

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediène

Faculté d'Électronique et d'Informatique



THÈSE

Présentée pour l'obtention du **grade** de **DOCTEUR EN SCIENCES**

En : ÉLECTRONIQUE

Spécialité : Instrumentation Électronique

Par : SNOUSSI Assya

**Apport des techniques de traitement et de classification
pour l'exploitation des signaux
d'électroencéphalogramme (EEG)**

Soutenue publiquement le 08/10/2015, devant le jury composé de :

M. ATTARI Mokhtar	Professeur à l'U.S.T.H.B	Président
Mme TALHA-KEDIR M.D.	Maître de Conférences/A à l'U.S.T.H.B	Directrice de thèse
M. GUESSOUM Abderezak	Professeur à l'Université de Blida	Examineur
M. MEKAOUI Slimane	Professeur à l'U.S.T.H.B	Examineur
M. GOUTAS Ahcène	Maître de Conférences/A à l'Université de Skikda	Examineur
M. ADNANE Mourad	Maître de Conférences/A à l'École Nationale Polytechnique	Examineur

A la mémoire de mon père.

Remerciements

Je remercie ma directrice de thèse docteur TALHA-KEDIR Malika pour ses judicieux conseils, pour son dynamisme et surtout pour m'avoir fait confiance en m'intégrant dans l'équipe du génie biomédicale.

Un grand merci au professeur ATTARI Mokhtar de m'avoir accepté dans son laboratoire (LINS) et donné l'occasion de réaliser cette thèse. Je le remercie aussi pour l'intérêt qu'il a bien voulu porter à ce travail en acceptant de présider le jury.

Je souhaite, également, remercier les professeurs GUESSOUM Abdelkader, MEKAOUI Slimane et les docteurs ADNANE Mourad et GOUTAS AHCÈNE pour m'avoir fait l'honneur d'accepter d'être examinateurs de ce travail.

Je dédie cette thèse à mon mari Hicham et à mes enfants Hana et Ryad. Sans leur soutien, ce travail n'aurait pas vu le jour. Merci.

Table des matières

Introduction	1
1 Contexte médical, état de l’art et problématique	5
I.1 Contexte médical	5
1.1 Introduction.....	5
1.2 Le cerveau.....	6
1.2.1 Le neurone	7
1.3 Transmission de l’influx nerveux	8
1.3.1 Transmission chimique	8
1.3.2 Transmission électrique	10
1.4 L’épilepsie	10
1.4.1 Type de crises	11
1.5 L’électro-encéphalogramme	12
1.5.1 Activités électriques cérébrales rythmiques	15
1.5.2 Mesure des signaux EEG	16
I.2 État de l’art et problématique.....	18
2 Espace de représentation à base d’ondelette	21
2.1 Représentation temps-fréquence.....	21
2.1.1 Principe d’Heisenberg-Gabor	21
2.1.2 Transformée de Fourier Fenêtrée.....	21
2.1.3 Transformée en ondelette continue.....	22
2.2 Analyse multi-résolution	25
2.3 Espace de représentation.....	25
2.3.1 Mise en œuvre de la DWT	28

2.3.2 Transformée par paquets d'ondelettes	29
3 Méthodes de réduction	31
3.1 Analyse en composantes principales	31
3.1 Analyse discriminante linéaire	32
4 Classification et performances de mesures	36
4.1 Introduction.....	36
4.2 Classification par machines à vecteurs support	37
4.3 Classification par k plus proches voisins	41
4.4 Classification par réseau de neurone artificiel	42
4.5 Performances statistiques.....	44
5 Résultats expérimentaux et évaluation des performances	47
5.1 Base de données.....	48
5.2 Prétraitement et extraction des caractéristiques.....	50
5.2.1 Transformée en ondelette discrète	50
5.2.2 Transformée en paquets d'ondelette	53
5.3 Extraction des paramètres caractéristiques	57
5.4 Réduction des données	57
5.5 Classification	60
5.5.1 Méthodes utilisant la classification par SVM	60
5.5.1.1 Validation croisée	60
5.5.1.2 Machines à vecteur support	61
5.5.2 Performances et évaluations statistiques.....	64
5.5.3 Méthodes utilisant la classification par k plus proches voisins	77
5.5.4 Méthodes utilisant la classification par réseaux de neurone artificiels.....	80
5.5.5 Comparaison des algorithmes de classification utilisés.....	84
6 Conclusion et perspectives	87
7 Bibliographie	90

Liste des figures

Figure 1.1 : Schéma des divisions de l'encéphale	7
Figure 1.2 : Représentation d'un neurone	8
Figure 1.3 : Transmission chimique	9
Figure 1.4 : Exemple d'un enregistrement d'EEG	14
Figure 1.5 : Système (10 – 20) : Placement de 21 électrodes	14
Figure 1.6 : Champs électriques générés par des courants synaptiques	15
Figure 2.1 : Signal EEG analysé à l'aide de la transformée en ondelette continue	24
Figure 2.2 : Décomposition par transformée en ondelette discrète	26
Figure 2.3 : Arbre de paquets d'ondelettes	30
Figure 2.4 : Décomposition par transformée en paquets d'ondelettes	30
Figure 3.1 : Réduction de dimensions par LDA et par ACP	33
Figure 4.1 : Classification binaire avec plusieurs hyperplans	38
Figure 4.2 : Hyperplan de séparation optimale avec marge souple	38
Figure 4.3 : Architecture d'un réseau MLP	42
Figure 5.1 : Schéma de la méthodologie proposée.....	48
Figure 5.2 : Représentations des sets (A-B-C-D).....	49
Figure 5.3 : Représentation des différentes ondelettes	51

Figure 5.4 : Décomposition par la DWT d'un signal EEG sain	52
Figure 5.5 : Décomposition par la DWT d'un signal EEG épileptique	53
Figure 5.6 : Décomposition par la WPT du signal EEG sain	54
Figure 5.7 : Décomposition par la WPT d'un signal EEG épileptique	55
Figure 5.8 : Distribution énergétique : méthode de réduction PCA.....	58
Figure 5.9 : Distribution énergétique : méthode de réduction LDA	59
Figure 5.10 : Classification par SVM (utilisant la PCA)	62
Figure 5.11 : Classification par SVM (utilisant la LDA)	63
Figure 5.12 : Histogrammes de la précision de classification des sets A et E	66
Figure 5.13 : Histogrammes de la précision d'une classification par la WPT	66
Figure 5.14 : Décomposition du signal EEG (set C) basée sur la DWT	69
Figure 5.15 : Décomposition du signal EEG (set C) basée sur la WPT	70
Figure 5.16 : Performances de classification (basée sur la DWT) du groupe (A-C)	72
Figure 5.17 : Performances de classification (basée sur la WPT) du groupe (A-C)	73
Figure 5.18 : Performances de classification (basée sur la DWT), du groupe (A-E).....	74
Figure 5.19 : Performances de classification (DWT) groupe A-(C-E)	76
Figure 5.20 : Performances de classification (WPT) groupe A-(C-E).....	76
Figure 5.21 : Boxplot des performances de classification par k-NN	79
Figure 5.22 : Performances de classification du set (A-C-E) par MLP	82
Figure 5.23 : Performances de classification du set (A-E) par MLP	82

Abréviations

ANN	Artificial Neural Network (<i>en français</i> : Réseau de Neurone Artificiel)
CWT	Continus Wavelet Transform (<i>en français</i> : Transformée en Ondelette Continue)
DWT	Discret Wavelet Transform (<i>en français</i> : Transformée en Ondelette Discrète)
E	Energie
ECoG	Electrocorticographie
EEG	Electro-Encéphalogramme
Ent	Entropie
FDA	Fisher Discriminant Analysis (<i>en français</i> : Analyse discriminante de Fisher)
Gam	Gamme
ICA	Independant Component Analysis (<i>en français</i> : Analyse en Composante Indépendante)
iEEG	intracrânien EEG

KKT	KarushKhun-Tucker
K-NN	K-Nearest Neighbors (<i>en français</i> : K-Plus Proches voisins)
Max	Maximum
MEG	Magnétoencéphalographie
Min	Minimum
MLP	Multi Layer Perceptron (<i>en français</i> : Perceptron Multi-Couche)
OPF	Optimum Path Forest
PCA	Principal Component Analysis (<i>en français</i> : Analyse en Composante Principale)
RBF	Radial Basis Function (<i>en français</i> :Fonction de Base Radiale)
RNN	Recurrent Neural Network (<i>en français</i> : Réseau de Neurone Récurent)
Std	Standard deviation (<i>en français</i> : Ecart-type)
STFT	Short Fourier Transform (<i>en français</i> : Transformée de Fourier Fenêtrée)
SVM	Support Vector Machine (<i>en français</i> : Machine à vecteur support)
TEP	Tomographie par Émission de Positons
TVN	Spécificité
TVP	Sensibilité
WPT	Wavelet Packet Transform (<i>en français</i> : Transformée par paquet d'ondelette)

Liste des tableaux

Tableau 4.1 : Matrice de confusion	45
Tableau 5.1 : Fréquences correspondant aux différents coefficients	51
Tableau 5.2 : Fréquences correspondant aux coefficients de décomposition WPT	56
Tableau 5.3 : Résultats de mesures des différentes méthodes utilisées	65
Tableau 5.4 : Valeurs des précisions obtenues dans la littérature	67
Tableau 5.5 : Taux des performances de classification du groupe A- (C-E)	75
Tableau 5.6 : Résultats de mesures de classification par l'algorithme k-NN	78
Tableau 5.7 : Résultats de la classification par MLP pour A-(C-E).....	81
Tableau 5.8 : Résultats de classification par MLP pour (A-E)	81
Tableau 5.9 : Valeurs des précisions obtenues dans la littérature	84

*La dernière chose qu'on trouve en faisant un ouvrage
est de savoir celle qu'il faut mettre la première.
Pascal, Pensées.*

Introduction

Le cerveau humain est le centre de contrôle du corps humain; il reste l'organe le plus complexe et le plus mystérieux de tous les organes. Pendant la lecture de cette phrase, plusieurs millions de neurones sont activés et stimulés. Ces réactions chimiques se manifestent par une création et propagation de signaux électriques permettant une communication des différents membres du corps humain. Ces signaux peuvent être exploités grâce à des techniques comme : l'électroencéphalographie (EEG), l'électrocorticographie (ECoG), l'imagerie par résonance magnétique (IRM), l'imagerie par résonance magnétique fonctionnelle (IRMf) et la tomographie par émission de positons (TEP). Toutes ces techniques nous aident à acquérir une meilleure compréhension du comportement du cerveau afin que l'individu effectue un certain nombre de tâches. L'EEG est une technique de mesure des bio-potentiels provenant de l'activité électrique du cerveau. Les électrodes de mesure peuvent être placées sur le scalp, c'est le cas de l'EEG de surface, ou déposées sur le cortex cérébral dans l'espace sous-dural, c'est le cas de l'ECoG.

De par sa facilité d'utilisation, son excellente résolution temporelle et son faible coût, l'électroencéphalographie est une méthode non-invasive de référence dans le

milieu médical, plus précisément dans le diagnostic, le traitement des maladies mentales et neuro-dégénératives (Alzheimer, Parkinson, Apnée du sommeil....), et plus particulièrement dans le cas de l'épilepsie. Après les accidents vasculaires, l'épilepsie est la deuxième cause d'hospitalisation. Elle se caractérise par la survenue brusque et répétée de crises provoquées par la manifestation physique de décharges électriques, excessives et soudaines, générées par une population de neurones. Les recueils des tracés d'EEG peuvent être de longue durée (dépassant plusieurs dizaines d'heures), et chaque enregistrement peut contenir un nombre important de canaux (jusqu'à 256 canaux). Par conséquent, bien que l'enregistrement d'EEG de surface reste un examen incontournable pour le diagnostic et la classification des épilepsies, son analyse par les neurologues est longue, fastidieuse et reste subjective dans certains cas (car il n'existe pas de critères standards).

La problématique scientifique qui découle de cet objectif médical est donc le développement de systèmes automatiques permettant le traitement et la classification d'EEG de surface. La conception d'un système efficace de classification est complexe, étant donné la variabilité et la non-stationnarité des signaux EEG de différents sujets appartenant à différentes classes.

Afin de répondre à cet objectif, le travail présenté dans cette thèse consiste à développer différents algorithmes permettant d'extraire l'information significative à partir d'enregistrements de signaux EEG issus d'une base de données réelle prétraitée afin d'identifier les différentes catégories de données (EEG non épileptique, EEG avec crises épileptiques et EEG entre deux crises épileptiques) et d'évaluer les performances pour une classification de signaux épileptiques. En effet, nous avons développé trois techniques de classification avec différentes méthodes de prétraitement des enregistrements d'EEG de surface dans le cadre des crises d'épilepsie, afin d'identifier trois catégories de population (des personnes non malades, des patients épileptiques pendant la crise d'épilepsie et d'autres patients, entre deux crises épileptiques). Les performances d'évaluation de différentes techniques de classification ont été validées par des mesures statistiques permettant de décrire leur fiabilité.

L'état de l'art sera développé dans la deuxième section du chapitre un, mais de par ses qualités intrinsèques, la transformée en ondelette associée à l'analyse multi-résolution a eu un intérêt particulier pendant les deux dernières décennies, et de nombreux chercheurs se sont intéressés au problème de détection et de classification des signaux EEG épileptiques [95, 96]. Dans le travail proposé, les distributions fréquentielles et énergétiques des ondelettes discrètes et des paquets d'ondelette sont des paramètres essentiels afin de classer les signaux

d'enregistrement EEG. A partir des distributions fréquentielles et énergétiques, nous avons utilisé différents paramètres caractéristiques statistiques et non statistiques, ce qui représente une étape importante avant d'entamer la classification.

Différentes études ont montré que le signal EEG est un processus non stationnaire et des caractéristiques non linéaires ont été extraites à partir d'enregistrements de l'activité cérébrale pour des paramètres spécifiques des signaux [33, 69, 95, 97]. Ensuite, ces fonctions sont utilisées comme entrées des classificateurs [98]. Bien que les enregistrements des signaux EEG contiennent des informations pertinentes représentant la fonction du cerveau, les techniques de classification et d'évaluation de ces enregistrements ne sont pas encore bien développées, et par conséquent, il existe un besoin de mettre au point des systèmes automatiques de classification de signaux EEG. Dans la littérature, *Subasi* [60] a utilisé les coefficients de la transformée en ondelette discrète (DWT) obtenus à partir de segments d'EEG normaux et épileptiques comme entrée d'un réseau de neurone modulable appelé 'a mixture of expert model'¹. Pour les mêmes données d'enregistrement EEG, *Polat et al.* [51] utilisent les méthodes de réduction comprenant la DWT, la transformée de Fourier discrète et la transformée autorégressive. Dans [58] *Subasi et al.*, réduisent les dimensions des caractéristiques des coefficients de la DWT par l'analyse en composantes principales (ACP), l'analyse en composantes indépendantes (ICA) et l'analyse discriminante linéaire (LDA). Les caractéristiques obtenues ont été utilisées pour classer les signaux des EEG normaux et des EEG épileptiques en utilisant une classification par machine à vecteur de support. La transformée par paquet d'ondelette (WPT) apparaît comme une méthode prometteuse, et dans [12], *Wang et al.* ont appliqué l'entropie aux coefficients de la WPT pour extraire les paramètres caractéristiques utilisés comme entrée du classificateur *K*-plus proche voisin (K-NN).

Le premier chapitre a pour but de présenter une analyse électro-physiologique par les signaux électro-encéphalographiques. Pour cela, une rapide description du contexte médical ainsi que des méthodes de diagnostic de crises épileptiques sera donnée.

Afin de positionner notre travail, la problématique scientifique ainsi qu'une synthèse des travaux de recherche relatifs à notre application sont formalisées dans ce premier chapitre.

¹ 'Mélange d'un modèle expert'

Dans **le deuxième chapitre**, la représentation temps-fréquence à base de transformée en ondelette est décrite et nous associons à cela l'analyse multi-résolution qui permet l'analyse à plusieurs échelles.

Le troisième chapitre expose deux méthodes de réduction, l'analyse en composante principale et l'analyse discriminante linéaire, qui sont des approches nous permettant d'explorer la structure des données en tenant compte de leur caractère multidimensionnel afin de manipuler une quantité d'information réduite avec une perte minimale d'information.

La classification des crises épileptiques est une étape importante pour l'aide au diagnostic et **le chapitre quatre** est consacré aux fonctions de décision qui doivent être mises en œuvre. Notre choix s'est porté sur les machines à vecteurs support, les K-plus proches voisins et le perceptron multicouche dont nous rappelons tout d'abord la théorie, puis nous détaillons les algorithmes permettant leur implémentation.

Enfin, nous présentons, dans **le chapitre cinq**, les résultats expérimentaux obtenus à partir d'une base de données d'EEG réels contenant des crises épileptiques expertisées dans un cadre d'une aide au diagnostic médical. La validation de ces résultats est basée sur le pourcentage du taux de la précision de classification de paramètres calculés à partir des performances statistiques évaluées.

Nous terminons cette thèse par la conclusion et les perspectives relatives à ce travail.

*De par sa complexité, le cerveau suscite autant
d'interrogations que de fascination !!*

Contexte médical, état de l'art et problématique

I.1 Contexte médical

1.1 Introduction

Le cerveau est l'organe le plus complexe du corps humain. Il possède des centaines de milliards de neurones qui communiquent entre eux. De ce fait, depuis une trentaine d'années la recherche en neurosciences connaît une expansion abordant différents aspects scientifiques, médicaux, biologiques, psychologiques,..., pour la compréhension approfondie de cet organe.

Dans ce chapitre, nous allons d'abord décrire très brièvement la constitution et le fonctionnement du système nerveux chez l'être humain [1, 2, 3]. Ces notions seront utiles afin de comprendre d'où proviennent les signaux enregistrés par l'électro-encéphalogramme (EEG). Nous aborderons les bases anatomiques du

cerveau, la maladie neurologique, et plus particulièrement, l'épilepsie qui concerne les différents signaux d'enregistrements EEG analysés, traités, classés et utilisés dans notre travail, ainsi que les caractéristiques du signal EEG et la façon d'effectuer sa mesure afin d'établir un diagnostic.

1.2 Le cerveau

L'encéphale est illustré par la figure 2.1. Le cerveau possède deux moitiés relativement symétriques appelées des hémisphères, l'un à droite et l'autre à gauche (figure 2.1(a)). Toute la couche externe du cerveau, la substance grise qui présente des circonvolutions complexes en périphérie, est appelée le cortex cérébral. Le cortex de chaque hémisphère est divisé en quatre lobes (figure 2.1(b)): le lobe pariétal, le lobe occipital, le lobe frontal et le lobe temporal dont l'appellation est directement inspirée du nom des os crâniens qui les recouvrent [1, 2, 4].

Les lobes sont impliqués dans différentes fonctions :

- Les lobes occipitaux, localisés à l'arrière du cortex, sont spécialisés dans la vision.
- Les lobes pariétaux sont situés à l'avant des lobes occipitaux. Ils sont impliqués dans le sens du toucher, dans la détection de mouvements dans l'environnement et la localisation des objets dans l'espace.
- Les lobes frontaux sont impliqués dans le mouvement, l'attention, la planification, les compétences sociales, la pensée abstraite, la mémoire et certains aspects de la personnalité.
- Les lobes temporaux sont importants pour l'audition, le langage et la reconnaissance visuelle des objets.

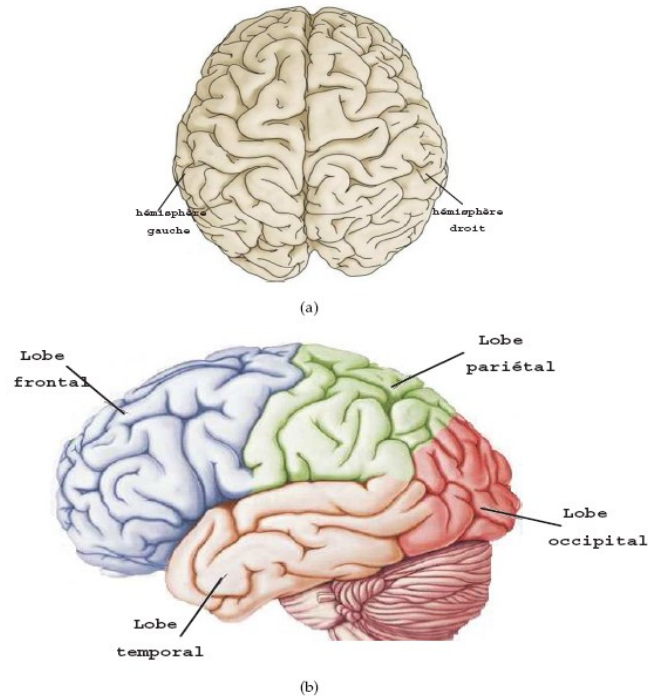


Figure 1.1 : Schéma des divisions de l'encéphale [2]: (a) : Hémisphères ; (b) : Lobes

1.2.1 Le neurone

Le neurone est l'unité cellulaire du système nerveux qui permet de traiter l'information au niveau du cerveau. A travers l'influx nerveux, les neurones peuvent assurer la génération, la communication et la transmission d'un signal.

A la base du système neuronal se trouvent des neurones interconnectés ou non (phase d'apprentissage) représentant les entités essentielles du système nerveux central. Ils forment un réseau pouvant aller jusqu'à cent milliards de cellules pour nous permettre de réaliser la totalité de nos activités conscientes ou non [7, 8, 10, 11]. Les réseaux se créent grâce à la propriété qu'ont les neurones à se connecter à d'autres via les dendrites, les axones et les synapses (figure 1.2).

Les échanges d'informations entre cellules sont réalisés au niveau des synapses. La synapse est l'unité structurale et fonctionnelle de connexions entre deux neurones. La membrane des neurones possède la propriété d'être excitable.

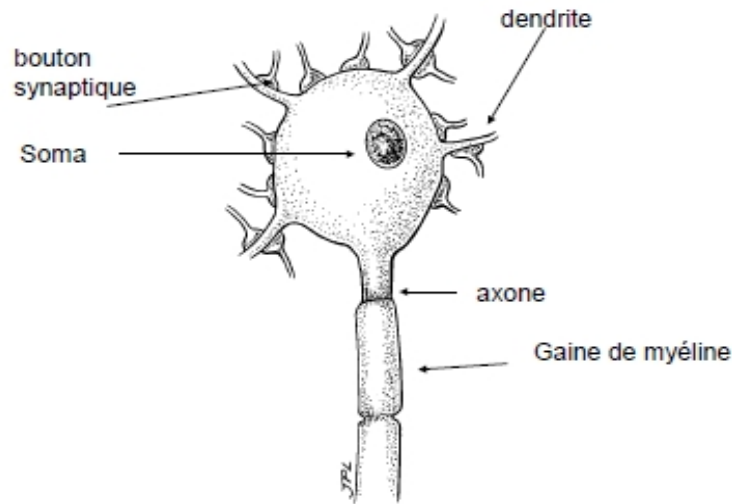


Figure 1.2 : Représentation d'un neurone. Reproduit avec autorisation.

Deux mécanismes sont à l'œuvre pour la communication neuronale :

- La transmission de l'influx nerveux le long des fibres nerveuses par le potentiel d'action.
- La transmission synaptique qui fait le lien entre deux neurones.

En se fondant sur le mécanisme de transmission utilisé par les synapses, on peut en distinguer deux types : les synapses électriques (transmission de l'information électriquement) et les synapses chimiques (transmission de l'information par le biais de neurotransmetteurs) [8, 10, 14].

1. 3 Transmission de l'influx nerveux

1.3.1 Transmission chimique : potentiel d'action

La transmission de l'influx nerveux est effectuée par des agents chimiques (neurotransmetteurs) via des synapses dits chimiques. Son principe général est illustré à la figure 1.3.

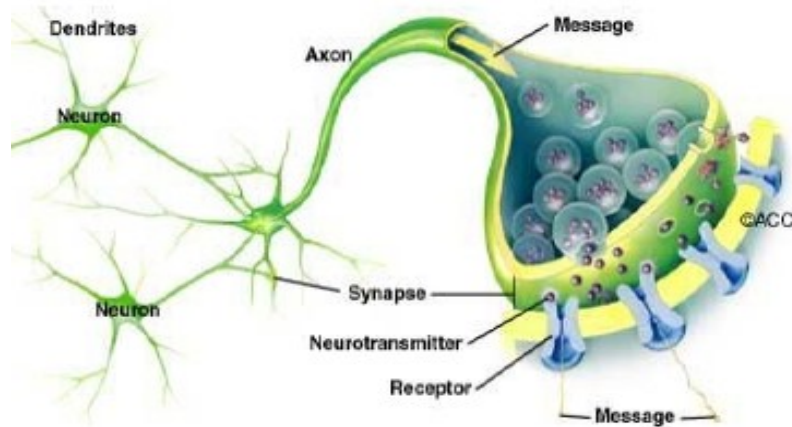


Figure 1.3 – Transmission chimique - Reproduite avec autorisation.

Synapse : C'est une région de contact entre deux neurones ou entre un neurone et une cellule effectrice (muscle ou glande). La cellule émettrice est dite pré-synaptique, la cellule réceptrice est quant à elle dite post-synaptique, les deux étant séparées sur une distance allant de 20 à 40 nm par l'espace inter-synaptique (ou fente synaptique) [9, 10].

Potentiel de repos : Au repos, il y a une différence de concentration d'ions des deux côtés de la membrane plasmique de la cellule neuronale : la charge est positive à l'extérieur (cations de sodium) tandis qu'elle est négative à l'intérieur (anions de potassium) ; cette distribution résulte en un potentiel dit de repos.

La pompe de sodium-potassium maintient cet état et réalise un transport actif (grâce à des protéines et à la consommation d'énergie), allant à l'encontre du gradient de concentration.

Potentiel d'action : Le potentiel d'action est une inversion de la polarité le long de la membrane plasmique durant quelques millisecondes suite à une stimulation de la cellule nerveuse.

La genèse a lieu au niveau du cône d'émergence (à l'origine de l'axone,

contient un grand nombre de canaux). Ce cône fait une synthèse de tous les potentiels aux niveaux des synapses dendrites/soma, et génère un potentiel d'action qui se déroule comme suit :

1. Ouverture des canaux de sodium, entrée d'ions Na^+ et génération du potentiel d'action ;
2. La dépolarisation suffit pour ouvrir tous les canaux ;
3. Ouverture des canaux de potassium (K^+), plus de gradient électrochimique ;
4. Ouverture des canaux de rectification du sodium, repolarisation progressive du neurone.

Le retour à l'état de repos et la repolarisation ont lieu grâce au travail de la pompe active évoquée plus haut.

La transmission du potentiel d'action autorise la libération de vésicules contenant des neurotransmetteurs dans la fente synaptique. Ces derniers se lient alors à l'élément post-synaptique au niveau de récepteurs prévus à cet effet et influent sur son potentiel, avec soit un effet inhibiteur PPSI (Potentiel Post-Synaptique Inhibiteur) soit un effet exciteur PPSE (Potentiel Post-Synaptique Excitateur), le premier inhibe la transmission de l'influx nerveux depuis la cellule post-synaptique vers d'autres cellules comme son nom l'indique, le second ayant un effet opposé [7, 11, 14].

1.3.2 Transmission électrique

Ce type de transmission est assuré par la circulation d'ions et de molécules. Il se distingue par sa rapidité en comparaison avec la transmission chimique.

Cette jonction permet le passage bidirectionnel de molécules d'un diamètre inférieur à 1.5 à 2 nm. Il s'agit en particulier de seconds messagers jouant le rôle d'intermédiaires et qui sont produits en réponse à des messagers primaires (hormones, neurotransmetteurs).

1.4 L'Épilepsie

A la fin du XIXe siècle, l'épilepsie prend sa place dans la famille des maladies neurologiques et c'est en 1873 que Hughlings Jackson a défini la crise épileptique comme la survenue soudaine d'une décharge excessive et rapide d'une population plus ou moins étendue de neurones qui constituent la substance grise de

l'encéphale. Il émet aussi l'idée que le type de crise varie en fonction de la localisation et de la nature de ces décharges.

Selon l'Organisation Mondiale de la Santé [5], l'épilepsie est une affection chronique du cerveau qui touche toutes les populations du monde. Elle se caractérise par des crises récurrentes qui sont la manifestation physique de décharges électriques excessives et soudaines, généralement brèves, dans un groupe de cellules cérébrales (neurones). Ces décharges peuvent avoir lieu dans différentes parties du cerveau. Les crises peuvent varier en intensité, allant de brèves pertes d'attention ou de petites secousses musculaires à des convulsions sévères et prolongées (cas du grand mal) [6]. Leur fréquence est également variable.

L'épilepsie affecte environ 50 millions de personnes dans le monde et selon la « Ligue Algérienne contre l'Epilepsie », cette maladie touche environ 350 000 à 400 000 personnes en Algérie ; soit 1 % de la population algérienne. Une fois sur deux, cette crise peut préjuger d'une maladie épileptique qui est, après la migraine, la première raison de consultation chez les neurologues.

1.4.1 Type de crises

Pendant les crises d'épilepsie, des changements brefs et soudains de l'activité électrique du cerveau se produisent. Cette décharge électrique anormale peut avoir lieu soit dans une zone précise du cerveau, soit dans son ensemble. La distinction entre crises partielles et crises généralisées repose sur le caractère localisé ou diffus de cette décharge à son origine. Pour certaines, les recherches se placent au niveau de l'évaluation de l'épilepsie d'un patient, le plus souvent dans le cadre d'un bilan pré chirurgical. Dans ce cas, les neurologues classent les crises suivant des critères établis en 1981 par la « International League Against Epilepsy » [5]. Cela permet d'avoir une idée précise du type de crise et le cas échéant, de savoir si un traitement chirurgical est possible afin de déterminer ses chances de réussite sur le long terme.

Lorsque la décharge est localisée, les signes sont en rapport avec les fonctions de la zone cérébrale concernée. De ce fait, les crises sont classées en deux types:

- ✓ **Les crises partielles** n'affectent que certaines parties du cerveau, et donc des commandes cérébrales. Elles peuvent se traduire par des troubles moteurs, des troubles sensoriels et sensitifs, des troubles de la mémoire ou

de la conscience. Certaines de ces crises partielles peuvent évoluer vers une crise tonico-clonique.

✓ **Les crises généralisées**, parmi lesquelles :

- les crises tonico-cloniques (appelées aussi Grand mal) : ce sont les plus connues, les plus impressionnantes mais non les plus fréquentes. Elles se manifestent par une perte de connaissance avec chute, mouvements convulsifs, morsure de la langue, etc....

- les absences : elles se manifestent par une brève rupture de contact (quelques secondes) se traduisant par une fixité du regard. Elles sont parfois accompagnées de mâchonnements ou de gestes involontaires et inadaptés appelés automatismes. Les absences peuvent se répéter d'une manière fréquente au cours de la journée.

L'enregistrement des données électro-encéphalographiques est l'unique moyen de mettre en évidence, directement, l'activité épileptique. En effet, l'électroencéphalogramme permet d'enregistrer de manière directe l'activité électrique produite au niveau des neurones. Cet examen est incontournable pour les neurologues afin d'établir un diagnostic et une classification des épilepsies. Deux états distincts peuvent être distingués dans l'EEG du patient épileptique : l'état ictal, c'est-à-dire la crise, et l'état inter-ictal qui est le tracé entre les crises.

1.5 L'Électro-encéphalogramme

L'histoire de l'électro-encéphalographie commence avec le biologiste Richard Caton qui, en 1875, détecta chez le singe et le lapin la présence de courants électriques attestée par des oscillations du galvanomètre. L'invention de l'électro-encéphalographie fût attribuée au psychologue allemand Hans Berger, qui commença à l'étudier dans les années 1920 en l'appliquant à l'homme, et enregistra le premier signal d'activité cérébrale sous forme de variations permanentes de potentiel enregistrées avec des électrodes impolarisables appliquées sur une lacune crânienne, ou à la surface du crâne intact. Ses travaux furent repris et complétés par le britannique Edgar Douglas Adrian, qui obtint en 1932 le prix Nobel de physiologie.

L'inscription à jet d'encre, introduite par Grass en 1935, a permis de visualiser les activités électriques sur papier. Les bases de l'examen furent posées dès 1945 et sont toujours appliquées aujourd'hui [1, 6, 7].

Ce n'est qu'au début des années 1960 que le développement de l'électro-encéphalographie en clinique médicale s'est répandu. Les électro-encéphalogrammes sont actuellement utilisés afin de diagnostiquer des foyers

épileptiques, des accidents vasculaires cérébraux, des tumeurs cérébrales, des lésions, Alzheimer et Parkinson... Ils permettent aussi de trouver l'origine de migraines, de problèmes d'étourdissements, ou encore de somnolence et d'apnée du sommeil.

Comparé aux autres techniques d'investigation en neurologie, l'EEG est le moyen le plus utilisé pour mettre en évidence l'activité électrique du cerveau. Il fournit des informations en temps réel avec une très bonne résolution temporelle de l'ordre de la dizaine de millisecondes (figure 1.4).

L'EEG du scalp est un examen indolore et non-invasif qui renseigne sur l'activité neurophysiologique du cerveau au cours du temps et en particulier du cortex cérébral. Cet enregistrement de l'activité électrique du cerveau se fait à l'aide d'électrodes placées sur le cuir chevelu suivant un système international 10-20 [7, 9, 15] (figure 1.5).

Le signal électrique à la base de l'EEG est la résultante de la sommation des potentiels d'action post-synaptiques synchrones issus d'un grand nombre de neurones. Il en existe néanmoins d'autres techniques d'investigation en neurologie, comme expliquées dans [1,11]:

- L'ECOG (électrocorticographie) est la technique d'enregistrement qui consiste à capter l'activité électrique neuronale en posant directement les électrodes sur le cortex du patient. Cette méthode est généralement utilisée en prévision d'opérations chirurgicales. La résolution spatiale est faible car les électrodes sont concentrées à un endroit bien précis.
- Le iEEG (intracrânien EEG) consiste dans l'implantation de sondes dans le cerveau même, par le neurochirurgien. La résolution spatiale est également faible.
- La MEG (la magnétoencéphalographie) est une technique permettant de capter le champ magnétique émis par l'activité électrique du cerveau. Elle nécessite un appareillage lourd et coûteux. C'est une technique souvent utilisée de pair avec l'EEG.

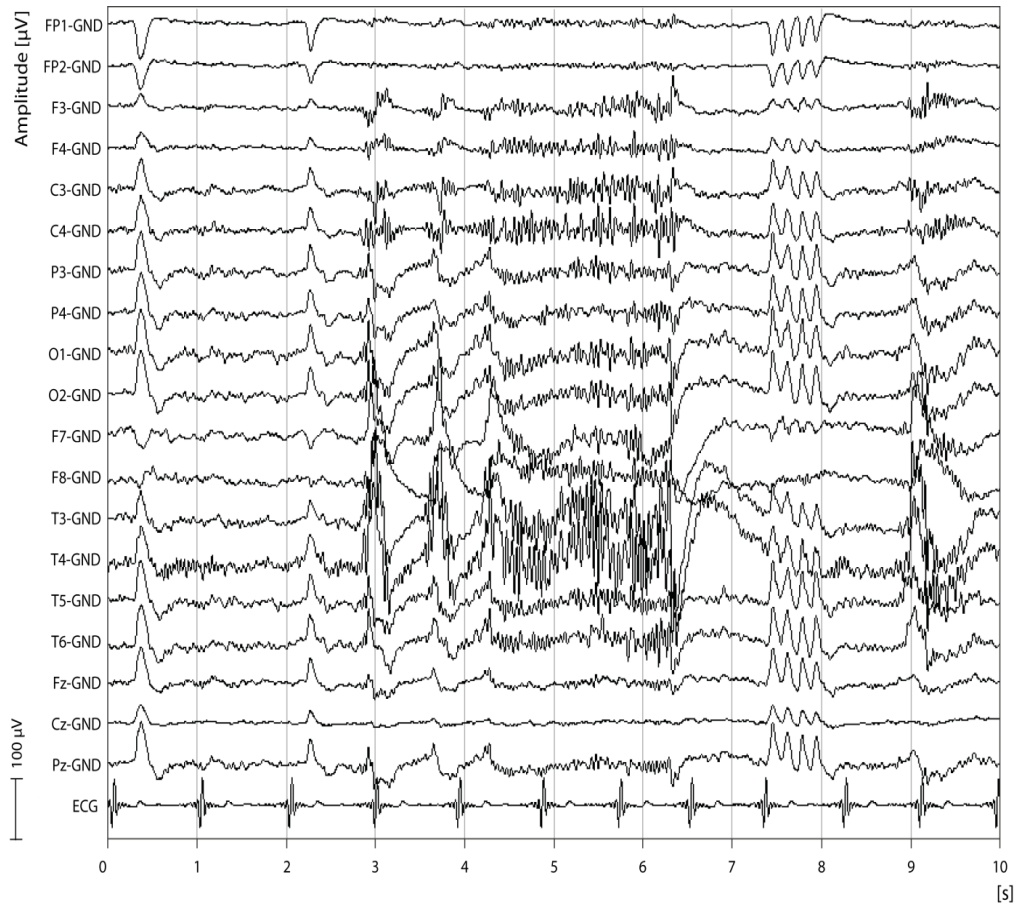


Figure 1.4 : Exemple d'un enregistrement d'EEG.

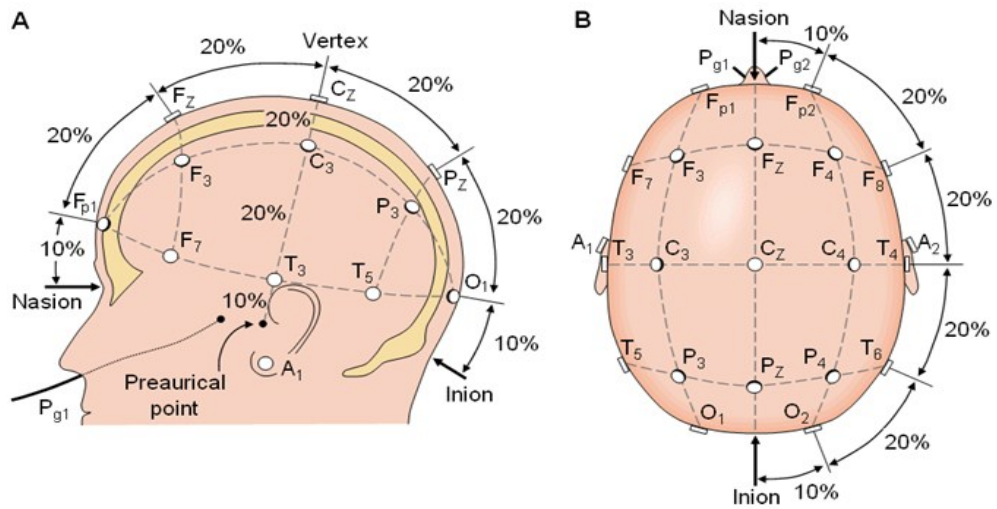


Figure 1.5 : Système (10 – 20) : Placement de 21 électrodes par rapport au nasion et à l'inion [15].

1.5.1 Activités électriques cérébrales rythmiques

La contribution au signal de chaque neurone est extrêmement faible et le signal doit traverser plusieurs couches de tissu non neuronal avant d'atteindre les électrodes (figure 1.6). Une population de neurones ne peut produire un signal d'amplitude mesurable à la surface du scalp que si les activités de ces éléments se somment de façon constructive, c'est-à-dire si une majorité de neurones se décharge en phase. Les oscillations des signaux EEG reflètent donc nécessairement une synchronisation oscillatoire sous-jacente. En conséquence, l'amplitude du signal EEG dépend fortement du degré de synchronisation de l'activité des neurones. Les signaux rythmiques de l'EEG sont souvent décrits en termes d'amplitude relative pour montrer le degré de synchronisation ou de désynchronisation cérébrale sous-jacente [7].

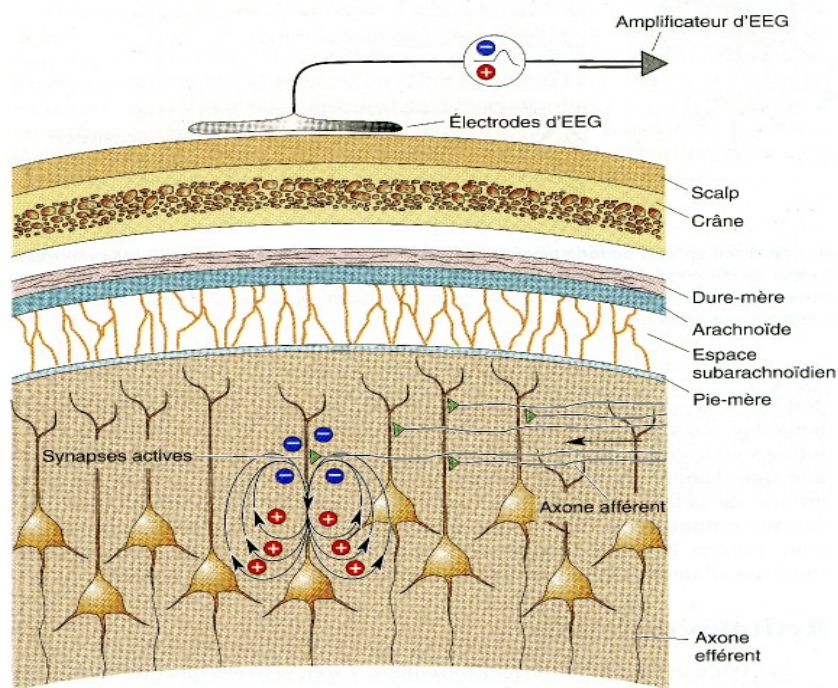


Figure 1.6 : Champs électriques générés par des courants synaptiques [7].

Dans ce cas, la synapse excitatrice se trouve dans la partie supérieure de la région dendritique. Lors de la transmission synaptique, des courants entrants positifs pénètrent dans la dendrite, conduisant à rendre le milieu extracellulaire légèrement négatif. L'EEG mesure les courants générés au niveau du dipôle électrique au

travers des différentes couches de tissus. Il est nécessaire que des milliers de cellules corticales soient activées en même temps pour atteindre une variation de potentiel détectable en surface [7, 12, 13].

L'analyse de l'activité électrique du cerveau permet d'observer des signaux contenus dans des bandes de fréquences données et correspondant à des états et des zones du cerveau déterminés. Il s'agit des rythmes cérébraux classés selon leurs bandes de fréquence :

1. Rythme **alpha**: ondes de 8 à 13Hz, caractéristique d'un adulte en veille diffuse (éveillé avec les yeux fermés), les plus fortes amplitudes sont enregistrées au niveau des zones occipitales et pariétales. Ces amplitudes (de 10 à 100 μ V) diminuent, en général, lorsque le sujet ouvre les yeux ;
2. Rythme **bêta** : ondes de 13 à 30Hz, localisées au niveau des zones frontales. Elles correspondent tout autant à une activité normale ou un effort mental intense qu'à la phase du sommeil paradoxal. D'amplitudes faibles (inférieure à 20 μ V), ces rythmes peuvent être masqués par les rythmes alpha (plus énergétiques).
3. Rythme **gamma** : à 40Hz, correspondent à une forte sollicitation du cerveau et à des processus cognitifs tels que la perception, le langage ou les émotions.
4. Rythme **delta**: en-dessous de 4Hz, ondes caractéristiques du sommeil profond. Dans les épilepsies partielles, elles sont fréquentes et souvent localisées dans la même région que le foyer épileptique inter-critique.
5. Rythme **thêta** : de 4 à 7Hz. D'amplitudes de 10 à 50 μ V, et localisé dans les zones temporales, il est lui aussi caractéristique du sommeil profond. Il tend d'ailleurs à apparaître de concert avec la diminution de l'amplitude des ondes alpha.

1.5.2 Mesure des signaux EEG

Plusieurs difficultés existent lors de l'acquisition d'un signal EEG avec des électrodes de surface :

- La faible amplitude du signal mesuré (de 10 à 200 μ V) causées par la mauvaise conduction du crâne.
- Le bruit de mesure dû à l'éloignement des sources.
- La présence d'artéfacts (connus sous le nom d'Electromyogrammes et d'Electro-oculogrammes) dont l'amplitude est supérieure à celle des signaux EEG.

Conclusion

L'EEG est utile au diagnostic et au suivi de différentes pathologies encéphaliques en montrant des anomalies différentes selon le stade de la maladie, certaines étant très évocatrices comme l'épilepsie. Il faut rappeler que si l'EEG est à la base de l'étude des différentes crises d'épilepsie, il n'en est pas moins extrêmement utile dans l'étude de nombreuses pathologies (maladies du sommeil, certaines démences, Alzheimer...), et plus généralement, dans les études cognitives du cerveau. L'électro-encéphalogramme est devenu un procédé standard en neurosciences. Par rapport à d'autres examens, il présente les avantages suivants :

- Résolution temporelle de l'ordre de la milliseconde, elle est, en générale, de l'ordre de la seconde pour les techniques d'imagerie ;
- Examen non-invasif, i.e. aucun passage à travers les tissus cutanés n'est nécessaire.
- Coût contenu.

I.2 État de l'art et problématique

Nous décrivons un état de l'art, non exhaustif, concernant le traitement, l'analyse, l'exploitation et la classification des signaux d'enregistrement EEG, particulièrement pour le cas de l'épilepsie. Nous y définissons, succinctement, les différentes méthodes de traitement les plus communément utilisées, l'extraction des paramètres caractéristiques ainsi que leurs classifications étudiées par la communauté scientifique.

Nous allons donc décrire les axes majeurs de recherche ayant trait à l'EEG et en se limitant au cas de l'épilepsie.

L'EEG de surface est l'examen le plus utilisé par les neurologues pour l'analyse et l'interprétation des signaux enregistrés sur des épileptiques, et par conséquent, le plus étudié par la communauté scientifique.

Le suivi de l'activité du cerveau des patients atteints d'épilepsie, à travers l'EEG, montre deux catégories d'activité anormale : des signaux inter-critiques, qui sont des signaux enregistrés entre deux crises d'épilepsie, et des signaux critiques, qui sont des signaux enregistrés pendant la crise d'épilepsie. De nombreuses méthodes de traitement du signal ont été mises en œuvre pour étudier les EEG enregistrés chez les épileptiques, des renseignements sur la maladie peuvent être obtenus aussi bien aux périodes critiques, ainsi qu'aux périodes inter-critiques. Lors de cette dernière, le signal EEG peut comporter des anomalies inter-critiques qui permettent d'identifier le type de crise et même la localisation du foyer épileptogène [35]. Mais avec l'avènement des enregistrements d'EEG ambulatoires, les patients ont eu des enregistrements dans leur environnement, sur une longue durée (supérieure à vingt-quatre heures). De nombreuses études ont montré que les systèmes portables d'enregistrements d'EEG ont augmenté le rendement d'un diagnostic utile afin d'améliorer la prise en charge du patient [36, 37]. Cependant, l'exploration visuelle de ces enregistrements par des médecins neurologues expérimentés devient un travail fastidieux, en particulier pour les

enregistrements de longue durée. Et un désaccord entre neurophysiologistes peut être possible sur le même enregistrement, en raison de la nature subjective et de la variété des pointes inter-critiques et de leurs morphologies. Et d'autre part, le signal EEG de surface est perturbé par de nombreux artefacts (les clignements des yeux, l'activité musculaire, l'électrocardiogramme, interférences électriques, les électrodes,...) existants sur le scalp, c'est pour cela que différentes équipes de recherche ont été mises en place afin d'étudier les nombreuses méthodes de débruitage en se basant sur l'analyse en composantes indépendantes [38, 39], ainsi que les méthodes de débruitage basées sur la transformée en ondelette [40, 41, 42].

En mettant à part, le prétraitement du signal et la modélisation de l'activité électrique épileptique, nous avons deux grandes tendances qui émanent de l'ensemble des recherches consacrées à l'EEG épileptique : celles qui cherchent à détecter ou à prédire les crises et celles qui tentent de les classer en vue d'une future intervention chirurgicale. C'est dans cette dernière section que nous allons positionner nos travaux de recherche par rapport à la communauté scientifique qui s'intéresse au traitement du signal des différents types d'EEG (de surface, invasif ou semi-invasif).

L'ensemble des méthodes développées, pour aboutir à un système de classification automatique des signaux EEG, utilisent tout le panel des réseaux de neurones aux ondelettes en passant par les méthodes temps-fréquence, ainsi que les modèles cachés de Markov ou la mesure de la non-linéarité dans les enregistrements EEG.

En effet, l'automatisation de l'analyse des crises épileptiques se réfère aux trois grands axes :

- ✓ Détection des crises
- ✓ Prédiction des crises, et
- ✓ Localisation de l'origine des crises.

Dans la littérature, de nombreux algorithmes de détection de crises d'épilepsie ont été proposés en utilisant des méthodes classiques de traitement de signal [44]. Toutes ces méthodes suggérées ont pour objectif de détecter les différentes formes qui sont l'indication d'une manifestation d'une crise d'épilepsie sur les enregistrements d'EEG. L'ensemble du processus de méthodes développées pour la détection automatique de crises d'épilepsie peut être divisé en deux principales étapes : a) extraction des caractéristiques, et b) classification.

a) La sélection des paramètres caractéristiques, discriminants et pertinents, est la base des méthodes de détection des crises épileptiques. Certains auteurs se basent sur les phénomènes physiologiques détectés au niveau du scalp. Pour mesurer cela, les auteurs utilisent : la détermination de la fonction d'autocorrélation [45], la vraisemblance de la synchronisation [46], ou la synchronisation de phase du plus proche voisin [47]. D'autres auteurs prennent en considération les caractéristiques morphologiques des enregistrements d'EEG épileptiques où les décharges rythmiques sont visibles sous formes de plusieurs pointes [48] en prenant l'amplitude et la durée des demi-ondes épileptiques.

Pour détecter les décharges rythmiques, les auteurs [49, 50, 51] se sont basés sur le calcul de la transformée de Fourier rapide, dans [52, 53, 54, 55, 56] ils utilisent le domaine fréquentiel et dans [33, 34, 57, 58] la représentation temps-fréquence; Les caractéristiques extraites des coefficients de décomposition d'ondelettes ont été utilisées par [59, 60, 61, 62, 63, 64, 65, 66]. La transformée en ondelette discrète a été utilisée dans [74, 75, 83, 91, 92, 98], la transformée en paquet d'ondelette dans [93] et la transformée en ondelette multiple dans [63].

Des paramètres statistiques ont été calculés dans [58, 59, 60, 94], des caractéristiques non linéaires et d'ordre supérieur dans [44, 55], et d'autres caractéristiques obtenues à partir des méthodes de vecteurs propres [54], de l'entropie [52, 63, 64, 68] et de fonction de cross-corrélation [56, 67], ont été proposées pour caractériser le signal EEG. Ainsi, différentes caractéristiques ont été extraites dans le but de détecter des crises d'épilepsie, mais ces paramètres caractéristiques n'ont pas de signification physiologique.

b) Une fois qu'un ensemble de paramètres a été obtenu pour caractériser un enregistrement EEG, une ou plusieurs méthodes de classification sont appliquées afin de décider si cet enregistrement fait partie d'une crise épileptique ou non. Dans la littérature, les chercheurs ont utilisé tout le panel des méthodes de classification, des classificateurs linéaires [56, 58, 62] aux réseaux de neurones artificiels [33,34, 52, 53,69].

Les méthodes de classification comme celle des machines à vecteur support utilisées dans les références [54, 58, 66, 67, 70, 81, 87, 89], des k-plus proches voisins [34, 56, 65], d'arbre de décision et des classificateurs « Naive Bayes » [34, 49, 56], les réseaux de neurones artificiels [35, 52, 53, 61, 63, 69, 82, 84, 85] ainsi que le mélange de modèle expert [60] ont été étudiées.

Espaces de représentation à base d'ondelettes

Les ondelettes sont des outils, relativement, récents de traitement du signal permettant l'analyse, à plusieurs échelles en temps et en fréquence, des signaux électro-physiologiques. Nous nous intéressons aux prétraitements des signaux d'enregistrement EEG qui sont de nature non-stationnaire et, par conséquent, seront analysés afin d'en extraire l'information pertinente. Leur représentation temps-fréquence utilisera les transformées en ondelettes discrètes (DWT)², ou bien, la décomposition en paquets d'ondelettes (WPT)³.

1. Représentation temps-fréquence

1.1 Principe d'Heisenberg-Gabor

Ce principe est décrit par l'équation (2. 1), dite d' « Heisenberg-Gabor », reliant la durée de l'observation σ_t à la largeur de bande σ_ω d'un signal :

$$\sigma_t \sigma_\omega \geq \frac{1}{2} \quad (2.1)$$

où ω est la pulsation.

² DWT :discret wavelet transform.

³ WPT :wavelet packet transform.

1.2 Transformée de Fourier fenêtrée

Afin d'analyser un signal dont les caractéristiques fréquentielles varient au cours du temps, on pourrait essayer d'appliquer la transformée de Fourier à des portions du signal, afin de localiser les fréquences contenues dans des fenêtres temporelles. C'est la transformée de Fourier fenêtrée ou STFT⁴ qui correspond à une projection du signal sur un espace de fenêtres translatées et modulées. La STFT est définie par :

$$STFT(f, \tau) = \int_{-\infty}^{+\infty} f(t) e^{-j2\pi t \tau} W(t - \tau) dt \quad (2.2)$$

où $W(t)$ est la fonction fenêtre et τ le paramètre de translation.

La STFT permet de représenter le signal dans le plan temps-fréquence. Néanmoins, cette représentation est très redondante d'une part, et d'autre part, elle ne permet pas d'obtenir la résolution maximale donnée par le principe d'Heisenberg. Si l'on cherche à établir une représentation temps-fréquence d'un signal, il est donc très souhaitable que les résolutions de cette représentation en temps et en fréquences soient liées par la relation (2.1), une petite résolution temporelle (σ_t petit) pour les hautes fréquences, et une grande résolution en temps (σ_t grand) pour les basses fréquences. Cette exigence n'est pas facile à satisfaire pour la transformée de Fourier par fenêtres glissantes. C'est une des raisons pour lesquelles la transformée en ondelettes s'est rapidement imposée comme outil de représentations temps-fréquence [16, 18].

1.3 Transformée en ondelettes continue

Une fonction est appelée ondelette si elle satisfait la condition suivante dite d'admissibilité :

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (2.3)$$

$$\int_{\mathbb{R}^+} \frac{|\hat{\psi}(w)|^2}{|w|} dw = \int_{\mathbb{R}^-} \frac{|\hat{\psi}(w)|^2}{|w|} dw < +\infty \quad (2.4)$$

La transformée continue en ondelette, ou CWT⁵, est une extension de l'analyse de Fourier. Elle remplace la sinusoïde de la Transformée de Fourier par une famille de translation et dilatation d'une même fonction. Cette fonction, nommé ondelette mère Ψ est une fonction de fenêtrage, oscillante, de dimension finie.

⁴STFT : Short Fourier Transform

⁵CWT : Continus Wavelet Transform

De plus, la fonction ondelette doit être de carré sommable : elle a donc une durée finie, contrairement à l'exponentielle complexe qui intervient dans la transformation de Fourier. Par dilatations et translations de l'ondelette Ψ [7, 8, 16, 92], on définit les atomes de la transformée en ondelettes. Pour toute échelle $a \in \mathbb{R}^{+*}$ et toute position $b \in \mathbb{R}$, on définit un atome de la transformée par :

$$\Psi_{a,b}(t) = n \Psi\left(\frac{t-b}{a}\right) \quad (2.5)$$

Où n est un facteur de normalisation (égal à $1/\sqrt{a}$), b est le facteur de translation et a est le facteur de dilatation. La dilatation change l'étendue temporelle de l'ondelette ($a > 1$ augmente son étendue temporelle, et $a < 1$ la réduit), tandis que la translation localise l'ondelette dans le temps ; la CWT notée $W_f(a, b)$ de $f(t)$ est alors une fonction des deux variables a et b définie par le produit scalaire :

$$W_f(a, b) = \int_{-\infty}^{+\infty} f(t) \Psi_{a,b}^* dt \quad (2.6)$$

où $\Psi_{a,b}^*$ est la fonction complexe conjuguée de $\Psi_{a,b}$.

Cette représentation du signal nous donne une fonction redondante du signal. Tout comme la transformée de Fourier, la transformée en ondelette conserve l'énergie (équation (2.7)) :

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |W_f(a, b)|^2 d\tau \frac{da}{a^2} = \int_{-\infty}^{+\infty} |f(t)|^2 dt \quad (2.7)$$

Les spectrogrammes, représentés sur la figure 2.1, montrent la distribution énergétique d'un signal EEG réel (extrait de notre base de données). L'application de la transformée en ondelette continue donne une représentation temps-fréquence en 2D sur la figure 2.1 (a), et l'espace des coefficients a été représenté en 3D sur la figure 2.1 (b).

2. Analyse multi-résolution

La théorie de l'analyse multi-résolution, développée par Mallat [16], Meyer [19], et Daubechies [20] donne un cadre formel à la construction de bases d'ondelettes présentant des propriétés d'orthogonalité.

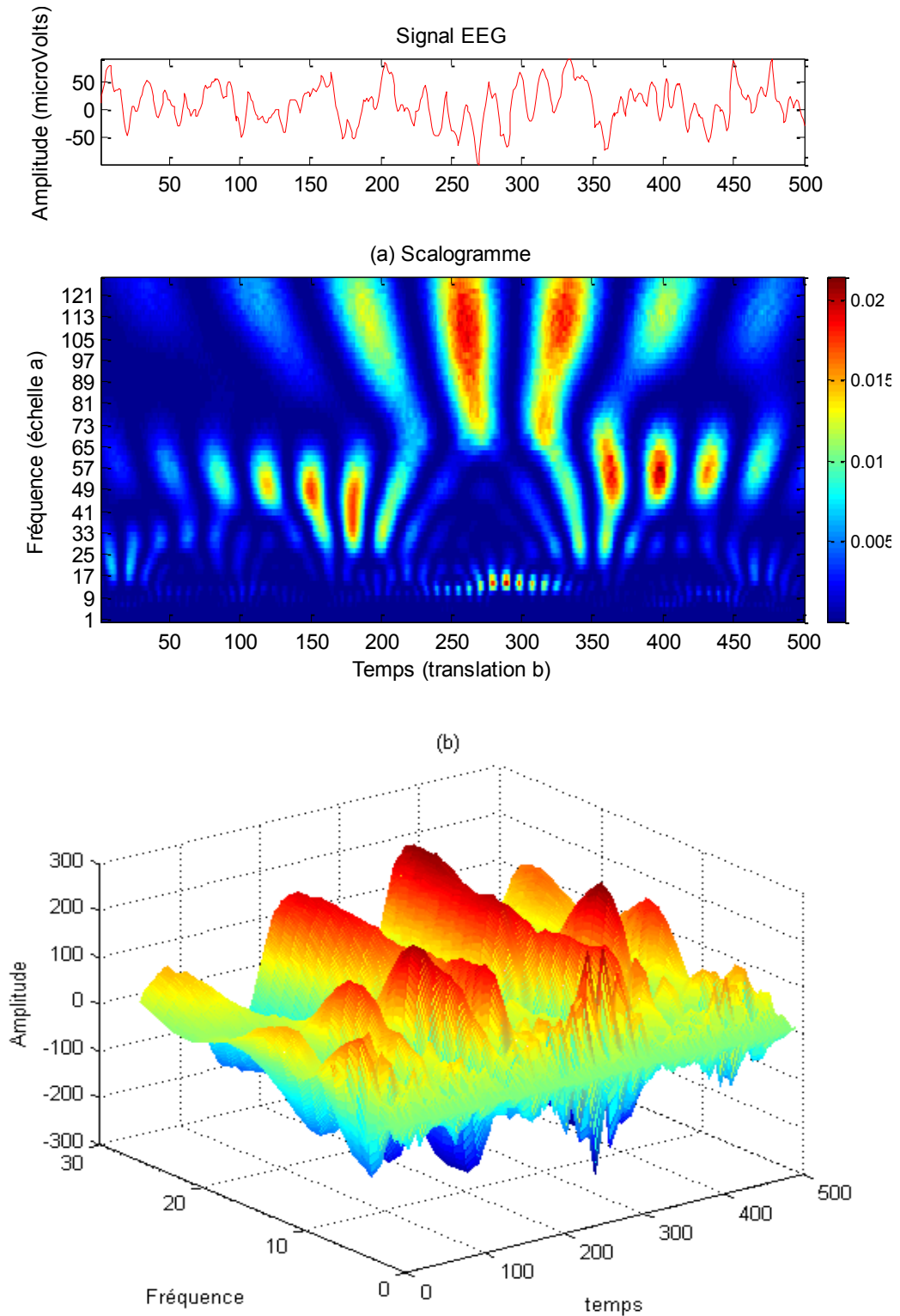


Figure 2.1 : Exemple d'un signal EEG à analyser à l'aide de la transformée en ondelette continue. (a) : Scalogramme (Pourcentage d'énergie de chaque d'ondelette) et (b) : représentation en 3 dimensions.

Elle permet notamment :

- la construction de base d'ondelettes $\{\psi_{j,k}\}_{(j,k) \in \mathbb{Z}^2}$ et leurs duales,
- la mise en œuvre pratique du calcul de la transformée en ondelette discrète (DWT) à base de bancs de filtres sans qu'il soit nécessaire de manipuler des produits scalaires.

Nous présentons dans cette partie les notions fondamentales sur la DWT qui sont utiles pour la définition et l'optimisation des espaces de représentation [12,16, 17, 18].

La DWT est une transformée non redondante qui décompose le signal $x(t)$ sur la base d'ondelettes discrètes correspondant à toutes les versions de ψ dilatées et décalées suivant une grille dyadique. Elle fournit un ensemble de coefficients :

$$DWT_x(j, k) = \langle x(t), \psi_{j,k}(t) \rangle \quad (2.8)$$

avec :

$$\psi_{j,k}(t) = 2^{-j/2} \psi\left(\frac{t-2^j k}{2^j}\right) \quad (2.9)$$

j et k correspondent, respectivement, aux indices d'échelle et de temps. L'ondelette $\psi_{j,k}$ analyse le signal à l'échelle 2^j . En termes fréquentiels, la DWT réalise un banc de filtre : l'analyse aux petites échelles met en évidence son comportement aux hautes fréquences. La longueur de bande des filtres d'analyse n'est pas uniforme mais correspond à un découpage dyadique de l'axe fréquentiel. La figure 2.2 donne une représentation de la décomposition, sur cinq niveaux, d'un signal $x(t)$ par la DWT.

3. Espace de représentation

- ✓ Les espaces d'approximation [18]: Une analyse multi-résolution de $\mathcal{L}^2(\mathbb{R})$ consiste en une suite croissante $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces fermés de $\mathcal{L}^2(\mathbb{R})$ tels que :

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad \bigcup_{j \in \mathbb{Z}} V_j = \mathcal{L}^2(\mathbb{R}) \quad \text{dans } \mathcal{L}^2(\mathbb{R}) \quad (2.10)$$

- La suite de $\{ V_j \}_{j \in \mathbb{Z}}$ est croissante et les V_j sont emboîtés les uns dans les autres :

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j \quad (2.11)$$

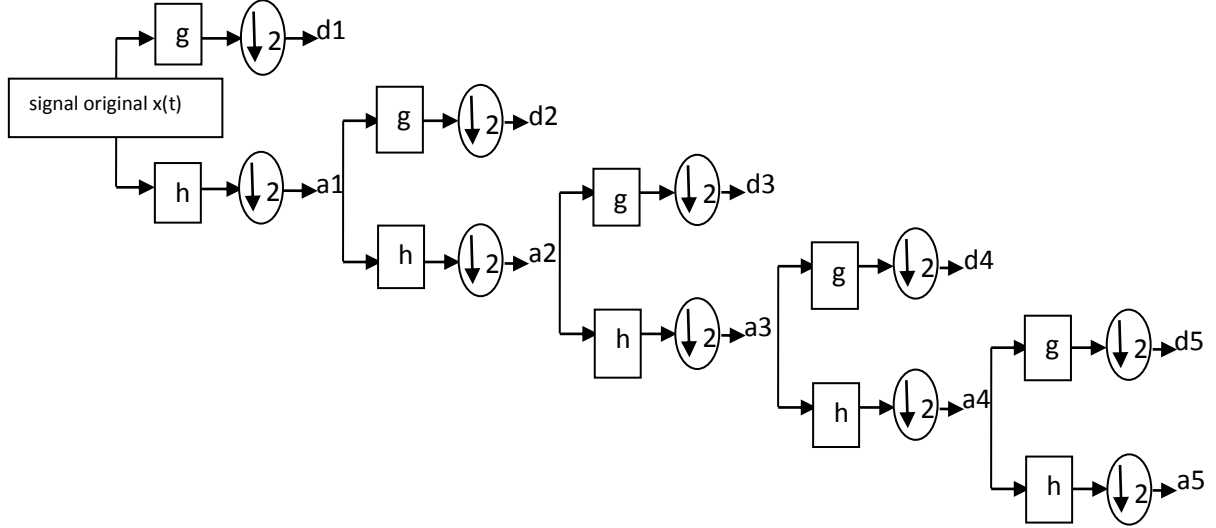


Figure 2.2 : Décomposition par transformée en ondelette discrète sur cinq niveaux. (d_j : coefficients de détails. a_j : coefficients approximations)

- La projection d'une fonction $f \in \mathcal{L}^2(\mathbb{R})$ sur un de ces V_j constitue une approximation plus ou moins grossière de f et on peut passer d'un espace à l'autre par changement d'échelle dyadique :

$$\forall f \in \mathcal{L}^2(\mathbb{R}), \forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1} \quad (2.12)$$

- Chaque espace est invariant par translation :

$$\forall f \in \mathcal{L}^2(\mathbb{R}), \forall (j, k) \in \mathbb{Z}^2, f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j \quad (2.13)$$

- Il existe une fonction ϕ telle que $\{\phi(t-n)\}_{n \in \mathbb{Z}}$, soit une base de Riesz⁶ de V_0 :

$$\forall x \in V_0, \quad x(t) = \sum_{k \in \mathbb{Z}} a[k] \phi(t - k) \quad (2.14)$$

Les propriétés vues précédemment, permettent de définir des familles $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$, bases de Riesz de chaque V_j où :

$$\phi_{j,k}(t) = 2^{-j/2} \phi\left(\frac{t-2^j k}{2^j}\right), (j, k) \in \mathbb{Z}^2 \quad (2.15)$$

On passe donc d'un espace d'approximation à l'autre, par changement d'échelle dyadique et chaque espace est invariant par translation. L'approximation d'un signal $x(t)$ au niveau j (i.e. sa projection sur un espace V_j) correspond à son analyse à la résolution (ou échelle) 2^j et est donnée par ses coefficients d'approximation :

$$a_j[k] = \langle x, \phi_{j,k} \rangle \quad (2.16)$$

En définissant la base duale $\{\tilde{\phi}_{j,k}\}_{k \in \mathbb{Z}}$ de la base $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$, le signal d'approximation $A_j(t)$ de $x(t)$ au niveau j est :

$$A_j(t) = \sum_k a_j[k] \tilde{\phi}_{j,k}(t) \quad (2.17)$$

- ✓ Les espaces détails : L'approximation A_j étant plus grossière que A_{j-1} , l'information perdue entre deux approximations successives constitue les signaux de détail :

$$D_j(t) = A_{j-1}(t) - A_j(t) \quad (2.18)$$

Les signaux de détails sont des projections du signal $x(t)$ dans des espaces $\{W_j\}_{j \in \mathbb{Z}}$, appelés espaces d'ondelettes, où W_j est le complémentaire de V_j dans V_{j-1} :

$$V_j + W_j = V_{j-1} \quad (2.19)$$

On peut construire pour chaque W_j une base d'ondelettes $\{\psi_{j,k}(t) = 2^{-j/2} \psi\left(\frac{t-2^j k}{2^j}\right)\}_{k \in \mathbb{Z}}$ par dilatation et translation d'une ondelette mère ψ , elle-même construite à partir de la fonction ϕ .

⁶ Une suite $\{\psi_j\}_{j \in \mathbb{Z}}$ est une base de Riesz [18] de $\mathcal{L}^2(\mathbb{R})$, s'il existe A et B tel que $A \sum_j |c_j|^2 \leq \|\sum_j c_j \psi_j\|^2 \leq B \sum_j |c_j|^2$ avec c_j quelconques. Dans le cas d'une base orthonormée (base hilbertienne), nous avons $A = B = 1$ et $\|\sum_j c_j \psi_j\|^2 = \sum_j |c_j|^2$ (égalité de Parseval).

A partir des ondelettes $\psi_{j,k}$, on définit les coefficients de détail d'un signal $x(t)$ au niveau j par :

$$d_j[k] = \langle x, \psi_{j,k} \rangle = \text{DWT}_x(j, k) \quad (2.20)$$

En définissant la base duale $\{\tilde{\psi}_{j,k}\}_{k \in \mathbb{Z}}$ de la base $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$, le signal de détail $D_j(t)$ de $x(t)$ au niveau j est :

$$D_j(t) = \sum_k d_j[k] \tilde{\psi}_{j,k}(t) \quad (2.21)$$

On peut écrire $\phi\left(\frac{t}{2}\right)$ comme une combinaison linéaire des $\phi(t - k)$. Cette combinaison linéaire introduit l'existence d'une séquence génératrice de la fonction échelle (ou filtre échelle) $h = \{h[k]\}_{k \in \mathbb{Z}}$ et nous obtenons la première équation à deux échelles :

$$\phi\left(\frac{t}{2}\right) = \sqrt{2} \sum_k h[k] \phi(t - k) \quad (2.22)$$

De plus, $\psi\left(\frac{t}{2}\right) \in W_1$ et $W_1 \subset V_j$. Ainsi, $\psi\left(\frac{t}{2}\right)$ peut s'écrire comme une combinaison linéaire des $\phi(t - k)$. En introduisant la séquence génératrice de la fonction ondelette (ou filtre) $g = \{g[k]\}_{k \in \mathbb{Z}}$, nous obtenons l'équation suivante, à deux échelles :

$$\psi\left(\frac{t}{2}\right) = \sqrt{2} \sum_k g[k] \phi(t - k) \quad (2.23)$$

Ces équations mettent en évidence l'existence de deux filtres discrets h et g qui définissent la DWT. Ce sont ces deux filtres (filtres miroirs en quadrature), et non les fonctions ψ et ϕ , qui sont utilisées en pratique pour le calcul et pour la conception d'ondelettes (Algorithme de Mallat [16]).

3.1 Mise en œuvre de la DWT

Pour un signal à temps continu $x(t)$, la DWT est la représentation des coefficients de la projection de x dans les différents espaces W_j :

$$\text{DWT}_x = \{d_j[k] = \langle x, \psi_{j,k} \rangle\}_{(j,k) \in \mathbb{Z}^2} \quad (2.24)$$

Les relations à deux échelles permettent d'exprimer le coefficient d'approximation et de détail à un niveau donné comme le filtrage des coefficients d'approximation du niveau précédent par les séquences génératrice $\bar{h}$ et $\bar{g}$ où $(\bar{h}[k] = h[k])$, suivi d'une décimation :

$$a_j[k] = \downarrow 2(\bar{h} * a_{j-1}[k]) \quad (2.25)$$

$$d_j[k] = \downarrow 2(\bar{g} * a_{j-1}[k]) \quad (2.26)$$

et de calculer la DWT par la mise en œuvre des bancs de filtres $(\bar{h}, \bar{g})$.

Pour un signal discret $x[k]$, le nombre de niveaux de décomposition est fini et limité par la longueur N du signal. La DWT est alors définie par :

$$DWT_x = \left\{ \left\{ \{a_j[k]\}_{k=0, \dots, \frac{N}{2^j}-1} \right\}_{j=1, \dots, J}, \left\{ \{d_j[k]\}_{k=0, \dots, \frac{N}{2^j}-1} \right\}_{j=1, \dots, J} \right\} \quad (2.27)$$

avec $J \leq \log_2(N)$ le niveau de décomposition maximum.

3.2 Transformée par paquets d'ondelettes

Les paquets d'ondelettes sont une généralisation des ondelettes orthogonales. Ils permettent une analyse plus fine en décomposant les espaces de détail qui ne sont pas décomposés dans le cas de la DWT.

Ils ont été introduits au début des années 1990 [21], afin de pallier le manque de résolution fréquentielle de l'analyse en ondelettes. Leur génération se fait par récurrence. On part des filtres miroirs conjugués h et g , de longueur N , associés à l'ondelette orthogonale à support compact ψ et la fonction d'échelle ϕ , issues d'une analyse multi-résolution de $\mathcal{L}^2$.

Par induction des équations (2.22) et (2.23), on définit la suite de fonctions $(w_p)_{p \in \mathbb{N}}$ [16], par :

$$w_{2p}(t) = \psi_{j+1}^{2p}(t) = \sum_{k=-\infty}^{+\infty} h[k] \psi_j^p(t - 2^j k) \quad (2.28)$$

$$w_{2p+1}(t) = \psi_{j+1}^{2p+1}(t) = \sum_{k=-\infty}^{+\infty} g[k] \psi_j^p(t - 2^j k) \quad (2.29)$$

$p = 0, \dots, 2^j - 1$

avec $\psi_0^0(t) = \phi(t)$. Ces fonctions sont organisées selon un arbre binaire où chaque nœud (paquet d'ondelettes) est une sous-base $\Psi_j^p = \{\psi_j^p(t - 2^j k)\}_k$ explorant la bande de fréquence normalisée $2^{-(j+1)} \cdot [p, p+1]$ (Figure 2.3). Les coefficients de la décomposition d'un signal discret x de longueur N sont obtenus par l'extension de l'algorithme de Mallat [16].

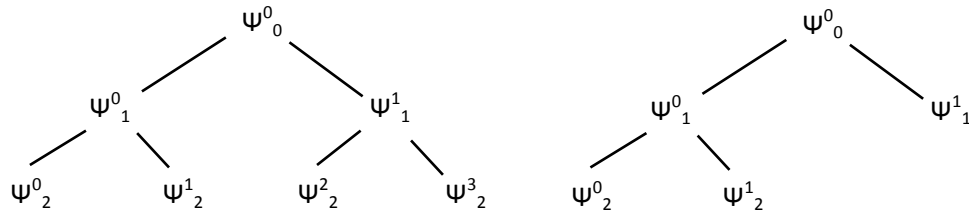


Figure 2.3 : A gauche, arbre de paquets d'ondelettes, avec $j=2$. A droite, l'arbre des sous-bases de la DWT.

La transformée par paquets d'ondelettes, en comparaison avec la DWT, décompose les informations détaillées du signal dans la région de haute-fréquence. La figure 2.4 illustre, schématiquement, un processus de décomposition du signal en WPT à trois niveaux, produisant un total de 8 sous-bandes, avec chaque sous-bande couvrant un huitième du spectre de fréquence du signal.

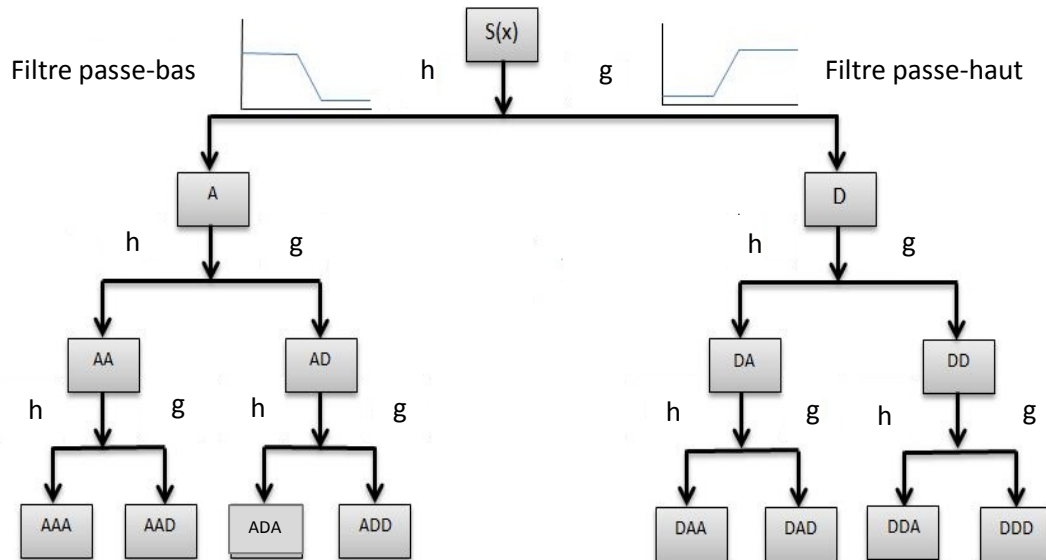


Figure 2.4 : Décomposition par transformée en paquets d'ondelettes sur trois niveaux.

La transformée en ondelette est couramment utilisée dans différents domaines du traitement de signal, incluant l'analyse des signaux biomédicaux, le traitement de la parole, ainsi que dans la compression et le débruitage des signaux [18, 101].

Méthodes de réduction

1. Analyse en Composantes Principales

L'analyse en composantes principales (ACP) est une des plus anciennes techniques de statistique exploratoire connues pour la réduction de dimension. C'est une méthode spectrale et projective. Son principe repose sur des idées énoncées par Pearson en 1901, et formalisées par Hotelling en 1933, Karhunen en 1947 et Loève en 1948. On parle également de transformée de Karhunen-Loève [86]. Cette analyse non-supervisée recherche les axes principaux (ou axes d'inertie) qui permettent une dispersion maximale des données.

L'ACP consiste à remplacer une famille de variables corrélées par de nouvelles variables décorrélées les unes des autres et qui sont des combinaisons linéaires des variables corrélées d'origine. Ces nouvelles variables (ou composantes principales) définissent des axes factoriels qui servent de base à une représentation graphique plane des variables initiales. L'interprétation des résultats se restreint généralement aux deux premiers plans factoriels, avec un minimum de perte d'information, sous réserve que ceux-ci expliquent la plus grande variance du nuage des variables initiales [22]. En termes de projection linéaire normalisée, l'ACP consiste à maximiser la variance dans l'espace projeté [24]. L'ACP utilise des matrices carrées, symétriques réelles et donc diagonalisables [86].

En effet, pour une dimension p d'ensemble de données X , les m axes principaux $W_1, \dots, W_m$ (où $1 \leq m \leq p$) sont des axes orthogonaux sur lesquels la variance est conservée au maximum dans l'espace prévu. Ces principales composantes sont définies par :

$$W_m = \arg_{\|W\|=1} \max E\{(W^T X)^2\} \quad (3.1)$$

La transformation de Karhunen-Loève équivaut donc à trouver la décomposition en valeurs singulières de la matrice des données X :

$$X = W \Sigma V^T \quad (3.2)$$

Nous obtenons la matrice de donnée à espace réduit Y , en projetant X sur l'espace réduit défini uniquement par les m premiers vecteurs singuliers W_m :

$$Y = W_m^T X = \Sigma V_m^T \quad (3.3)$$

Les m composantes principales du vecteur d'observation X_i sont données par le vecteur de caractéristique réduit, défini par :

$$Y_i = W^T X_i \quad (3.4)$$

$W_1, \dots, W_m$ sont représentés par les m premiers vecteurs propres de la matrice de covariance S :

$$S = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)(X_i - \mu)^T \quad (3.5)$$

où μ représente la moyenne et N le nombre d'échantillons, de sorte que :

$$S W_i = \lambda_i W_i \quad (3.6)$$

où λ_i représente la i ème plus grande valeur propre de S .

Quant à l'opération inverse, X_i est donné par :

$$X_i = W Y_i \quad (3.7)$$

2. Analyse Discriminante Linéaire

Les données de grande dimension présentent deux particularités qui rendent difficiles l'exploitation et l'extraction de l'information qu'elles contiennent : d'une part, les données de grande dimension vivent dans des espaces de dimension intrinsèque plus petite que la dimension de l'espace original et, d'autre part, ces données peuvent être bruitées. A cela s'ajoutent une difficulté de représentation graphique et de visualisation de ces données complexes. En effet, l'ACP, qui est une technique non-supervisée, n'utilisent pas les classes associées aux données. Contrairement à l'ACP, l'objectif de l'analyse discriminante linéaire LDA⁷ est de réduire le nombre de dimensions tout en préservant au maximum les classes. Pour cela, elle cherche les axes tels que la projection des données dans l'espace engendré par ces axes donne une plus grande séparation des classes.

⁷ Linear discriminant analysis

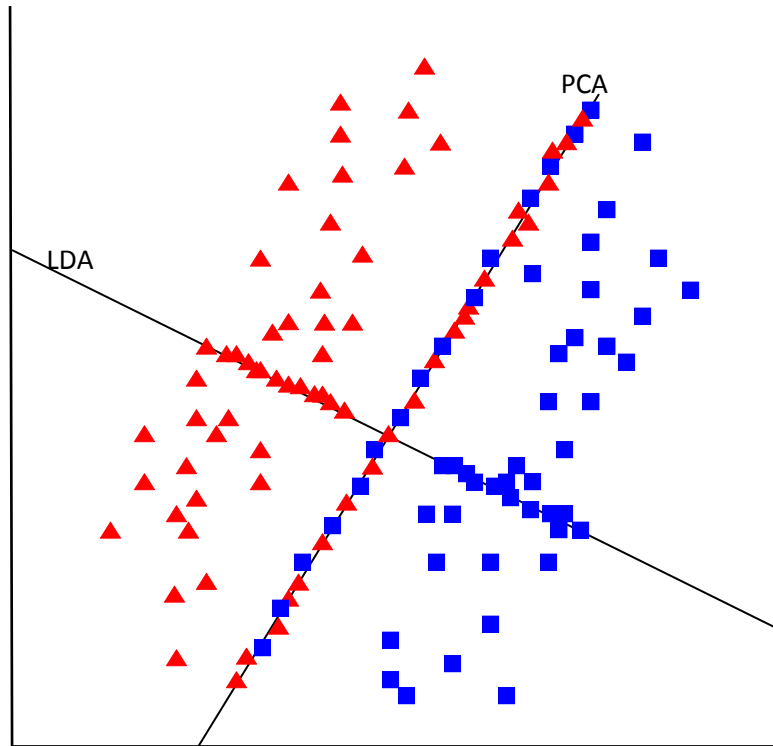


Figure 3.1 : Réduction de dimensions par LDA et par ACP.

L'analyse discriminante linéaire LDA et les discriminants de Fisher FDA⁸ relatifs, sont utilisés pour trouver la combinaison linéaire des caractéristiques qui séparent au mieux les classes [25]. La figure 3.1 compare les axes discriminants choisis par la LDA et l'ACP pour les mêmes données. Les deux méthodes permettent de réduire de 2 à 1 dimension en trouvant, soit l'axe qui décrit le mieux les données pour l'ACP, soit l'axe qui sépare le mieux les classes pour la LDA.

L'analyse discriminante de Fisher consiste dans un premier temps à projeter les données de l'espace initial sur des axes discriminants qui maximisent le rapport de la variance interclasse sur la variance intra-classe, puis à classer les données dans cette nouvelle représentation.

⁸ Fisher discriminant analysis

En effet la LDA utilise le critère de classification du signal afin de trouver le meilleur vecteur de projection. Supposons que nous disposons d'ensemble d'apprentissage $\{X_1, X_2, \dots, X_N\}$, où chaque document appartient à l'une des deux classes c_1 et c_2 (caractérisées par leurs centres de gravité ou centroïdes). Nous cherchons l'axe discriminant W tel que les projetés des éléments soient le mieux séparés en classes.

Soit le scalaire Y_i représentant le projeté des X_i sur cet axe, on a :

$$Y_i = W^T X_i \quad (3.8)$$

La mesure de la séparation entre les points projetés est la différence des vecteurs moyens μ_i de chaque classe c_i . Les vecteurs moyens sont définis par :

$$\mu_i = \frac{1}{N_i} \sum_{X \in c_i} X \quad i = 1, 2 \quad (3.9)$$

et

$$\hat{\mu}_i = \frac{1}{N_i} \sum_{Y \in c_i} Y = \frac{1}{N_i} \sum_{X \in c_i} W^T X = W^T \mu_i \quad (3.10)$$

La matrice W (variance-covariance) peut-être décomposée et caractérisée par la somme de deux matrices : S_B (matrice de dispersion inter-classe⁹) et S_w (matrice de dispersion intra-classe¹⁰).

La fonction linéaire (équation 3.8), permettant la réduction des dimensions, cherche à maximiser le critère de Fisher $J(W)$ ¹¹:

$$J(W) = \frac{W^T S_B W}{W^T S_W W}, \quad (3.11)$$

avec la matrice de dispersion inter-classe définie comme la moyenne des dispersions de chaque classe :

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T, \quad (3.12)$$

⁹ Between-class= matrice variance-covariance pondérée des i centroïdes.

¹⁰ Within-class= moyenne des i matrices variance-covariance des classes.

¹¹ Le critère de Fisher peut aussi être utilisé sous la forme : $J(W) = \frac{W^T S_B W}{W^T S_T W}$ (S_T : matrice de dispersion totale $S_T = S_w + S_B$) ou $(W) = \frac{W^T S_w W}{W^T S_B W}$.

où :

μ est le vecteur moyen de l'ensemble des données.

μ_i est le vecteur moyen des éléments de la classe c_i .

Et la matrice de dispersion intra-classe définie comme la dispersion des moyennes de chaque classe :

$$S_W = \sum_{i=1}^c \sum_{j=1}^{N_i} (X_j^{(i)} - \mu_i) (X_j^{(i)} - \mu_i)^T, \quad j = 1, \dots, i \quad (3.13)$$

La matrice de dispersion totale S_T est définie par :

$$S_T = \sum_{X_i} (X_i - \mu) (X_i - \mu)^T \quad (3.14)$$

Conclusion

De manière plus générale, l'analyse en composantes principales et l'analyse discriminante linéaire appartiennent à la même famille de l'analyse factorielle. Ces techniques partagent toutes le même principe de diagonalisation, mais cette matrice n'est pas constituée des mêmes éléments, selon la méthode utilisée et la nature des données. L'approche définie par la méthode de réduction dans les sous-espaces discriminants de Fisher se révèle être peu sensible à l'augmentation de la dimension et aux données bruitées. De plus, le nombre de paramètres à estimer est indépendant de la dimension originale des données. L'algorithme de l'analyse discriminante linéaire est un peu plus complexe que celui de l'analyse en composantes principales en raison de sa métrique particulière.

Classification et performances de mesures

1. Introduction

Notre travail consiste à identifier une méthode de traitement des signaux EEG, permettant de réaliser le meilleur taux de classification du signal EEG. Il existe plusieurs méthodes permettant de calculer des fonctions discriminantes qui permettent d'estimer la classe d'un nouvel exemple. Dans le cadre de ce travail, nous disposons d'un ensemble d'apprentissage labélisé. Nous sommes, donc, dans un contexte de classification supervisée dont les méthodes les plus connues, en génie-biomédical, sont: les réseaux de neurones artificiels (ANN: Artificial Neural Network) [82], les K-plus proches voisins (K-NN: K-Nearest Neighbors) [65], les réseaux de neurones à fonction de base radiale (RBFNN: Radial Basis Function Neural Network) [84], les réseaux de neurones récurrents (RNN: Recurrent Neural Network) [85] et les machines à vecteurs supports (SVM: Support Vector Machine) [86].

Dans ce travail, nous avons utilisé trois méthodes de classification : SVM, K-NN et ANN ; mais c'est sur la méthode des machines à vecteurs supports que nous nous sommes focalisés car elle présente différents avantages. Cette méthode de classification par SVM :

- ✓ ne fait aucune hypothèse forte sur la forme des classes,
- ✓ fonctionne avec de faibles populations d'apprentissage,

- ✓ possède une excellente capacité de généralisation,
- ✓ permet de s'affranchir des problèmes des espaces de représentation de trop grande dimension
- ✓ évite le sur-apprentissage,
- ✓ possède une solution unique.

2. Classification par machines à vecteurs support

Les machines à vecteurs de support ont été introduites en 1995 par Cortes et Vapnik [27, 28, 29]. Elles sont largement utilisées dans de nombreux problèmes d'apprentissage : diagnostic médical, reconnaissance de forme, ou encore catégorisation de texte [102].

Les machines à vecteurs support forment une classe d'algorithmes d'apprentissage supervisés. Nous nous intéressons à une fonction f qui à toute entrée x fait correspondre une sortie $y=f(x)$.

Soit une population d'apprentissage $X = x_1, \dots, x_m$ répartie suivant deux classes c_{-1} et c_1 , où $x_i \in \mathbb{R}^d$ est le vecteur descripteur de l'individu i . L'ensemble d'étiquettes associées $\{y_i\}$ où y_i décrit la classe de x_i . Dans ce problème, les machines à vecteurs support vont être utilisées pour classer une nouvelle observation x_i en se limitant à deux classes $y_i \in \{-1, 1\}$; $y_i = -1$ si $x_i \in c_{-1}$ ou $y_i = +1$ si $x_i \in c_1$.

$$f: \mathbb{R}^d \rightarrow \{-1, 1\}, f(x_i) = y_i \quad (4.1)$$

Le but est d'essayer d'apprendre f à partir d'un ensemble du couple (x_i, y_i) , donc à estimer cette fonction suivant la règle de décision :

$$y = \text{signe}(f(X)) \quad (4.2)$$

Dans le cas linéaire, i.e. où les données sont linéairement séparables (voir figure 4.1), une fonction discriminante f est obtenue par combinaison linéaire du vecteur d'entrée X et s'écrit :

$$f(X) = w^t X + b = \langle w, X \rangle + b \quad (4.3)$$

De plus, l'équation (4.2) nous donne les conditions suivantes :

$$\begin{cases} \langle w, x_i \rangle + b > 0 & \text{si } y_i = +1 \\ \langle w, x_i \rangle + b < 0 & \text{si } y_i = -1 \end{cases} \quad (4.4)$$

L'équation $\langle w, x_i \rangle + b = 0$ définit l'hyperplan séparateur.

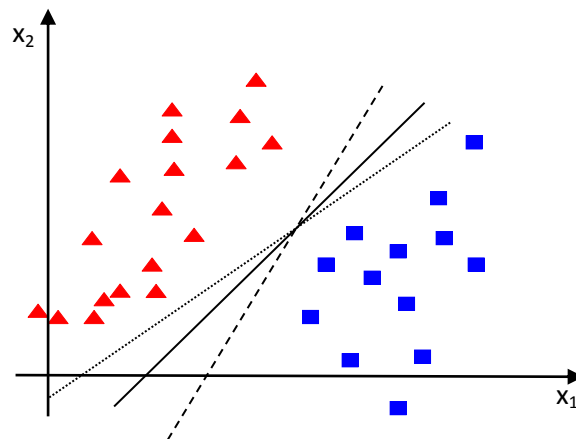


Figure 4.1 : Classification binaire avec plusieurs hyperplans.

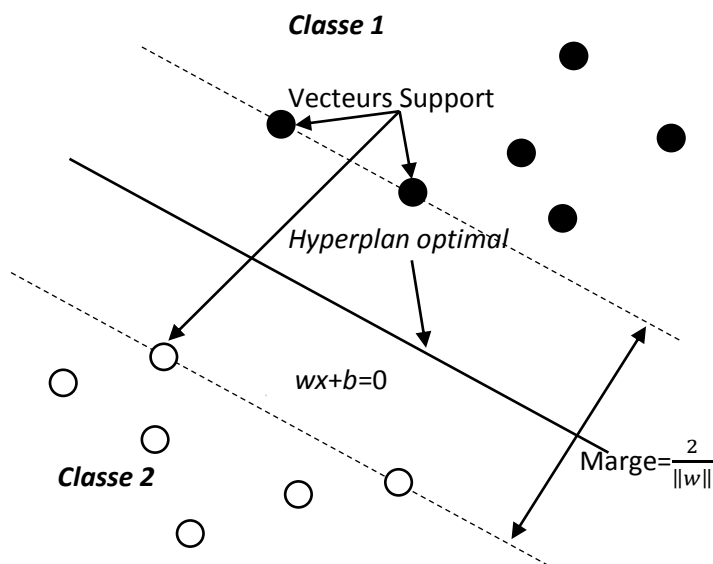


Figure 4.2 : Hyperplan de séparation optimale avec marge souple dans un cas non linéairement séparable.

Sur la figure 4.1, il existe plusieurs hyperplans séparateurs satisfaisants les équations (4.3) et (4.4), ceci montre qu'il n'existe pas de solution unique. Selon la théorie de Vapnick [27], les SVM reposent sur deux notions: celle de la marge maximale et celle de la fonction noyau. Elles permettent de résoudre des problèmes de discrimination non linéaire.

Cette marge est définie comme la distance entre un hyperplan et les points échantillons les plus proches appelés vecteurs support et représentés par la figure 4.2. La distance d'un point à la séparatrice est $(w^t X + b)/\|w\|$, soit $2/\|w\|$. Le problème des SVM revient à maximiser la marge, i.e. à minimiser $\|w\|$. Nous sommes confrontés à un problème de minimisation quadratique sous contrainte qui s'écrit :

$$\begin{cases} \min(\frac{1}{2}\|w\|^2) \\ y_i(\langle w, x_i \rangle + b) - 1 \geq 0 \quad i = 1, \dots, m \end{cases} \quad (4.5)$$

Pour résoudre ce problème d'optimisation, on utilise les multiplicateurs de Lagrange notés :

$$\mathcal{L}(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^m \alpha_i [y_i(\langle w, x_i \rangle + b) - 1], \quad \alpha_i \geq 0 \quad \forall i \quad (4.6)$$

Cette équation vérifie les conditions de Karush Khun-Tucker (KKT) :

$$\alpha_i^0 [y_i(\langle w^0, x_i \rangle + b^0) - 1] = 0 \quad (4.7)$$

Les multiplicateurs de Lagrange α_i sont donc positifs si la contrainte est saturée (ce sont les vecteurs supports qui seuls contribuent à définir l'hyperplan optimal) et nuls ailleurs.

La résolution de l'équation (4.6) revient à la résolution du problème dual pour des fonctions convexes. On peut l'écrire sous la forme suivante :

$$\min_{w, b} \left[\max_{\alpha} \mathcal{L}(w, b, \alpha) \right] = \max_{\alpha} \left[\min_{w, b} \mathcal{L}(w, b, \alpha) \right] \quad (4.8)$$

La formulation de Lagrange permet de trouver les extremums en annulant les dérivées partielles de la fonction $\mathcal{L}(w, b, \alpha)$. Le lagrangien doit-être minimisé par rapport à w et b et maximisé par rapport à α [102], i.e. le minimum doit annuler le gradient en w et b du Lagrangien :

$$\frac{\partial \mathcal{L}}{\partial b} = 0 \Rightarrow \sum_{i=1}^m \alpha_i y_i = 0 \quad (4.9)$$

$$\frac{\partial \mathcal{L}}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^m \alpha_i y_i x_i \quad (4.10)$$

Suivant les équations (4.9) et (4.10), le problème dual se simplifie comme suit :

$$\max_{\alpha} \left[\sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \right] \text{ avec } \begin{cases} 0 \leq \alpha_i \quad \forall i \\ \sum_{i=1}^m \alpha_i y_i = 0 \end{cases} \quad (4.11)$$

La fonction discriminante f (équation 4.3) s'écrit alors :

$$f(X) = \langle \hat{w}, X \rangle + \hat{b} \quad (4.12)$$

avec les paramètres suivants :

$$\hat{w} = \sum_{i=1}^m \hat{\alpha}_i y_i x_i \quad (4.13)$$

$$\text{et} \quad y_i (\langle \hat{w}, x_i \rangle + \hat{b}) - 1 = 0 \quad \alpha_i \neq 0 \quad (4.14)$$

Dans le cas non séparable, les conditions du problème seront reformulées comme suit :

$$\begin{cases} \langle w, x_i \rangle + b \geq 1 - \xi_i \text{ si } y_i = +1 \\ \langle w, x_i \rangle + b \leq -1 + \xi_i \text{ si } y_i = -1 \end{cases} \quad \text{avec } \xi_i \geq 0 \quad \forall i \quad (4.15)$$

La variable de relaxation ξ_i (dite « ressort ») est introduite dans le concept de marge souple, autorisant les exemples mal classés. Un autre paramètre de régularisation C est introduit afin de contrôler la pénalité et limiter l'augmentation des ξ_i . L'équation (4.5) devient alors [99]:

$$\begin{cases} \min(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i) \\ y_i (\langle w, x_i \rangle + b) - 1 + \xi_i \geq 0 \quad i = 1, \dots, m \end{cases} \quad (4.16)$$

En plus, une fonction noyau K remplace tous les produits scalaires, telle que $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$. Donc, comme exprimé précédemment (équation (4.6)), le problème revient à maximiser le multiplicateur de Lagrange $\mathcal{L}(\alpha)$:

$$\mathcal{L}(w, b, \alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i - \sum_{i=1}^m \alpha_i [y_i (\langle w, x_i \rangle + b) - 1 + \xi_i] \quad (4.17)$$

$$\mathcal{L}(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \alpha_i \alpha_j y_i y_j K(x_i, x_j), 0 \leq \alpha_i \leq C \text{ et } \sum_i y_i \alpha_i = 0 \quad (4.18)$$

La fonction de décision est sous la forme :

$$f(x) = \sum_{i=1}^m \hat{\alpha}_i y_i K(x, x_i) + \hat{b} \quad (4.19)$$

et la règle de décision :

$$y = \text{signe}(f(x)) \quad (4.20)$$

Il existe de nombreuses fonctions noyau prédéfinies [99], telles que :

- Noyau linéaire :

$$K(x_i, x_j) = \langle x_i, x_j \rangle \quad (4.21)$$

- Noyaux polynomiaux :

$$K(x_i, x_j) = (1 + \langle x_i, x_j \rangle)^p \quad (4.22)$$

- Noyau gaussien (RBF) :

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (4.23)$$

où σ contrôle la largeur de la fonction RBF¹².

3. Classification par K plus proches voisins

La classification par la méthode des K plus proches voisins (K-NN¹³) est une des méthodes d'apprentissage artificiel les plus simples à mettre en œuvre. L'idée repose sur le fait de déterminer la classe d'une nouvelle observation X en fonction de la classe majoritaire parmi les K plus proches voisins de X [99]. Cette méthode utilise deux paramètres: le nombre K et les fonctions de similarité (mesure de distance) pour comparer le nouveau cas à ceux déjà classés.

Afin de minimiser l'erreur de classification, le choix du nombre de voisins K peut-être optimisée par des méthodes heuristiques comme la validation croisée :

Les étapes de cet algorithme sont données comme suit :

1. Choix de $K \in N$;

¹² Radial Basis Function

¹³ K- Nearest Neighbor

2. Calculer les distances $d(X, X_i)$, $i=1,2,...,n$ où d est la métrique choisie (distance euclidienne ou distance de Mahalanobis¹⁴).
3. Retenir les k observations $X_{(1)}, ..., X_{(k)}$ pour lesquelles d est la plus petite.
4. Compter le nombre de fois $k_1, ..., k_j$ que ces observations apparaissent dans chacune des classes.
5. Prédiction des observations appartenant à la classe qui apparaît le plus fréquemment dans le voisinage de K .

La valeur par défaut de K est égale à 1, et le paramètre de similarité (plus proche voisin), par défaut, est la distance euclidienne¹⁵. Dans ce travail, la valeur de K est réglée à 5 (où l'erreur de classification est minimale).

4. Classification par réseau de neurone artificiel

Dans cette étude, nous avons également utilisé les réseaux de neurone artificiel ANN¹⁶ [34, 63, 99] pour une comparaison avec les autres méthodes. Parmi les réseaux de neurone artificiel le plus couramment utilisé, particulièrement dans le domaine biomédical, on trouve comme variante le perceptron multicouche MLP¹⁷ [60, 97, 100]. Le schéma de la figure 4.3 représente un réseau de neurone MLP à trois couches.

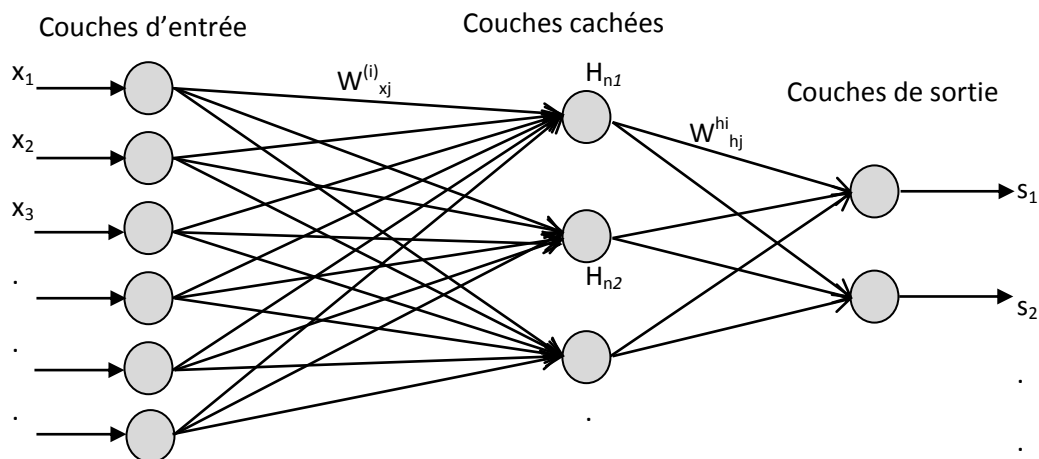


Figure 4.3 : Architecture d'un réseau MLP.

¹⁴ Matrice inverse de la matrice de variance.

¹⁵ $d(X_i, X_j) = \sqrt{\sum_i (X_i - X_j)^2}$

¹⁶ Artificial Neural Network

¹⁷ Multilayer Perceptron

Les différentes étapes de l'algorithme général de rétro-propagation sont données comme suit :

1. Soient les vecteurs de données d'entrée : $x_1(k), x_2(k), \dots, x_i(k) \in \mathbb{R}$.
2. Sélectionner : le nombre total de couche M , le nombre n_i ($i=1,2,\dots, M-1$) de neurones de la couche cachée, et le paramètre d'erreur de tolérance $\varepsilon > 0$.
3. Sélection des valeurs initiales du poids du vecteur $w_{xj}^{(i)}$ pour $i=1,2,\dots, n_i$.
4. Initialisation
5. Les couches de sortie $\begin{cases} s_j^{(i)} = (w_{xj}^{(i)})^T x^{(i-1)} \\ x_j^{(i)} = \sigma(s_j^{(i)}) \end{cases} \quad i = 1,2,\dots,M \text{ et } j = 1,2,\dots,n_i.$
6. L'erreur en sortie : $e_j = s_j - x_j^{(M)} \quad j = 1,2,\dots,m$
7. Les erreurs de propagation des neurones cachées :

$$e_j^{(i)} = \sum_{l=1}^{n_{i+1}} \delta_l^{(i+1)} w_{lj}^{(i+1)} \quad \text{de la couche } M-1, \dots, \text{à la couche } 1.$$
8. Les valeurs delta de neurones cachées : $\delta_l^{(i)} = e_j \sigma'(s_j^{(i)})$
9. Réactualiser les vecteurs poids : $w_{xj}^{(i)} = w_{xj}^{(i)} + \eta \delta_l^{(i)} x^{(i-1)}$
10. Calcul de la fonction erreur : $E = E + \frac{1}{k} \sum_{j=1}^m e_j^2$
11. Si $k=K$ alors continuer en 12 ; sinon ($k=k+1$) reprendre en 5.
12. Si $E \leq \varepsilon$ continuer en 13 ; sinon reprendre en 4.
13. Fin d'apprentissage. Le MLP est prêt pour le test.

La fonction d'activation σ est généralement choisie comme la fonction sigmoïde définie par [99] :

$$\sigma(s) = \frac{1}{(1+e^{-s})} \quad (4.24)$$

Cette fonction est celle utilisée dans notre travail ; cette fonction est non-linéaire, continue et dérivable, ce qui permet d'utiliser le gradient d'erreur afin d'actualiser les poids [99]. Parfois, on peut aussi utiliser σ comme une fonction de tangente hyperbolique ou une fonction à base radiale [60]. La fonction d'activation utilisée pour la couche de sortie est la fonction linéaire $\sigma(s)=s$.

L'algorithme de rétro propagation du gradient est utilisé pour l'ajustement des poids entre couches.

La complexité du réseau de neurones est caractérisée par le nombre de couches cachées [103]. Il n'y a aucune règle générale pour la sélection du nombre approprié de couches cachées. D'une part, un réseau neuronal ayant un faible nombre de neurones peut ne pas être suffisamment puissant pour modéliser une fonction complexe et d'autre part, un réseau de neurones avec trop de neurones peut conduire à un *overfitting*¹⁸ des ensembles d'apprentissage et par conséquent, perdre sa capacité de généralisation qui est la principale caractéristique souhaitée d'un réseau de neurone. Dans notre étude, l'approche élaborée pour trouver le nombre optimal de couches cachées se fait sur la base des essais et de la fonction erreur. La configuration du réseau la plus appropriée a été obtenue pour dix neurones de la couche cachée, trois neurones de la couche d'entrée et deux neurones pour la couche de sortie.

5. Performances statistiques

Evaluer les performances d'un système de classification est un enjeu de grande importance car ces performances peuvent être utilisées pour l'apprentissage en tant que tel ou pour optimiser les valeurs des hyper paramètres du classificateur.

La première mesure à laquelle nous allons nous intéresser, est le taux de bonne classification ou la précision. Il s'agit de l'indicateur le plus évident permettant d'évaluer les performances d'un système de classification. Cette valeur correspond au nombre d'éléments correctement identifiés par le système et elle est définie par la formule suivante:

$$PRECISION = \frac{\text{Nombre d'éléments correctement identifiés}}{\text{Nombre d'éléments total}}$$

Nous nous intéressons également à la sensibilité et la spécificité. La sensibilité d'un test ou d'un examen diagnostique est sa capacité de donner un résultat positif lorsque la maladie est présente. Elle s'oppose à la spécificité, qui est la capacité d'un test ou d'un examen de donner un résultat négatif lorsque la maladie n'est pas présente.

¹⁸ Sur-apprentissage

Tableau 4.1 Matrice de confusion.

	CLASSE REELLE	
	SUJETS SAINS	SUJETS EPILEPTIQUES
TESTS POSITIFS	VP	FP
TESTS NEGATIFS	FN	VN

Pour évaluer la qualité de détection de l'algorithme proposé, il est convenu d'utiliser les indices de performances largement répandues dans le domaine biomédical [70], en adoptant les abréviations suivantes (Voir tableau 4.1) :

FN (Faux Négatif): désigne un test négatif d'un patient atteint d'une maladie.

VP (Vrai Positif) : désigne un test positif d'un patient atteint d'une maladie.

VN (Vrai Négatif) : désigne un test négatif pour une personne saine.

FP (Faux Positif) : désigne un test positif pour une personne saine.

Un bon test de diagnostic minimise le nombre de FN et de FP. Les critères retenus sont:

- La précision mesure le taux de classification correcte des cas Normaux et Épileptiques :

$$\text{PRECISION} = \frac{VP + VN}{VP + FN + FP + VN} 100\% \quad (4.25)$$

- La sensibilité (TVP), ou la probabilité que le test soit positif si la maladie est présente, se mesure chez les malades seulement. Elle est donnée par :

$$\text{SENSIBILITE} = \frac{VP}{VP + FN} 100\% \quad (4.26)$$

- Une mesure de la sensibilité s'accompagne toujours d'une mesure de la spécificité (TVN). Cette dernière se mesure chez les non-malades seulement. Ainsi, la spécificité, ou la probabilité d'obtenir un test négatif chez les non-malades, est donnée par :

$$\text{SPECIFICITE} = \frac{VN}{VN + FP} 100\% \quad (4.27)$$

- Le taux d'erreur mesure l'erreur commise lors de la classification des cas Normaux et Épileptiques :

$$\text{TAUX D'ERREUR} = \left(1 - \frac{VP + VN}{VP + FN + FP + VN}\right) 100\% \quad (4.28)$$

Conclusion

Dans ce chapitre, nous avons décrit trois techniques de classification utilisées dans cette thèse, i.e. les machines à vecteur support, les K-plus proches voisins, ainsi que le perceptron multicouche à retro propagation. La stabilité de ces algorithmes de classification est évaluée par l'utilisation de différentes mesures statistiques représentées par des indices de performance calculés à partir des éléments donnés par la matrice de confusion. La mise en œuvre de ces différentes techniques est présentée dans le prochain chapitre.

Résultats expérimentaux et évaluation des performances

Ce dernier chapitre a pour objectif de mettre en évidence la méthodologie énoncée précédemment. La méthode proposée sera testée sur une base de données réelle d'enregistrement EEG; d'une part, la première application traitera des signaux EEG enregistrés sur des personnes saines ne présentant aucune pathologie et des signaux critiques EEG enregistrés au cours de la crise d'épilepsie; d'autre part, dans la seconde méthodologie, nous inclurons des enregistrements inter-critiques d'EEG réels. Pour cela, nous illustrons le schéma de la méthodologie proposée sur la figure 5.1. Le développement des différents algorithmes a été réalisé sous la plateforme du logiciel *Matlab(2012b)*.

Le système étudié nous permet :

- ✓ de mesurer l'activité cérébrale (EEG épileptiques et normaux)
- ✓ de prétraiter à l'aide de deux transformées en ondelette (DWT et WPT)
- ✓ d'extraire les paramètres caractéristiques et de réduire ceux qui sont redondants (à l'aide de la PCA et la LDA)
- ✓ d'utiliser ces caractéristiques pour une classification (Normale/ Pathologique)

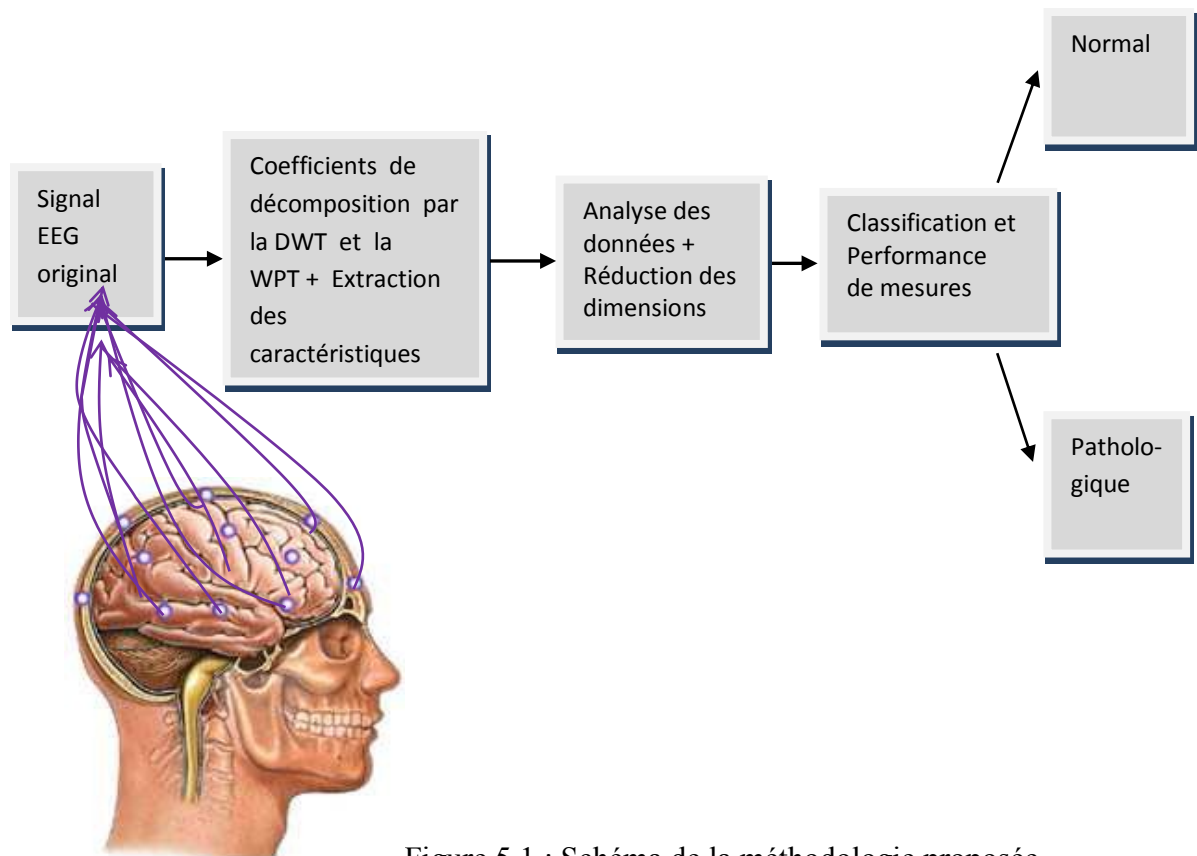


Figure 5.1 : Schéma de la méthodologie proposée.

1. Base de données

Nous avons utilisé les données d'enregistrement EEG, sans artefacts, disponibles au département d'Épileptologie de l'Université de Bonn [71]. C'est une base de données comprenant cinq ensembles (sets : A-B-C-D-E). Chaque ensemble est constitué de 100 signaux EEG monocanaux d'une durée de 23,6 s. Les données représentées par l'ensemble (A) ont été obtenues à partir des enregistrements d'EEG de cinq volontaires ne présentant aucune pathologie neurologique et se trouvant dans un état détendu avec les yeux ouverts. Ces signaux ont été enregistrés à partir d'enregistrements EEG de surface (extra-crâniens) conformément au système international (10-20) décrit au chapitre 1, section 3. Le set (B) concerne les enregistrements EEG des cinq personnes de l'ensemble (A), dans les mêmes conditions d'enregistrement, mais avec les yeux fermés.

Les sets (C), (D) et (E) proviennent des archives de diagnostic préopératoire. Ils ont été enregistrés par voie endocrânienne. L'ensemble des signaux du set D ont été enregistrés au sein de la zone épileptogène et ceux du set C à partir de la zone hippocampique de l'hémisphère opposé du cerveau. Les enregistrements de l'ensemble (C) représentent cent signaux EEG mesurant l'activité électrique du cerveau pendant des intervalles de repos entre deux crises. Il n'y a que le set (E) qui contient l'activité épileptique. Il a été choisi parmi tous les enregistrements contenant des crises de patients épileptiques.

Tous les signaux EEG ont été recueillis avec le même système d'amplificateur de 128 canaux et numérisés avec une fréquence d'échantillonnage de 173,61 Hz. Une conversion analogique-numérique à 12 bits et un filtre passe-bande, de 12dB/octave avec une bande passante (0,53-40Hz), ont été utilisés. Pour une description plus détaillée, le lecteur peut se référer à [72].

La figure 5.2 illustre l'ensemble des signaux EEG ;

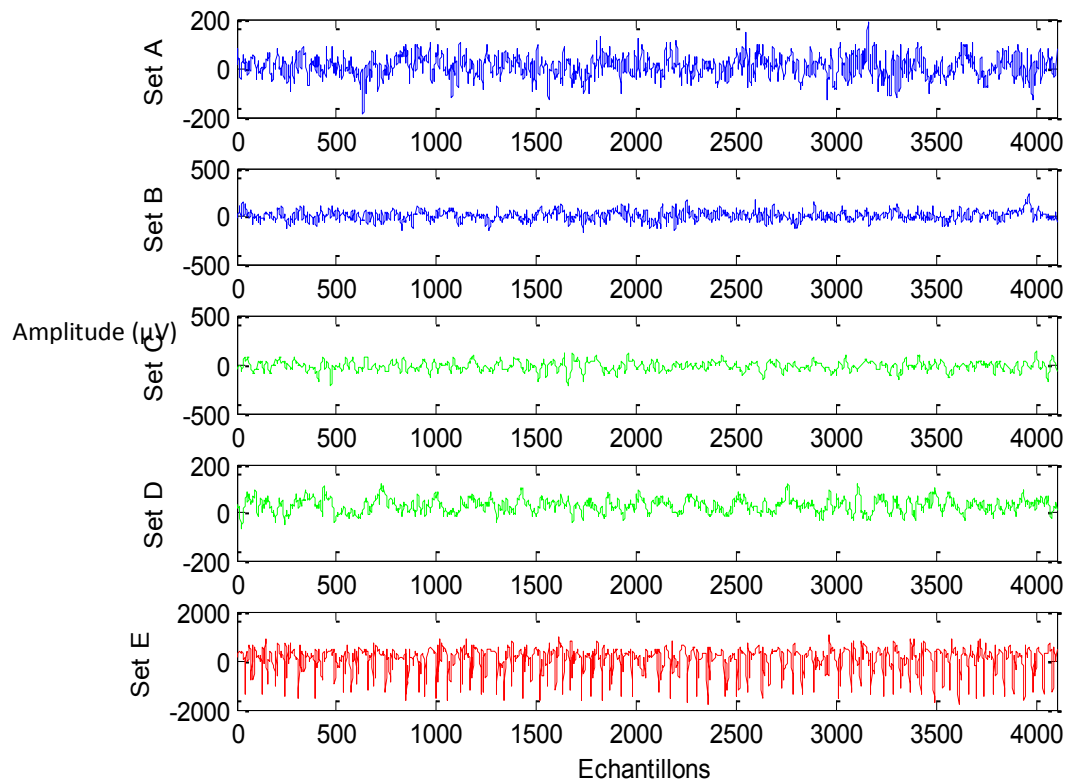


Figure 5.2 : Représentations des sets (A-B-C-D) issus de la base de données [71].

2. Prétraitement et extraction des caractéristiques

2.1 Transformée en ondelette discrète

Nous avons extrait différents paramètres statistiques et non statistiques des coefficients issus de la décomposition du signal EEG par la DWT et la WPT, et cela a été fait pour chaque niveau de décomposition. Le choix de l'ondelette est crucial afin de déterminer la meilleure ondelette mère pour l'analyse des signaux [73, 74, 91, 92, 101, 105, 106]. Nous avons évalué les performances de notre système en utilisant plusieurs types d'ondelette mère : *Daubéchies*, *Coiflet*, *Symlet*, *Biorthogonal*, *Discret Meyer*. Cette évaluation est représentée par la figure 5.3 où nous montrons la précision, la sensibilité et la spécificité de ces différentes ondelettes. La meilleure ondelette donnant le meilleur taux de précision est celle de *Db2*, *Db4*, *coif3* et *Bior1.1*.

Le choix de l'ondelette s'est porté sur l'ondelette de *Daubéchies* où la longueur du filtre est égale à $2N$ (N étant le nombre d'échantillons), tandis que la longueur du filtre de l'ondelette *Coiflet* est $6N$, et celle de l'ondelette *Biorthogonale* est égale à $(2N+2)$.

La figure (5.3) représente l'ondelette *Db2* donnant moins d'erreur de classification par rapport aux autres types d'ondelette, et par conséquent, elle sera choisie pour la suite de notre système avec l'analyse en ondelette discrète.

Le niveau de décomposition est tout aussi important pour la décomposition des signaux biomédicaux, en particulier les signaux EEG [58, 75, 91] ; et notre choix s'est porté sur cinq niveaux de décomposition de façon à traiter toutes les bandes de fréquences qui correspondent à celles des rythmes d'un signal EEG (delta (<4 Hz), thêta (4-8 Hz), alpha (8-13 Hz), beta (14-20 Hz) et gamma (20-60 Hz)). Une étude dans [91] montre que la précision de classification atteint les 99,33% à partir du cinquième niveau et reste inchangée et que les fréquences, utilisées dans l'interprétation d'un EEG, sont en dessous de 30 Hz [60].

Les coefficients d'ondelettes, qui sont extraits, fournissent une représentation compacte montrant la distribution de l'énergie du signal EEG en temps et en fréquence. Le tableau 5.1 représente les fréquences qui correspondent aux différents niveaux de décomposition pour des ondelettes *Daubechies 2* avec une fréquence d'échantillonnage de 173,6 Hz.

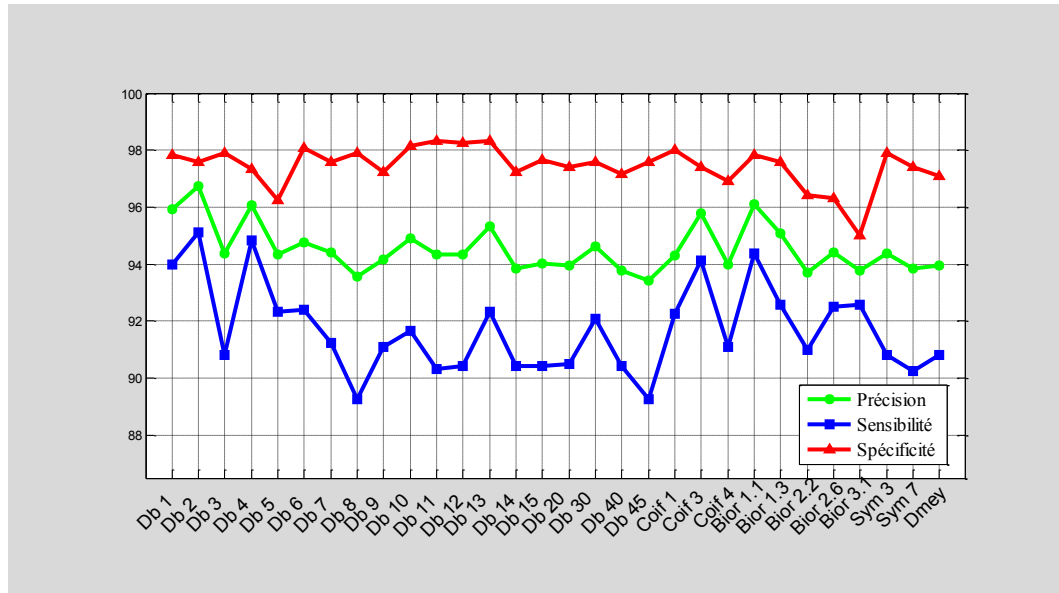


Figure 5.3 : Représentation des différentes ondelettes (*Daubéchies*, *Coiflet*, *Symlet*, *Biorthogonal*, *Discret Meyer*) utilisées: taux de précision, de sensibilité et de spécificité de deux cent signaux du set (A) et (E).

Tableau 5.1 : Fréquences correspondantes aux différents coefficients

Coefficient de décomposition	Bande de fréquence (Hz)
D_1	43.4 – 86.8
D_2	21.7 – 43.4
D_3	10.8 – 21.7
D_4	5.4 – 10.8
D_5	2.7 – 5.4
A_5	0 – 2.7

L'extraction des caractéristiques a pour but de décrire les signaux mesurés par des valeurs pertinentes. Afin de réduire davantage la dimension des vecteurs de caractéristiques extraites, deux méthodes de réduction (PCA et LDA) sur l'ensemble des coefficients d'ondelettes ont été utilisées.

Les figures 5.4 et 5.5 représentent la décomposition en coefficients d'ondelette des signaux EEG sains (set A) et des signaux EEG épileptiques (set E), respectivement, en utilisant la DWT.

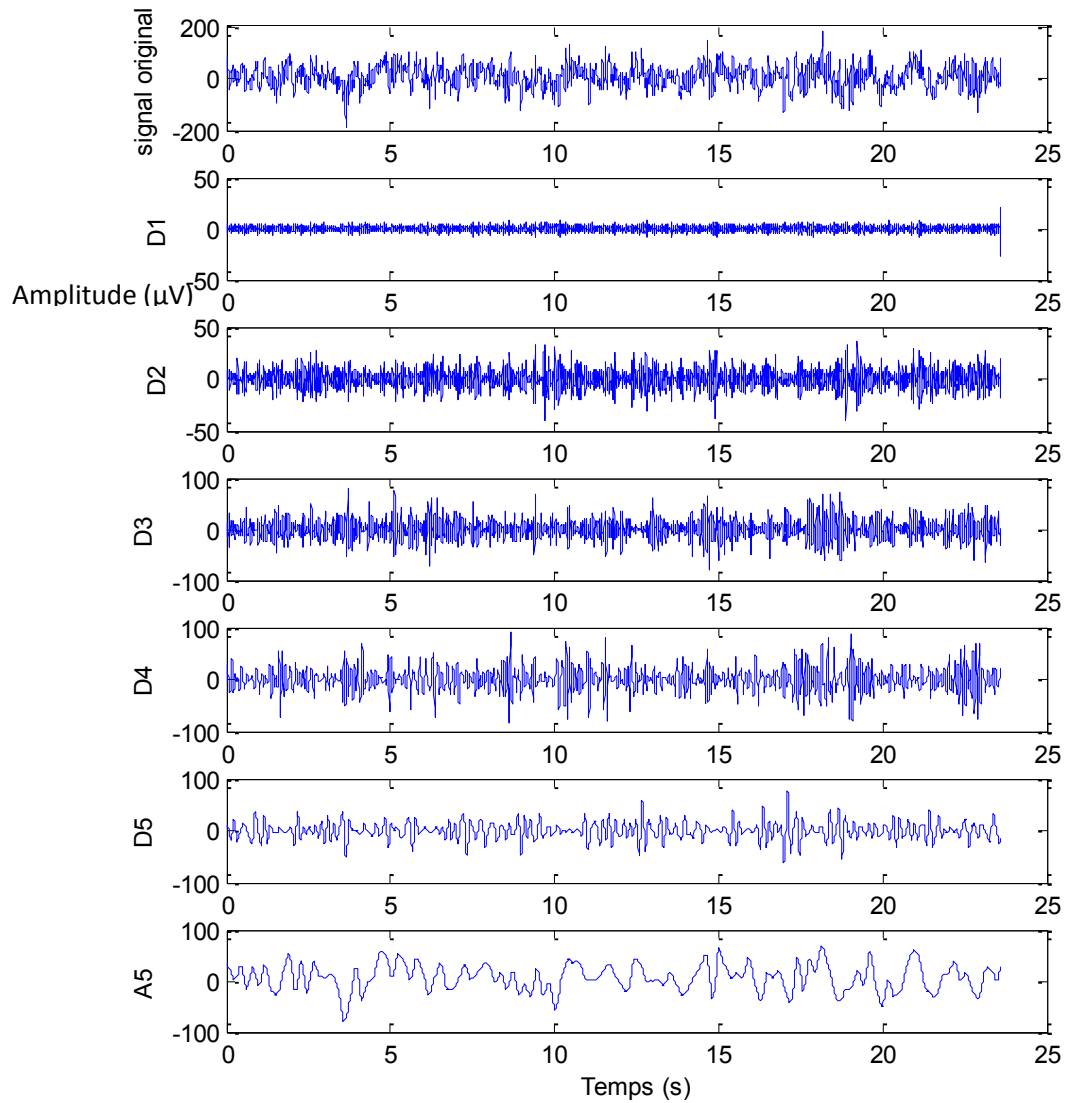


Figure 5.4 : Décomposition par la DWT d'un signal EEG sain (Set A).

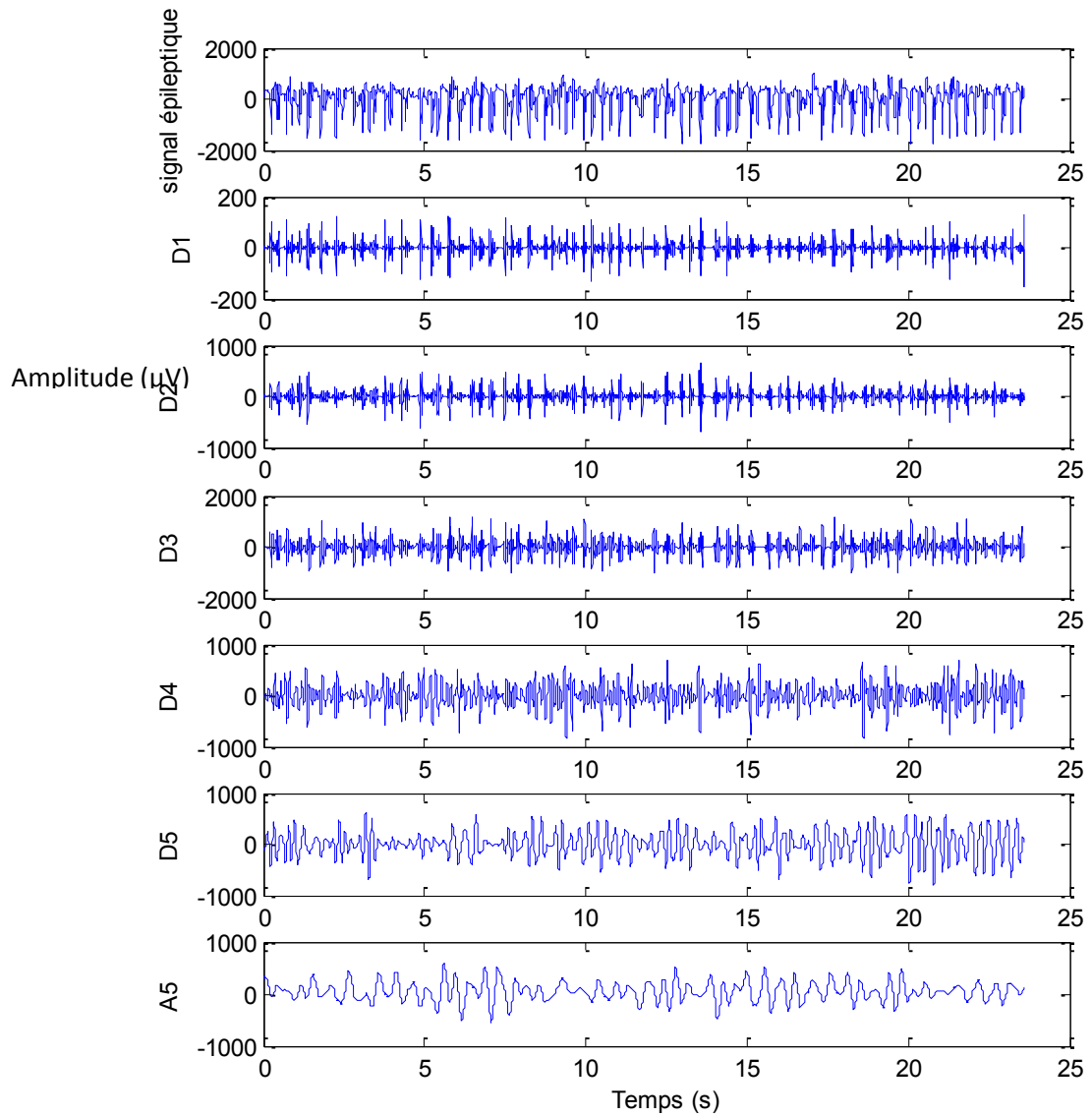


Figure 5.5 : Décomposition par la DWT d'un signal EEG épileptique (Set E).

Après décomposition par la DWT, les coefficients obtenus (D1, D2, D3, D4, D5, A5) représentent le signal original ; nous utiliserons ces coefficients pour extraire l'information en leur appliquant quelques paramètres caractéristiques.

2.2 Transformée en paquet d'ondelette

Comme mentionné précédemment pour l'analyse par DWT, nous avons besoin de choisir une ondelette mère et un niveau de décomposition optimal pour analyser et traiter les signaux EEG par la transformée en paquet d'ondelette; avec la même démarche que pour l'analyse par la transformée en ondelette discrète, nous trouvons que l'ondelette mère *Daubéchies* d'ordre 2 et d'ordre 4 avec un niveau de décomposition égal à cinq, nous donne une analyse optimale lors de la décomposition par paquet d'ondelette du signal EEG.

La figure 5.6 représente la décomposition en coefficients d'ondelette des signaux EEG sains (set A) par la WPT et la figure 5.7 montre la décomposition par la WPT des signaux EEG épileptiques (set E).

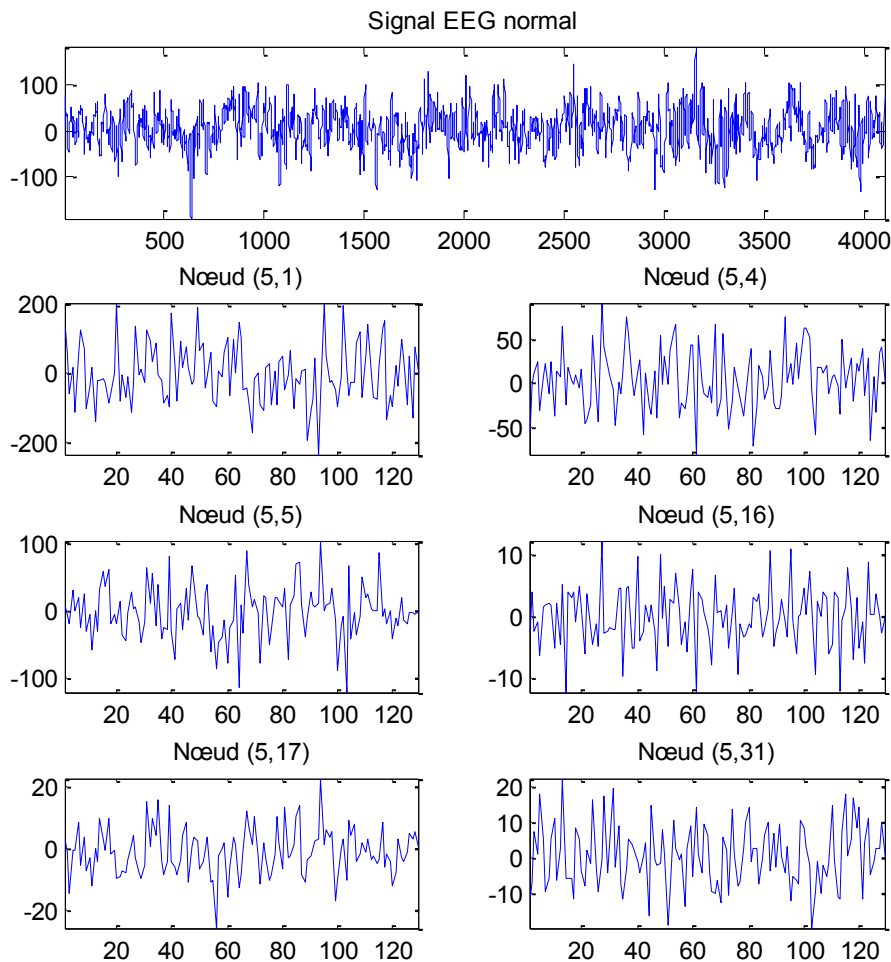


Figure 5.6 : Décomposition par la WPT du signal EEG sain (amplitude en μV).

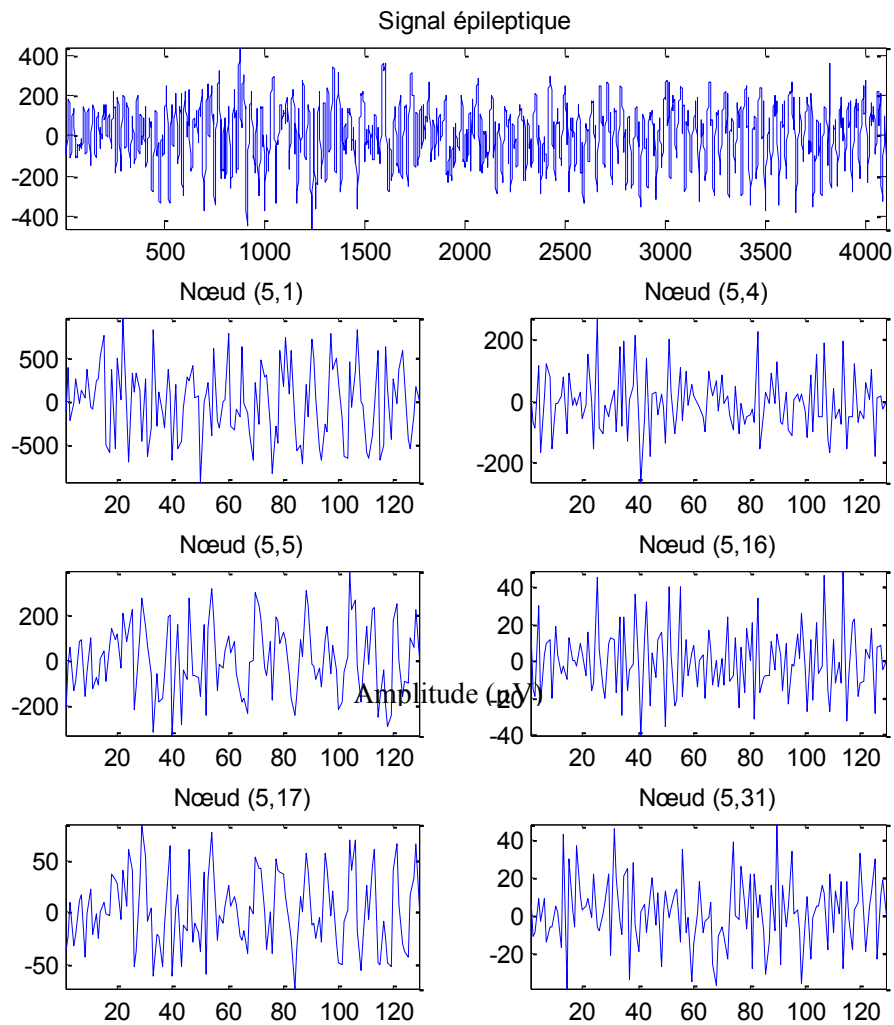


Figure 5.7 : Décomposition par la WPT d'un signal EEG épileptique.

L'ensemble des données EEG (set A et E) regroupent deux cent signaux d'enregistrement, et chaque signal enregistré a une durée de 23,6s et se compose de 4096 échantillons.

Le tableau 5.2 nous donne des bandes de fréquence qui correspondent aux différents niveaux de décomposition (WPT) pour des ondelettes *Daubechies 2* avec une fréquence d'échantillonnage de 173,6 Hz.

Tableau 5.2 : Fréquences correspondantes aux différents coefficients de décomposition basés sur la WPT.

n	Bandes de fréquence (Hz)	
	$l[n]$	$h[n]$
1	0-86.8	86.8-173.6
2	0-43.5; 86.3-130.2	43.5-86.8; 130.2-173.6
3	0-21.75; 43.5- 54.375; 86.3-108.05; 130.2 -151.95	21.75-43.5; 54.375- 86.3; 108.05-130.2; 151.95-173.6
4	0-10.875; 21.75-32.625; 43.5-54.375; 65.25-76.125; 87-97.875; 108.75-119.625; 130.5 141.375; 152.25-163.125	10.875-21.75; 32.625-43.5; 54.375-65.25; 76.125-87; 97.875-108.75; 119.625-130.5; 141.375-152.25; 163.125-173.6
5	0-5.44; 10.875-16.31; 21.75-27.19; 32.625-38.06; 43.5-48.94; 54.375-59.81; 65.25-70.69; 76.125-81.56; 87-92.44; 97.87-103.3; 108.75-114.19; 119.625-125.06; 130.5-135.94; 141.38-146.81; 152.25-157.69; 163.125-168.56	5.44-10.875; 16.31-21.75; 27.19-32.625; 38.06-43.5; 48.94-54.375; 59.81-65.25; 70.69-76.125; 81.56-87; 92.44-97.87; 103.3-108.75; 114.19-119.625; 125.06-130.5; 135.94-141.38; 146.81-152.25; 157.69-163.125; 168.56-173.6

3. Extraction des paramètres caractéristiques

Les paramètres suivants ont été utilisés pour représenter la distribution en temps-fréquence des signaux EEG:

- L'énergie du signal :

$$E_n = \sum_{k=1}^N |D_{n,k}|^2 \quad (5.1)$$

- L'entropie :

$$Ent_n = - \sum_{k=1}^N |D_{n,k}|^2 \log |D_{n,k}|^2 \quad (5.2)$$

- La déviation standard :

$$Std_n = \left(\frac{1}{N-1} \sum_{k=1}^N (D_{n,k} - \mu_n)^2 \right)^{1/2} \quad (5.3)$$

où : n est le niveau de décomposition ($n=1, 2, \dots, l$).

N est le nombre de coefficients détail ou approximation à différents niveaux.

$D_{n,k}$ étant les coefficients détails définis au chapitre 2.

μ_n est la moyenne donnée par :

$$\mu_n = \frac{1}{N} \sum_{k=1}^N D_{n,k} \quad (5.4)$$

4. Réduction des données

Dans ce paragraphe, nous utilisons les deux méthodes de réduction (PCA et LDA), décrites au chapitre 3, et représentées, respectivement par les figures 5.8 et 5.9.

Pour chaque signal EEG, nous procédons à l'extraction de paramètres caractéristiques de 6 coefficients de décomposition par la DWT et de 32 coefficients de décomposition par la WPT. Nous représentons les distributions d'énergie en utilisant les différentes combinaisons avec les deux paramètres caractéristiques (déviations standard, énergie) à partir des coefficients de décomposition : par la DWT (voir figure 5.8), et la WPT (voir figure 5.9), suivant les premières et les deuxièmes composantes principales.

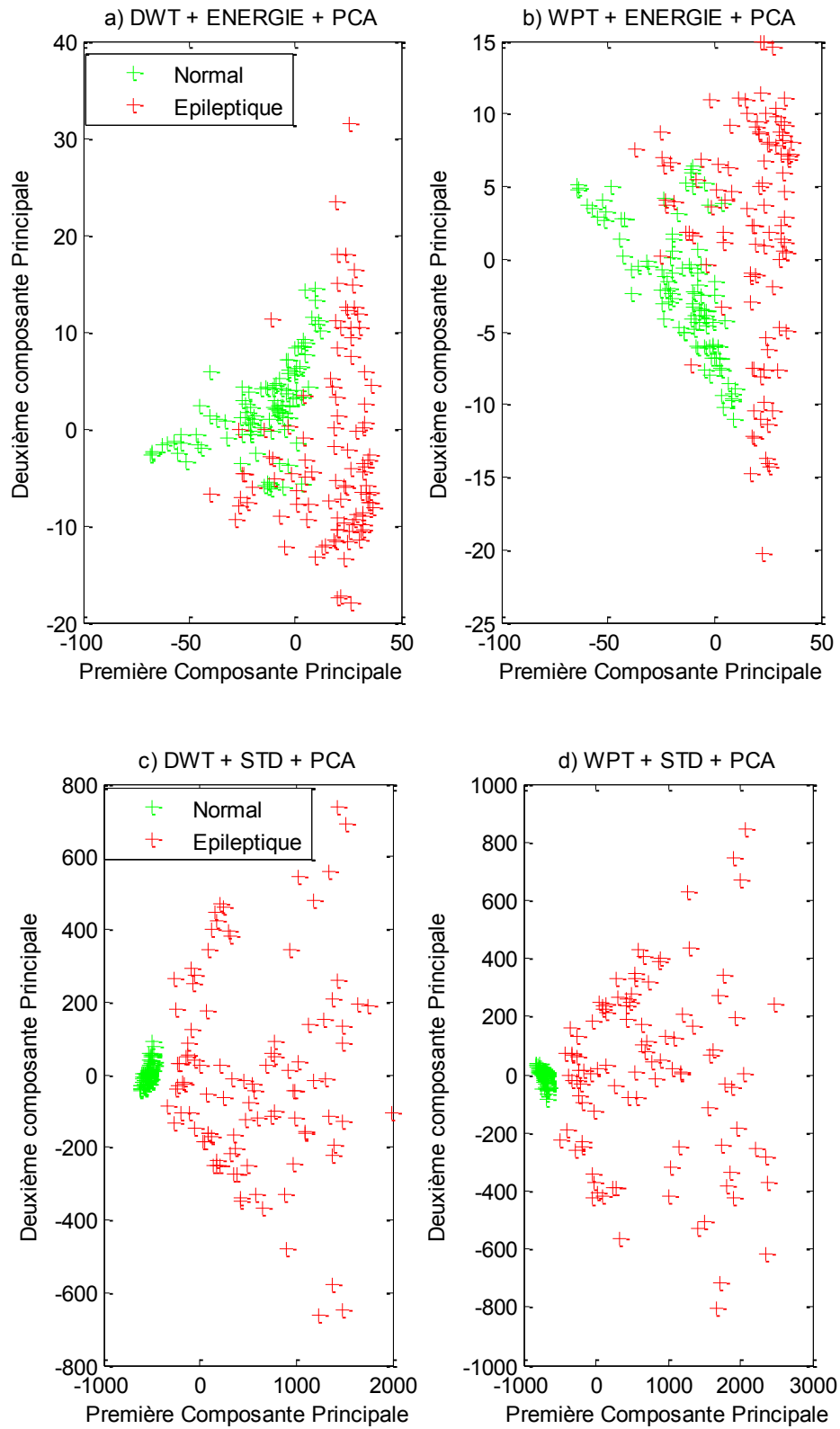


Figure 5.8 : Distribution énergétique (méthode de réduction PCA sur les paramètres caractéristiques statistiques et non-statistiques).

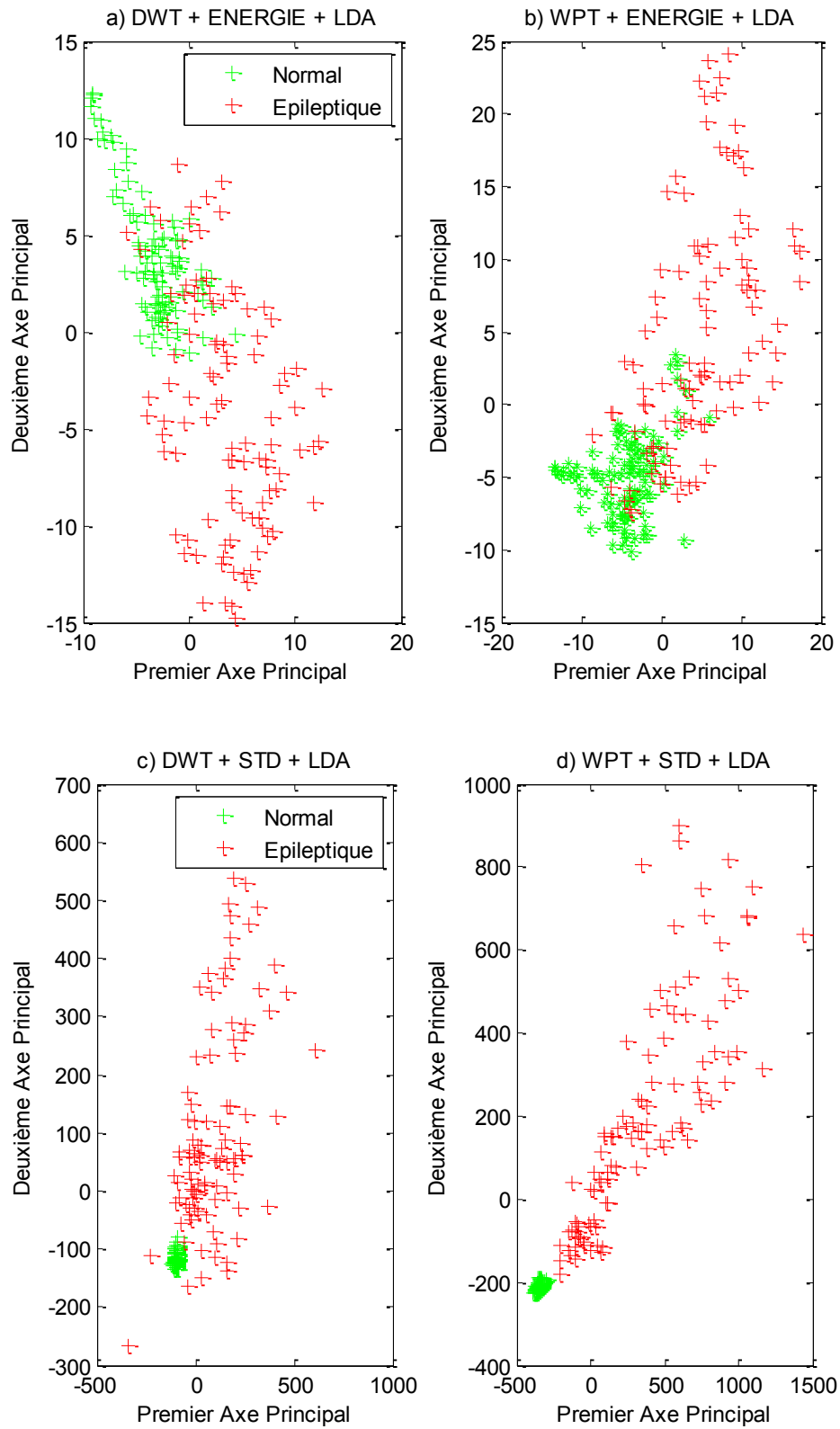


Figure 5.9 : Distribution énergétique (méthode de réduction LDA sur les paramètres caractéristiques statistiques et non-statistiques).

5. Classification

5.1 Méthodes utilisant la classification par SVM

Depuis une vingtaine d'années, les applications des techniques d'apprentissage artificiel ont connus un intérêt croissant dans la classification des signaux EEG. Elles permettent d'assister les médecins à établir un diagnostic approprié des pathologies neurologiques. Les techniques avancées, concernant la classification pour une analyse des signaux biomédicaux EEG, sont essentielles dans le développement et la compréhension de la recherche biomédicale cognitive.

Ces techniques de classification fonctionnent suivant deux étapes importantes : l'extraction des paramètres caractéristiques des données EEG, et la prise de décision en utilisant ces paramètres obtenus comme des entrées du procédé de classification. Il est important de noter que ces paramètres caractérisent le comportement des données EEG originales. Les machines à vecteur support, mentionné au chapitre 4, sont utilisées pour la classification (suivant deux classes) des deux groupes set (A) et set (E), utilisant les paramètres statistiques et non statistiques réduits comme entrées du classificateur. L'approche proposée se définit par rapport au taux de la précision de cette classification [76]. Afin d'obtenir un taux de précision élevé, nous utilisons une méthode de la validation croisée (*k-fold cross-validation* mentionnée au paragraphe 5.1.1) et nous évaluons statistiquement les performances de la méthode de classification proposée.

5.1.1 Validation croisée

La validation croisée (*cross-validation*) est une méthode d'estimation de fiabilité d'un modèle [77]. C'est une pratique de partitionnement statistique d'un échantillon de données en sous-ensembles, i.e. elle permet de tester un modèle d'apprentissage. La validation croisée se présente en trois sous-méthodes (*holdout method*, *k-fold cross-validation* et *leave-one-out cross-validation*). La plus répandue est la méthode *k-fold*, avec $k \in [4 \text{ à } 10]$.

Soit une base d'apprentissage A_j contenant j éléments : $A_j = \{x_1; \dots; x_j\}$, la validation croisée consiste à appliquer les cinq étapes suivantes :

1. Diviser l'ensemble des exemples en k sous-ensembles disjoints de taille j/k
2. Apprendre sur les $(k - 1)$ sous-ensembles restants.
3. Calculer l'erreur sur la $k^{\text{ième}}$ partie.
4. Répéter le processus j fois.

5. Obtenir l'erreur finale en calculant la moyenne des k erreurs précédentes.

La validation croisée utilise toutes les données. Elle permet d'obtenir une estimation de l'erreur de généralisation, et cela permet d'éviter le sur-apprentissage [78]. Dans cette section, nous utilisons une méthode de *10-fold cross validation*. Autrement dit, l'ensemble est divisé en 10, ensuite on sélectionne un des dix échantillons comme un ensemble pour la validation et les 90% autres échantillons constitueront l'ensemble d'apprentissage. L'erreur est estimée en calculant l'erreur quadratique moyenne. Ensuite, on répète l'opération en sélectionnant un autre échantillon de validation parmi les $(k-1)$ échantillons restants qui n'ont pas été utilisés pour la validation du modèle. L'opération se répète ainsi *dix* fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. La moyenne des dix erreurs quadratiques moyennes est enfin calculée pour estimer l'erreur de prédiction.

5.1.2 Machines à vecteur support

Après avoir obtenu les différentes caractéristiques représentatives de toutes les données des différents canaux EEG, nous utilisons ces paramètres comme entrées pour l'algorithme SVM. Après des tests sur les fonctions noyaux, décrites au chapitre 4, la fonction du noyau RBF est utilisée comme une fonction noyau optimale pour la classification par SVM. Dans la littérature [67, 80, 81, 88], différents noyaux, plus particulièrement le noyau gaussien, ont été testés et utilisés dans la classification des signaux biomédicaux. Et dans [78], *Hsu et al.* donnent des exemples de calcul exhaustif du taux de validation croisée pour une grille des valeurs des hyper-paramètres (C, σ) . Nous avons utilisé plusieurs valeurs de C et de σ . Les résultats des représentations de la classification SVM utilisant trois paramètres caractéristiques et deux méthodes de réduction sont illustrés sur les figures 5.10 et 5.11, avec une sélection des deux hyper-paramètres $C=100$ et $\sigma=10$.

Ces deux figures montrent les signaux EEG sains (figure 5.10) et les signaux EEG épileptiques (figure 5.11) pendant l'apprentissage et pendant le test, ainsi que leur frontière séparatrice (comme indiqué au chapitre 4, paragraphe 2). Les vecteurs supports sont aussi représentés sur ces deux figures.

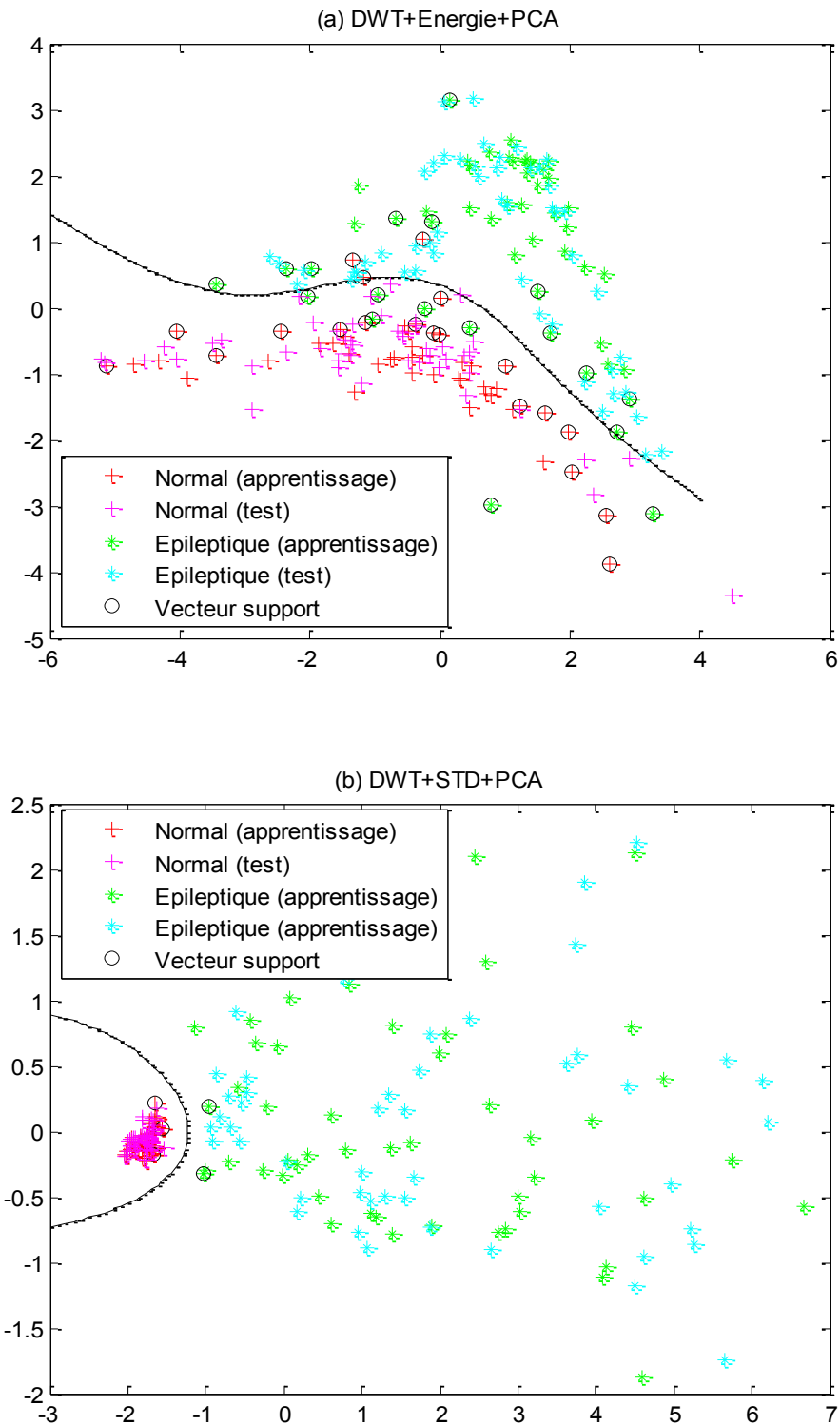


Figure 5.10 : Classification par SVM (distribution énergétique du signal en utilisant la PCA).

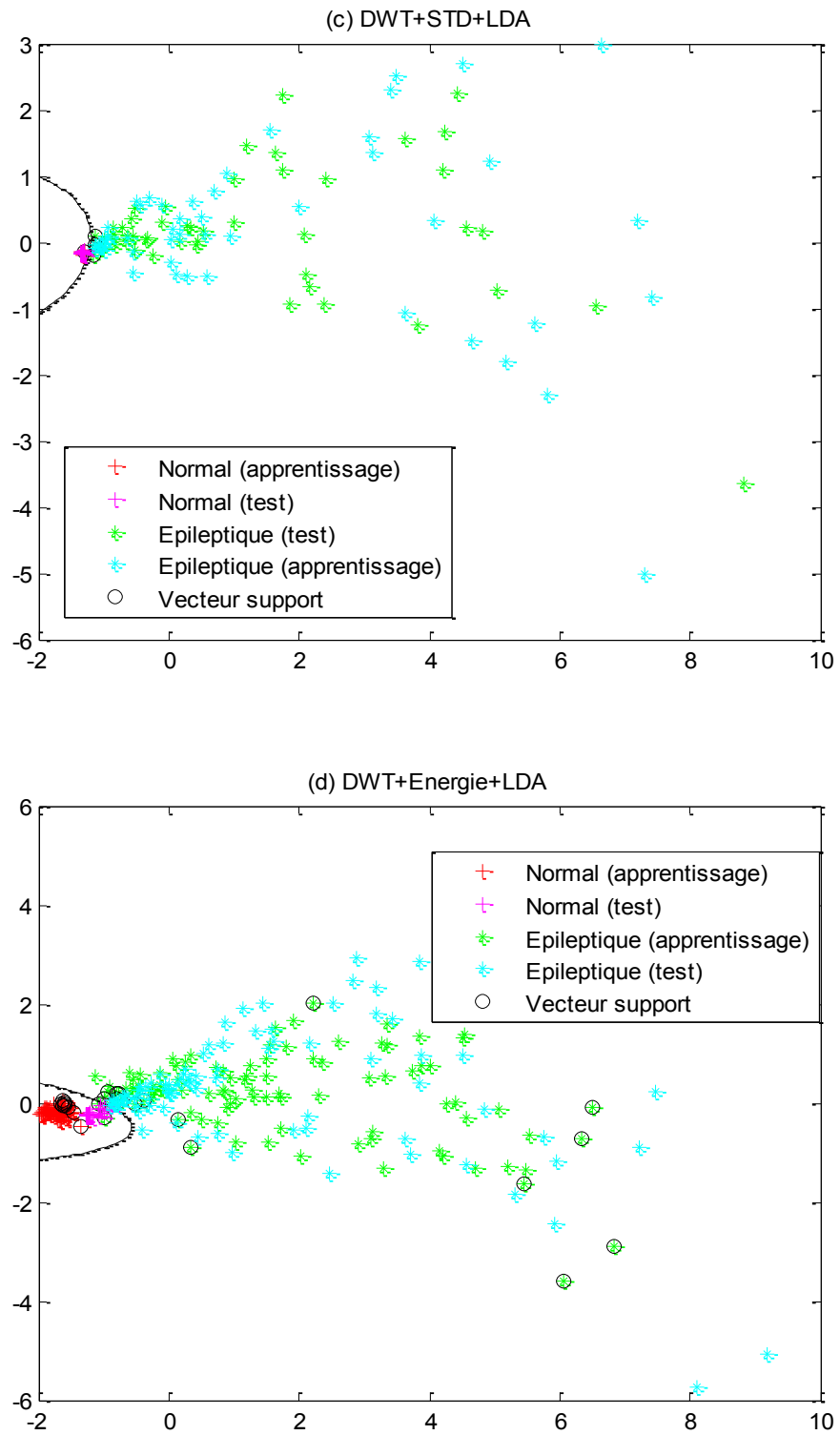


Figure 5.11 : Classification par SVM (distribution énergétique du signal en utilisant la LDA).

5.2 Performances et évaluations statistiques

Les trois paramètres caractéristiques, i.e. $\{E, Ent, Std\}$, sont calculés à partir des coefficients de décomposition de la DWT et de la WPT sur les cinq niveaux de décomposition. Afin d'évaluer les performances de mesure de notre système de classification, nous avons calculé le taux de la sensibilité du système, de sa spécificité ainsi que la précision de la classification.

Le tableau 5.3 représente la moyenne des taux de pourcentage (TVP, TVN, Précision) obtenus en utilisant les différentes méthodes citées précédemment. La méthode proposée nous a permis d'obtenir un taux optimal de précision de classification de 100% en utilisant un prétraitement par la WPT et en se basant sur la LDA comme méthode de réduction du vecteur caractéristique de la déviation standard de tous les coefficients de décomposition obtenus.

Dans cette partie, nous étudions les potentiels de l'application de l'algorithme SVM à noyau RBF pour obtenir la meilleure caractéristique représentative de tous les canaux d'enregistrement EEG; ces vecteurs caractéristiques sont utiles pour l'entrée du classificateur.

Comme représenté sur le tableau 5.3 et pour le même groupe de données, le classificateur SVM donne une précision de classification de 96,80% et 90,59%, respectivement, pour une utilisation de la distribution énergétique (E) et de la réduction de donnée PCA, obtenue à partir de la DWT et de la WPT. Un taux de précision de classification de 99% et 91%, respectivement, pour une utilisation de la distribution entropique (Ent) et de la PCA, obtenue à partir de la DWT et de la WPT.

Pour l'étude correspondant à la LDA, le tableau 5.3 montre un taux de précision de classification de 86,70% et 88,01%, respectivement, pour une utilisation de la distribution énergétique obtenue à partir de la DWT et de la WPT; le taux de précision est de 99,50% et 97% pour une utilisation de l'entropie des signaux décomposés par la DWT et la WPT, respectivement.

Nous remarquons que le traitement fait à base de la déviation standard nous permet d'extraire les informations représentatives et pertinentes des signaux d'enregistrement EEG (figures 5.12 et 5.13).

Tableau 5.3 : Résultats de mesures obtenues à partir de différentes méthodes utilisées. (TVP: sensibilité ; TVN: spécificité).

Méthode utilisée	TVP (%)	TVN (%)	Précision (%)
DWT+Std+PCA	98.78	99.30	99.04
DWT+Std+LDA	99.70	99.90	99.80
DWT+E+PCA	95.90	97.70	96.80
DWT+E+LDA	85.30	88.10	86.70
DWT+Ent+PCA	98.35	99.65	99.00
DWT+Ent+LDA	99.40	99.41	99.50
WPT+Std+PCA	99.88	100	99.94
WPT+Std+LDA	100	100	100
WPT+E+PCA	95.38	85.80	90.59
WPT+E+LDA	98.42	77.60	88.01
WPT+Ent+PCA	85.70	96.30	91.00
WPT+Ent+LDA	96.30	97.70	97.00

Les figures 5.12 et 5.13 représentent les histogrammes des différents taux de précision de la classification du set A (personnes saines avec yeux ouverts) et du set E (patients épileptiques en pleine crise) obtenus en utilisant, respectivement, la DWT et la WPT. Ces deux figures illustrent les résultats obtenus sur le tableau 5.3.

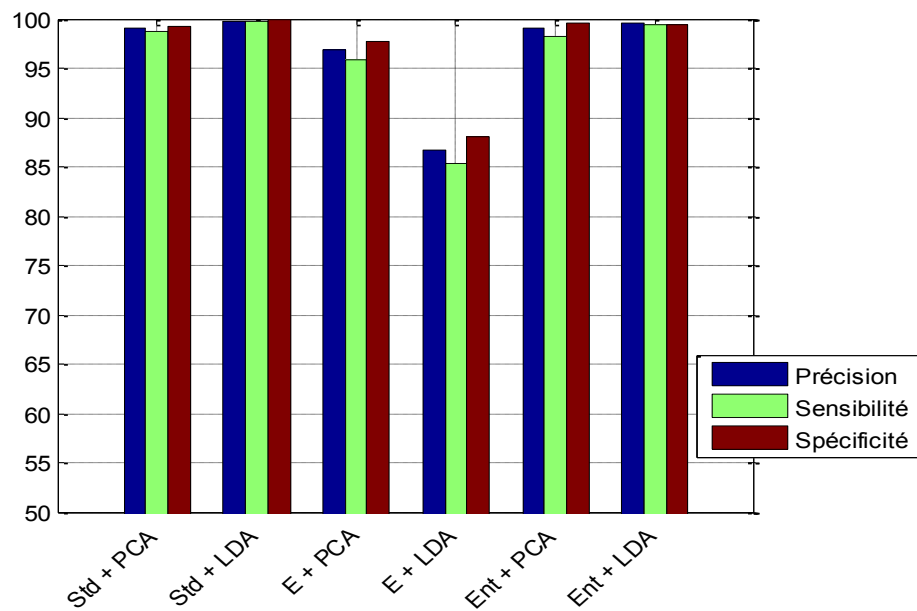


Figure 5.12 : Valeur moyenne des taux de précision de classification des sets A et E basé sur l'utilisation de l'énergie, de l'entropie et la déviation standard appliqués aux coefficients d'ondelette DWT.

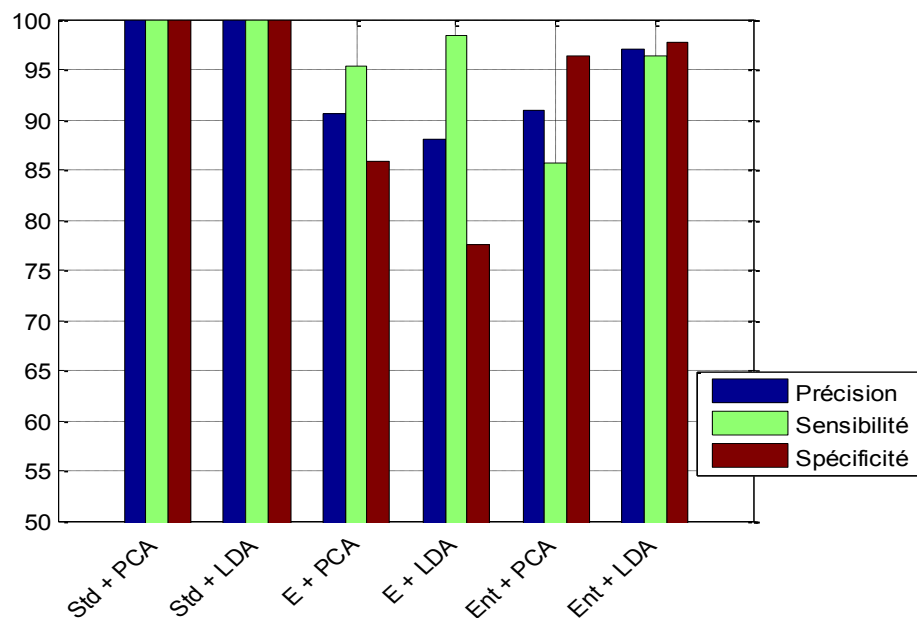


Figure 5.13 : Histogramme des taux de précision de classification (par SVM) des sets A et E basé sur l'utilisation de la WPT, l'énergie, de l'entropie et la déviation standard.

Dans la littérature, de nombreuses techniques ont été utilisées pour obtenir des représentations en temps-fréquence et des caractéristiques pertinentes ont été utilisées à des fins de classification. Sur le tableau 5.4, nous comparons les résultats obtenus avec d'autres méthodes existantes dans la littérature et qui utilisent les mêmes groupes (set A et set E) de la base de données EEG. Aucune méthode de classification de signaux EEG n'a utilisé, dans la littérature, la combinaison : prétraitement par la WPT, réduction par la LDA des vecteurs caractéristiques présentés comme entrée au classificateur SVM [76]. Cette méthode atteint une sensibilité, une spécificité et une précision de classification d'un taux atteignant 100% pour les données EEG épileptiques (critiques).

Tableau 5.4 : Valeurs des précisions obtenues dans la littérature et utilisant la même base de données (set A et set E).

Auteurs	Méthodes utilisées	Précision(%)
[60] Subasi (2007)	DWT + Mixture of Expert classification	94.50
[51] Polat and Günes (2008)	DWT+DFT+ Auto-regressive model + Decision Tree	99.32
[58] Subasi and Gursoy (2010)	DWT+PCA+ LDA+ICA + classification SVM	98.75(PCA) 100(LDA) 99.5(ICA)
[93] Wang, Miao and Xie (2011)	WPT+ Entropie-hiérarchique+ classification K-NN	99.44
[54] Übeyli (2010)	Burg autorégressive + classification LS-SVM	99.56
[91] Gandhi , Panigrahi and Anand (2011)	DWT+ Energie + classification PNN	99.35
La méthode proposée	WPT + STD + LDA + classification SVM	100

Une autre approche a été effectuée en déterminant différents ensembles et d'autres caractéristiques statistiques extraites de chaque groupe. De la base de données originale, nous avons choisi les sets A, C et E, i.e. des groupes d'enregistrements EEG sains, inter-critiques et critiques, respectivement. Les caractéristiques statistiques suivantes ont été utilisées pour représenter la distribution en temps-fréquence des signaux EEG :

- Le maximum (*Max*) : traduit la plus grande valeur qu'atteint le signal.
- Le minimum (*Min*) : traduit la plus petite valeur qu'atteint le signal.
- La gamme (*Gam*): traduit la distance entre la valeur maximale et minimale du signal.
- La déviation standard (*Std*) : mesure la dispersion d'une variable aléatoire réelle.

Comme décrits précédemment, les paramètres caractéristiques non statistiques, i.e. l'énergie (*E*) et l'entropie (*Ent*), ont aussi été utilisés.

La même démarche, que dans les paragraphes 2 et 3 de ce chapitre, a été adoptée pour l'analyse des trois cent signaux EEG issus des sets A, C et E. Chaque canal se compose de 4096 échantillons avec une durée de 23,6s. Nous utilisons une ondelette *Db2* pour la décomposition en cinq niveaux par la DWT et avec trois niveaux de décomposition par la WPT. Les figures 5.14 et 5.15 représentent, respectivement les coefficients de décomposition par la DWT et par la WPT du signal EEG inter-critique. L'ensemble de six paramètres caractéristiques $\{Max, Min, Gam, Std, E, Ent\}$ sont calculés à partir des coefficients de décomposition de la transformée en ondelette basée sur la DWT et sur la WPT. Notre étude concernera trois groupes de données EEG : le groupe 1 constitué de l'ensemble (set A-C), le groupe 2 composé de l'ensemble (set A-E) et du groupe 3 formé du set A et du set (C-E). Une réduction de dimension de ces paramètres par la méthode PCA a été introduite avant de classer, en une classe de sujets sains et une autre classe de patients épileptiques, par l'algorithme SVM à noyau *RBF* (*Radial Basis Function*) décrit dans la section 5.4 du chapitre 4.

Les coefficients de décomposition de la DWT sont au nombre de six et ceux de la WPT sont égaux à huit. Par conséquent, nous obtenons, pour chaque canal, une matrice de taille $[4096 \times 36]$ concernant la DWT et de taille $[4096 \times 48]$ concernant la WPT, ensuite une réduction de dimension par PCA a été appliquée afin d'éliminer la redondance d'information produite sur les enregistrements EEG et qui risque de diminuer les performances du classificateur.

Les figures 5.14 et 5.15 montre cette décomposition d'un enregistrement EEG du groupe de données inter-critiques effectuée à l'aide de la DWT et de la WPT.

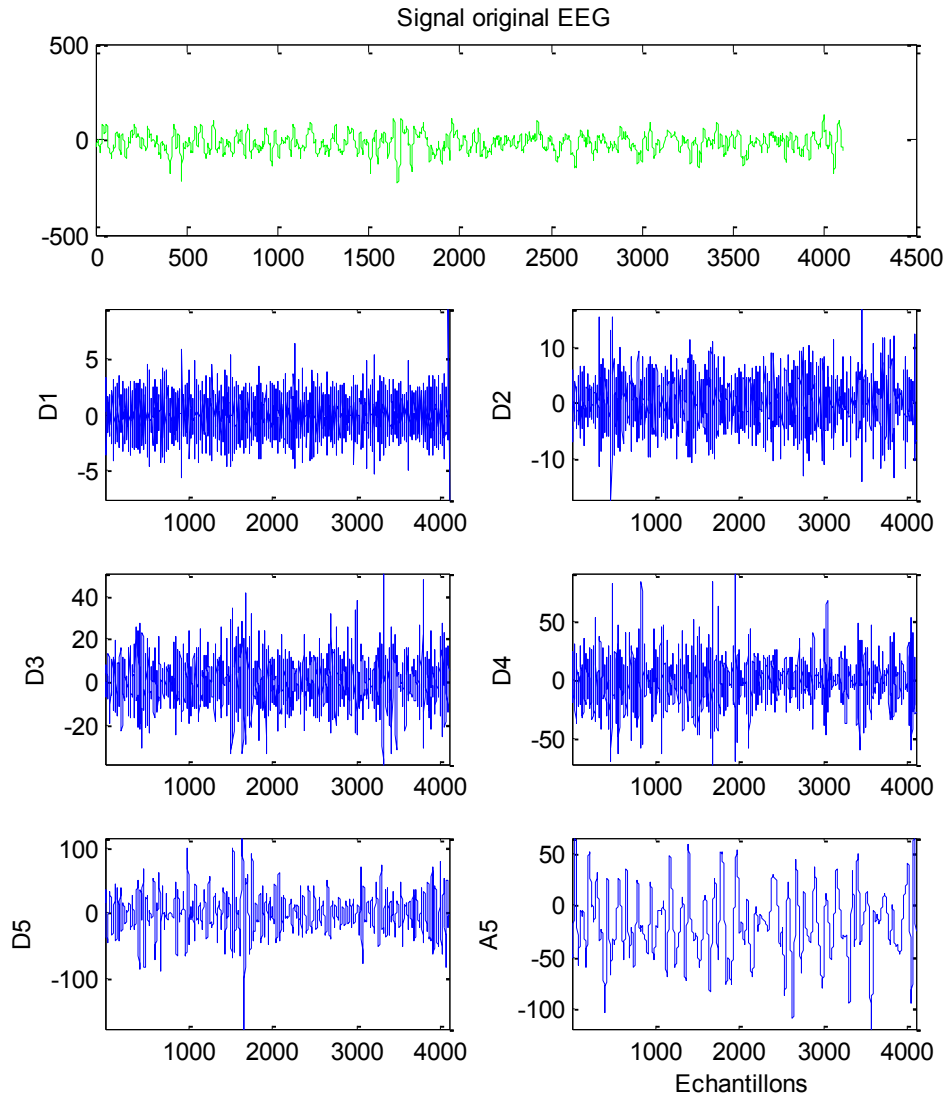


Figure 5.14 : Décomposition du signal EEG (set C) basée sur la DWT (les amplitudes sont données en μV).

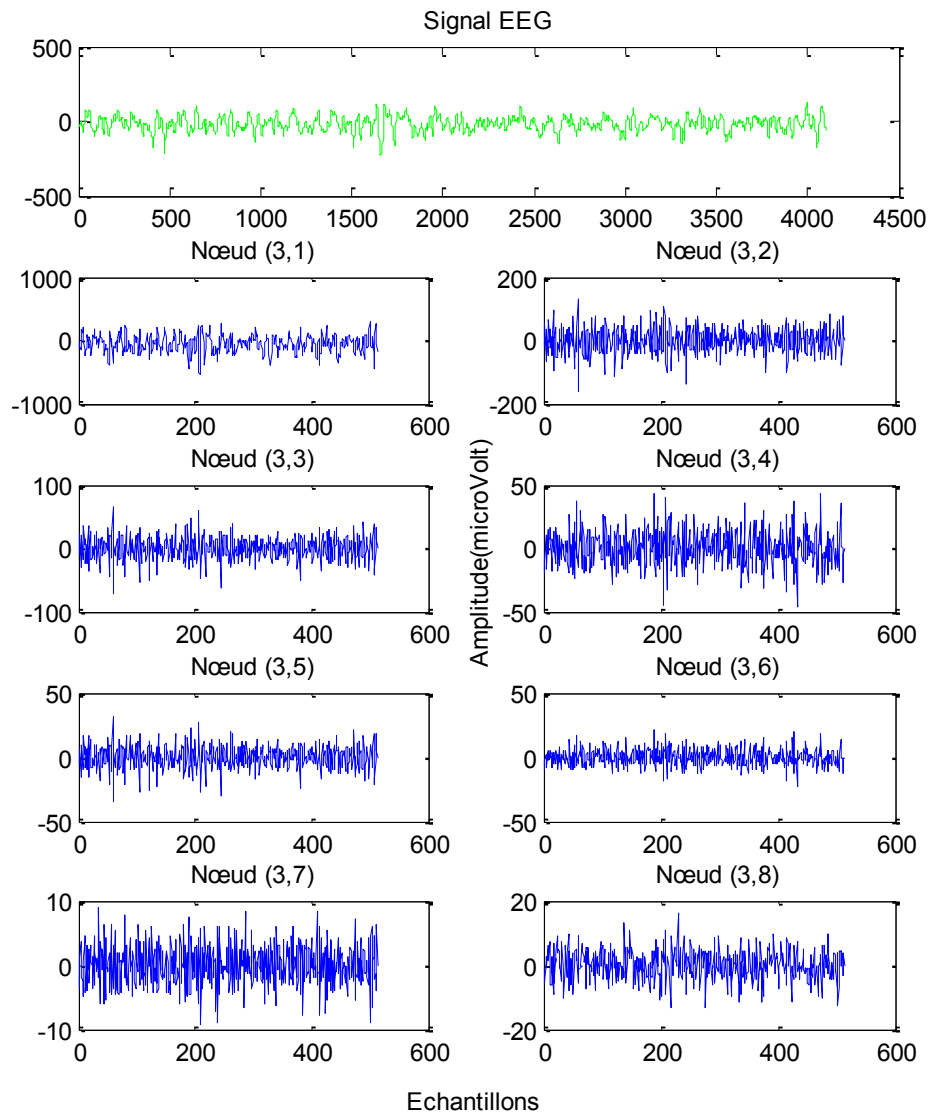


Figure 5.15 : Décomposition du signal EEG (set C) basée sur la WPT et représentant les coefficients de décomposition aux différents nœuds des paquets d'ondelette du troisième niveau.

Dans cette étude, nous avons évalué la stabilité des performances de l'algorithme SVM utilisé en fonction de différentes mesures statistiques comme la sensibilité, la spécificité et le taux de précision de la classification. En effet, nous utilisons la

méthode de validation croisée (*k-fold cross validation*) décrite plus haut au paragraphe 5.1.1. Cette méthode est appliquée avec différentes valeurs de pourcentages (*Apprentissage-Test*) : 50%-50%, 80%-20% et 90%-10% qui correspondent, respectivement, aux valeurs $k=2$, $k=5$ et $k=10$. Les résultats concernant le groupe 1 sont représentés par la figure 5.16. Les taux de la sensibilité, de la spécificité et de la précision de la classification de la méthode proposée sont donc calculés à partir de l'ensemble du test. La figure 5.16 représente les performances de la classification, en utilisant la DWT et en faisant varier k ; les performances de la classification, utilisant la WPT, concernant le groupe 1 sont représentées par la figure 5.17.

Les résultats obtenus des taux de précision de la classification du groupe 1 donnent des valeurs de {92.88%, 93.55%, 94.58%, 98.61%, 97.55%, 96.51%} pour les paramètres {*Max*, *Min*, *Gam*, *Std*, *Ent*, *E*} respectivement, en prenant $k=2$; des valeurs de {93.03%, 93.93%, 94.55%, 98.70%, 98.09%, 96.87%} pour les paramètres {*Max*, *Min*, *Gam*, *Std*, *Ent*, *E*} respectivement, en prenant $k=5$; et des valeurs de {93.07%, 94.01%, 94.75%, 98.95%, 98.24%, 96.95%} pour le même ensemble des paramètres caractéristiques en prenant $k=10$; ces valeurs sont représentées par la figure 5.16 où nous remarquons que pour $k=10$, les taux de précision de la classification du groupe 1 ont une valeur plus élevée par rapport aux valeurs où $k=2$, ou à $k=5$; ceci est dû au fait que la procédure *10-fold cross validation* utilise 90% des ensembles de vecteurs de caractéristique pour la phase d'apprentissage, et par conséquent, les résultats expérimentaux obtenus montrent la performance du classificateur SVM à noyau RBF.

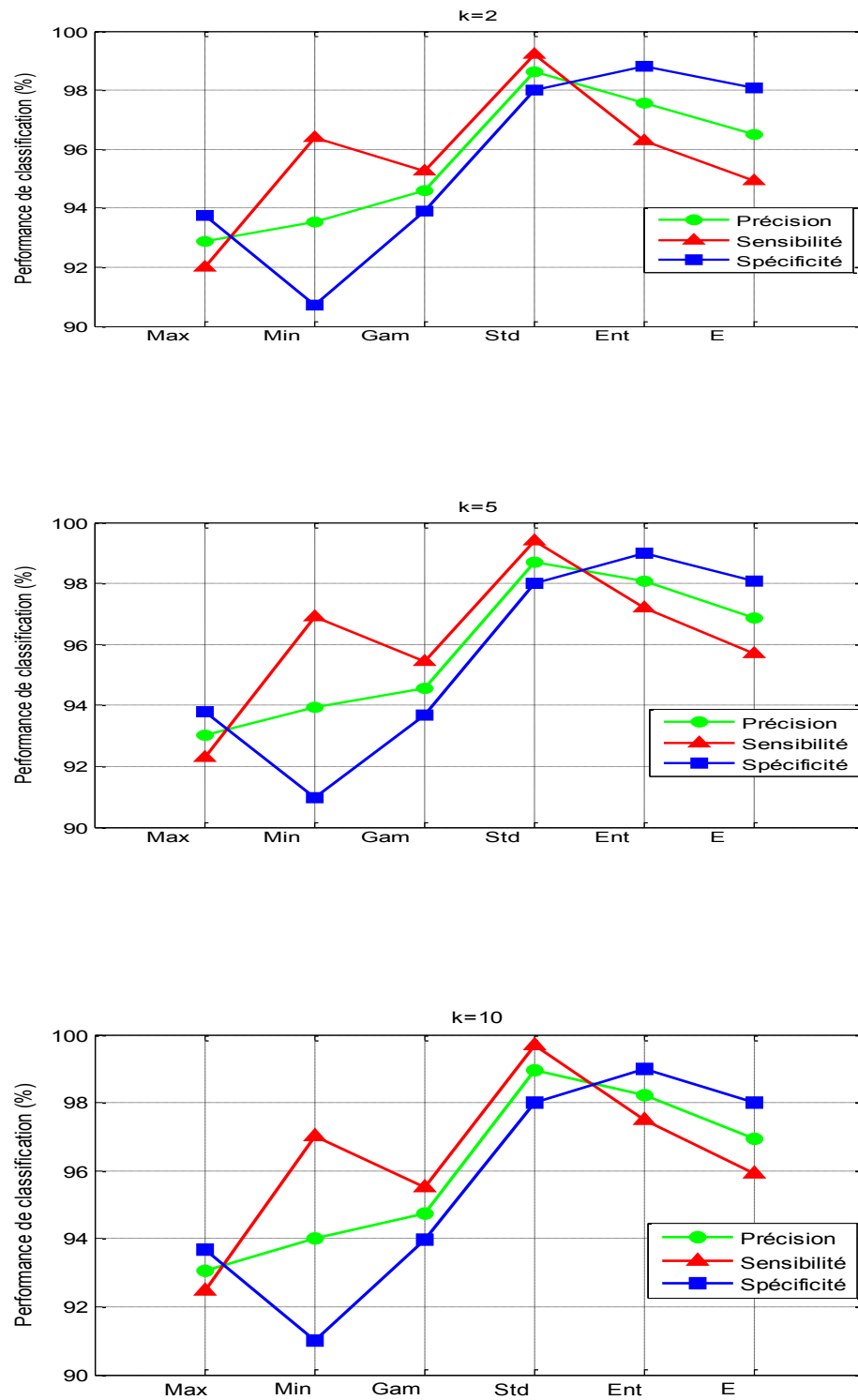


Figure 5.16 : Performances de classification (basée sur la DWT) du groupe (A-C), utilisant la méthode 'k-fold cross validation' avec k= 2, 5, 10.

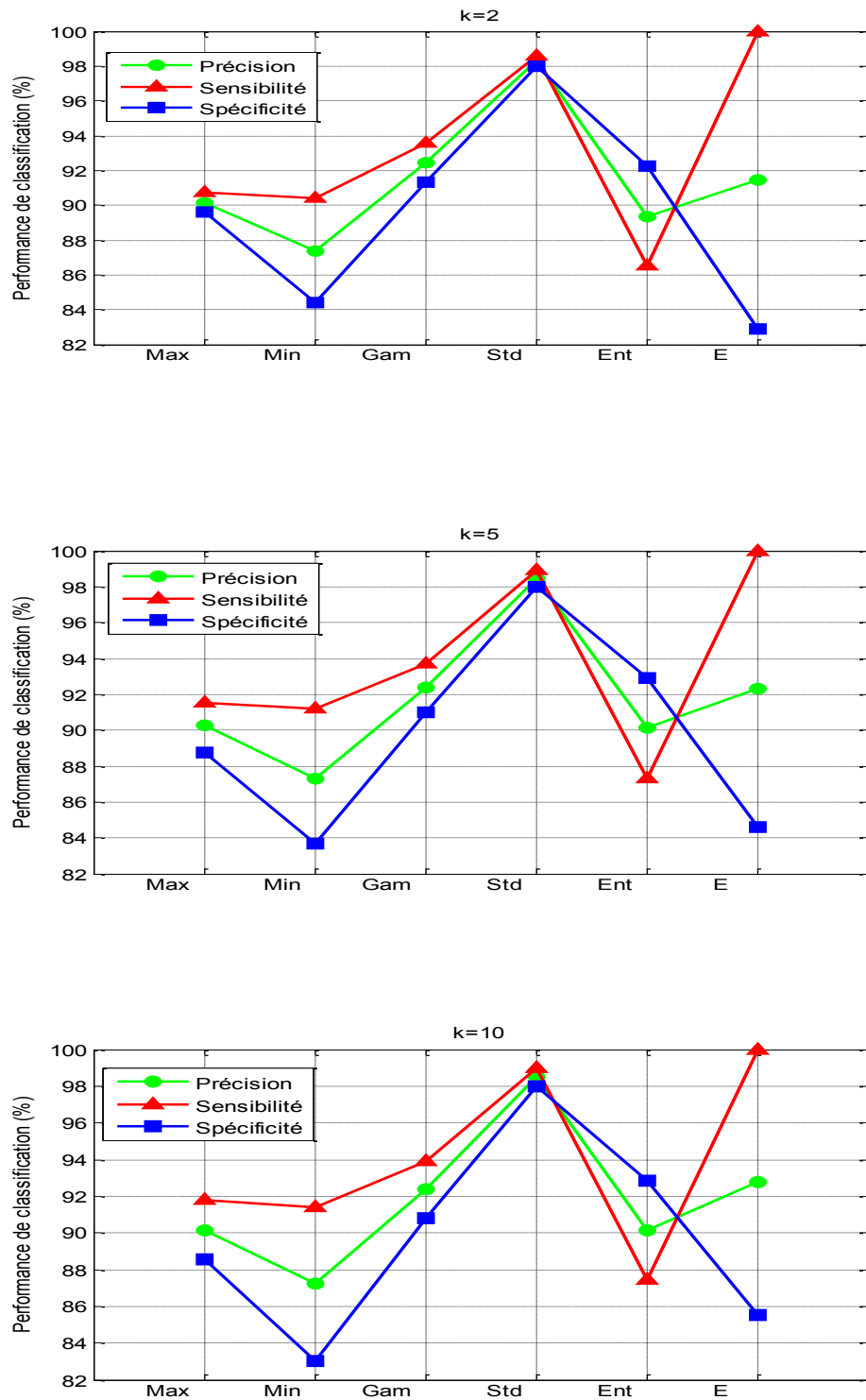


Figure 5.17 : Performances de classification (basée sur la WPT) du groupe (A-C), utilisant la méthode 'k-fold cross validation' avec k= 2, 5, 10.

Le taux de précision de la classification du groupe 2 est représenté par la figure 5.18. Nous avons aussi utilisé trois valeurs de k ($k=2, 5, 10$), ensuite calculé l'exactitude de la classification. Sur la figure 5.18, la méthode 10-fold cross validation est celle qui donne le meilleur taux. Cette procédure divise les ensembles de vecteurs de caractéristique dans dix partitions distinctes. Une partition est utilisée pour le test, tandis que les neuf autres partitions sont utilisées pour l'apprentissage du modèle. Pour améliorer encore cette estimation, la procédure est répétée dix fois. La précision moyenne sur les dix estimations obtenues à partir des résultats des tests est considérée comme le critère d'évaluation des performances de l'algorithme de classification.

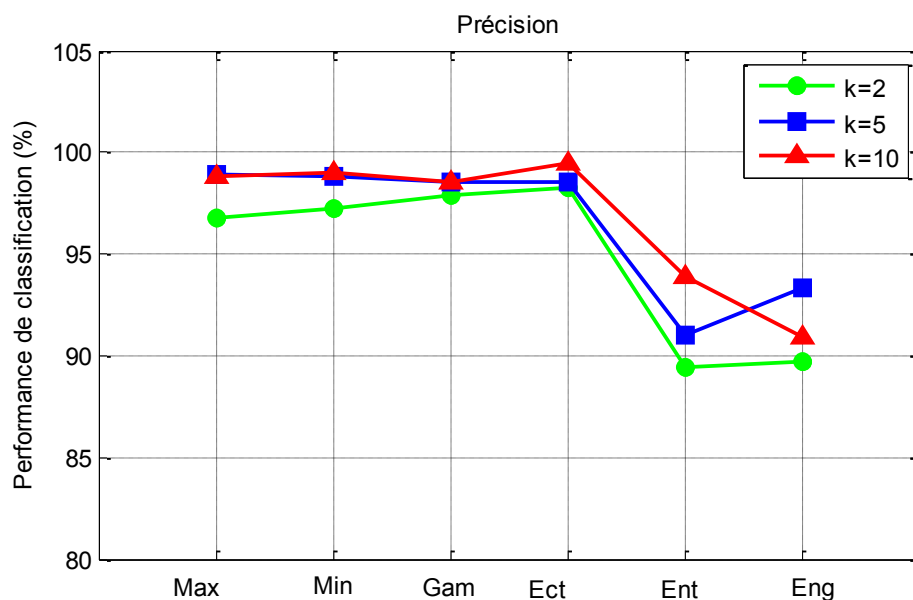


Figure 5.18 : Performances de classification (basée sur la DWT), du groupe (A-E), utilisant 'k-fold cross validation' avec $k= 2, 5, 10$.

Les résultats obtenus pour le groupe 3, concernant les sujets sains et les patients épileptiques (en pleine crise ou entre deux crises), sont représentés sur le tableau 5.5. Les trois cent signaux EEG sont prétraités à l'aide de la DWT et la WPT, et pour les différents coefficients de décomposition, six vecteurs caractéristiques sont calculés.

L'algorithme utilisé pour le troisième groupe a un comportement satisfaisant en fonction du taux de précision, de la sensibilité et de la spécificité (Tableau 5.5).

Tableau 5.5 : Taux des performances de classification du groupe (A- (C-E)), en utilisant la DWT, La WPT, et la méthode '10-fold cross validation'.

Transformée	Performances	Max	Min	Gam	Std	Ent	E
DWT	Précision (%)	94.01	94.55	94.75	99.07	98.11	93.86
	Sensibilité (%)	95.15	95.45	95.54	99.90	99.99	93.16
	Spécificité (%)	94.34	93.66	93.95	98.24	96.23	94.55
WPT	Précision (%)	91.33	88	93.25	99.77	96	90.17
	Sensibilité (%)	100	100	100	100	100	100
	Spécificité (%)	82.66	76	86.50	99.54	92	80.34

Pour le groupe 3, les figures 5.19 et 5.20 montrent la variation du pourcentage du taux de classification en fonction des six paramètres caractéristiques $\{Max, Min, Gam, Std, Ent, E\}$, utilisant la DWT et la WPT pour le prétraitement des signaux EEG.

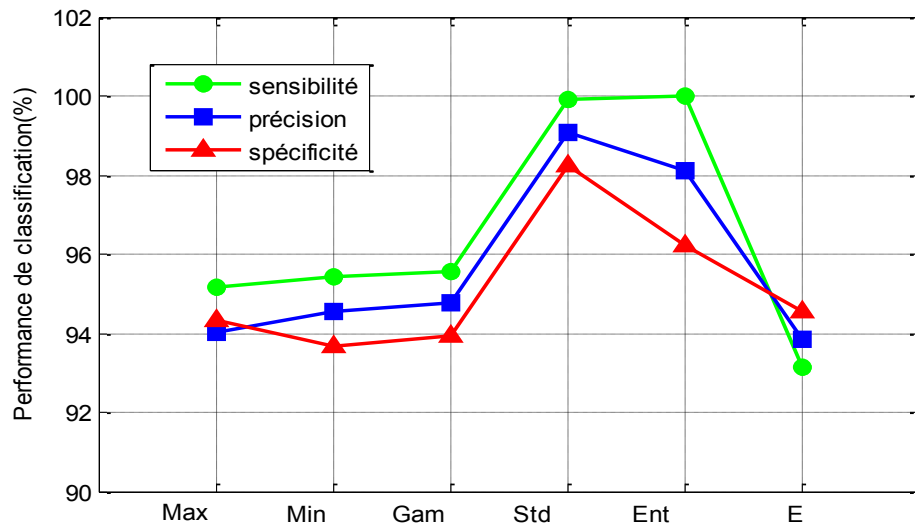


Figure 5.19 : Performances de classification (basées sur la DWT), du groupe (A-(C-E)), en fonction de six caractéristiques.

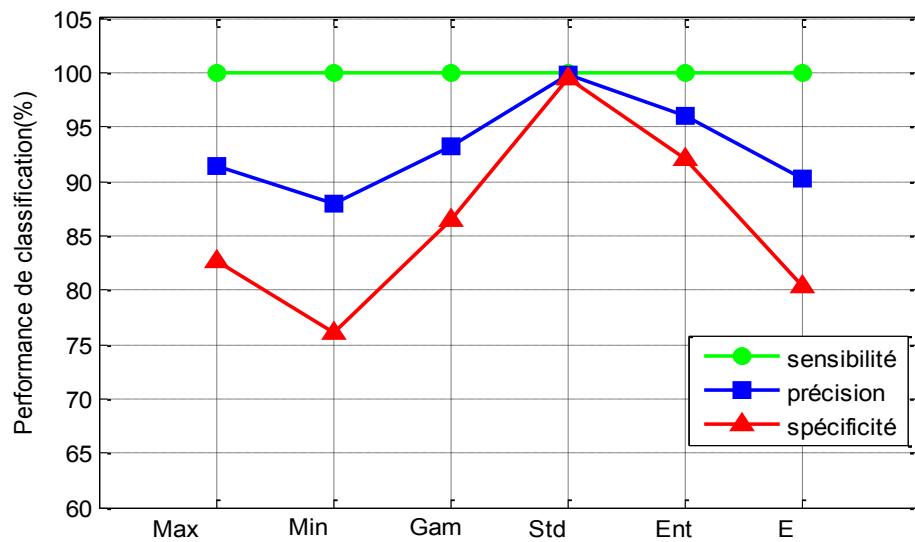


Figure 5.20 : Performances de classification (basées sur la WPT), du groupe (A-(C-E)), en fonction de six caractéristiques.

Nous avons étudié les potentiels de l'application de l'algorithme SVM à noyau RBF utilisant, comme entrée, six paramètres caractéristiques et calculés à partir des coefficients de décomposition par la transformée en ondelette (DWT et WPT). Les résultats du groupe 3, i.e. les signaux EEG appartenant à des personnes saines, épileptiques entre deux crises et en pleine crise obtenus et montrés sur le tableau 5.5 (où la précision optimale est 99,07%), sont relativement plus faibles que ceux obtenus pour le groupe (A-E), i.e. personnes saines et épileptiques, dans la première partie (la précision optimale est 99,50% ; voir la figure 5.18). En revanche, l'extraction du paramètre de la déviation standard comme caractéristique statistique et pertinente, a été démontrée suivant différents groupes et différentes combinaisons étudiées, pour obtenir les meilleurs résultats d'une classification de données EEG épileptiques.

5.3 Méthodes utilisant la classification par k-plus proches voisins

D'autres travaux ont été menés sur les mêmes groupes des signaux EEG étudiés dans la section 5.2, i.e. set A, set C et Set E. Dans cette étude, la stabilité des performances du classificateur K-plus proches voisins (K-NN) sera évaluée en fonction de différentes mesures statistiques, comme la sensibilité, la spécificité et la précision de la classification. Le paramètre de similarité (plus proche voisin), par défaut, est la distance euclidienne (définie au chapitre 4, paragraphe 3) et la valeur de K est réglée à 5 (erreur de classification minimale).

Les résultats obtenus donnent les valeurs moyennes du pourcentage du taux de précision de classification, de sa sensibilité et de sa précision; ces valeurs sont représentées sur le tableau 5.6.

Tableau 5.6 : Résultats de mesures obtenues à partir de la classification par l'algorithme K-NN pour le set (A-E) et le set (A-(C-E). (TVP : sensibilité ; TVN : spécificité).

Base de données	Performances	Max	Min	Gam	Std	E	Ent
A-E	Précision (%)	96.79	97.23	98.50	98.50	90.90	93.93
	TVP (%)	99.69	98.61	98	99.21	93.05	98.43
	TVN (%)	93.89	95.85	98.91	97.79	88.75	89.43
A-(C-E)	Précision (%)	97.50	97.23	97.97	98.30	90.90	92.97
	TVP (%)	96.79	96.5	99.03	97.35	92.05	96.51
	TVN (%)	98.20	97.96	96.91	99.25	89.75	89.43

La figure 5.21 montre la performance de cet algorithme présentée sous forme de *boxplot*¹⁹ où la ligne centrale représente la médiane, i.e. la centrale des 50% des données. Les lignes représentant la frontière basse et la frontière haute indiquent, respectivement, 25% et 75% des données à représentées. Les extensions de moustaches sont des points de données les plus extrêmes et les points '+' représentent les points aberrants (outliers).

¹⁹ Appelées aussi boîtes à moustaches.

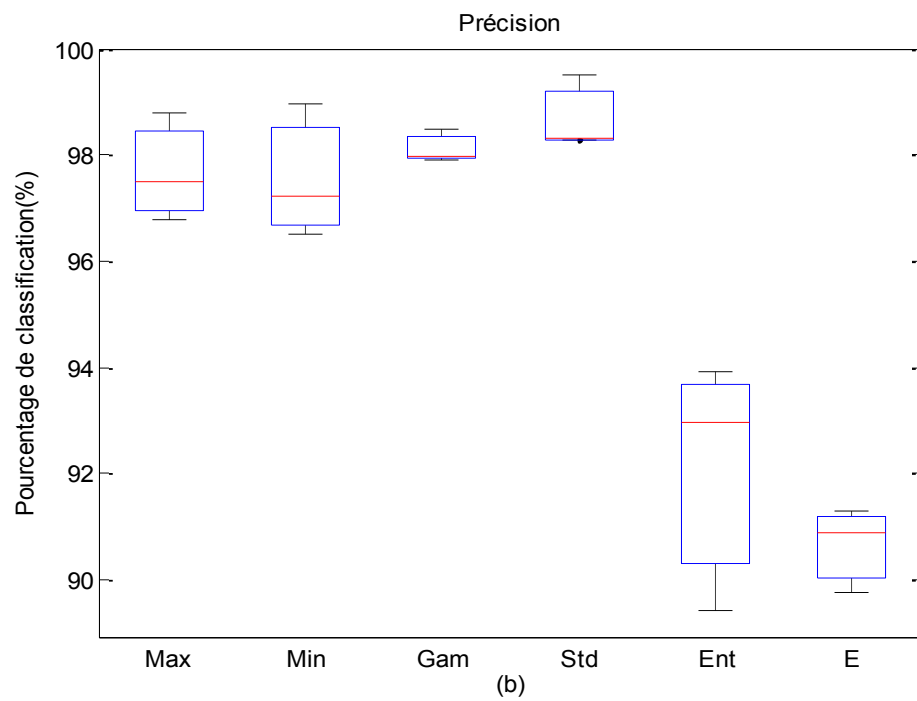
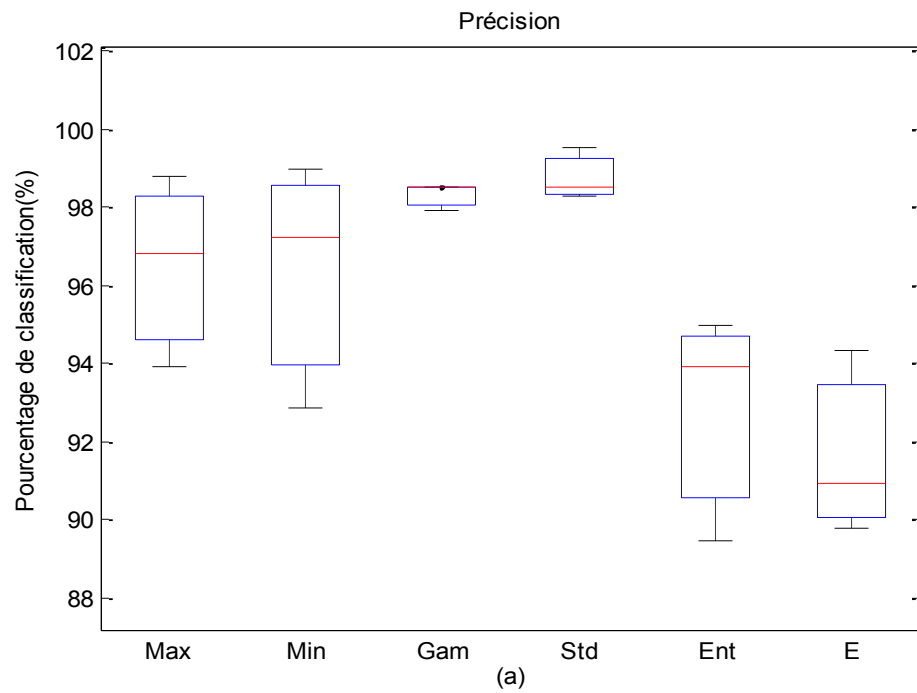


Figure 5.21 : Boxplot montrant les performances de classification utilisant les K-plus proches voisins ; (a): Set (A-E), (b): Set (A-(C-E)).

Les résultats obtenus sur la précision de classification avec la méthode des K-plus proches voisins sont plus faibles que ceux obtenus avec un classificateur SVM à noyau RBF. Comme le montre la figure 5.18, le taux de précision pour le groupe 2, utilisant deux cent signaux, est de 99,50% pour le classificateur SVM, tandis que cette valeur est de 98,50% (voir figure 5.21) pour le classificateur K-NN. Le taux de précision montré sur la figure 5.19 pour le groupe 1, utilisant les sets A-C-E, est de 99,07% pour le classificateur SVM, tandis que cette valeur est de 98,30% (voir figure 5.21) pour le classificateur K-NN, ceci reflète la performance du classificateur SVM à noyau RBF par rapport au classificateur K-NN.

5.4 Méthodes utilisant les réseaux de neurone artificiels

Au cours des deux dernières décennies, un intérêt croissant apparaît pour la classification des signaux EEG utilisée comme système d'aide au diagnostic de certaines pathologies du cerveau. Les techniques avancées de classification pour l'analyse des enregistrements EEG sont essentielles pour le développement de la recherche biomédicale. La technique de classification est basée d'une part, sur le choix des caractéristiques extraites des données d'enregistrement EEG et d'autre part, sur la performance du classificateur utilisé.

Les travaux présentés dans cette partie portent sur l'utilisation du perceptron multicouche (MLP) à retro-propagation du gradient (vu au chapitre 4, section 4). Ils seront analysés et comparés à d'autres travaux dans la littérature. La spécification du réseau de neurone est déterminée par le nombre de couches cachées et le nombre de neurone choisi.

Le perceptron multicouche utilisé dans notre travail est formé d'une couche d'entrée, d'une couche cachée composée de 10 neurones et d'une couche de sortie représentant les deux classes.

Pendant la phase d'apprentissage, chaque neurone de sortie compare son activation calculée à la valeur désirée pour déterminer l'erreur associée et ajuste, par la suite, les poids synaptiques afin de minimiser l'erreur. La fonction d'activation utilisée est la fonction sigmoïde pour la couche cachée, et la fonction linéaire pour la couche de sortie.

Dans cette étude, la stabilité des performances du classificateur MLP est évaluée en fonction des différentes mesures statistiques (sensibilité, spécificité, précision de classification), en utilisant la validation croisée (*10-fold cross validation*).

Cette méthode a permis d'obtenir des valeurs optimales de 99,50%, 93% et 96,25% concernant, respectivement, la sensibilité, la spécificité, et la précision de classification des données d'enregistrement EEG du groupe (set A-E); pour le groupe (set A-(C-E)), les résultats obtenus sont 98,65%, 99,17% et 98,91%, respectivement, pour la sensibilité, la spécificité, et la précision.

Sur les tableaux 5.7 et 5.8, les résultats représentés donnent les pourcentages de la classification basée sur le réseau de neurone MLP et utilisant trois paramètres caractéristiques (la déviation standard, l'énergie et l'entropie).

Les figures 5.22 et 5.23 représentent le pourcentage de la précision de la classification (a), de la sensibilité (b) et de la spécificité (c).

Tableau 5.7 : Résultats de mesures obtenues à partir de la classification par réseau de neurone MLP pour le Set (A-(C-E)). (TVP: sensibilité ; TVN: spécificité).

Classification	Performances	Std	E	Ent
MLP	Précision (%)	98.91	90.75	95.75
	TVP (%)	98.65	91.45	96.54
	TVN (%)	99.17	90.05	94.96

Tableau 5.8 : Résultats de mesures obtenues à partir de la classification par réseau de neurone MLP pour le Set (A-E). (TVP: sensibilité ; TVN: spécificité).

Classification	Performances	Std	E	Ent
MLP	Précision (%)	96.25	86.15	92.65
	TVP (%)	99.5	90.75	90.45
	TVN (%)	93	81.55	94.85

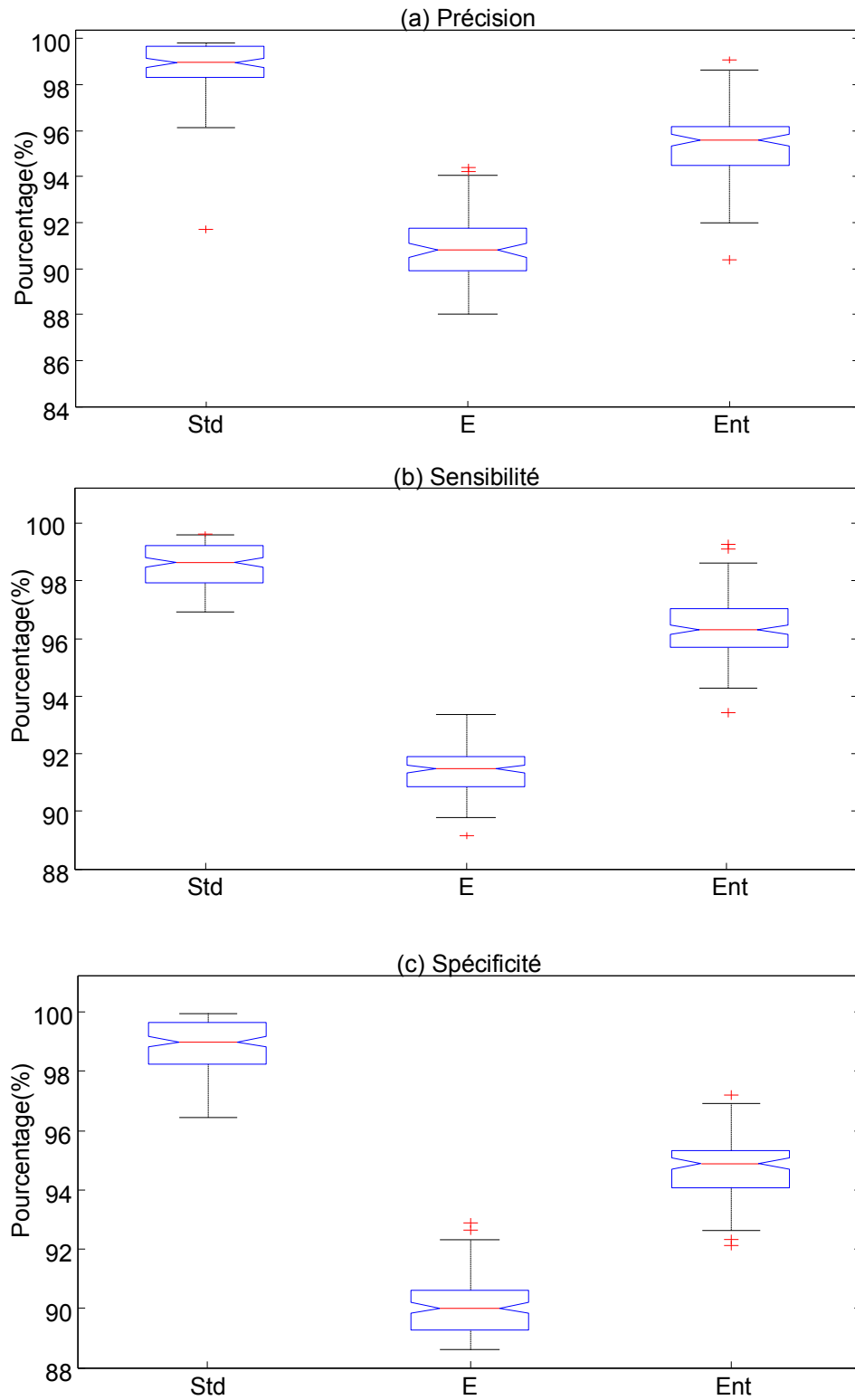


Figure 5.22 : Performances de classification ((a): précision, (b): sensibilité et (c): spécificité) du set (A-(C-E)) utilisant le réseau de neurone MLP.

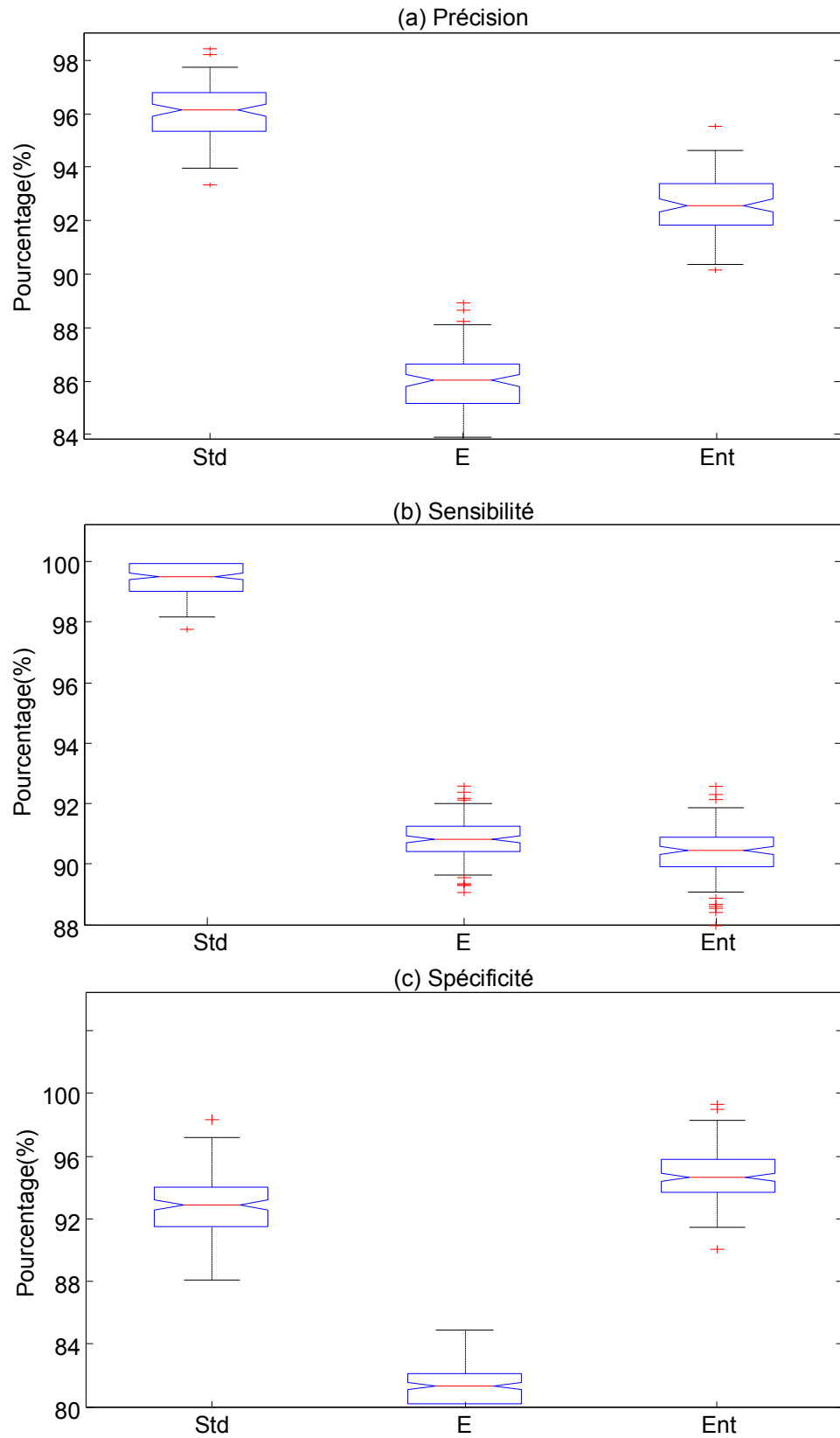


Figure 5.23 : Performances de classification ((a): précision, (b): sensibilité et (c): spécificité) du set (A-E) utilisant le réseau de neurone MLP.

Les performances des algorithmes utilisant le MLP pour la détection des EEG épileptiques et provenant de la littérature sont comparées dans le tableau 5.9. Nous comparons les résultats obtenus avec d'autres méthodes existantes, uniquement, avec des travaux qui utilisent les mêmes groupes de la base de données EEG.

Tableau 5.9 : Valeurs des précisions obtenues dans la littérature et utilisant les mêmes groupes de données.

Auteurs	Méthodes utilisées	EEG	Précision (%)
[70] Guler <i>et al.</i> (2007)	DWT + Lyapunov exponent+ MLP classification	(A-B-C-D-E)	93.63
[103] Naghsh-Nilchi <i>et al.</i> (2010)	EigenMethod + MUSIC+ MLP	(A-C-E)	97.50
[69] Ubeyli <i>et al.</i> (2010)	Lyapunov exponents + PNN	(A-B-C-D-E)	98.50
[63] Guo <i>et al.</i> (2011)	Multi-wavelet + Approximate entropy + MLP	(ABCD-E)	98.20
[65] Orhan <i>et al.</i> (2011)	Wavelet transform+ K-means clustering+ MLP	(AB-CD-E)	95.60
la méthode proposée	DWT+ STD + PCA + MLP	(A-CE)	98.91

5.5 Comparaison des algorithmes de classification utilisés

Les résultats obtenus sur la précision de classification avec la méthode de classification par réseau de neurone MLP sont moins élevés que ceux obtenus avec le classificateur SVM à noyau RBF. Les tableaux 5.7 et 5.8 montrent des

taux de précision de 98,91% pour le groupe 1 et de 96,25% pour le groupe 2. Comme nous l'avons montré sur le tableau 5.5 et concernant le groupe 1, les résultats expérimentaux obtenus montrent un taux d'erreur de 0,93% pour une classification par SVM à base de noyau RBF, par rapport à un taux d'erreur de 1,09% (voir tableau 5.7) pour une classification par réseau de neurone MLP et ce taux d'erreur est de 1,7% pour le classificateur K-NN. Ceci est aussi valable pour le groupe 2, où le taux d'erreur obtenu est de 0,96% pour une classification par SVM, d'un taux d'erreur de 3,96% pour le MLP et de 1,5% pour une classification par K-NN.

L'étape d'extraction des caractéristiques, statistiques ou non, est très importante, car les performances de la classification sont dégradées si les caractéristiques ne sont pas bien choisies. Les mêmes paramètres caractéristiques, les plus représentatives, ont été pris en compte pour constituer l'entrée des classificateurs. Cette approche a été effectuée sur les mêmes groupes d'enregistrement EEG issus d'une base de données réelle. Les résultats montrent que l'algorithme basé sur les SVM à noyau RBF donne les meilleurs résultats comparativement à l'algorithme des réseaux de neurone MLP et à celui des K-NN.

Dans [58], Subasi *et al.* (2010) présentent, dans leurs travaux, une méthodologie basée sur l'analyse, en temps-fréquence, par la transformée en ondelette discrète. Quatre caractéristiques, i.e. {Moyenne, Puissance, Déviation standard, Rapport des moyennes}, sont calculées pour chaque coefficient de décomposition par la transformée en ondelette des deux sets (A et E) de la base de données [71].

Pour évaluer leur méthodologie, les auteurs utilisent trois méthodes de réduction (PCA, LDA et ICA) appliquées aux coefficients d'ondelettes; les vecteurs caractéristiques sont ensuite introduits comme entrée du classificateur SVM. Ils obtiennent des résultats, en termes de précision globale, variant de 98,75% à 100%.

Wang *et al.* (2011) [93], introduisent dans leur article la décomposition en paquet d'ondelette sur les données EEG, set A et set E. Cette étude utilise l'entropie pour l'extraction des caractéristiques de la meilleure base dans l'étape de l'algorithme d'apprentissage pour la classification des signaux EEG par la méthode des K -plus proches voisins. La précision maximale de l'algorithme de classification obtenue est de 99,44%.

Dans des travaux plus récents [104], Nunes *et al.* (2014) proposent un algorithme d'aide au diagnostic de l'épilepsie permettant l'évaluation systématique des

performances du classificateur *Optimum Path Forest* (OPF) appliqué aux cinq sets (A-B-C-D-E), normalisés ou non normalisés, de la même base de données utilisée dans nos travaux. L'extraction et la sélection des paramètres caractéristiques {Minimum, Maximum, Moyenne, Puissance, Déviation standard, Rapport des moyennes absolues} sont basées sur les coefficients de prétraitement à l'aide de quatre familles d'ondelette (Haar, Daubéchies, Coiflets et Symlet). Les auteurs utilisent différentes méthodes de réduction, entre autre l'algorithme génétique, afin d'éliminer les fonctions redondantes et par la suite, comparent les performances du classificateur OPF à celles des classificateurs supervisés SVM à base de noyau RBF, le MLP et le bayésien. Les scores de performance du classificateur OPF obtenus ne dépassent pas la moyenne de précision de 89,2% ; par contre, la moyenne du temps d'exécution (apprentissage/test) est beaucoup plus faible par rapport à celle requise par le SVM et le MLP.

Conclusion et Perspectives

Ce travail de thèse s'inscrit dans le cadre de développement des systèmes médicaux d'aide au diagnostic. Que ce soit dans le diagnostic ou dans la recherche en sciences neurocognitives, l'EEG constitue une indispensable source d'informations. L'EEG mesure l'activité du cerveau, permettant de déceler des pathologies comme l'épilepsie. La classification de ces enregistrements est d'une grande importance dans la recherche biomédicale pour l'identification et l'évaluation de l'activité du cerveau. Cet examen produit, néanmoins, de longs enregistrements sur de nombreux canaux avec une redondance élevée; la lecture et l'interprétation de ces signaux pouvant dépasser plusieurs jours, voire plusieurs semaines, d'enregistrement est une tâche difficile et fastidieuse pour les neurologues. Cette difficulté a motivé le développement de méthodes automatiques permettant d'analyser et de classer les événements épileptiques.

L'objectif de ce travail de thèse est d'étudier et de développer les techniques de traitement et de classification des enregistrements de signaux électro-encéphalographiques incluant une pathologie très répandue qui est l'épilepsie, dans le but d'identifier les différents types de signaux EEG.

Pour cela, nous avons traité et analysé des enregistrements EEG issus d'une base

de données réelle comprenant cinq cent canaux EEG. Une première étape consistait à utiliser l'analyse multi-résolution en appliquant la transformée en ondelette discrète qui est une représentation de la distribution d'énergie en temps-fréquence, très utilisée pour l'analyse des signaux biomédicaux, et la transformée par paquet d'ondelette permettant une analyse en ondelette plus fine des signaux EEG. L'élaboration d'un algorithme de sélection de l'ondelette mère a été mise en œuvre en utilisant cinq familles d'ondelette (*Daubéchies*, *Coiflet*, *Symlet*, *Biorthogonal*, *Discret Meyer*), car on sait que le choix de l'ondelette est une étape cruciale dans l'analyse des signaux EEG.

La deuxième étape concernait l'extraction des six paramètres caractéristiques statistiques et non statistiques. L'examen EEG, produit sur trois cent canaux d'enregistrement, présente une redondance, et par conséquent, risque de réduire les performances des classificateurs. Pour remédier à cela, nous avons appliqué deux méthodes bien connues de réduction de dimension, i.e. l'analyse par composante principale et l'analyse discriminante linéaire.

Après développement de ces méthodes de traitement de signaux EEG et pour atteindre notre objectif, nous avons utilisé trois techniques de classification : machine à vecteur support, les K -plus proches voisins et le perceptron multicouche. Ces techniques de classification ont été appliquées à des signaux EEG issus de cent personnes saines, de cent patients souffrant de l'épilepsie en pleine crise et de cent signaux d'autres patients épileptiques enregistrés entre les crises. Nous avons également utilisé la méthode de '*k-fold* validation croisée' avec différentes valeurs de k pour évaluer les performances des classificateurs, en se basant sur le calcul des mesures statistiques de la précision, de la sensibilité et de la spécificité de classification.

La classification par SVM étudiée tout au long de ce travail offre des performances supérieures à celles des deux autres algorithmes auxquels elle a été comparée. En tirant profit de ces apports d'outils issus du traitement de signal, nous avons obtenu des taux de classification atteignant jusqu'à 100% sur la base de données d'EEG épileptiques en pleine activité. Nous nous sommes limités, pour cette étude, à une seule base de données réelle de l'Université de Bonn; néanmoins, afin de comparer les performances d'évaluation et vérifier l'efficacité de l'algorithme proposé, la méthode proposée pourrait s'appliquer, d'une part, sur différents signaux biomédicaux conçus sur une base de données issue d'un service de neurologie hospitalier d'Alger.

D'autre part, souvent les signaux d'enregistrement EEG sont fortement contaminés par divers artefacts provenant à la fois du sujet et/ou des interférences des équipements. Dans notre thèse, nous n'avons pas développé de méthodes de suppression d'artefacts à partir des signaux de la base de données EEG. Une autre étude approfondie pourrait-être envisagée afin d'éliminer ces interférences sans contaminer les signaux EEG.

Notre travail pourrait s'appliquer, aussi, à d'autres pathologies comme la maladie d'Alzheimer en utilisant différents algorithmes de classification, et pour la sélection du meilleur classificateur, appliquer l'algorithme génétique comme outil heuristique. Du point de vue instrumentation électronique, les résultats obtenus sont un prélude à des travaux futurs en expérimentant une architecture embarquée sur carte FPGA afin d'utiliser un outil de traitement du signal permettant de faire une prédiction, en temps réel, de la crise épileptique.

Bibliographie

- [1] Clarac F., Ternaux J.P. *Encyclopédie historique des neurosciences, du neurone à l'émergence de la pensée*. Editions De Boeck, Paris, France. 2008.
- [2] Kolb, B. and Whishaw I. *Cerveau et comportements*. Editions De Boeck, Paris, France. 2002.
- [3] Budzynski T. H., Kogan Budzynski H. , Evans J. R. and Abarbanel A. *Introduction to quantitative EEG and neurofeedback*. Second Edition: Advanced Theory and Applications. 2008.
- [4] Costes Nicolas. *Neurotransmission sérotoninergique 5-HT1A : approche méthodologique de la mesure in vivo par le [18F] MPPF en tomographie par émission de positons*. Thèse de doctorat. Université Claude Bernard - Lyon 1. 2007.
- [5] World Health Organisation (WHO) statistics on epilepsy, International League against Epilepsy. *Atlas: Epilepsy care in the world*. WHO Press. 2011.
- [6] Adam C., Arzimanoglou A., Baulac M., Bendriem B., Biraben A., Chiron C., Dupont S., Hirsch E., Marescaux C., Semah F., Spelle L. *Epilepsies : de la clinique aux examens complémentaires (Tome 1)*. Expressions Santé Editions. 2000.
- [7] Bernard M. *Méthodologie d'analyse des synchronisations neuronales dans les signaux EEG à l'aide de graphes d'informations temps-fréquence*. Thèse de doctorat Université de Poitiers. 2006.
- [8] Oussar Y. *Réseaux d'ondelettes et réseaux de neurones pour la modélisation statique et dynamique de processus*. Thèse de doctorat de l'université Pierre et Marie Curie. Paris, France. Juillet 1998.

- [9] Williams, R. W. and Herrup, K. The control of neuron number. *Annual Review of Neuroscience*, 11:423–53. 1988.
- [10] Quotb A. *Méthodes et systèmes pour la détection adaptative et temps réel d'activité dans les signaux biologiques*. Thèse de doctorat de l'université de Bordeaux I. Octobre 2012.
- [11] Hocepied Gatien. *Détection précoce de crises d'épilepsie à l'aide d'une modélisation du comportement oscillatoire neuronal*. Thèse de doctorat, école polytechnique de Bruxelles. 2012.
- [12] Vialatte F. B. *Modélisation en bosses pour l'analyse de motifs oscillatoires reproductibles dans l'activité de populations neuronales: applications à l'apprentissage olfactif chez l'animal et à la détection précoce de la maladie d'Alzheimer*. Thèse de doctorat, Université Paris 6 - Pierre et Marie Curie – Neurosciences. 2006.
- [13] Daunizeau J. *Localisation et dynamique des sources d'activité cérébrale par fusion d'informations multimodales EEG/IRMf*. Thèse de doctorat, Université de Montréal, Canada. 2005.
- [14] Miltner W., Braun C., Johnson R., Simpson G.V., Ruchkin D.S. A test of brain electrical source analysis (BESA): A simulation study. *Electroencephalography and Clinical Neurophysiology* 91:295-310. 1991.
- [15] Jasper, H. The ten-twenty electrode placement system of the International Federation. *Electroencephalography and Clinical Neurophysiology*. 10:371-375. 1958.
- [16] Mallat Stéphane. *A Wavelet Tour of Signal Processing*. 3rd Ed. Academic Press; Burlington, USA. 2009.
- [17] Burrus C.S., Ramesh A. Gopinath and Haitao Guo. *Introduction to wavelets and wavelet transform: A primer*. Prentice Hall. 1997.
- [18] Misiti M., Misiti Y., Oppenheim G. et Poggi J.M. *Les ondelettes et leurs applications*. Paris : Lavoisier, Hermes Sciences. 2003.
- [19] Meyer Y. *Ondelettes et opérateurs*. Hermann Editions. 1990.
- [20] Daubechies I. *Ten lectures on wavelets*. Society for industrial and applied mathematics. 2006.
- [21] Coifman R. R., Wirckerhauser M.V. Entropy-based algorithms for best basis

- selection. *IEEE Trans. on Inf. Theory*, vol.38, n°2, p. 713-718. 1995.
- [22] Gilbert Saporta. *Probabilités, analyse des données et statistique*. Editions TECHNIP. 2006.
- [23] Lebart Ludovic, Morineau Alain et Piron Marie. *Statistique exploratoire multidimensionnelle*. Dunod, Paris. 1995.
- [24] Hervé Abdi and Lynne J. Williams. *Principal component analysis*. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2010.
- [25] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley Interscience, New York. 2000.
- [26] Fuchs S. *Conception et réalisation de méthodes de détection de polluants gazeux atmosphériques à l'aide d'un nez électronique portable*. Thèse de doctorat de l'Université de Metz. 2008.
- [27] Vapnik V. *The nature of statistical learning theory*, Springer-Verlag, New York, 1995.
- [28] Burges CJC, *A tutorial on support vector machines for pattern recognition*, Kluwer Academic Publishers, Boston, 1998.
- [29] Bottou L. and Chih-jen L. *Support Vector Machine Solvers in Large Scale Kernel Machines*. MIT Press. 2007.
- [30] Rémi Flamary. *Apprentissage statistique pour le signal*. Thèse de Doctorat en Informatique de l'Université de Rouen, Laboratoire d'Informatique, de Traitement de l'Information et des Systèmes. 2012.
- [31] Cornuéjols A., Miclet L. et Kodratoff Y. *Apprentissage artificiel: concepts et algorithmes*. Ed. Eyrolles. 2002.
- [32] Mormann F., Andrzejak R. G., Elger C. E. and Lehnertz K. Seizure prediction: the long and winding road. *Brain*, 130, 314-333. 2007.
- [33] Tzallas A. T., Tsipouras M. G. and Fotiadis D. I. Automatic seizure detection based on time-frequency analysis and artificial neural networks. *Computational Intelligence and Neuroscience*. 2007.
- [34] Tzallas A. T., Tsipouras M. G. and Fotiadis D. I. Epileptic seizure detection in EEGs using time-frequency analysis. *IEEE Transaction on Information Technology in Biomedicine*, 13(5), 703-710. 2009.

- [35] McGrogan N. *Neural Network Detection of Epileptic Seizures in the Electroencephalogram*. Oxford University, Oxford. 2001.
- [36] Casson A., Yates D., Smith S., Duncan J. and Rodriguez-Villegas, E. Wearable Electroencephalography. What is it, why is it needed, and what does it entail? *IEEE Engineering in Medicine and Biology Magazine*, 29(3), 44-56. 2010.
- [37] Waterhouse, E. New horizons in ambulatory electroencephalography. *IEEE Engineering in Medicine and Biology Magazine*, 22(3), 74-80. 2003.
- [38] Iriarte J., Urrestarazu E., Valencia M., Alegre M., Malanda A., Viteri C., and Artieda J. Independent component analysis as a tool to eliminate artifacts in EEG: A quantitative study. *Journal of Clinical Neurophysiology*, 20(4):249–257. 2003.
- [39] Bousbia-Salah A., Belouchrani A. and Cichocki A. Application of time-frequency distributions to the independent component analysis of ECG signals. *Conference proceeding of IEEE- International Symposium on Signal Processing and its Applications*. August 2001.
- [40] Hesse, C. W. and James, C. J. Tracking and detection of epileptiform activity in multi-channel ictal EEG using signal subspace correlation of seizure source scalp topographies. *Medical & Biological Engineering & Computing*, 45(10), 909-916. 2007.
- [41] Rosso O.A., Blanco S. and Rabinowicz A. Wavelet analysis of generalized tonic-clonic epileptic seizures. *Signal Processing*, 83:1275–1289. 2003.
- [42] Saab M.E. and Gotman J. A system to detect the onset of epileptic seizures in scalp EEG. *Clinical Neurophysiology*, 116(2):427–442. 1997.
- [43] Shoeb A. *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*. PhD Thesis, Massachusetts Institute of Technology, September 2009.
- [44] Mc Sharry P. E., He T., Smith L. A. and Tarassenko L. Linear and non-linear methods for automatic seizure detection in scalp electro-encephalogram recordings. *Medical & Biological Engineering & Computing*, 40(4), 447-461. 2002.
- [45] Liu H. S., Zhang T. and Yang F.S. A multistage, multi-method approach for automatic detection and classification of epileptiform EEG. *IEEE Transaction on Biomedical Engineering*, 49, 1557-1566. 2002.

- [46] Altenburg J., Vermeulen R. J., Strijers R. L., Fetter W. P. and Stam C. J. Seizure detection in the neonatal EEG with synchronization likelihood. *Clinical Neurophysiology*, 114(1), 50-55. 2003.
- [47] Van Putten M. J. Nearest neighbor phase synchronization as a measure to detect seizure activity from scalp EEG recordings. *Journal of Clinical Neurophysiology*, 20(5), 320-325. 2003.
- [48] Gotman J. Non-invasive methods for evaluating the localization and propagation of epileptic activity. *Epilepsia*, 44, 21-29. 2003.
- [49] Polat K. and Gunes S. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. *Applied Mathematics and Computation*, 187(2), 1017–1026. 2007.
- [50] Polat K. and Gunes S. Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals. *Expert Systems with Applications*, 34(3), 2039-2048. 2008.
- [51] Polat K. and Gunes S. A novel data reduction method: Distance based data reduction and its application to classification of epileptiform EEG signals. *Applied Mathematics and Computation*, 200(1), 10-27. 2008.
- [52] Srinivasan V., Eswaran C. and Sriraam N. Artificial neural network based epileptic detection using time-domain and frequency-domain features. *Journal of Medical Systems*, 29(6), 647-660. 2005.
- [53] Alkan A., Koklukaya E. and Subasi A. Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods*, 148(2), 167-176. 2005.
- [54] Ubeyli E.D. Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals. *Expert Systems with Applications*, 37(1), 233–239. 2010.
- [55] Chua K. C., Chandran V., Acharya R. and Lim C. M. Automatic identification of epilepsy by HOS and power spectrum parameters using EEG signals: a comparative study. *Conference Proceeding of IEEE Engineering in Medicine and Biology Society*. 3824-3827. 2008.
- [56] Iscan Z., Dokur Z. and Tamer D. Classification of electroencephalogram signals with combined time and frequency features. *Expert Systems with Applications*, 38, 10499–10505. 2011.

- [57] Martinez-Vargas J. D., Avendano-Valencia L. D., Giraldo E. and Castellanos-Dominguez G. Comparative analysis of Time Frequency Representations for discrimination of epileptic activity in EEG Signals. *Proceeding of IEEE-EMBS Conference on Neural Engineering*. 2011.
- [58] Subasi A. and Gursoy I. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications* 37, 8659–8666. 2010.
- [59] Subasi A. Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction. *Computers in Biology and Medicine*, 37(2), 227-244. 2007.
- [60] Subasi A. Signal Classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 32(4), 1084-1093. 2007.
- [61] Subasi A., Alkan A., Koklukaya E. and Kiymik M. K. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. *Neural Networks*, 18(7), 985-997. 2005.
- [62] Adeli H., Ghosh-Dastidar S. and Dadmehr N. A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy. *IEEE Transaction of Biomedical Engineering*, 54(2), 205-211. 2007.
- [63] Guo L., Rivero D. and Pazos A. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. *Journal of Neuroscience Methods*, 193(1), 156-163. 2010.
- [64] Ocak H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications*. 36(2), 2027– 2036. 2009.
- [65] Orhan U., Hekim M. and Ozer M. EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. *Expert Systems with Applications* (38), 13475–13481. 2011.
- [66] Lima C. A., Coelho A. L. and Eisencraft M. Tackling EEG signal classification with least squares support vector machines: a sensitivity analysis study. *Computers in Biology and Medicine*, 40(8), 705-714. 2010.
- [67] Chandaka S., Chatterjee A. and Munshi S. Cross-correlation aided support vector machine classifier for classification of EEG signals. *Expert Systems with Applications* 36(2), 1329-1336. 2009.

- [68] Kannathal N., Acharya U. R., Lim C. M. and Sadasivan P. K. Characterization of EEG: a comparative study. *Computer Methods and Programs in Biomedicine*, 80(1), 17-23. 2005.
- [69] Ubeyli E. D. Lyapunov exponents/probabilistic neural networks for analysis of EEG signals. *Expert Systems with Applications* 37, 985–992. 2010.
- [70] Guler I. and Ubeyli E. D. Multiclass support vector machines for EEG signals classification. *IEEE Transaction on Information Technology in Biomedicine*, 11(2), 117-126. 2007.
- [71] <<http://www.meb.uni-bonn.de/epileptologie/science/physik/eegdataold.html>> Accessed October 2012.
- [72] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David & C.E. Elger, Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E* 64:061907, 2001.
- [73] A.Bousbia-Salah, M.A.Ait-Ameur and M.Talha-Kedir. Biorthogonal Wavelet for EEG Signal Compression ». ACM-proceeding of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies, October 26-29, 2011, ISABEL'11, Barcelona, Spain.
- [74] Subasi A. Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines. *Computers in Biology and Medicine*. 42, 806–815. 2012.
- [75] Donna Giri, U. Rajendra Acharya, Roshan Joy Martis, S. Vinitha Sree, Teik-Cheng Lim, Thajudin Ahamed and Jasjit S. Suri. Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA and Discrete Wavelet Transform. *Knowledge-Based Systems* 37, 274–282. 2013.
- [76] A. Bousbia-Salah and M. Talha-Kedir. Time-Frequency Processing Method of Epileptic EEG Signal. *Biomedical Engineering: Applications, Basis and Communications*, 27 (2), 1550015(1-8). 2015. DOI:10.4015/S1016237215500155.
- [77] Abdulkadir S. Multiclass least-square support vector machines for analog modulation classification. *Expert System with Applications*, Vol. 36, pp. 6681–6685. 2009.

- [78] Hsu C-W., Chang C.C. and Lin C.J. A practical guide to support vector classification. Rapport technique, National Taiwan University. 2010.
- [79] Devyver P. A. and Kittler J. *Pattern recognition: A Statistical Approach*. London: Prentice-Hall. 1982.
- [80] Lima C.A.M. and Coelho A.L.V. Kernel machines for epilepsy diagnosis via EEG signal classification: A comparative study. *Artificial Intelligence in Medicine*, 53, pp 83–95. 2011.
- [81] Lei Guo, Youxi Wu, Lei Zhao, Ting Cao, Weili Yan and Xueqin Shen. Classification of Mental Task from EEG Signals Using Immune Feature Weighted Support Vector Machines. *IEEE Transactions on Magnetics*, 47(5), pp 866-869. 2011.
- [82] Alkan A., Koklukaya E. and Subasi A. Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods*, 148(2), pp. 167-76, 2005.
- [83] Guler I. Ubeyli E.D. Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. *Journal of Neuroscience Methods*, 148(2), pp. 113-21. 2005.
- [84] Aslan K., Bozdemir H., Sahin C., Ogulata S.N. and Erol R. A Radial Basis Function Neural Network Model for Classification of Epilepsy Using EEG Signals. *Journal of Medical Systems*, 32(5). 2008.
- [85] Petrosian A., Prokhorov D., Homan R., Dasheiff R. and Wunsch D. Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG. *Neurocomputing*, vol. 30, pp. 201-218. 2000.
- [86] Lefèvre T. *Formes et fonctions, en théories et en pratiques : l'intérêt d'une approche morpho dynamique fondée sur l'information et la géométrie*. Thèse de doctorat de l'Université de Bretagne Occidentale. Mai 2014
- [87] Siuly Y. Li. A novel statistical algorithm for multiclass EEG signal classification. *Engineering Applications of Artificial Intelligence*, 34, pp154–167. 2014.
- [88] Suykens J.A.K., Gestel T.V., Brabanter J.D., Moor B.D., Vandewalle J. Least Square Support Vector Machine. WorldScientific. Singapore. 2002.
- [89] Joshi V., Pachori R.B., Vijesh A. Classification of ictal and seizure-free EEG signals using fractional linear prediction. *Biomedical Signal Processing and Control*. 9, pp1– 5. 2014.

- [90] Khandoker A.H., Lai D.T.H., Begg R.K., Palaniswami M., Wavelet-based for support vector machines for screening balance impairments in the elderly. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 15(4), pp 587–597. 2007.
- [91] Gandhi T., Panigrahi B.K., Anand S. A comparative study of wavelet families for EEG signal classification. *Neurocomputing*. 74, 3051–3057. 2011.
- [92] Adeli H. Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of Neuroscience Methods*. 123, 69-87. 2003.
- [93] Wang D., Miao D., Xie C. Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection. *Expert Systems with Applications*. 38(11). 2011.
- [94] Hadj-youcef M.A., Adnane M. and Bousbia-Salah A. Detection of Epileptics during seizure free periods. *IEEE- WOSSPA2013, the 9th International Workshop on System, Signal Processing and their applications*. 12-15 May 2013, Algiers.
- [95] Adeli H., Zhou Z. and Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. *Journal of Neuroscience Method*. 123(1). 2003.
- [96] Güler NF., Übeyli E.D. and Güler I. Recurrent neural networks employing Lyapunov exponents for EEG signal classification. *Expert System with Applications*. (29)3. 2005.
- [97] Übeyli E.D. Statistics over features: EEG signals analysis. *Computers in Biology and Medicine*. 39(8). 2009.
- [98] Acharya U.R., VinithaSree S., Alvin A.P.C and Suri J.S. Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. *Expert System with Applications*. 39(10). 2012.
- [99] Hastie T., Tibshirani R. and Friedman J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer Series in Statistics. Second edition. Springer-Verlag. 2009.
- [100] Victor Marcos J., Hornero R., Alvarez D. , Aboy M. and Felix Del Campo. Automated Prediction of the Apnea-Hypopnea Index from Nocturnal Oximetry Recordings. *IEEE Transactions On Biomedical Engineering*, 59(1). 2012.

- [101] Chendeb M. *Détection et classification des signaux non stationnaires par utilisation des ondelettes. Application aux signaux électromyographiques utérins*. Thèse de doctorat de l'Université de technologies de Troyes. Mars 2006.
- [102] Mercier L. *Les machines à vecteurs support pour la classification en imagerie hyper spectrale : implémentation et mise en œuvre*. Février 2010.
- [103] Naghsh-Nilchi A.R., Aghashahi M. Epilepsy seizure detection using eigen-system spectral estimation and Multiple Layer Perceptron neural network. *Biomedical Signal Processing and Control*. 5, pp147– 157. 2010.
- [104] Nunes T.M., Coelho A. L., Lima C.A.M., Papa J.P., de Albuquerque V.C. EEG signal classification for epilepsy diagnosis via optimum path forest – A systematic assessment. *Neurocomputing*. 136, pp103–123. 2014.
- [105] A.Bousbia-Salah, M.A.Ait-Ameur and M.Talha-Kedir « Compression de Signaux Biomédicaux EEG par la Transformée en ondelette ». Instrumentation Laboratory Workshop JLINS'2010, 25-26 October 2010, USTHB, Algiers.
- [106] A. Bousbia-Salah and M.Talha-Kedir « Automated diagnosis of Epileptic EEG signals ». Instrumentation Laboratory Workshop JLINS'2012, 30-31 October 2012, USTHB, Algiers.

الملخص

هذا العمل يتعلق بتحليل وتصنيف إشارات تسجيل EEG. تصميم نظام فعال للتصنيف معقد نظرا لتقلب وعدم استقرار إشارات EEG لمختلف الحالات المنتمية لأقسام مختلفة. لقد قمنا بتطوير برامج وخوارزميات لاستخراج المعلومات من إشارات EEG من قاعدة بيانات حقيقية لتحديد فئات مختلفة من البيانات (EEG من غير الصرع، EEG مع الصرع و EEG بين الصرع).

أخيرا وبعد تحديد الخصائص الإحصائية وغير الإحصائية، قمنا أداء المقاييس للتحقق من صحة أفضل معدل لتصنيف إشارات EEG.

Résumé

Ce travail concerne l'analyse et la classification de signaux d'enregistrement électro-encéphalographiques. La conception d'un système efficace de classification est complexe, étant donné la variabilité et la non-stationnarité des signaux d'encéphalogramme (EEG) de différents sujets appartenant à différentes classes.

Nous avons développé différents algorithmes permettant d'extraire les informations à partir des signaux EEG issus d'une base de données réelle afin d'identifier les différentes catégories de données (EEG non épileptique, EEG avec crises épileptiques et EEG entre deux crises épileptiques).

Enfin, après identification des caractéristiques statistiques et non statistiques des EEG, nous avons évalué les performances de mesures pour les valider par le meilleur taux de classification des signaux d'enregistrement EEG.

Abstract

This work concerns the analysis and classification of electroencephalographic recording signals. The design of an effective classification system is complex, given the variability and non-stationary of EEG signals of different subjects belonging to different classes.

We have developed algorithms for extracting information from the EEG signals from a real database to identify different categories of data (normal EEG, critical EEG and inter-critical EEG).

Finally, after identification of the statistical and non-statistical characteristics of this signal, we evaluated the performance metrics to validate the best EEG signals classification rate.