

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Master

Filière Électronique
Spécialité Instrumentation

présenté par

Benamar Dounia

&

Ahmed seghir manel

Analyses des signaux cardiaques par la méthode d'intelligence artificielle « réseaux de neurones convolutifs CNN »

Proposé par : Nom et prénom Enseignant 1 & Nom et prénom Enseignant 2

Année Universitaire 2019-2020

Nous remercions d'abord ALLAH le tout puissant qui nous a guidés et donnés la force et la volonté de réaliser cette mémoire. Nôtres pensés vont vers nos parents, qui ont toujours cru en nous. C'est grâce à leurs soutient et prières que nous avons accomplis ce travail.

Nous remercions notre promoteur Mr Ykhlef Farid d'avoir pris en charge et aidé tout au long du projet. Et de nous avoir orientés avec ces précieux conseils et son soutient moral vue ces dernières conséquences du COVID_19.

Remerciements les plus sincères à toutes les personnes qui ont de près ou de loin à l'évaluation de cette mémoire. Enfin, nous tenons aussi à remercier les jurés pour avoir accepté et de juger notre travail.

Listes des acronymes et abréviations

OMS : Organisation **M**ondiale de la **S**anté.

MCV : **M**aladie **C**ardio-**v**asculaire.

E.C.G. : Electro**C**ardio**G**amme.

STEMI : Infarctus du **M**yocarde avec élévation du **ST**.

PVC : Contraction **V**entriculaire **P**rématuration.

TV : Tachycardie **V**entriculaire.

FV : Fibrillation **V**entriculaire.

SVT : Tachycardie **S**upra**v**entriculaire.

V : **V**entriculaire.

Q.R.S : durée de complexe **QRS**.

NS : **N**œud **S**inusal.

NAS : **N**œud **A**uriculo **V**entriculaires.

aVF : **aVL** : **aVR** : **D**érivation électrocardiographie unipolaire .

ARN : **R**éseau de **N**eurones **A**rtificielle.

CVD : **S**ystème **C**ardio-**V**asculaire.

CNN : **C**onvolutional **N**eural **N**etwork.

ML : **M**achine **L**earning.

AI : **A**rtificielle **I**ntelligence.

MLP : **M**ulti **L**ayer **P**erceptron.

RNA : **A**rtificiel **N**eural **N**etwork.

RNN : **R**écurrent **N**eural **N**etwork.

RCE : **R**estricted **C**oulomb **E**nergy.

AM : **A**pprentissage **M**achine.

SVM : Support Vecteur **M**achine.

FC : Fully Connected.

Table de matière

Chapitre I : Généralité sur les signaux cardiaque

1.1 Introduction	4
1.1.1 Système cardiovasculaire.....	4
1.2 Définition de l'électrocardiogramme	9
1.2.1 Histoire de l'ECG.....	9
1.2.2 But de l'ECG.....	11
1.2.3 Principe de fonctionnement	11
1.3 Technique de l'enregistrement	14
1.3.1 L'enregistrement	14
1.3.2 Les dérivations de l'ECG.....	15
1.4 Analyser l'ECG.....	17
1.4.1 Les phases de l'ECG.....	17
1.4.2 Les caractéristiques de l'ECG.....	18
1.4.3 La fréquence de l'ECG.....	19
1.5 Conclusion	19

Chapitre II : Technique de l'apprentissage machine

2.1 Introduction.....	20
2.2 L'intelligence artificielle.....	21
2.2.1 Le but de l'intelligence artificielle.....	22
2.2.2 Le domaine de l'intelligence artificielle.....	22
2.2.3 L'utilisation de l'intelligence artificielle pour classifier l'ECG.....	22
2.3 Représentation de l'apprentissage machine.....	23
2.3.1 Apprentissage machine.....	23
2.3.2 Type de l'apprentissage machine.....	24
2.4 Architectures de réseaux neurones profonds.....	26
2.4.1 Réseaux de neurones.....	26
2.5 Réseaux de neurones conventionnels.....	27
2.5.1 Architecture de réseau de neurones conventionnels.....	28
2.5.2 Exemples de modèles de CNN.....	32
2.6 Choix de paramètres.....	33
2.6.1 Nombre de filtre.....	33
2.6.2 Forme de filtre.....	33
2.6.3 Forme du Max Pooling.....	34
2.7 Méthode de régularisation.....	34
2.7.1 Empirique.....	34
2.7.2 Explicite.....	34
2.8 Histogramme 2d.....	35
2.9 Conclusion.....	36

Chapitre III : Méthodologie de recherche

3.1 Introduction.....	37
3.2 Logiciel et librairies utilisées dans l'implémentation.....	37
3.2.1 Python	37
3.2.2 Keras.....	38
3.2.3 Numpy.....	39
3.2.4 Pandas.....	39
3.2.5 Matplotlib.....	39
3.2.6 Scikit-learn.....	39
3.2.7 Seaborn : visualisations de données statiques.....	40
3.3 La base de données.....	40
3.3.1 Lecture de base de données.....	41
3.3.2 Représentation des classes.....	42
3.3.3 Rééchantillonnage pour équilibrer l'ensemble de données.....	42
3.4 Tracé des classes ECG.....	46
3.5 Retraitement des données ECG.....	49
3.6 Model training.....	55
3.7 Résultats.....	56
3.8 Conclusion.....	59
Conclusion Générale.....	60
 Bibliographie.....	 61

Liste des Figures

➤ Chapitre I : Généralité sur les signaux cardiaque

Figure 1.1 : Représentation schématique du système cardiovasculaire	5
Figure 1.2 : Représentation du cœur humaine	6
Figure 1.3 : Schéma du circuit sanguin. La couleur bleue indique que le sang est pauvre en oxygène tandis que la couleur rouge correspond Au sang chargé en oxygène.....	7
Figure 1.4 : Principes des systoles auriculaire et ventriculaire.....	8
Figure 1.5 : représentation de l'activité électrique.....	9
Figure 1.6 : Electrocardiogramme ECG Contec100G et Edan PADECG Numérique sans fil avec interprétation et tablette tactile androïde	11
Figure 1.7 : L'électrophysiologie cardiaque du cœur	12
Figure 1.8 : Intervalles de l'ECG normal, Incluant l'onde U.....	13
Figure 1.9 : dérivations précordiales.....	16

➤ Chapitre II : Technique de l'apprentissage machine

Figure 2.1 : La relation entre l'apprentissage machine et l'intelligence artificielle.....	20
Figure 2.2 : Apprentissage machine.....	24
Figure 2.3 : l'apprentissage supervisé.....	25
Figure 2.4 : Le mode faciles non-supervisé	26
Figure 2.5 : Une couche du CNN en 3 dimensions (Vert =Volume d'entrée, bleu = volume du champ récepteur, gris = couche de CNN, cercles = neurones artificiels indépendants)	29
Figure 2.6 : Ensemble de neurones (cercle) créant la profondeur d'une couche de convolution (bleu). Ils sont liés à un même champ récepteur	30
Figure 2.7 : Pooling avec un filtre 2*2 et pas de 2	31
Figure 2.8 : Structure de CNN	32

➤ **Chapitre 3 : 1. Méthodologie de recherche**

Figure3.1 : logo de python.....	38
Figure 3.2 : logo de Keras.....	38
Figure 3.3 : Logo de NumPy.....	39
Figure 3.4 : Représentation des classes dans un cercle.....	43
Figure 3. 5: Méthode de l'échantiollnage.....	44
Figure 3. 6: Méthode de sous échantiollnage.....	45
Figure 3. 7: Réprésentation des classes équilibré dans un cercle.....	46
Figure 3. 8: Battement normal.....	47
Figure 3. 9: Battement inconnu.....	47
Figure 3. 10: Battement ectopique ventriculaire.....	48
Figure 3. 11: Battement ectopique supraventriculaire.....	49
Figure 3. 12: Battement de fusion.....	49
Figure 3. 13: Représentation de transformation la classes 0 en histogramme de deux dimension.....	52
Figure 3. 14 : Représentation de transformation la classes 1 en histogramme de deux dimension.....	52
Figure 3. 15: Représentation de transformation la classes 2 en histogramme de deux dimension.....	53
Figure 3. 16: Représentation de transformation la classes 3 en histogramme de deux dimension.....	54
Figure 3. 17: Représentation de transformation la classes 4 en histogramme de deux dimension.....	54
Figure 3. 18: Représentation de signal ECG dans le cas parfait.....	55
Figure 3. 19: Représentation de signal ECG avec bruit.....	55
Figure 3. 20: Le model accuracy pour l'apprentissage et le teste.....	59

Figure 3.21: Le model loss pour l'apprentissage et teste.....59

Introduction générale

Nous vivons dans un environnement qui évolue à un rythme rapide, et les mêmes forces efficaces façonnent les caractéristiques de la santé humaine. Dans toutes les régions du monde. L'un des exemples les plus évidents de cette transformation est le fait que les maladies non transmissibles, comme les maladies cardiovasculaires, ont dépassé les maladies infectieuses pour devenir la principale cause de décès dans le monde.

Et l'hypertension artérielle, également connue sous le nom d'hypertension artérielle, est l'un des principaux facteurs de risque d'insuffisance cardiaque congestive. Et les vaisseaux sanguins. Qu'un milliard de personnes dans tous les pays du monde souffrent d'hypertension artérielle, ce qui conduit au développement de maladies qui tuent 9 millions de personnes chaque année à l'heure actuelle.

L'hypertension artérielle provoque rarement des symptômes et elle n'est pas diagnostiquée chez de nombreuses personnes atteintes de cette maladie, et celles qui reçoivent un diagnostic de cette maladie peuvent ne pas être en mesure de guérir. Selon l'organisation mondiale de la santé (OMS) les maladies cardiovasculaires (MCV) sont aujourd'hui la première cause de décès. plus de 17.7 millions de personnes sont décédées de maladies cardiovasculaires en 2017 dans le monde, soit environ 31% de intermédiaire. L'arythmie est un type représentatifs de maladies cardiovasculaire qui fait référence à tout changement irrégulier par rapport aux rythmes cardiaques normaux. Il existe plusieurs types d'arythmie, notamment la fibrillation auriculaire, la contraction prématurée, la fibrillation ventriculaire et les tachycardies. Bien que le rythme cardiaque d'arythmie unique puisse ne pas avoir un impact sérieux sur la vie, des battements d'arythmies continues peuvent entraîner des circonstances mortelles. Par exemple, des battements prolongés de contraction ventriculaire prématurée (PVC) se transforment parfois en une tachycardie ventriculaire (TV) ou des battements de

fibrillation ventriculaire (FV) qui peuvent immédiatement conduire à une insuffisance cardiaque. Ainsi, il est important de surveiller périodiquement les rythmes cardiaques pour gérer et prévenir les maladies cardiovasculaires. L'Electrocardiogramme (ECG) est un outil médical non invasif qui affiche le rythme et l'état du cœur. Cependant, des études antérieures chez des médecins et des ambulanciers paramédicaux ont montré que la précision du diagnostic par ECG n'est pas suffisante, Ainsi, nous proposons un algorithme de détection basé sur un réseau de neurones à convolution (CNN) pour la détection de l'infarctus du myocarde avec élévation du ST (STEMI) est une maladie aigue potentiellement mortelle.

Le reste de notre recherche est abordé en trois chapitres principaux :

- Le premier chapitre traite des notions théoriques générales pour donner l'aide à la compréhension sur le système cardiovasculaire et l'activité cardiaque en particulier et les principes de l'électrocardiographie
- Le deuxième chapitre nous expliquons l'intelligence artificielle et les méthodologies utilisées pour la classification des arythmies ECG, y compris le prétraitement des données ECG et le classificateur de réseau se neuronal convolutif
- Le troisième chapitre présente notre application finale, à savoir la conception d'un réseau de neurones

Enfin, nous clôturons notre travail par une conclusion générale

OBJECTIFS

- Argumenter les principales indications de l'ECG et discuter l'interprétation des résultats.
- Connaître les dérivations, les paramètres de quantification de l'ECG normal, la définition des caractéristiques et des phases de l'ECG.
- Connaître les différents territoires explorés par les différentes dérivations.
- Savoir diagnostiquer une hypertrophie atriale ou ventriculaire, droite ou gauche.
- L'apprentissage automatique extraire et exploiter automatiquement l'information présente dans un jeu de données.

Chapitre 1 Généralité sur les signaux cardiaque

1.1 Introduction

Dans ce chapitre , nous traitons la morphologie du cœur humain et apprenons la caractéristique , la fonction et l'importance du système cardiovasculaire , le système cardiovasculaire se compose du cœur et des vaisseaux sanguins , et sa fonction principale est d'assurer que le flux sanguin continue de manière suffisantes et sous une pression suffisante sur les organes et les tissus pour répondre aux besoins énergétiques et régénérer les cellules .Indépendamment des conditions environnementales et de l'activité individuelles , Selon l'organisation mondiale de la santé (OMS) les maladies cardiovasculaire (MCV) sont la base de la mort dans le monde entier, l'électrocardiogramme (ECG) est une méthode peu couteuse , bien connue et non invasive pour évaluer l'activité électrique du cœur.

Cette première partie de la thèse introduit des concepts généraux des notions d'anatomie et d'électrophysiologie, il décrira le fonctionnement du cœur et du système vasculaire , ses principales fonctions et ses principaux composants .les caractéristiques du cycle cardiaque et du processus de prolifération cardiaque seront : ensuite , l'accent sera mise sur ECG: sa genèse , les différents types d'enregistrement , et ses composantes qui seront étudiées dans les parties suivantes .En fin , nous fournirons quelque concepts de physiologie qui nous seront utiles pour la comprendre les résultats de l'analyse des périodes cardiaques.

1.1.1 Système cardiovasculaire

Le système cardiovasculaire se compose du cœur et du système, et leurs fonctions sont schématisées sur la figure, et sa fonction principale est d'assurer la circulation de la sanguine dans le corps afin de répondre aux besoins énergétiques et de régénérer les cellules quelles que soient les conditions environnantes et l'activité de l'individu afin de réaliser cette activité cardiaque.

La pression artérielle doit être soumise à une Régulation permanente, qui définit le concept d'homéostasie [1].

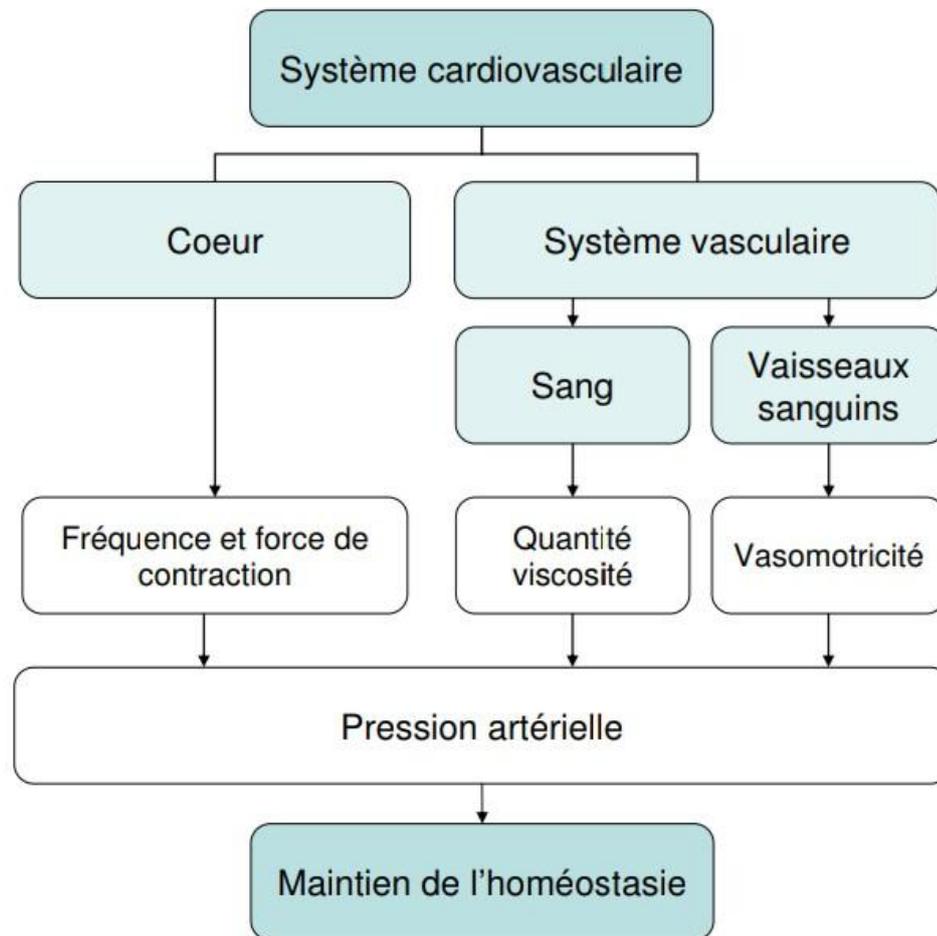


Figure 1.1 : Représentation schématique du système cardiovasculaire.

a Anatomie du cœur humaine

Le cœur est un organe contractile qui assure la circulation sanguine, C'est le composant principale du system cardiovasculaire, C'est un muscle strié creux qui se divise en deux moitiés indépendantes, il permet au sang d'être de pompé le sang en le faisant circuler dans tous les corps tissus corporels.

Le travail du cœur dépend de la chronologie d'événements mécaniques et électriques complexes qui sont régulés dynamiquement. L'activité cardiaque est cyclique ; la séquence de tous les évènements qui la composent dure environ un second. Répété tout au long de la vie, assurant une circulation sanguine et une oxygénation continue. Ainsi, le cœur moyenne environ 6 litres d'un sang par minute, soit environ 220 millions de litres de sang en 70 ans.

Le cœur se compose de deux parties : le côté gauche est séparé par un muscle, qui est le septum : chaque partie se compose d'une cavité supérieure « oreillette » et d'une cavité inférieure « ventricule », la partie droite contient du sang pauvre en oxygène et assure la circulation pulmonaire et ses branches vers les poumons (petite circulation) et la partie gauche contient du sang riche en oxygène qui le pousse à travers tous les tissus (grande circulation) chaque oreillette est séparée des ventricules par une valve d'échappement qui assurent la connexion entre le ventricule droit (valve pulmonaire), ainsi qu'entre le ventricule gauche et l'artère aorte (valve aortique). Dans la partie supérieure de l'oreillette droite se trouve un petit morceau de tissu cardiaque spéciale appelé nœud sino-auriculaire (nœud sinusal) qui contrôle l'ensemble du mécanisme de régulation du rythme cardiaque [2].

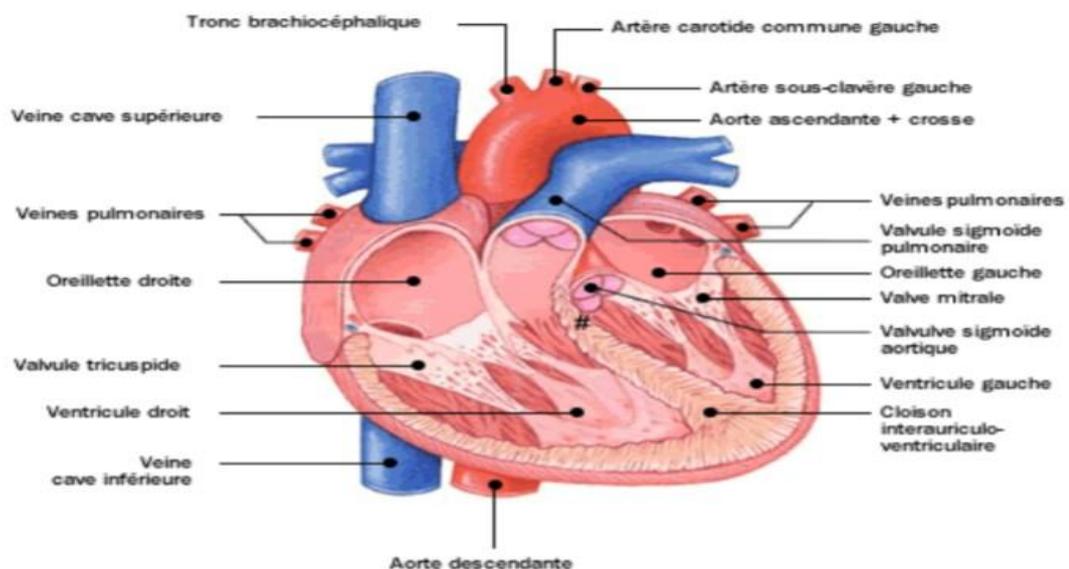


Figure 1.2 : Représentation du cœur humaine

Le cœur comprend un système de conduction électrique automatique qui en assure chaque impulsion, et la contraction du muscle cardiaque régit une impulsion électrique provenant du nœud sinusal situé au niveau de la veine cave supérieure, dans l'oreillette droite, ce nœud se compose d'un groupe de cellules automatiquement excitées qui génèrent un courant de dépolarisation 60 à 100 fois par minute. Ce nœud est considéré comme un stimulateur cardiaque du cycle cardiaque.

Cette excitation électrique se propage dans aux oreillettes, provoquant la contraction des oreillettes et jusqu'au nœud auriculo-ventriculaire, qui est le seul point de passage entre les oreillettes et les ventricules situé au niveau du septum ventriculaire.

b Le circuit sanguin

Le cercle sanguin est présent schématisé sur la figure 1.3 Son passage travers le corps entre dans l'oreillette droite à travers la veine cave supérieure et inférieure. Puis traverse le ventricule droit d'où il est éjecté dans l'artère pulmonaire vers les poumons, Ce cycle correspond à ce qu'on la circulation dite pulmonaire, ou la petite circulation (microcirculation).

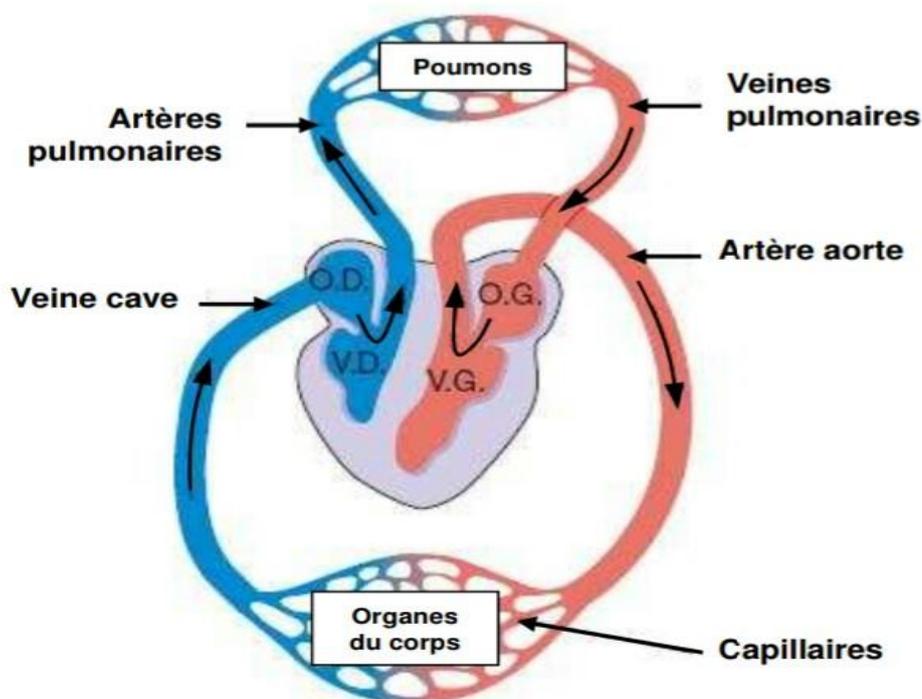


Figure 1.3 : Schéma du circuit sanguin. La couleur bleue indique que le sang est pauvre en oxygène tandis que la couleur rouge correspond Au sang chargé en oxygène

c Le battement cardiaque

Chaque battement cardiaque provoque une chaîne combinée d'évènements mécaniques et électriques appelé battements cardiaques. Cela se compose de trois phases ventriculaires .il comprend la contraction des ventricules et l'expulsion du sang vers les oreillettes. Le système circulatoire ferme les deux valves sigmoïdes – la valve pulmonaire à droite et la valve aortique à gauche –une fois que le sang est sorti. En fin de compte la diastole est la relaxation de toutes les parties du cœur, permettant le remplissage passif du ventricule et l'arrivée de nouveau sang.

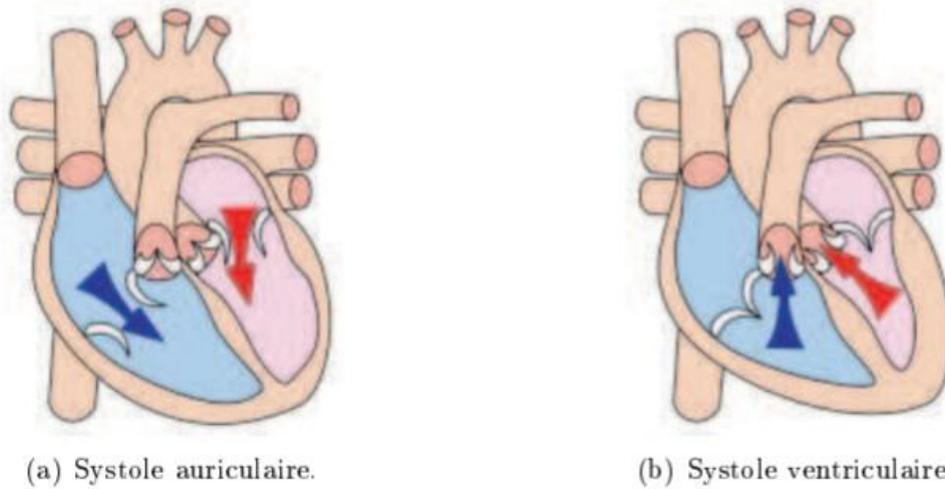


Figure 1.4 : Principes des systoles auriculaire et ventriculaire

d **Activité électrique du cœur humaine**

La stimulation électrique du cœur est contrôlée par des cellules spécialisées dans les fonctions d'excitation et de conduction (Gulizia), Les cellules responsables de l'excitation sont automatisées et génèrent automatiquement un potentiel d'action (stimulateur physiologiques), Les cellules responsables de la propagation de ces micro-courants ou impulsion sont des fibres spécialisées pour la conduction intracardiaque (fibres lentes du nœud AV et fibres rapides du réseau de His-Purkinje).

Les courants microscopiques qui sont prélevés sont responsables de la dépolarisation des cellules musculaires, permettant au calcium de pénétrer dans les cellules et provoquant la contraction, puis il commence à se dépolariser, Le cycle de stimulation cardiaque se traduit dans l'ECG au moyen de signaux électrique qui témoignent de l'activation successive des différents sites anatomiques fonctionnels.

Ce cycle commence par l'activation du nœud sinusal (sans translation sur l'électrocardiogramme), puis l'activation des oreillettes (onde P) et du nœud AV (auriculo-ventriculaire) et la descente de l'impulsion dans son tronc de son faisceau de His (segment PQ) puis dans ses branches de faisceau de His avant l'activation ventriculaire des fibres de Purkinje (complexe QRS), Cette activation ventriculaire génère une phase de contractile qui continue à travers le segment ST et se termine à la fin de l'onde T qui annonçant la relaxation [3].

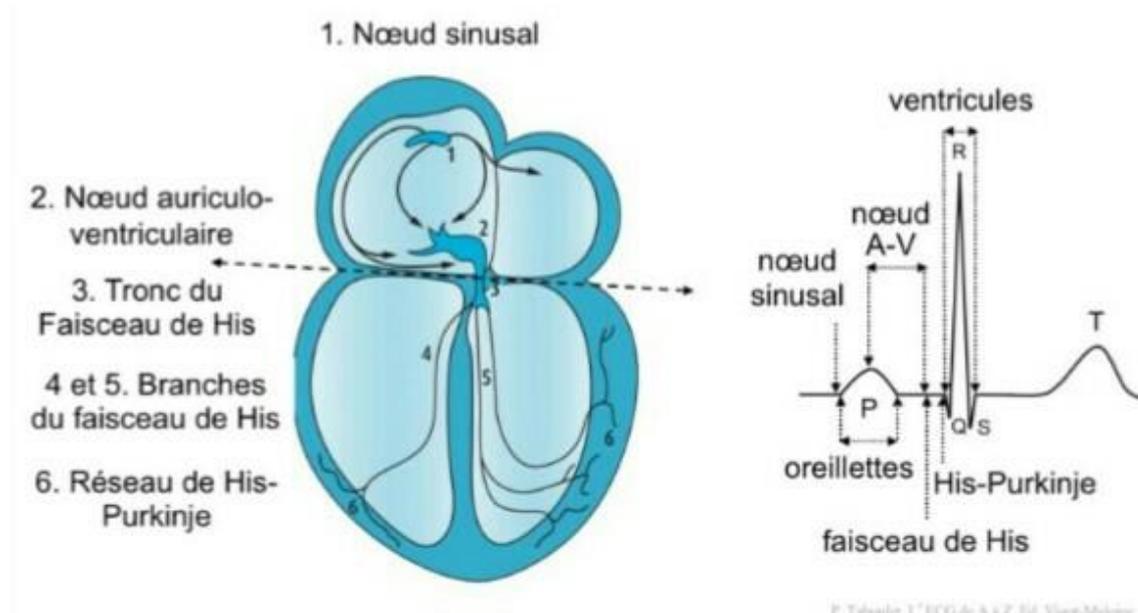


Figure 1.5: Représentation de l'activité électrique

1.2 Définition de l'électrocardiogramme

L'électrocardiogramme (ECG) est l'enregistrement sur un support papier standard la stimulation électrique, et ces enregistrements permettent aux cardiologues de mesurer la fréquence cardiaque et de détecter les anomalies. Un électrocardiogramme est une représentation graphique de la tension électrique et de l'activité cardiaque qui contrôle l'activité des muscles cardiaques sous la forme d'un électrocardiogramme. C'est une méthode utilisée pour analyser le système de conduction du cœur permettant « d'obtenir des informations sur l'activité électrique du cœur ». Ils voyagent à travers les oreillettes et les ventricules sur papier, donnant à la médecine la possibilité de juger de l'état structurel et fonctionnel du cœur d'un patient, L'électrocardiogramme ECG est obtenu à partir d'une machine. Il, qui enregistre l'activité mécanique du cœur sous la forme d'un signal électrique. Ce signal électro-physiologique est collecté par des électrodes métalliques bien placée à la surface de la peau [4].

1.2.1 Histoire de l'ECG

L'ECG a traversé plusieurs étapes historiques pour atteindre sa forme actuelle [5]

1842 : Le physicien italien Carlo Matteucci montre le phénomène physiologique produit par le muscle de contraction, chaque contraction du cœur étant accompagnée d'un courant électrique.

1843 : Le physiologiste allemand Emil Dubois Reymond démontre une possibilité réelle de contraction musculaire, confirmant les recherches de Carlo Matteucci.

1870 : Alexander Moorhead, de l'hôpital St. Bartholomew de Londres, aurait enregistré un plan cardiaque électrique.

1878 : Les physiciens britanniques John Borden Sanderson et Frederick Big ont enregistré l'énergie électrique du cœur de la grenouille avec une échelle poétique à deux niveaux (appelée plus tard P, QRS et T).

1884 : John Burden Sanderson et Frederick Page publient certains de ces enregistrements sur les phénomènes d'excitation électrique au cœur de la tortue.

1887 : Le physicien britannique Augustus de. Walter, du Collège de médecine St. Mary à Londres, a publié le premier schéma électrique du cœur humain. Il a été enregistré sur Thomas Joswell, un technicien de laboratoire.

1891 : Les physiciens britanniques William Baillis et Edward Starling de Collège College London améliorent le niveau poétique. Reliez la pointe au bras droit et au cuir chevelu et visualisez le "contraste triphasé associé à chaque contraction du cœur". Ces écarts seront appelés plus tard P, QRS, T. Bayliss, WM et Starling EH. Sur les différences électriques du cœur chez l'homme.

1895 : En utilisant une formule de correction électrique améliorée développée indépendamment par Porsche, Ainhoffen met en évidence cinq écarts décrits par P, Q, R, S et T.

1906 : Antofen publie le premier classement du GEC. Ventricule hypertrophique gauche et droit, hypertrophie auriculaire gauche et droite, éléments d'onde U sur QRS, Contractions ventriculaires prématurées, ventricule jumeau ventriculaire, flutter auriculaire et masse auriculaire ventrale complète : normales et anormales.

1908 : Edward Schaffer, de l'Université d'Édimbourg, est le premier à acheter le Galvanomètre à usage médical.

1910 : Le premier magazine américain sur les électrocardiogrammes, Walter James de L'Université Columbia et Harthy Williams de la Faculté de médecine de l'Université Cornell à New York. Ils décrivent l'hypertrophie ventriculaire, les spectres ventriculaires et ventriculaires et la fibrillation auriculaire ventriculaire.

1911 : Thomas Lewis publie The Heartbeat Machine à Londres : Chau and Sons, dédié à Einhoven.

1920 : Harold Hardy, New York, publie le premier infarctus du myocarde avec Électrocardiogramme chez l'homme, décrivant l'onde T comme élargie. Hardy Hipp. Signe ECG de blocage des artères coronaires.

1924 : Willem Einthofen, **prix Nobel** pour l'invention d'un dispositif ECG.

1928 : Le changement ECG de Frank Sanborn dans la première version portable de 50 livres fonctionne sur une batterie de voiture de 6 volts.

1938 : L'American Heart Association et la Heart Association de Grande-Bretagne déterminent la position standard des branches V1-V6 précondiales. AR Barnes, Pardee HEB, PD blanche. Et d'autres Unification des indices

1942 : Emmanuel Goldberger ajoute aux dérivés introvents de aRr, aVI, aVf. Cela lui permet, avec 6 V1-V6 pré-filetés, de réaliser le premier planificateur de cœur à 12 canaux, qui est encore utilisé aujourd'hui (**Figure I.6** [6]).



Figure 1.6 : Electrocardiographe ECG Contec100G et Edan PADECG Numérique sans fil avec interprétation et tablette tactile androïde

1.2.2 But de l'ECG

Le but de l'électrocardiogramme est de mesurer et d'enregistrer l'activité électrique du système cardiaque. Cet enregistrement est représenté comme une illustration du rythme cardiaque, il vise au diagnostic précoce de nombreuses maladies, crises cardiaques, et identifie également les problèmes musculaires cardiaques ainsi que les diagnostics d'électrocardiogramme comme l'angine de poitrine et les crises cardiaques. L'électrocardiogramme a pour but de surveiller le bon fonctionnement du système cardiaque et peut être pratiqué au repos, à l'effort ou pendant plus de 24 heures, selon le médecin. En l'absence de symptômes, l'ECG vise à détecter un problème cardiaque qui aurait pu passer inaperçu.

1.2.3 Principe de fonctionnement de l'ECG

Nous devons être conscients que le diagnostic ECG est outil et non une fin en soi. Pour lui l'analyse est basée sur quelques principes de base et des règles simples.

Électrocardiogrammes La contraction des muscles du corps est causée par des changements électriques appelés la double polarisation, avec des électrodes attachées au corps, peut détecter ces changements. Le muscle cardiaque, semblable aux muscles du corps, peut détecter ses changements, A condition que le corps du patient soit complètement détendu électriquement, afin de l'éviter modifications électriques des muscles squelettiques [5].

➤ Cycle d'électrophysiologie cardiaque

L'impulsion initiale de l'ECG par cycle cardiaque provient des ondes conduction (poutre de support), puis doublée pour faire une systole ventriculaire, suivie d'une repolarisation [5].

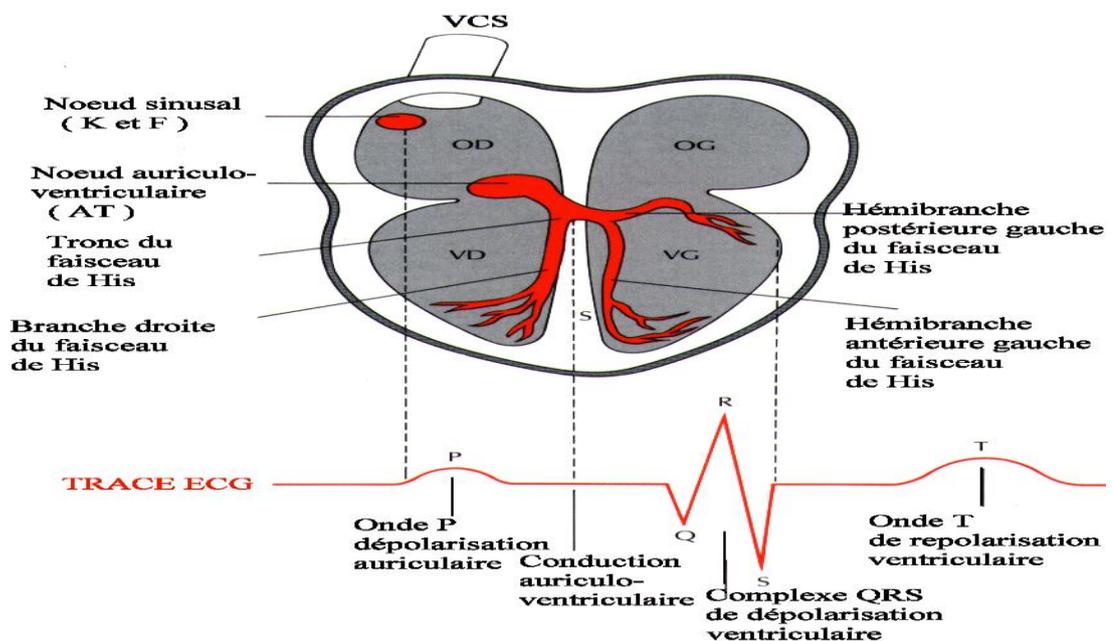


Figure 1.7 : L'électrophysiologie cardiaque du cœur

➤ Les intervalles de l'ECG

L'analyse de l'électrocardiogramme comprend la mesure des amplitudes et durée ainsi que l'examen de la morphologie de l'onde P, du complexe QRS, de l'onde T, de l'intervalle PR, du segment ST, de l'intervalle QT, En plus des différentes ondes qui sont les paramètres de base pour une bonne caractérisation d'un signal ECG [5].

Les intervalles et les segments les plus importantes sont :

L'onde P : l'onde de dépolarisation auriculaire, souvent positive, de faible amplitude (1 à 3m volts) et <0.12 seconde en D2.

Le complexe QRS : c'est l'activation et à la dépolarisation des ventricules

- L'onde Q : première déflexion négative : activation septale
- L'onde R : première déflexion positive : activation pariétale du VG
- L'onde S : déflexion négative qui suit l'onde R : activation basale du VG

L'intervalle de la durée QRS varie de 0.06 à 0.1 secondes (3 à 5 petite carreaux) et se mesure du début jusqu'à la fin QRS.

L'onde T : la durée de repolarisation ventriculaire. Son amplitude est inférieure à 2 mm.

Segment PR : (pause nœud AV) : le segment PR correspond au délai entre la fin de la dépolarisation des oreillettes et le début de celle des ventricules

L'onde U : repolarisation tardive de zones myocardiques d'amplitude inscrite entre de l'onde P et de l'onde T < à quartier de l'amplitude de l'onde T.

Intervalle PR : (durée de conduction auriculo-ventriculaire) l'intervalle PR correspond à la durée de propagation de l'onde de dépolarisation du nœud sinusal jusqu' aux cellules myocardiques.

Intervalle QT : (durée de systole ventriculaire) Cet intervalle correspond au temps de systole ventriculaire, qui va du début de l'excitation des ventricules jusqu' la fin de leur relaxation, mesuré du début du QRS à la fin de l'onde T correspond l'ensemble de la dépolarisation et de la repolarisation cardiaque.

Segment ST : (durée de stimulation complété des ventricules) le segment ST correspond à la phase pendant laquelle les cellules ventriculaires sont toutes dépolarisés, le segment est alors isoélectrique, le segment ST est la partie du tracé comprise entre la fin du complexe QRS et le début de l'onde T, le segment ST normal peut être légèrement décalé, vers le haut au repos, ou vers le bas à l'effort [5].

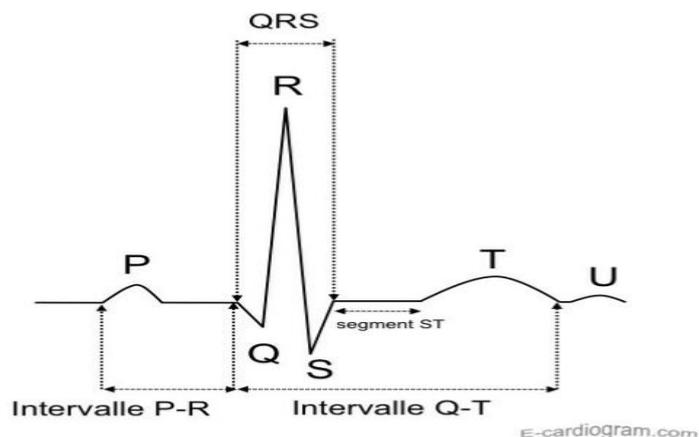


Figure 1.8 : Intervalles de l'ECG normal, Incluant l'onde U

➤ **Calibrage de l'enregistrement ECG**

Avec la hauteur d'onde **P**, des complexes **QRS** et d'onde **T**, nous pouvons fournir une quantité limitée d'informations, mais nous devons nous concentrer sur l'étalonnage des performances. La tête d'enregistrement de 1 cm (2 grandes carreaux) doit être remplie d'un signal standard de 1 mV [7].

➤ **Calcul de la fréquence cardiaque (FC)**

Les changements d'activité électrique sont enregistrés en plaçant une ligne sur une bande de papier millimétrique. Le taux d'enregistrement ECG standard est de 25 mm / s. Chaque grand carreau (5 mm) représente 0,2 seconde ou 200 millièmes de seconde (200 ms). En conséquence, il y a 5 grandes carreaux par seconde et 300 par minute [7].

➤ **La méthode des 300**

Un petit carré correspond à 0,04 seconde et 60 secondes font 1500 petits carrés. Chaque bloc de 5 petits carrés est marqué en trait gras. On repère une onde R coïncidant avec un trait gras qui deviendra un multiple des autres traits gras suivants. Entre 2 traits gras il y a 5 petits carreaux. Donc l'onde R qui coïncide au trait gras, départ des 1500, le prochain, on aura $1500/5$ soit 300, au suivant $1500/10$ soit 150 puis 100, 75, 60, L'onde R suivante permet d'indiquer la fréquence cardiaque en relation avec le trait gras correspondant. C'est la méthode la plus rapide et la plus utilisée [5].

1.3 Technique de l'enregistrement

1.3.1 L'enregistrement

Le signal électrique peut être enregistré soit en externe à l'aide d'électrodes se faire soit par voie externe à l'aide d'électrodes placées à la surface du corps, en interne en enregistrant des impulsions électriques à la surface du cœur, soit directement en insérant la pointe de la sonde cavitaire.

L'enregistrement de l'électrocardiographie se fait selon 2 plans :

- ✓ Frontal par les électrodes périphériques (pôles terminaux).
- ✓ Horizontal par les électrodes précordiales.

1.3.2 Les Dérivations de l'ECG

La dérivation électrocardiographie est déterminée par deux d'observation de l'activité électrique du cœur dans lesquelles la différence de potentiel électrique est mesurée, En générale, les appareils électrocardiographiques peuvent enregistrer de nombreuses variations de potentiel en même temps, en fonction de l'emplacement et du nombre d'électrodes réparties sur la poitrine et les extrémités, chaque mesure de ces potentiels correspondant à la dérivation se compose d'un réseau cohérent de dérivations et les électrodes sont situées pour explorer presque tout le champ électrique du cœur.

Dans l'ECG, le fil correspond à une ligne droite reliant deux points d'observation de l'activité électrique du cœur et à travers laquelle on mesure une différence de potentiel électrique, Nous appelons le circuit défini par deux paires d'électrodes, dans lesquelles les électrodes sont placées à des emplacements bien définis du cœur, Si nous ne mesurons le vecteur cardiaque que dans une seule direction, nous ne pouvons pas le caractériser complètement. Il est donc important de disposer d'un critère de positionnement d'électrodes pour l'évaluation clinique du signal ECG.

Les dérivations standards forment un système de 12 dérivations complémentaires :

- Trois dérivations unipolaires bipolaires.
- Trois dérivations unipolaires.
- Six dérivations précordiales.

a Dérivation bipolaire

La disposition des électrodes dans les trois conducteur dipôles standard suit une convention établie par Einthoven, le système Einthoven est basé sur les hypothèses suivant :

- L'activité électrique du cœur est équivalente à celle d'un dipôle, dont la direction, la direction et le moment diffèrent pendant la systole mais dont l'origine reste constante.
- Les membres sont de simples conducteurs linéaires.
- Le corps constituée un milieu résistif homogène.

La dérivation à trois bipolaire, I, II, III s'appelle dérivation d'Einthoven

- **I** : enregistre les différences de potentiel électrique entre le poignet droit et le poignet gauche.
- **II** : enregistre les différences de potentiel électrique entre le poignet droit et la jambe gauche.

- **III** : enregistre les différences de potentiel électrique entre le poignet gauche et la jambe gauche.

Déterminé par Einthoven et obtenu à partir des potentiels DI, DII, DIII, obtenus en commutant des électrodes placées sur le bras droit, le bras gauche et la jambe gauche de la manière suivante :

$$\mathbf{DI = VL - VR}$$

$$\mathbf{DII = VF - VR}$$

$$\mathbf{DIII = VF - VL}$$

VR correspond au potentiel au bras droit, VL correspond au potentiel au bras gauche et VF correspond au potentiel dans la jambe gauche. On a la relation suivante entre la dérivation :

$$\mathbf{DII = DI + DIII}$$

b Dérivation unipolaire des membres

Les dérivations unipolaires des membres permettent l'étude de l'activité électrique du cœur au niveau du plan antérieur (frontal), et ont été identifiées par Wilson, et ces dérivations permettent de mesurer la tension entre un point de référence et le bras droit, le bras gauche et la jambe gauche respectivement. Le point de référence obtenu en faisant la moyenne des signaux apparaissent sur les deux organes d'autre qui ne sont pas sur notre note.

Plus les dérivations unipolaires aVR, aVL, aVF dite de Wilson :

- **aVL** : pour l'avant-bras gauche.
- **aVR** : pour l'avant-bras droit.
- **aVF** : pour la jambe gauche.

c Dérivation unipolaires précordiales

Pour mesurer les potentiels quasi cardiaque , Wilson a introduit les sondes à partir du plan horizontal V1, V2, V3, V4, V5 et V6 .et ces six dérivations sont situés sur le côté gauche de la cage thoracique , placés sur le poitrine et sont identifiés par la lettre V suivie de leur numéro emplacement .les potentiels des électrodes d'exploration sont pris par rapport à la moyenne des potentiels VL à VR et VF , Six points, définis par Wilson, permettent d'obtenir les dérivations V1 – a V6, les dérivation précordiales mesurent l'activité électrique du cœur dans le plan horizontal et sont placées près du cœur.

- **V1** : électrode placée à droit du sternum.

- **V2** : électrode placée à gauche du sternum.
- **V3** : électrode placée à mi-distance entre V2 et le mamelon.
- **V4** : électrode placée sous le mamelon.
- **V5** : électrode placée à mi-distance entre le mamelon et la ligne axillaire.
- **V6** : électrode placée sur la ligne axillaire [8].

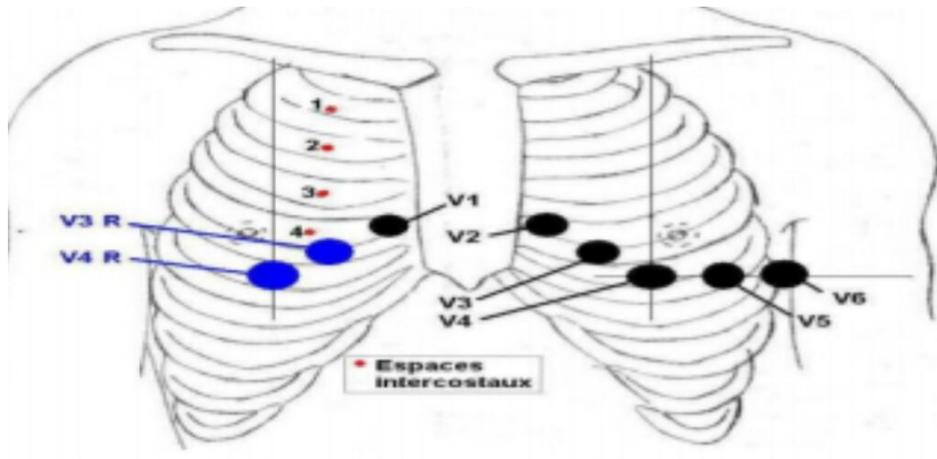


Figure 1.9 : dérivations précordiales

1.4 Analyse de l'ECG

L'analyse du rythme et le diagnostic automatique des troubles du rythme représentent un champ spécifique qui complète l'analyse de la circonférence des ondes, et la fréquence est liée au moment de l'origine de l'activité électrique du cœur et du rythme, ou non à sa propagation, quand on parle de fréquence cardiaque, on parle à la fois de l'association de la genèse de l'activité électrique du cœur et de la régularité de sa propagation. Ainsi, on parle de rythme sinusal régulier quand il est régulier, l'intervalle R_R est presque constant sur toute trace avec des complexes QRS similaires. savoir que l'activité électrique est créée par le nœud sinusal [9].

1.4.1 Les phases de l'ECG

Les différentes ondes constituantes de l'ECG sont en relation directe avec la différente phase du cycle cardiaque :

- Une phase de contraction auriculaire.
- Une phase de contraction ventriculaire.
- Un relâchement total du cœur.
- Les contractions des oreillettes pendant le passage de l'onde électrique laissent sur l'ECG, le tracé d'une petite déflexion positive : l'onde P.

- La contraction représente la pause de l'impulsion électrique au niveau du nœud auriculo-Ventriculaire (AV).
- Une succession de trois ondes (QRS) correspond à la phase de contraction et de décontraction du ventricule. Le complexe QRS constitué de ces trois ondes est visible sur l'ECG :
 - La première onde Q est dirigée vers le bas.
 - La deuxième onde R dirigée vers le haut est de grande amplitude.
 - L'onde S est dirigée vers le bas.
- La repolarisation des fibres musculaires est représentée par l'onde T. c'est une période de l'inactivité cardiaque, l'étude de segment ST reste très importante pour l'identification de certaines pathologies.

1.4.2 Les Caractéristiques de l'ECG

L'électrocardiogramme représente l'activité électrique de cœur humaine, il se compose 5 ondes, et la détection des ondes P et T et des complexes ORS est une ces pics, il devient très facile d'extraire un autre paramètre important du signal tel que la fréquence cardiaque, la durée du segment ST, etc.

La forme de ces ondes est instable, elle peut varier d'un sujet à l'autre, et même au sein d'un même individu, elle peut varier d'un sujet à l'autre ; cela rend l'étape de détection automatique des pics très difficile, et le premier objectif du médecin est également de

de 1.5 mV et durée maximale 70 ms, il se trouve entre les deux pic négative déterminer la fréquence cardiaque et de reconnaître les ondes P, Q, R, S et T,z.

- **L'onde R** : c'est le pic le plus élevé de l'électrocardiogramme et d'amplitude maximale QS dans le cas d'un fonctionnement cardiaques normale.
- **L'onde QRS** : il commence au début de l'onde Q et se termine à l'extrémité de l'onde S .elle représente la durée de dépolarisation ventricule précédant l'effet mécanique de contraction et il posséd la plus grande amplitude de l'ECG il constitue de trois onde consécutives, l'onde Q qui est négative et l'onde R qui est positive dans un ECG normal et l'onde S qui est négative. sa durée normale est comprise entre 85 et 95 ms. si nous nous and pic et d'un autre creux. Cette série est généralement considérée ensemble, et elle s'appelle la vague QRS. Vous remarquez que cela place les vagues dans l'ordre alphabétique. L'onde QRS est parfois appelée complexe QRS et représente la dépolarisation des ventricules. Cela conduit rapidement à la contraction des ventricules et l'éjection du sang hors du cœur et dans les grandes artères sortant du cœur.
- **L'onde T** : c'est la déviation positive après chaque complexe QRS elle représente la repolarisation ventriculaire. C'est la récupération électrique ou le retour à un

état de repos des ventricules, l'onde T normale a une amplitude plus faible et elle est plus étendue que le complexe QRS parce que la repolarisation se fait plus lentement que la dépolarisation.

- **L'intervalle PR** : c'est la durée entre le début de l'onde P et le début de Q ou de R, sa durée est comprise entre 120 ms et 200 ms, il correspond à la pause qui s'écoule entre le début de l'excitation auriculaire et le début de l'excitation ventriculaire. L'intervalle P-Q correspond au temps requis pour permettre à l'influx électrique de se propager à travers les oreillettes, le nœud auriculo-ventriculaire et le reste des fibres du système de conduction.
- **L'intervalle RR** : c'est la durée entre deux pics R successifs d'un cycle cardiaque. Les moyennes des intervalles RR détermine le rythme cardiaque (70 battements par minute en moyenne) et son écart-type détermine la dispersion ou l'irrégularité du rythme cardiaque.

1.4.3 La fréquence cardiaque

La fréquence cardiaque est le nombre de cycles cardiaques par unité de Temps (par minute). Il est très rapide chez le nouveau-né, rapide chez l'enfant et légèrement plus lente chez la personne âgée, et c'est un concept quantitatif qui correspond au nombre de cycles cardiaques par unité de temps, et varie selon plusieurs critères physiologiques comme l'âge, La fréquence cardiaque est également le nombre de contractions ventriculaire par unité de temps qui est calculée sur l'ECG à partir du nombre de complexes QRS [10].

1.5 Conclusion

Ce chapitre est une générique basé sur les signaux cardiaque issus d'une description détaillée de l'anatomie cardiaque humaine et du fonctionnement du système cardiovasculaire, Nous présentons les bases de l'électrocardiogramme, l'électrocardiographie est l'étude différences d'enregistrement de l'activité électrique des cellules cardiaque , Cette technique permet de surveiller l'activité électrique du muscle cardiaque de l'extérieur du corps , l'électrocardiogrammes est une technique largement utilisé pour le bon diagnostic du muscle cardiaque le nombre de maladies cardiovasculaires , s'est-il fait bien son travail et fournit un signal fiable avec lequel nous pouvons appliquer des méthodes de traitement du signal appropriés.

Chapitre 2 Technique de l'apprentissage machine

2.1 Introduction

Le domaine médicale a actuellement besoin de nouvelles techniques et technologiques, Afin d'évaluer l'information d'une manière objective, l'intelligence artificielle est devenue un sujet brulent dans les médias et les revues scientifique, et l'intelligence artificielle est une discipline scientifiques qui a ses racines principalement dans la psychologie, les mathématiques et de l'informatique et qui vise à comprendre et à développer les systèmes qui présentent des « propriétés de l'intelligence ».

Le domaine de l'apprentissage automatique pour étudier l'intelligence artificielle, par conséquent, l'apprentissage machine est un outil très puissant qui permet l'effet de multiple actions telles que la catégorisation des données , faire apprendre au programme des expérience ou jusqu' à présent pour créer un programme évolutif qui continue de s'améliore , l'apprentissage automatique doit etre utilisé comme un outil car tous les problèmes ne nécessitent pas un programme apprentissage automatique complexes.

Dans ce chapitre, nous présenterons une méthode de classification du mécanisme des réseaux de neurones convolution (CNN) et de la relation entre l'intelligence et apprentissage machine dans l'ECG.

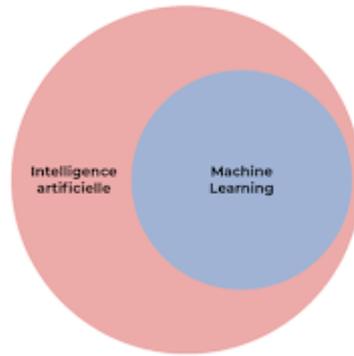


Figure 2.1 : la relation entre l'apprentissage machine et l'intelligence artificielle

2.2 L'intelligence artificielle

L'intelligence artificielle peut être définie comme une branche de l'informatique qui comprend le fait de donner à une machine une partie de l'intelligence humaine qui est la production de programmes et de machines intelligentes dans un ordinateur. Programme classique que nous programmons à chaque pas que nous programmons à chaque pas que vous faites. Dans les ordinateurs et les systèmes, nous dépendons de la capacité d'un ordinateur automatisé ou d'un esprit de machine à apprendre de l'environnement qui l'entoure sans avoir à le programmer à l'avance et littéralement, à travers des algorithmes. Données : Les données sont de la formation et de l'éducation, mais en ce qui concerne les algorithmes, c'est un moyen de performance par lequel un esprit mécanique apprend, et c'est un ensemble de procédures ou d'étapes qu'un programmeur peut programmer pour obtenir ou accéder à une tâche spécifique. Autrement dit, nous essayons de prendre sa propre décision en transférant la capacité à inférer, déduire, percevoir et comprendre en plus de la capacité à apprendre et à résoudre des problèmes, nous trouvons des exemples d'apprentissage automatique un modèle pour discrimination visuelle, car la machine distingue plusieurs images différentes et cela se fait en deux étapes, la première consiste à construire le modèle, la seconde à enseigner et à construire un modèle qui donne des informations sur le problème et de nombreuses images de la machine pour créer ses propres concepts, puis la machine généralise cela. Pour chaque modèle est proche ou similaire aux concepts que la machine a dérivés, exemple: un dispositif pour distinguer une image de chat afin qu'elle puisse distinguer toute nouvelle image contenant un chat ou ne pas qu'elle. L'intelligence artificielle est un roman de science-fiction sur des hypothèses et les illusions que de nombreux scientifiques ont poursuivies.

2.2.1 Le but de l'intelligence artificielle

- L'intelligence artificielle vise à concevoir des systèmes capables de reproduire le comportement humain dans leurs activités logiques.
- L'intelligence artificielle est un groupe qui vise à modéliser l'intelligence comme un phénomène (en plus de la physique, de la chimie ou de la biologie) qui vise à modéliser d'autres phénomènes.
- L'intelligence artificielle ne vise pas à remplacer les humains, mais plutôt à les aider pour qu'ils puissent se concentrer sur des tâches plus créatives ou amusantes.

2.2.2 Le Domaine de l'intelligence artificielle

- ✚ Apprentissage créé des programmes qui génèrent leurs propres connaissances en se modifiant à partir de leurs succès et leurs erreurs.
- ✚ Les domaines de l'intelligence artificielle (IA) visent à donner à un système informatique la capacité de penser comme une capacité humaine, il y a donc concurrence dans la réflexion entre les humaines et l'intelligence artificielle, en particulier dans la modélisation et comme le clonage.

2.2.3 L'utilisation de l'intelligence artificielle pour classifier l'ECG

Plusieurs études récentes révèlent que l'intelligence artificielle peut prédire si une personne mourra dans l'année en regardant les résultats, les tests cardiaques et même lorsque les lectures semblent normales, les médecins est Une étude sur l'intelligence artificielle a été menée sur 1.77 million d'électrocardiogrammes (ECG), sur environ 400000 personnes .Pour prédire s'il risque de mourir d'ici un an un électrocardiogramme (ECG) enregistre l'activité électrique du cœur est ses changements de modèle dans des conditions telles que les crises cardiaques et la fibrillation auriculaire, Les études ont formé deux versions de l'intelligence artificielle, la premier sur l'algorithme de donnée ECG qui mesure l'effort au fil du temps et la seconde, les données ECG ont été alimentées en conjonction avec l'âge et le sexe du patient, et dans cette expérience, une échelle appelé AUC a été utilisé qui mesure dans laquelle le modèle distingue deux group de personne.

L'intelligence artificielle prédit avec précision le risque de décès, même pour les personnes que les médecins considèrent comme ayant un ECG normal.

Nous trouvons beaucoup technique d'application approche de l'intelligence artificielle et en personnalisé les réseaux de neurones, pour l'analyse automatique de l'ECG.

Le travail de Mark et al montre les techniques de réseaux de neurones et spécialement, le réseau multi layer perception et la carte Auto-Organisatrices sont utilisés à des fins de classification. Une autre approche, comme l'approche bayésienne et heuristique et des modèles de Markov. Ils ont aussi une expérience pour la classification des signaux ECG, ART a été appliquée à une variété de domaines allant des applications médicales comme la classification des modèles ECG. Au traitement sémantique des données Barro et al ont utilisé la théorie MART pour classer les signaux ECG. Mettre en œuvre des résultats montre que ce classifieur peut discriminer les battements normaux des battements ventriculaires avec une précision de plus de 90 %.

Olmez, Son travail à classer quatre types de battement comme les battements normaux, le bloc de branche gauche, la contraction ventriculaires prématurées, et les battements de type P en utilisant un réseau de neurones RCE (Restricted Coulomb Energy) et les algorithmes génétiques. Il a obtenu une précision de classification de plus de 94%. De nombreuses recherches et études confirment que les performances des réseaux de neurones artificiels (RNA) permettent de détecter et d'identifier les ECG anormaux.

Les neurones d'analyse du signal ECG présentent de nombreux avantages des techniques classiques, et c'est ici que sont apparus les réseaux de neurones qui ont remplacé les humains afin qu'ils puissent se concentrer sur des tâches plus créatives et amusantes. Différents domaines d'application des réseaux de neurones artificiels dans la distinction et la classification des maladies.

Les RNA sont utilisés pour distinguer les grandes complexes QRS ainsi que pour diagnostiquer la tachycardie ventriculaires (TV) et la tachycardie supraventriculaire (SVT), Il détecte une fibrillation ventriculaire précoce (V) critique en cas de risque de mort subite, il y a plusieurs techniques utilisant les RNA ont été développées pour établir des critères permettant la différenciation de la vit parmi des battements normaux. AL-Fahoum et al [1221] ont utilisé des transformations en ondelette (continue et discrète) dans la phase de l'extraction des paramètres. Le vecteur de paramètres extrait par ondelette est utilisé ensuite comme entrée d'un réseau de neurones de type KBF dans le but de la détection automatique des arythmies.

2.3 Représentation de l'apprentissage machine

2.3.1 Apprentissage machine

Les études d'apprentissage machine utilisent des techniques pour donner à la machine la capacité d'apprendre à partir d'expériences passées, Il s'agit d'un des domaines formant l'intelligence artificielle. En informatique "traditionnelle", on résout un problème en donnant explicitement à la machine les instructions qu'elle doit exécuter, pour ce qui est de l'apprentissage machine, on cherche plutôt à donner à la machine la capacité d'apprendre à résoudre ce problème à

partir d'exemples d'entrées/sorties formant l'ensemble d'entraînement, en plus de la possibilité d'apprendre à partir des données c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et l'implémentation de programmes pour chacune plus. Et Plus vous avez d'exemples d'un problème particulier, meilleurs sont vos résultats. Et à partir de là, nous disons Plus vous avez d'exemples d'un problème particulier, meilleurs sont vos résultats [11].

Il Ya une relation avec l'apprentissage artificiel (machine Learning en anglais) et la classification automatique, est un thème d'une grande importance et présente de vastes champs d'application qui ne cessent de croître d'année en année. Il est clair qu'une brève description de ce domaine permet de mettre en évidence son importance dans les différents champs d'application [12].

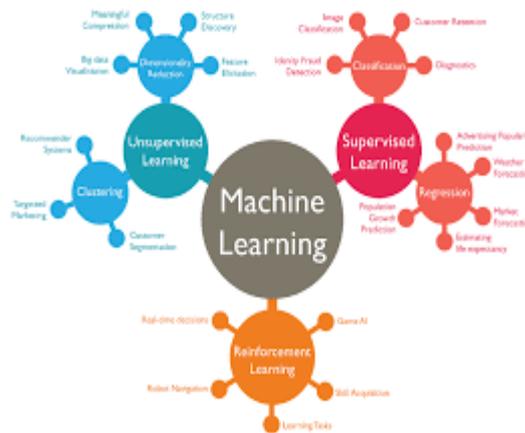


Figure 2.2 : L'apprentissage automatique

2.3.2 Les Types d'apprentissage Machine

Dans le domaine de l'intelligence artificielles (IA), il existe de nombreuses approches de l'apprentissage machine (AM). Celles –ci peuvent être divisées en plusieurs sous –de catégories, dont deux sont : l'apprentissage machine supervisé et l'apprentissage machine non supervisée.

a Apprentissage supervisé

Apprentissage supervisé (en anglais : Supervised Learning) est le paradigme d'apprentissage le plus populaire en Machine Learning et en Deep Learning. Pour que lui de nombreuses applications et est beaucoup plus souvent utilisé que l'apprentissage non supervisé. Le système apprend à classer selon un modèle de classification ou de

classement et c'est à travers les instructions, les données et exemples donnés précédemment l'un des exemples les plus marquants d'apprentissage supervisé est la traduction automatique à l'aide de l'IA. Doit préalablement étiqueter des exemples. Le processus se passe en deux phases. Lors de la première phase (hors ligne, dite d'apprentissage), il s'agit de déterminer un modèle à partir des données étiquetées. La seconde phase (en ligne, dite de test) consiste à prédire l'étiquette d'une nouvelle donnée, connaissant le modèle préalablement appris. Parfois il est préférable d'associer une donnée non pas à une classe unique, mais une probabilité d'appartenance à chacune des classes prédéterminées (on parle alors d'apprentissage supervisé probabiliste). Ex. : L'analyse discriminante linéaire ou les SVM en sont des exemples typiques [13].

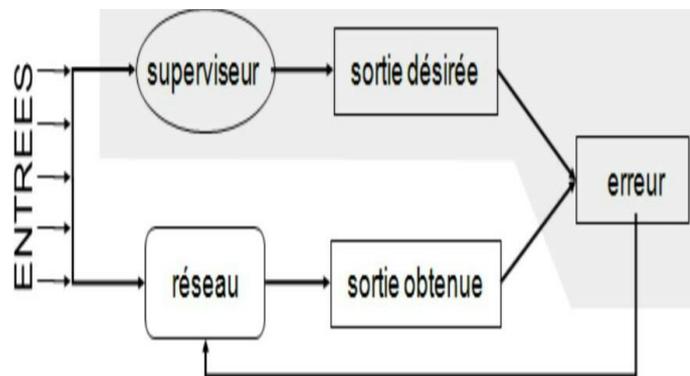


Figure 2 .3 : l'apprentissage supervisé

b Apprentissage non supervisé

On l'appelle apprentissage non supervisé car, contrairement à l'apprentissage supervisé ci-dessus, puisqu'il n'y a pas de réponse correcte ni d'enseignant. Les algorithmes sont laissés à leurs propres données. L'apprentissage non supervisé contient deux catégories d'algorithmes : Algorithmes de regroupement et d'association. Lorsque le système ou l'opérateur ne contient que des exemples, sans étiquettes et que le nombre et la nature des classes n'ont pas été spécifiés à l'avance. Nous parlons ici d'apprentissage non supervisé, c'est-à-dire que l'algorithme doit lui-même découvrir dans une certaine mesure la structure des données masquées. L'algorithme d'apprentissage non supervisé. Le système doit ici — dans l'espace de description (la somme des données) — cibler les données selon leurs attributs disponibles, pour les classer en groupes homogènes d'exemples. . La similarité est généralement calculée selon une fonction de distance entre paires d'exemples. C'est ensuite à l'opérateur d'associer ou déduire du sens pour chaque groupe et pour les motifs (patterns en anglais) d'apparition de groupes, ou de groupes de groupes, dans leur « espace ». Divers outils mathématiques

et logiciels peuvent l'aider. On parle aussi d'analyse des données en régression (ajustement d'un modèle par une procédure de type moindres carrés ou autre optimisation d'une fonction de coût) [14].

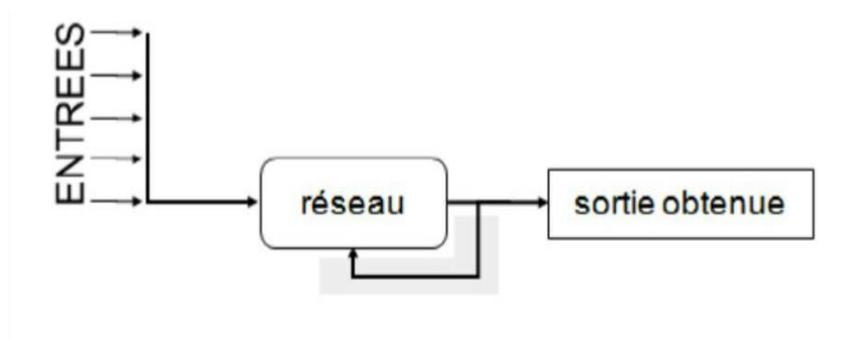


Figure 2.4 : Le mode faciles non-supervisé

***c* Apprentissage par renforcement**

Le renforcement Learning ou apprentissage par renforcement est l'une des méthodes d'apprentissage de la machine Learning pour faire simple, cette méthode consiste à laisser l'algorithme apprendre de ses propres erreurs.

2.4 Architecture de réseaux neurones profonds

Réseau de neurones avec plusieurs couches de neurones artificiel. La sortie d'une couche (la sortie d'un neurone) est introduite en entrée dans la couche suivante (un autre neurone) peut avoir des poids arbitraires.

Deep Learning est un ensemble de technique d'apprentissage automatique qui est permet des avancées importantes en intelligence artificielle dans les dernières années.

2.4.1 Réseaux de neurones

Les réseaux de neurones sont considères comme l'un des éléments de base et une composante importante dans les domaines de l'intelligence artificielle , et ils sont un ensemble d'algorithmes qui connaissent et découvrent leur méthode de travail à partir de neurones biologiques .Ces réseaux ont été développé pour résoudre les problèmes de reconnaissance du contrôle des formes , des mots , de la décision , de la mémorisation .Les réseaux de neurones sont concernés par le stockage des connaissances et leur mise à disposition.

Les réseaux de neurones les plus utilisées dans la littérature, comme perceptron monocouche et multicouche.

a Application de réseaux de neurones :

Ces réseaux jouent le rôle d'accomplir des tâches pour l'homme, mais difficile pour une machine :

Espace – avion pilote automatique, détection de crash d'avion, voiture, système de direction automatique militaire – point et point d'arme, suivie cible, reconnaissance, identification du signal / image.

Prédiction de séquence de code électronique, disposition des puces IC, analyse des défaillances de puces, machine, vision, installation sonore

Parole – reconnaissance vocale, classification vocale.

Compression d'images et de données de télécommunication, services d'informations automatisés, véritable transmission de la langue parlée.

Transport –diagnostiquer le système de freinage du camion, la planification du véhicule et les systèmes de direction.

Prédiction des séries chronologique - ANNS est utilisé pour faire des prédictions sur les stocks et les normales.

Les réseaux neuronaux de traitement du signal audio peuvent être formés pour traiter et filtrer le signal audio.

ANNS est souvent utilisé pour prendre des décisions de direction physique du véhicule, contrôle.

2.5 Réseaux de neurones conventionnels

Un réseau neurones convulsifs (CNN : convolutional Neural Network) est un type d'apprentissage profond (réseaux de neurones profonds) et est un algorithme de Deep Learning et est similaire à un réseau de perceptron multicouche (MLP). Ce réseau diffère dans ce qu'il apprend, comment il est construit et ce qu'il vise à faire . Habituellement, les réseaux CNN sont utilisés dans les applications de vision par ordinateur et l'analyse de visualisation visuelle, car ils sont caractérisés par la présence d'une ou plusieurs couches cachées qui extraient les caractéristiques trouvées dans les images ou les clips vidéo, et une couche entièrement connectée pour produire la sortie requise [15].

2.5.1 Architecture de Réseau de neurones conventionnel

Les réseaux de neurones convolutifs visent à apprendre en profondeur sans surveillance. Cette technique d'apprentissage a été le premier modèle. Un modèle artificiel utilisé pour la reconnaissance afin d'obtenir de grande performance dans diverses tâches.

Le but de la création de ce réseau est de traiter différentes données qui se présentent sous la forme de tableaux numérique, par exemple, une image couleur constituée de trois matrices couches bidimensionnelles, qui contient également différentes intensités de pixels dans ces trois couches colorées.

Les réseaux de neurones convolutifs reposent sur une perception multicouche (MLP) inspirée du comportement du cortex visuel du vertébré.

MLP est difficile à gérer de grandes images malgré son efficacité et sa capacité à traiter les images .Cela est dû à l'intersection exponentielle du nombre de connexion avec la taille de l'image .Par exemple si nous prenons une image d'une taille de 32 x 32 x 3 (52 largeurs 32 hauteur , 3 canaux de couleur) ce sera pour des neurones uniques entièrement connecté dans la première couche cachée de 3072 entrées MLP (32 x 32 x 3) , de sorte que l'image 200x200 entrainera le traitement de 120000 entrées par neurones , CNN est caractérisé par [15] :

- 1. Volumes 3D de neurones :** Des volumes tridimensionnels de neurones, la couche neurale n'est plus juste une surface. Ça devient volume à profondeur. Si on considère un seul champ récepteur du CNN, les n neurones associée (sur la profondeur) forment l'équivalent de première couche d'un MLP.
- 2. Connectivité locale :** Grace au champ récepteur qui limite le nombre d'entrée de neurones, tout en conservant l'architecture MLP, les CNN assurent ainsi que les « filtre » produisent la réponse la plus forte à un motif d'entrée spatialement localisé, ce qui conduit à une représentation parcimonieuse de l'entrée. Une telle représentation occupe moins d'espace en mémoire. De plus, le nombre de paramètre à estimer étant réduit, leur estimation (statistique) est plus robuste pour un volume de données fixé (comparé à un MLP).
- 3. poids partagés :** Dans les CNN, les paramètres de filtrage d'un neurones (pour un champs récepteur donnée) sont identique pour tous les autres neurones d'un même noyau (traitant tous les autres champs récepteurs de l'image) .Ce paramétrage (vecteur de poids et biais) est défini dans une « carte de fonction » .Cela signifie que tous les neurones dans une couche de convolution donnée détecter exactement la même caractéristique .En multipliant les champs récepteurs , il devient possible de détecter des éléments indépendamment de

leur position dans le champ visuel .ce qui induit une propriété d'invariance par translation [15].

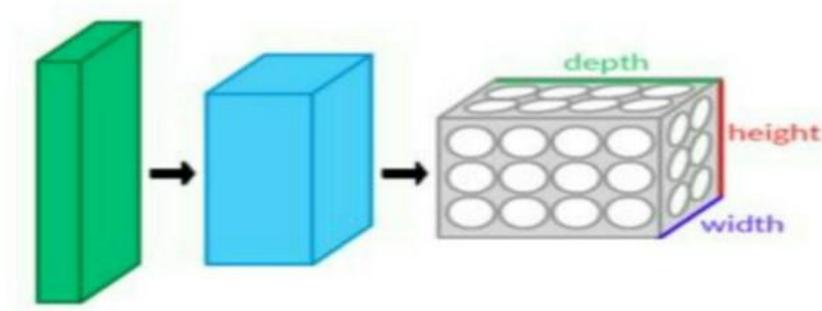


Figure 2.5 : Une couche du CNN en 3 dimensions (Vert =Volume d'entrée, bleu = volume du champ récepteur, gris = couche de CNN, cercles = neurones artificiels indépendants)

Ces caractéristiques permettent un meilleur apprentissage sur les problèmes de vision d'un réseau de neurones convolutifs, ainsi qu'une réduction du nombre de paramètre libre pour apprentissage meilleur et plus grand et ce dont la mémoire a besoin pour faire fonctionner le réseau, car la réduction de l'espace mémoire permet un apprentissage plus grand et plus puissant par rapport aux réseaux.

Un réseau de neurones à convolution est formé par un empilement de couches de traitement indépendantes :

- La couche de Convolution (CONV) qui traite les données d'un champ récepteur.
- La couche de correction (ReLU), souvent appelée par abus « ReLU » en référence à la fonction d'activation (Unité de rectification linéaire).
- La couche de Pooling (POOL), qui permet de compresser l'information en réduisant la taille de l'image intermédiaire (souvent par sous-échantillonnage).
- La couche « entièrement connecté » (FC), qui est une couche de type perceptron.
- La couche de perte (LOSS) [15].

a Couche de convolution (CONV)

Lors de la soumission d'une nouvelle image, CNN ne sait pas si les caractéristiques seront ailleurs ou dans une image. C'est pourquoi il cherche à la recherche dans n'importe quelle situation. En calculant dans n'importe quelle image si la caractéristique est présente, Le processus convolutif est processus mathématique que nous utilisons pour y parvenir, dont les réseaux de neurones convolutif tirent leur nom. La couche convolutif

est le composant principale des réseaux de neurones convolutifs, et elle forme toujours au moins la première couche. Cette couche vise à tirer un avantage des images.

Une série de filtre sont passés, ou des billes de torsion, et nous les obtenons pour chaque paire (image, filtre), carte d'activation ou carte de fonction qui nous indique où se trouvent les caractéristiques de l'image ayant la valeur la plus élevée.

Enfin, les valeurs des dernières cartes de caractéristique sont liées dans un vecteur, le vecteur déterminé la sortie du premier bloc et l'entrée du second, trois hyper paramètres sont utilisés pour déterminer la taille de la couche de chaîne.

1. **Profondeur de la couche** : nombre de noyaux de convolution (ou nombre de neurones associées à un même champ récepteur).
2. **Le pas** : contrôle le chevauchement des champs récepteur. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie sera grand.
3. **La marge (à 0) ou 'zéro padding'** : cette image permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée [15].

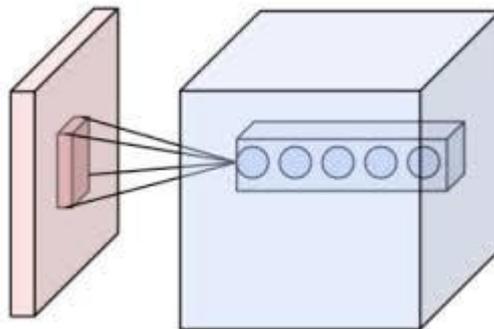


Figure 2.6 : Ensemble de neurones (cercle) créant la profondeur d'une couche de convolution (bleu). Ils sont liés à un même champ récepteur

b Couche de Pooling (POOL)

Un autre concept important des CNN est le Pooling, ce qui est une forme de sous-échantillonnage de l'image. L'image d'entrée est découpée en une série de rectangles de n pixels de côté ne se chevauchant pas (Pooling). Chaque rectangle peut être vu comme une tuile. Le signal en sortie de tuile est défini en fonction des valeurs prises par différents pixels de la tuile.

Le Pooling réduit la taille spatiale d'une image intermédiaire, réduisant ainsi la quantité de paramètres et de calculs dans le réseau. Il est donc fréquent d'insérer périodiquement

une couche de Pooling entre deux couches convolutés successive d'un architecture CNN pour contrôler l'overfitting (sur apprentissage).

La couche de Pooling fonctionne indépendamment sur chaque tranche de profondeur de l'entrée et la redimensionne uniquement au niveau de la surface. La forme la plus courante est une couche de mise en commun avec des tuiles de taille 2 x 2 (largeur/hauteur) et comme valeur de sortie la valeur maximale en entrée. On parle dans ce cas de « Max-Pool 2 x 2 ». Il est possible d'utiliser d'autres fonctions de Pooling que le maximum. On peut utiliser un « average Pooling » (la sortie est la moyenne des valeurs du patch d'entrée), du « L2-norm Pooling ».

Il est aussi possible d'éviter la couche de Pooling mais cela implique un risque sur-apprentissage plus important [15].

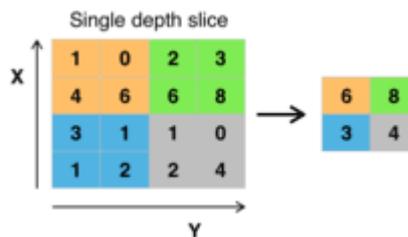


Figure2.7 : Pooling avec un filtre 2 x 2 et pas de 2

c Couches de correction (RELU)

Une opération supplémentaire appelée ReLU est utilisé après chaque opération de convolution, ReLU signifie unité linéaire rectifiée est une opération non linéaire .sa formule de sortie est : $Sortie = \text{Max}(\text{Zéro}, \text{Entrée})$.

ReLU est une opération appliquer par pixel et remplacer toutes les valeurs de pixels négatifs dans la Feature map par zéro. Le but de ReLU est d'introduire la notion de non-linéarité [15].

d Couche entièrement connectée (FC) ou Fully Connected

La couche Fully Connected est configurée exactement comme son nom l'indique : elle est entièrement connectée à la sortie de la couche précédente. Une couche entièrement connectée prend tous les neurones de la couche précédente (qu'elle soit entièrement connectée, en pool ou convolutionnelle) et la connecte à chaque neurone qu'elle possède [15].

e Couche de perte (LOSS)

La couche de perte est la dernière couche dans le réseau. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. La fonction « Softmax » permet de calculer la distribution de probabilités sur les classes de sortie [15].

2.5.2 Exemples de modèles de CNN

La forme la plus courante et la plus connue de la structure CNN combine quelques couches Conv-ReLu, suivies des couches d'assemblage (couche pool), et cette étape est ramenée dans le but de réduire la rentrée dans une zone d'une seule taille suffisamment petite et on la connaît et on place des couches entièrement connectées FC liées vers la sortie. Voici quelques architectures communes CNN qui suivent ce modèle :

- INPUT -> CONV -> RELU -> FC
- INPUT -> [CONV -> RELU -> POOL] x 2 -> FC -> RELU -> FC Ici, il y a une couche de CONV unique entre chaque couche POOL
- INPUT -> [CONV -> RELU -> CONV -> RELU -> POOL] x 3 -> [FC -> RELU] x 2 -> FC Ici il y a deux couches CONV empilées avant chaque couche POOL

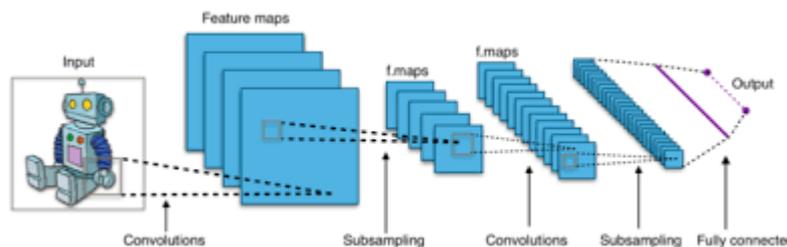


Figure 2.8 : Structure de CNN

L'assemblage de couche de CONV permet l'utilisation de petit filtre de collecte (plutôt qu'un grand filtre) pour une extension de traitement plus puissante avec moins de paramètres, Le modèle a formé le réseau de neurones convolutifs (CNN) à partir d'un ensemble d'impulsions électriques du cœur, puis a testé sa capacité à classer et à distinguer le rythme cardiaque d'une personne en bonne santé et le rythme cardiaque d'une personne souffrant d'insuffisance cardiaque congestive. Le modèle développé a montré des niveaux élevés de précision dans les performances lors de la détection d'une maladie cardiaque congestive [15].

2.6 Choix des paramètres

Les CNN utilisent beaucoup de paramètres, tous les aspects des CNN ne sont pas faciles à comprendre et à apprendre comme nous l'avons vu jusqu'à présent, et pour cela, il

existe une longue liste de paramètre qui peut être définis manuellement pour activer le fonctionnement du CNN et obtenir de meilleurs résultats.

2.6.1 Nombre de filtres

Dans les applications pratique, CNN apprend les valeurs de ces filtres pendant le processus de formation et sa limite. Comme le nombre de paramètre tels que le nombre de filtre est déterminé avant de commencer le processus de formation. Plus il Ya des filtres, plus le nombre de filtre augmente, l'extraction des propriétés d'image augmente et la capacité du réseau à extraire des caractéristiques et des caractéristiques augmente.

Classification des images la taille des images intermédiaires contraste avec la profondeur de traitement, les couches proches de l'entrée ont tendance et préfèrent avoir un filtre plus bas, tandis que les couches proches de la sortie ont tendance à en avoir plus pour atteindre l'équilibre dans chaque couche, la sortie du nombre de propriétés et le nombre de pixels de traitement sont choisis pour être presque constants à travers les couches.

2.6.2 Forme du filtre

Les formes de filtre varient grandement dans la littérature. Ils sont généralement choisis en fonction de l'ensemble de donnée. Les meilleurs résultats sur les images de MNIST (28 x 28) sont habituellement dans la gamme de 5 x 5 sur la première couche, tandis que l'ensemble de données d'image naturelle (souvent avec des certaine de pixels dans chaque dimensions) ont tendance à utiliser de plus grand filtre de première couche de 12 x 12, voire 15 x 15.

Le défi est donc de trouver le bon niveau de granularité de manière à créer des abstractions à l'échelle approprié et adapté à chaque cas [15].

2.6.3 Forme du Max Pooling

Les valeurs typiques sont 2 x 2. De très grands volumes d'entrée peuvent justifier un Pooling 4 x 4 dans les premières couches. Cependant, le choix de formes plus grandes va considérablement réduire la dimension du signal, et peut entraîner la perte de trop d'information.

2.7 Méthodes de régularisation

2.7.1 Empirique

a Dropout

Les couches « FC » ‘Fully connected’ occupent la majeure partie de la mémoire du CNN. D’ailleurs le concept de FC crée un problème exponentiel mémoire appelé « overfitting » (‘sur-connexion’ conduisant au sur-apprentissage) ralentissant le traitement de l’information. Pour prévenir cela, la méthode du dropout est utilisée pour « éteindre » les neurones aléatoirement (avec une probabilité prédéfinie, souvent un neurone sur deux) ainsi que les neurones périphériques. Ainsi, avec moins de neurones, le réseau est le plus réactif et peut donc apprendre plus rapidement. À la fin de la séance d’apprentissage, les neurones ‘éteints’ sont ‘rallumés’ (avec leurs poids originaux). Plus la couche FC est proche de l’image source, moins on éteindra de neurones.

L’objectif est d’éteindre et rallumer les neurones aléatoires, dans le cadre d’entraînement successifs. Une fois les séries d’entraînement terminées, on rallume tous les neurones et on utilise le réseau comme d’habitude. Cette technique a montré non seulement un gain dans la vitesse d’apprentissage, mais en déconnectant les neurones, on a aussi limité des effets marginaux, rendant le réseau plus robuste et capable de mieux généraliser les concepts appris.

b Dropconnect

Le DropConnect est une évolution du dropout, ou on ne va non plus éteindre un neurone, mais une connexion (l’équivalent de la synapse), et ce de manière toujours aléatoire. Les résultats sont similaires (rapidité, capacité de généralisation de l’apprentissage), mais présentent une différence au niveau de l’évolution des poids des connexions. Une couche FC avec un DropConnect peut s’apparenter à une couche à connexion ‘diffuse’.

c Pooling stochastique

Le Pooling stochastique reprend le même principe que le Max-Pooling, mais la sortie choisie sera prise au hasard, selon une distribution multinomiale définie en fonction de l’activité de la zone adressée par le pool. Dans les faits, ce système s’apparente à faire du Max-Pooling avec un grand nombre d’images similaire, qui ne varient que par des déformations localisées. On peut aussi considérer cette méthode comme une adaptation à des déformations élastique de l’image. C’est pourquoi cette méthode est très efficace sur les images MNIST (base de données d’images représentant des chiffres

manuscrits). La force du Pooling stochastique est de voir ses performances croître de manière exponentielle avec le nombre de couche du réseau.

2.7.2 Explicite

a Taille du réseau

Pour limiter la surcharge d'apprentissage, il faut spécifier le nombre de couche du réseau et éditer les connexions libres au réseau, ce qui limite et réduit la puissance et le potentiel prédictif du réseau.

b Dégradation du poids

Il s'agit de regarder le vecteur de la liste des poids associés aux signaux d'entrée, et d'ajouter un vecteur d'erreur proportionnel à la somme des poids ou au carré des poids, puis on multiplie ce vecteur d'erreur par le facteur de proportionnalité que l'on ajoutera pour augmenter la punition des vecteurs de poids fort.

- **Organiser selon les critères 1 :**

1. La particularité de cette liste est de réduire le poids des entrées irrégulières et faibles et d'augmenter le poids des entrées importantes car le système devient moins sensible au bruit.

- **Organiser selon les critères 2 :**

2. règle euclidienne : la particularité de cette liste est de réduire le poids des intrants puissants, et de forcer le neurone à prendre plus de poids en entrée avec peu de considération.

L'organisation du réseau est la somme entre les deux bases 'c'est la « régularisation de réseau élastique » (Elastic net régulation) [15].

2.8 Histogramme 2D

Graphiques 2D (le terme a été utilisé première fois par Pearson en 1895) représenter graphiquement une distribution fréquences des variables spécifiées, auquel les colonnes sont connectées à des intervalles de classe et les hauteurs sont proportionnelles à la fréquence de la classe, pour créer un histogramme 2d en python et le tracer avec matplotlib il existe plusieurs solutions. Le plus simple est d'utiliser la fonction matplotlib his2d.

En imagerie numérique, l'histogramme représente la distribution des intensités (ou des couleurs) de l'image, c'est un outil fondamental du traitement d'images, avec de très nombreuses applications, un graphique est utilisé pour résumer des données discrètes ou continues.

2.9 Conclusion

Dans ce chapitre, nous avons présenté et expliqué les concepts importants liés à l'apprentissage automatique et à l'intelligence artificielle ainsi qu'aux réseaux de neurones, tout en expliquant et en expliquant la méthode détaillée dans laquelle nous choisissons de travailler.

Pour mettre en œuvre notre travail de classification, nous avons utilisé l'apprentissage automatique, la méthode d'apprentissage qui a fait ses preuves ces dernières années, et avons choisi la méthode CNN, En tant que méthode de classification, C'est un choix, et c'est parce qu'il s'agit d'une méthode efficace facile à utiliser et simplifiée et on a discuté sur l'apprentissage machine et sa relation avec la classification dans l'ECG. C'est un sujet est très importance.

Chapitre 3 Méthodologie de recherche

3.1 Introduction

Python c'est le langage le plus parlé et le plus utilisé pour la construction et la formation de réseaux de neurones, alors Dans ce chapitre nous présenterons la partie pratique où nous verrons la classification des arythmies ECG à l'aide de la technologie CNN et son test d'apprentissage. Les signaux que nous allons classifier font partie de données d'arythmie ECG, et l'ensemble de données se compose de 10946 signaux, de sorte que chaque classe a de nombreux signaux, donc nous prenons 20% de chaque classe, qui est à l'origine 5 classes en gros (rythme normal, Beat inconnu, Battement ectopique ventriculaire, Battement ectopique supraventriculaire, Battement de fusion), Ceci afin d'équilibrer les données et d'obtenir un bon apprentissage. Le modèle Cnn que nous avons utilisé utilise les données comme un fichier d'entrée, il s'agit donc d'une image 2D, nous avons donc converti les signaux, c'est-à-dire chaque impulsion ECG en une image, Étant donné que les couleurs ne sont pas importantes dans les signaux ECG, nous les avons converties en images en niveaux de gris, D'où l'utilisation de l'histogramme 2d qui fonctionne sur le compte de l'intensité de chaque pixel.

3.2 Logiciel et librairies utilisé dans l'implémentation

3.2.1 Python

Python est un langage de programmation au même titre comme le C, C++, java, crée en 1989 par Guido Van Rossum, il se caractérise simple très lisibles, l'utilisation gratuitement et les fichiers sources sont disponibles et modifiables, aussi doté d'une bibliothèque disponibles et typage dynamique grande portabilité, fonctionnelle et orienté objet .python présent un support pour l'intégration d'autres langage , L'activité essentielle de ce programme consiste à résoudre des problèmes, Compétence de haut niveau .

La dernière version de python est la version 3, plus précisément, la version 3.7, il Ya plusieurs utilisation de langage python utilisé dans de nombreux contexte et s'adapter à tout type d'utilisation grâce à des bibliotheque spécialisé, et aussi langage de

développement de prototype lorsque on a besoin d'une application fonctionnelle avant l'optimiser avec un langage de plus bas niveau [16].



Figure3.1 : Logo de python

3.2.2 Keras

Keras est une api de réseaux de neurones de haut niveau, écrit en python et capable de fonctionner sur TensorFlow ou Theano .il a été développé en mettant l'accent sur l'expérimentation rapide. Etre capable d'aller de l'idée à un résultat avec le moins de délai possible est la clé pour faire de bonnes recherches. Il a été développé dans le cadre de l'effort de recherche du projet ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System), et son principal auteur et mainteneur est François Chollet, un ingénieur Google.

En 2017, l'équipe TensorFlow de Google a décidé de soutenir Keras dans la bibliothèque principale de TensorFlow. Chollet à expliquer que Keras a été conçue comme une interface plutôt que comme un cadre d'apprentissage end to end .il présente un ensemble d'abstraction de niveau supérieur et plus intuitif qui facilitent la configuration des réseaux neuronaux indépendamment de la bibliothèque informatique de backend. Microsoft travaille également à ajouter un backend CNTK à Keras aussi [17].



Figure 3.2 : logo de Keras

3.2.3 Numpy

Est une bibliothèque permettant d'effectuer des calculs numériques avec python. Elle introduit une gestion facilitée des tableaux de nombres, des fonctions sophistiquées (diffusion), on peut aussi l'intégrer le code C/C++ et Fortran [18].



Figure 3. 3: Logo de NumPy

3.2.4 Pandas

Pandas est une bibliothèque open source sous licence BSD fournissant des structures de données hautes performances et faciles à utiliser, ainsi que des outils d'analyse des données pour le langage de programmation python.

Pandas est un projet sponsorisé par NumFOCUS. Cela contribuera au succès du développement de pandas en tant que projet open source de classe mondiale et permettra de faire un don au projet [19].

3.2.5 Matplotlib

Est une bibliothèque de traçage pour le langage de programmation python et son extension mathématique numérique Numpy .il fournit une API orientée objet permettant d'incorporer des graphiques dans des applications à l'aide de kits d'outils d'interface graphique à usage général tels que Tkinter, wxPython, Qt ou GTK+ [20].

3.2.6 Scikit-learn

Scikit-learn est une bibliothèque libre Python dédiée à l'apprentissage automatique. Elle est développée par de nombreux contributeurs notamment dans le monde académique par des instituts français d'enseignement supérieur et de recherche comme Inria et Télécom ParisTech.

Elle comprend notamment de fonctions pour estimer des forêts aléatoires, des régressions logistiques, des algorithmes de classification, et les machines à vecteurs de

support. Elle est conçue pour s'harmoniser avec des autres bibliothèques libre python, notamment Numpy et SciPy [21].

3.2.7 Seaborn : Visualisation de données statistiques

Seaborn est une bibliothèque de visualisation de données Python basée sur matplotlib. Il fournit une interface de haut niveau pour dessiner des graphiques statistiques attrayants et informatifs.

Pour une brève introduction aux idées derrière la bibliothèque, vous pouvez lire les notes d'introduction. Visitez la page d'installation pour voir comment télécharger le package et commencer à l'utiliser [22].

3.3 La base de données

On a choisi pour notre projet comme base de données contiennent les signaux de pulsations de cœur, Dans lequel Cet ensemble de données est composé de deux collections de signaux de battement de cœur dérivés de deux célèbres ensembles de données dans la classification des battements de cœur, le jeu de données d'arythmie MIT-BIH et la base de données PTB Diagnostic ECG. Le nombre d'échantillons dans les deux collections est suffisamment grand pour entraîner un réseau neuronal profond.

Cet ensemble de données a été utilisé pour explorer la classification des pulsations cardiaques à l'aide d'architectures de réseaux neuronaux profonds et pour observer certaines des capacités d'apprentissage par transfert. Les signaux correspondent aux formes d'électrocardiogramme (ECG) des battements cardiaques pour le cas normal et les cas affectés par différentes arythmies et infarctus du myocarde. Ces signaux sont prétraités et segmentés, chaque segment correspondant à un battement cardiaque [23].

❖ Fichiers de données :

Cet ensemble de données se compose d'une série de fichiers CSV. Chacun de ces fichiers CSV contient une matrice, chaque ligne représentant un exemple dans cette partie de l'ensemble de données. Le dernier élément de chaque ligne indique la classe à laquelle appartient cet exemple.

❖ Le jeu de données d'arythmie contient :

- Nombre d'échantillons : 109446
- Nombre de catégories : 5
- Fréquence d'échantillonnage : 125Hz

- Source de données : ensemble de données d'arythmie MIT-BIH de Physio net
 - Classes : [0, 1, 2, 3, 4]
 - 1 = Beat inconnu
 - 2 = Battement ectopique ventriculaire
 - 3 = Battement ectopique supraventriculaire
 - 4 = Battement de fusion
 - Le nom des classes en programmation est :
 - 0 = normal beat
 - 1 =unknown beat
 - 2 =Ventricular ectopic beat
 - 0 = 0 = rythme normal
 - 3 =Supraventricular ectopic beat
 - 4 =Fusion beat
- ✓ Premièrement on a téléchargé la base de données sur notre PC, après nous l'avons transformé en fichier zip et on a téléchargé ce dernier en ligne sous le logiciel python (Colab), l'étape suivante consiste à ouvrir la base de données avec un code spécifique.

```

▶ from zipfile import ZipFile
  file_name = "mitbih_train.zip"
  with ZipFile (file_name,'r') as zip:
    zip.extractall()
    print('done')

```

3.3.1 Lecture de la base de données

- ✓ On crée un code spécial pour la lecture des entrées, des sorties et des mémoires utilisées de notre base de données (train et test).
- Mémoire utilisée pour « Train » : 31.4 MB.
 - Mémoire utilisée pour « Test » : 125.6 MB.

In put :

```

▶ import numpy as np
import pandas as pd

mit_test_data = pd.read_csv("mitbih_test.csv", header=None)
mit_train_data = pd.read_csv("mitbih_train.csv", header=None)

print("MIT test dataset")
print(mit_test_data.info())
print("MIT train dataset")
print(mit_train_data.info())

```

Output :

```

↳ MIT test dataset
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21892 entries, 0 to 21891
Columns: 188 entries, 0 to 187
dtypes: float64(188)
memory usage: 31.4 MB
None
MIT train dataset
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 87554 entries, 0 to 87553
Columns: 188 entries, 0 to 187
dtypes: float64(188)
memory usage: 125.6 MB
None

```

- ✓ Dans cette partie on a installé sous Colab les bibliothèques nécessaires pour faire les opérations concernant notre application, plus l'utilisation des fonctions d'API pour le bon fonctionnement.

```

▶ import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import classification_report
from sklearn.model_selection import train_test_split
from sklearn.metrics import f1_score
from sklearn.metrics import confusion_matrix
from keras.utils.np_utils import to_categorical
from sklearn.utils import class_weight
import warnings
warnings.filterwarnings('ignore')

```

3.3.2 Représentation des classes

Dans cette partie on a dessiné les classes dans un cercle proportionnel pour obtenir le pourcentage de chaque classe présentée sous notre base de données, après on va faire une comparaison entre les différentes classes présentées sous ce cercle.

In put :

```
per_class = train_df[187].value_counts()
plt.figure(figsize=(20,10))
my_circle=plt.Circle( (0,0), 0.7, color='white')
plt.pie(per_class, labels=['normal beat','unknown Beats','Ventricular ectopic beats','Supraventricular ectopic beats','Fusion Beats'],
        colors=['tab:blue','tab:orange','tab:purple','tab:olive','tab:green'],autopct='%1.1f%%')
p=plt.gcf()
p.gca().add_artist(my_circle)
plt.show()
```

Output :

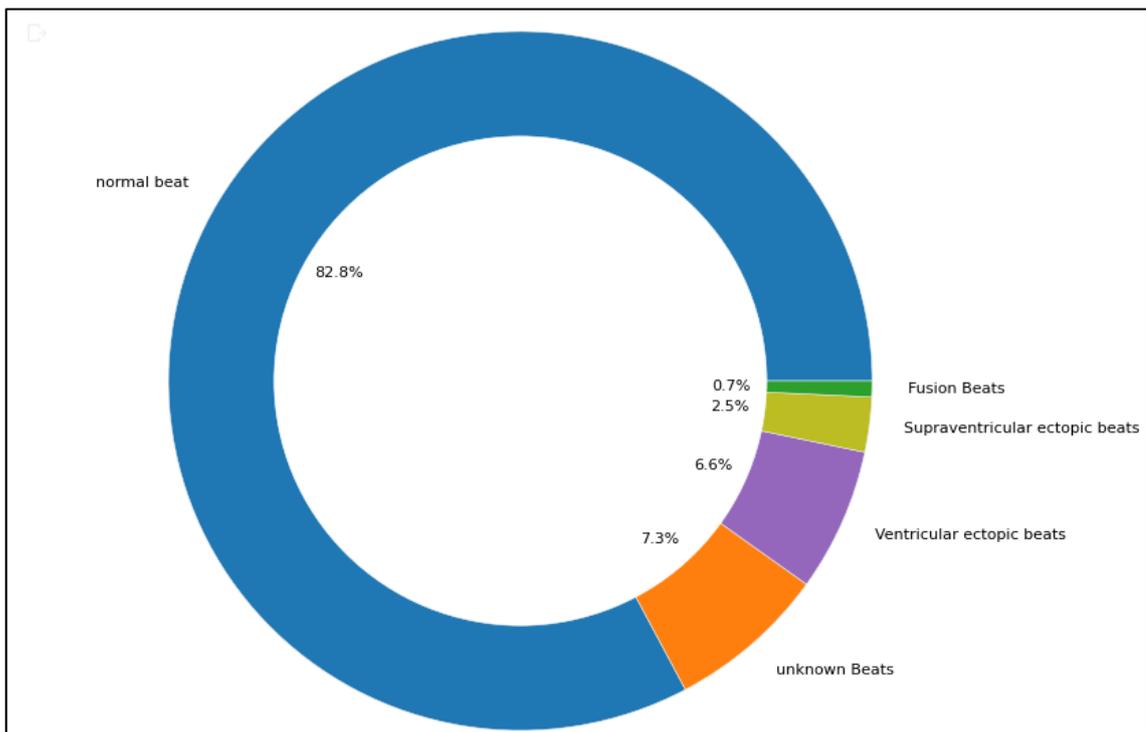


Figure 3.4 : Représentation des classes dans un cercle

3.3.3 Rééchantillonnage pour équilibrer l'ensemble de données

Ré-échantillonnage pour gérer correctement les ensembles de données non équilibrés :

Il existe plusieurs techniques pour gérer le déséquilibre d'une base de données, donc dans notre cas nous avons utilisé la technique de ré échantillonnage.

Le ré échantillonnage transforme l'ensemble de données en un ensemble plus équilibré en ajoutant des instances à la classe minoritaire ou en supprimant celles de la classe majoritaire, de cette façon nous construisons de meilleurs modèles d'apprentissage automatique.

La manière d'introduire ces changements dans un ensemble de données donné est obtenue via deux méthodes principales : **le sur échantillonnage** et **le sous-échantillonnage** [24].

- ❖ **Méthode de sur échantillonnage** : cette méthode ajoute des copies des instances la classe sous-représentée (classe minoritaire) pour obtenir un ensemble de données équilibré. Il existe plusieurs façons de sur échantillonner un ensemble de données.

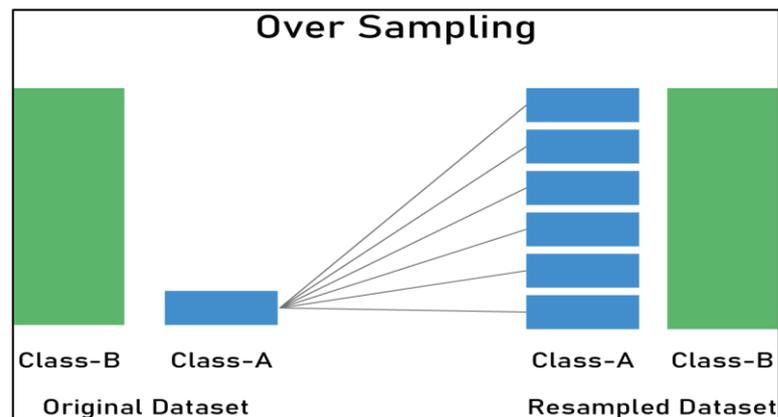


Figure 3. 5: Méthode de l'échantiollnage

- ❖ **Méthodes de sous-échantillonnage** : ces méthodes suppriment simplement les instances de la classe surreprésentée (classe majoritaire) de différentes manières. Le moyen le plus évident est de supprimer des instances de manière aléatoire.

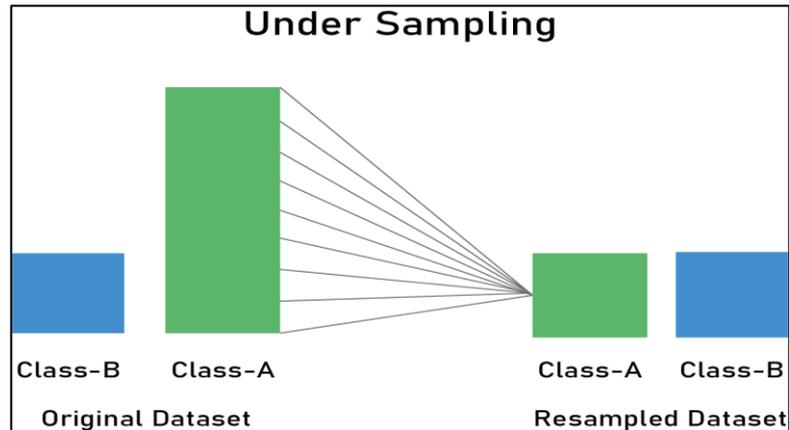


Figure 3. 6: Méthode de sous échantiollnage

Alors :

- ✓ Pour obtenir un meilleur résultat de l'apprentissage et pour établir l'équilibre dans notre base de données, on a prélevé de nouveau des échantillons pour chaque classe.

In put:

```

from sklearn.utils import resample
df_1=train_df[train_df[187]==1]
df_2=train_df[train_df[187]==2]
df_3=train_df[train_df[187]==3]
df_4=train_df[train_df[187]==4]
df_0=(train_df[train_df[187]==0]).sample(n=20000,random_state=42)

df_1_upsample=resample(df_1,replace=True,n_samples=20000,random_state=123)
df_2_upsample=resample(df_2,replace=True,n_samples=20000,random_state=124)
df_3_upsample=resample(df_3,replace=True,n_samples=20000,random_state=125)
df_4_upsample=resample(df_4,replace=True,n_samples=20000,random_state=126)

train_df=pd.concat([df_0,df_1_upsample,df_2_upsample,df_3_upsample,df_4_upsample])

per_class = train_df[187].value_counts()
plt.figure(figsize=(20,10))
my_circle=plt.Circle( (0,0), 0.7, color='white')
plt.pie(per_class, labels=['normal beat','unknown Beats','Ventricular ectopic beats','Supraventricular ectopic beats','Fusion Beats'],
        colors=['tab:blue','tab:orange','tab:purple','tab:olive','tab:green'],autopct='%1.1f%%')
p=plt.gcf()
p.gca().add_artist(my_circle)
plt.show()

```

Output:

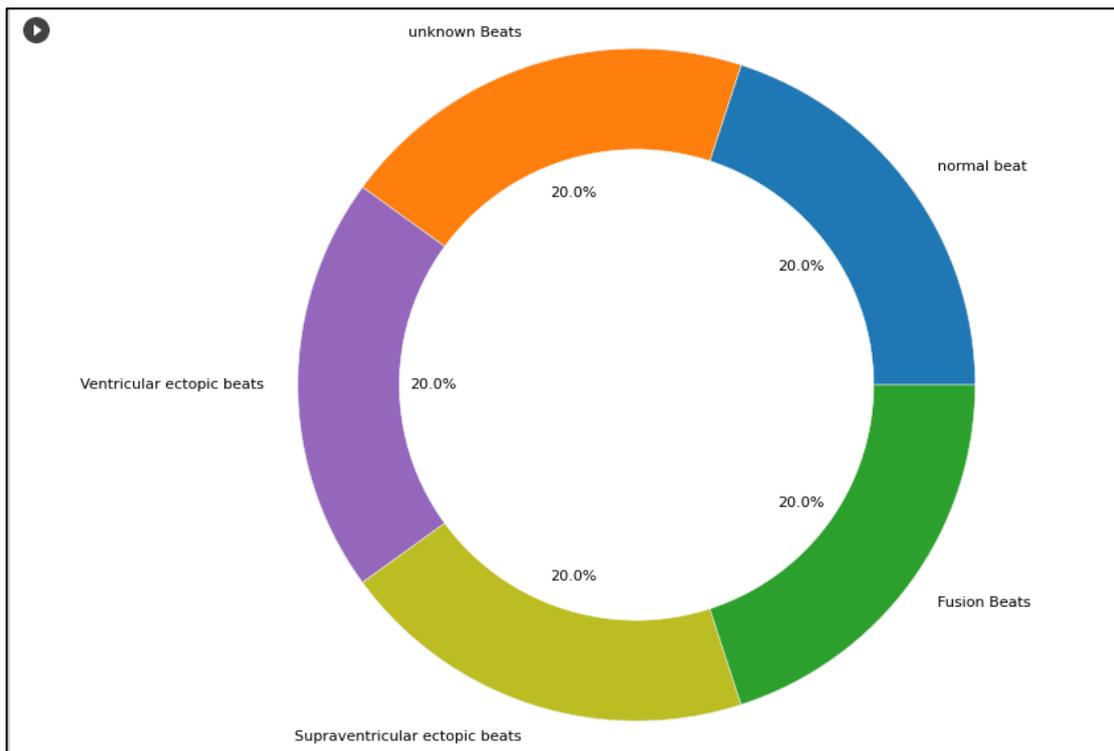


Figure 3. 7: Représentation des classes équilibré dans un cercle

3.4 Tracé des classes ECG

- ✓ Dans cette partie on a dessiné toutes les classes qu'on a sous notre base de données.

```
[ ] # plotting classes ECG  
plt.figure(figsize=(16,8))
```

- **Normal beat :**

In put :

```
# normal  
plt.subplot(2, 3, 1)  
plt.plot(classes.iloc[0,:186])  
plt.title('normal beat')
```

Output :

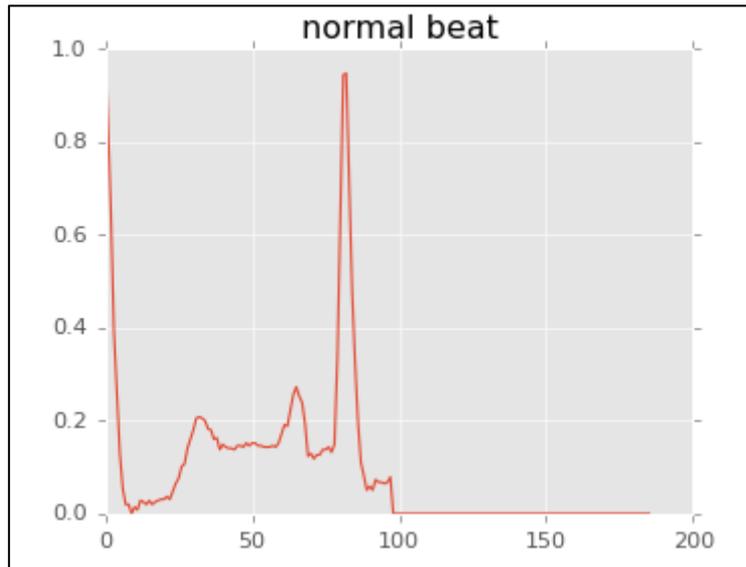


Figure 3. 8: Battement normal

- Unknown beat :

In put :

```
# unknown
plt.subplot(2, 3, 2)
plt.plot(classes.iloc[1,:186])
plt.title('unknown Beat')
```

Output :

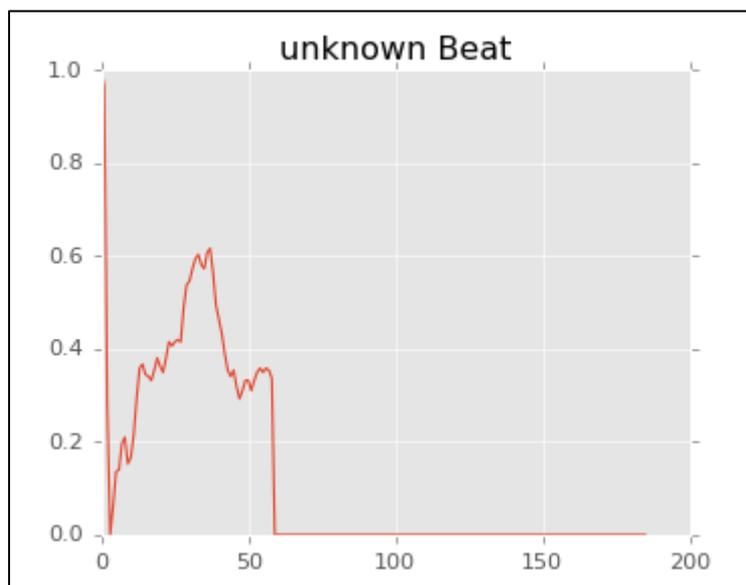


Figure 3. 9: Battement inconnu

- **Ventricular ectopic beat :**

In put :

```
# veb
plt.subplot(2, 3, 3)
plt.plot(classes.iloc[2,:186])
plt.title('Ventricular ectopic beat')
```

Output :

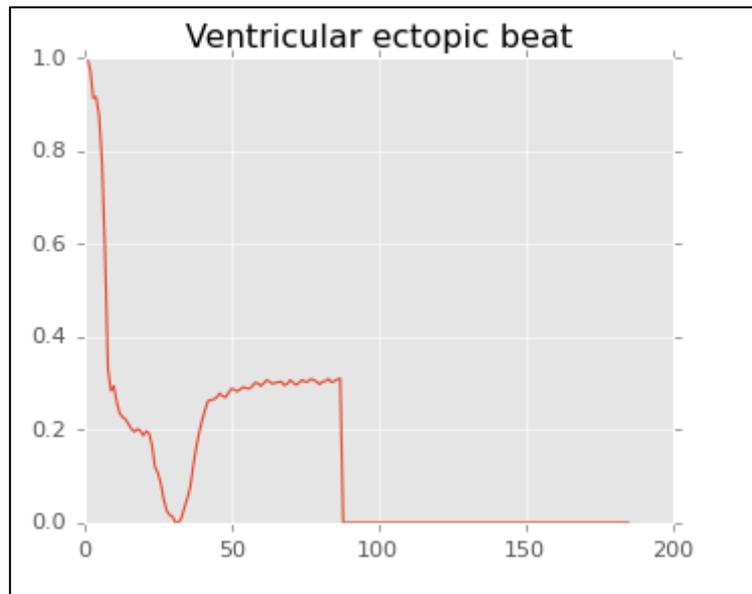


Figure 3. 10: Battement ectopique ventriculaire

- **Supraventricular ectopic beat :**

In put:

```
# seb
plt.subplot(2, 3, 4)
plt.plot(classes.iloc[3,:186])
plt.title('Supraventricular ectopic beat')
```

Output:

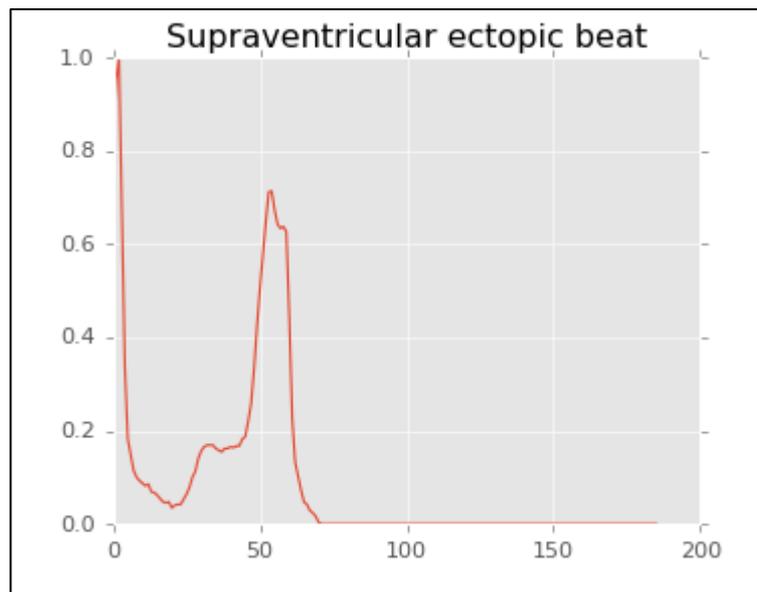


Figure 3. 11: Battement ectopique supraventriculaire

- **Fusion beat :**

In put :

```
# fusion
plt.subplot(2, 3, 5)
plt.plot(classes.iloc[4, :186])
plt.title('Fusion Beat')
```

Output :

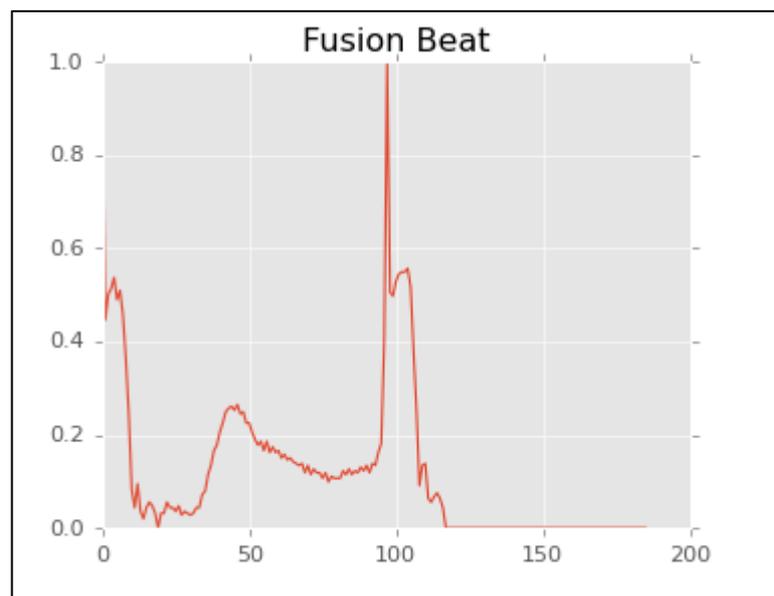


Figure 3. 12: Battement de fusion

3.5 Retraitement des données ECG

Le modèle CNN que nous avons utilisé nécessite des images de deux dimensions comme un fichier de données d'entrée, nous avons donc converti les signaux ECG en images ECG, Ceci est fait lors de l'étape de prétraitement des données, De sorte que nous traçons chaque impulsion ECG comme une seule image en niveaux de gris. Dans la base de données d'arythmie de chaque fichier, la fréquence cardiaque est segmentée en fonction de l'heure de pointe

Q-wave, Il s'agit de trouver et de déterminer le type d'arythmie au moment du pic de l'onde Q pour chaque battement cardiaque dans l'ECG.

$$\{T(Q_{peak}(n-1) + 20 < T(n) < T(Q_{peak}(n+1) - 20)\}$$

- **L'image au niveau de gris :**

L'image est une matrice composée de pixels, Cela signifie que l'image est en niveaux de gris, Autrement dit, l'image se compose de [0,255] couleurs où 0=noire et 255=blanc et Chaque pixel a une intensité entre [0 à 255].

- **Histogramme 2d :**

L'histogramme est la distribution de l'intensité en niveau de gris.

- ✓ Afin d'utiliser la technique de CNN, nous devons convertir les signaux ECG en images, Donc Nous avons utilisé un histogramme de deux dimensions pour que chaque classe a ses propres images.
- ✓ Convertissez tous les signaux ECG en histogramme de deux dimensions :

In put :

```

▶ def plot_hist(class_number,size,min_):
    # get the class data
    img=train_df.loc[train_df[187]==class_number].values
    # crop it
    img=img[:,min_:size]
    # creates a continous steam of data
    img_flatten=img.flatten()

    final1=np.arange(min_,size)
    for i in range (img.shape[0]-1):
        tempo1=np.arange(min_,size)
        final1=np.concatenate((final1, tempo1), axis=None)
    print(len(final1))
    print(len(img_flatten))
    plt.hist2d(final1,img_flatten, bins=(80,80),cmap=plt.cm.jet)
    plt.show()

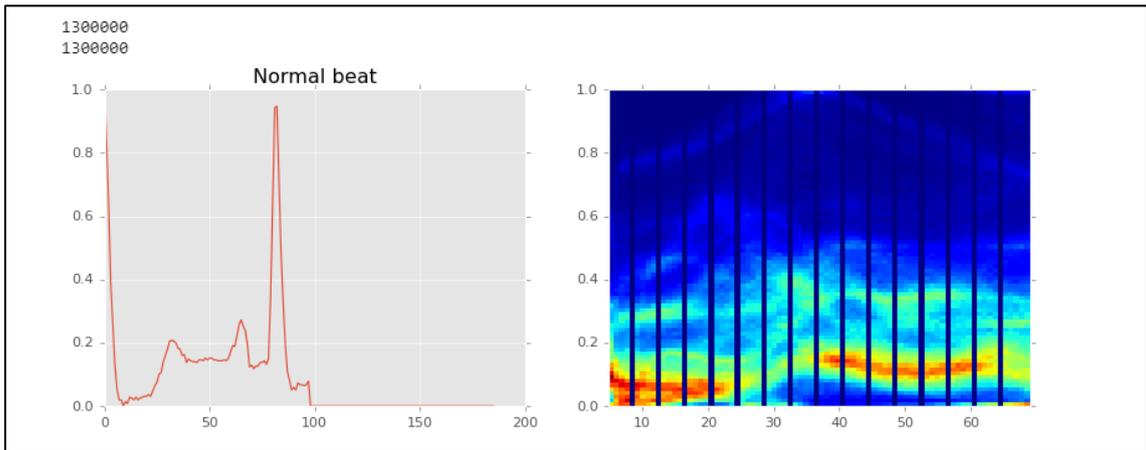
    # Normal Beat
plt.figure(figsize=(12,4))
plt.subplot(1, 2, 1)
plt.plot(classes.iloc[0,:186])
plt.title('Normal beat')
plt.subplot(1, 2, 2)
plot_hist(0,70,5)

```

Output :

- **Normal beat :**

Figure 3. 13: Représentation de transformation la classes 0 en histogramme de deux dimension



- **Unknown beat :**

```

▶ # unknown
plt.figure(figsize=(12,4))
plt.subplot(1, 2, 1)
plt.plot(classes.iloc[1,:186])
plt.title('unknown Beat')
plt.subplot(1, 2, 2)
plot_hist(1,50,5)

```

Output :

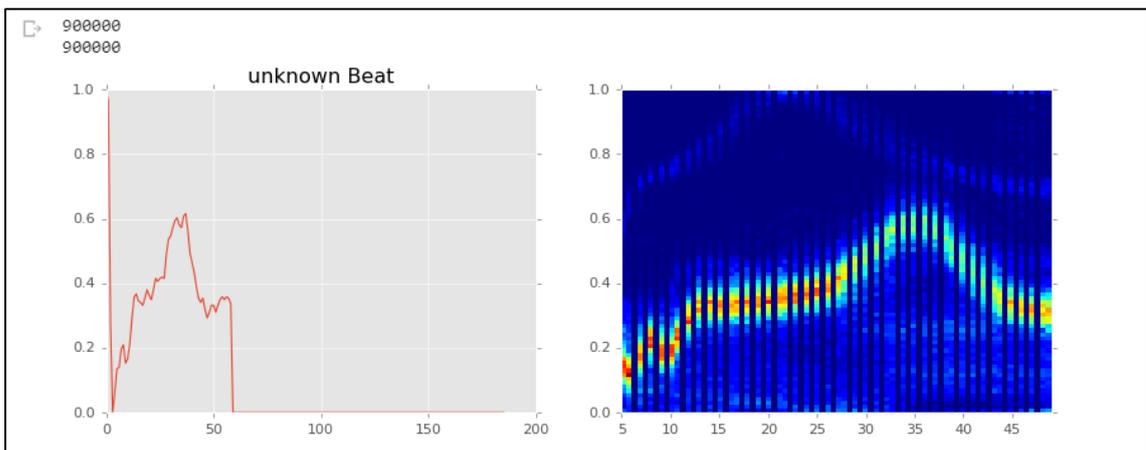


Figure 3. 14: Représentation de transformation la classes 1 en histogramme de deux dimension

- **Ventricular ectopic beat :**

```
[ ] # veb
plt.figure(figsize=(12,4))
plt.subplot(1, 2, 1)
plt.plot(classes.iloc[2,:186])
plt.title('Ventricular ectopic beat')
plt.subplot(1, 2, 2)
plot_hist(2,60,30)
```

Output :

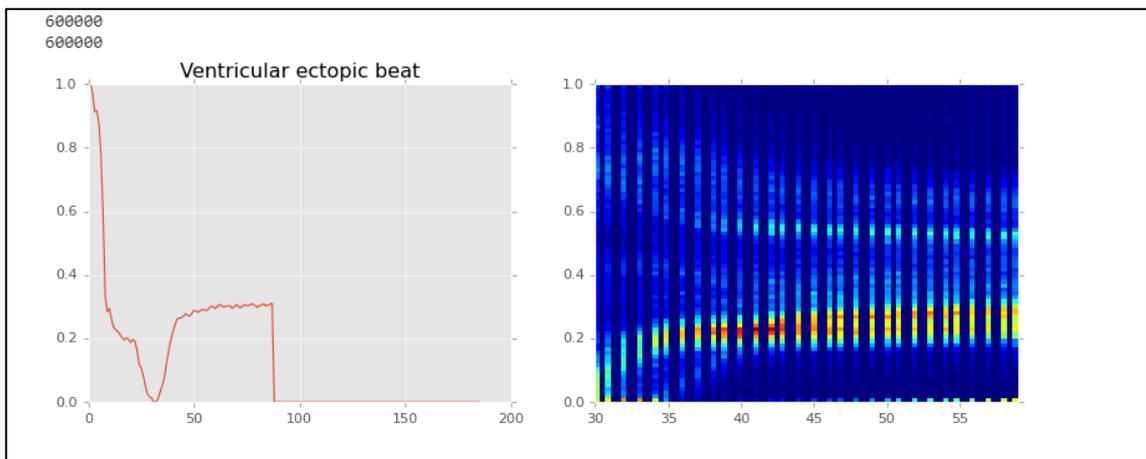


Figure 3. 15: Représentation de transformation la classes 2 en histogramme de deux dimension

- **Supraventricular ectopic best :**

```
[ ] # seb
plt.figure(figsize=(12,4))
plt.subplot(1, 2, 1)
plt.plot(classes.iloc[3,:186])
plt.title('Supraventricular ectopic beat')
plt.subplot(1, 2, 2)
plot_hist(3,60,25)
```

Output :

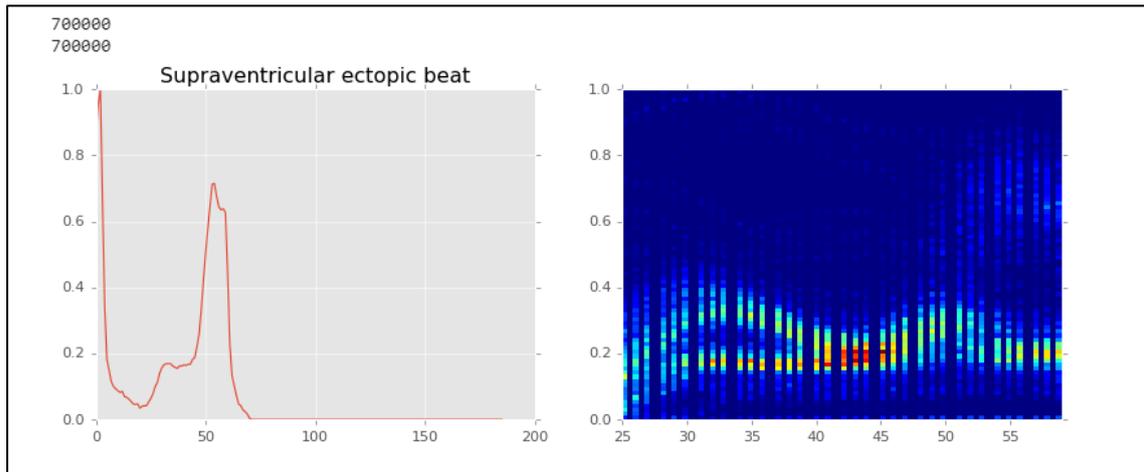


Figure 3.16: Représentation de transformation la classes 3 en histogramme de deux dimension

- **Fusion beat :**

```

# fusion
plt.figure(figsize=(12,4))
plt.subplot(1, 2, 1)
plt.plot(classes.iloc[4,:186])
plt.title('Fusion Beat')
plt.subplot(1, 2, 2)
plot_hist(4,50,18)

```

Output :

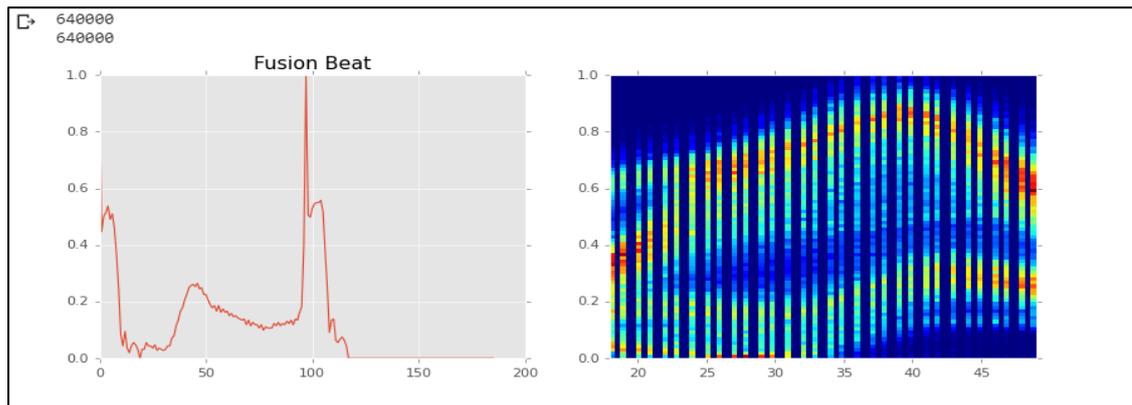


Figure 3.17: Représentation de transformation la classes 4 en histogramme de deux dimension

✓ Nous avons utilisé gaussien pour rapprocher les signaux ECG de la réalité.

Car nous avons travaillé dans le cas parfait, où il n'y a pas de bruit, ce qui signifie que le signal est complètement isolé, Nous avons également utilisé le gaussien pour déterminer la quantité de bruit que notre signal peut gérer.

In put :

```
def add_gaussian_noise(signal):  
    noise=np.random.normal(0,0.05,186)  
    return (signal+noise)  
tempo=classes.iloc[0,:186]  
bruite=add_gaussian_noise(tempo)  
  
plt.subplot(2,1,1)  
plt.plot(classes.iloc[0,:186])  
  
plt.subplot(2,1,2)  
plt.plot(bruite)  
  
plt.show()
```

Output :

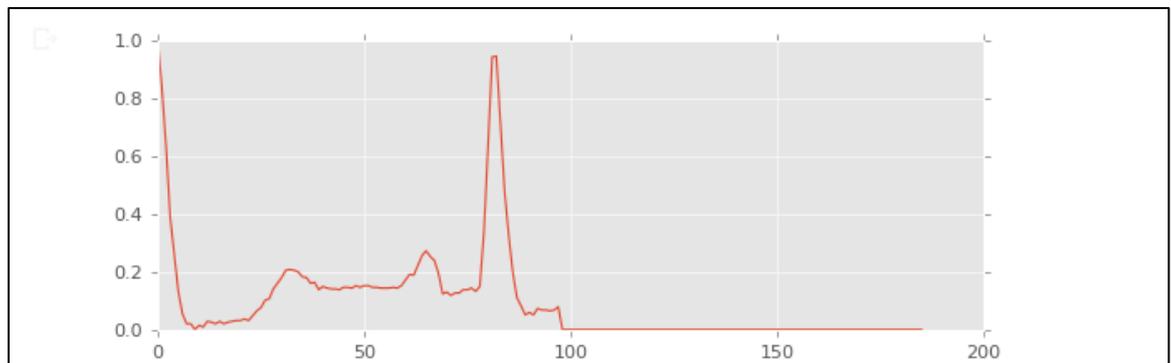


Figure 3. 18: Représentation de signal ECG dans le cas parfait

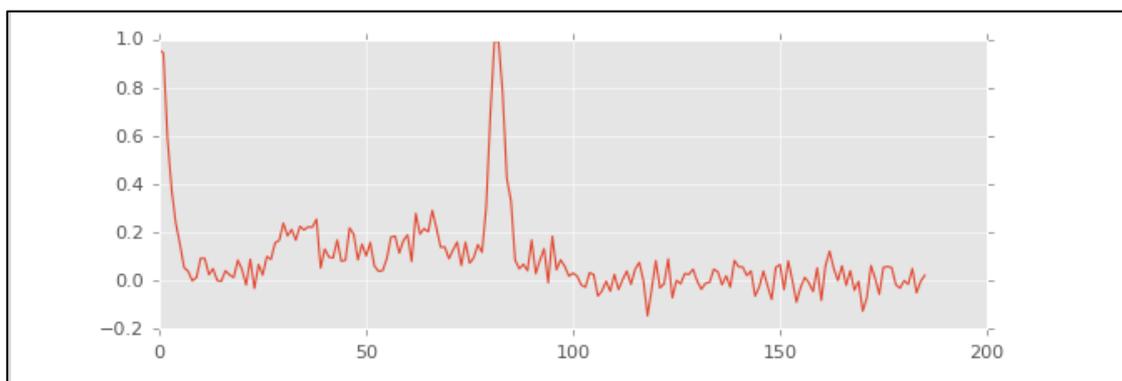


Figure 3. 19: Représentation de signal ECG avec bruit

3.6 Model Training

Dans la dernière partie de notre code (programme) on a utilisé la technique CNN, le fonctionnement de cette technique consiste à classifier les images et à la fin elle va nous donner 5 classes.

Le modèle de CNN (L'architecture) proposé dans notre travail contient 3 couches convolutives se rejoindre avec 3 couches de max-pooling .la sortie de ces couches va concaténer en vecteur qui devient l'entrée de ANN, ce dernier comporte 3 couches entièrement connectées (couche d'entrée, couche caché et couche de sortie).

1. Les couches de CNN (Convolutional Neural Network)

La 1^{ère} Couche (Couche convolutive 1) : Cette couche composée de 64 filtres de taille (6x6), opération d'activation ReLu (fonction $h(x) = \max(0, x)$).

La 2^{ème} Couche (Couche max-pooling 1) : Comporte un filtre de taille (3x3) et un stride avec une matrice de (2x2), suivie avec l'opération de padding.

La 3^{ème} Couche (Couche convolutive 2) : Cette couche composée aussi de 64 filtres de taille (3x3), Ceci est suivi d'un ReLU fonction d'activation (fonction $h(x) = \max(0, x)$).

La 4^{ème} Couche (Couche max-pooling 2) : contient un filtre de taille (2x2) et aussi un stride avec une matrice de (2x2), suivie avec l'opération de padding.

La 5^{ème} Couche (Couche convolutive 3) : Elle est composée aussi de 64 filtres de taille (3x3), et une fonction d'activation ReLU.

La 6^{ème} Couche (Couche max-pooling 3) : Cette couche est la dernier dans le CNN comporte aussi 64 filtres de taille (2x2) et un stride avec une matrice de (2x2), suivie par une opération d'activation.

2. Les couches d'ANN (Artificiel Neural Network)

La 1^{ère} Couche (Couche d'entrée) : Cette couche est la première couche de ANN, Elle contient 64 neurones suivi avec la fonction d'activation ReLU.

La 2^{ème} Couche (Couche caché) : La couche caché dans ce Modèle contient 32 neurones qui se rejoindre une fonction d'activation ReLU.

La 2^{ème} Couche (Couche de sortie) : La couche de sortie contient 5 neurones, suivis avec la fonction d'activation SOFTMAX.

En résumé, Nous appliquons une série des couches de notre modèle CNN, pour former à la sortie un vecteur de 64 (Flattening), qui devient l'entrée de la première couche d'ANN. Nous appliquons ensuite une couche cachée de 32 neurones avec la fonction d'activation ReLU. La couche finale est une couche SOFTMAX avec 5 neurones car nous avons 5 classes qui sont normales : beat, unknown Beat, Ventricular ectopic beat, Supraventricular ectopic beat, Fusion Beat.

In put :

```
# Network
def network(X_train,y_train,X_test,y_test):
    im_shape=(X_train.shape[1],1)
    inputs_cnn=Input(shape=(im_shape), name='inputs_cnn')
    conv1_1=Convolution1D(64, (6), activation='relu', input_shape=im_shape)(inputs_cnn)
    conv1_1=BatchNormalization()(conv1_1)
    pool1=MaxPool1D(pool_size=(3), strides=(2), padding="same")(conv1_1)
    conv2_1=Convolution1D(64, (3), activation='relu', input_shape=im_shape)(pool1)
    conv2_1=BatchNormalization()(conv2_1)
    pool2=MaxPool1D(pool_size=(2), strides=(2), padding="same")(conv2_1)
    conv3_1=Convolution1D(64, (3), activation='relu', input_shape=im_shape)(pool2)
    conv3_1=BatchNormalization()(conv3_1)
    pool3=MaxPool1D(pool_size=(2), strides=(2), padding="same")(conv3_1)
    flatten=Flatten()(pool3)
    dense_end1 = Dense(64, activation='relu')(flatten)
    dense_end2 = Dense(32, activation='relu')(dense_end1)
    main_output = Dense(5, activation='softmax', name='main_output')(dense_end2)

    model = Model(inputs= inputs_cnn, outputs=main_output)
    model.compile(optimizer='adam', loss='categorical_crossentropy',metrics = ['accuracy'])

    callbacks = [EarlyStopping(monitor='val_loss', patience=8),
                 ModelCheckpoint(filepath='best_model.h5', monitor='val_loss', save_best_only=True)]

    history=model.fit(X_train, y_train,epochs=5,callbacks=callbacks, batch_size=32,validation_data=(X_test,y_test))
    model.load_weights('best_model.h5')
    return(model,history)
```

3.7 Résultats

In put :

```
[ ] model = Model(inputs= inputs_cnn, outputs=main_output)
model.compile(optimizer='adam', loss='categorical_crossentropy',metrics = ['accuracy'])

callbacks = [EarlyStopping(monitor='val_loss', patience=8),
             ModelCheckpoint(filepath='best_model.h5', monitor='val_loss', save_best_only=True)]

history=model.fit(X_train, y_train,epochs=5,callbacks=callbacks, batch_size=32,validation_data=(X_test,y_test))
model.load_weights('best_model.h5')
return(model,history)
```

```
[ ] def evaluate_model(history,X_test,y_test,model):
    scores = model.evaluate((X_test),y_test, verbose=0)
    print("Accuracy: %.2f%%" % (scores[1]*100))

    print(history)
    fig1, ax_acc = plt.subplots()
    plt.plot(history.history['accuracy'])
    plt.plot(history.history['val_accuracy'])
    plt.xlabel('Epoch')
    plt.ylabel('Accuracy')
    plt.title('Model - Accuracy')
    plt.legend(['Training', 'Validation'], loc='lower right')
    plt.show()
```

```
fig2, ax_loss = plt.subplots()
plt.xlabel('Epoch')
plt.ylabel('Loss')
plt.title('Model- Loss')
plt.legend(['Training', 'Validation'], loc='upper right')
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.show()
target_names=['0','1','2','3','4']

y_true=[]
for element in y_test:
    y_true.append(np.argmax(element))
prediction_proba=model.predict(X_test)
prediction=np.argmax(prediction_proba,axis=1)
cnf_matrix = confusion_matrix(y_true, prediction)
```

Output :

```
Epoch 1/5
3125/3125 [=====] - 96s 31ms/step - loss: 0.1936 - accuracy: 0.9309 - val_loss: 0.1708 - val_accuracy: 0.9441
Epoch 2/5
3125/3125 [=====] - 91s 29ms/step - loss: 0.0778 - accuracy: 0.9730 - val_loss: 0.1342 - val_accuracy: 0.9611
Epoch 3/5
3125/3125 [=====] - 92s 29ms/step - loss: 0.0504 - accuracy: 0.9833 - val_loss: 0.1329 - val_accuracy: 0.9625
Epoch 4/5
3125/3125 [=====] - 92s 29ms/step - loss: 0.0393 - accuracy: 0.9869 - val_loss: 0.0949 - val_accuracy: 0.9761
Epoch 5/5
3125/3125 [=====] - 92s 29ms/step - loss: 0.0310 - accuracy: 0.9902 - val_loss: 0.1892 - val_accuracy: 0.9529
```

❖ Résultat finale :

```
Accuracy: 97.61%
<tensorflow.python.keras.callbacks.History object at 0x7fa3f7b75f60>
.. . . .
```

✓ Après quelques minutes d'attente, nous avons obtenu successivement les résultats finals avec des degrés de précision très satisfaisants, 98,69% pour les données de l'apprentissage et 97,61% pour les données du test.

- Les résultats sont affichés dans les dernières images :

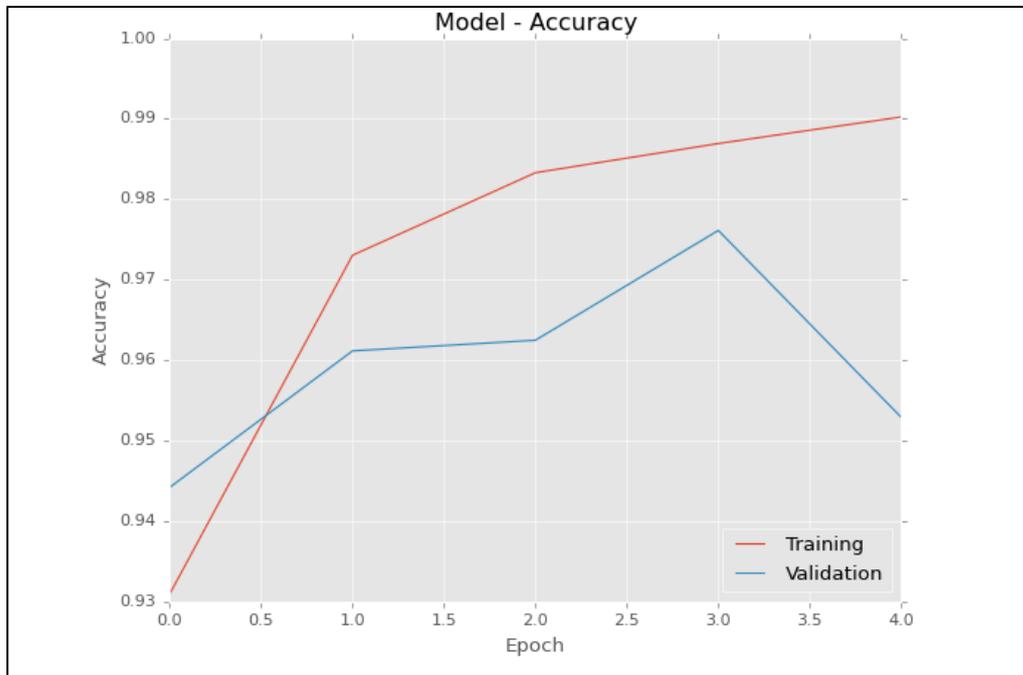


Figure 3. 20: Le model accuracy pour l'apprentissage et le teste.

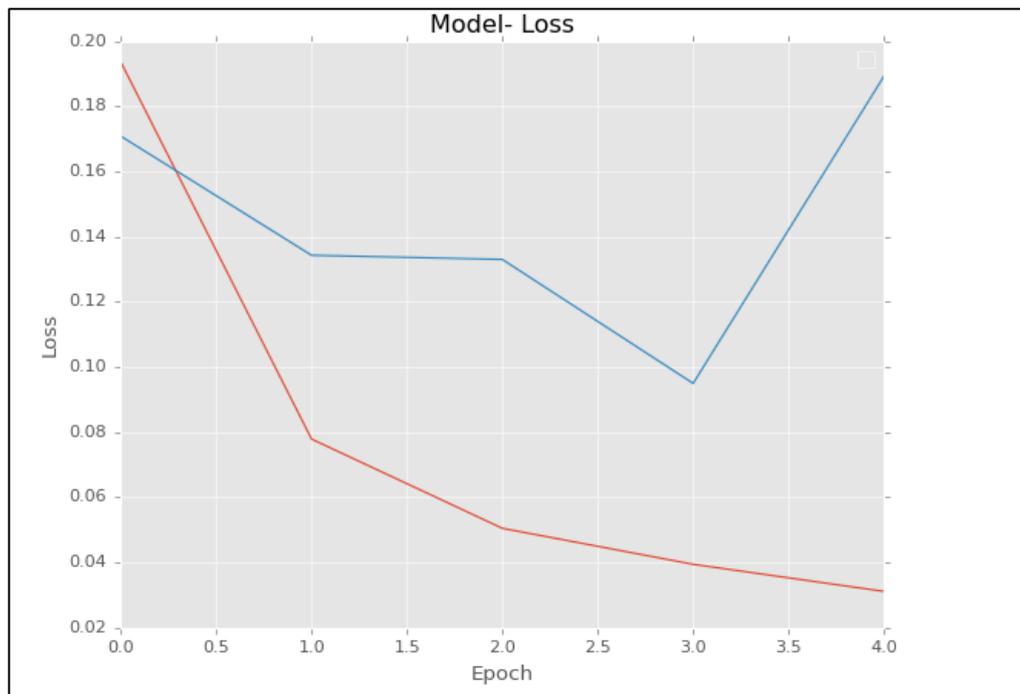


Figure 3.21: Le model loss pour l'apprentissage et le teste.

- A partir du schéma de la perte et de la précision on conclut que notre modèle marche très bien, parce que cela nous donne un taux de précision très élevé, 98,69% pour l'apprentissage et 97,61% pour le test, On outre, la valeur de la perte était très basse et presque constante, mais elle diminuait avec le temps, De cela, nous disons que le classificateur CNN a bien fonctionné et il n'a pas eu de sur-traitement.

3.8 Conclusion

Ce travail a démontré la capacité à stabiliser les arythmies ECG, pour cela on a utilisé la technique de de classification qui utilise à son tour certaines fonctionnalités et algorithmes d'apprentissage automatique. Le modèle CNN avec lequel nous avons travaillé traite les images de deux dimensions comme des données d'entrée, pendant le Retraitement des données ECG Les signaux ECG sont convertis en images ECG, étant donné que les couleurs ne sont pas importantes pour les signaux ECG, donc nous les avons converties en une image en niveaux de gris par un histogramme de deux dimensions. En utilisant ces images obtenues à partir de l'ECG, nous pouvons classifier 5 types d'ECG, L'objectif principal de notre travail et d'utiliser les données du pc pour faire un apprentissage pour que l'ordinateur fonctionne comme un système nerveux humain, puis en faisant un test pour voir la fiabilité du travail que nous avons effectué. Malgré les complications de l'arythmie ECG, les résultats étaient satisfaisants et très élevés et le taux de réussite était de 97,61 %.

Conclusion générale

L'avenir de l'intelligence artificielle est en grand développement et est en cours de développement pour son utilisation à l'heure actuelle, et des formations sont constamment faites pour l'améliorer et augmenter sa précision.

Dans ce projet, nous avons discuté de la manière de classer à l'aide de réseaux de neurones généraux. Pour effectuer cette classification, nous avons utilisé l'apprentissage automatique avec la méthode CNN car elle se caractérise par sa simplicité et sa précision.

Ce travail vise à être en mesure de classer les rythmes cardiaques ainsi que d'étudier le signal d'électrocardiogramme (ECG), en construisant un système permettant de classer les signaux ECG dans la catégorie normale ou arythmique. De plus, l'ensemble du processus peut être utilisé pour la classification avec d'autres types d'ensemble de données.

On peut conclure de ce travail que l'utilisation de réseaux de neurones convolutifs avec apprentissage automatique pour détecter les arythmies est très précise.

Bibliographie

- [1]: S. Allender, P. Scarborough, V. Peto, M. Rayner, J. Leal, R. Luengo- Fernandez, and A. Gray, "European cardiovascular disease statistics," 2008.
- [2]: Mansier, J. Clairambault, N. Charlotte, C. Médigue, C. Vermeiren, G. LePape, F. Carré, A. Gounaropoulou, B. Swynghedauw, Linear and nonlinear analyses of heart rate variability : a minireview, *Cardiovascular Research* 31 (1996)371.
- [3]: D. Lloyd-Jones, R. J. Adams, T. M. Brown, M. Carnethon, S. Dai, G. De Simone, T. B. Ferguson, E. Ford, K. Furie, C. Gillespie et al., "Heart disease and stroke statistics 2010 update," *Circulation*, vol. 121, no. 7, pp. e46–e215, 2010.
- [4]: Gavin P. Shorten; Martin J. Burke "A precision ECG signal generator providing full Lead II QRS amplitude variability and an accurate timing profile" 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
- [5]: miscauthor = A.L.I.E.2, title = Anatomie et lectrophysiologie du coeur, réalisation, langage et anomalies des lectrocardiogrammes., howpublished.
- [6] : misc author = Cardiostore, title = Le matériel de cardiologie, plus pro, moins cher., howpublished = "<https://www.cardiostore.fr/>", note = " [Online ; accessed 10/04/2019]".
- [7]: book author=John R. Hampton, title=The ECG Made Easy, publisher=Elsevier Masson SAS, year=2013.
- [8]: Arzeno NN, Deng ZD, Poon CS. 2008. Analysis of first-derivative based QRS detection algorithms. *IEEE Transaction on Biomedical Engineering* 55(2) :478–484 DOI 10.1109/tbme.2007.912658. J. P. ABENSTEIN, W. J. TOMPKINS "A New data.
- [9]: Reduction Algorithm for Real Time ECG Analysis" *IEEE Trans. Biomed. Eng.*
- [10]: Vol. BME-29, NO 1 Jan. 1982, pp43-48P. AUGUSTYNIAK "Time-Frequency and Discrimination of Noise in the Electrocardiogram", Physiological Measurement Institute of Physics Publishing, N°24, 2003 PP. 753-767.

[11]:http://www.iro.umontreal.ca/~aimeur/cours/ift6261/Survol_apprentissage_mach_ine.pdf

[12]:http://www.iro.umontreal.ca/~aimeur/cours/ift6261/Survol_apprentissage_mach_ine.pdf

[13] : https://www.elementai.com/fr/news/2019/apprentissage-machine-supervise-oppose-a-non-supervise,c'est_consulter_le_19/08/2020

[14] : https://le-datascientist.fr/apprentissage-supervise-vs-non-supervise.c'est_consulter_le_19/08/2020

[15] : Mr. Mokri Mohammed Zakaria., » Classification des images avec les réseaux de neurones convolutionnels « : mémoire de projet de fin d'étude. Université Abou Bakr Belkaid Tlemcen. Page 48

[16] : [https://www.python.org/\(2019\)](https://www.python.org/(2019)).

[17] : [https://www.keras.org/\(2019\)](https://www.keras.org/(2019)).

[18] :[https://www.Numpy.org/\(2019\)](https://www.Numpy.org/(2019)).

[19] : [https://www.pandas.org/\(2019\)](https://www.pandas.org/(2019)).

[20] : [https://www.matplotlib.org/\(2019\)](https://www.matplotlib.org/(2019)).

[21] : [https://www.scikit-learn.org/\(2019\)](https://www.scikit-learn.org/(2019))

[22] : [https://www.Seaborn.org/\(2019\)](https://www.Seaborn.org/(2019))

[23] : https://www.kaggle.com/shayanfazeli/heartbeat#mitbih_train.csv

[24] : <https://heartbeat.fritz.ai/resampling-to-properly-handle-imbalanced-datasets-in-machine-learning>.