fin**

**Pour chaque**  $\{x' \in kppv(x)\}$  **faire**

Compter le nombre d'occurrences de chaque classe

**Fin**

Attribuer à  $x$  la classe la plus fréquente

**Fin**

## Conclusion

Dans ce chapitre, nous avons présenté l'historique de l'ECG ainsi qu'un bref rappel anatomique du système cardiovasculaire et de l'ECG. Nous avons présenté les grandes catégories d'extraction des caractéristiques d'ECG mentionnés dans la littérature. Les méthodes de classification sont considérées comme une tâche importante dans n'importe quel système biométrie, on présente les méthodes de classification (ML) le plus utilisées dans notre contexte à savoir, RN et KNN.