

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université 8Mai 1945 – Guelma
Faculté des sciences et de la Technologie
Département d'Electronique et Télécommunications



**Mémoire de fin d'étude
Pour l'obtention du diplôme de Master Académique**

Domaine : **Sciences et Technologie**
Filière : **Electronique**
Spécialité : **Instrumentation**

**Classification des signaux cardiaques en utilisant les arbres de
décision**

Présenté par :

Bessioud alaa eddine

Zenalabidine Mahamat Tahir

Sous la direction de :

Dr.nemessi mohammed

Septembre 2020

Remerciements

Nos remerciements vont premièrement à Allah tout puissant pour la volonté, la santé et la patience qu'il nous a donné durant toutes ces longues années.

Ainsi, nous tenons aussi à exprimer nos remerciements à notre encadreur le Dr Nemessi pour nous avoir proposé un tel sujet d'actualité, pour sa disponibilité tout au long de la réalisation de ce mémoire et pour ses conseils et remarques.

Nos remerciements vont aussi à tous les enseignants et enseignantes du département d'électronique qui ont contribués à notre formation.

Ont tiens à remercier toutes les personnes qui nous ont soutenues durant la période de nos études.

On remercie également nos parents pour leurs conseils et soutiens financières et morales.

Résumé

L'électrocardiogramme (ECG) est un test non invasif largement utilisé pour analyser les arythmies cardiaques. L'application des méthodes informatisé d'aide à la décision constitue des solutions permettant aux cardiologues de mieux analyser les enregistrements ECG. Ce travail s'inscrit dans ce contexte, notre objectif est de développer un système de reconnaissance automatique des battements ECG. Dans ce système, nous utilisons les arbres de décision en tant que classificateur basé sur cinq caractéristique temporels. D'autre part, ce système vise à classifier quatre classes de rythmes cardiaques : Battement normal, Extrasystole ventriculaire, Bloc de branche droit et Bloc de branche gauche.

Abstract :

The electrocardiogram (ECG) is a non-invasive test widely used to analyze cardiac arrhythmias. The application of computerized decision support methods constitutes solutions that allow cardiologists to better analyze ECG recordings. In this context, our goal is to develop an automatic recognition system for ECG beats. In this system, we use decision trees as a classifier based on five temporal characteristics. This system aims to classify four classes of heart rhythms: Normal beat, Ventricular extrasystole, Right bundle branch block and Left bundle branch block.

ملخص

مخطط كهربية القلب (ECG) هو اختبار غير جراحي يستخدم على نطاق واسع لتحليل عدم انتظام ضربات القلب. يشكل تطبيق طرق دعم القرار الأوتوماتيكية حلاً لتسمح لأطباء القلب بتحليل تسجيلات مخطط كهربية القلب بشكل أفضل. في هذا السياق، هدفنا هو تطوير نظام التعرف التلقائي على دقات تخطيط القلب. في هذا النظام، نستخدم أشجار القرار كمصنف بناءً على خمس خصائص زمنية. يهدف هذا النظام إلى تصنيف أربع فئات من إيقاعات القلب: النبض الطبيعي، والانقباض البطيني الإضافي، وإحصار الحزمة اليمنى وإحصار الحزمة اليسرى.

Table des matières

Introduction générale	1
Chapitre I	
I.1 Introduction sur le système cardiovasculaire.....	4
I.2 Anatomie du cœur	4
I.2.1 Fonctionnement mécanique (La conduction mécanique du cœur).....	5
I.2.2 Fonctionnement électrique (La conduction électrique du cœur).....	6
I.3 L'électrocardiographie	7
I.3.1 Principe.....	7
I.3.2 Les dérivations	8
I.4 Intervalles temporelles (Ondes constituant un ECG).....	10
I.4.1 L'onde P.....	11
I.4.2 Le complexe QRS.....	11
I.4.3 L'onde T.....	11
I.5 Les Rythmes cardiaques	12
I.5.1 Le Rythme sinusal	12
I.5.2 Battements prématurés.....	13
I.5.3 Arythmie auriculaire.....	13
I.5.4 Arythmies ventriculaires	14
I.5.5 Trouble de la conduction cardiaque.....	14
Chapitre II	16
II.1 Catégories des appareils d'enregistrement des ECG.....	16
II.2 Acquisition des signaux ECG.....	17
II.2.1 Schéma général.....	17
II.2.2 Amplificateur de mesure.....	18
II.2.3 Amplificateur d'isolation.....	19
II.2.4 Filtre passe haut.....	19
II.2.5 Filtre passe bas.....	19
II.3 Les électrodes.....	20
II.3.1 L'interface peau-électrode.....	20

II.3.2 Types d'électrodes.....	20
a) Electrodes à plaques.....	20
a) Electrodes d'aspiration.....	21
c) Electrode à colonne de fluide.....	22
d) Electrodes secs.....	22
e) Electrodes actives.....	23
II.4 Artéfacts et bruits dans l'ECG.....	23
II.4.1 Bruit dû au secteur.....	23
II.4.2 Mouvements de la ligne de base.....	24
II.4.3 Bruit musculaire.....	24
Chapitre III	
III.1 Introduction au processus général de la reconnaissance des formes.....	27
III.2 Processus la classification des ECG.....	27
III.2.1 Prétraitement du signal ECG.....	28
III.2.2 Détection des Complexe QRS et segmentation.....	28
III.2.2 Extraction des caractéristiques.....	31
III.2.3 Classification.....	31
III.3 Classification par l'arbre de décision.....	32
III.3.1 Principe.....	32
III.3.2 Critères de division.....	32
a) Indice de Gini.....	32
b) Entropie.....	33
III.3.3 Critères d'arrêt.....	34
III.3.4 Algorithmes d'arbre de décision.....	34
a) ID3.....	34
b) IC4.5.....	34
c) CART.....	34
III.3.5 Exemple d'arbre de décision.....	35
III.3.6 Avantages et inconvénients des arbres de décision.....	36
Chapitre IV	
IV. 1 Introduction.....	39

IV. 2 Processus.....	39
IV.3 Exemple de classification avec les arbres de décision.....	41
IV.4 Description de la base de données.....	42
IV.5 Résultats.....	43
IV.5.1 1er test : Classe N et Classe V.....	43
IV.5.2 2ème test : Classe N et Classes R et L.....	46
IV.5.3 3ème test : Classe N et Classes R et L et V.....	47
IV. Conclusion.....	49
Conclusion générale.....	50

Liste des figures

Figure I.1	Anatomie du cœur.....	5
Figure I.2	Activité électrique du cœur.....	7
Figure I.4	Les dérivations bipolaires des membres.....	8
Figure I.5	Les dérivations unipolaires des membres.....	9
Figure I.6	Les dérivations précordiales.....	10
Figure I.7	Ondes du signal ECG.....	10
Figure I.8	Rythme sinusal normal.....	12
Figure II.1	Exemple de quelques Catégories des appareils.....	16
Figure II.2	Chaîne du traitement de l'ECG.....	17
Figure II.3	Amplificateur d'instrumentation ECG à deux étages.....	18
Figure II.4	Deux exemples d'électrodes en plaque utilisées en électrocardiographie.	21
Figure II.5	Electrode d'aspiration.....	22
Figure II.6	Signal électrocardiographie perturbé par le secteur	23
Figure II.7	Mouvement de la ligne de base.....	24
Figure II.8	Bruit musculaire.....	25
Figure III.1	Schéma de base d'un processus de RDF.....	27
Figure III.2	Un diagramme du système de classification des arythmies.....	27
Figure III.3	Point de référence et intervalles des battements cardiaques	30
Figure III.4	Exemple d'échantillons, de la courbe de l'ECG, utilisées en tant que caractéristiques	31
Figure III.5	L'arbre de décision correspondante à l'exemple du joueur.....	36
Figure IV.1	Processus générale.....	39
Figure IV.2	Les 5 caractéristiques utilisées	40
Figure IV.3	Détection des Ondes R et P et de début et fin du complexe QRS.....	40
Figure IV.4	Classification de l'exemple synthétique, frontières induites par l'arbre de décision	41
Figure IV.5	L'arbre de décision pour la classification de l'exemple 2D.....	41
Figure IV.6	Exemples de battements de la base de données MIT-BIH.....	43
Figure IV.7	un exemple d'arbre de décision correspondant au premier test.....	45
Figure IV.8	Les taux d'apprentissage et de test en fonction du nombre de nœuds.....	46

Liste des tableaux

Tableau III.1	Les valeurs standards pour les intervalles d'un ECG normal.....	30
Tableau III.2	Données d'apprentissage de l'exemple du joueur.....	35
Tableau IV.1	les résultats obtenus pour le premier test.....	44
Tableau IV.2	les résultats obtenus pour le 2 ^{ème} test.....	47
Tableau IV.3	les résultats obtenus pour le 3 ^{ème} test	48

Introduction générale

L'électrocardiogramme (ECG) est un test standard pour surveiller l'activité du cœur. De nombreuses anomalies cardiaques se manifestent dans l'ECG, y compris l'arythmie qui est un terme général caractérisant les rythmes cardiaques anormaux. La base du diagnostic d'arythmie est l'identification des battements cardiaques individuels normaux et anormaux, et leur classification correcte en différents diagnostics, en fonction de la morphologie de l'ECG.

Le processus d'identification et de classification des arythmies peut être très difficile pour les cardiologues car il est parfois nécessaire d'analyser chaque battement cardiaque de l'enregistrement ECG, acquis par un moniteur holter par exemple, pendant des heures, voire des jours. De plus, il existe un risque d'erreur humaine lors de l'analyse des enregistrements ECG, en raison de la fatigue. Une alternative consiste à utiliser des techniques classifications automatiques.

L'objectif de ce travail est de proposer un système de reconnaissance automatique des battements ECG. Dans ce système, les caractéristiques utilisées sont basées les ondes et les intervalles temporels. Pour l'étape de classification nous utilisons les arbres de décision. En effet, les arbres de décision présentent un avantage majeur qui consiste en leur explicité, ils peuvent être facilement interprétés par des utilisateurs non professionnels. De plus, les arbres de décision peuvent être convertis en un ensemble de règles compréhensible.

Le présent mémoire comporte quatre chapitres :

Le premier chapitre constitue une introduction à l'électrocardiographie. Il donne de brefs aperçus sur le système cardiovasculaire, l'électrocardiographie et les rythmes cardiaques.

Le deuxième chapitre s'attache aux aspects techniques de l'électrocardiographie. Il décrit les principaux composants des systèmes d'acquisition des signaux ECG. Un accent particulier est mis sur les électrodes et leurs types. Ce chapitre donne également un aperçu sur les artefacts et bruits dans l'ECG.

Le troisième chapitre concerne le processus général de la classification du battement ECG. La première partie de ce chapitre décrit les différentes étapes de ce processus : le prétraitement, la segmentation, l'extraction des caractéristiques et la classification. La deuxième partie de ce chapitre décrit les arbres de décision en évoquant leurs principes et algorithmes ainsi que leurs avantages et inconvénients.

Le quatrième chapitre présente le système proposé, classification des battements ECG en utilisant les arbres de décision, et analyse les résultats obtenus sur la base de données des arythmies (MIT-BIH).

Finalement, une conclusion conclue ce mémoire

Chapitre I :

Introduction à

l'électrocardiographie

I.1 Introduction sur le système cardiovasculaire

Un électrocardiogramme (ECG) décrit l'activité électrique du cœur enregistré par des électrodes placées sur la surface du corps. Les variations de tension mesurées par les électrodes sont causées par les potentiels d'action des cellules cardiaques excitable en contractant. Le rythme cardiaque résultant dans l'ECG se manifeste par une série d'ondes dont la morphologie et le timing transmettent des informations permettant de diagnostiquer les maladies qui se reflètent par des perturbations de l'activité électrique du cœur. Le modèle temporel caractérisant la succession de battements cardiaques est également très important pour le diagnostic des maladies [1].

I.2 Anatomie du cœur

Le cœur est un organe musculaire de la taille d'un gros poing dont la fonction principale est de pomper du sang riche en oxygène dans tout le corps. Son anatomie est divisée en deux côtés "en miroir", gauche et droite, qui soutiennent de différents systèmes circulatoires mais qui pompent de manière synchronisée et rythmée. Chaque côté du cœur se compose de deux chambres, l'oreillette où le sang entre et le ventricule où le sang est forcé à circuler plus loin (figure I.1).

Les deux côtés sont séparés par une paroi musculaire appelée septum. La direction du flux sanguin est contrôlée par quatre valves différentes qui sont situées entre les oreillettes et les ventricules (valves auriculo-ventriculaires) et entre les ventricules et les artères (valves pulmonaires et aortiques).

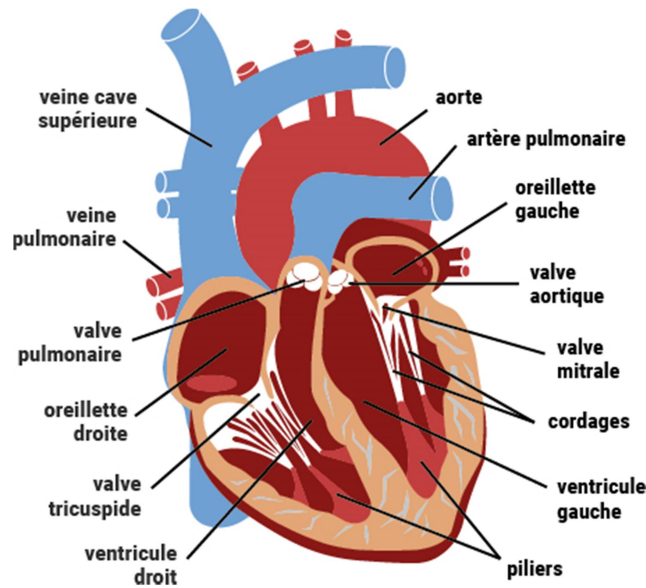


Figure I.1 : Anatomie du cœur [2]

I.2.1 Fonctionnement mécanique (La conduction mécanique du cœur)

La paroi du cœur est appelée le myocarde et est principalement composée des cellules musculaires qui produisent une force mécanique lors de la contraction du cœur. Le myocarde contient également des cellules musculaires spécialisées qui sont connecté à un réseau (système de conduction) qui permet une impulsion électrique pour se propager rapidement dans tout le cœur. Un cycle cardiaque est créé quand une telle impulsion se propage à travers le système de conduction. L'électricité l'impulsion est l'événement qui déclenche la force mécanique, et donc là un événement électrique précède la contraction cardiaque.

La séquence d'événements mécaniques qui définit un cycle cardiaque peut être supposée commencer dans l'oreillette droite où le sang est collecté de tous les veines du corps sauf celles des poumons. Lorsque l'oreillette droite est déclenchée pour se contracter, il force le sang dans le ventricule droit. Quand le ventricule droit soit rempli, il se contracte et force le sang dans les poumons, où l'excès de dioxyde de carbone est remplacé par de l'oxygène. Les veines pulmonaires renvoient le sang oxygéné dans l'oreillette gauche qui se vide à son tour dans le ventricule gauche. En tant que pompe haute pression, le ventricule gauche force le sang à tous les organes et tissus du corps (sauf les poumons) à travers les vaisseaux artériels qui évoluent en capillaires et enfin en système veineux de retour.

I.2.2 Fonctionnement électrique (La conduction électrique du cœur)

L'initialisation d'un cycle cardiaque se produit dans une masse de cellules de stimulation cardiaque avec la capacité de déclencher spontanément une impulsion électrique. Ces cellules sont collectivement appelées le nœud sino-auriculaire (SA) et sont situées dans la partie supérieure de l'oreillette droite. L'impulsion électrique se propage ensuite à travers le système de conduction afin que la contraction et la relaxation auriculaire et ventriculaire puissent avoir lieu avec le bon timing. Après activation électrique des oreillettes droite et gauche, l'impulsion est recueillie et retardée au niveau du nœud auriculo-ventriculaire (AV) avant qu'elle n'entre dans les ventricules. Le retard permet la contraction de l'oreillette pour augmenter encore le volume sanguin dans les ventricules avant que la contraction ventriculaire se produise. Le retard dans le nœud AV est causé par une conduction ralentie de l'impulsion par le tissu musculaire dans cette zone. L'impulsion entre dans le mur entre les deux ventricules au faisceau de His, c'est le seul endroit qui connecte électriquement les oreillettes, le nœud AV et les ventricules.

La voie est ensuite divisée en faisceaux à conduction rapide avec des branches aux ventricules gauche et droit, puis plus loin dans un vaste réseau de fibres de conduction spécialisées appelées « fibres de Purkinje ». La grande taille des deux ventricules nécessite que l'impulsion électrique se propage rapidement pour initier une contraction unifiée.

Le nœud SA est le stimulateur cardiaque naturel du cœur qui détermine le taux de battement. Étant donné que les cellules du nœud SA ont le stimulateur cardiaque le plus rapide, toutes les autres cellules suivent en synchronie.

L'activité électrique du cœur peut être caractérisée par des mesures acquises au niveau cellulaire ainsi qu'à la surface du corps.

L'ECG décrit les différentes phases électriques d'un cycle cardiaque et représente une sommation dans le temps et l'espace des potentiels d'action générés par des millions de cellules cardiaques.

Avant qu'un nouveau rythme cardiaque soit initié par le nœud SA, toutes les cellules cardiaques sont au repos.

Pendant la dépolarisation auriculaire, une onde auriculaire à polarité positive est générée dans l'ECG enregistré à la position de l'électrode d'exploration (figure). L'amplitude de l'onde résultante est faible car la masse musculaire des oreillettes qui produit l'onde électrique est relativement petite.

Une fois la dépolarisation des oreillettes terminée, l'ECG revient à la ligne isoélectrique où elle reste jusqu'à ce que les ventricules se dépolarisent. La dépolarisation du nœud AV et le faisceau de His commencent vers la fin de l'onde auriculaire mais ne produit aucune onde visible à l'ECG à cause des petites masses musculaires.

Les ondes associées à la dépolarisation ventriculaire sont beaucoup plus grandes que l'onde auriculaire puisque les ventricules ont une masse musculaire beaucoup plus importante.

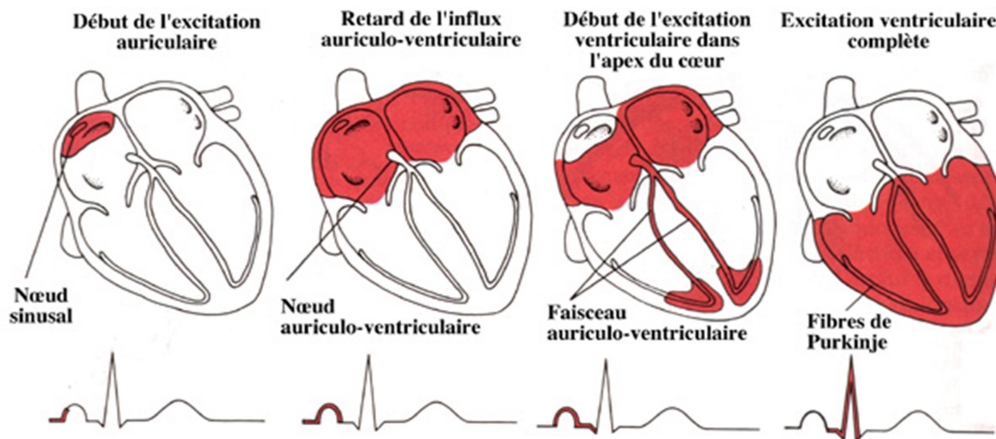


Figure I.2 : Activité électrique du cœur [3]

I.3 L'électrocardiographie

I.3.1 Principe

L'activité électrique du cœur est mesurée à la surface du corps en attachant un ensemble d'électrodes à la peau. Les électrodes sont positionnées de manière à ce que les variations spatio-temporelles du champ électrique cardiaque soient suffisamment bien réfléchies. Pour un enregistrement ECG, la différence de tension entre une paire d'électrodes est appelée une dérivation.

L'ECG est généralement enregistré avec une configuration multi-dérivations qui comprend des dérivations unipolaires ou bipolaires, ou les deux. Une dérivation unipolaire reflète la variation de tension d'une seule électrode et est mesurée par rapport à une électrode de référence qui est positionnée de sorte que la tension reste presque constante dans tout le cycle cardiaque. Une dérivation bipolaire reflète la différence de tension entre deux électrodes, par exemple, entre le bras gauche et droit.

I.3.2 Les dérivations

L'ECG standard à 12 dérivations est le plus utilisé dans la routine clinique et est défini par une combinaison de trois différentes configurations de dérivations : les dérivations bipolaires des membres, les dérivations-augmentés unipolaires des membres et les dérivations précordiales unipolaires.

L'ECG à 12 dérivations est enregistré en plaçant 10 électrodes à des positions normalisées sur la surface du corps. Les trois dérivations des membres bipolaires sont notées I, II et III et sont obtenues en mesurant la différence de tension entre le bras gauche, le bras droit et la gauche jambe dans les combinaisons suivantes:

- $DI = VL - VR$ (I.1)

- $DII = VF - VR$ (I.2)

- $DIII = VF-VL$ (I.3)

Où VLA, VRA et VLL indiquent la tension enregistrée sur le bras gauche, bras droit et jambe gauche, respectivement.

Étant donné que ces trois positions d'électrodes peuvent être vu comme les coins d'un triangle qui-angulaire ("triangle d'Einthoven") avec le cœur en son centre, les dérivations des membres qui en résultent décrivent activité dans trois directions différentes du plan frontal chaque direction est ainsi séparée d'un angle de 60°

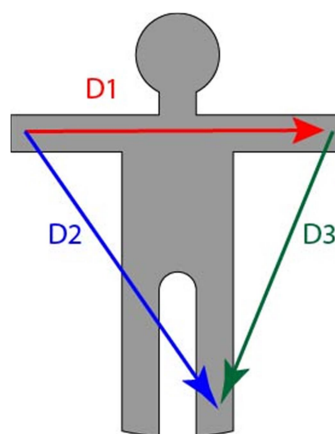


Figure I.4 : Les dérivations bipolaires des membres [4]

Les dérivations augmentées unipolaire des membres (aVF, aVL et aVR) ont été introduites pour compléter les 60 degré dans les directions des dérivations bipolaires des membres. Ces dérivations utilisent les mêmes électrodes que les dérivations des membres bipolaires, mais sont définies comme les différences de tension entre un coin du triangle et la moyenne des deux coins restants :

$$\bullet \quad aVR = VRA - \frac{VLA+V} {2} \quad (I.4)$$

$$\bullet \quad aVL = VLA - \frac{VRA+VLL} {2} \quad (I.5)$$

$$\bullet \quad aVF = VLL - \frac{VLA+VRA} {2} \quad (I.6)$$

Par conséquent, les dérivations des membres augmentés décrivent des directions qui sont décalés de 30 degrés de ceux des dérivations des membres bipolaire

Les dérivations augmentées des membres sont considérées comme unipolaires parce qu'une électrode explore alors que la moyenne des deux autres servent d'électrode de référence.

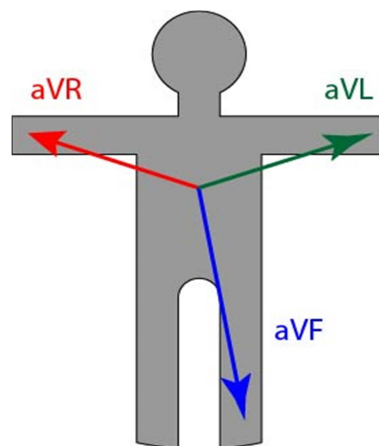


Figure I.5 : Les dérivations unipolaires des membres [4]

Les dérivations précordiales sont positionnées successivement sur les côtés avant et gauche de la poitrine afin de fournir une vue plus détaillée du cœur que les dérivations des membres. Les six dérivations précordiales, par convention étiquetés : V1, ..., V6, sont unipolaires et liés à une borne centrale qui est définie par la moyenne des tensions mesurées sur les bras droit et gauche et la jambe gauche.

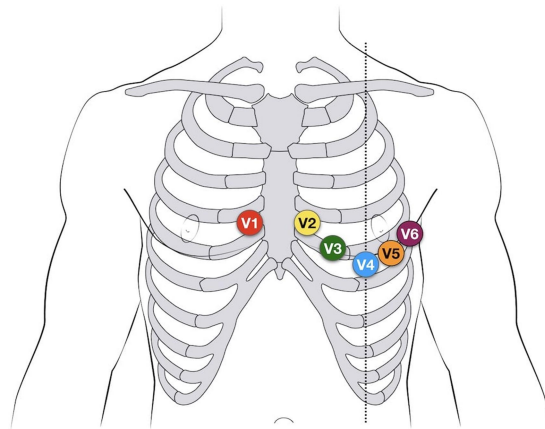
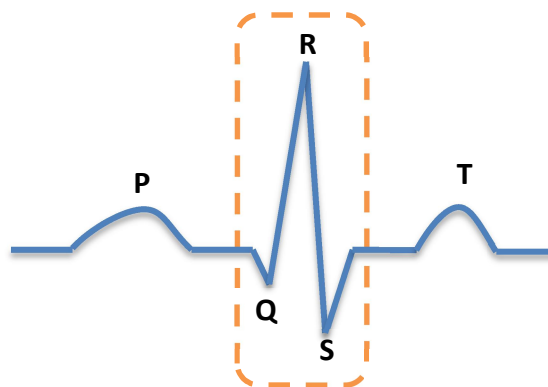


Figure I.6 : Les dérivations précordiales [5]

I.4 Intervalles temporelles (Ondes constituant un ECG)

Cette partie décrit quelques caractéristiques importantes des ondes ECG. La dépolarisation auriculaire est reflétée par l'onde P et ventriculaire la dépolarisation est reflétée par le complexe QRS, tandis que l'onde T reflète repolarisation ventriculaire, La repolarisation auriculaire ne peut généralement pas être discerné de l'ECG car il coïncide avec le QRS beaucoup plus grand complexe. L'amplitude d'une onde est mesurée par référence au niveau de base de l'ECG, généralement défini par la ligne isoélectrique qui immédiatement précède le complexe QRS. La durée d'une onde est définie par les deux instants temporels auxquels l'onde s'écarte sensiblement de la ligne de base ou le traverse.



Complexe QRS

Figure I.7 : Ondes du signal ECG

I.4.1 L'onde P

L'onde P reflète la dépolarisation séquentielle des oreillettes droite et gauche. Dans la plupart des dérivations, l'onde P a une polarité positive. Son amplitude est normalement inférieure à 300 μV , et sa durée est inférieure à 120 ms.

Il est parfois problématique de déterminer les instants qui définissent le début et la fin d'une onde P en raison de sa faible amplitude et de sa lisse morphologie. Par conséquent, l'analyse individuelle des ondes P est exclue de certaines applications ECG où la présence de bruit est considérable.

I.4.2 Le complexe QRS

Le complexe QRS reflète la dépolarisation des ventricules droit et gauche. Dans le cœur normal, la durée du complexe QRS est environ 70-110 ms. La première déviation négative du complexe QRS est appelée l'onde Q, et la première déviation positive est noté l'onde R, tandis que la déviation négative qui suit l'onde R est noté l'onde S.

La morphologie du complexe QRS est très variable et dépend de l'origine du rythme cardiaque.

Étant donné que le complexe QRS a la plus grande amplitude des formes d'onde ECG, atteignant parfois 2-3 mV, c'est la première onde identifiée dans tout système d'analyse.

En raison de ses pentes dures, le contenu en fréquence du complexe QRS est considérablement plus élevé que celui des autres ondes ECG et est principalement concentré dans l'intervalle 10-50 Hz

I.4.3 L'onde T

L'onde T reflète la repolarisation ventriculaire et s'étend sur environ 300 ms après le complexe QRS. La position de l'onde T est fortement dépendante de la fréquence cardiaque, devenant plus étroit et plus proche du complexe QRS dans le cas des taux cardiaques rapides. Cette propriété de "contraction" ne s'applique pas à l'onde P ou au complexe QRS. L'onde T normale a une morphologie lisse et arrondie et est associée à un seul pic positif.



I.5 Les Rythmes cardiaques

Le rythme cardiaque normal est contrôlé par les impulsions électriques formé dans le nœud SA et produit une fréquence cardiaque comprise entre 50 et 100 battements / minute pendant le repos. Une déviation ou une perturbation du rythme sinusal normale est appelé arythmie. Une arythmie peut survenir lorsque la dépolarisation est déclenchée par d'autres cellules du stimulateur cardiaque que celles du nœud SA, altérant ainsi la formation des impulsions électriques. Un autre mécanisme qui produit une arythmie est lorsque la conduction des impulsions électriques est altérée.

La classification de l'arythmie cardiaque se base sur l'emplacement de son origine.

Une arythmie est également classée en fonction de son taux : Bradyarythmie a une fréquence cardiaque lente de moins de 60 battements / minute, tandis que la tachyarythmie à un rythme rapide de plus de 100 battements / minute.

L'importance d'une arythmie diffère largement et peut être bénigne, symptomatique, dangereuse ou mortel. Sa conséquence ne dépend pas seulement de sa manifestation, mais aussi sur la présence de conditions structurelles anormales du cœur.

I.5.1 Le Rythme sinusal

Le rythme sinusal normal provient du nœud SA et a un taux compris entre 50 et 100 battements / minute au repos. Le rythme est appelé bradycardie sinusale lorsque le taux est inférieur à la limite inférieure et sinus tachycardie lorsqu'il est supérieure à la limite supérieure. Au repos, la fréquence cardiaque est essentiellement régulière mais pas totalement, même en absence des perturbations externes qui forment de stress physique ou mental.

La variabilité de la fréquence cardiaque a été de plus en plus étudiée récemment pour diagnostiquer et prédire les maladies cardiovasculaires.

I.5.2 Battements prématurés

Le rythme sinusal normal est parfois interrompu par un battement survenu avant le temps normal du prochain battement de sinus et est donc appelé battement prématuré ; les termes "battement ectopique" et "extrasystole" sont des synonymes fréquemment utilisés. Le foyer ectopique produisant le battement prématuré peut être situé dans n'importe quelle partie du cœur autre que le nœud SA.

Un battement prématuré est soit appelé battement prématuré supra ventriculaire (SVPB) si son origine est au-dessus des ventricules, c'est-à-dire dans les oreillettes ou le nœud AV, ou un battement ventriculaire prématuré (VPB) si son origine est dans les ventricules.

Le type SVPB le plus courant est associé à une morphologie anormale des ondes P et à une morphologie du complexe QRS ressemblant à celle d'un battement sinusal normal.

Un VPB peut provenir de n'importe quelle zone au-delà du point où le bundle s'est ramifié dans les branches gauche et droite. Du fait que l'impulsion électrique d'une source ectopique ventriculaire ne suit pas les voies de la conduction normale, le complexe QRS produit est anormalement prolongé et a une morphologie qui s'écarte considérablement de celle d'un battement de sinus, il est souvent beaucoup plus grand et bizarre.

I.5.3 Arythmie auriculaire

Diverses perturbations du rythme proviennent d'un ou de plusieurs foyers ectopiques dans les oreillettes. Les arythmies qui en résultent sont caractérisées soit par la présence d'ondes P anormales ou d'une absence totale d'ondes P distinguables. Des ondes P anormales apparaissent lorsqu'un foyer ectopique est situé loin du nœud SA, c'est-à-dire, plus proche du nœud AV, car l'impulsion électrique se propage alors dans une direction opposée à la normale ; une onde P normale devient ainsi négative. Si le foyer est près du nœud AV et des ventricules, la dépolarisation des ventricules commence à peu près en même temps que la dépolarisation des oreillettes. En conséquence, l'apparition de l'onde P coïncide avec celui du complexe QRS et ne peut pas être discerné dans l'ECG.

La tachycardie auriculaire est une arythmie produite par une automaticité accrue dans les cellules de stimulation d'un ou de plusieurs foyers des oreillettes. La plupart des impulsions électriques sont transmises aux ventricules, entraînant une fréquence cardiaque de 140 à 220 battements / minute. En raison de la vitesse rapide, les ondes P sont souvent masquées par l'onde T ou même le complexe QRS du cycle cardiaque précédent.

Le flutter auriculaire et la fibrillation auriculaire sont deux tachyarythmies dans lesquelles les oreillettes ne sont pas synchronisées avec les ventricules et battent à un rythme qui est

beaucoup plus rapide que celle des ventricules. Cela se manifeste dans l'ECG par une ligne de base ondulée qui remplace les ondes P, et, par conséquent, la ligne isoélectrique n'est plus bien définie.

I.5.4 Arythmies ventriculaires

Les tachyarythmies ventriculaires les plus courantes sont : la tachycardie ventriculaire, le flutter ventriculaire et la fibrillation ventriculaire. Les mécanismes derrière ces trois arythmies sont analogues aux trois tachyarythmies auriculaires susmentionnées.

I.5.5 Trouble de la conduction cardiaque

La propagation d'une impulsion électrique peut être perturbée par un bloc le long son sa voie de conduction normale. Le bloc provoque des dépolarisations et repolarisations anormales, perturbant la fonction du cœur.

L'un des blocs de conduction est lié au nœud AV et produit une connexion anormale entre les oreillettes et les ventricules. La gravité du bloc AV est classée comme mineure, lorsque toutes les impulsions sont effectuées avec retard, comme modérée, lorsque certaines impulsions n'atteignent pas les ventricules, et comme complète, quand aucune impulsion n'est conduite.

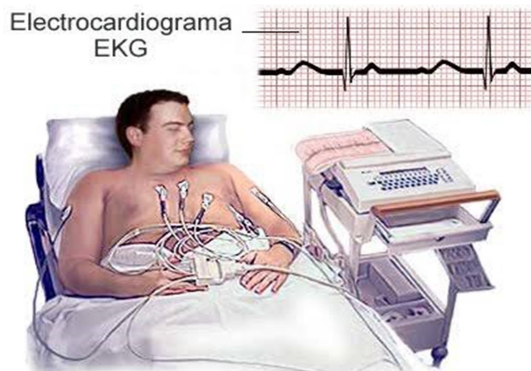
Le bloc AV Complet se manifeste par des ondes P et des complexes QRS à deux taux différents et indépendants ; les ondes P sont produites par le nœud SA, tandis que les complexes QRS ont leur origine dans un foyer ectopique ventriculaire. Beaucoup d'autres types des blocs de conduction peuvent se produire, par exemple, dans les branches gauche ou droite du faisceau [1].

Chapitre II : Aspect technique de l'électrocardiographie

II.1 Catégories des appareils d'enregistrement des ECG

Il existe différentes catégories d'appareils ECG. Leur domaine d'application est différent, ils doivent donc se conformer à de différentes exigences. Les principales catégories d'appareils sont [7] :

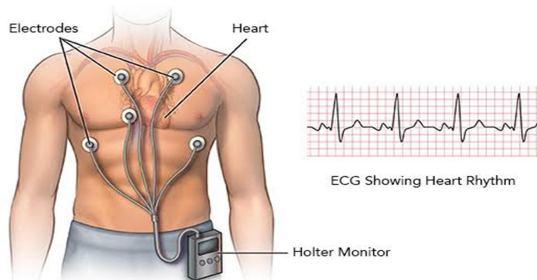
- ECG de diagnostic pour enregistrer les signaux au repos, (voir figure a)
- ECG d'effort (exercice), (voir figure b)
- ECG Holter (24 heures, ambulatoire), (voir figure d)
- Module ECG dans un moniteur patient de soins intensifs,
- Module ECG dans un défibrillateur automatique implantable,
- ECG de télémétrie,
- ECG fœtal. (Voir figure c)



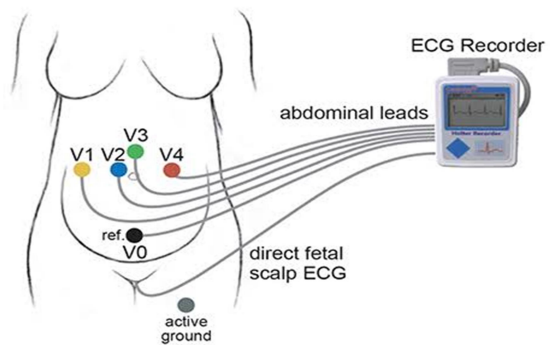
a : ECG de diagnostic au repos



b : ECG d'effort



d : ECG Holter



d : ECG fœtal

Figure II.1 : Exemple de quelques Catégories des appareils

II.2 Acquisition des signaux ECG

Pour la plupart des configurations de dérivations, la différence de potentiel générée par le cœur est généralement de l'ordre de 1mV et, pour l'affichage et l'interprétation ultérieurs du signal, une amplification est nécessaire. La connexion des deux fils à un amplificateur électronique moderne n'affectera généralement pas la différence de tension présente.

II.2.1 Schéma général

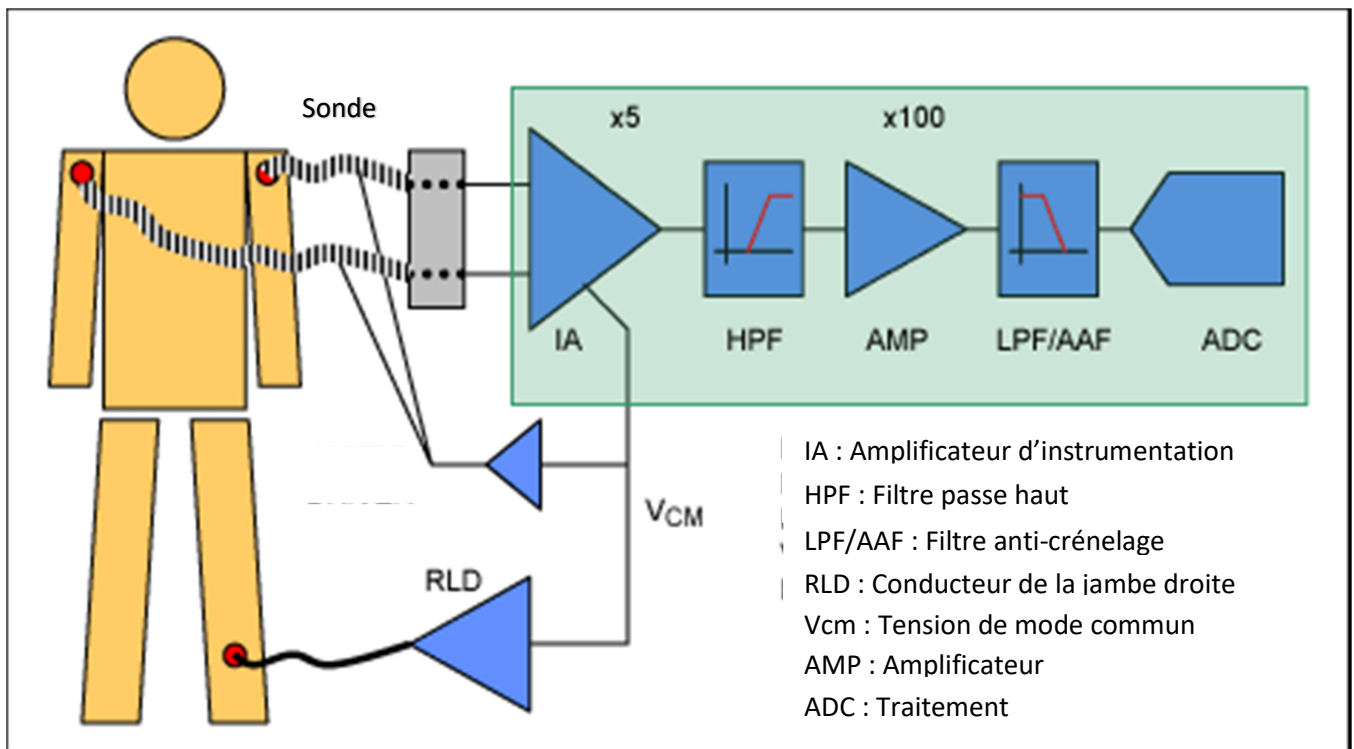


Figure II.2 : Chaîne du traitement de l'ECG

Les biopotentaux mesurés par les capteurs sont appliqués à l'avant d'un amplificateur d'instrumentation (IA). Pour des raisons de sécurité, la sortie de l'amplificateur est souvent isolé des autres circuits de traitement par un amplificateur d'isolement qui limite le flux de courant de l'équipement d'enregistrement vers le patient. Le signal ECG est limité en bande par une combinaison d'un filtre passe-bas (LP) et d'un filtre passe-haut (HP) qui minimise l'aliasing du signal et couper la tension continue, respectivement. Un filtre coupe-bande de 50 – 60Hz est fréquemment utilisé pour minimiser les interférences de la ligne électrique.

Outre l'amplification, les systèmes de mesure et d'enregistrement ECG exécutent certaines étapes de prétraitement du signal de base. Ils peuvent être exécutés dans des modules de traitement de signaux analogiques et / ou numériques.

La différence de potentiel entre les deux bornes d'entrée de l'amplificateur ne signale pas le champ potentiel généré par l'activité électrique du cœur. Il comprend également les effets de plusieurs autres sources, les plus importants étant les effets d'autres sources bioélectriques au sein du patient (activation musculaire) et les interférences des champs électromagnétiques entourant le patient, par exemple générées par la ligne électrique, les fréquences RF ou radio, les appareils d'électrochirurgie, etc. [8].

II.2.2 Amplificateur de mesure

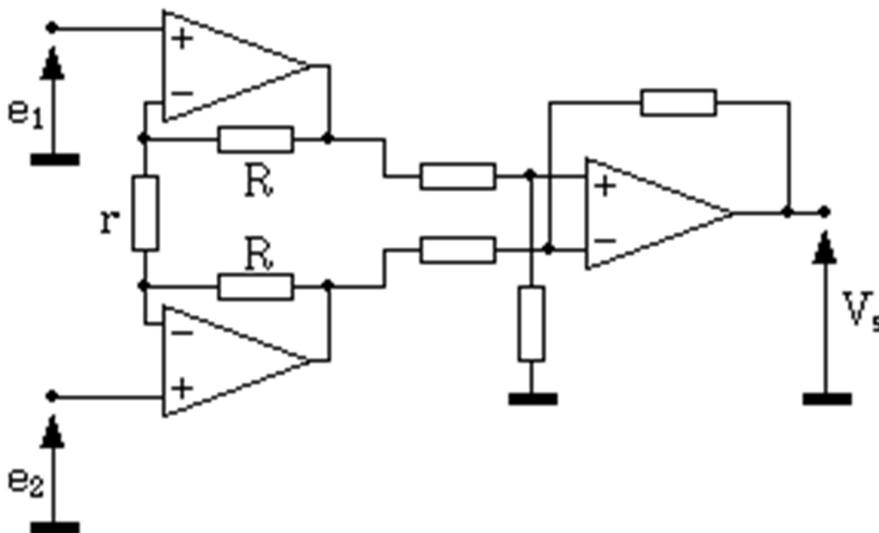


Figure II.3 : Amplificateur d'instrumentation ECG à deux étages

Le premier étage de l'amplificateur d'instrumentation (IA Fig. II.3), sert à élever les amplitudes du signal à un certain niveau souhaité et également à réduire la représentation de tout potentiel que les deux capteurs peuvent avoir en commun par rapport à la masse de l'alimentation électrique de l'appareil d'enregistrement.

Les amplificateurs opérationnels ont un facteur d'amplification très élevé ; Idéalement, le potentiel à la sortie du deuxième étage indiqué doit être proportionnel à $V_s = e_1 - e_2$ et être complètement indépendant de la valeur moyenne de e_1 et e_2 , les valeurs que les électrodes de détection e_1 et e_2 ont en commun. En utilisant résistances très précises avec des valeurs identiques, il est possible d'obtenir un rejet élevé du potentiel moyen commun.

II.2.3 Amplificateur d'isolation

L'amplificateur d'isolement indiqué sur la figure II.2 est inclus pour des raisons de sécurité. Il isole le patient des autres circuits de traitement en limitant le flux de courant entre l'appareil d'enregistrement et le patient. Il nécessite que sa partie d'entrée soit alimenté par une batterie isolé galvaniquement. Le couplage peut être réalisé en utilisant une transmission en fibre de verre ou une transmission RF.

II.2.4 Filtre passe haut

La perturbation la plus courante rencontrée dans le signal ECG est celle causée par des changements dans l'interface capteur-peau. Cela peut être dû à des artefacts de mouvement, des changements de température, des changements d'humidité de la peau (transpiration) et des fuites de gel ou de pâte. Ceci est particulièrement important dans les applications de test d'effort, de respiration et de surveillance, car les capteurs ont tendance à se détacher et le contact cutané sous les capteurs change lorsque les patients bougent. Les signaux ECG au repos montrent souvent les artefacts causés par la respiration ou par des tremblements musculaires à basse température. Une mauvaise préparation de la peau sur des sujets à poitrine velue peut créer du bruit en raison d'une mauvaise adhérence du capteur et d'une interface d'impédance peau-capteur élevée.

Même pendant l'état de repos, les changements au niveau des capteurs peuvent donner lieu à des fluctuations potentielles de l'ECG qui sont plus importantes que l'ECG lui-même. Heureusement, dans ces conditions, les changements ont un spectre de basses fréquences, avec des fréquences inférieures à celles présentes dans l'ECG. L'effet observé sur l'ECG est mouvement lent de la ligne de base. Ceci permet de réduire ce que l'on appelle le mouvement de la ligne de base au moyen d'un filtrage passe-haut dans le domaine fréquentiel du signal. La forme de mise en œuvre la plus simple du filtre passe-haut est celle d'inclure des condensateurs dans les fils reliant les capteurs à l'amplificateur. La fréquence de coupure utilisée dépend de l'application particulière impliquée.

II.2.5 Filtre passe bas

Un composant indésirable supplémentaire de l'ECG est celui généré par l'activité électrique des muscles squelettiques. Cette interférence gêne en particulier l'interprétation de l'ECG lors des tests d'effort, ou de surveillance à long terme (holter). Le spectre se situe au-dessus de ceux des signaux ECG et un filtre passe-bas peut être utilisé pour réduire leur effet.

II.3 Les électrodes

II.3.1 L'interface peau-électrode

Il existe une différence fondamentale dans la génération et la représentation des signaux électriques dans les tissus biologiques et les conducteurs métalliques. Dans les tissus biologiques, les champs électriques sont générés par des processus biochimiques dans lesquels les ions sont séparés, concentrés et déplacés en raison des forces thermodynamiques, des gradients de concentration ou des gradients de potentiel électrique imposés. D'un point de vue électrique, le tissu biologique se comporte comme un électrolyte. Dans les conducteurs métalliques, le courant électrique est représenté par des impulsions d'électrons et des mouvements d'électrons.

II.3.2 Types d'électrodes

Différents types d'électrodes pour l'enregistrement de signaux bioélectriques ont été développés en fonction de leur application spécifique. Nous discutons seulement les électrodes utilisées pour l'électrocardiographie clinique. Le besoin d'électrodes maniables et faciles à appliquer avec une faible tension de consigne et une faible impédance, une faible collecte des artefacts, une stabilité élevée des propriétés électriques et une irritation cutanée minimale a conduit à la conception d'un certain nombre de types d'électrodes différents avec différents modes de fonctionnement.

En raison des coûts, presque toutes les électrodes pour les enregistrements de routine sont passives. Pour des applications spécifiques, par exemple l'analyse de la surface corporelle, des amplificateurs plus puissants ont été intégrés dans le boîtier d'électrode. Ces électrodes sont alors appelées électrodes actives. Malgré les progrès considérables dans la compréhension des phénomènes de l'interface peau-électrode et les progrès de la technologie des électrodes, cette interface reste le maillon le plus faible dans la chaîne de mesure. Pour une analyse quantitative informatisée de l'ECG, il faut faire très attention à cette étape pour obtenir des signaux ECG reproductibles de haute-fidélité.

a) Electrodes à plaques

La figure 2.1 montre des électrodes à plaque typiques appliquées à l'électrocardiographie de surface corporelle. Le type illustré en 2.1-a est utilisé pour les dérivations des membres, tandis que pour les dérivations thoraciques (1-2 cm), le type indiqué en 2.1-b est utilisé. Les grandes électrodes à plaques, qui ont été introduites en 1917, sont en argent germanique (alliage de

nickel, cuivre et zinc), en acier inoxydable, en nickel ou en acier nickelé. L'électrode métallique de ce type d'électrode doit être séparée de la peau par un film de pâte électrolytique ou par un papier humide afin d'obtenir une tension constante stable et une faible impédance de peau d'électrode. Les plus petites électrodes thoraciques sont en nickel, un alliage d'argent parfois recouvert de chlorure d'argent ou d'un matériau fritté contenant $\text{Ag} | \text{AgCl}$. Les électrodes en plaque sont généralement fixées par des sangles en caoutchouc. Cette méthode convient bien aux dérivations des membres, mais pas aux dérivations thoraciques.



Figure II.4 : Deux exemples d'électrodes en plaque utilisées en électrocardiographie

1. une électrode de dérivation membre qui doit être appliquée en utilisant un papier humide entre la peau et l'électrode;
2. une électrode thoracique typique qui est appliquée avec un gel d'électrode

b) Electrodes d'aspiration

Il y a deux types d'électrodes d'aspiration. Le type (a), électrode à ventouse, (figure II.4) peut être localisée avec précision et appliquée rapidement. Il est également bien adapté à la fixation sur des surfaces et des tissus plats et mous. Il est maintenu par un vide produit après avoir appuyé et relâché l'ampoule en caoutchouc, tandis que le type (b) est maintenu par un vide produit par un courant d'air traversant l'électrode. Le matériau d'électrode de type a est du nickel ou un alliage. Alors que l'électrode de type b peut être en nickel, en alliage ou en $\text{Ag} | \text{AgCl}$. En raison de la petite surface de contact, l'impédance de l'électrode de type a est élevée et, en raison du contact étroit entre l'électrode et la peau, l'impédance et la tension de consigne sont sensibles au mouvement. Etant donné que la zone de contact est plus grande dans l'électrode de type (b) que celle du type (a), l'impédance cutanée est plus petite. La fixation des électrodes de type (b) est beaucoup plus robuste que pour les électrodes à ventouse et est fiable même sur peau sèche ou velue. Un inconvénient des électrodes d'aspiration est qu'ils ne peuvent être laissés sur le sujet que pendant une durée limitée afin d'éviter de graves irritations cutanées.

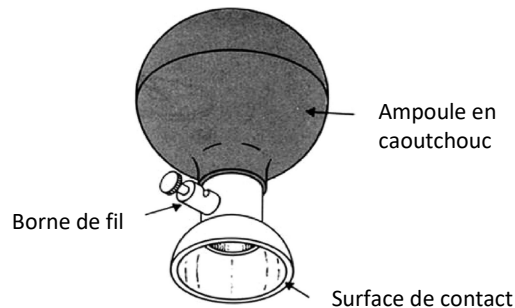


Figure II.5: Electrode d'aspiration

c) Electrode à colonne de fluide

Dans les électrodes à colonne fluide, l'électrode consiste en un disque métallique et un élément $\text{Ag} | \text{AgCl}$ qui sont encastrés de sorte qu'il n'entre pas en contact avec la peau elle-même. La cavité pourrait être remplie de pâte qui fournit un pont électrolytique entre l'électrode et la peau en maintenant la distance entre eux.

En cas de léger mouvement de l'électrode par rapport à la surface du corps, l'interface électrode-électrolyte n'est pas modifiée de manière significative et les artefacts de mouvement sont donc minimisés. L'électrode peut être fixée sur la peau par un anneau de ruban adhésif double face ou par le vide.

d) Electrodes secs

L'utilisation de pâte d'électrode en électrocardiographie clinique est une procédure lente. La préparation de la peau et l'application de pâte sur chaque patient pour chaque électrode prennent du temps, et les électrodes à usages multiples doivent être nettoyées régulièrement pour maintenir un faible bruit et une faible impédance peau-électrode. Dans les applications à long terme, la pâte a tendance à se dessécher ou peut irriter la peau. Des efforts ont été faits pour développer des électrodes sans pâte qui pourraient être directement mises sur la peau. Il existe deux types d'électrodes sèches: les électrodes à plaque métallique qui captent l'ECG de manière conductrice; et des électrodes isolées, où la surface de l'électrode métallique est recouverte d'un diélectrique et le corps agit comme l'autre plaque d'un condensateur, de sorte que le signal est capté capacitivement [8]. Le problème avec les deux types d'électrodes est l'impédance d'entrée élevée et son instabilité. Les électrodes en métal sec, ainsi que les

électrodes isolées, nécessitent des amplificateurs transformateurs d'impédance directement fixés à l'électrode.

e) Electrodes actives

En raison de l'impédance élevée et instable des électrodes ECG conventionnelles et du problème existant de blindage des câbles entre l'électrode et le préamplificateur, un amplificateur tampon doit être fixé directement à l'électrode. La microélectronique permet désormais l'intégration d'amplificateurs dans le boîtier d'électrode. Ces électrodes sont alors appelées électrodes actives.

II.4 Artéfacts et bruits dans l'ECG

Le signal ECG comprend de nombreux artéfacts. L'origine et la nature de ces artéfacts sont d'un intérêt considérable, en particulier pour les applications de surveillance à long terme. Certains des artéfacts sont dus à des raisons physiologiques comme le bruit de l'électromyographie (EMG ou activité musculaire) et mouvement de la ligne de base due à la respiration. Certains artéfacts sont dus à des raisons non physiologiques, par exemple, des interférences de 50 ou 60 Hz sur les lignes électriques et des artéfacts de mouvement dans l'ECG. La présence des artéfacts rendra problématique tout diagnostic basé sur la morphologie.

II.4.1 Bruit dû au secteur

Les réseaux de distribution électrique peuvent brouiller le signal électrocardiographique avec une onde dont l'harmonique principale est à 50 Hz. Ce type de bruit se produit sur tout l'enregistrement et peut être assez fort mais il s'élimine facilement avec un filtre sélectif car c'est un bruit haute fréquence a bande étroite.

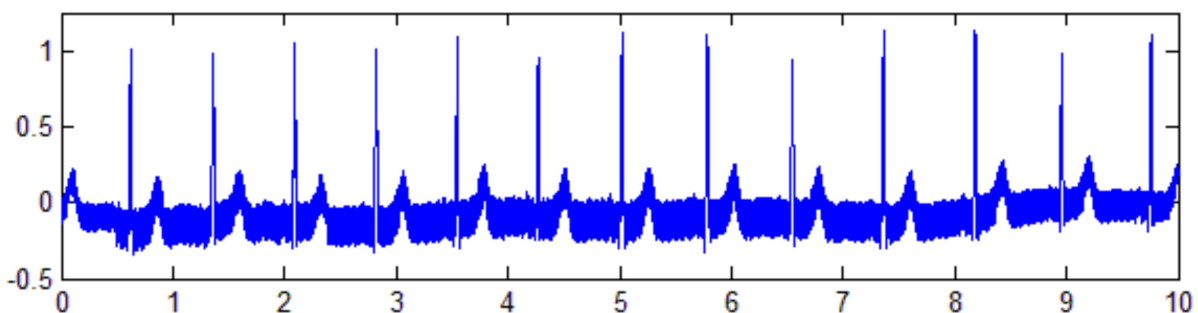


Figure II.6: Signal électrocardiographique perturbé par le secteur

Lorsque les électrodes sont connectées incorrectement, des sauts brusques de la ligne de base apparaissent. L'effet sur le trace peut aller de la simple diminution d'amplitude à l'apparition de pics lorsque les électrodes sont en contact intermittent avec la peau. Ces pics peuvent parfois être confondus avec les ondes du tracé normal. Ce type de bruit intermittent à bande spectrale large s'élimine difficilement car son énergie se trouve dans la même gamme de fréquence que le complexe QRS.

II.4.2 Mouvements de la ligne de base

Lors de l'enregistrement de l'électrocardiogramme, l'activité respiratoire peut faire osciller la ligne de base de l'ECG. D'autres perturbations peuvent avoir pour effet de déplacer temporairement la ligne de base comme, par exemple, les mauvais contacts entre la peau et les électrodes. Dans une surveillance de courte durée dans les cliniques, le mouvement du patient est limité et, par conséquent, l'artefact de mouvement n'est pas fréquent. Cependant, les mouvements de ligne de base sont assez courants dans la surveillance ECG à long terme.

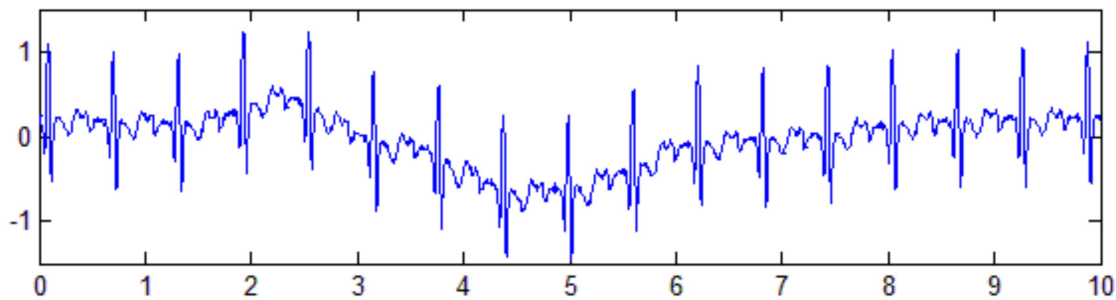


Figure II.7 : Mouvement de la ligne de base

Ces perturbations sont généralement peu gênantes pour l'analyse de l'ECG et peuvent être en grande partie filtrées car leur énergie se situe dans une bande de fréquence basse, qui empiète peu sur celle de l'ECG normal.

II.4.3 Bruit musculaire

Toute activité musculaire dans le corps produit un signal bio-potentiel qui est également connu sous le nom de signal d'électromyographie (EMG). L'amplitude de pic du signal EMG à la surface du corps est comprise entre 0,1 et 1 mV et le spectre est concentré sur la plage de fréquences de 5 à 500 Hz. Puisque les signaux EMG et ECG ont des spectres qui se chevauchent partiellement, l'activité musculaire peut provoquer des interférences dans le signal ECG. Ce type de bruit est appelé bruit EMG. Au cours des tests cliniques, le patient est généralement dans un état de repos et, par conséquent, les risques de bruit EMG sont rares.

Pour les surveillances de courte durée, seuls quelques cycles de signaux ECG sont nécessaires. Cependant, pour la surveillance à long terme dans des conditions ambulatoires, la présence de bruit EMG à haute fréquence est problématique.[9]

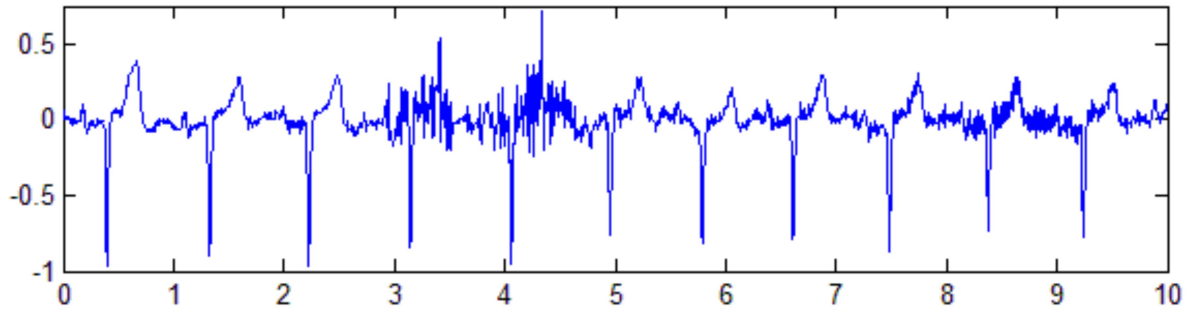


Figure II.8 : Bruit musculaire

Chapitre III :

Classification des signaux ECG avec les arbres de décision

III.1 Introduction au processus général de la reconnaissance des formes

La reconnaissance des formes (RDF) consiste essentiellement en l'extraction de l'information utile à partir de données brutes. La RDF s'intéresse donc au développement des systèmes intelligents qui remplissent des tâches de perception et de prise de décision. La RDF englobe une grande gamme de problèmes de traitement d'information ayant de grandes importances pratiques. Les formes à identifier dépendent de l'application : ils peuvent être des images, des lettres, des paroles, des cibles militaires, des formes d'onde d'un signal... etc [10].

Un système de RDF de base se compose principalement d'un capteur, d'un dispositif d'extraction des caractéristiques et d'un classificateur (Fig. III.1). Un système plus complet inclura également des étapes de prétraitement, de sélection des caractéristiques d'évaluation et de post traitement.

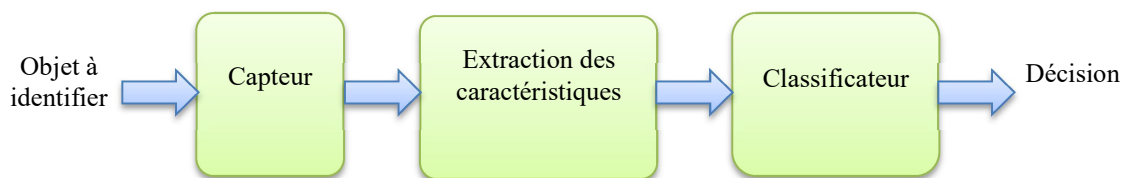


Figure III.1 : Schéma de base d'un processus de RDF

III.2 Processus la classification des ECG

Un système entièrement automatique de classification des arythmies à partir de signaux acquis par un appareil ECG peut être divisé en quatre étapes (voir Fig. III.11), comme suit : (1) prétraitement du signal ECG; (2) la segmentation du rythme cardiaque; (3) extraction de caractéristiques; et (4) apprentissage / classification [11].

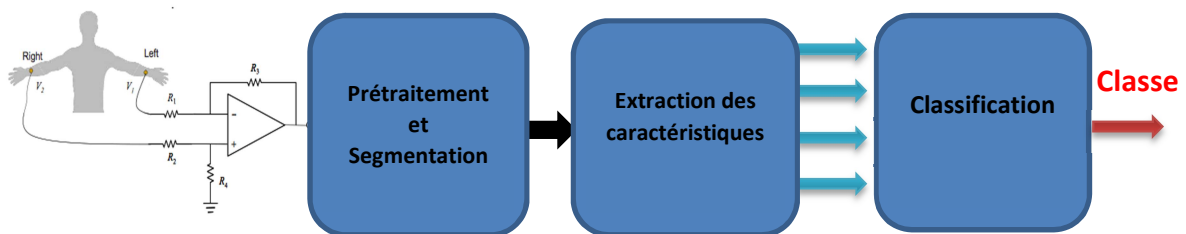


Figure III.1: Processus général de classification des signaux ECG

III.2.1 Prétraitement du signal ECG

Afin de réduire les bruits dans les signaux ECG, plusieurs méthodes ont été proposées. Parmi ces propositions, la plus simple et la plus largement utilisée est la mise en œuvre de filtres numériques récurrents de la réponse impulsionnelle finie (FIR), qui a été rendue possible avec les progrès technologiques des microcontrôleurs et microprocesseurs. Ces méthodes fonctionnent bien pour l'atténuation des bandes de fréquences connues, telles que le bruit provenant du réseau électrique (50 Hz ou 60 Hz), car ils permettent une application rapide et facile du filtre à bande [11].

Au cours de la dernière décennie, de nombreuses méthodes basées sur les transformations en ondelettes ont été employées pour éliminer le bruit. Ces méthodes permettent de préserver les propriétés du signal ECG et en pas perdre ses importants détails physiologiques ; de plus ces méthodes sont simples d'un point de vue informatique.

III.2.2 Détection des Complexe QRS et segmentation

L'étape suivante du traitement du signal ECG consiste à détecter tous les battements cardiaques de l'enregistrement ECG. En raison du fait que le complexe QRS est généralement plus large que les ondes P et T, l'identification des battements cardiaques est généralement réalisée avec un détecteur QRS. De plus, le complexe QRS a une composante de fréquence élevée pouvant être facilement distingués des basses fréquences des ondes P et T.

Le détecteur QRS doit être capable de détecter un grand nombre de morphologies QRS différentes afin être cliniquement utile et capable de suivre des changements soudains ou progressifs du rythme cardiaque. Par conséquent, les performances de tout système d'analyse automatique ECG dépendent d'une bonne détection de tous les complexes QRS dans l'enregistrement ECG.

Dans la littérature, plusieurs méthodes de détection de QRS ont été proposées. Parmi ces méthodes, l'approche la plus utilisée, a été proposé par Pan & Tompkins [13]. Sa mise en œuvre est simple et pouvant atteindre de bon performances en matière de sensibilité et de prédictivité (tous deux > 99,5%). Cette méthode comporte les étapes suivantes:

- Un filtre passe-bande qui maintient la partie spectrale où se concentre l'énergie QRS. Ce filtre permet de : (i) atténue les composantes basses fréquences des ondes P et T ;(ii) supprime les changements de la ligne de base ; (iii) réduit les interférences de ligne 50/60 Hz et le bruit EMG haute fréquence.

- Un différenciateur sélectionne les pics du QRS, qui sont très différentes des autres composants plus lisses.
- La valeur quadratique moyenne de chaque échantillon de signal est calculée pour n'obtenir que des valeurs positives et pour souligner le QRS avec le composant élevé de fréquence.
- Un intégrateur de fenêtres mobiles additionne les zones sous le quadratique signal pour produire des impulsions et éliminer les artefacts de courte durée. Une telle sortie va à l'unité de décision où chaque impulsion est comparée au seuil préétabli en la distinguant ou non et le situant à sa place temporelle relative appropriée.

Des méthodes plus sophistiquées ont également été utilisées, telles que les méthodes basées sur les réseaux de neurones, les algorithmes génétiques, la transformée en ondelettes, les bancs de filtres.

III.2.2 Extraction des caractéristiques

L'étape d'extraction des caractéristiques est la clé du succès dans la classification du rythme cardiaque de l'arythmie à l'aide du signal ECG. Toute information extraite du signal utilisée pour discriminer son type peut être considérée comme une caractéristique. Les caractéristiques peuvent être extraites sous diverses formes directement à partir de la morphologie du signal ECG dans le domaine temporel et /ou dans le domaine de fréquence ou à partir du rythme cardiaque.

Les caractéristiques les plus courantes dans la littérature sont calculées à partir du rythme cardiaque (ou intervalle des battements cardiaques). L'intervalle RR est le temps entre le pic R d'un battement de cœur par rapport à un autre battement de cœur, qui pourrait être son prédécesseur ou son successeur.

Les caractéristiques de l'intervalle RR ont une grande capacité à distinguer les types de battements cardiaques et certains auteurs ont basé leurs méthodes uniquement sur l'utilisation des caractéristiques d'intervalle RR.

D'autres caractéristiques extraites des intervalles de battements cardiaques sont les distances entre les points de référence d'un battement cardiaque (appelés intervalles ou segments ECG). Parmi ces intervalles, l'intervalle QRS, ou la durée du complexe QRS, est le plus utilisé. Certains types d'arythmies provoquent des variations de l'intervalle QRS, ce qui en fait une bonne caractéristique de discrimination.

Tableau III.1 : illustre les valeurs standards pour ces intervalles, en considérant un être humain en bonne santé sans anomalies cardiaques.

Caractéristique	Valeur normale	Variation normale
Onde P	110 ms	±20 ms
Intervalle PQ / PR	160 ms	±40 ms
Largeur QRS	100 ms	±20 ms
Intervalle QT	400 ms	±40 ms
Amplitude de P	0,115 mV	±0,05 mV
Amplitude du QRS	1,5 mV	± 0,5 mV
Niveau ST	0 mV	±0,1 mV
Amplitude de T	0.3 mV	± 0.2 mV

Tableau III.1 : Les valeurs standards pour les intervalles d'un ECG normal.

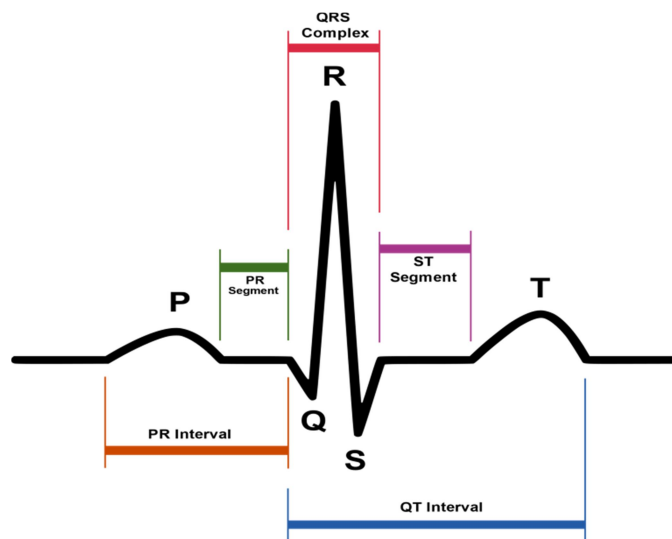


Figure III. 3 : Point de référence et intervalles des battements cardiaques

La manière la plus simple d'extraire des caractéristiques dans le domaine temporel consiste à utiliser des points de la courbe ECG segmentée (le battement cardiaque) comme caractéristiques. Cependant, l'utilisation d'échantillons de la courbe en tant que caractéristiques est une technique peu efficace, car en plus de produire un vecteur des caractéristiques de grande dimension, il souffre de plusieurs problèmes liés à l'échelle ou au déplacement du signal par rapport au point central (pic R). Afin but de réduire la dimension du vecteur de caractéristiques, diverses techniques ont été appliquées sur les échantillons qui

représentent le rythme cardiaque, comme analyse en composantes principales (ACP), ou analyse en composantes indépendantes (ICA), dans lesquels de nouveaux coefficients sont extraits pour représenter le rythme cardiaque. Des techniques plus simples, telles que l'interpolation, ont également été utilisées pour réduire le nombre de points représentant le rythme cardiaque.

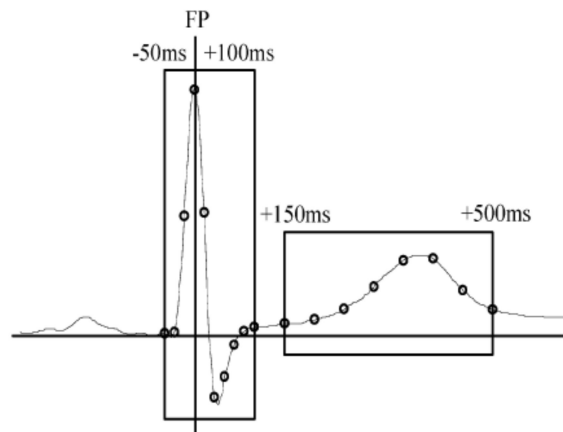


Figure III. 4: Exemple d'échantillons, de la courbe de l'ECG, utilisées en tant que caractéristiques

Bien que diverses techniques aient été envisagées, la plupart des études présentées dans la littérature utilisent des transformées d'ondelettes. Les chercheurs affirment qu'il s'agit de l'une des meilleures méthodes pour extraire les caractéristiques du signal ECG. La transformation en ondelettes permet l'extraction d'informations à la fois dans les domaines fréquentiel et temporel, ce qui diffère de ce qui est habituellement réalisé par la transformée de Fourier traditionnelle qui permet l'analyse du domaine fréquentiel seulement.

III.2.3 Classification

Une fois que l'ensemble de caractéristiques a été défini à partir des battements cardiaques, des modèles de classification peuvent être construits à partir de ces données en utilisant des algorithmes d'intelligence artificielle et d'apprentissage automatique.

Les méthodes les plus utilisées pour cette tâche sont [11] :

- Les machines à vecteurs de support (SVM, Support vector machines)
- Les réseaux de neurones artificiels (ANN, Artificial Neural Networks)
- La discrimination linéaire (LD, Linear discrimination)
- Les arbres de décision (DT, Decision trees)
- Les voisins les plus proches (NN, Nearest Neighbors)

III.3 Classification par l'arbre de décision

III.3.1 Principe

Dans l'exploration de données, un arbre de décision est un modèle prédictif qui peut être utilisé pour représenter à la fois des classificateurs et des modèles de régression. Les arbres de décision sont utilisés pour classer un objet ou une instance en fonction des valeurs de leurs attributs. Les arbres de décision sont fréquemment utilisés dans des domaines appliqués tels que le financement, le marketing, l'ingénierie et la médecine.

L'utilisation d'un arbre de décision est une technique très populaire dans l'exploration de données. De l'avis de nombreux chercheurs, les arbres de décision sont populaires en raison de leur simplicité et de leur transparence. Les arbres de décision sont explicites; il n'est pas nécessaire d'être un expert en exploration de données pour suivre un certain arbre de décision. Les arbres de décision sont généralement représentés graphiquement sous forme de structures hiérarchiques, ce qui les rend plus faciles à interpréter que les autres techniques. En effet, Cet outil offre la possibilité de visualiser graphiquement les différentes solutions possibles : Chaque extrémité de l'arbre, les branches, représentent un résultat potentiel.

III.3.2 Critères de division

Dans la plupart des arbres de décision, les fonctions de séparation sont univariées, c'est-à-dire qu'un nœud interne est divisé en fonction de la valeur d'un seul attribut. Par conséquent, l'inducteur recherche le meilleur attribut sur lequel effectuer la division. Les sections suivantes décrivent les deux critères les plus courants utilisés dans la littérature.

a) Indice de Gini

L'indice de Gini est un critère basé sur les impuretés qui mesure les divergences entre les distributions de probabilité des valeurs d'attributs cibles. L'indice Gini a été utilisé dans divers travaux.

Soit un problème de classification de K classes : c_1, c_2, \dots, c_K . Soit S l'ensemble de tous les exemples d'apprentissage. L'indice de Gini pour un ensemble S est défini par:

$$Gini(S) = 1 - \sum_{S_k \in S} (P(S_k))^2 \quad (III.1)$$

Où $P(S_k)$ est le pourcentage du nombre d'exemple ($|N_k|$) d'une certaine classe C_k sur le nombre de tous les exemples d'apprentissage ($|N|$) de l'ensemble S :

$$P(S_k) = \frac{|N_k|}{|N|} \quad (\text{III.2})$$

Etant donnée un vecteur de tous les caractéristiques X , en utilisant les valeurs d'une caractéristique x_j , on peut diviser l'ensemble de donnée S en plusieurs sous ensembles groupés dans un ensemble S_j . Le gain d'information est mesuré en se basant sur la différence entre l'indice de Gini originale de S et celle après sa division en se basant sur une caractéristique x_j . Le gain d'information est ainsi donné par :

$$IG_gini(S, x_j) = Gini(S) - \sum_{S_{jk} \in S_j} P(S_{jk}) Gini(S_{jk}) \quad (\text{III.3})$$

b) Entropie :

L'entropie, originaire de la théorie d'information, est une mesure d'impureté largement utilisée. L'entropie pour un ensemble S est défini par:

$$Entropie(S) = \sum_{S_k \in S} -P(S_k) \cdot \log_2(P(S_k)) \quad (\text{III.4})$$

Le gain d'information, correspondant à la caractéristique x_j , est cette fois-ci donné par :

$$IG_entropie(S, x_j) = Entropie(S) - \sum_{S_{jk} \in S_j} P(S_{jk}) Entropie(S_{jk}) \quad (\text{III.5})$$

III.3.3 Critères d'arrêt

La phase de croissance se poursuit jusqu'à ce qu'un critère d'arrêt soit déclenché. Les conditions suivantes sont des règles d'arrêt courantes [14]:

- Toutes les instances de l'ensemble d'apprentissage appartiennent à une seule classe.
- La profondeur maximale de l'arbre a été atteinte.
- Le nombre de cas dans le nœud terminal est inférieur au nombre minimal de cas pour les nœuds parents.
- Le nombre de cas dans un ou plusieurs nœuds enfants serait inférieur au nombre minimal de cas pour les nœuds enfants.
- Le meilleur critère de division n'est pas supérieur à un certain seuil.

III.3.4 Algorithmes d'arbre de décision

a) ID3

L'algorithme ID3 est considéré comme un algorithme d'arbre de décision très simple [14]. En utilisant le gain d'informations comme critère de division, l'ID3 cesse de croître lorsque toutes les instances appartiennent à une seule valeur de sortie désirée ou lorsque le meilleur gain d'informations n'est pas supérieur à zéro. ID3 n'applique aucune procédure d'élagage ni ne gère les attributs numériques ou les valeurs manquantes.

b) IC4.5

C4.5, une évolution de ID3, présentée par le même auteur [16], utilise le rapport de gain comme critère de division. Le fractionnement cesse lorsque le nombre d'instances à diviser est inférieur à un certain seuil. L'élagage basé sur les erreurs est effectué après la phase de croissance. C4.5 peut gérer les attributs numériques. Il peut également procéder sur un ensemble d'apprentissage qui incorpore des valeurs manquantes, et ce en utilisant des critères de rapport de gain corrigés.

c) CART

CART signifie arbres de classification et de régression (Classification And Regression Trees). Il a été développé par [17] et se caractérise par le fait qu'il construit des arbres binaires, à savoir que chaque nœud interne a exactement deux branches sortantes. Les divisions sont sélectionnées en utilisant les critères de Gini ou bien de « Twoing ». Une caractéristique importante de CART est sa capacité à générer des arbres de régression. Dans les arbres de

régression, les feuilles prédisent un nombre réel et non une classe. En cas de régression, CART recherche des divisions qui minimisent l'erreur quadratique de prédiction (la déviation la moins carrée). La prédiction dans chaque feuille est basée sur la moyenne pondérée pour le nœud.

III.3.5 Exemple d'arbre de décision

Pour expliquer le principe des arbres de décision soit l'exemple suivant : l'objectif est d'estimer une décision (jouer ou non) en se basant sur 4 caractéristiques : temps, température, humidité et vent. Les données d'apprentissage sont données dans le tableau (III.2).

Tableau (III.2) : Données d'apprentissage de l'exemple du joueur

Temps	Température	Humidité	Vent	Jouer
ensoleillé	chaude	haute	non	<i>non</i>
ensoleillé	chaude	haute	oui	<i>non</i>
nuageux	chaude	haute	non	<i>oui</i>
pluvieux	douce	haute	non	<i>oui</i>
pluvieux	fraîche	normale	non	<i>oui</i>
pluvieux	fraîche	normale	oui	<i>non</i>
nuageux	fraîche	normale	oui	<i>oui</i>
ensoleillé	douce	haute	non	<i>non</i>
ensoleillé	fraîche	normale	non	<i>oui</i>
pluvieux	douce	normale	non	<i>oui</i>
ensoleillé	douce	normale	oui	<i>oui</i>
nuageux	douce	haute	oui	<i>oui</i>
nuageux	chaude	normale	non	<i>oui</i>
pluvieux	douce	haute	oui	<i>non</i>

L'arbre de décision correspondante à cet exemple est donné dans la figure (III. 5). La première caractéristique vérifiée est « temps » qui divise les données sur 3 sous-ensembles. Comme l'entropie du temps pour la valeur « nuageux » est 0, cet ensemble contient des exemples de la même classe. Donc, cet ensemble forme une feuille. Puis les deux autres caractéristiques sont testées de la même façon.

Les règles obtenues de cet arbre sont :

- Si « temps » est « nuageux » alors « non »
 Si « temps » est « humidité » et « Température » est « haute » alors « non »
 et « Température » est « normale » alors « oui »
 Si « temps » est « vent » et « Température » est « non » alors « oui »
 et « Température » est « oui » alors « non »

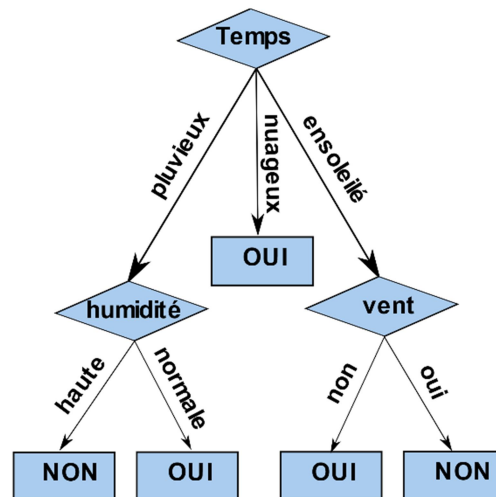


Figure III. 5 : L'arbre de décision correspondante à l'exemple du joueur

III.3.6 Avantages et inconvénients des arbres de décision

Plusieurs avantages de l'arbre de décision en tant qu'outil de classification apparaissent dans la littérature [14]:

- Les arbres de décision sont explicites et lorsqu'ils sont compactés, ils sont également faciles à suivre. C'est-à-dire que si l'arbre de décision a un nombre raisonnable de feuilles, il peut être saisi par des utilisateurs non professionnels. De plus, comme les arbres de décision peuvent être convertis en un ensemble de règles, ce type de représentation est considéré comme compréhensible. Les arbres de décision peuvent gérer à la fois les attributs d'entrée nominaux et numériques.
- La représentation de l'arbre de décision est suffisamment riche pour représenter n'importe quel classificateur à valeurs discrètes.

Parmi les inconvénients des arbres de décision figurent :

- La plupart des algorithmes (comme ID3 et C4.5) exigent que l'attribut cible n'ait que des valeurs discrètes.

- Comme les arbres de décision utilisent la méthode «diviser pour régner», ils ont tendance à bien fonctionner s'il existe quelques attributs très pertinents, mais moins si de nombreuses interactions complexes sont présentes.
- La sensibilité excessive à l'ensemble d'apprentissage, aux attributs non pertinents et au bruit rend les arbres de décision particulièrement instables ; un changement mineur dans une division près de la racine changera le sous-arbre entier ci-dessous.
- Le problème de fragmentation provoque le partitionnement des données en fragments plus petits. Cela se produit généralement si de nombreuses caractéristiques sont testées.

Chapitre IV : Tests et résultats

IV.1 Introduction

Dans ce chapitre nous présentons le système proposé de reconnaissance automatique des battements ECG et les résultats obtenus sur la base de données MIT-BIH. Pour les caractéristiques, nous utilisons les ondes et les intervalles temporels. Pour l'étape de classification nous utilisons les arbres de décision. En effet, les arbres de décision présentent les avantages suivants : ils sont explicites et lorsqu'ils sont compactés, ils sont faciles à suivre. C'est-à-dire que si l'arbre de décision a un nombre raisonnable de feuilles, il peut être saisi par des utilisateurs non professionnels. De plus, comme les arbres de décision peuvent être convertis en un ensemble de règles, ce type de représentation est considéré comme compréhensible. Les arbres de décision peuvent gérer à la fois des caractéristiques d'entrée nominaux et numériques.

IV.2 Processus

Le système proposé se base essentiellement sur l'extraction des caractéristiques temporels et la classification en utilisant les arbres de décision. Le figure IV.1 schématise se système. D'abord, le signal ECG est enregistré. Puis, les différentes ondes sont déterminées, en détectant les pics R et en localisant le début, le pic et la fin des ondes. Ensuite, les caractéristiques sont calculées en se basant sur les points fiduciels de l'étape précédente. Finalement, ce signal ECG sera classé, en utilisant les arbres de décision, à l'une de quatre classes.

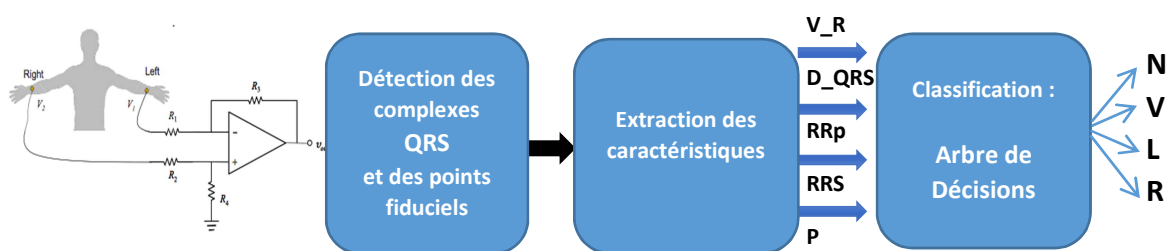


Figure IV.1 : Processus proposé de classification des battements ECG

Nous utilisons cinq caractéristiques parmi les plus utilisées pour la classification des signaux ECG (figure IV.2). Il s'agit de : Amplitude du pic R (V_R) , Intervalle RR précédant (RRp) Intervalle RR suivant (RRs), Durée du complexe QRS (D_{QRS}) et la présence de l'onde P (Onde).

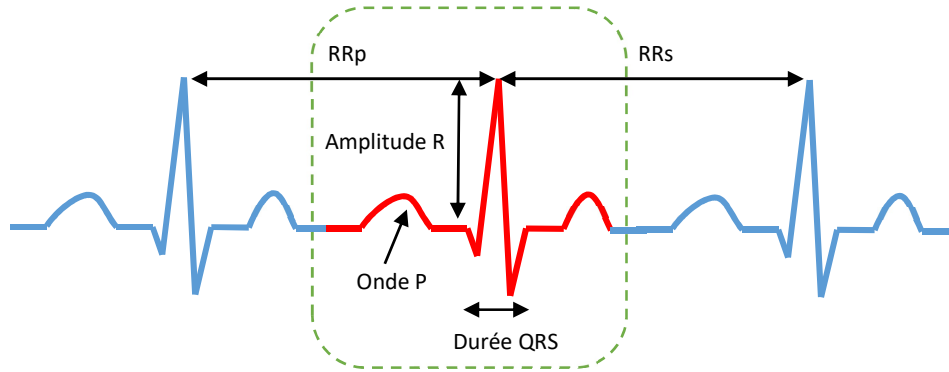


Figure IV.2 : Les 5 caractéristiques utilisées : Amplitude du pic R (V_R) , Intervalle RR précédant (RRp) Intervalle RR suivant (RRs), Durée du complexe QRS (D_{QRS}) et la présence de l'onde P (Onde_P).

Pour l'extraction des caractéristiques nous utilisons le la boîte à outils (WFDB toolbox) [19] pour MATLAB et Octave qui consiste ne un ensemble de fonctions de lecture, d'écriture et de traitement de signaux physiologiques et de séries chronologiques dans les formats utilisés par les bases de données Physio Bank. Nous utilisons les fonctions :

- Rdsamp : permet de lire le signal ECG
Donne la fréquence, l'enregistrement et le temps
- Rdann : lire les annotations (N, R, L, V,)
- Ecgpuwave Permet de détecter le début et la fin des ondes dans le signal ECG

La figure (IV.3) illustre un exemple de détection des ondes R et P et de début et fin du complexe QRS

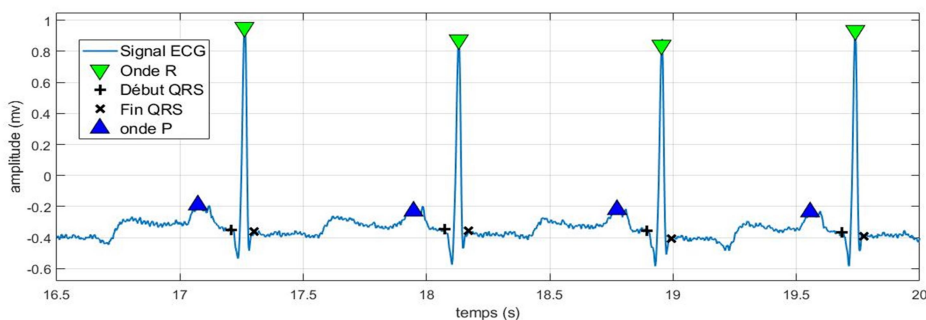


Figure IV.3 : Détection des Ondes R et P et de début et fin du complexe QRS

IV.3 Exemple de classification avec les arbres de décision

Pour expliquer le fonctionnement des arbres de décision, nous considérons un exemple de synthétique classification 2D. C'est un problème à quatre classes (C1,C2,C3,C4) dont chacune est représentée par cinq exemples (Figure IV.4).

Nous effectuons l'apprentissage d'un arbre de décision avec tous les exemples de ce problème. L'arbre ainsi obtenu est illustré dans la figure (IV.5). Cet arbre se compose de 7 nœuds : 3 branches et 4 feuilles. Cet arbre permet de classer tous les exemples de ce problème de classification. Les frontières de décision obtenue sont représentées sur la figure (IV.6).

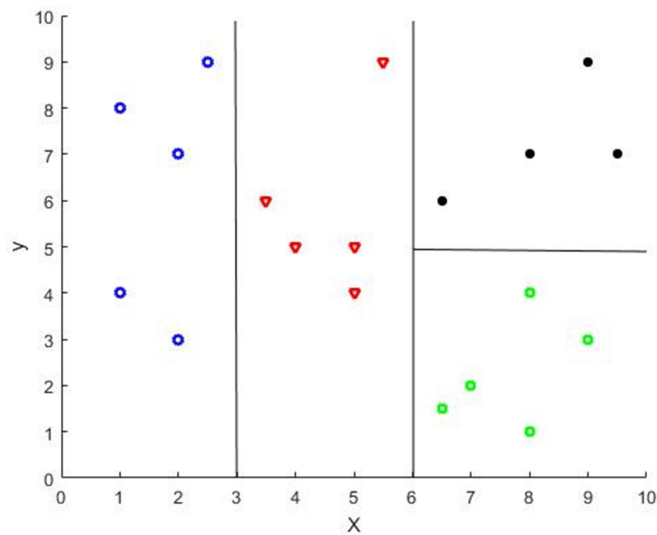


Figure IV.4 : Classification de l'exemple synthétique, frontières induites par l'arbre de décision

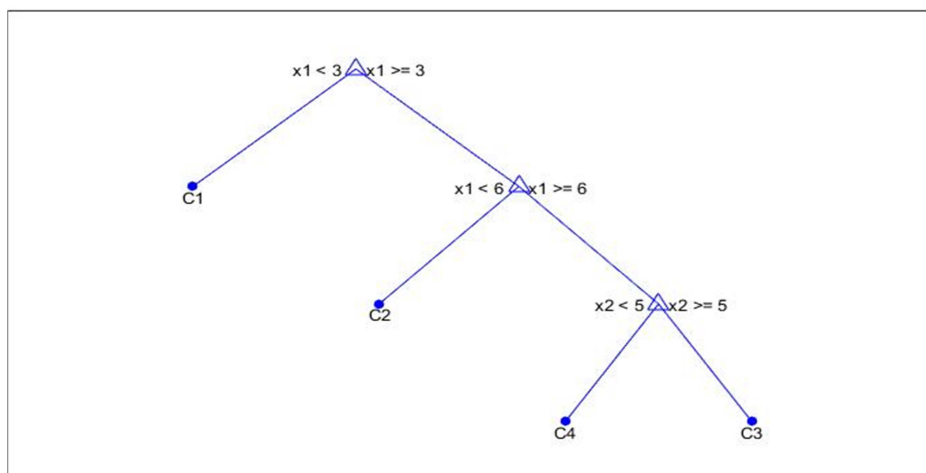


Figure IV.5 : L'arbre de décision pour la classification de l'exemple 2D

IV.4 Description de la base de données

La base de données d'arythmie du MIT-BIH [18] contient 48 extraits d'une demi-heure d'enregistrements d'ECG ambulatoires à deux canaux, obtenus auprès de 47 sujets étudiés par le laboratoire d'arythmie BIH entre 1975 et 1979. Vingt-trois enregistrements ont été choisis au hasard parmi un ensemble de 4000 24- enregistrements ECG ambulatoires d'une heure collectés auprès d'une population mixte de patients hospitalisés (environ 60%) et ambulatoires (environ 40%) à l'hôpital Beth Israel de Boston; Les 25 enregistrements restants ont été sélectionnés dans le même ensemble pour inclure des arythmies moins courantes mais cliniquement significatives qui ne seraient pas bien représentées dans un petit échantillon aléatoire.

Les enregistrements ont été numérisés à 360 échantillons par seconde par canal avec une résolution de 11 bits sur une plage de 10 mV. Deux cardiologues ou plus annotaient indépendamment chaque enregistrement ; Les désaccords ont été résolus pour obtenir les annotations de référence lisibles par ordinateur pour chaque battement (environ 110 000 annotations au total) incluses dans la base de données.

Dans ce travail on s'intéresse aux quatre classes les plus présentes dans cette base de données, à savoir :

- N : Battement Normal (Normal beat)
- V : Extrasystole ventriculaire (PVC, Premature ventricular contraction)
- R : Bloc de branche droit (Right bundle branch block)
- L : Bloc de branche gauche (Left bundle branch block)

La figure (IV. 5) illustre des exemples de ces quatre classes. D'autre part, on considère six enregistrement qui correspondent aux patients ayant ces types d'arythmies : 100, 109,111, 118, 200 et 208.

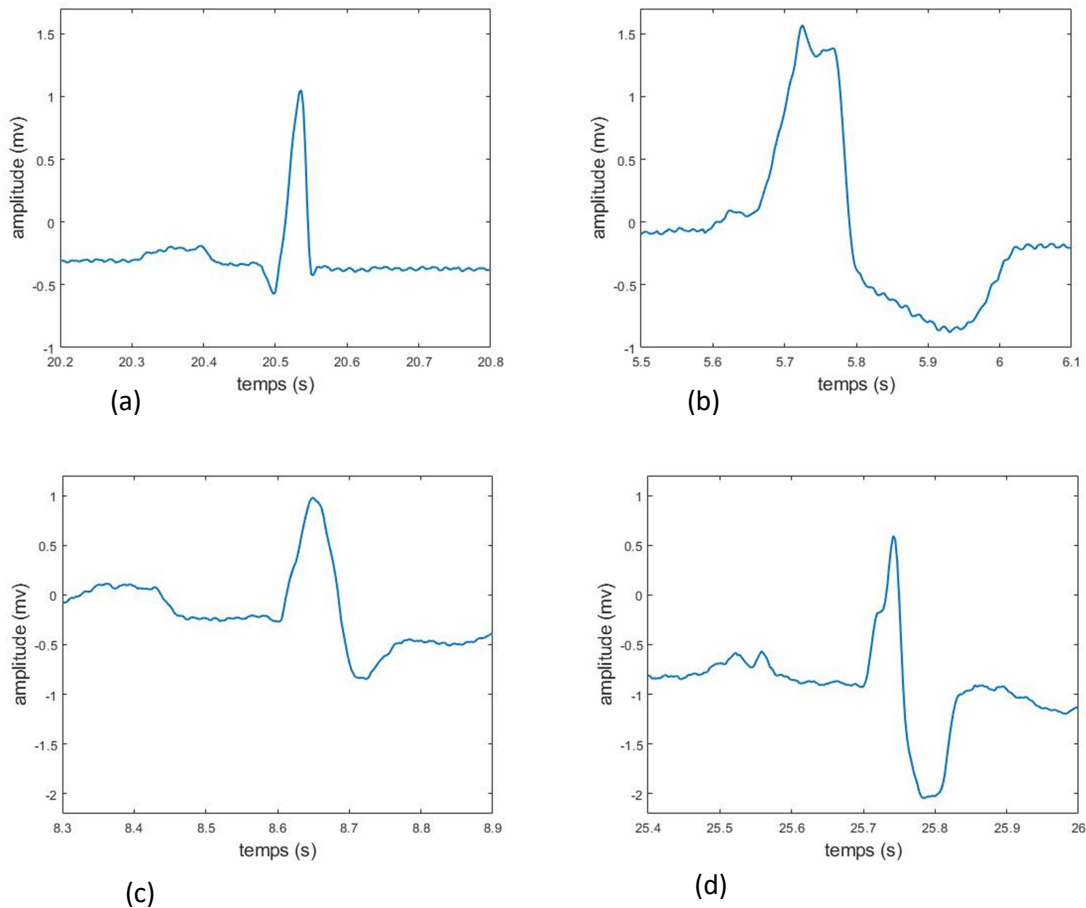


Figure IV.6 : Exemples de battements de la base de données MIT-BIH

(a) Classe 'N' (a) Classe 'V'
 (c) Classe 'L' (d) Classe 'R'

IV.5 Résultats

IV.5.1 1^{er} test : Classe N et Classe V

Dans ce test, nous allons classer les exemples des classe N et V. nous utilisons les Trois enregistrements : 100, 200, 208. Nous effectuons des tests en utilisant deux, trois, quatre et cinq caractéristiques.

Le tableau IV.1 illustre les résultats obtenus pour le premier test avec différents nombre de caractéristiques

Caractéristiques	Nombre de nœuds	Taux apprentissage (%)	Taux de test (%)
2 caractéristiques Durée_QRS RRp	5	91.1555	91.4286
	9	91.5359	91.4286
	19	92.4394	91.6190
	19	92.4394	91.6190
	39	92.8198	91.0476
	55	93.8659	90.4762
2 caractéristiques Valeur R Durée_QRS	5	82.7389	78.6667
	9	83.7375	80.1905
	25	93.1051	91.4286
	33	93.5330	91.8095
	45	93.7233	91.6190
3 caractéristiques Valeur R Durée_QRS RRp	11	94.5792	93.7143
	19	96.0533	95.2381
	27	96.7665	94.8571
	41	97.1469	94.6667
	53	97.7651	95.2381
4 caractéristiques Valeur R Durée_QRS RRp RRs	11	94.5792	93.7143
	19	96.0057	94.6667
	29	97.1945	95.0476
	39	97.6700	95.8095
	49	98.1455	95.4286
	53	98.2882	95.4286
	53	98.2882	95.4286
5 caractéristiques Valeur R Durée_QRS RRp RRs Onde P	11	94.5792	93.7143
	29	97.1945	95.0476
	37	97.6700	95.8095
	51	98.2406	95.4286
	60	98.2882	95.4286
	73	98.4784	95.0476
	73	98.4784	95.0476
	93	98.7161	94.8571

Tableau IV.1 : les résultats obtenus pour le premier test

La figure (IV.7) illustre un exemple d'arbre de décision correspondant au premier test. Cet arbre est obtenu avec 3 caractéristiques et il comprend 7 nœuds. On remarque que cet arbre est facile à suivre.

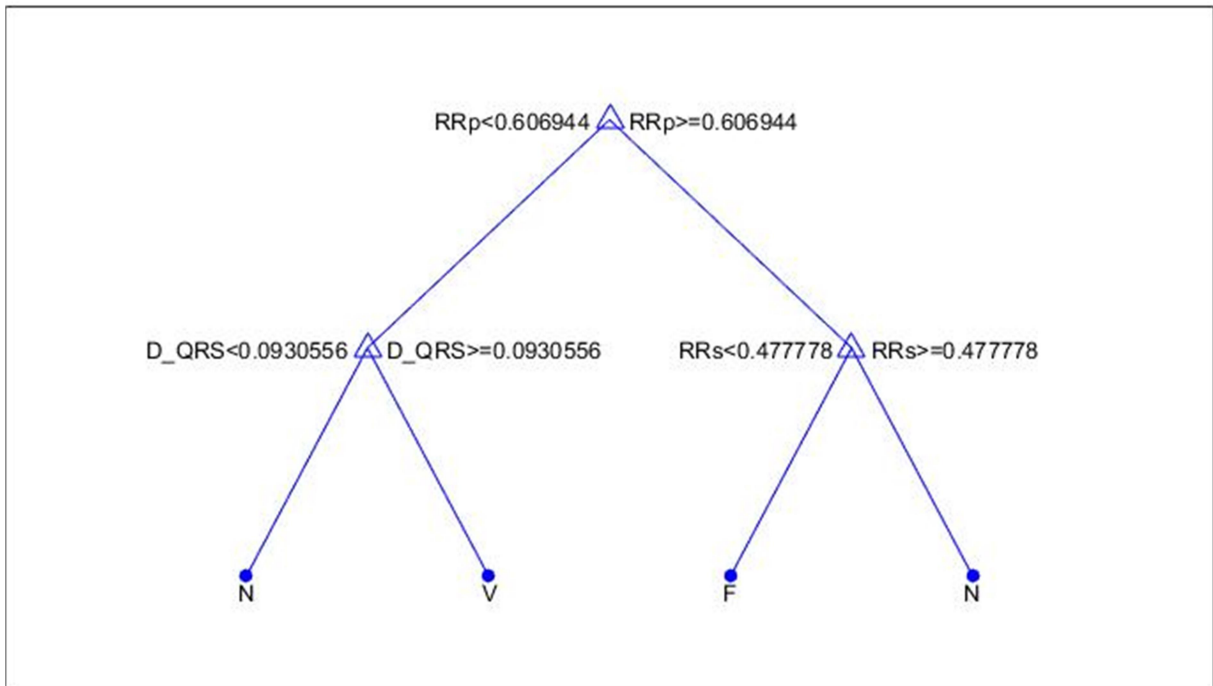


Figure IV.7 : un exemple d'arbre de décision correspondant au premier test

Les règles obtenues à partir de cet arbre sont :

Si $RRp < 0.60$ et Si $D_QRS < 0.09$ Alors la classe est « N »
et Si $D_QRS > 0.09$ Alors la classe est « V »
Si $RRp > 0.60$ et Si $RRs < 0.47$ Alors la classe est « F »
et Si $RRs > 0.47$ Alors la classe est « N »

On remarque que le maximum de taux de test (95.81 %) est le même avec quatre ou cinq caractéristiques et avec 39 et 37 nœuds, respectivement. La cinquième caractéristique (présence de l'onde P) n'a pas donc d'effet dans le premier test, classer N contre V.

On remarque aussi que nous pouvant avoir des résultats acceptable avec seulement deux caractéristiques : Intervalle RR précédant (RRp) et Durée du complexe QRS (D_QRS). Un arbre avec 39 nœuds a permet d'avoir un taux de classification de 91.43 %.

Pour expliquer le problème de sur-apprentissage (over-training), nous représentons les taux d'apprentissage et de test en fonction du nombre de nœuds dans la figure (IV. 8). On remarque clairement que le taux d'apprentissage augment lorsque le nombre de nœuds augmente, mais ce n'est pas le cas pour le taux de test. On note, qu'à partir de 37 nœuds, le

taux de test commence à diminuer. Ceci représente le problème du sur-apprentissage de ce modèle de classification

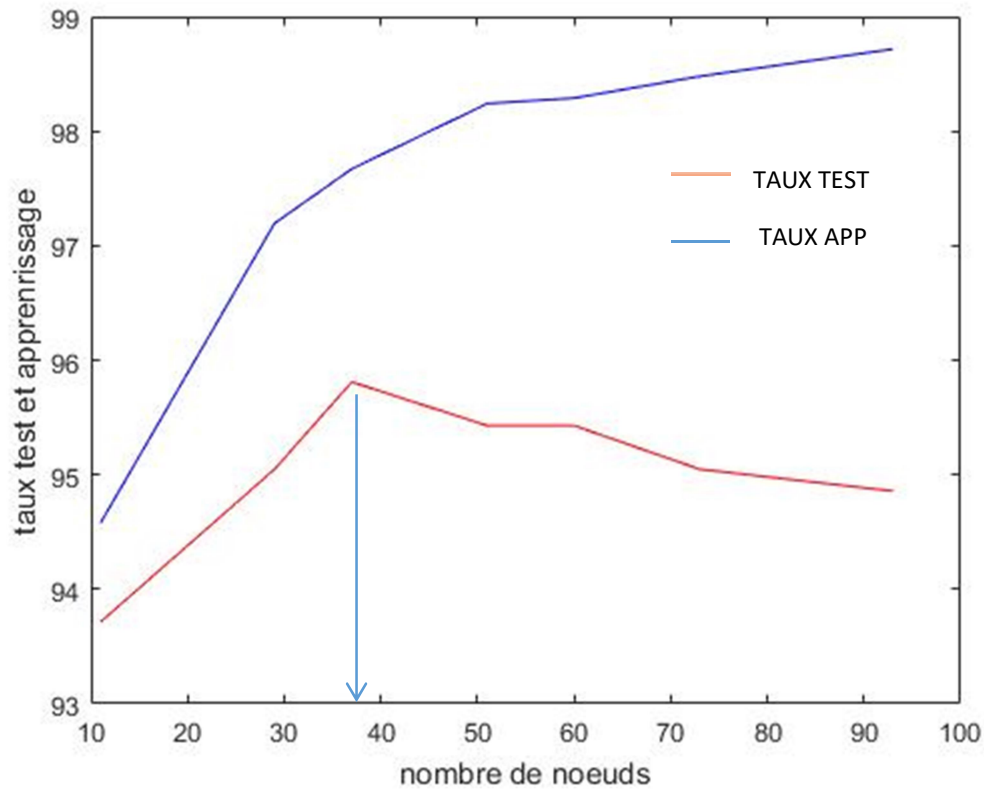


Figure IV.8 : Les taux d'apprentissage et de test en fonction du nombre de nœuds

IV.5.2 2^{ème} test : Classe N et Classes R et L

Dans ce test, nous allons classer les exemples des classe N , R et L. Nous utilisons les quatre enregistrements : 100, 109, 111 et 118.

Caractéristiques	Nombre de nœuds	Taux apprentissage (%)	Taux de test (%)
2 caractéristiques Durée_QRS RRp	7	77.8548	80.0979
	17	80.6688	82.5449
	27	81.0767	82.7080
	45	83.2790	83.0343
	71	85.1142	84.3393
3 caractéristiques Valeur R Durée_QRS RRp	17	85.7667	84.1762
	37	88.0098	87.9282
	55	90.0897	89.5595
	67	90.7423	90.2121
	83	91.7618	91.5171
4 caractéristiques Valeur R Durée_QRS RRp RRs	15	85.7667	86.7863
	41	89.3964	89.8858
	55	90.9462	89.0701
	99	93.0669	89.7227
	133	94.9021	90.7015
5 caractéristiques Valeur R Durée_QRS RRp RRs Onde P	15	86.7047	87.7651
	35	89.8042	90.8646
	59	92.3328	90.5383
	85	93.5563	92.1697
	105	94.7798	91.6803

Tableau IV.2 : les résultats obtenus pour le 2^{ème} test

On remarque que le maximum de taux de test (92.17%) a été obtenu avec cinq caractéristiques et 85 nœuds.

On remarque aussi que nous pouvons avoir des résultats acceptable avec seulement trois caractéristiques : Valeur du pic R, Intervalle RR précédant (RRp) et Durée du complexe QRS (D_QRS). Un arbre avec 85 nœuds a permet d'avoir un taux de classification de 91.52%.

IV.5.3 3ème test : Classe N et Classes R et L et V

Dans ce test, nous allons classer les exemples des classe N, R et L et V Nous utilisons les six enregistrements : 100, 109, 111 et 118 ,200 et 208.

Caractéristiques	Nombre de nœuds	Taux apprentissage (%)	Taux de test (%)
2 caractéristiques Durée_QRS RRp	7	68.6329	68.2877
	17	76.8101	75.1773
	33	78.7089	75.7852
	49	79.1139	75.7852
	75	80.2278	77.0010
3 caractéristiques Valeur R Durée_QRS RRp	11	72.9873	72.4417
	29	79.1139	78.6221
	43	83.4177	81.1550
	63	86.5823	84.9037
	101	88.7089	85.3090
	141	90.1772	86.2209
5 caractéristiques Valeur R Durée_QRS RRp RRs Onde P	165	90.9873	86.6261
	11	75.6203	75.8865
	21	82.8354	82.5735
	75	88.5570	85.9169
	99	89.8228	87.6393
	133	91.0886	88.9564
	165	92.3038	89.4630
	209	93.6456	89.8683
	242	94.4557	89.6657
	289	95.4937	90.2736
	315	95.8481	90.4762
343	96.2025	90.4762	

Tableau IV.3 : les résultats obtenus pour le 3ème test

D'après le tableau III. 3, On peut noter que le maximum de taux de test (90.48 %) a été obtenu avec cinq caractéristiques et 315 nœuds. Ceci peut être expliqué par le fait que la classification de tous les classes et avec tous les patients complique le problème de classification.

IV. Conclusion :

Les tests de classification effectués sur la base de données des arythmies MIT-BIH, nous permettent de noter les points suivants :

- Dans le premier test, classe N et V, le maximum de taux de classification a été obtenu avec quatre ou cinq caractéristiques. Nous pouvons constater que la présence ou non de l'onde P n'a pas d'effet sur la détection des battements V (Extrasystole ventriculaire)
- Les deux caractéristiques : intervalle RR précédant et durée du complexe QRS sont très important pour la détection des battements Extrasystole ventriculaire (classe V)
- Dans le deuxième et le troisième test, tous les cinq caractéristiques sont nécessaires.
- Les arbres de décision compacts (avec peu de nœuds) sont faciles à interpréter et suivre. Mais, lorsque l'arbre devient grand il difficile à suivre, et il présente également le problème de sur-apprentissage.

Conclusion générale

L'application des méthodes informatiques constituent des solutions permettant aux cardiologues d'observer les maladies cardiaques, dans des enregistrements ECG à long terme, avec meilleur résultat de diagnostic. Ainsi, dans ce mémoire, nous avons développé un système de reconnaissance automatique des battements ECG. Dans ce système, nous avons utilisé les arbres de décision en tant que classificateur des battements ECG en se basant sur cinq caractéristiques temporelles : amplitude du pic R, intervalle RR précédant, intervalle RR suivant, durée du complexe QRS et la présence de l'onde P. D'autre part, on s'est intéressé au quatre classes : Battement Normal, Extrasystole ventriculaire, Bloc de branche droit et Bloc de branche gauche.

Pour évaluer notre système, nous avons effectué des tests sur la base de données d'arythmie du MIT-BIH qui contient 48 extraits d'une demi-heure d'enregistrements d'ECG ambulatoires obtenus auprès de 47 sujets. Nous avons effectué trois tests. Dans le premier, la classification des Extrasystole ventriculaire des battement normaux. Dans le deuxième, classification des Bloc de branche droit, Bloc de branche gauche des battement normaux. Dans le troisième, la classification des quatre classes. Les taux de classification abstenus pour ces trois test sont, respectivement, 95.81 %, 92.17 % et 90.48 %.

Les travaux menés dans ce travail montrent la difficulté de la classification des battement ECG. Les résultats obtenue sont acceptables pour la classification des Extrasystole ventriculaire des battement normaux, mais pour les autres tests ne sont pas satisfaisants. Ce modèle nécessite donc beaucoup d'amélioration, notamment sur les caractéristiques. En effet, dans ce travail nous avons utilisé seulement cinq caractéristique temporels, mais pour améliorer les performances de classification d'autres caractéristiques doivent être considérés, comme les caractéristiques fréquentielle et tempo-fréquentielles.

Références

- [1] SÖRNMO, Leif et LAGUNA, Pablo. *Bioelectrical signal processing in cardiac and neurological applications*. Academic Press, 2005.
- [2] <https://www.fedecardio.org/Je-m-informe/Le-coeur/le-fonctionnement-du-coeur>
- [3] <https://www.afblum.be/bioafb/coeur/coeur.htm>
- [4] <https://ecgformation.com/blog/electrodes-derivations>
- [5] <https://litfl.com/ecg-lead-positioning/>
- [6] <https://www.ilephysique.net/sujet-oscilloscope-et-ecg-tp-seconde-260081.html>
- [7] JOBBÁGY, Ákos, VARGA, Sándor, *Biomedical instrumentation, : Biotechnológia és bioinformatika tananyagok*, 2013.
- [8] MACFARLANE, Peter W., VAN OOSTEROM, Adriaan, PAHLM, Olle, *et al.* (ed.). *Comprehensive electrocardiology*. Springer Science & Business Media, 2010.
- [9] CHAUDHURI, Subhasis, PAWAR, Tanmay D., et DUTTAGUPTA, Siddhartha. *Ambulation analysis in wearable ECG*. Springer, 2009.
- [10] M.Nemessi, «Classification et reconnaissance des formes par algorithmes hybrides,» These de doctorat, Université de Guelma, 2009.
- [11] LUZ, Eduardo José da S., SCHWARTZ, William Robson, CÁMARA-CHÁVEZ, Guillermo, *et al.* ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer methods and programs in biomedicine*, 2016, vol. 127, p. 144-164.
- [12] MILLIS, Richard (ed.). *Advances in Electrocardiograms: Methods and Analysis*. BoD–Books on Demand, 2012.
- [13] PAN, Jiapu et TOMPKINS, Willis J. A real-time QRS detection algorithm. *IEEE transactions on biomedical engineering*, 1985, no 3, p. 230-236.
- [14] ROKACH, Lior et MAIMON, Oded Z. *Data mining with decision trees: theory and applications*. World scientific, 2008.
- [15] QUINLAN, J.. Ross . *Induction of decision trees*. Machine learning, 1986, vol. 1, no 1, p. 81-106.
- [16] QUINLAN, John. *C4. 5: programs for machine learning*. Elsevier, 2014.
- [17] BREIMAN, Leo, FRIEDMAN, Jerome, OLSHEN, Richard, et al. *Classification and regression trees*. Wadsworth Int. Group, 1984, vol. 37, no 15, p. 237-251.
- [18] MOODY, George B. et MARK, Roger G. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 2001, vol. 20, no 3, p. 45-50.
- [19] SILVA, Ikaro et MOODY, George B. An open-source toolbox for analysing and processing physionet databases in matlab and octave. *Journal of open research software*, 2014, vol. 2, no 1.