

République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la Recherche Scientifique
Université de Mohamed El Bachir El Ibrahimi de Bordj Bou Arréridj
Faculté des Mathématiques et d'Informatique
Département d'informatique



MEMOIRE

Présenté en vue de l'obtention du diplôme

Master en informatique

Spécialité : Réseau et multimédia

THEME

Classification des IRM cérébrales pathologiques basée sur
les techniques d'apprentissage profond

Présenté par :

Boussoualim Wiam

Idir Sonia

Soutenu publiquement le : 22 / 06 / 2023

Devant le jury composé de:

Président : Ms Attia Abdelouaheb

Examineur : Mme Fares Nour El Houda

Encadreur : Mme Chellakh Hafida

2022/2023

Remerciement :

En préambule à ce mémoire nous remercions **ALLAH** qui nous a aidés et nous a donné la patience et le courage durant ces longues années d'études d'accomplir ce modeste travail.

Nous tiendrons d'abord à remercier très chaleureusement :

- **Mme Chellakh Hafida** qui nous a permis de bénéficier de son
- encadrement , ses conseils et son orientation .
- **Mr Attia Abdelouahab** pour ses conseils
- **Les membres du jury** qui nous feront l'honneur d'estimer et juger ce travail.

Ensuite, Nous adressons nos remerciements les plus sincères à nos **très chers PARENTS** qui ont apporté leurs aides durant ces longues années .

Merci à **tous** ceux qui nous ont **aidé**, de près ou de loin, de contribuer à la réalisation de ce travail .

Dédicace :

Je tiens à dédier ce modeste travail à tous ceux qui m'ont encouragé durant toute la période de réalisation de ce travail.

En particulièrement :

- A mes chers parents qui se sacrifient pour me voir réussir
- A mes chères sœurs **Asma, Rayane, Maria** et mon cher frère **Yassine** pour leurs soutiens.
- A toutes mes familles **Boussoualim** et **Helilou** ;
- A mes chers amis, merci à tous et à toutes avec qui je partage

des moments de ma vie au fil du temps ; **Hind , Boutheyne , Sonia , khalissa , dounia**

- Mes collègues : **Idir Sonia , Aggoune Manel** ;
- Toute la promotion d'informatique réseaux et multimédia 2022/2023 ;
- Tous mes enseignants pour le savoir, la connaissance et surtout la générosité .

 **Boussoualim
Wiam**

Dédicace :

Premièrement et avant tout « الحمد لله على تمام النعمة »

Je dédie ce travail à :

- A ma source de vie, à mon support qui toujours été à mes coté pour me soutenir et m'encourager, à **mon cher papa**.
- A mon paradis, à la prunelle de mes yeux, à ma source d'amour, de joie et de bonheur, à **ma moitié maman**.
- A mes très chers frères **Mohamed et Djibrail**.
- A mes trois chers **grands parents** .
- A **mon grand-père** maternelle pais à son âmes, que dieu le garde dans son vaste Paradie.

J'espère sincèrement que vous êtes fiers de moi.

- A toute ma famille paternelle et maternelle (**IDIR , DRAHMOUNE**).
- A mes cousines **Nessrine, Houda, Karima et Isra**.
- A mes chères amies **Wiam, Manel, Fatima, Lina**.
- A toutes mes copines.
- A tous mes camarades et mes collègues qui ont été mes compagnons de route durant ces années d'étude.
- Au personnes qui sont absentes de mes yeux et présentes dans mon cœur.
- A toutes les personnes que j'aime et qui m'aiment et qui ont été à mes cotées et m'ont soutenue .

 **Idir Sonia**

Résumé :

Le but de cette étude était de proposer un nouveau modèle de classification pour l'IRM cérébrale pathologique. Pour cette raison, il est essentiel de concevoir et de développer un système d'aide au diagnostic pour aider les médecins à analyser ces images IRM. Avec notre contribution, nous développons une procédure de classification d'images cérébrales saines et pathologiques. L'image est prétraitée, redimensionnée, puis l'étape d'extraction des caractéristiques est terminée à l'aide de deux extracteurs de caractéristiques de bas niveau : Hog, Gist et un autre ResNet-50 de haut niveau. Un mécanisme d'optimisation est ensuite appliqué aux vecteurs de caractéristiques produits par l'étape d'extraction afin de : minimiser la taille de ce vecteur, utilisez l'algorithme d'optimisation Particle Swart Optimization (PSO) pour sélectionner les meilleures caractéristiques et améliorer la précision. Les fonctionnalités optimisées deviennent l'entrée du classifieur sélectionné pour notre contribution "Deep Rule Based" DRB. Pour évaluer ce classificateur, des expériences ont été menées avec deux autres algorithmes d'optimisation (GW et GA) et deux autres classificateurs (KNN et SVM). Nous obtenons des résultats montrant que DRB a des performances élevées avec une précision de 99,67 % , et on a aussi utilisé la technique de changement d'angle pour augmenter le taux d'accuracy.

Mots clés :Imagerie par résonance; tumeur cérébrale ; extraction de caractéristiques ; Deep Rule Based ; classification.

Abstract :

The aim of this study was to propose a new classification model for pathological brain MRI. For this reason, it is essential to design and develop a diagnostic support system to help doctors analyze these MRI images. With our contribution, we are developing a procedure for classifying healthy and pathological brain images. The image is pre-processed, resized, and then the feature extraction step is completed using two low-level feature extractors: Hog, Gist, and another high-level ResNet-50. An optimization mechanism is then applied to the feature vectors produced by the extraction step in order to: minimize the size of this vector, use the Particle Swarm Optimization (PSO) optimization algorithm to select the best features and improve the precision. The optimized features become the input of the selected classifier for our "Deep Rule Based" DRB contribution. To evaluate this classifier, experiments were conducted with two other optimization algorithms (GW and GA) and two other classifiers (KNN and SVM). We obtain results showing that DRB has high performance with an accuracy of 99.67%, and we also use the angle change technique to increase the accuracy rate.

Keywords: Magnetic resonance imagerie ; brain tumor ; feature extraction; Deep-Rule based; classification.

الملخص :

كان الهدف من هذه الدراسة هو اقتراح نموذج تصنيف جديد للتصوير بالرنين المغناطيسي للدماغ المرضى. لهذا السبب ، من الضروري تصميم وتطوير نظام دعم تشخيصي لمساعدة الأطباء على تحليل صور التصوير بالرنين المغناطيسي هذه. من خلال مساهمتنا ، نقوم بتطوير إجراء لتصنيف صور الدماغ الصحية والمرضية . من خلال مساهمتنا ، نقوم بتطوير إجراء لتصنيف صور الدماغ الصحية والمرضية. تتم معالجة الصورة مسبقا وتغيير حجمها ثم تكتمل خطوة استخراج المعالم باستخدام اثنين من مستخلصات الميزات منخفضة المستوى : Gist و Hog آخر عالي المستوى ResNet-50. ثم يتم تطبيق آلية التحسين على متجهات الميزات التي تنتجها خطوة الاستخراج من أجل: تقليل حجم هذا المتجه ، واستخدام (PSO) لتحديد أفضل الميزات وتحسين الدقة. تصبح الميزات المحسنة مدخلات خوارزمية تحسين سرب الجسيمات DRB المصنف المحدد لمساهمة "القائمة على القواعد العميقة".

لتقييم هذا المصنف ، أجريت تجارب مع اثنين من خوارزميات التحسين (GW,GA) ومصنفين آخرين (SVM,KNN) نحصل على نتائج توضح DRB يتمتع بأداء عال بدقة 99,67% و نستخدم أيضا تقنية تغيير الزاوية لزيادة معدل الدقة.

الكلمات المفتاحية: صور الرنين المغناطيسي ; ورم في المخ ; استخراج الميزة على أساس القاعدة العميقة ؛ الامثل تصنيف .

Liste de figures :

Figure 1 LES COMPOSANTS DE LA STRUCTURE TISSULAIRE DU CERVEAU HUMAIN.....	14
Figure 2 Présentation des différents lobes du cerveau.....	15
Figure 3 Exemple d'une tumeur Gliomes à gauche et une tumeur Méningiomes à droite.....	17
Figure 4 (a) Plan axial, (b) Plan coronal et (c) Plan sagittal de l'image cérébrale.	21
Figure 5 (a) Plan axial, (b) Plan sagittal et (c) Plan coronal du cerveau observé par l'IRM.....	21
Figure 6 Les domaines d'application de l'ia en médecine	22
Figure 7 le processus typique du ml	24
Figure 8 La relation entre l'intelligence artificielle, le ML et le deep learning.	25
Figure 9 PROCESSUS DE CLASSIFICATION.....	31
Figure 10 Le paramètre de l'orientation du HOG.....	33
Figure 11 la distribution des pixels dans une cellule d'un bloc.....	33
Figure 12 Méthode de k plus proches voisins.....	38
Figure 13 Machine à vecteurs de support (SVM).....	40
Figure 14 Structure d'un neurone artificiel.....	41
Figure 15 exemple d'arbre de décision.....	42
Figure 16 Architecture générale du classificateur DRB.	43
Figure 17 Processus d'optimisation.	46
Figure 18 : Représentation en 3D de la fonction à optimiser.....	48
Figure 19 : Organigramme de fonction de la méthode GWO.....	50
Figure 20 : Schéma général d'un PSO.....	51
Figure 21 Processus de classification avec optimisation.	52
Figure 22 MATLAB Logo.....	56
Figure 23 exemple de courbe ROC (AUC).	60
Figure 24 Architecture de modèle ResNet-50.....	62
Figure 25 Courbes Roc des classifieurs DRB,KNN et SVM et l'extracteur HOG avec les trois Data-sets.....	63
Figure 26 Matrices de confusion du KNN et l'extracteur HOG avec les trois Data-sets.....	64
Figure 27 Matrices de confusion du DRB et l'extracteur HOG avec les trois Data-sets.....	64

Figure 28 Courbes Roc des classifieurs DRB,KNN et SVM et l'extracteur GIST avec les trois Data-sets.....	65
Figure 29 Matrices de confusion du SVM et l'extracteur GIST avec les trois Data-sets	66
Figure 30 Matrices de confusion du KNN et l'extracteur GIST avec les trois Data-sets	66
Figure 31 Matrices de confusion du DRB et l'extracteur GIST avec les trois Data-sets	66
Figure 32 Courbes Roc des classifieurs DRB,KNN et SVM et l'extracteur ResNet-50 avec les trois Data-sets	67
Figure 33 Matrices de confusion du SVM et l'extracteur ResNet-50 avec les trois Data-set	68
Figure 34 Matrices de confusion du KNN et l'extracteur ResNet-50 avec les trois Data-set	68
Figure 35 Matrices de confusion du DRB et l'extracteur ResNet-50 avec les trois Data-set.	68
Figure 36 ResNet-50 avec DRB,SVM,KNN et Dataset1+2 optimisé par PSO,GW,GA	70
Figure 37 HOG avec DRB,SVM,KNN et Dataset1+2 optimisé par PSO,GW,GA.....	71
Figure 38 GIST avec DRB,SVM,KNN et Dataset1+2 optimisé par PSO,GW,GA.....	73
Figure 39 Comparaison des courbes entre l'angle 90 et 120 sur les trois classifieur avec HOG	74
Figure 40 Comparaison des matrices de confusion entre l'angle 90 et 120 sur le classifieur SVM avec HOG.....	75
Figure 41 Comparaison des matrices de confusion entre l'angle 90 et 120 sur le classifieur KNN avec HOG.....	75
Figure 42 Comparaison des matrices de confusion entre l'angle 90 et 120 sur le classifieur DRB avec HOG	76

Liste des tableaux :

Tableau 1 Matrice de confusion.....	58
Tableau 2 Résultats comparatifs de l'extracteur Hog avec les classifieurs DRB,KNN,SVM	63
Tableau 3 Résultats comparatifs de l'extracteur GIST avec les classifieurs DRB,KNN,SVM	65
Tableau 4 Résultats comparatifs de l'extracteur ResNet-50 avec les classifieurs DRB,KNN,SVM.....	67
Tableau 5 Résultats comparatifs de l'extracteur ResNet-50 avec les classifieurs DRB,KNN,SVM optimisé par PSO,GW,GW	69
Tableau 6 Résultats comparatifs de l'extracteur HOG avec les classifieurs DRB,KNN,SVM optimisé par PSO,GW,GW	71
Tableau 7 Résultats comparatifs de l'extracteur GIST avec les classifieurs DRB,KNN,SVM optimisé par PSO,GW,GW	72
Tableau 8 Résultats comparatifs entre l'angle 90 et 120 sur les trois classifieur avec HOG .	74
Tableau 9 Evolution des résultats du DRB	77

Liste des abréviations :

IRM	Imagerie par Résonance Magnétique
TEP	Tomographie par Émission de Positons
TR	le temps de répétition des séquences d'impulsion
TE	le temps d'écho
DRB	Deep Rule Based
KNN	K-nearest neighbors algorithm
SVM	Support Vector Machine
PSO	Particles Swarm Optimisation
GW	Grey Wolf
GA	Genetic Algorithm
IA	Intelligence Artificielle
ML	Machine Learning
DL	Deep Learning
FS	Feature Selection
ResNet	Residual Network
HOG	Histogram of Oriented Gradients
GIST	Global Image Structure
CNN	Convolution Neural Network
AUC	Area Under Curv

Table de matières :

Liste de figures

Liste des tableaux

Liste des abréviations

Introduction Générale :	10
Chapitre 1 : Notion d'étude	13
1- Introduction :	13
2- Description du cerveau :	13
2-1 A quoi sert le cerveau ?	13
2-2 De quoi est composé le cerveau ?	13
2-3 L'anatomie du cerveau :	15
3- Tumeur cérébrale :	16
3-1 Les tumeurs cérébrales primaires :	16
3-2 Les tumeurs cérébrales métastatiques :	17
3-3 Tumeurs cérébrales bénignes et tumeurs cérébrales malignes :	18
3-4 Symptômes des tumeurs cérébrales :	18
3-5 Diagnostic des tumeurs cérébrales :	18
4- L'Imagerie par Résonance Magnétique IRM	19
4-1 C'est quoi l'IRM ?	19
4-2 Utilisations de l'IRM :	19
4-3 Les inconvénients de l'IRM :	20
4-4 Les avantages de l'IRM :	20
4-5 Dépistage du cancer du cerveau basé imagerie par résonance magnétique :	20
4-6 Limites des IRM durant le diagnostic du cancer du cerveau :	22
5- l'intelligence artificielle :	22
5-1 C'est quoi l'Intelligence artificielle ?	22

5-2	l'intelligence artificielle et son impact sur l'aspect médical :.....	22
5-3	Une aide précieuse pour la détection de tumeurs cérébrales :	23
5-4	Apprentissage automatique :.....	23
6-	Deep Learning :.....	24
6-1	Apprentissage supervisé et non supervisé en Deep Learning :.....	25
6-2	Applications du Deep Learning	26
6-3	Défis et considérations du Deep Learning :.....	26
7-	Conclusion :.....	27
Chapitre 2 : Classification.....		29
1-	Introduction :.....	29
2-	La classification	29
2-1	Définitions :.....	29
2-2	Domaines d'application :	30
2-3	Etapes de la classification automatique :	30
3-	Extraction de caractéristiques :	31
3-1	Extraction de bas niveau :.....	31
a)	Histogramme des gradients orientés (HOG):.....	32
b)	GIST (Global Image Structure) :	34
3-2	Extracteurs de plus haut niveau :.....	35
c)	ResNet-50 :	35
4-	Approches de classification :	36
4-1	Classification Non-supervisée (Clustering) :	36
4-1-1	Algorithmes de classification non-supervisée :	37
a)	Méthode k-means :.....	37
b)	Fuzzy C-means :.....	37
c)	Hierarchical clustering :	37
4-2	Classification supervisée :.....	38

4-2-1	Algorithmes de classification supervisée	38
a)	Méthode de k plus proches voisins :	38
b)	Machine à vecteurs de support (SVM).....	39
c)	Réseau de neurones :.....	40
d)	Arbre de décision :	41
e)	Deep Rule-Based (DRB) :.....	42
e.1)	Définition :.....	42
e.2)	Architecture générale du classificateur DRB :	43
e.3)	Caractéristiques de DRB :.....	44
5-	Optimisation :.....	45
5-1	Le processus d'optimisation :	45
5-2	Le processus d'optimisation :	45
5-3	Les problèmes d'optimisation :.....	46
5-4	Les algorithmes d'optimisation :.....	47
5-5	Classification avec optimisation :	52
6-	Conclusion :.....	53
Chapitre 3 : Analyse et Résultats		55
1-	Introduction :.....	55
2-	Environnement de travail :.....	55
2-1	Ressources matériel :.....	55
2-3	Langages de programmation :	55
2-4	Les Modèles et architectures implémentés :.....	56
2-5	Data-set utilisés :	57
3-	Mesures de performance des algorithmes intelligents :.....	57
3-1	La matrice de confusion :	58
3-2	La surface sous la courbe ROC (AUC, Area Under the Curve) :	60
4-	Les étapes d'implémentation :.....	61

4-1	Première étape : Classification sans optimisation (basique):	61
4-2	Résultats et discussions :	63
4-3	Changement d'Angle :	74
5-	L'évolution du classifieur DRB avec et sans optimisation et avec le changement d'angle :.....	77
6-	Conclusion :	77
	Conclusion Générale :	78
	Bibliographies :	79

Introduction Générale :

- **Contexte :**

Les maladies neurologiques telles que les tumeurs cérébrales, les accidents vasculaires cérébraux et la maladie d'Alzheimer ont un impact considérable sur la santé et la qualité de vie des patients, ainsi que sur leur entourage. L'imagerie par résonance magnétique (IRM) cérébrale est une technique d'imagerie non invasive qui peut aider à diagnostiquer et à suivre ces maladies. Cependant, l'interprétation des images d'IRM cérébrales peut être difficile, en particulier pour les cliniciens inexpérimentés, et la précision du diagnostic peut varier considérablement.

L'intelligence artificielle (IA) est également de plus en plus utilisée dans le domaine de l'imagerie médicale, y compris l'IRM. Les algorithmes d'apprentissage automatique, tels que les réseaux de neurones convolutifs (CNN).

L'utilisation de l'IA pour l'interprétation d'IRM peut aider à améliorer la précision du diagnostic, à réduire les erreurs et à accélérer le processus d'analyse. Les réseaux de neurones convolutifs ont montré des performances prometteuses dans la classification des tumeurs cérébrales, la détection des anévrismes cérébraux, la segmentation du cerveau et l'identification des régions d'intérêt.

- **Problématique :**

Cependant, le choix de l'approche de classification la plus appropriée pour les images d'IRM cérébrales reste un défi. Les méthodes traditionnelles de classification, telles que les classificateurs basés sur les statistiques, peuvent ne pas être suffisamment précises pour distinguer les nuances subtiles dans les images.

En outre, les résultats des réseaux de neurones convolutifs peuvent être difficiles à interpréter, ce qui peut être un problème dans un contexte clinique.

L'analyse des données médicale pour l'aide au diagnostic et à la thérapeutique peut se faire grâce aux techniques de l'Intelligence Artificielle et plus particulièrement l'aide de l'apprentissage automatique et profond . Par voie de conséquence, la problématique traitée dans ce mémoire est la détection intelligente d'une pathologie cérébrale choisie, et ce, pour l'aide à la décision pour le diagnostic et la thérapeutique médicaux .

- **Objectifs :**

L'objectif de notre travail est une étude sur les techniques de l'Intelligence Artificielle IA en utilisant différents algorithmes de classification et concevoir une méthode de classification des images médicales , plus particulièrement des IRMs cérébrales basée sur les techniques de l'apprentissage profond à base de DRB normales ou pathologique .

Il existe plusieurs méthodes de classification chaque méthode possède ses avantages et ses limites d'utilisation. Pour cela nous avons tester et comparer l'approche proposée avec d'autres methods.

Ce mémoire s'articule autour de trois chapitres qui nous permettront de présenter les différents aspects de notre travail.

Chapitre 1 : Exclusivement pour la portée de ce rapport. Nous proposons des concepts de recherche tels que : la définition de l'IRM, son fonctionnement et la séquence d'acquisition de l'IRM, et nous définissons ce qu'est le cerveau, de quoi il est fait et ce qu'est une tumeur. Des sections dédiées à l'intelligence artificielle, au machine learning, au deep learning, à leurs algorithmes, leurs forces et leurs faiblesses.

Chapitre 2 : Nous avons défini le processus de classification et ses différentes étapes. Nous introduisons ensuite les principaux algorithmes de classification. Ce chapitre présente en détail la DRB dans notre domaine de recherche.

Chapitre 3 : Une étude comparative des différentes méthodes proposées, des données et méthodes utilisées pour résoudre les problèmes de l'environnement de travail et la comparaison des résultats obtenus ont été réalisées ;

Enfin, les résultats des travaux menés, les résultats obtenus et surtout les perspectives qui en découlent sont présentés.

CHAPITRE 1 :

Notion d'étude

Chapitre 1 : Notion d'étude

1- Introduction :

Dans ce chapitre préliminaire, nous d'écrivons le cadre biomédical de notre travail, nous allons évoquer l'état de l'art pour le dépistage des tumeurs cérébrales. Nous commençons par description macroscopique et microscopique du cerveau, nous présentons, ensuite, les tumeurs cérébrales, leurs types, leurs symptômes ainsi que leurs diagnostic. Ainsi de suite, nous allons avoir une présentation du principe de base du système d'acquisition permettant d'obtenir d'excellentes images cérébrale : il s'agit de l'imagerie par résonance magnétique nucléaire (IRM). Nous terminons ce chapitre par l'intelligence artificielle et son impact sur l'aspect médical dans le diagnostic et la prédiction des tumeurs.

2- Description du cerveau :

Le cerveau est un organe complexe et vital du système nerveux central qui contrôle toutes les fonctions du corps humain. Il est constitué de milliards de cellules nerveuses appelées neurones qui communiquent entre elles à travers des connexions appelées synapses. Le cerveau se compose de plusieurs parties distinctes qui ont chacune leur propre fonction.

2-1 A quoi sert le cerveau ?

Protégé par la boîte crânienne, le cerveau est l'organe qui fait office de chef d'orchestre pour tous les membres et organes du corps humain. Il centralise les informations et renvoie des messages aux différents membres. Il se compose de deux hémisphères reliés: l'hémisphère gauche et l'hémisphère droit, qui contrôlent chacun la partie du corps qui lui est opposée. La main droite sera donc contrôlée par une partie de l'hémisphère gauche.

2-2 De quoi est composé le cerveau ?

Le cerveau humain a trois parties importantes en termes de structure tissulaire [1], à savoir la matière grise, la matière blanche et le liquide céphalo-rachidien (CSF) comme on peut le voir sur la figure 1

La matière grise est faite des corps cellulaires des neurones, de leurs dendrites et d'autres cellules. Elle est responsable de notre activité sensori-motrice et de nos fonctions cognitives comme la lecture, le calcul, l'attention, la mémoire...

Les neurones sont des cellules du cerveau qui servent à transmettre les informations. Ils sont tous interconnectés et communiquent entre eux par messages électriques et chimiques au travers de milliers de petites branches appelées dendrites sur lesquelles se terminent les axones, prolongement des neurones pour transmettre l'information à distance.

La matière blanche est, quant à elle, constituée de ces axones, enveloppés d'un manchon graisseux de myéline, reliant les différentes régions de matière grise afin qu'elles échangent leur information.

Le liquide céphalo-rachidien (LCR) est un liquide clair et incolore qui circule dans les cavités du cerveau appelées ventricules, ainsi que dans les espaces qui entourent la moelle épinière et le cerveau.

Le LCR est produit principalement par les plexus choroïdes, des structures vasculaires situées dans les ventricules cérébraux. Il joue plusieurs rôles importants dans le système nerveux central, notamment :

- Protéger le cerveau et la moelle épinière en amortissant les chocs et en agissant comme un coussin de protection contre les dommages physiques.
- Fournir des nutriments et de l'oxygène aux cellules nerveuses du cerveau et de la moelle épinière.
- Éliminer les déchets métaboliques et les produits résiduels du cerveau et de la moelle épinière.[2]

Le LCR est également utilisé pour diagnostiquer certaines maladies neurologiques, telles que la méningite, l'encéphalite et la sclérose en plaques, en prélevant un échantillon de liquide dans la colonne vertébrale lors d'une ponction lombaire.

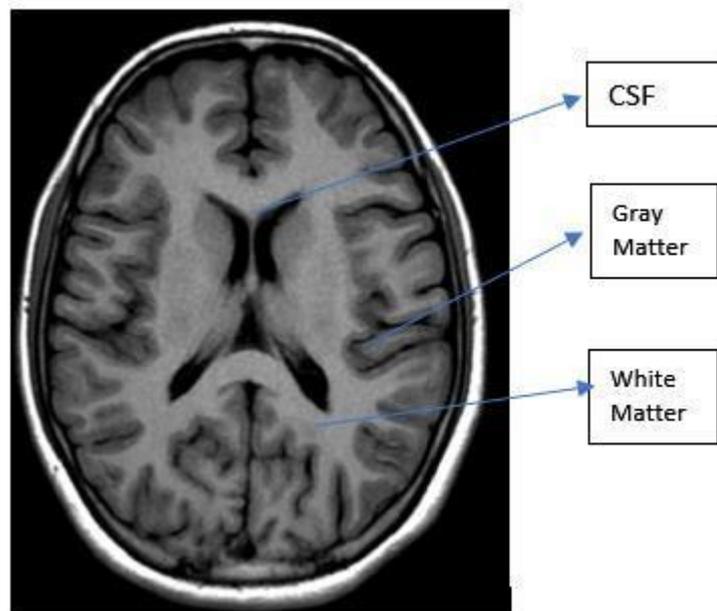


FIGURE 1 LES COMPOSANTS DE LA STRUCTURE TISSULAIRE DU CERVEAU HUMAIN. [1]

Ces trois parties ainsi que la tumeur présentent un contraste différent lorsqu'elles sont imagées sous différentes caractéristiques physiques et jouent un rôle clé dans l'imagerie IRM pour la détection des tumeurs cérébrales .

2-3 L'anatomie du cerveau :

Chacun des hémisphères du cerveau est divisé en cinq régions (quatre extérieures et une enfouie : le cortex insulaire ou Insula) on peut les voir sur la figure 2 . Ces lobes sont composés de zones plus petites qui gèrent des fonctions précises. Elles sont appelées aires cérébrales. On en dénombre aujourd'hui près de 200 par hémisphère[3].

Dans ces zones, les neurones sont spécialisés dans une fonction précise comme transmettre un message visuel, sonore, sensitif. Les zones des différents lobes coopèrent pour réaliser les tâches complexes. Par exemple, le langage fait intervenir plusieurs zones de différents lobes pour nous permettre de parler ou lire.

- **Le lobe frontal** est le siège de la parole, du langage et du raisonnement. Il a également la fonction de gérer les mouvements des membres.
- **Le lobe pariétal** est la partie qui va s'occuper du repérage dans l'espace, des sens et de la lecture
- **Le lobe occipital** est dédié à la vision
- **Le lobe temporal** est la zone où se situent le langage, la mémoire et l'émotivité.
- **Le cortex insulaire ou Insula** est spécialisé dans la perception de soi/sa conscience, dans la socialisation et impacte également les émotions.

Les régions associées à certaines fonctions sont localisées à des endroits variant légèrement d'un individu à l'autre. Par ailleurs, la spécialisation hémisphérique de certaines fonctions comme le langage varie : elle est majoritairement située dans l'hémisphère gauche chez les droitiers mais peut se situer dans l'hémisphère droit, comme chez la plupart des gauchers.

Ces différentes régions du cerveau sont connectées pour combiner les messages. C'est cette coopération des zones qui permet, par exemple, la reconnaissance de visages ou de lieux.

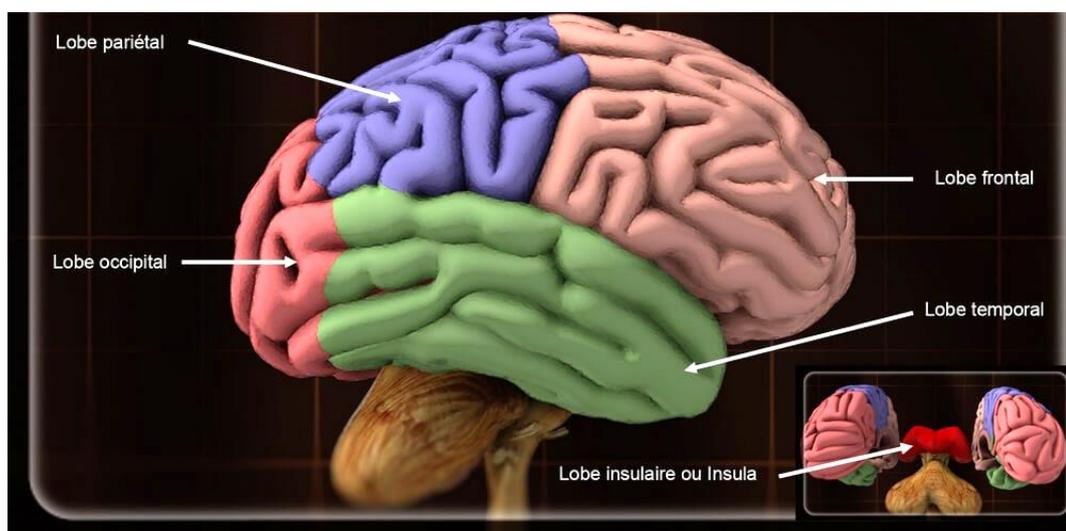


FIGURE 2 PRESENTATION DES DIFFERENTS LOBES DU CERVEAU. [3]

3- Tumeur cérébrale :

Après l'âge de vingt ans, nous perdons chaque jour des milliers de neurones. Il existe en outre un certain nombre de maladies, telles que la maladie de Parkinson, la maladie d'Alzheimer et les tumeurs du cerveau, qui augmentent cette dégénérescence cellulaire.

Une tumeur cérébrale, connue sous le nom de tumeur intracrânienne, est une masse anormale de tissu dans laquelle les cellules se développent et se multiplient de manière incontrôlable dans le cerveau, apparemment non contrôlée par les mécanismes qui contrôlent les cellules normales. Plus de 150 tumeurs cérébrales différentes ont été documentées, mais les deux principaux groupes de tumeurs cérébrales sont appelés primaires et métastatiques ou secondaire.

Les causes exactes des tumeurs cérébrales ne sont pas toujours connues, mais il existe des facteurs de risque qui peuvent augmenter la probabilité de développer une tumeur cérébrale, tels que l'exposition aux radiations, les antécédents familiaux de tumeurs cérébrales, les troubles génétiques et les antécédents de certaines maladies.

3-1 Les tumeurs cérébrales primaires :

Les tumeurs cérébrales primaires comprennent les tumeurs qui proviennent des tissus du cerveau ou de l'environnement immédiat du cerveau. Les tumeurs primaires sont classées comme gliales (composées de cellules gliales) ou non gliales (développées sur ou dans les structures du cerveau, y compris les nerfs, les vaisseaux sanguins et les glandes) et bénignes ou malignes.

Les tumeurs primaires les plus fréquentes :

- **Gliomes** : Les gliomes représentent 65 % de toutes les tumeurs cérébrales primitives [4]. Ils ont des tumeurs malignes intrinsèques à la cavité cérébrale et peuvent affecter le système nerveux central provoquant un œdème infiltrant la tumeur comme mentionné sur la figure 1.3
- **Méningiomes** : les méningiomes sont pour la plupart des tumeurs cérébrales bénignes et extrinsèques dans les méninges (la membrane qui sépare le cerveau de la moelle épinière) comme mentionné sur la figure 3

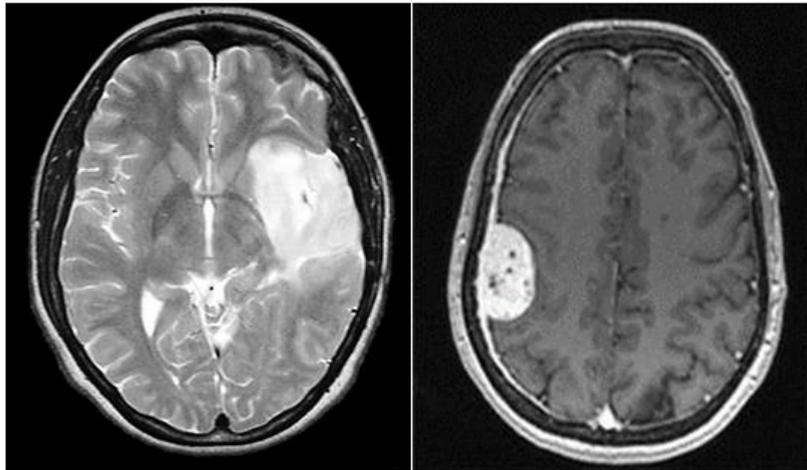


FIGURE 3 EXEMPLE D'UNE TUMEUR GLIOMES A GAUCHE ET UNE TUMEUR MENINGIOMES A DROITE.[4]

3-2 Les tumeurs cérébrales métastatiques :

Les métastases cérébrales sont beaucoup plus fréquentes que les tumeurs primitives. Plus de 80 % des personnes atteintes de métastases cérébrales ont plus d'une métastase.[4]

Les tumeurs cérébrales métastatiques comprennent les tumeurs qui surviennent ailleurs dans le corps et migrent vers le cerveau, généralement par la circulation sanguine. Les tumeurs métastatiques sont considérées comme cancéreuses et malignes. De nombreux types de cancers peuvent se propager vers le cerveau. À savoir :

- Cancer du sein
- Cancer du poumon
- Cancer du rein
- Mélanome
- Cancer de la thyroïde
- Lymphome et leucémie

Les tumeurs métastatiques au cerveau affectent près d'un patient sur quatre atteint de cancer, soit environ 150 000 personnes par an. Jusqu'à 40% des personnes atteintes d'un cancer du poumon développeront des tumeurs cérébrales métastatiques. Dans le passé, le résultat pour les patients diagnostiqués avec ces tumeurs était très médiocre, avec des taux de survie typiques de quelques semaines seulement. Des outils de diagnostic plus sophistiqués, en plus d'approches chirurgicales et radiologiques novatrices, ont contribué à l'augmentation des taux de survie jusqu'à plusieurs années; et a également permis d'améliorer la qualité de vie des patients après le diagnostic.[5]

3-3 Tumeurs cérébrales bénignes et tumeurs cérébrales malignes :

- **Les tumeurs cérébrales bénignes :**

Elles sont constituées de cellules qui croissent lentement et qui à l'analyse microscopique n'ont pas de caractère d'activité importante.

Une tumeur cérébrale bénigne peut être soit simplement surveillée radiologiquement soit être extirpée chirurgicalement complètement ou de façon partielle. Si une tumeur bénigne n'est pas complètement enlevée (non cancéreuse), elle peut récidiver.

- **Les tumeurs cérébrales malignes :**

Elles sont classées « cancéreuses », définissent certaines tumeurs primaires ainsi que toutes les lésions cérébrales métastatiques. Elles sont constituées de cellules qui se divisent relativement vite. Ces tumeurs croissent donc rapidement et peuvent envahir et endommager les structures cérébrales importantes. Elles peuvent être traitées par la chirurgie, la radiothérapie, la chimiothérapie ou la combinaison de celles-ci.

3-4 Symptômes des tumeurs cérébrales :

Les symptômes varient en fonction de l'emplacement de la tumeur cérébrale, mais les éléments suivants peuvent accompagner différents types de tumeurs cérébrales:[6]

- Maux de tête qui peuvent être plus sévères le matin ou réveiller le patient la nuit
- Convulsions ou convulsions
- Difficulté à penser, à parler ou à articuler
- Changements de personnalité
- Faiblesse ou paralysie d'une partie ou d'un côté du corps
- Perte d'équilibre ou vertiges
- Changements de vision
- Changements auditifs
- Engourdissement ou picotement du visage
- Nausées ou vomissements, difficultés à avaler
- Confusion et désorientation

3-5 Diagnostic des tumeurs cérébrales :

Des techniques d'imagerie sophistiquées peuvent identifier les tumeurs cérébrales. Les outils de diagnostic comprennent la tomographie par ordinateur (TDM ou tomographie assistée par ordinateur) et l'imagerie par résonance magnétique (IRM). D'autres séquences d'IRM peuvent aider le chirurgien à planifier la résection de la tumeur en fonction de l'emplacement des voies nerveuses normales du cerveau. L'IRM peropératoire est également utilisée pendant la chirurgie pour guider les biopsies tissulaires et l'ablation de la tumeur.

La spectroscopie par résonance magnétique (SRM) est utilisée pour examiner le profil chimique de la tumeur et déterminer la nature des lésions observées à l'IRM.

La tomographie par émission de positons (TEP) peut aider à détecter les tumeurs cérébrales récurrentes.

Parfois, la seule façon de poser un diagnostic définitif d'une tumeur cérébrale est par une biopsie. Le neurochirurgien effectue la biopsie et le pathologiste pose le diagnostic final, déterminant si la tumeur semble bénigne ou maligne et la classant en conséquence.

4- L'Imagerie par Résonance Magnétique IRM

4-1 C'est quoi l'IRM ?

Un examen radiologique, l'Imagerie par Résonance Magnétique (IRM) est une technique d'imagerie médicale non-invasive, indolore et fiable qui permet d'obtenir des vues en 2D et 3D de l'ensemble des tissus du corps humain, qu'il s'agisse des tissus durs (les os) ou des tissus mous (moelle osseuse, système nerveux central, muscles, cœur). L'IRM s'avère souvent plus précise que la radiographie, l'échographie ou le scanner.[7]

L'IRM utilise des ondes magnétiques et des champs radiofréquences pour produire des images du cerveau. Les images IRM sont utilisées pour diagnostiquer et évaluer une grande variété de conditions médicales du cerveau, notamment les accidents vasculaires cérébraux, les tumeurs cérébrales, les maladies neurodégénératives et les lésions cérébrales traumatiques.[7]

4-2 Utilisations de l'IRM :

L'IRM est préférée à la tomographie par ordinateur (TDM) lorsque les médecins ont besoin de plus de détails pour les tissus mous, par exemple pour visualiser des anomalies du cerveau, de la colonne vertébrale, des muscles et du foie. L'IRM est particulièrement utile pour identifier les tumeurs dans ces tissus.

L'IRM est également utilisée pour :

- Mesurer certaines molécules cérébrales qui aident à distinguer les tumeurs cérébrales des abcès cérébraux.
- Identifier les anomalies dans les organes reproducteurs féminins et les fractures de la hanche et du bassin.
- Pour aider les médecins à évaluer les anomalies articulaires (telles que les déchirures des ligaments ou du cartilage du genou) et les entorses.
- Aider les médecins à évaluer les saignements et les infections.

L'IRM est également utilisée lorsque les risques de la TDM sont élevés. Par exemple, l'IRM peut être préférée chez les personnes ayant déjà eu une réaction aux agents de contraste iodés utilisés dans la TDM, et chez la femme enceinte (en raison des risques que les radiations peuvent comporter pour le fœtus).

Une IRM réalisée après l'injection d'un agent de contraste au gadolinium dans une veine aide les médecins à évaluer l'inflammation, les tumeurs et les vaisseaux sanguins.

L'injection de ce médicament dans l'articulation aide les médecins à voir plus clairement les anomalies articulaires, surtout si elles sont complexes (telles que blessure, dégénérescence des ligaments et cartilage du genou). [8]

4-3 Les inconvénients de l'IRM :

L'IRM prend plus de temps que la TDM. De plus, l'IRM est généralement moins susceptible d'être immédiatement disponible que la tomodensitométrie. Par conséquent, la tomodensitométrie peut être préférable dans des situations d'urgence telles qu'une blessure grave ou un accident vasculaire cérébral. L'IRM est également plus chère que la tomodensitométrie.

Les autres inconvénients sont :

- La claustrophobie et parfois la difficulté à faire tenir une personne dans l'appareil d'IRM parce que c'est un espace étroit et clos
- Les effets du champ magnétique sur les dispositifs métalliques implantés dans le corps
- Les réactions à l'agent de contraste

4-4 Les avantages de l'IRM :

Le principal avantage de l'IRM est qu'elle n'utilise pas de radiations pour obtenir les images. Grâce à sa technologie de pointe avec des aimants et des systèmes informatisés, les résultats sont de meilleure qualité. Ainsi, les médecins peuvent observer la structure interne du corps de manière plus détaillée. Cependant, cela la rend considérablement plus chère qu'une radiographie ou un scanner conventionnel.

Comme autres avantages pratiques :

- Elle n'utilise pas de rayonnements ionisants
- Jusqu'à présent, aucun effet indésirable n'a été constaté
- Les images peuvent être obtenues dans différents plans
- Elle fournit des informations sur le métabolisme et la composition des tissus
- L'agent de contraste utilisé, le gadolinium, ne produit pas d'effets indésirables significatifs.

4-5 Dépistage du cancer du cerveau basé imagerie par résonance magnétique :

Les symptômes qui se manifestent sont typiques du type de tumeur cérébrale du patienta et aussi dépendent de la localisation de la tumeur. Si le médecin soupçonne un patient portant une tumeur cérébrale, il est généralement conseillé au patient de subir soit une imagerie médicale (comme l'imagerie par résonance magnétique (IRM), la tomographie par ordinateur (CT), angiographie cérébrale, etc.) ou d'autres tests comme l'audiométrie, mesure de l'acuité du champ visuel ou évaluation endocrinienne.

L'IRM est utilisée pour imager les tissus mous du cerveau à l'aide d'un champ magnétique et d'une radio ondes de fréquence [6].

L'IRM est plus efficace que la tomodensitométrie engendrant de cette imagerie, la plus demandée par les neurologues durant le processus de dépistage du cancer du cerveau [9].

En effet, l'IRM fournit un contraste notable pour différents tissus cérébraux par rapport à la tomodensitométrie [10]. De plus, l'IRM peut reconnaître le sang qui coule et les dysfonctionnements vasculaires cryptiques, elle peut détecter les maladies nerveuses, et fut réalisée sans rayonnement ionisant [10]. L'IRM offre le meilleur contraste en termes de structure tissulaire car elle a la capacité d'afficher des images jusqu'à 65535 niveaux de gris, ce qui est impossible à être visualisé à l'œil nu [1]. Pour visualiser les caractéristiques détaillées du cerveau et d'autres structures crâniennes, l'IRM bénéficie de la visualisation de l'anatomie dans trois plans différents ; axiale, coronale et sagittale. La Figure 4 et la Figure 5 montrent les trois anatomies différentes du cerveau humain ainsi que les vues axiales, sagittales et coronales du cerveau tiré de l'imagerie par résonance magnétique [10].

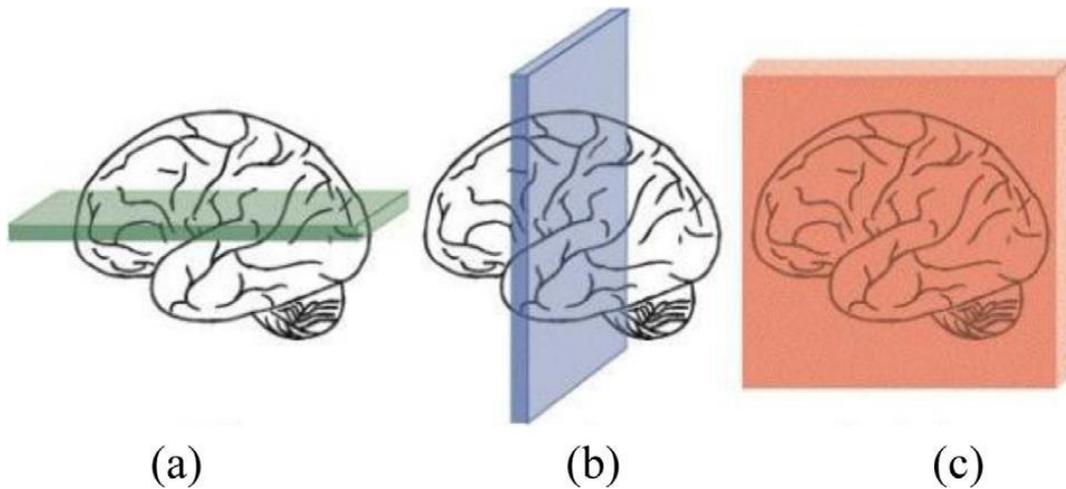


FIGURE 4 (A) PLAN AXIAL, (B) PLAN CORONAL ET (C) PLAN SAGITTAL DE L'IMAGE CEREBRALE.[10]

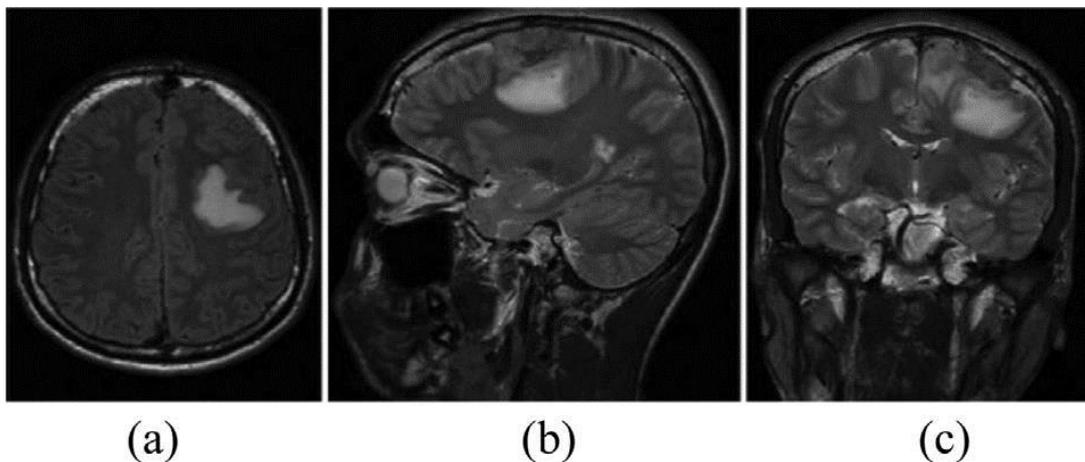


FIGURE 5 (A) PLAN AXIAL, (B) PLAN SAGITTAL ET (C) PLAN CORONAL DU CERVEAU OBSERVE PAR L'IRM. [10]

4-6 Limites des IRM durant le diagnostic du cancer du cerveau :

Nous distinguons quelques facteurs dégradants la puissance des images IRM [1], en effet, les équipements impliqués dans la génération des IRM produisent un bruit qui perturbe le processus d'acquisition d'images.

En raison du son, les images IRM intègrent du bruit sous forme d'artefacts physiques. Des artefacts de mouvement sont également introduits si le sujet examiné bouge pendant l'acquisition de l'image. Ces bruits sont inévitables et nécessitent un moyen standard pour les supprimer.

En vue de détecter une tumeur cérébrale à son stade naissant, l'expérience du radiologue compte beaucoup [6]. Depuis, le diagnostic de la tumeur et l'évaluation de son type, désignent les tâches principales à portée de main avant de suggérer tout autre médicament. Voir plus, une biopsie doit être effectuée où une chirurgie cérébrale définitive est essentielle dans ce sens, contrairement à d'autres parties du corps. En effet, la subjectivité potentielle de la biopsie est indésirable pour les patients.

5- l'intelligence artificielle :

5-1 C'est quoi l'Intelligence artificielle ?

L'IA peut se définir comme un ensemble de technologies visant à réaliser informatiquement des tâches cognitives traditionnellement effectuées par l'humain. Le grand public imagine l'IA comme une machine aussi intelligente qu'un humain, qui a conscience d'elle-même et peut faire des choix en toute autonomie. Cette vision est très loin de la réalité.

5-2 l'intelligence artificielle et son impact sur l'aspect médical :

L'IA peut avoir un impact positif sur la pratique de la médecine de différentes façons, que ce soit en accélérant le rythme de la recherche ou en aidant les cliniciens à prendre de meilleures décisions. Voici quelques exemples d'utilisation de l'IA dans la Figure 6



FIGURE 6 LES DOMAINES D'APPLICATION DE L'IA EN MEDECINE .[11]

L'intelligence artificielle (IA) est de plus en plus utilisée dans le domaine médical pour aider à diagnostiquer et à prédire les tumeurs. Cette technologie peut aider les médecins à détecter les tumeurs plus rapidement et plus précisément, ce qui peut améliorer les résultats pour les patients.

Les systèmes d'IA utilisés pour diagnostiquer les tumeurs sont souvent basés sur des algorithmes d'apprentissage automatique. Ces algorithmes peuvent apprendre à partir de grandes quantités de données, ce qui leur permet de détecter des modèles subtils que les médecins pourraient manquer. Les données utilisées pour entraîner ces algorithmes peuvent inclure des images de rayons X, des scans de tomographie par ordinateur (CT) et des images de résonance magnétique (IRM).

5-3 Une aide précieuse pour la détection de tumeurs cérébrales :

La nouvelle technique, appelée "histologie Raman stimulée" (SRH), peut révéler l'infiltration tumorale dans les tissus humains en collectant la lumière laser diffusée. Ainsi, il élucide les caractéristiques essentielles qui ne sont normalement pas vues dans les images histologiques standard.[11]

Pour construire le nouvel outil d'IA, les chercheurs ont demandé à l'algorithme d'analyser plus de 2,5 millions d'échantillons de tumeurs provenant de 415 patients. Ces échantillons de tissus ont ensuite été classés en 13 catégories représentant les tumeurs cérébrales les plus courantes. Ils ont ensuite recruté 278 patients devant subir une résection de tumeur cérébrale ou une chirurgie de l'épilepsie. Des échantillons de tumeurs cérébrales de ces patients ont ensuite été biopsiés, divisés en échantillons égaux et assignés au hasard à un groupe expérimental ou témoin.[11]

En utilisant des outils d'intelligence artificielle pour traiter et analyser les images microscopiques capturées par l'imagerie, en moins de deux minutes et demie,[11] les chirurgiens ont pu diagnostiquer avec précision une tumeur cérébrale jusqu'alors indétectable et donc pouvoir l'enlever.

“Le SRH va révolutionner le domaine de la neuropathologie en améliorant la prise de décision pendant la chirurgie et en fournissant une évaluation de niveau expert dans les hôpitaux où les neuropathologistes formés ne sont pas disponibles”[12].

5-4 Apprentissage automatique :

L'apprentissage automatique ou Le ML est une discipline de l'IA qui offre aux ordinateurs la possibilité d'apprendre à partir d'un ensemble d'observations que l'on appelle ensemble d'apprentissage.

Chaque observation, comme par exemple « j'ai mangé tels et tels aliments à tel moment de la journée pendant telle période ce qui a causé telle maladie » est décrite au moyen de deux types de variables :

Les premières sont les variables prédictives (ou attributs ou caractéristiques), dans notre exemple mon âge, mon dossier médical, mes antécédents médicaux. Ce sont les variables à partir desquelles on espère pouvoir faire des prédictions. les n variables prédictives associées à une observation seront notées comme un vecteur $x=(x_1, \dots, x_n)$ à n composantes. Un ensemble de M observations sera constitué de M tels vecteurs $x(1), \dots, x(M)$.

Une variable cible dont on souhaite prédire la valeur pour des événements non encore observés. Dans notre exemple, il s'agirait de la maladie contractée. on notera y cette variable cible.

En résumé, la valeur de la variable y dépend de :

- Une fonction $F(x)$ déterminée par les variables prédictives.
- Un bruit (ϵ) qui est le résultat d'un des paramètres dont on ne peut pas tenir compte.

Aussi bien F que ϵ ne seront jamais connues mais l'objectif d'un modèle de ML est d'obtenir la meilleure approximation possible de F à partir des observations disponibles. Cette approximation sera notée f , on l'appelle la fonction de prédiction.

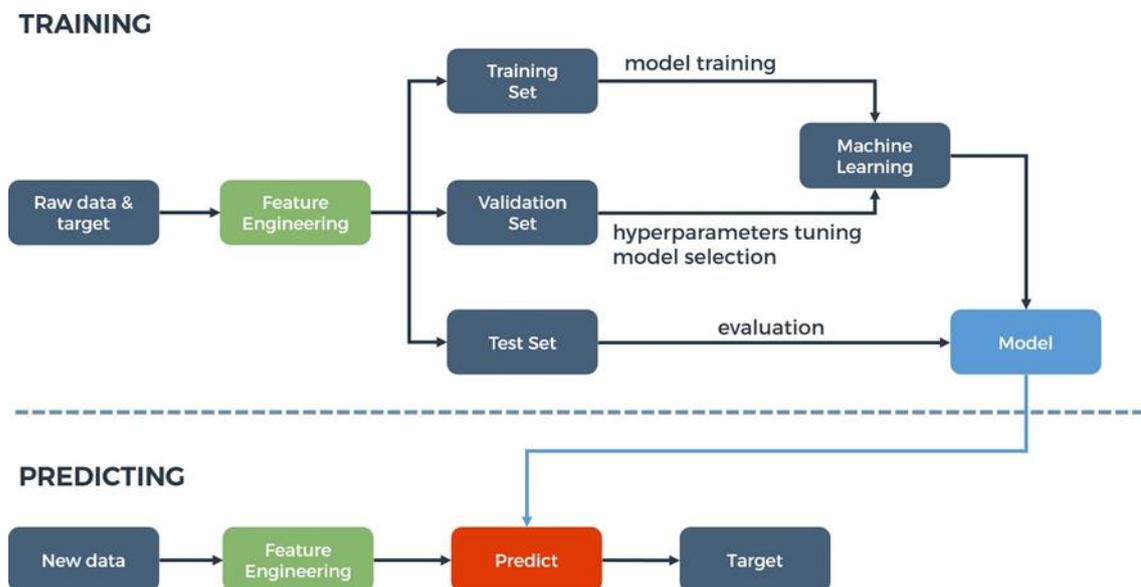


FIGURE 7 LE PROCESSUS TYPIQUE DU ML .[30]

6- Deep Learning :

Le Deep Learning est une sous-branche de l'apprentissage automatique (machine learning) qui vise à modéliser et à simuler le fonctionnement du cerveau humain pour résoudre des problèmes complexes d'apprentissage et de reconnaissance de motifs. Il est basé sur des réseaux de neurones artificiels profonds, qui sont capables d'apprendre des représentations de données hiérarchiques à partir de grandes quantités de données non structurées.

Dans ce rapport, nous explorerons les principes fondamentaux du Deep Learning, ses applications et les défis associés.

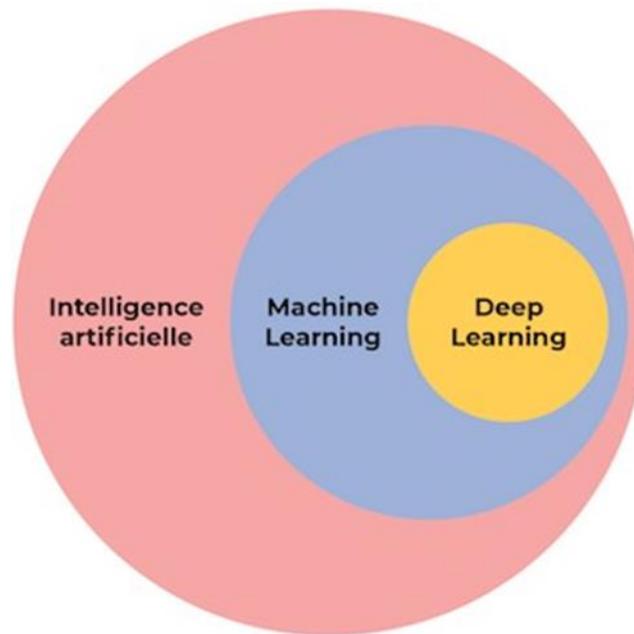


FIGURE 8 LA RELATION ENTRE L'INTELLIGENCE ARTIFICIELLE, LE ML ET LE DEEP LEARNING.[30]

Les réseaux de neurones profonds sont composés de couches de neurones interconnectés, avec une architecture profonde comprenant plusieurs couches cachées entre la couche d'entrée et la couche de sortie. Chaque neurone dans une couche est connecté aux neurones des couches adjacentes par des poids de connexion. L'apprentissage dans les réseaux de neurones profonds se fait par la propagation avant (forward propagation) et la rétropropagation du gradient (backpropagation), où les poids des connexions sont ajustés pour minimiser une fonction de perte prédéfinie.

6-1 Apprentissage supervisé et non supervisé en Deep Learning :

Le Deep Learning peut être effectué en utilisant des techniques d'apprentissage supervisé et non supervisé.

- **Apprentissage supervisé :** Dans l'apprentissage supervisé, les réseaux de neurones profonds sont entraînés sur un ensemble de données étiquetées, où les exemples d'entrée sont associés à des étiquettes de sortie correspondantes. Le réseau apprend à généraliser à de nouvelles données en ajustant les poids des connexions à partir des erreurs de prédiction. Les exemples d'apprentissage supervisé comprennent la classification d'images, la reconnaissance vocale, la traduction automatique, etc.
- **Apprentissage non supervisé :** L'apprentissage non supervisé implique l'utilisation de données non étiquetées pour découvrir des motifs et des structures sous-jacentes. Les réseaux de neurones profonds peuvent être utilisés pour effectuer des tâches telles que la réduction de dimension, la classification non supervisée (clustering), la génération de contenu, etc.

6-2 Applications du Deep Learning

Le Deep Learning a été appliqué avec succès dans de nombreux domaines, notamment :

- **Vision par ordinateur** : La reconnaissance d'objets, la détection d'objets, la segmentation sémantique, la reconnaissance faciale, etc.
- **Traitement du langage naturel** : La traduction automatique, la génération de texte, la compréhension du langage naturel, le résumé automatique, etc.
- **Reconnaissance vocale** : La conversion de la parole en texte, la synthèse vocale, l'identification des locuteurs, etc.
- **Bio-informatique** : L'analyse de séquences génétiques, la prédiction de structures de protéines, la classification de tumeurs, etc.
- **Finance** : La prédiction de séries temporelles financières, la détection de fraudes, l'analyse de sentiments sur les marchés, etc.
- **Robotique** : La perception visuelle, la manipulation d'objets, la navigation autonome, etc.

6-3 Défis et considérations du Deep Learning :

Malgré ses succès, le Deep Learning présente également des défis et des considérations importantes :

- **Besoin de grandes quantités de données d'entraînement** : Les réseaux de neurones profonds nécessitent souvent de vastes ensembles de données pour apprendre des représentations efficaces, ce qui peut être un obstacle dans des domaines où les données sont limitées ou coûteuses à obtenir.
- **Complexité computationnelle** : Les modèles de Deep Learning sont généralement complexes et nécessitent une puissance de calcul importante pour l'entraînement et l'inférence. Cela peut entraîner des exigences élevées en matière de ressources matérielles et de temps de calcul.
- **Interprétabilité** : Les réseaux de neurones profonds sont souvent considérés comme des "boîtes noires" en raison de leur complexité. L'interprétation des décisions et des prédictions peut être difficile, ce qui limite leur applicabilité dans certains domaines sensibles où l'explicabilité est primordiale.
- **Régularisation et sur ajustement** : Les réseaux de neurones profonds peuvent être sujets au sur ajustement (overfitting) si les données d'entraînement sont insuffisantes ou si le modèle est trop complexe. Des techniques de régularisation telles que le dropout, la normalisation de lot (batch normalization) et la régularisation L1/L2 sont utilisées pour atténuer ce problème.

7- Conclusion :

Aujourd'hui, l'adoption croissante de techniques automatisées et de l'apprentissage profond a conduit à un changement de paradigme dans la classification des tumeurs cérébrales. En effet, des études récentes ont montré qu'il était prometteur d'utiliser des paradigmes d'IA pour détecter la présence et le type de tumeurs cérébrales directement à partir d'images médicales, réduisant ainsi le temps d'intervention humaine et de diagnostic.

Dans le chapitre suivant, nous présenterons les concepts issus de ce domaine ainsi que le cadre méthodologique utilisé pour mettre en œuvre l'approche proposée dans le but ultime de diagnostiquer les tumeurs cérébrales.

CHAPITRE 2 :

Classification et optimisation

Chapitre 2 : Classification et optimisation

1- Introduction :

La classification de documents est l'activité du Traitement Automatique des Langues naturelles qui consiste à classer de façon automatique des ressources documentaires, généralement en provenance d'un corpus.

Cette classification peut prendre une infinité de formes, dont, la classification par genre, par thème, par opinion, ou encore par rapport au contenu du document. Elle connaît ces derniers temps un fort regain d'intérêt. Cela est dû essentiellement à la forte croissance des documents numériques disponibles et à la nécessité de les organiser de façon rapide. Et puis, regrouper des documents par catégories améliore la cohérence et la lisibilité d'une collection.

La tâche de classification est réalisée avec des algorithmes spécifiques, mis en œuvre par des systèmes de traitement de l'information. C'est une tâche d'automatisation d'un processus de classement, qui fait le plus souvent appel à des méthodes numériques (c'est-à-dire des algorithmes de classification de type mathématique). Bien que les premières bases de l'approche algorithmique de la classification automatique soient relativement anciennes, ce n'est qu'avec le développement de l'informatique qu'il est devenu possible de les mettre en œuvre.

Dans ce chapitre, nous allons définir la classification, ses méthodes, ses techniques et ses différents algorithmes en se basant essentiellement sur l'approche supervisée.

2- La classification

2-1 Définitions :

- En son sens générique, la classification est une répartition systématique en classes, en catégories, d'êtres, d'objets ou de notions ayant des caractères communs notamment afin d'en faciliter l'étude ; résultat de cette opération ». [13] Elle répond à un besoin constant de l'esprit humain d'organiser les savoirs pour les articuler entre eux, et de catégoriser les savoirs au sein d'une hiérarchie ou d'une chronologie. [14]
- La classification est une discipline reliée à plusieurs domaines, elle est connue sous divers appellations (classification, catégorisation, clustering, segmentation,...) selon les objets qu'elle traite et les objectifs qu'elle vise à atteindre. Pour attribuer une définition au terme « classification », il faudrait d'abord définir ses racines, ça vient du verbe « classer » qui désigne plus une action qu'un domaine, ou plutôt une série de méthodes qu'une théorie unifiée [15]. Le terme « classification » en anglais fait référence à l'affectation d'un individu à une classe (existant à priori) dans le cadre de l'analyse discriminante. [16]
- En mathématique, on appelle classification, la catégorisation algorithmique d'objets. Elle consiste à attribuer une classe ou catégorie à chaque objet (ou individu) à classer, en se basant sur des données statistiques. [17]

- La classification est une méthode mathématique d'analyse de données, pour faciliter l'étude d'une population d'affectif important, on les regroupe en plusieurs classes de telle sorte que les individus d'une même classe soient le plus semblables possible et que les classes soient les plus distinctes possibles. [18]
- En règle générale, La classification automatique consiste à regrouper divers objets (les individus) en sous-ensembles d'objets (les classes) dans le but d'identifier les classes auxquelles appartiennent des objets à partir de traits descriptifs (attributs, caractéristique, features). Elle peut être :
 - **Supervisée** : les classes sont connues a priori, elles ont en général une sémantique associée.
 - **Non-supervisée (clustering)** : les classes sont fondées sur la structure des objets, la sémantique associée aux classes est plus difficile à déterminer.

2-2 Domaines d'application :

- **Affectation de documents à des thèmes prédéfinis** :
 - Traitement des e-mails (redirection, filtrage) ;
 - Organisation des documents par catégories ;
 - Classement de contacts au centre d'appel.
- **Recherche d'information (Information retrieval)** :
Interrogation de textes par concepts, mots-clés, sujets, phrases visant à obtenir des résultats triés par ordre de pertinence.

2-3 Etapes de la classification automatique :

Le processus de classification passe par plusieurs étapes qui sont :

- **Extraction des données** : extraire les données et en distinguer des objets spécifiques (individus). C'est ici que les données sont préparées, nettoyées.
- **Représentation des données** : obtenir une représentation des données adaptée à l'algorithme, choisir une fonction de similarité suivant la nature de cette représentation, appréhender les différents critères sur ces données.
- **Stratégie de classement** : choisir l'algorithme suivant les exigences (performances, représentation des résultats, types de données) et fixer les éventuels paramètres.
- **Validation** : estimer la qualité de la construction produite.
- **Interprétation et utilisation** : juger si la classification répond bien aux besoins exprimés en utilisant la classification avec les données réelles .

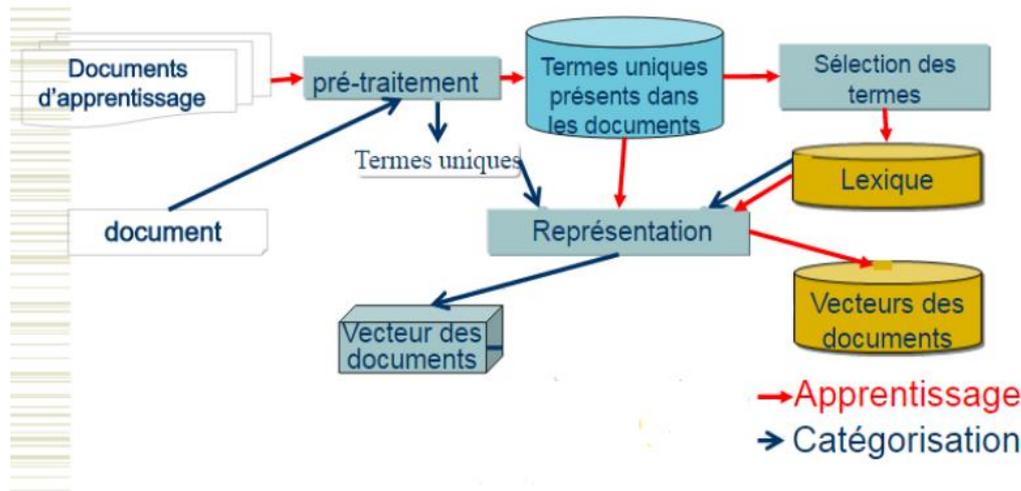


FIGURE 9 PROCESSUS DE CLASSIFICATION. [18]

3- Extraction de caractéristiques :

3-1 Extraction de bas niveau :

L'extraction de bas niveau, également connue sous le nom d'extraction de caractéristiques de bas niveau, fait référence à la première étape de traitement dans le cadre de l'analyse des données. Elle vise à extraire des informations brutes et élémentaires à partir des données d'entrée, sans y ajouter de sens ou de contexte plus élevé.[19]

L'extraction de bas niveau est souvent utilisée dans différents domaines, tels que la vision par ordinateur, le traitement d'images, le traitement du signal, etc.

- **Traitement d'images médicales :**

Extraction de caractéristiques anatomiques : Détection et extraction de caractéristiques anatomiques spécifiques dans des images médicales, comme les contours du cœur dans une imagerie par résonance magnétique cardiaque ou les lésions dans une image de scanner cérébral.

Extraction de caractéristiques de texture : Calcul de descripteurs de texture, tels que la matrice de co-occurrence de niveaux de gris (GLCM) ou les histogrammes de gradients orientés (HOG), pour capturer les motifs texturaux dans les images médicales.[20]

L'extraction de bas niveau est une étape cruciale dans le traitement des données, car elle permet de réduire la complexité des données brutes en représentations plus abstraites et informatives. Ces caractéristiques extraites peuvent ensuite être utilisées comme entrées pour des algorithmes de classification, de segmentation ou d'autres tâches d'analyse de données de niveau supérieur.

Dans notre travail, deux descripteurs de caractéristiques de bas niveau (GIST et HOG) sont utilisés.

a) Histogramme des gradients orientés (HOG):

L'histogramme de gradients orientés HOG (en anglais Histogram of Oriented Gradients) est un descripteur de caractéristiques utilisé dans le traitement d'image, principalement pour la détection d'objets. Le principe sous-jacent de HOG est que l'apparence et la forme locales de l'objet dans une image peuvent être décrites par la distribution des gradients d'intensité ou des directions de bord.[30]

Voici comment fonctionne le processus HOG :

- **Prétraitement de l'image** : L'image d'entrée est généralement convertie en niveaux de gris pour simplifier le calcul des gradients. Des opérations de normalisation peuvent également être effectuées pour égaliser l'éclairage et réduire les variations d'intensité.
- **Calcul des gradients** : Les gradients de luminance de l'image sont calculés pour capturer les variations d'intensité et les contours. Cela peut être fait en utilisant des opérateurs de détection de gradients, tels que l'opérateur de Sobel, qui calcule les dérivées partielles de l'image selon les axes x et y.
- **Calcul de l'orientation et de l'amplitude des gradients** : Pour chaque pixel, l'orientation et l'amplitude du gradient sont calculées à partir des composantes x et y des gradients. L'orientation du gradient représente la direction du changement d'intensité, tandis que l'amplitude représente l'intensité du changement.
- **Création des cellules** : L'image est divisée en cellules rectangulaires qui recouvrent l'ensemble de l'image. Les pixels appartenant à chaque cellule contribuent à l'histogramme des gradients orientés de cette cellule.
- **Construction de l'histogramme** : Pour chaque cellule, un histogramme des orientations des gradients est construit en répartissant les orientations des gradients en plusieurs bins (intervalles) et en accumulant les amplitudes correspondantes. Cela permet de capturer les informations sur la distribution des orientations des gradients dans la cellule.
- **Normalisation** : Pour améliorer la robustesse aux variations d'éclairage et de contraste, les histogrammes des cellules voisines sont regroupés en blocs et normalisés. Cela signifie que les histogrammes sont pondérés en fonction de leur contribution relative à l'intérieur du bloc.
- **Vecteur de caractéristiques HOG** : Finalement, les histogrammes normalisés des blocs sont concaténés pour former un vecteur de caractéristiques HOG global représentant l'image. Ce vecteur peut être utilisé comme entrée pour des algorithmes de classification, tels que des classifieurs linéaires ou des machines à vecteurs de support (SVM), pour la détection d'objets.

L'histogramme des gradients orientés est particulièrement efficace pour la détection d'objets dans des images, en particulier pour des objets ayant des contours bien définis, tels que des visages, des véhicules ou des personnes. Il a été largement utilisé dans des applications telles que la détection de piétons, la reconnaissance de gestes, la surveillance de vidéos, etc.

➤ **Les paramètres de HOG :**

• **Orientation (nombre de bins d'orientation) :**

L'orientation fait référence à la plage complète des angles possibles pour les gradients. Le nombre de bins d'orientation détermine la précision de la représentation des orientations.

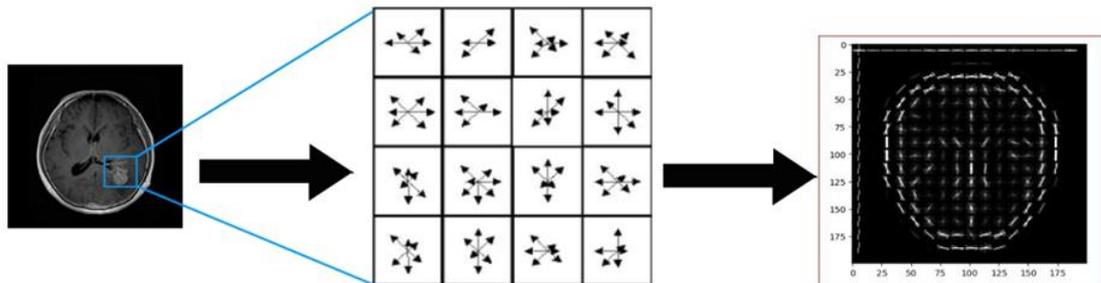


FIGURE 10 LE PARAMETRE DE L'ORIENTATION DU HOG [29]

• **Cellules par bloc (cells per block) :**

Les cellules sont les régions carrées sur lesquelles les gradients sont calculés. Les cellules par bloc font référence au nombre de cellules incluses dans un bloc rectangulaire. Un bloc est une région rectangulaire composée de plusieurs cellules adjacentes. Les cellules par bloc peuvent être réglées en fonction de la taille souhaitée pour les caractéristiques extraites. Une valeur courante pour les cellules par bloc est de 4x4, ce qui signifie que chaque bloc est composé de quatre cellules[29].

• **Pixels par cellule :**

Les pixels par cellule font référence à la taille de la région carrée sur laquelle les gradients sont calculés. Une valeur typique pour les pixels par cellule est de 8x8, ce qui signifie que les gradients sont calculés sur une région de 8x8 pixels.[29]

Ces paramètres déterminent la granularité de la représentation des caractéristiques extraites par HOG. Ils peuvent être ajustés en fonction des spécificités de l'application et des objets à détecter pour obtenir de bons résultats en termes de précision et de performance.

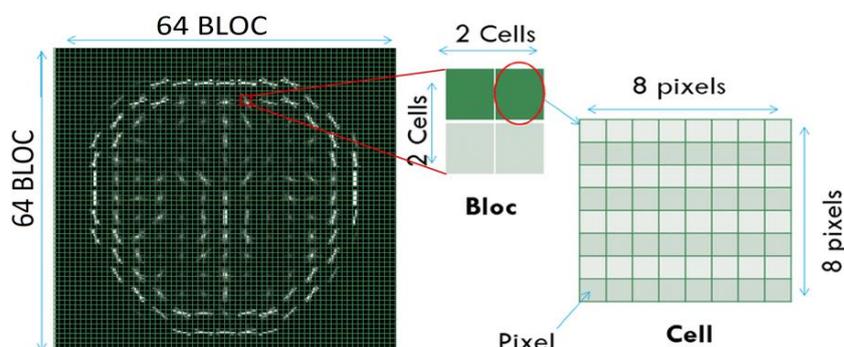


FIGURE 11 LA DISTRIBUTION DES PIXELS DANS UNE CELLULE D'UN BLOC [29]

b) GIST (Global Image Structure) :

GIST est un extracteur de caractéristiques utilisé dans le domaine de la vision par ordinateur pour représenter l'information de structure globale d'une image. Il vise à capturer les propriétés holistiques et intuitives de l'image plutôt que des détails spécifiques tels que les textures ou les contours.

Le GIST extracteur est basé sur l'hypothèse selon laquelle les humains sont capables d'avoir une perception rapide et approximative de la scène visuelle en se concentrant sur les informations structurelles globales plutôt que sur les détails fins. L'idée principale derrière GIST est de modéliser cette perception globale pour obtenir une représentation compacte et informative de l'image.(Choutri,Mehdi2022)

Voici les principales étapes de l'extraction de caractéristiques GIST :

- **Prétraitement de l'image** : L'image est généralement convertie en niveaux de gris pour simplifier le traitement. Des opérations de normalisation peuvent également être appliquées pour égaliser l'éclairage et réduire les variations d'intensité.
- **Découpage de l'image en plusieurs blocs** : L'image est divisée en plusieurs blocs de taille fixe. La taille et le nombre de blocs dépendent du paramétrage spécifique de l'algorithme GIST.
- **Calcul des histogrammes d'orientation** : Pour chaque bloc, des orientations sont calculées en utilisant des filtres de Gabor, qui sont des filtres adaptés pour détecter les caractéristiques de texture. Les réponses aux filtres de Gabor sont utilisées pour construire un histogramme d'orientation qui représente la distribution des orientations dans le bloc.
- **Agrégation des histogrammes** : Les histogrammes d'orientation des différents blocs sont agrégés pour former une représentation globale de l'image. Cela peut être fait en concaténant les histogrammes ou en utilisant des méthodes plus sophistiquées telles que la moyenne ou la max-pooling.
- **Normalisation** : Pour rendre la représentation GIST invariante à l'échelle et aux variations d'éclairage, une normalisation peut être appliquée aux valeurs des histogrammes.

Le résultat final est un vecteur de caractéristiques GIST qui représente la structure globale de l'image. Ce vecteur peut être utilisé comme entrée pour des tâches de classification, de recherche d'images similaires, de détection d'objets ou d'autres tâches de vision par ordinateur.

L'extraction de caractéristiques GIST est souvent utilisée dans des scénarios où l'analyse de la structure globale de l'image est pertinente, tels que la reconnaissance de scènes, la navigation robotique, la recherche d'images et la compréhension de l'environnement visuel.

3-2 Extracteurs de plus haut niveau :

Les méthodes d'extraction de plus haut niveau tiennent compte des formes et des structures dans l'image, des relations spatiales entre les pixels ou ces structures. Les propriétés les plus recherchées dans ces extracteurs sont, outre leur pouvoir descriptif, l'invariance et la robustesse à différentes transformations pouvant affecter l'image . Ainsi, la description obtenue demeure relativement inchangée face à ces transformations pouvant plus ou moins affecter des contenus identiques ou similaires dans les images. On retrouve couramment les invariances au changement d'échelle, de perspective, aux transformations affines comme la translation, la rotation et la robustesse au changement de luminosité ou de contraste. On retrouve particulièrement les méthodes : Scale Invariant Feature Transform, et d'autre basées sur des réseaux de neurones à convolution profonde(DCNN), Par rapport aux premières catégories, les méthodes de haut niveau peuvent effectuer une classification avec la plus grande précision.[31]

Dans notre travail, Un descripteurs de caractéristiques de plus haut niveau (ResNet-50) est utilisé .

c) ResNet-50 :

Contrairement aux architectures de réseau séquentielles traditionnelles telles que AlexNet, OverFeat et VGG, ResNet est plutôt une forme d '« architecture exotique» qui s'appuie sur des modules de micro-architecture (également appelés «architectures réseau dans réseau»).Une collection de blocs de construction de micro- architecture (avec des couches CONV, POOL, etc.) mène à la macro-architecture (le réseau final lui-même).[20]

Voici les principales caractéristiques de ResNet-50 :

- **Architecture résiduelle** : ResNet-50 utilise une architecture résiduelle qui permet de résoudre le problème du vanishing gradient (gradient qui tend à disparaître) lors de l'entraînement de réseaux profonds. L'idée clé est d'introduire des connexions résiduelles, également appelées skip connections, qui permettent de propager l'information directement à travers les différentes couches du réseau. Ces connexions permettent de résoudre le problème du dégradé de la performance en facilitant le flux du gradient et en permettant un apprentissage plus profond.
- **Profondeur du réseau** : ResNet-50 est un réseau relativement profond, composé de 50 couches. Il est constitué de plusieurs blocs de convolution et de pooling, suivis de couches de convolution résiduelle. Ces blocs permettent d'apprendre des caractéristiques de plus en plus complexes à mesure que l'information se propage dans le réseau.
- **Couches de convolution** : ResNet-50 utilise des couches de convolution pour extraire les caractéristiques des images en appliquant des filtres à travers les différentes couches du réseau. Ces couches de convolution sont capables de capturer des motifs visuels de plus en plus complexes à mesure que l'information se propage dans le réseau.
- **Utilisation de blocs résiduels** : ResNet-50 utilise des blocs résiduels, également appelés blocs ResNet, qui sont composés de plusieurs couches de convolution avec des

connexions résiduelles. Ces blocs permettent d'apprendre des représentations de plus haut niveau en conservant les informations résiduelles de l'entrée d'un bloc à sa sortie.

- **Classification à l'aide de couches fully-connected** : Après les couches de convolution, ResNet-50 utilise des couches fully-connected pour effectuer la classification finale. Ces couches prennent les caractéristiques extraites par les couches de convolution et les transforment en une probabilité de classe pour les tâches de classification.

ResNet-50 a été pré-entraîné sur de grands ensembles de données, tels qu'ImageNet, qui contiennent des millions d'images réparties sur différentes catégories. En raison de sa profondeur et de sa capacité à capturer des caractéristiques complexes, ResNet-50 est souvent utilisé comme modèle de base pour de nombreuses tâches de vision par ordinateur, notamment la classification d'images, la détection d'objets et la segmentation sémantique. Il a démontré des performances remarquables et est couramment utilisé dans les travaux de recherche et les applications pratiques.

4- Approches de classification :

Les méthodes de classification ont pour objet d'identifier la classe d'appartenance d'objets définis par leur description. Il existe deux types d'approches de classification, une approche non-supervisée et une approche supervisée. La différence entre les deux situations est la connaissance des classes. Dans le cas non-supervisée, on fait une recherche à l'aveugle alors qu'en supervisée, on a un échantillon d'apprentissage. En supervisé on connaît un indicateur de performance : le taux de mal classé, mais en non-supervisé, difficile de valider les résultats

4-1 Classification Non-supervisée (Clustering) :

La classification non-supervisée, ou clustering est l'une des techniques fondamentales de l'extraction de données structurées ou non structurées. Elle est utilisée lorsque l'on possède des documents qui ne sont pas classés et dont on ne connaît pas de classification, autrement dit, dont les classes et leur nombre sont inconnus. Dans ce cas, l'apprentissage se ramène alors à cibler les groupes homogènes d'exemples existants dans les données, c'est-à-dire à identifier des groupes appelés cluster, et que les exemples les plus différents soient séparés dans des groupes distincts. A la fin du processus de classification non-supervisée, les documents doivent appartenir à l'une des classes générées par la classification. On distingue deux catégories de classifications non-supervisées : hiérarchiques et non-hiérarchique

- **Classification hiérarchique (CH)** Dans la classification hiérarchique (CH), les sous-ensembles créés sont emboîtés de manière hiérarchique les uns dans les autres. On distingue la CH descendante qui part de l'ensemble de tous les individus et les fractionne en un certain nombre de sous-ensembles, chaque sous-ensemble étant alors fractionné en un certain nombre de sous-ensembles, et ainsi de suite. Et la CH ascendante qui part des individus seuls que l'on regroupe en sous-ensembles, qui sont à leur tour regroupés, et ainsi de suite. Pour déterminer quelles classes on va fusionner, on utilise le critère d'agrégation. [21]

- **Classification non-hiérarchique** Dans la classification non-hiérarchique, les individus ne sont pas structurés de manière hiérarchique. Si chaque individu ne fait pas partie que d'un sous-ensemble, on parle de partition. Si chaque individu peut appartenir à plusieurs groupes, avec la probabilité P_i d'appartenir au groupe i , alors on parle de recouvrement.[21]

4-1-1 Algorithmes de classification non-supervisée :

a) Méthode k-means :

L'algorithme k-means, créé par MacQueen en 1976 est l'algorithme de Clustering le plus connu et le plus utilisé car il s'avère être très simple à mettre en œuvre et efficace. Il suit une procédure simple de classification d'un ensemble d'objets en un certain nombre k de clusters, k fixé à priori. Dans cette algorithme, chaque cluster est caractérisé par son centre qui se trouve être la moyenne des éléments composant le cluster.[21]

▪ Avantages et Inconvénients

Avantages :

- Simple et compréhensible ;
- Éléments affectés automatiquement aux classes.

Inconvénients :

- Nombre de classes fixé à l'avance ;
- Chaque élément doit être dans une classe ;
- Très sensible aux éléments marginaux (intérêt d'une normalisation).

b) Fuzzy C-means :

Fuzzy C-means (FCM) est une méthode de clustering qui permet à un objet de données d'appartenir à deux ou plusieurs clusters. Cette méthode dérivée de l'algorithme *c-means* identique à l'algorithme *k-means* décrit précédemment, elle a été développée par Dunn en 1973 et améliorée par Bezdek en 1981. [15]

c) Hierarchical clustering :

Le processus basique des méthodes hiérarchiques a été donné par Johnson, Ce type de clustering consiste à effectuer une suite de regroupements en Clusters de moins en moins fines en agrégeant à chaque étape les objets (simple élément) ou les groupes d'objets (un Cluster-partition-) les plus proches. Ce qui nous donne une arborescence de clusters. Cette approche utilise la mesure de similarité pour refléter l'homogénéité ou l'hétérogénéité des classes. [15]

Dans ce chapitre, nous nous intéressons plus particulièrement à l'étude de la classification supervisée.

4-2 Classification supervisée :

La classification supervisée consiste à inférer, à partir d'un échantillon de données classées, une procédure de classification. Selon Gérard Govaert : « la classification supervisée a pour objectif « d'apprendre » par l'exemple. Elle cherche à expliquer et à prédire l'appartenance de documents à des classes connues a priori ». [19]

Autrement dit l'objectif de la classification supervisée est principalement de définir des règles permettant de classer des objets dans des classes à partir de variables qualitatives ou quantitatives caractérisant ces objets. Ainsi, on dispose au départ d'un échantillon dit d'apprentissage dont le classement est connu. Cet échantillon est utilisé pour l'apprentissage de règles de classement.

4-2-1 Algorithmes de classification supervisée

a) Méthode de k plus proches voisins :

La méthode des plus proches voisins (noté parfois **k-PPV** ou **k-NN** pour **Nearest-Neighbor**) Il s'agit d'un algorithme d'apprentissage paresseux qui stocke toutes les instances correspondant aux données d'apprentissage dans un espace à n dimensions. Il s'agit d'un algorithme d'apprentissage paresseux car il ne se concentre pas sur la construction d'un modèle interne général, mais sur le stockage d'instances de données d'entraînement.

Le classement est calculé à partir d'un vote à la majorité simple des k plus proches voisins de chaque point. Il est supervisé et prend un tas de points étiquetés et les utilise pour étiqueter d'autres points. Pour étiqueter un nouveau point [20], il examine les points étiquetés les plus proches de ce nouveau point, également connus sous le nom de ses voisins les plus proches. Il fait voter ces voisins, donc quel que soit l'étiquette que la plupart des voisins ont est l'étiquette pour le nouveau point. Le « k » est le nombre de voisins qu'il vérifie.

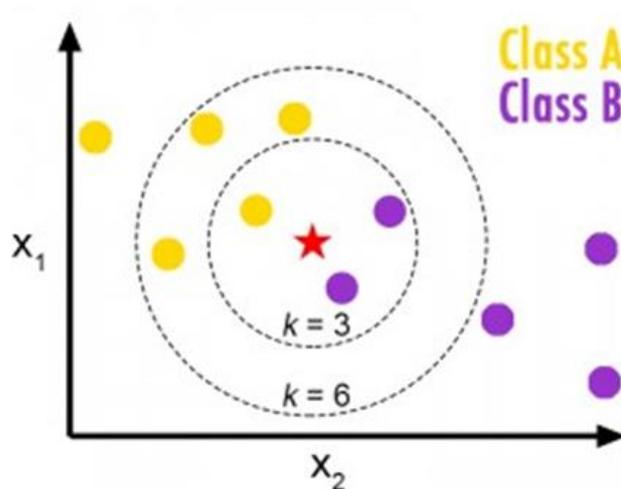


FIGURE 12 METHODE DE K PLUS PROCHES VOISINS . [20].

▪ **Avantages et Inconvénients :**

Avantages :

- Apprentissage rapide ;
- Méthode facile à comprendre ;
- Adapté aux domaines où chaque classe est représentée par plusieurs prototypes et où les frontières sont irrégulières (*Exemple* : reconnaissance de chiffre manuscrit ou d'images satellites).

Inconvénients :

- Prédiction lente car il faut revoir tous les exemples à chaque fois ;
- Méthode gourmande en place mémoire ;
- Sensible aux attributs non pertinents et corrélés.

b) Machine à vecteurs de support (SVM)

La machine à vecteurs de support (*Support Vector Machine, SVM*) est une technique de classification binaire par apprentissage supervisé. Elle fut introduite par Vladimir Vapnik en 1995. Elle repose sur l'existence d'un classificateur linéaire appelé hyperplan dans un espace approprié. Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe « au milieu » des points des deux classes d'exemples. L'hyperplan calculé permet ainsi de séparer l'espace en deux zones. Pour classer les nouveaux documents, on calcule dans quelle région de l'espace ils se situent et on leur attribue la classe correspondante.[21]

La SVM est basée sur l'utilisation de fonctions dites noyau (*kernel*) qui permet une séparation optimale des données. Pour deux classes de données, le but de SVM est de retrouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes

▪ **Notions de base**

Hyperplan : sépare les deux ensembles de points (de documents).

Vecteurs de support : points les plus proches qui déterminent l'hyperplan.

Marge : hyperplan dont la distance minimale aux exemples est maximale.

▪ **Principe**

Cette technique est une méthode de classification à deux classes qui tente de séparer les exemples positifs des exemples négatifs dans l'ensemble des exemples. Cela garantit une généralisation du principe car de nouveaux exemples pourront ne pas être trop similaires à ceux utilisés pour trouver l'hyperplan mais être situés d'un côté ou l'autre de la frontière. Les exemples utilisés lors de la recherche de l'hyperplan ne sont alors plus utiles et seuls ces

vecteurs supports sont utilisés pour classer un nouveau cas, ce qui peut être considéré comme un avantage pour cette méthode.[15]

▪ **Exemples d'applications :**

- Classification de données biologiques/ physique
- Classification de documents numériques
- Classification d'expressions faciales

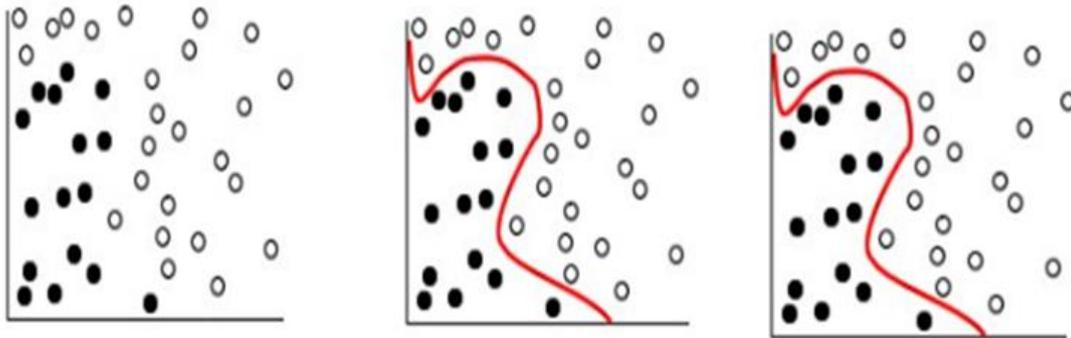


FIGURE 13 MACHINE A VECTEURS DE SUPPORT (SVM)

▪ **Avantages et Inconvénients**

Avantages :

- Minimisation de l'erreur empirique et structurelle ;
- Algorithmes optimisés ;
- Simple, peu de paramètres à régler ;
- Les SVM possèdent des fondements mathématiques solides ;
- Les exemples de test sont comparés juste avec les supports vecteurs et non pas avec tous les exemples d'apprentissage ;
- Décision rapide. La classification d'un nouvel exemple consiste à voir le signe de la fonction de décision $f(x)$.

Inconvénients :

- Les données dans des espaces de grande dimension sont souvent non linéairement séparables ;
- Classification binaire, d'où la nécessité d'utiliser l'approche un-contre-un ;
- Grande quantité d'exemples en entrées implique un calcul matriciel important;
- Temps de calcul élevé lors d'une régularisation des paramètres de la fonction noyau.

c) Réseau de neurones :

Un réseau de neurones artificiel est un ensemble d'algorithmes dont la conception est à l'origine très schématiquement inspirée du fonctionnement des neurones biologiques, et qui par la suite s'est rapproché des méthodes statistiques. Les approches neuronales furent les premières à être utilisées afin de réaliser un apprentissage de type statistique grâce à leur

capacité de classification et de généralisation. Un réseau de neurones artificiel est en général composé d'une succession de couches dont chacune prend ses entrées sur les sorties de la précédente. A chaque synapse est associé un poids synaptique, de sorte que les N_{i-1} sont multipliés par ce poids, puis additionnés par les neurones de niveau i , ce qui est équivalent à multiplier le vecteur d'entrée par une matrice de transformation. Nous pouvons ainsi voir un réseau de neurones artificiels comme un réseau ou graphe orienté dont les nœuds sont les neurones artificiels. Le but va être d'attribuer des poids synaptiques à chaque neurone afin d'obtenir le résultat voulu en sortie.[22]

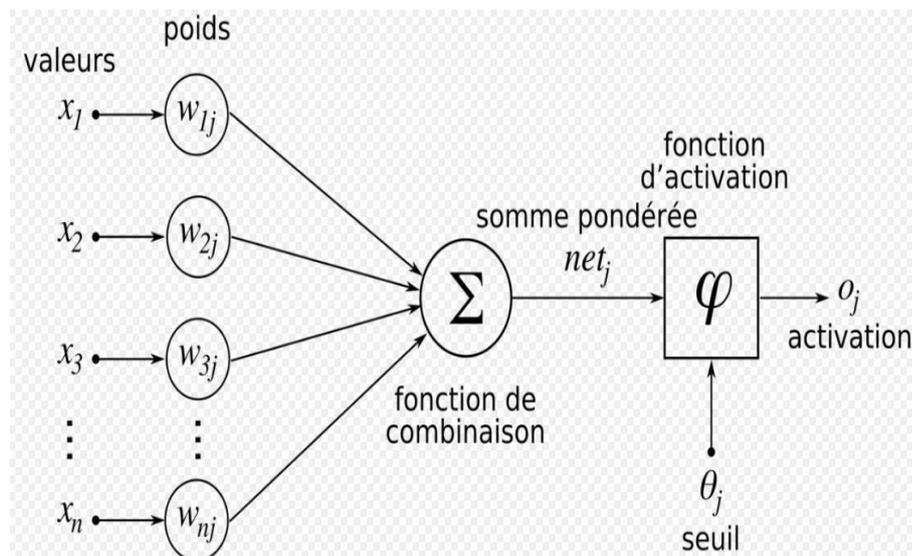


FIGURE 14 STRUCTURE D'UN NEURONE ARTIFICIEL. [22]

Le neurone calcule la somme de ses entrées puis cette valeur passe à travers la fonction d'activation pour produire sa sortie.

▪ **Avantages et Inconvénients :**

Avantages :

- Classification très précise (si bien paramétré) ;
- Résistance aux pannes (si un neurone ne fonctionne plus, le réseau ne se perturbe pas).

Inconvénients :

- La détermination de l'architecture du réseau est complexe ;
- Paramètres difficiles à interpréter (boite noire) ;
- Difficulté de paramétrage surtout pour le nombre de neurone dans la couche cachée.

d) Arbre de décision :

Les arbres de décision sont des modèles de ML supervisés, pouvant être utilisés pour la classification que pour la régression.

Un arbre de décision représente une fonction qui prend comme entrée un vecteur d'attributs et retourne une décision qui est une valeur unique. Les entrées et les sorties peuvent être discrètes ou continues.

Un arbre de décision prend ses décisions en exécutant une séquence de test, chaque nœud interne de l'arbre correspond à un test de la valeur d'un attribut et les branches qui sortent du nœud sont les valeurs possibles de l'attribut. La classe de la variable cible est alors déterminée par la feuille dans laquelle parvient l'observation à l'issue de la séquence de test.[15]

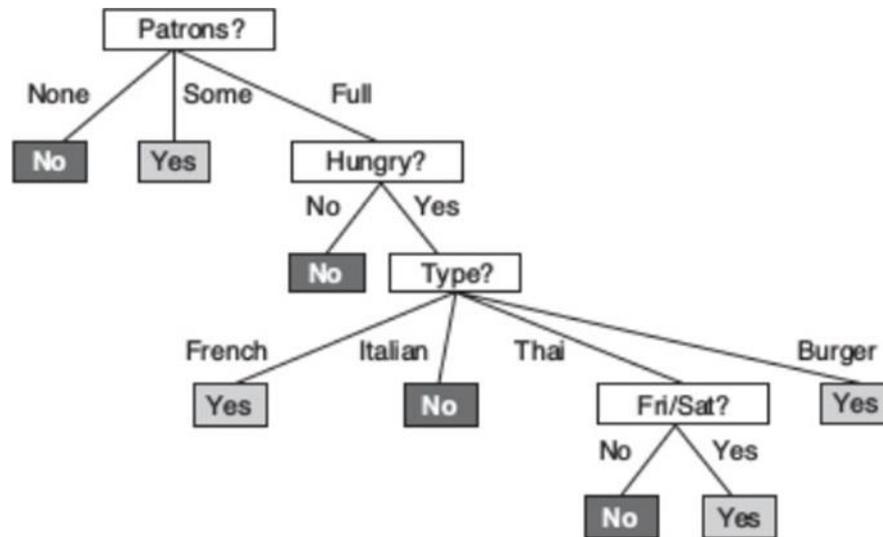


FIGURE 15 EXEMPLE D'ARBRE DE DECISION. [15]

▪ **Avantages et Inconvénients :**

Avantages :

- Leur capacité à travailler sur des données symboliques ;
- Leur grande capacité et efficacité à faire de la classification ;
- Leur facilité d'apprentissage et d'utilisation.

Inconvénients :

- Leur sensibilité au changement des données ;
- Une détection difficile des interactions entre les variables.

e) Deep Rule-Based (DRB) :

e.1) Définition :

L'algorithme de classification par Deep Rules-Based (DRB) est une approche qui combine les avantages des algorithmes de classification basés sur des règles et des réseaux de neurones profonds (deep learning). L'objectif principal du DRB est de capturer des règles complexes et hiérarchiques à partir des données d'entraînement et d'utiliser ces règles pour effectuer des prédictions sur de nouvelles données.

Un nouveau type de classification basé sur des règles multicouche est appliqué aux problèmes de classification d'images. L'approche DRB [23] est entièrement basée sur les données et entièrement automatique. Il est générique et peut être appliqué à divers problèmes de classification et de prédiction, en particulier.

L'approche DRB utilise un ensemble massivement parallèle de règles floues (Fuzzy Rules) d'ordre 0, comme moteur d'apprentissage, qui s'auto-organise une structure de système SI. . . ALORS FRB transparente et compréhensible par l'homme. Chaque règle floue SI. . . ALORS du système DRB se compose d'un (grand) nombre de prototypes, qui ne sont pas prédéterminés, mais sont identifiés par un processus d'apprentissage entièrement autonome, en ligne, non itératif et non paramétrique. Ces prototypes sont les échantillons de données (images) réels les plus représentatifs auxquels la densité de données obtient des maxima locaux (les images locales les plus typiques) ; ils sont utilisés pour former automatiquement des nuages de données (regroupements de données en cluster avec des propriétés similaires) en attirant les autres échantillons de données (images) vers eux . Le processus de formation du système DRB peut commencer «à partir de zéro» et, plus important encore, il peut commencer la classification à partir de la première image de chaque classe de la même manière que les humains, et est capable de s'auto-évoluer et de se mettre à jour de manière cohérente. structure et méta-paramètres avec des images d'apprentissage nouvellement observées, ce qui rend le classificateur proposé adapté aux applications en temps réel.[24]

e.2) Architecture générale du classificateur DRB :

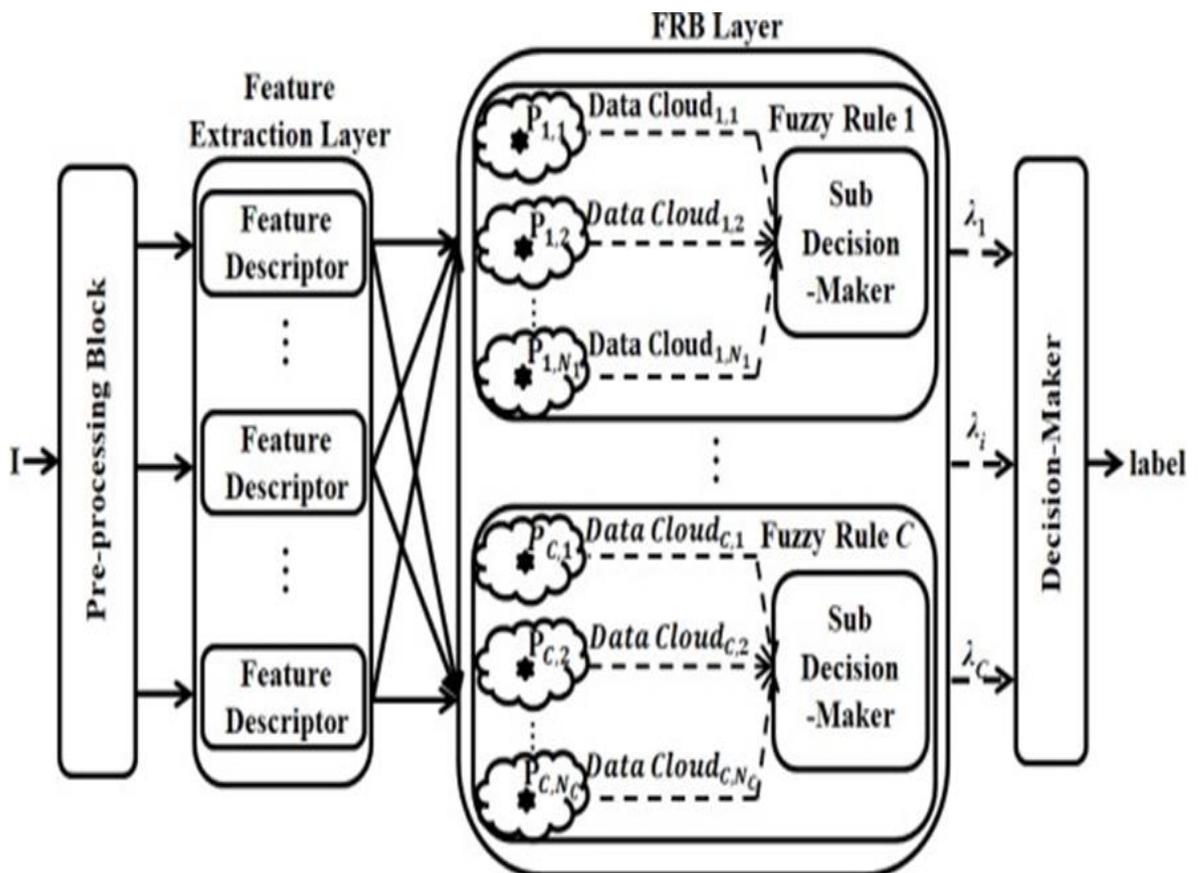


FIGURE 16 ARCHITECTURE GENERALE DU CLASSIFICATEUR DRB. [23]

- **Prétraitement des données** : Tout d'abord, les données sont prétraitées pour assurer leur qualité et leur cohérence. Cela peut inclure des étapes telles que l'élimination des valeurs manquantes, la normalisation des variables, la réduction de dimension, etc.
- **Extraction de règles** : L'étape suivante consiste à extraire des règles à partir des données d'entraînement. Différentes techniques peuvent être utilisées, telles que l'algorithme Apriori pour extraire des règles d'association ou l'algorithme de recherche de règles d'association fréquentes (Frequent Pattern Growth) pour extraire des règles fréquentes.
- **Construction d'un réseau de neurones** : Les règles extraites sont ensuite utilisées pour construire un réseau de neurones. Chaque règle peut être convertie en une unité de neurone, et la combinaison des règles forme l'architecture du réseau.
- **Entraînement du réseau de neurones** : Le réseau de neurones est entraîné sur les données d'entraînement en utilisant des techniques d'apprentissage supervisé. Les poids des connexions entre les neurones sont ajustés pour minimiser une fonction de perte prédéfinie.
- **Prédiction** : Une fois le réseau de neurones entraîné, il peut être utilisé pour effectuer des prédictions sur de nouvelles données. Les données de test sont alimentées dans le réseau, qui applique les règles apprises pour générer des prédictions.

e.3) Caractéristiques de DRB :

Les caractéristiques clés de l'algorithme de classification par Deep Rules-Based (DRB) sont les suivantes :[25]

- **Intégration de règles explicables** : Le DRB vise à extraire des règles à partir des données d'entraînement, ce qui permet d'obtenir une représentation explicite des motifs et des relations présents dans les données. Cela rend les résultats du modèle plus interprétables, ce qui est crucial dans certaines applications où la transparence est requise.
- **Apprentissage hiérarchique** : Le DRB est capable de capturer des règles complexes et hiérarchiques à partir des données. Il peut identifier des motifs à différents niveaux d'abstraction, ce qui lui permet de modéliser des relations non linéaires et d'obtenir une représentation plus riche des données.
- **Combinaison de techniques** : Le DRB combine les avantages des algorithmes de classification basés sur des règles et des réseaux de neurones profonds. Il utilise les règles extraites pour construire l'architecture d'un réseau de neurones, ce qui permet d'exploiter les capacités de généralisation et de modélisation non linéaire des réseaux de neurones tout en bénéficiant de l'interprétabilité des règles.
- **Adaptabilité aux données déséquilibrées** : Le DRB peut être particulièrement utile lorsque les ensembles de données sont déséquilibrés, c'est-à-dire lorsque certaines classes sont sous-représentées par rapport à d'autres. Les règles extraites peuvent aider à mieux représenter les classes minoritaires, ce qui améliore les performances de classification dans de tels scénarios.

- **Capacité à gérer des ensembles de données de petite taille :** Le DRB peut être efficace même avec des ensembles de données de petite taille, où les approches de deep learning traditionnelles peuvent rencontrer des problèmes de surajustement. En utilisant les règles extraites pour guider l'apprentissage, le DRB peut obtenir de bonnes performances même avec un nombre limité d'exemples d'entraînement.
- **Interprétation des prédictions :** Étant donné que le DRB utilise des règles explicables, les prédictions générées par le modèle peuvent être interprétées et justifiées en termes de règles spécifiques. Cela peut être utile pour comprendre les décisions du modèle, diagnostiquer les erreurs et établir la confiance dans les prédictions.

Il convient de noter que les caractéristiques du DRB peuvent varier en fonction de sa mise en œuvre spécifique et des techniques utilisées pour l'extraction des règles et la construction du réseau de neurones.

5- Optimisation :

5-1 Le processus d'optimisation :

Le processus d'optimisation dans la classification fait référence à la recherche des meilleurs paramètres ou configurations pour un modèle de classification donné. L'objectif est de trouver les valeurs optimales des paramètres qui permettent au modèle de fournir les meilleures performances de prédiction sur un ensemble de données.

Le processus d'optimisation se fait en trois étapes: analyse, synthèse et évaluation[27]. Tout d'abord, il convient d'analyser le problème et d'opérer un certain nombre de choix préalables :

- Variable du problème quels sont les paramètres intéressants à faire varier?
- Espace de recherche dans quelles limites faire varier ces paramètres?
- Fonction objective quels sont les objectifs à atteindre?
- Méthode d'optimisation quelle méthode choisir?

Une fois effectuée ces différents choix, la méthode choisie synthétise des solutions potentielles qui sont évaluées, puis éliminées jusqu'à obtention d'une solution acceptable. Si nécessaire, le problème peut être redéfini à partir des solutions déjà obtenues.

5-2 Le processus d'optimisation :

Le processus d'optimisation dans la classification fait référence à la recherche des meilleurs paramètres ou configurations pour un modèle de classification donné. L'objectif est de trouver les valeurs optimales des paramètres qui permettent au modèle de fournir les meilleures performances de prédiction sur un ensemble de données.

Le processus d'optimisation se fait en trois étapes: analyse, synthèse et évaluation[27]. Tout d'abord, il convient d'analyser le problème et d'opérer un certain nombre de choix préalables :

- Variable du problème quels sont les paramètres intéressants à faire varier?
- Espace de recherche dans quelles limites faire varier ces paramètres?
- Fonction objective quels sont les objectifs à atteindre?
- Méthode d'optimisation quelle méthode choisir?

Une fois effectuée ces différents choix, la méthode choisie synthétise des solutions potentielles qui sont évaluées, puis éliminées jusqu'à obtention d'une solution acceptable. Si nécessaire, le problème peut être redéfini à partir des solutions déjà obtenues.

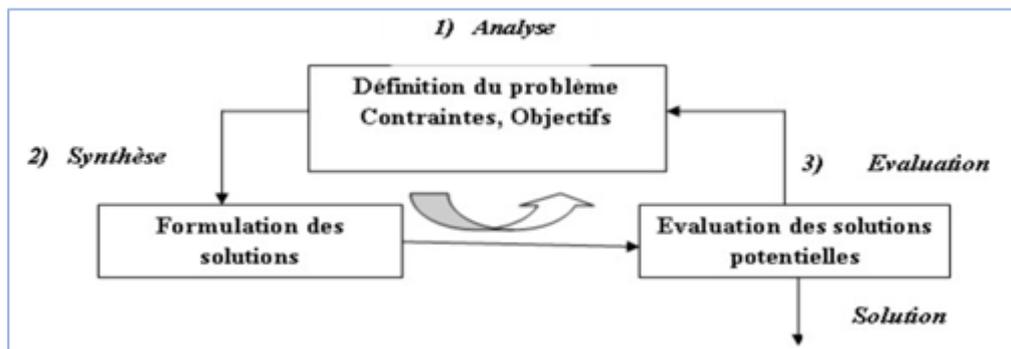


FIGURE 17 PROCESSUS D'OPTIMISATION. [27]

5-3 Les problèmes d'optimisation :

Il existe différents problèmes d'optimisation qui peuvent se poser lors de l'optimisation dans la classification. Voici quelques-uns des problèmes couramment rencontrés :

- **Minimum local** : L'algorithme d'optimisation peut converger vers un minimum local plutôt que vers le minimum global souhaité. Un minimum local est un point où la fonction d'objectif atteint un minimum, mais pas nécessairement le minimum absolu.
- **Sur-ajustement (overfitting)** : L'optimisation peut conduire à un modèle qui s'adapte trop étroitement aux données d'entraînement, ce qui peut entraîner une performance médiocre sur de nouvelles données. Cela se produit lorsque le modèle capture le bruit ou les spécificités uniques des données d'entraînement plutôt que les modèles généraux.
- **Sous-ajustement (underfitting)** : Le modèle peut être trop simplifié ou avoir une capacité limitée pour capturer les motifs et les relations dans les données. Cela conduit à une performance médiocre tant sur les données d'entraînement que sur les nouvelles données.
- **Dimensionnalité élevée** : Lorsque le nombre de paramètres du modèle est élevé, l'optimisation peut devenir plus difficile en raison de la complexité de l'espace de recherche. Il peut être difficile d'explorer efficacement l'espace des paramètres et de trouver les valeurs optimales.
- **Temps de calcul élevé** : Certains algorithmes d'optimisation peuvent être très coûteux en termes de temps de calcul, surtout lorsque l'espace de recherche est grand ou que les données sont volumineuses. Cela peut limiter la capacité d'explorer l'espace des paramètres de manière exhaustive.

- **Inter-prétabilité** : Lorsque des techniques d'optimisation avancées sont utilisées, comme l'apprentissage profond, les modèles résultants peuvent être très complexes et difficiles à interpréter. Cela peut rendre difficile la compréhension des facteurs qui contribuent aux prédictions du modèle.
- **Dépendance aux valeurs initiales** : Certains algorithmes d'optimisation peuvent être sensibles aux valeurs initiales des paramètres. Des valeurs initiales inappropriées peuvent conduire à une convergence vers des solutions sous-optimales.

Il est important de prendre en compte ces problèmes lors de l'optimisation d'un modèle de classification et d'utiliser des techniques appropriées pour les atténuer, comme la régularisation, la validation croisée, l'ajustement des hyperparamètres et la sélection du modèle.

5-4 Les algorithmes d'optimisation :

➤ Les algorithmes génétiques :

Les algorithmes génétiques (GA, Genetic Algorithms), également connus sous le nom d'algorithmes évolutionnaires, sont une technique d'optimisation inspirée de la théorie de l'évolution biologique. Les GA sont utilisés pour résoudre des problèmes d'optimisation complexes où une recherche exhaustive de l'espace des solutions est difficile ou impossible.

Les GA sont basées sur les phases suivantes :

- Initialisation : Une population initiale de N chromosomes est tirée aléatoirement.
- Évaluation : Chaque chromosome est décodé, puis évalué.
- Sélection : Création d'une nouvelle population de N chromosomes par l'utilisation d'une méthode désélection appropriée.
- Reproduction : Possibilité de croisement et mutation au sein de la nouvelle population. Retour à la phase d'évaluation jusqu'à l'arrêt de l'algorithme.

Avantages :

- Prise en compte de plusieurs critères simultanément.
- Parvient à trouver de bonnes solutions sur des problèmes très complexes.
- Combinaison de l'exploration de l'espace de recherche, et de l'exploitation des meilleures solutions disponibles à un moment donné.

Inconvénients:

- Coût d'exécution important par rapport à d'autres méta-heuristiques.
- L'ajustement d'un algorithme génétique est délicat.

Exemple d'application :

Cet exemple est dédié à la minimisation d'une fonction mathématique nommée DEJONG5FCN. Le code source de cette fonction en Matlab est présenté sous dessous. Les paramètres de GA utilisés sont : la population 50, la Génération maximale 100. La fonction objective à optimiser est composée de deux variables. La représentation en 3D de la fonction à optimiser est présentée dans la Figure (18).

```
%DEJONG5FCN Compute DeJongs fifth function.  
  
a =[-32, -16, 0, 16, 32, -32, -16, 0, 16, 32, -32, -16,  
0, 16, 32 -32, -16, 0, 16, 32, -32, -16, 0, 16, 32; 0,  
-32, -32, -32, -32, -32, -16, -16, -16, -16, -16, 0,  
0, 0, 0, 16, 16, 16, 16, 16, 32, 32, 32, 32, 32...];  
  
if nargin == 0  
    plotobjective(@dejong5fcn, 65.536 * [-1 1; -1 1]);  
    return  
end  
  
scores = zeros(size(xin,1),1);  
for i = 1:size(xin,1)  
    p = xin(i,:);  
    p = max(-65.536, min(65.536, p));  
    k = 0.002;  
    for j = 1:25  
        k = k + 1 / (j + (p(1) - a(1,j))^6 + (p(2) - a(2,j))^6);  
    end  
    scores(i) = 1/k;  
end
```

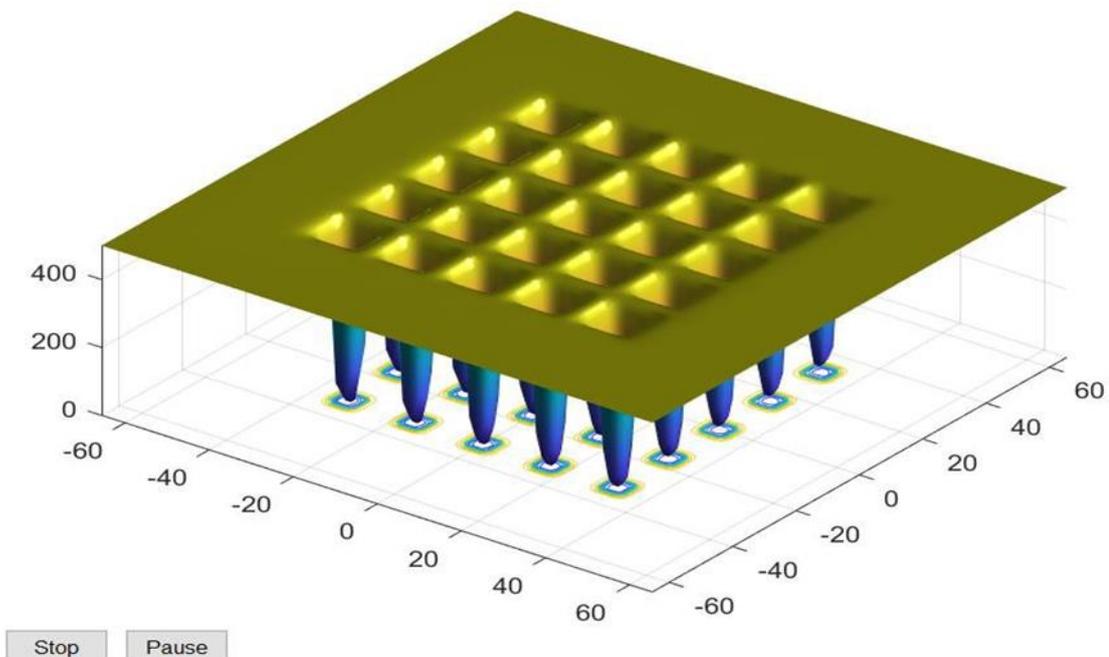


FIGURE 18 : REPRESENTATION EN 3D DE LA FONCTION A OPTIMISER.[33]

➤ Les algorithmes du loup gris :

L'algorithmes du loup gris (Grey Wolf Optimization en anglais) (*GWO*) est l'un des algorithmes d'optimisation inspirés biologiquement, proposé par Mirjalili et al en 1914[28]. C'est une méthode d'optimisation inspirée du comportement social et de la hiérarchie de meute des loups gris. Ces algorithmes imitent les interactions entre les loups pour résoudre des problèmes d'optimisation complexes.

L'algorithme GWO présente plusieurs avantages que nous résumons comme suit :

- Le groupe des loups gris comprend de 5 à 12 individus. Ils ont une hiérarchie sociale stricte à 4 niveaux.
- Les loups gris sont classés hiérarchiquement en alpha (α), bêta (β), delta (δ) et oméga (ω) dans la meute classés dans le groupe des loups .
- Les leaders sont un mâle et une femelle, appelés alphas. Le leader alpha décide de l'action des meutes. L'ensemble du pack reconnaît l'alpha en maintenant ses queues. Seul l'alpha le chef dans le pack. L'alpha domine en tant que bon administrateur. Ce n'est peut-être pas le membre le plus fort de la meute, mais il a l'expérience.
- Le deuxième niveau consiste en bêta. Sont des loups subordonnés, qui aident l'alpha dans la prise de décision ou d'autres activités de groupe. Le bêta-loup peut être un mâle ou une femelle et est le meilleur successeur au cas où l'un des loups alpha cesserait. Il joue le rôle d'un conseiller de l'alpha et du discipliner pour la meute.
- Les omégas sont au troisième niveau. Ils doivent se soumettre à tous les autres loups dominants. Ce sont les derniers loups autorisés à manger. Bien que les omégas soient moins bien classés, la meute entière est confrontée à des combats internes et à des problèmes sans eux. Ils aident à satisfaire l'ensemble et à maintenir la structure de dominance. Si le loup n'est pas de type alpha, bêta ou oméga, il est appelé sous (subordonné) ou delta dans certaines références. Les loups du delta doivent se soumettre aux alphas et aux betas, mais ils dominent les omégas .

L'algorithme GWO présente de nombreux avantages qui peuvent être résumés comme suit :

- C'est un algorithme à dérivée libre
- Il a des paramètres inférieurs à ajuster
- Il a des opérateurs inférieurs par rapport à d'autres algorithmes évolutifs.
- Il peut conserver des informations sur l'espace de recherche avec des itérations et enregistrera la meilleure solution obtenue
- il est simple et facile à mettre en œuvre.

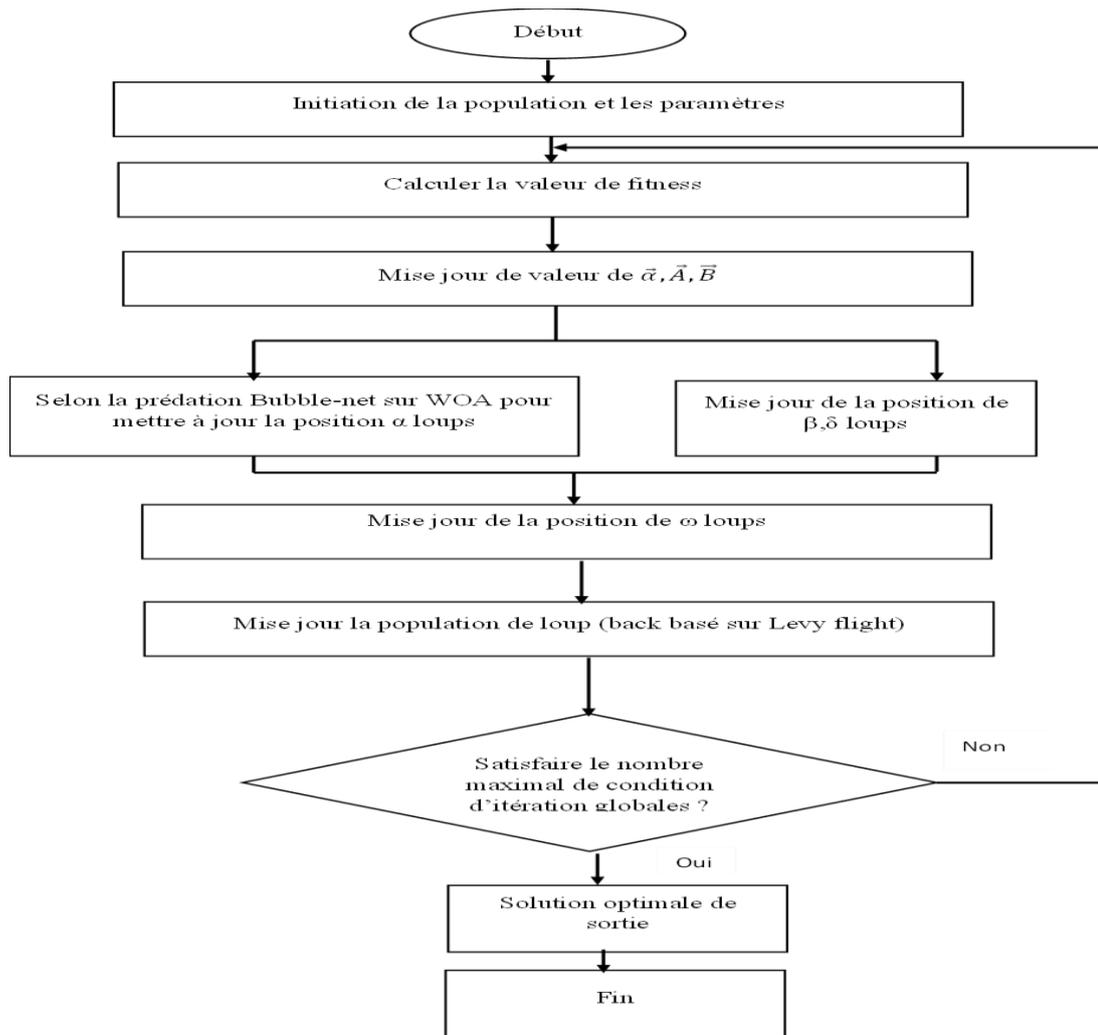


FIGURE 19 : ORGANIGRAMME DE FONCTION DE LA METHODE GWO

➤ Les algorithmes par essaim de particules:

L’algorithme par Essaim de Particules (OEP), ou « Particle Swarm Optimization (PSO) en anglais» a été développé par les inventeurs, Jim Kennedy et Russell Eberhart en 1995[28]. Est une méthode d’optimisation inspirée du comportement des essaims d’oiseaux ou de bancs de poissons. Ces algorithmes cherchent à trouver la meilleure solution possible en explorant et en exploitant l’espace des solutions de manière collaborative.

Le principe de l’algorithme PSO s’est éloigné du comportement (trop complexe) des animaux, pour ne conserver qu’une modélisation basée sur des agents simples, appelés particules. Chaque particule est considérée comme une solution du problème, où elle possède une position et une vitesse. De plus, chaque particule possède une mémoire lui permettant de se souvenir de sa meilleure performance et de la meilleure performance atteinte par les particules « voisines » ; chaque particule dispose en effet d’un groupe ’informatrices, historiquement appelé son voisinage[28].

Avantage :

- Peut converger rapidement vers des bonnes solutions
- Implémentation simple, avec peu de paramètres
- Possibilité de résoudre beaucoup de différents problèmes

Inconvénients:

- PSO peut prématurément converger dans deux cas :
 - L'essaim s'est attiré vers un optimum local.
 - Mauvaise initialisation de la population.
- Problème de dépendance qui résulte habituellement des arrangements de paramètres de chaque algorithme. Ainsi, en utilisant de différents arrangements de paramètres pour un seul algorithme de recherche stochastique on obtient des résultats de haute performance. Mais en général, l'arrangement d'un paramètre ne peut être appliqué à tous les problèmes[31].

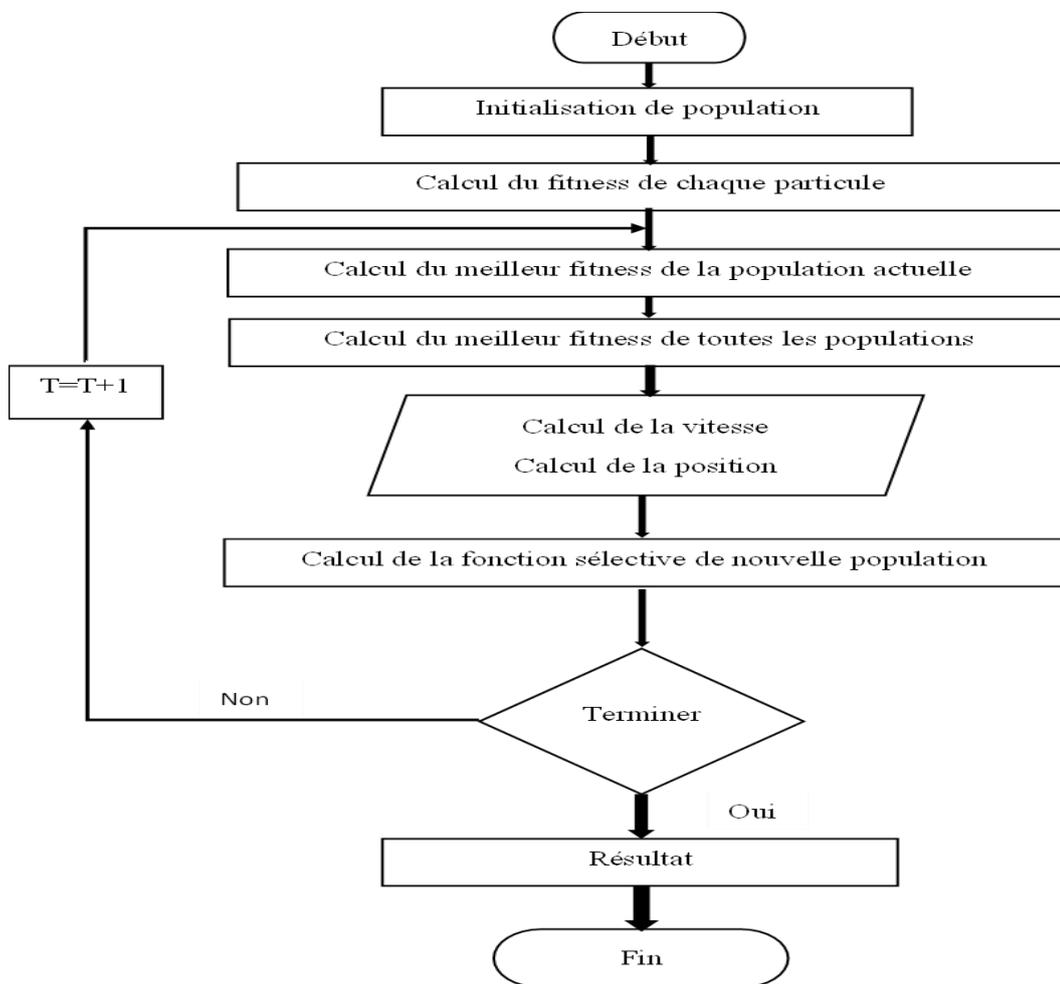


FIGURE 20 : SCHÉMA GÉNÉRAL D'UN PSO

5-5 Classification avec optimisation :

Voici un exemple simplifié de l'organigramme du processus de classification avec optimisation :

- **Collecte des données** : Rassembler un ensemble de données comprenant des exemples étiquetés, où chaque exemple est composé de caractéristiques (variables indépendantes) et d'une classe cible (variable dépendante).
- **Prétraitement des données** : Effectuer des étapes de prétraitement sur les données, telles que la normalisation, la réduction de dimension, le nettoyage des données, etc., afin de préparer les données pour la phase de classification.
- **Division des données** : Diviser l'ensemble de données en ensembles d'entraînement, de validation et de test. L'ensemble d'entraînement est utilisé pour entraîner le modèle, l'ensemble de validation est utilisé pour ajuster les hyperparamètres et l'ensemble de test est utilisé pour évaluer les performances finales du modèle.
- **Sélection du modèle** : Choisir un modèle de classification approprié pour le problème donné, tel que DRB, KNN, SVM. Le choix du modèle dépend des caractéristiques des données et des objectifs de classification.
- **Optimisation des hyperparamètres** : Utiliser une méthode d'optimisation, telle que GAO, GWO ou PSO, pour ajuster les hyperparamètres du modèle.
- **Entraînement du modèle** : Utiliser l'ensemble d'entraînement pour entraîner le modèle en utilisant les hyperparamètres optimisés. Le modèle apprend à partir des exemples étiquetés en ajustant ses paramètres internes pour mieux prédire les classes cibles.
- **Validation et évaluation** : Utiliser l'ensemble de validation pour évaluer les performances du modèle en termes de métriques de classification telles que la précision, le rappel, le score F1, etc. Cela permet d'ajuster davantage les hyperparamètres ou de prendre des décisions sur le modèle en fonction des performances obtenues.
- **Test final** : Évaluer les performances du modèle final en utilisant l'ensemble de test, qui représente de nouvelles données non vues par le modèle pendant l'entraînement et la validation. Cela donne une estimation de la capacité du modèle à généraliser et à bien se comporter sur de nouvelles données.
- **Utilisation du modèle** : Une fois le modèle finalisé, il peut être utilisé pour prédire les classes cibles des nouvelles données non étiquetées.

Il convient de noter que cet organigramme est une représentation générale du processus de classification avec optimisation, et que les détails spécifiques peuvent varier en fonction du problème et des techniques utilisées.

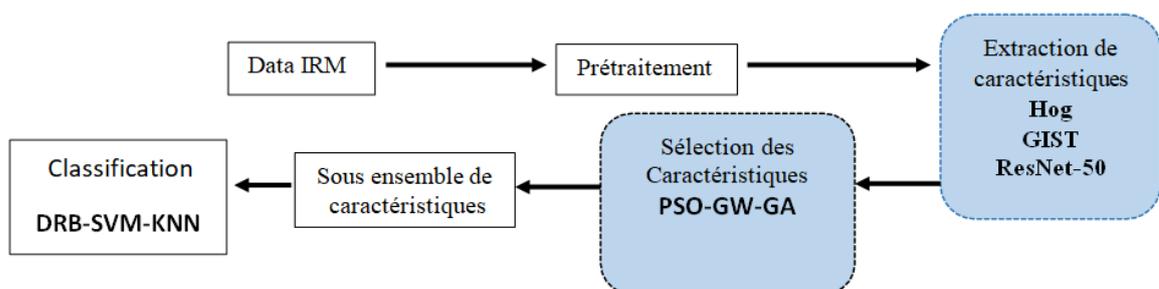


FIGURE 21 PROCESSUS DE CLASSIFICATION AVEC OPTIMISATION. [29]

6- Conclusion :

Dans ce chapitre nous nous sommes penchés sur le processus de classification. On a vu les différentes phases de classification : le prétraitement, l'extraction des caractéristiques et ses différents niveaux ensuite on a vu le classifieur DRB, son architecture et dérouler le processus de classification par ce dernier, et nous avons fini ce chapitre par les mesures de performance.

CHAPITRE 3 :

Analyse et Résultats

Chapitre 3 : Analyse et Résultats

1- Introduction :

Dans le chapitre précédent, nous avons examiné en détail la classification et les méthodes associées. Ce chapitre met en œuvre et compare plusieurs modèles basés sur DRB , SVM et KNN en termes de précision. Nous explorons l'optimisation des paramètres HOG,GIST et ResNet-50 avec PSO,GW et GA pour améliorer la précision. Les outils, modèles et résultats sont présentés. Notre approche se concentre sur l'implémentation et la performance des modèles, ainsi que sur l'optimisation des paramètres pour une meilleure précision. Aussi , le changement d'angle nous permet de pouvoir l'image et ce qu'elle contient plus clairement , ce qu'il nous donne de meilleurs résultats , comme on à remarquer une augmentation de l'accuracy .

2- Environnement de travail :

2-1 Ressources matériel :

Ce travail a été effectué sur un ordinateur personnel doté des caractéristiques suivantes :

- Processeur : Intel Core i5 de 8e génération
- Mémoire RAM : 16 Go
- Stockage : Un disque dur SSD de 256 Go
- Système d'exploitation : Windows 11 Profissionnel 64 bits

2-3 Langages de programmation :

➤ MATLAB :

MATLAB est une abréviation de Matrix LABoratory. Écrit à l'origine, en Fortran, par C. Moler à la fin des années 1970 . MATLAB était destiné à faciliter l'accès au logiciel matriciel développé dans les projets LINPACK et EISPACK. La version actuelle, écrite en C par the MathWorks Inc., existe en version professionnelle et en version étudiant. Sa disponibilité est assurée sur plusieurs platesformes : Sun, Bull, HP, IBM, compatibles PC (DOS, Unix ou Windows), Macintosh, iMac et plusieurs machines parallèles.

MATLAB est un environnement de programmation et de développement de logiciels utilisé principalement pour le calcul numérique, l'analyse de données et la visualisation. Le terme "MATLAB" est une contraction de "MATrix LABoratory", faisant référence à la capacité du logiciel à manipuler efficacement les opérations sur les matrices, qui constituent une structure de données fondamentale en mathématiques et en science de l'ingénieur.

MATLAB est développé par MathWorks et offre un langage de programmation de haut niveau qui permet d'effectuer facilement des opérations mathématiques complexes, de résoudre des problèmes d'algèbre linéaire, de réaliser des analyses statistiques, de tracer des graphiques et de créer des interfaces utilisateur. Il est largement utilisé dans divers domaines,

tels que les mathématiques, l'ingénierie, les sciences physiques, la finance et l'apprentissage automatique.

Voici quelques caractéristiques et fonctionnalités clés de MATLAB :

- **Manipulation de matrices** : MATLAB offre des fonctionnalités puissantes pour créer, manipuler et effectuer des opérations sur des matrices et des tableaux multidimensionnels.
- **Calcul numérique** : Il fournit une large gamme de fonctions mathématiques prédéfinies pour effectuer des calculs numériques avancés, y compris l'optimisation, l'algèbre linéaire, l'analyse numérique, l'intégration et la résolution d'équations différentielles.
- **Visualisation** : MATLAB permet de créer des graphiques et des visualisations de données de manière interactive. Il propose des fonctions pour tracer des graphiques en 2D et en 3D, des diagrammes, des images, des animations et des interfaces utilisateur graphiques (GUI).
- **Développement d'applications** : MATLAB permet de créer des applications et des interfaces utilisateur personnalisées en utilisant des outils tels que GUIDE (Graphical User Interface Development Environment).
- **Traitement du signal et de l'image** : MATLAB fournit des fonctionnalités spécialisées pour le traitement du signal, le traitement de l'image, la vision par ordinateur et le traitement du signal audio.
- **Apprentissage automatique** : MATLAB propose des outils et des boîtes à outils pour l'apprentissage automatique, y compris des algorithmes de classification, de régression, de regroupement et de réseaux de neurones.

MATLAB est largement utilisé dans la recherche universitaire, l'industrie et d'autres domaines où des calculs numériques avancés et des analyses de données sont nécessaires. Il offre également des intégrations avec d'autres langages de programmation tels que C, C++, Python, Java, et des outils complémentaires pour des domaines spécifiques comme Simulink pour la modélisation et la simulation de systèmes dynamiques.

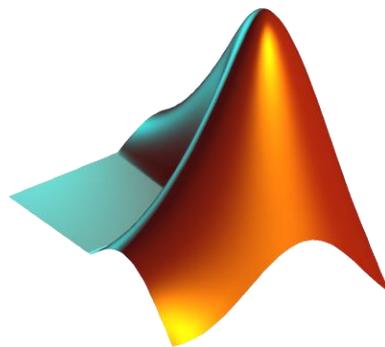


FIGURE 22 MATLAB LOGO

2-4 Les Modèles et architectures implémentés :

Dans cette section, nous allons présenter les bases de données utilisé et les architectures développées pour réaliser ce travail. nous nous sommes d'abord concentrés sur les quatre classificateurs **SVM**, **DRB**, **KNN** et nous avons utilisé trois réseaux de neurones à convolution pré-entraînés. Ces réseaux étaient **ResNet-50**, **HOG** et **GIST**. Ces réseaux CNN

pré-formés ont été utilisés pour extraire des caractéristiques d'image appropriées et les utiliser dans l'étape de classification. De plus Nous explorons l'optimisation des paramètres HOG,GIST et ResNet-50 avec **PSO,GW** et **GA** pour améliorer la précision .

2-5 Data-set utilisés :

Nous avons utilisé trois Data-sets pour réaliser ce travail, les deux premiers sont de type binaire « deux classe Yes \ No » contient 253, 3000 images respectivement. Le troisième Data-set est de type multi-classes il contient quatre classes dont le nombre des images est plus de 7000 images. Ils sont présentés ci-dessous.

➤ **Dataset1 , Dataset2 pour la classification binaire :**

L'ensemble de données a été téléchargé à partir du site **WEB Kaggle**. Il consiste en des examens IRM de deux classes :

- NO, pas de tumeur
- YES, pathologique

➤ **Dataset3 pour la classification multi classes :**

La même chose pour le troisième ensemble de données, il est téléchargé à partir du site **WEB Kaggle**. Il comporte les quatre classes:

- Glimora Tumor
- Meningioma tumor
- No tumor
- Pituitary tumor

3- Mesures de performance des algorithmes intelligents :

L'objectif principal de notre analyse repose sur l'évaluation des classifieurs et des algorithmes d'apprentissage. L'évaluation des performances d'un système de classification est essentielle, car elle peut être utilisée dans l'apprentissage pour optimiser les valeurs des paramètres du classifieur, elle permet aussi de montrer à quel point un algorithme ou un classifieur est performant et elle pourrait aussi faciliter la sélection des algorithmes ou des classifieurs les plus appropriés pour un problème particulier. Il est soutenu que l'évaluation présente la clé pour une estimation plus fiable de la qualité de généralisation d'un classifieur. Des questions importantes généralement soulevées sont de savoir comment évaluer correctement les algorithmes d'apprentissage et comment évaluer la performance prédictive des classifieurs générée par ces algorithmes [32].

Les mesures de performance nous fournissent un moyen d'évaluer notre système et de savoir à quel point il est bon. C'est la façon la plus simple de mesurer la «qualité» de notre modèle. Dans la littérature, plusieurs métriques ont été proposées, nous citons : **La matrice de confusion**.

3-1 La matrice de confusion :

Le Machine Learning consiste à nourrir un algorithme à l'aide de données afin qu'il apprenne par lui-même à effectuer une certaine tâche. Dans les problèmes de classification, il prédit des résultats que l'on doit comparer à la réalité pour mesurer son degré de performance. On utilise généralement **la matrice de confusion**, appelée aussi **tableau de contingence**. Elle mettra non seulement en valeur les prédictions correctes et incorrectes mais nous donnera surtout un indice sur le type d'erreurs commises. Pour calculer une matrice de confusion, on a besoin **d'un ensemble de données de test et un autre de validation qui contient les valeurs des résultats obtenus**.

Chaque colonne du tableau contient une classe prédite par l'algorithme et les lignes des classes réelles.

On classe les résultats en 4 catégories :

- **True Positive (TP)** : la prédiction et la valeur réelle sont positives.

Exemple : Une personne malade et prévu malade.

- **True Negative (TN)** : la prédiction et la valeur réelle sont négatives.

Exemple : Une personne saine et prévu saine.

- **False Positive (FP)** : la prédiction est positive alors que la valeur réelle est négative.

Exemple : Une personne saine et prévu malade.

- **False Negative (FN)** : la prédiction est négative alors que la valeur réelle est négative.

Exemple : Une personne malade et prévu saine.

TABLEAU 1 MATRICE DE CONFUSION

		Valeurs Prédites	
		0	1
Valeurs Réelles	0	TN	FP
	1	FN	TP

Les résultats présentés dans la matrice de confusion nous permettent de calculer d'autres métriques qui ont plus de sens en termes de meilleure interprétation des performances des algorithmes intelligents tels que :

- **Exactitude (en anglais Accuracy)** : C'est le taux de succès mesure combien de fois le modèle a fait la bonne prédiction. Voici l'équation utilisée pour calculer la précision du modèle :

$$\text{Accuracy (Exactitude)} = \frac{\text{Prédictions Correctes}}{\text{Nombre Total d'Exemples}} = \frac{TP + TN}{TP + FP + TN + FN}$$

- **Taux d'erreur (Error rate)** : C'est l'erreur globale de la classification. Le taux d'erreur est calculé comme suit :

$$(E) = 1 - \text{Taux de succès (Acc)}$$

- **Rappel (Sensivicity ou Racall)** : Appelé aussi sensibilité c'est le pourcentage des instances positives correctement identifiées, c'est-à-dire parmi les patients malades, combien a le système correctement diagnostiqué comme malades? Le rappel est calculé comme suit :

$$\text{Sensibilité}(Sn) = \frac{TP}{TP + FN}$$

- **Précision(Précision)** :c'est l'opposé du rappel. Elle dit essentiellement : sur tous les patients prévus comme positifs, combien d'entre eux étaient réellement positifs (combien étaient réellement malades?) En d'autres termes, quel est le taux de faux positifs de ce modèle ?. La précision est calculée comme suit :

$$\text{Précision} = \frac{TP}{TP + FP}$$

- **Spécificité (Specificity)** :c'est le pourcentage d'instances négatives correctement identifiées et elle est calculée comme suit :

$$\text{Spécificité} = \frac{TN}{TN + FP}$$

- **Mesure F (F-measure ou F-Score)** : c'est la mesure qui combine la précision et le rappel et leur moyenne harmonique. Cette mesure prend des valeurs entre 0 et 1. Elle est utilisée quand on cherche une balance entre la précision et le rappel. Plus que la valeur de cette mesure s'approche de 1, plus que notre modèle fait une bonne décision, et le contraire est vrai, quand elle s'approche de zéro(Mifdal, 2019) . La mesure F est calculée par l'équation suivante :

$$F1 \text{ Score} = 2 * \frac{\text{Précision} * \text{Rappel}}{\text{Précision} + \text{Rappel}}$$

- **Moyenne géométrique (G-Mean)** : C'est la moyenne de la spécificité et le rappel qui mesure la performance de globale du classifieur. Elle est calculée comme suit :

$$G - \text{Mean} = \sqrt{Sp * Sn}$$

3-2 La surface sous la courbe ROC (AUC, Area Under the Curve) :

Est une mesure de performance couramment utilisée pour évaluer la capacité d'un modèle de classification binaire à distinguer les échantillons positifs des échantillons négatifs. Elle représente la surface entre la courbe ROC et l'axe diagonal du graphique, qui correspond à une performance aléatoire.

L'AUC mesure la capacité du modèle à classer correctement les échantillons positifs avant les échantillons négatifs, quelle que soit la valeur de seuil de classification. Elle varie de 0 à 1, où une valeur de 1 indique une performance de classification parfaite et une valeur de 0,5 indique une performance aléatoire. [26]

En d'autres termes, l'AUC représente la probabilité que le modèle classe correctement un échantillon positif au hasard par rapport à un échantillon négatif au hasard. Plus l'AUC est élevée, meilleure est la capacité du modèle à discriminer les échantillons positifs des échantillons négatifs.

Les règles pour évaluer la performance des classificateurs en utilisant la courbe ROC et l'AUC sont les suivantes :

- Plus l'AUC est proche de 1, meilleure est la performance du classificateur.
- Si l'AUC est supérieure à 0,5, le classificateur est meilleur que le hasard, c'est-à-dire qu'il est capable de distinguer les échantillons positifs des échantillons négatifs.
- Si l'AUC est égale à 0,5, le classificateur est équivalent à une décision aléatoire.
- Si l'AUC est inférieure à 0,5, le classificateur est pire que le hasard et doit être évité.

La position de la courbe ROC sur le graphique et sa forme peuvent fournir des informations supplémentaires sur la performance du classificateur. Une courbe qui se rapproche du coin supérieur gauche du graphique indique une meilleure performance, tandis qu'une courbe qui se rapproche de la ligne diagonale de 45 degrés indique une performance aléatoire.

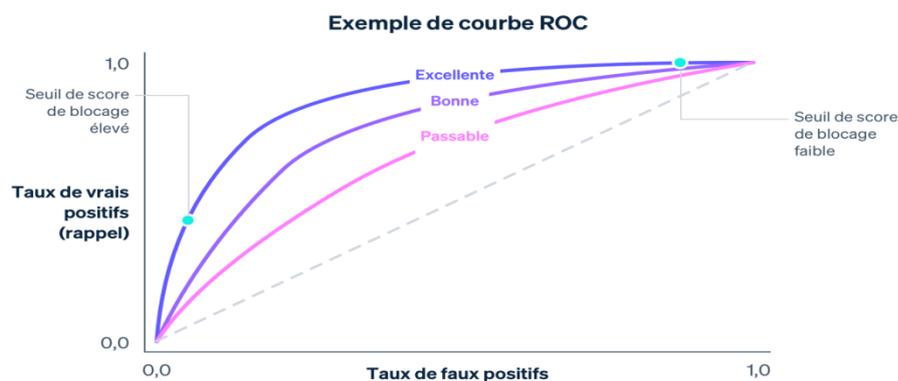


FIGURE 23 EXEMPLE DE COURBE ROC (AUC).[30]

4- Les étapes d'implémentation :

4-1 Première étape : Classification sans optimisation (basique):

Dans cette étape nous allons présenter le processus de classification effectué pour classer les différentes images des différentes Data-sets par les classifieurs DRB, SVM et KNN.

➤ Modèle 1 : HOG avec les trois différents Classifieurs :

Le modèle Feature HOG (Histogram of Oriented Gradients) combiné aux classifieurs SVM (Support Vector Machine) et KNN (K-Nearest Neighbors) avec DRB (Dynamic Region of Interest) est une approche courante utilisée dans la détection d'objets et la reconnaissance de formes.

Le modèle commence par initialiser des listes vides pour stocker les données des images et les étiquettes correspondantes. Ensuite, il boucle sur chaque classe d'images de tumeurs cérébrales et lit les images à partir des répertoires correspondants. Pour chaque image .

Ensuite, les données et les étiquettes sont converties en tableaux NumPy. Le code analyse ensuite les classes uniques et leur répartition dans le jeu de données.

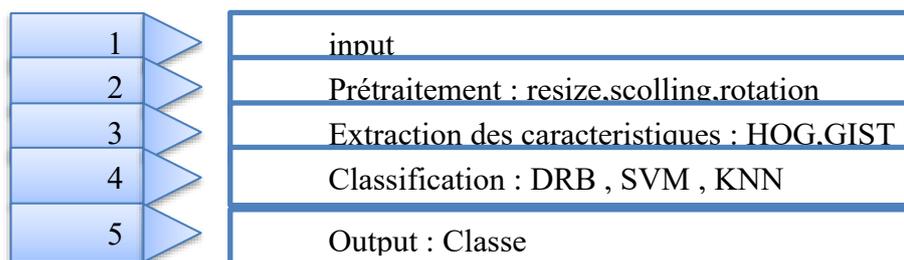
Les données d'image sont remodelées en un vecteur de caractéristiques unidimensionnel. Le jeu de données est divisé en ensembles d'apprentissage et de test en utilisant un pourcentage spécifié pour les tests.

Enfin, ce modèle calcule la spécificité (taux de vrais négatifs) et visualise la matrice de confusion ainsi que la courbe ROC (Receiver Operating Characteristic).

Le modèle effectue donc les étapes suivantes : prétraitement des images, extraction des caractéristiques HOG, division des données en ensembles d'apprentissage et de test, entraînement et évaluation de plusieurs modèles SVM , KNN et DRB , et affichage des performances et des visualisations associées.

Dans ce modèle, l'extracteur Hog est utilisé avec les trois classifieurs, la Figure 3.2 montre le processus de classification. L'étape de pré-traitement nécessite de retailler toutes les images en même taille 280×250 avant de passer à la normalisation. Ensuite on a fait l'extraction de caractéristiques, le vecteur résultat est utilisé pour la classification

Etapes :



➤ **Modèle 2 : GIST avec les trois différents Classifieurs :**

Le deuxième modèle c'est l'extracteur Gist qui nécessite d'ajuster toutes les images en taille 256x256 puis les normaliser pour l'étape d'extraction de caractéristiques ensuite la classification. Ce modèle a la même architecture que le modèle 1. (Figure 3.2).

➤ **Modèle 3 : ResNet-50 avec les trois différents Classifieurs :**

les caractéristiques d'image apprises à partir des images de formation ont été extraites et suivies par les quatre classificateurs . Dans cette architecture on utilise le réseau de neurones ResNet50 préformé comme extracteur de caractéristiques et pas comme classifieur. Les caractéristiques extraites seront utilisées pour entraîner et tester les modèles des différents classifieurs DRB , SVM et KNN. Notons que lors l'utilisation de ResNet50 les images seront [224x224], ensuite une augmentation des données sera faite par la conversion de toutes les images du niveau de gris en RGB (afin d'augmenter le data-set) pour des raisons de formation de ce dernier. ResNet50 se compose de nombreuses couches mais il est important de noter que toutes ces couches sont essentielles pour l'extraction des caractéristiques. La couche fortement connectée nommée « fc1000 » avant la couche de classification, est utilisée pour extraire les caractéristiques à l'aide de la méthode d'activation. Ces caractéristiques sont ensuite utilisées pour entraîner et tester les trois classifieurs.

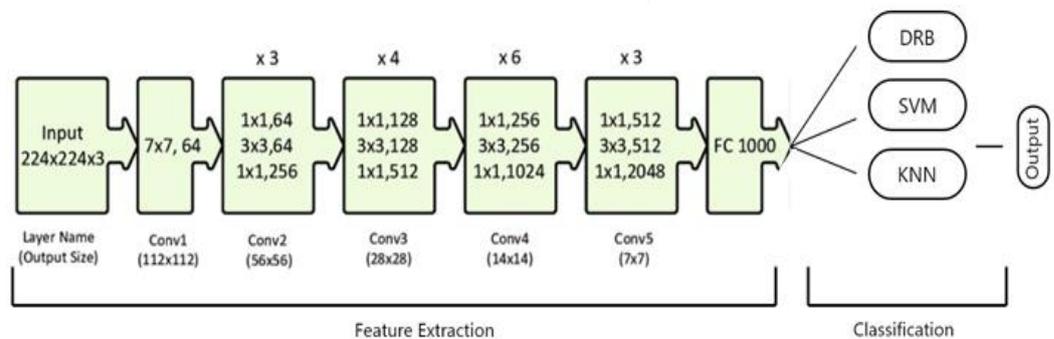


FIGURE 24 ARCHITECTURE DE MODELE RESNET-50.[30]

4-2 Résultats et discussions :

Plusieurs modèles sont implémentés, par la suite nous présentons une comparaison des résultats des différents modèles des trois Data-sets.

➤ Résultats du modèle 1 : HOG avec les trois différents Classifieurs :

a. l'évaluation de la performance :

TABLEAU 2 RESULTATS COMPARATIFS DE L'EXTRACTEUR HOG AVEC LES CLASSIFIEURS DRB,KNN,SVM

DataSets	Classifier	Accuracy	Sensitivity	Specifi- city	Precision	Recall	F_measure	G-mean
DataSet 1	SVM	0.8235	0.7500	0.8710	0.7895	0.7500	0.7692	0.8082
	KNN	0.8627	0.9500	0.8065	0.7600	0.9500	0.8444	0.8753
	DRB	0.8627	0.9500	0.8065	0.7600	0.9500	0.8444	0.8753
DataSet 2	SVM	0.9533	0.9467	0.9600	0.9595	0.9467	0.9530	0.9533
	KNN	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967
	DRB	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967
DataSet 3	SVM	0.8902	0.7967	0.9179	0.7422	0.7967	0.7685	0.8551
	KNN	0.9901	0.9967	0.9881	0.9614	0.9967	0.9787	0.9924
	DRB	0.9901	0.9967	0.9881	0.9614	0.9967	0.9787	0.9924

b. Courbes Roc :

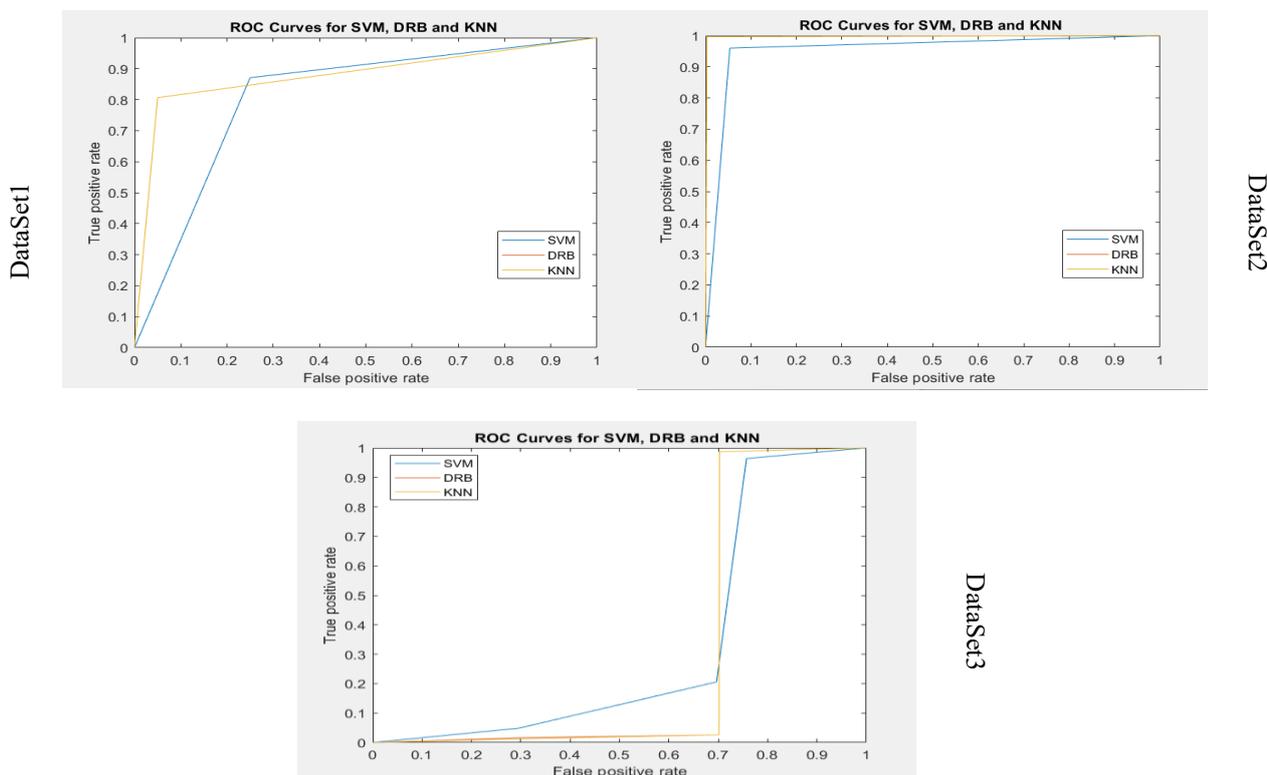


FIGURE 25 COURBES ROC DES CLASSIFIEURS DRB,KNN ET SVM ET L'EXTRACTEUR HOG AVEC LES TROIS DATA-SETS

c. les matrices de confusion :

▪ **SVM :**

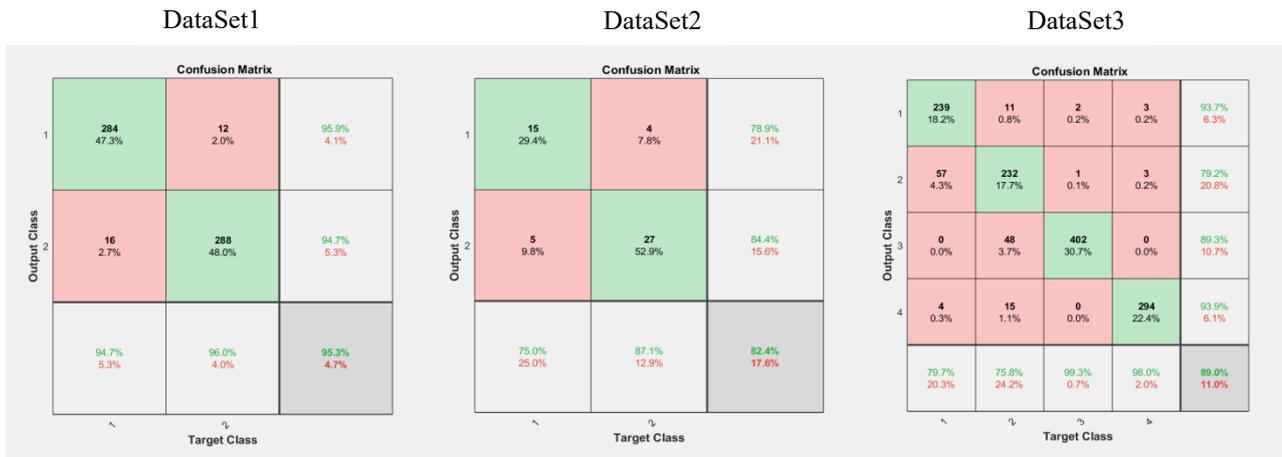


FIGURE 3.6 - MATRICES DE CONFUSION DU SVM ET L'EXTRACTEUR HOG AVEC LES TROIS DATA-SETS

▪ **KNN :**

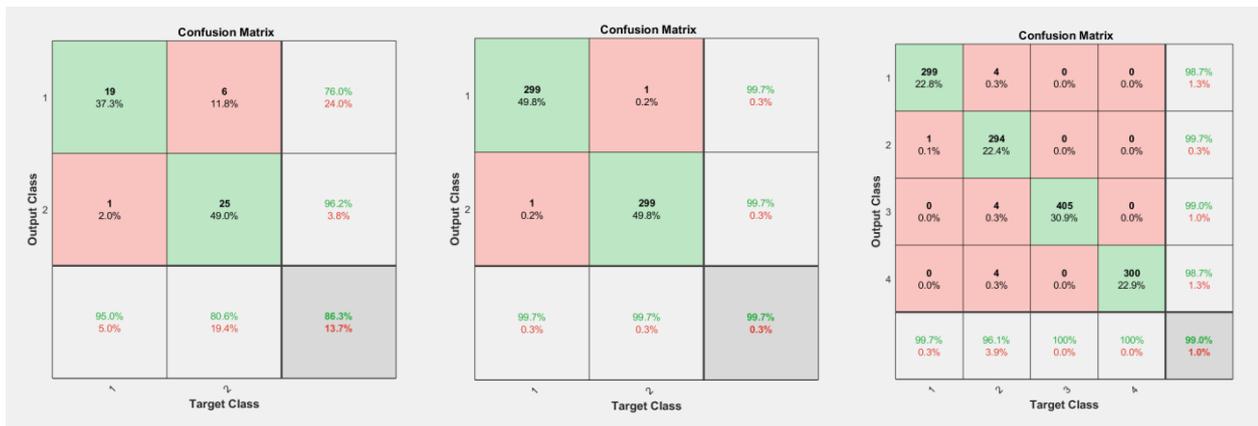


FIGURE 26 MATRICES DE CONFUSION DU KNN ET L'EXTRACTEUR HOG AVEC LES TROIS DATA-SETS

▪ **DRB :**

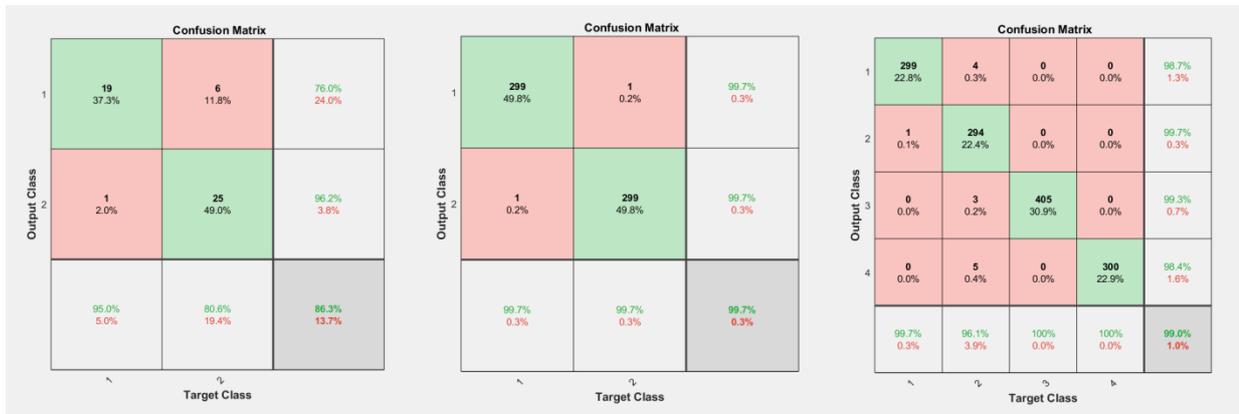


FIGURE 27 MATRICES DE CONFUSION DU DRB ET L'EXTRACTEUR HOG AVEC LES TROIS DATA-SETS

Discussion :

- Avec le modèle **HOG** le classifieur **DRB** donne une accuracy maximale de **86.27%**, **99.67%** et de **99.01%** avec les trois bases de données respectivement Dataset1, Dataset2 et Dataset3. Même chose pour le classifieur KNN.
- Avec le modèle **HOG** l'**AUC** du classifieur **DRB** est proche de **1** meilleure est la performance du classificateur. Et sa courbe se rapproche du **coin supérieur gauche** du graphique ca indique une **meilleure performance** .

➤ Résultats du modèle 2 : GIST avec les trois différents Classifieurs :

a. l'évaluation de la performance :

TABLEAU 3 RESULTATS COMPARATIFS DE L'EXTRACTEUR GIST AVEC LES CLASSIFIEURS DRB,KNN,SVM

DataSet	Classifieur	Accuracy	Sensitivity	Specifi- city	Precision	Recall	F_measure	Gmean
DataSet 1	SVM	0.7647	0.6000	0.8710	0.7500	0.6000	0.6667	0.7229
	KNN	0.7843	0.9000	0.7097	0.6667	0.9000	0.7660	0.7992
	DRB	0.7647	0.8500	0.7097	0.6538	0.8500	0.7391	0.7767
DataSet 2	SVM	0.6450	0.5833	0.7076	0.6654	0.5833	0.6217	0.6420
	KNN	0.8967	0.8933	0.9000	0.8993	0.8933	0.8963	0.8967
	DRB	0.8850	0.8467	0.9233	0.9170	0.8467	0.8804	0.8842
DataSet 3	SVM	0.8101	0.7533	0.8269	0.5636	0.7533	0.6448	0.7893
	KNN	0.9847	0.9667	0.9901	0.9667	0.9667	0.9667	0.9783
	DRB	0.9703	0.9400	0.9792	0.9307	0.9400	0.9353	0.9594

b. Courbes Roc :

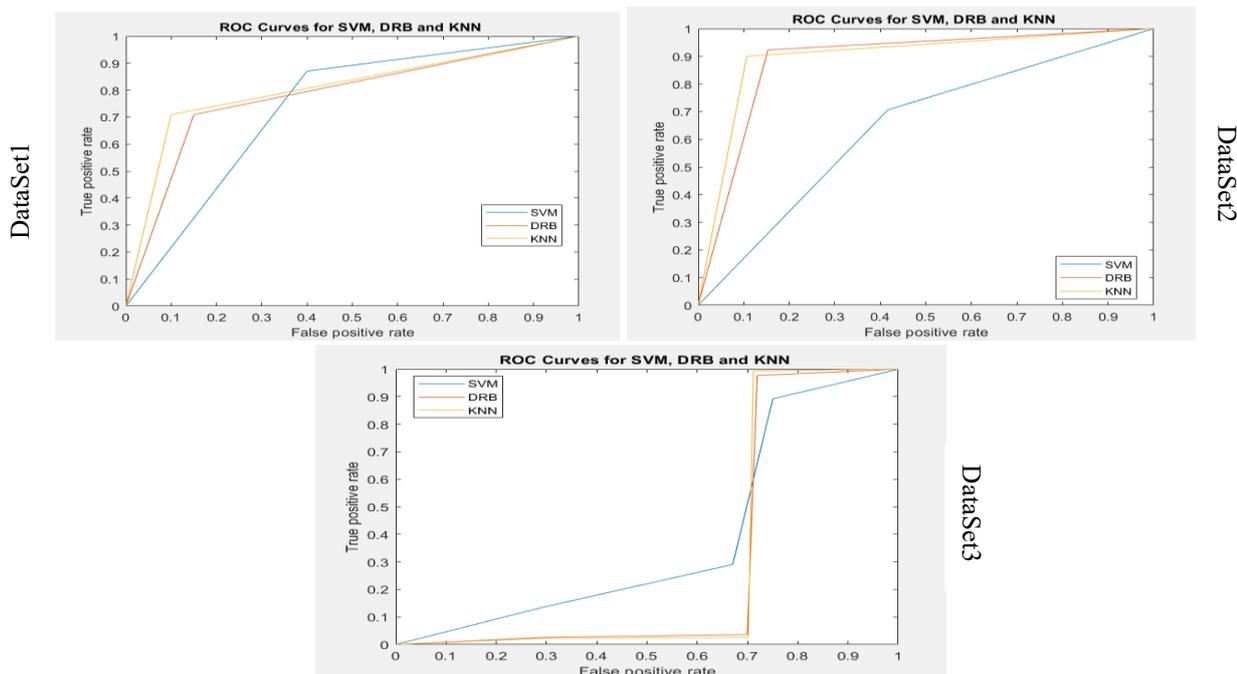


FIGURE 28 COURBES ROC DES CLASSIFIEURS DRB,KNN ET SVM ET L'EXTRACTEUR GIST AVEC LES TROIS DATA-SETS

c. Matrice de Confusion :

▪ **SVM :**

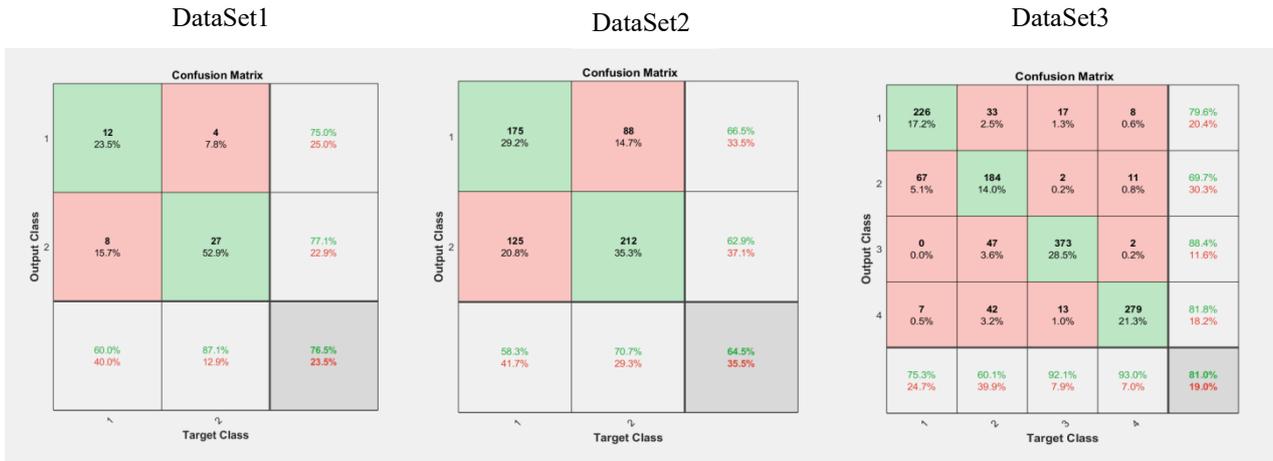


FIGURE 29 MATRICES DE CONFUSION DU SVM ET L'EXTRACTEUR GIST AVEC LES TROIS DATA-SETS

▪ **KNN :**

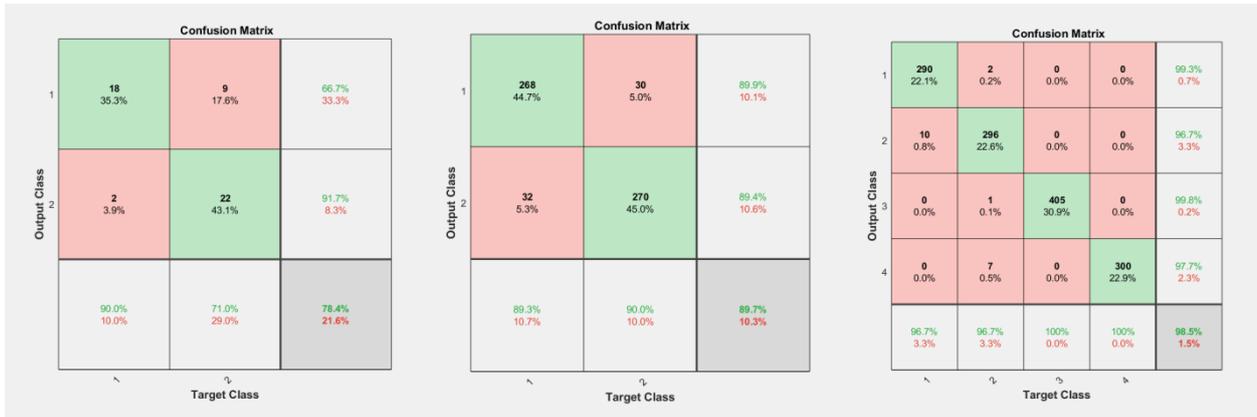


FIGURE 30 MATRICES DE CONFUSION DU KNN ET L'EXTRACTEUR GIST AVEC LES TROIS DATA-SETS

▪ **DRB :**

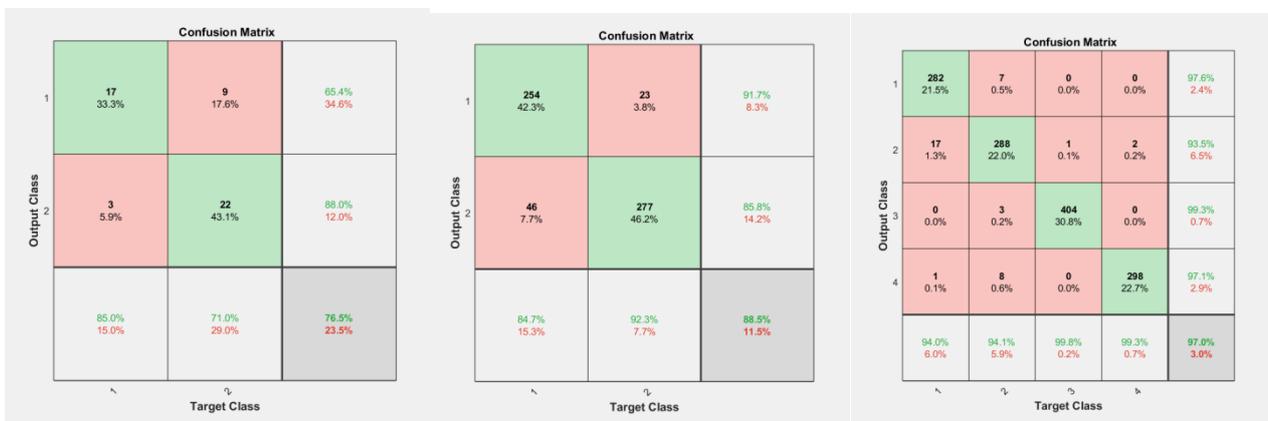


FIGURE 31 MATRICES DE CONFUSION DU DRB ET L'EXTRACTEUR GIST AVEC LES TROIS DATA-SETS

Discussion :

- Avec le modèle **HOG** le classifieur **KNN** donne une accuracy maximale de **78,34%**, **89.67%** et de **98.47%** avec les trois bases de données respectivement Dataset1, Dataset2 et Dataset3 .
- Le classifieur DRB donne presque les mêmes résultats .
- Avec le modèle GIST l'AUC du classifieur KNN et DRB est proche de 1 meilleure est la performance du classificateur. Et sa courbe se rapproche du coin supérieur gauche du graphique ca indique une meilleure performance .

➤ Résultats du modèle 3 : ResNet-50 avec les trois différents Classifieurs :

a. l'évaluation de la performance :

TABLEAU 4 RESULTATS COMPARATIFS DE L'EXTRACTEUR RESNET-50 AVEC LES CLASSIFIEURS DRB,KNN,SVM

DataSet	Classifieur	Accuracy	Sensitivity	Specifi- city	Precision	Recall	F_measure	Gmean
DataSet 1	SVM	0.8235	1.000	0.7097	0.6897	1.0000	0.8163	0.8424
	KNN	0.8627	1.000	0.7742	0.7407	1.000	0.8511	0.8799
	DRB	0.9020	1.000	0.8387	0.8000	1.000	0.8889	0.9158
DataSet 2	SVM	0.9367	0.8867	0.9867	0.9852	0.8867	0.9333	0.9353
	KNN	0.9767	0.9700	0.9833	0.9831	0.9700	0.9765	0.9766
	DRB	0.9750	0.9667	0.9833	0.9831	0.9667	0.9748	0.9750
DataSet 3	SVM	0.8902	0.7967	0.9179	0.7422	0.7967	0.7685	0.8551
	KNN	0.9901	0.9967	0.9881	0.9614	0.9967	0.9787	0.9924
	DRB	0.9901	0.9967	0.9881	0.9614	0.9967	0.9787	0.9924

b. Courbes Roc

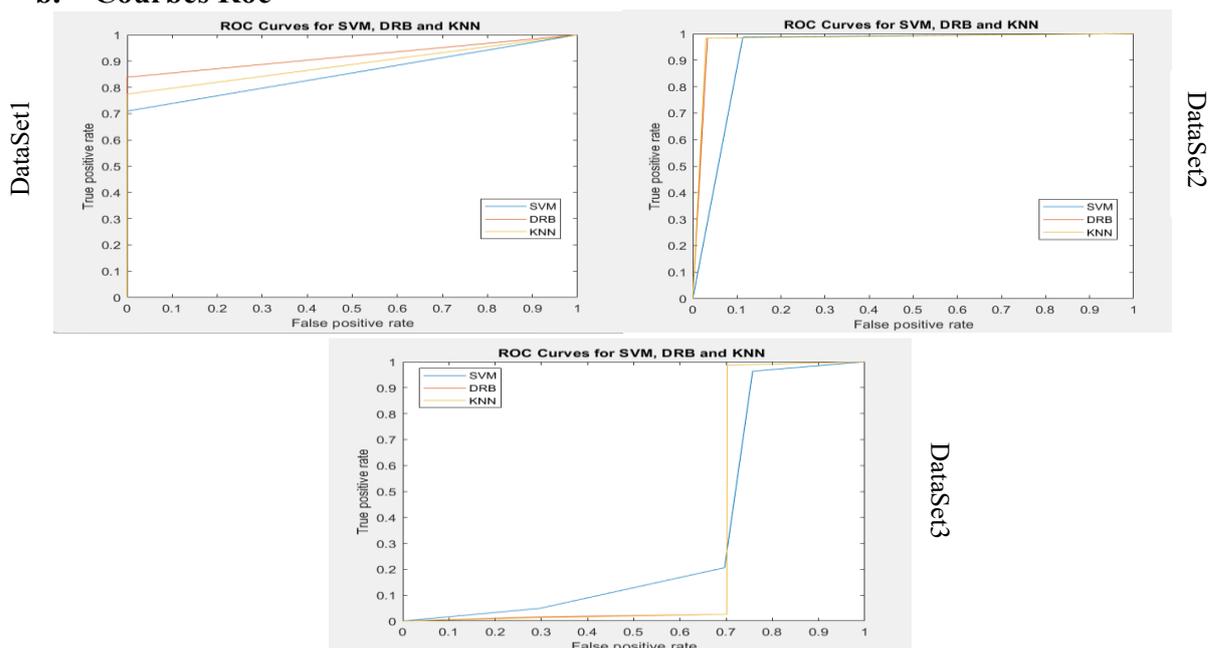


FIGURE 32 COURBES ROC DES CLASSIFIEURS DRB,KNN ET SVM ET L'EXTRACTEUR RESNET-50 AVEC LES TROIS DATA-SETS

c. Matrice de Confusion :

▪ **SVM :**

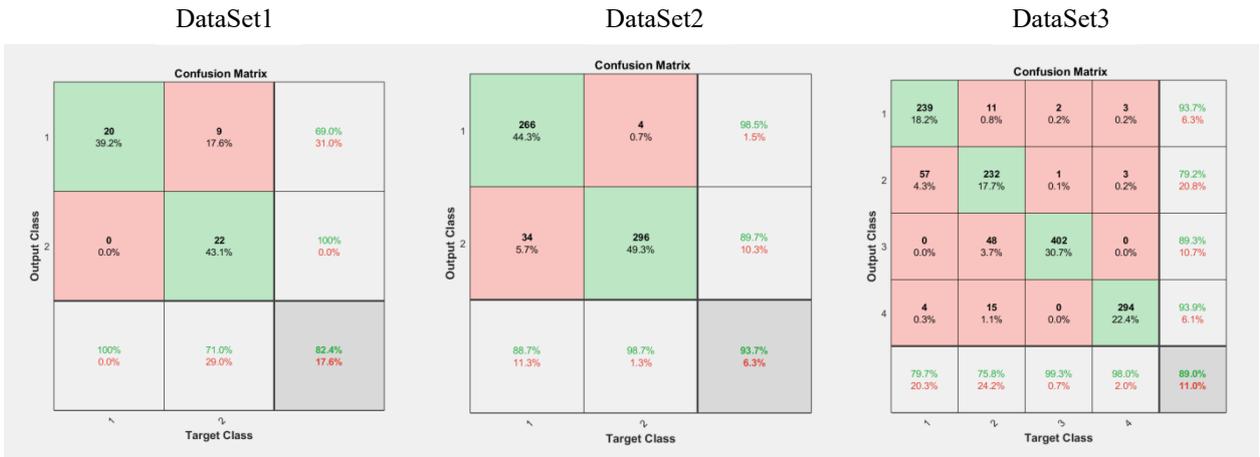


FIGURE 33 MATRICES DE CONFUSION DU SVM ET L'EXTRACTEUR RESNET-50 AVEC LES TROIS DATA-SET

▪ **KNN :**

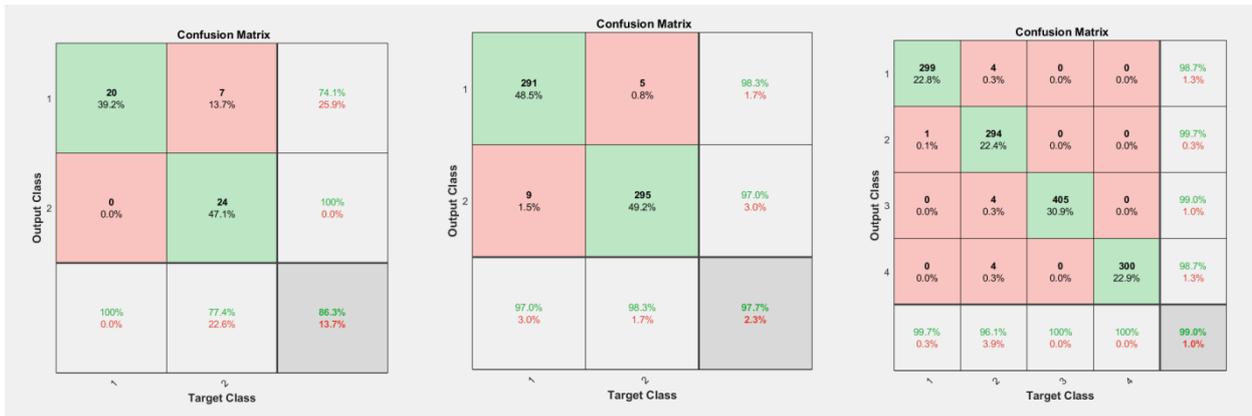


FIGURE 34 MATRICES DE CONFUSION DU KNN ET L'EXTRACTEUR RESNET-50 AVEC LES TROIS DATA-SET

▪ **DRB :**

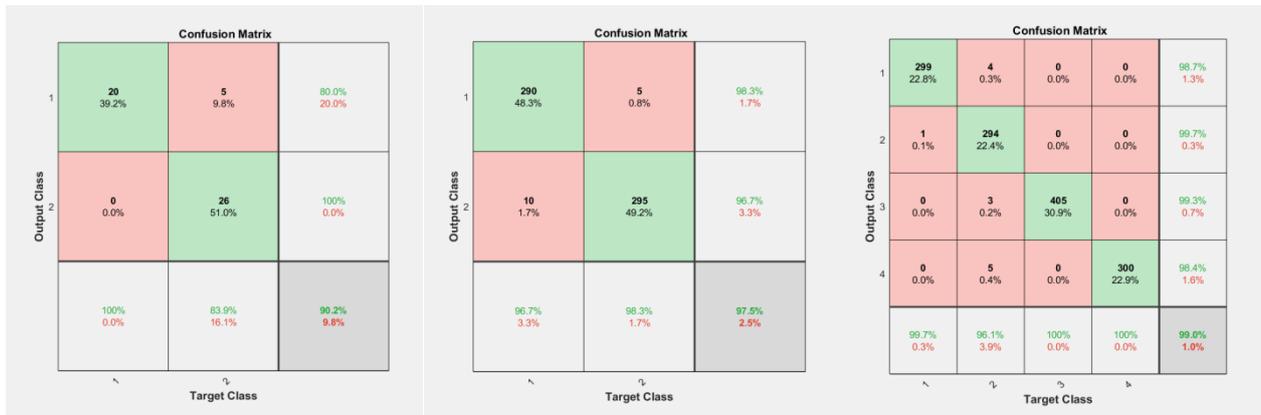


FIGURE 35 MATRICES DE CONFUSION DU DRB ET L'EXTRACTEUR RESNET-50 AVEC LES TROIS DATA-SET

Discussion :

- Le modèle ResNet-50 prouve que la performance des classifieurs **DRB** est supérieure dans Dataset1 **90.20%** à celle de SVM et KNN , alors que celle du KNN est **97.67%** avec le Dataset2 , par contre avec le Dataset3 le **KNN** et le **DRB** ont la meme performance **99,01%** .
- Dans les trois DataSet les courbes de les classifieurs **DRB** et **KNN** se rapproche du coin supérieur gauche du graphique , et sont très proches de **1** .

➤ Modèle 4 : ResNet-50 avec les trois différents Classifieurs optimisé par PSO,GW,GA (dataset1) :

a. l'évaluation de la performance :

TABLEAU 5 RESULTATS COMPARATIFS DE L'EXTRACTEUR RESNET-50 AVEC LES CLASSIFIEURS DRB,KNN,SVM OPTIMISE PAR PSO,GW,GW

Dataset	L'algorithme d'optimisation	Nombrede features	Classifieur	Accuracy	Sensitivity	Specificity	Precision	Recall	F_measure	Gmean
Dataset 1	PSO	484	SVM	0.8435	1.0000	0.7419	0.7143	1.0000	0.8333	0.8614
			KNN	0.9027	1.0000	0.7419	0.7143	1.0000	0.8333	0.8614
			DRB	0.9635	1.0000	0.7597	0.8897	1.0000	0.8463	0.8824
	GW	640	SVM	0.8627	1.0000	0.7742	0.7407	1.0000	0.8511	0.8799
			KNN	0.9216	1.0000	0.7419	0.7143	1.0000	0.8333	0.8614
			DRB	0.9235	1.0000	0.8710	0.8333	1.0000	0.9091	0.9333
	GA	486	SVM	0.8235	0.9500	0.7419	0.7037	0.9500	0.8085	0.8395
			KNN	0.8827	1.0000	0.7419	0.7143	1.0000	0.8333	0.8614
			DRB	0.9035	0.9500	0.9032	0.8636	0.9500	0.9048	0.9263
Dataset 2	PSO	514	SVM	0.9200	0.8800	0.9600	0.9565	0.8800	0.9167	0.9191
			KNN	0.9867	0.9633	0.9833	0.9830	0.9633	0.9731	0.9733
			DRB	0.9867	0.9633	0.9833	0.9830	0.9633	0.9731	0.9733
	GW	683	SVM	0.9369	0.8767	0.9833	0.9813	0.8767	0.9261	0.9285
			KNN	0.9735	0.9600	0.9833	0.9829	0.9600	0.9713	0.9716
			DRB	0.9833	0.9733	0.9933	0.9932	0.9733	0.9832	0.9833
	GA	484	SVM	0.9433	0.8800	0.9667	0.9635	0.8800	0.9199	0.9223
			KNN	0.9783	0.9500	0.9867	0.9862	0.9500	0.9677	0.9682
			DRB	0.9800	0.9700	0.9900	0.9898	0.9700	0.9798	0.9799

b. Matrice de Confusion :

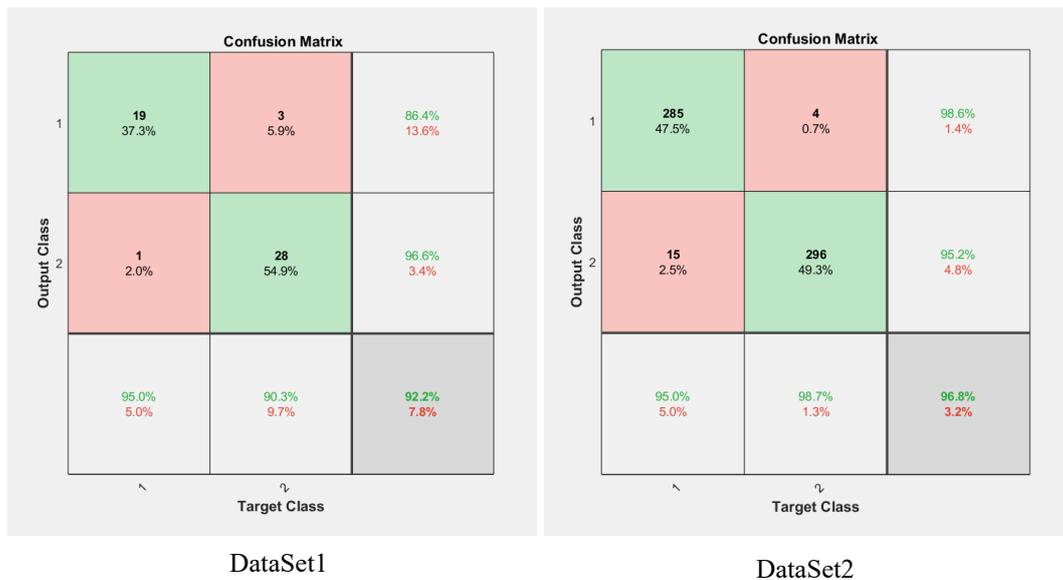


FIGURE 36 RESNET-50 AVEC DRB,SVM,KNN ET DATASET1+2 OPTIMISE PAR PSO,GW,GA

Discussion :

- Une comparaison des performances des trois algorithmes d'optimisation avec **Dataset1** et **ResNet50**. Dans ce modèle PSO surpasse GW et GA en termes de performance de classification et de nombre de caractéristiques. Ce dernier est moins de la moitié (**484 sur 1000**) et le classifieur **DRB** marque une Accuracy meilleur que **KNN** et **SVM** et plus que sans optimisation **96,35 %** sur **90,20 %**. Notons aussi que dans ce modèle **l'Accuracy du DRB** est **supérieure** par les trois algorithmes optimisation PSO, GW et GA que les autres classifieurs KNN et SVM.
- Le résultat des performances du classifieur **DRB** et l'extracteur **ResNet-50** optimisé par les trois algorithmes d'optimisation avec **Dataset2** :
 Dans ce modèle **PSO** a marqué avec les classifieurs **DRB** et **KNN** une meilleure accuracy améliorée par rapport au modèle basique **98.67%** sur **93.67%** avec une optimisation de l'extracteur presque la moitié (**514** sur 1000). Bien que l'accuracy du DRB avec les autres algorithmes GW et GA est la meilleur par rapport aux SVM ; KNN , le nombre GA reste toujours le petite par rapport aux autres algorithmes.

➤ **Modèle 5 : HOG avec les trois différents Classifieurs optimisé par PSO,GW,GA (dataset1+2) :**

a. l'évaluation de la performance :

TABLEAU 6 RESULTATS COMPARATIFS DE L'EXTRACTEUR HOG AVEC LES CLASSIFIEURS DRB,KNN,SVM OPTIMISE PAR PSO,GW,GW

Dataset	L'algorithme d'optimisation	Nombre de features	Classifieur	Accuracy	Sensitivity	Specificity	Precision	Recall	F_measure	Gmean
Dataset 1	PSO	18412	SVM	0.8235	0.7000	0.9032	0.8235	0.7000	0.7568	0.7951
			KNN	0.8727	0.9500	0.7419	0.7037	0.9500	0.8085	0.8395
			DRB	0.8831	0.9500	0.7742	0.7308	0.9500	0.8261	0.8576
	GW	25505	SVM	0.8235	0.7000	0.9032	0.8235	0.7000	0.7568	0.7951
			KNN	0.8727	0.9500	0.8065	0.7600	0.9500	0.8444	0.8753
			DRB	0.8831	0.9500	0.7742	0.7308	0.9500	0.8261	0.8576
	GA	18400	SVM	0.8235	0.7000	0.9032	0.8235	0.7000	0.7568	0.7951
			KNN	0.8727	0.9500	0.8387	0.7917	0.9500	0.8636	0.8926
			DRB	0.8831	0.9500	0.7742	0.7308	0.9500	0.8261	0.8576
Dataset 2	PSO	18405	SVM	0.8950	0.8833	0.9067	0.9044	0.8833	0.8938	0.8949
			KNN	0.9960	0.9967	0.9933	0.9934	0.9967	0.9967	0.9950
			DRB	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967
	GW	25447	SVM	0.9217	0.9067	0.9367	0.9347	0.9067	0.9205	0.9215
			KNN	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967
			DRB	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967
	GA	18502	SVM	0.8900	0.8533	0.9267	0.9209	0.8533	0.8858	0.8892
			KNN	0.9950	0.9967	0.9933	0.9934	0.9967	0.9950	0.9950
			DRB	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967

b. Matrice de confusion :

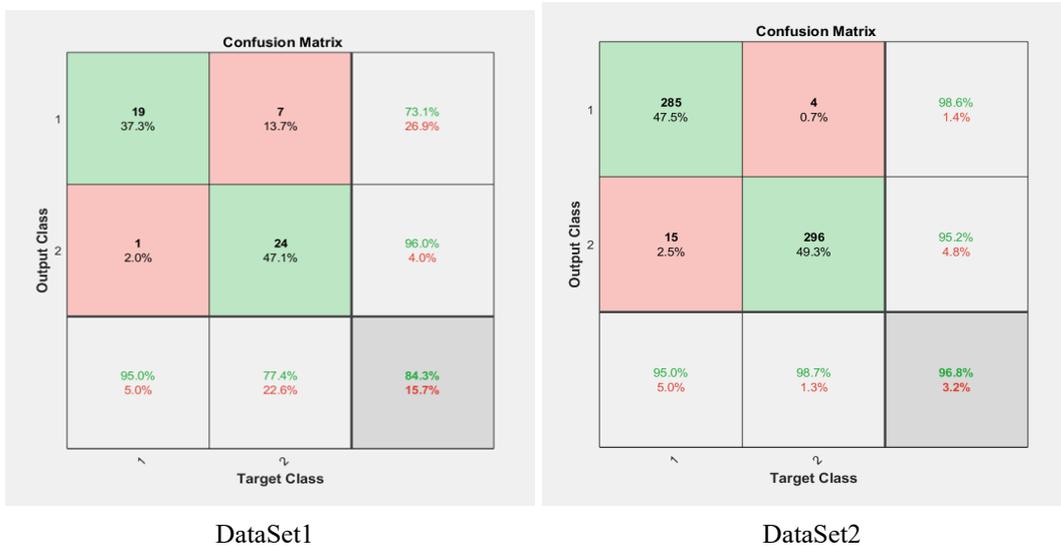


FIGURE 37 HOG AVEC DRB,SVM,KNN ET DATASET1+2 OPTIMISE PAR PSO,GW,GA

Discussion :

- La comparaison des performances des trois algorithmes d'optimisation avec **Dataset1** et Hog. PSO a surpassé GW et GA en termes de performances de classification et de nombre de caractéristiques :

Il a sélectionné presque la moitié des caractéristiques (**18412** sur 36720) et a obtenu une classification similaire ou nettement meilleure que l'utilisation de toutes les caractéristiques. **DRB-PSO** ont donné accuracy égale à **88,31%** par rapport à **DRB** basique **86.27%**, seule l'accuracy **SVM-PSO** est similaire à **SVM** basique 82.35%. On remarque aussi que le nombre de caractéristiques de GA est le plus petit mais l'accuracy est moins que par PSO et GW .

- le résultat des performances du classifieur **DRB** et l'extracteur **Hog** optimisé par les **trois algorithmes d'optimisation** avec **Dataset2** :

Dans ce modèle l'accuracy du DRB est la meilleure **99.67%** avec **PSO** (similaire au modèle basique). Une autre fois le classifieur DRB prouve sa robustesse avec PSO.

- **Modèle 6 : GIST avec les trois différents Classifieurs optimisé par PSO,GW,GA (dataset1+2) :**

a. l'évaluation de la performance :

TABLEAU 7 RESULTATS COMPARATIFS DE L'EXTRACTEUR GIST AVEC LES CLASSIFIEURS DRB,KNN,SVM OPTIMISE PAR PSO,GW,GW

Dataset	L'algorithme d'optimisation	Nombre de features	Classifieur	Accuracy	Sensitivity	Specificity	Precision	Recall	F_measure	Gmean
Dataset 1	PSO	18412	SVM	0.7843	0.6000	0.9032	0.8000	0.6000	0.6857	0.7362
			KNN	0.7843	0.9000	0.7097	0.6667	0.9000	0.7660	0.7992
			DRB	0.7451	0.8500	0.6774	0.6296	0.8500	0.7660	0.7588
	GW	25505	SVM	0.8235	0.6000	0.9677	0.9231	0.6000	0.7273	0.7620
			KNN	0.8235	1.0000	0.7097	0.6897	1.0000	0.8163	0.8424
			DRB	0.7843	0.9500	0.6774	0.6552	0.9500	0.7755	0.8022
	GA	18400	SVM	0.8039	0.5500	0.9677	0.9167	0.5500	0.6875	0.7296
			KNN	0.7647	0.9000	0.6774	0.6429	0.9000	0.7500	0.7808
			DRB	0.7255	0.7500	0.7097	0.6250	0.7500	0.6818	0.7296
Dataset 2	PSO	18405	SVM	0.6376	0.5967	0.6767	0.6486	0.5967	0.6215	0.6354
			KNN	0.8967	0.8900	0.9067	0.9051	0.8900	0.8975	0.8983
			DRB	0.8776	0.8300	0.9200	0.9121	0.8300	0.8691	0.8738
	GW	25447	SVM	0.5983	0.5733	0.6233	0.6035	0.5733	0.5880	0.5978
			KNN	0.9883	0.8767	0.9167	0.9132	0.8767	0.8946	0.8964
			DRB	0.8817	0.8267	0.9367	0.9288	0.8267	0.8748	0.8799
	GA	18402	SVM	0.6133	0.5600	0.6667	0.6269	0.5600	0.5915	0.6110
			KNN	0.9850	0.8900	0.9000	0.8990	0.8900	0.8945	0.8950
			DRB	0.8750	0.8433	0.9100	0.9036	0.8433	0.8724	0.8760

b. Matrice de confusion :

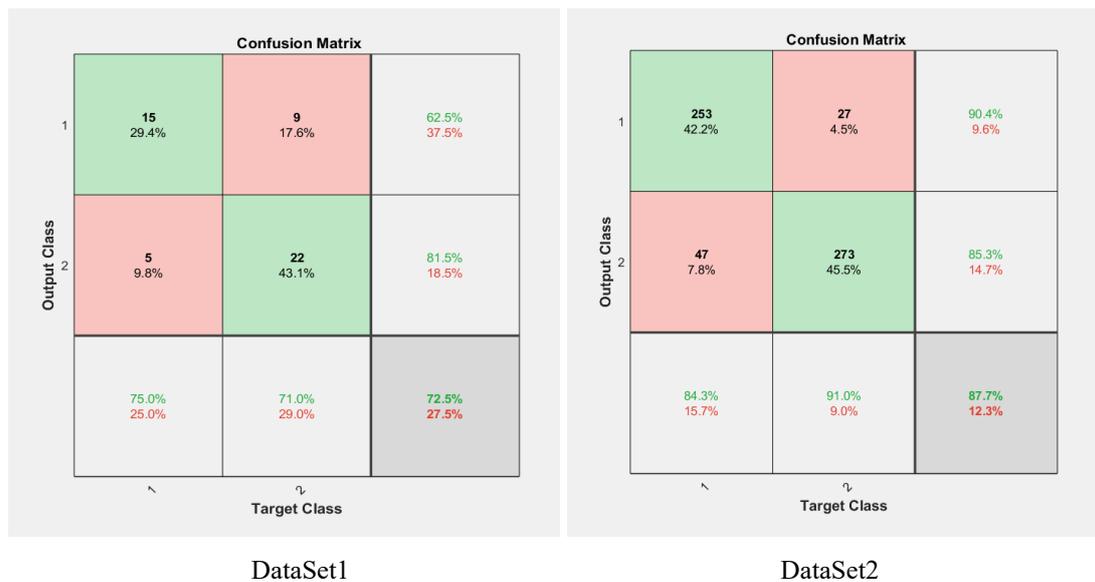


FIGURE 38 GIST AVEC DRB,SVM,KNN ET DATASET1+2 OPTIMISE PAR PSO,GW,GA

Discussion:

- La comparaison des performances des trois algorithmes d'optimisation avec Dataset1 et Gist :

Dans ce modèle PSO a surpassé GW et GA en termes de performances de classification et de nombre de caractéristiques. Il a sélectionné des caractéristiques 18412 sur 30000 des caractéristiques et a obtenu une classification d'accuracy SVM 78,43% .

- La comparaison des performances des classifieurs **DRB, KNN et SVM** avec les trois algorithmes d'optimisation sur l'extracteur **Gist** et **Dataset2** :

Avec cette approche l'**accuracy** est augmentée comparant à la classification sans optimisation. Premièrement on note que **GA** est meilleur en termes de nombre de caractéristiques optimisées **18402 sur 30000** , mais l'accuracy obtenue est la mauvaise par rapport à **GW** et **PSO**, ce dernier a surpassé les deux algorithmes en termes de performances de classification l'accuracy **KNN** est **98.50 %** , Dans cette approche **GW** se classe le premier, accuracy de **KNN** égale à **98.88%** .

4-3 Changement d'Angle :

l'angle joue un rôle essentiel dans le HOG en fournissant des informations sur les orientations des contours et des caractéristiques locales, ce qui permet de représenter efficacement les objets dans l'image pour leur détection.

Dans cette étape Nous avons effectué un changement de l'angle du Hog (120° au lieu du 90°) au niveau de la dataset 1, et ce changement a entraîné à une vaste augmentation de l'accuracy des trois classifieurs DRB SVM et KNN avec le Hog.

➤ Changement D'angle de 90° à 120° :

On l'as appliquer que sur le code du hog car l'angle accepte le changement que sur le HOG de 90° à 120° de la dateset 1 car sont les résultats les plus minimales .

a. l'évaluation de la performance :

TABLEAU 8 RESULTATS COMPARATIFS ENTRE L'ANGLE 90 ET 120 SUR LES TROIS CLASSIFIEUR AVEC HOG

Angle	Classifieur	Accuracy	Sensitivity	Specificity	Precision	Recall	F_measure	Gmean
Angle 90°	SVM	0.8235	0.7500	0.8710	0.7895	0.7500	0.7692	0.8082
	KNN	0.8627	0.9500	0.8065	0.7600	0.9500	0.8444	0.8753
	DRB	0.8627	0.9500	0.8065	0.7600	0.9500	0.8444	0.8753
Angle 120°	SVM	0.9417	0.9267	0.9567	0.9553	0.9267	0.9408	0.9415
	KNN	0.9954	0.9954	0.9954	0.9954	0.9954	0.9954	0.9954
	DRB	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967	0.9967

b. Courbes Roc :

L'angle 90° :

L'angle 120° :

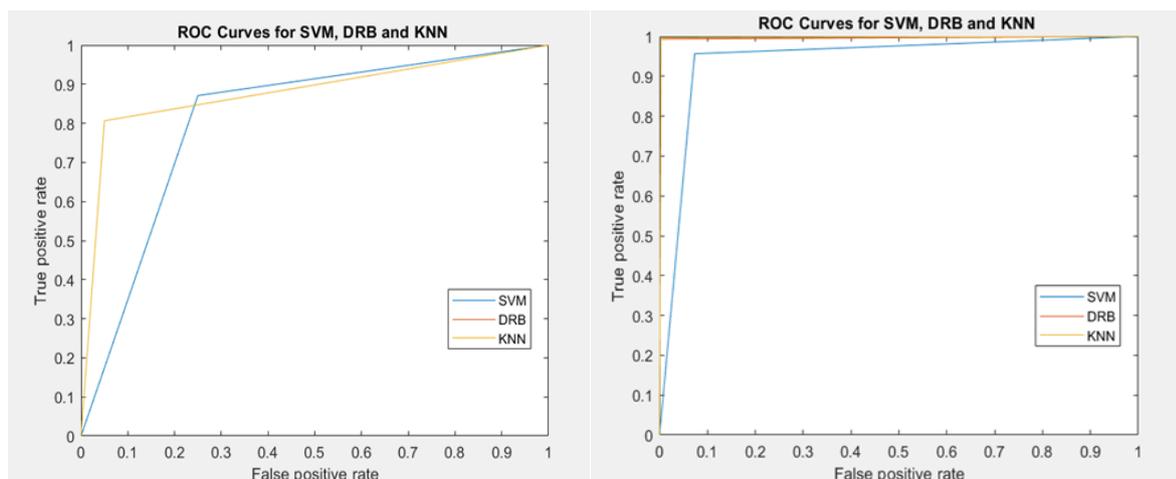


FIGURE 39 COMPARISON DES COURBES ENTRE L'ANGLE 90 ET 120 SUR LES TROIS CLASSIFIEUR AVEC HOG

c. Matrice de confusion :

▪ SVM :

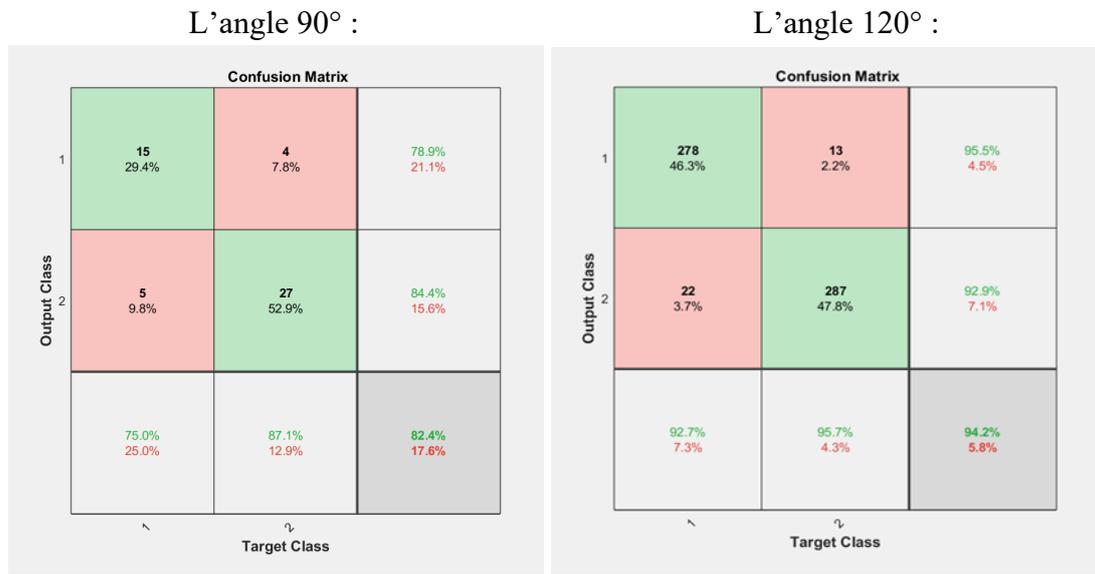


FIGURE 40 COMPARISON DES MATRICES DE CONFUSION ENTRE L'ANGLE 90 ET 120 SUR LE CLASSIFIEUR SVM AVEC HOG

▪ KNN :

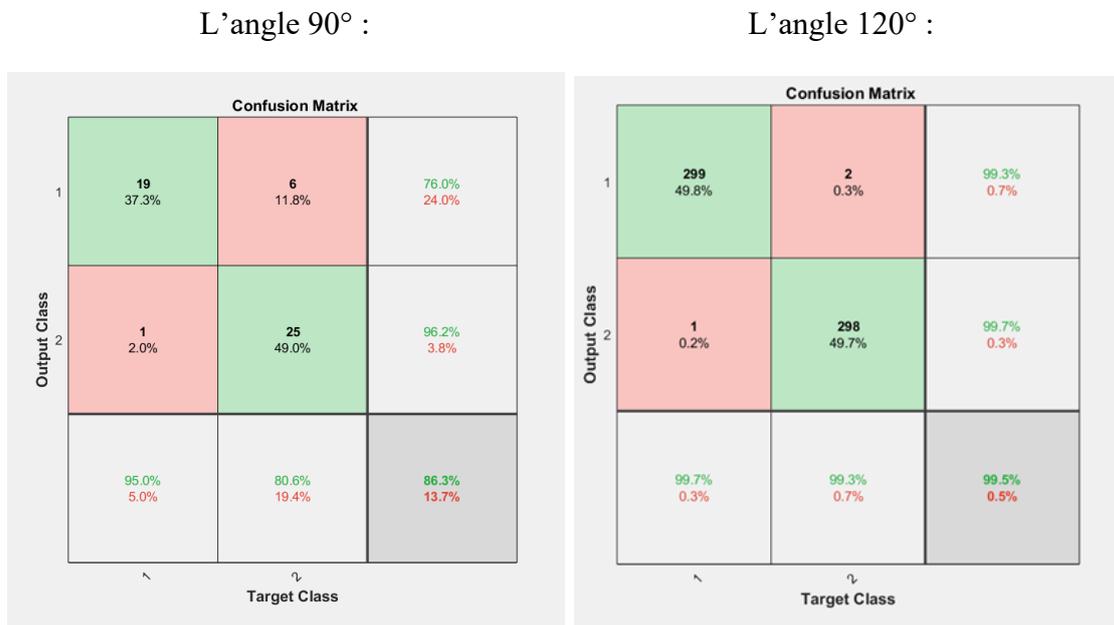
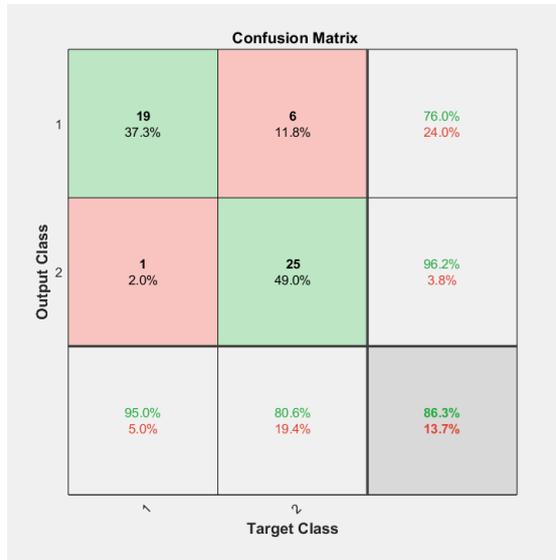


FIGURE 41 COMPARISON DES MATRICES DE CONFUSION ENTRE L'ANGLE 90 ET 120 SUR LE CLASSIFIEUR KNN AVEC HOG

▪ **DRB :**

L'angle 90° :



L'angle 120° :

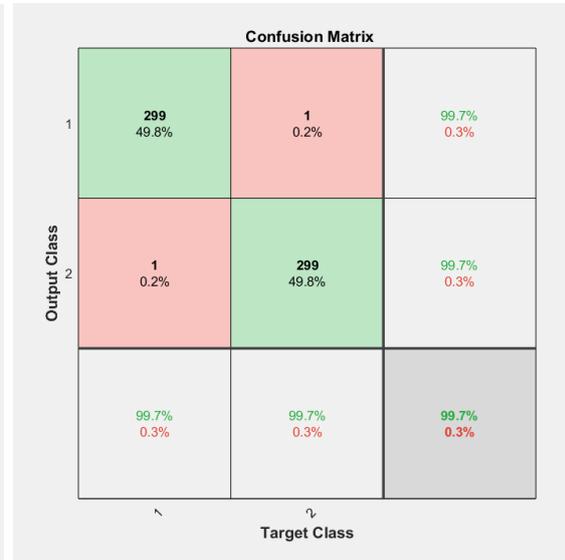


FIGURE 42 COMPARISON DES MATRICES DE CONFUSION ENTRE L'ANGLE 90 ET 120 SUR LE CLASSIFIEUR DRB AVEC HOG

Discussion :

- On remarque que l'angle de 120 augmentait significativement la précision par rapport à l'angle de 90 .
- DRB est le meilleur algorithme de classification car il a la plus grande précision 99,67%.

5- L'évolution du classifieur DRB avec et sans optimisation et avec le changement d'angle :

TABLEAU 9 EVOLUTION DES RESULTATS DU DRB

		DataSet1	DataSet2
HOG	Basique	86,27%	99,67%
	PSO	88,31%	99,67%
	GW	88,31%	99,67%
	GA	88,31%	99,67%
L'angle	90	86,27%	
	120	99,67%	
GIST	Basique	76,47%	88,50%
	PSO	74,51%	87,76%
	GW	78,43%	88,17%
	GA	72,55%	87,50%
ResNet-50	Basique	90,20%	97,50%
	PSO	96,35%	98,67%
	GW	92,35%	98,33%
	GA	90,35%	98,00%

6- Conclusion :

Dans ce chapitre, nous avons présenté tout d'abord les différents outils, data-sets, et algorithmes utilisés pour réaliser ce travail. L'implémentation des modèles est faite selon deux étapes principales : une sans optimisation et l'autre avec optimisation. Par la suite nous avons fait un changement d'angle sur le code du hog. En fin nous avons présenté les différents résultats des deux étapes avec une comparaison des mesures de performance des différentes approches pour l'optimisation et la classification.

Conclusion Générale :

Le travail réalisé dans le cadre de notre étude s'inscrit dans le domaine de classification des images d'IRM cérébrales. La classification des IRM cérébrales basée sur les techniques de l'apprentissage profond. Dans ce rapport, nous avons mis en évidence un certain nombre de notions et de définitions concernant l'imagerie médicale ; l'IRM, l'IA et ses approches ML et le Deep Learning. Ainsi on a cité les domaines d'application de ce dernier dont la classification.

Ainsi, nous avons exploré différentes méthodes de classification des IRM cérébrales et présenté leurs avantages et limitations, telles que les arbres de décision, SVM, KNN, DRB, K-means. Cela est particulièrement important dans le domaine médical, où la transparence et la justification des décisions sont cruciales.

Nous avons discuté de l'importance de la préparation des données et du prétraitement des images IRM cérébrales, notamment la normalisation, la réduction de dimension et le nettoyage des données. Ces étapes sont essentielles pour garantir la qualité des données et améliorer les performances du modèle de classification.

Ensuite, nous avons examiné différents extracteurs de connaissances, telles que le HOG, GIST, RESNET-50 avec les classificateurs SVM, KNN ainsi que l'utilisation de l'algorithme DRB. Chacune de ces méthodes présente ses avantages et ses inconvénients, et il est important de choisir la méthode la plus adaptée au problème spécifique et aux données disponibles.

Nous avons également souligné l'importance de l'optimisation dans le processus de classification. L'optimisation des hyperparamètres et la recherche des meilleures configurations sont des étapes critiques pour améliorer les performances du modèle. Différentes méthodes d'optimisation, telles que PSO, GW, GA, peuvent être utilisées pour trouver les meilleurs paramètres du modèle.

Un autre module est introduit dans notre étude, qui est la comparaison entre différentes méthodes de distance dans le classifieur DRB sur le Dataset1. Avec un changement de quelques paramètres : l'angle du Hog (120° au lieu de 90°).

Les résultats montrent que l'optimisation de classifieur DRB avec PSO et HOG dans l'angle juste (120°) améliore l'accuracy à 99.67 %.

En conclusion, la classification des IRM cérébrales basée sur des méthodes explicables est un domaine de recherche en plein essor qui offre des opportunités passionnantes pour améliorer les diagnostics et la prise de décision médicale. En combinant des méthodes explicables avec des techniques d'optimisation appropriées, il est possible de développer des modèles de classification précis, interprétables et robustes, ouvrant ainsi la voie à une meilleure compréhension des maladies cérébrales et à des soins de santé améliorés.

Bibliographies :

- [1] Nazir M, Shakil S, Khurshid K, Role of Deep Learning in Brain Tumor Detection and Classification (2015 to 2020): A Review, Computerized Medical Imaging and Graphics , (2021) doi: <https://doi.org/10.1016/j.compmedimag.2021.101940>
- [2] Séverine BOILLEE , COMPRENDRE LE CERVEAU, SON DÉVELOPPEMENT, SON FONCTIONNEMENT, SA PLASTICITÉ ET SES PATHOLOGIES , site web de : Institut du Cerveau, Hôpital Pitié-Salpêtrière
- [3] Essentiel sur le cerveau , Publié le 16 mars 2023 , récupéré sur <https://www.cea.fr/comprendre>
- [4] Steven A. Goldman , MD, PhD, Sana Biotechnology c, Présentation des tumeurs cérébrales Revue/Révision complète mars 2021
- [5] American Association of Neurological Surgeons , Patients, Neurosurgical Conditions and Treatments , Brain Tumors
- [6] Jyotishka Das, Suvadeep Ghosh, Rivu Chakraborty, and Ankita Pramanik. Chapter 16 Deep Learning Based Classification of Brain Tumor Types from MRI Scans. Advanced Machine Learning Approaches in Cancer Prognosis, Intelligent Systems Reference Library 204, 2021
- [7] Thomas Bourdrel, radiologue, IRM : préparation, déroulé, risques Article mis à jour le 08/06/22
- [8] D. Hartmann (2002) Revue Neurologique , examine l'utilisation de l'IRM dans la détection du cancer du pancréas , doi : [https://doi.org/10.1016/S0035-3787\(04\)70848-6](https://doi.org/10.1016/S0035-3787(04)70848-6)
- [9] Minakshi Sharma and Neha Miglani. Automated Brain Tumor Segmentation in MRI Images Using Deep Learning: Overview, Challenges and Future. Deep Learning Techniques for Biomedical and Health Informatics, Studies in Big Data 68, https://doi.org/10.1007/978-3-030-33966-1_16 2020
- [10] Arti Tiwari a , *, Shilpa Srivastava b , Millie Pant. Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019. Pattern Recognition Letters 131 (2020) 244–260.
- [11] Charlotte Arce , L'intelligence artificielle et la médecine , doi : <https://www.pourquoidoctor.fr/Articles/>
- [12] Matija Snuderl, Tumeurs au cerveau : quand l'intelligence artificielle permet leur détection "presque en temps réel" , Intelligence Artificielle , 456-36
- [13] Trésor de la Langue Française Informatisé
- [14] www.enssib.fr/le-dictionnaire/classification

- [15] D. Bensalem, C. Bounouar, Z. Boudia, Classification automatique de documents: de la classification classique à la classification utilisant une ressource externe, Département Informatique de l'UMMTO, 2014.
- [16] www.maths.cnam.fr
- [17] www.fr.wikipédia.
- [18] www.math.unice.fr
- [19] J.J. Rocchio, Jr, The SMART Retrieval System: Experiments in Automatic Document Processing, Prentice-Hall, 1971.
- [20] S.Douadi (2016), Comparaison des différentes approches de classification automatique des documents textuels. UNIVERSITE MOULOUD MAMMERI DE TIZI-OUZOU
- [21] "Math Behind SVM(Kernel Trick) " ,Figure, 16 février 2020, disponible sur :<https://mc.ai/math-behind-svmkernel-trick/>
- [22] NORBERT BERTRAND SANKA (mars 2020) ,ÉTUDE COMPARATIVE ET CHOIX OPTIMAL DU NOMBRE DE CLASSES EN CLASSIFICATION ET RÉSEAUX DE NEURONES : APPLICATION EN SCIENCE DES DONNÉES , L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES
- [23] A. Courville Y. Bengio and P. Vincent. Representation learning : A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8) :1798–1828, 2013.
- [24] Liu, X. L. (2007). Image segmentation algorithm based on simulated annealing and fuzzy c-means clustering. Journal of Engineering Graphics, 1:89{93.
- [25] MENOUEUR T, DERMOUCHE M, 2010.Mémoire fin d'étude.Systèmes Informatiques. Application de techniques de data mining pour la classification.ESI
- [26] Hand, D. J., & Till, R. J. (2001). A simple generalization of the area under the ROC curve for multiple class classification problems. Machine learning, 45(2), 171-18657
- [27] W.Beight, P. (1979). Foundation of optimization 2nd d. Englewood Cliffs.
- [28] Xue, B. (2014). Particle Swarm Optimisation for Feature Selection in Classification.Victoria University of Wellington.
- [29] "Computer Vision: Algorithms and Applications" de Richard Szeliski.
- [30] Z.HIRECHE. (2020). Extraction d'informations pertinentes guidée par expertise : application à la détection de pathologies cérébrales.Bordj Bou Arreridj: Université Mohammed El Bachir El Ibrahimi.

- [31] Choutri.Mehdi (2022) . Classification des IRMS cérébrales pathologiques basée sur l'extraction de caractéristiques avec optimisation . Bordj Bou Arreridj: Université Mohammed El Bachir El Ibrahimi.
- [32] Mifdal, R. (2019). Application des techniques d'apprentissage automatique pour la prédiction de la tendance des titres financiers. ÉCOLE DE TECHNOLOGIE SUPÉRIEURE UNIVERSITÉ DU QUÉBEC.
- [33] Hasni Aissa Haroun (2020) . Gestion optimale de l'intégration des sources PV au réseau distribution par utilisation des méthodes : GA, PSO, GWO . Université Mohamed Khider de Biskra .