

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

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

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Projet de Fin d'Études

Présenté par

HADJI Khadidja Imane

&

KAIBICHE Hind

pour l'obtention du diplôme de Master en Électronique

Option

Automatique et informatique industrielle

Thème

Détection de pathologies mammaires par les techniques bio-inspirées

Proposé par : Dr. REGUIEG F. Zohra & Pr. BENBLIDIA Nadja

Année Universitaire 2019-2020

Remerciements

Nous tenons à remercier tout d'abord DIEU le tout puissant, qui nous a donné durant toutes ces années la santé, le courage et la foi en nous-mêmes pour arriver à ce niveau-là.

Nous exprimons nos sincères remerciements à l'égard de Mesdames Reguieg F. Zohra et Benblidia Nadja, pour avoir accepté de nous encadrer, pour leur aide et leurs précieux encouragements dans la mise en œuvre, de ce modeste travail.

Nous exprimons notre gratitude aux membres du jury, pour l'intérêt porté à ce travail.

Notre infinie gratitude va à tous les enseignants du département d'électronique, qui ont contribué à notre formation.

Un grand merci à toute notre famille.

Finalement, nous remercions tous ceux qui ont participé de près ou de loin à l'achèvement de ce travail.

ملخص: التصوير الشعاعي للثدي هو أكثر تقنيات التصوير استخدامًا للكشف عن الأورام في مرحلة مبكرة. إنه حاليًا التحقيق الرئيسي في فحص سرطان الثدي.

يعد الكشف الشامل خطوة مهمة وحساسة للتشخيص المبكر لسرطان الثدي، وتتخلص هذه الصعوبة إلى حد كبير في تعقيد الصور الشعاعية للثدي وتنوع الجماهير المراد تقسيمها. مع الأخذ في الاعتبار هذه الصعوبة، ثلاث خوارزميات مختلفة مقترحة في هذه الأطروحة: خوارزمية FCM وخوارزمية الوسيلة-k-means المُحسَّنة باستخدام الخوارزمية الجينية وتجميع الوسائل k-means المحسَّنة باستخدام خوارزمية DPSO.

النتائج التي تم الحصول عليها واضحة جدًا من الناحية المرئية، وتُظهر أن الطرق المقترحة في هذا العمل تجعل من الممكن بنجاح استخراج الكتل من صور الثدي الشعاعية المرجعية لقاعدة بيانات MIAS .

كلمات المفاتيح: التصوير الشعاعي للثدي، التجزئة، k-means، FCM، DPSO،

Résumé : La mammographie est la technique d'imagerie la plus utilisée, pour détecter des tumeurs à un stade précoce. Elle est actuellement la principale investigation dans le dépistage du cancer du sein. La détection de masses est une opération importante et délicate, pour le diagnostic précoce du cancer du sein. Cette difficulté se présente, pour une large part, dans la complexité et la diversité des masses à segmenter. Tenant compte de cette difficulté, trois différents algorithmes sont proposés dans ce mémoire : l'algorithme du fuzzy c-means (FCM), l'algorithme k means optimisé avec l'algorithme génétique (AG) et le clustering k-means amélioré avec la DPSO. Les résultats obtenus sont visuellement très clairs et montrent que les méthodes proposées dans ce travail, permettent d'extraire avec succès les masses, à partir des mammographies référentes de la base de données MIAS.

Mots clés : Cancer du sein, Mammographie, segmentation, k-means, FCM, DPSO, AG.

Abstract: Mammography is the most widely used imaging technique, to detect tumors at an early stage. It is currently the main investigation in breast cancer screening.

Mass detection is an important and delicate operation, to do for the early diagnosis of breast cancer. This difficulty is summed up, largely, in the complexity of the mammographic images and the diversity of the masses to be segmented. Taking into account this difficulty, three different algorithms are proposed in this report : the fuzzy c means (FCM) algorithm, the k means algorithm optimized with the genetic algorithm and k-means clustering improved with the DPSO algorithm.

The obtained results are visually very clear and show that the methods proposed, in this work make it possible to successfully extract masses, from the referent mammograms of the MIAS database.

Keywords : Breast cancer, Mammography, segmentation, k-means, FCM, DPSO, AG.

Listes des acronymes et abréviations

AG : Algorithmes génétiques

DPSO : Darwinian Particle Swarm Optimization

FCM : Fuzzy C means

GAK-means : algorithme génétique avec l'algorithme k means

IM : Image mammographique

MIAS : Mammography Image Analysis Society

PSO : Particle swarm optimization

PSOG : Version PSO globale

Table des matières

Introduction générale.....	1
Chapitre 1 :contexte mammographique.....	3
1.1 Introduction.....	3
1.2 Anatomie du sein.....	4
1.3 Introduction au cancer du sein.....	5
1.4 Imagerie mammographique.....	7
1.4.1 Mammographie.....	7
1.4.2 Âge de pratique d'une mammographie.....	8
1.4.3 Constitution de la salle mammographie.....	8
1.4.5 Pathologies décelées par la mammographie.....	9
1.5 Anomalies mammaires.....	11
1.5.1 Calcifications.....	11
1.5.2 Masses.....	13
1.6 Classification ACR.....	18
1.7 Avantages et les inconvénients de la mammographie.....	19
1.7.1 Avantages.....	19
1.7.2 Inconvénients.....	20
1.8 Conclusion.....	21
Chapitre 2 :Techniques de la segmentation par classification et les algorithmes bio-inspirés.....	22
2.1 Introduction.....	22
2.2 Introduction à la segmentation.....	23
2.3 Segmentation par classification.....	25
2.3.1 Introduction au Clustering.....	25
2.3.2 Méthodologies du clustering.....	26
2.4 Clustering classique.....	29
2.4.1 Etapes du K-means.....	30
2.4.2 Fonctionnalités du k-means.....	31
2.5 Approche floue du clustering.....	32
2.5.1 Fonctionnement du FCM.....	33
2.5.2 Différence entre les k means et le fuzzy c means.....	35
2.6 Introduction aux techniques bio inspirées.....	36
2.6.1 Introduction à l'optimisation.....	36
2.6.2 Catégories des méthodes d'optimisation.....	37

➤ Paradigme des techniques bio inspirées	38
2.7 Algorithmes génétiques.....	39
2.7.1 Fonctionnement	39
2.7.2 Opérateurs d'évolution	40
2.7.3 Avantages et inconvénients des algorithmes génétiques	42
2.8 Optimisation par essaim de particules (PSO)	43
2.8.1 Mode de fonctionnement de l'algorithme PSO	44
2.8.2 Avantages et inconvénients de la PSO	47
2.9 Conclusion	48
Chapitre 3 : Clustering optimisé par les algorithmes génétiques et les essais de particules.....	49
3.1 Introduction.....	49
3.2 Logique floue	50
3.2.1 Notions d'ensemble et de sous-ensemble flou.....	50
3.2.2 Valeur d'appartenance	50
3.2.3 Fonctions d'appartenance.....	51
3.3 Organigramme de l'approche Fuzzy C-Means.....	52
3.4 Clustering optimisé par les AG	54
3.4.1 Principe général des AG	54
3.5 Clustering k means optimisé par la DPSO	60
3.5.1 Introduction à la DPSO	60
3.8 Conclusion	65
Chapitre 4 Mise en œuvre et résultats.....	66
4.1 Introduction.....	66
4.2 Environnement de travail	66
4.2.1 Matériel utilisé	66
4.2.2 Langage utilisé	66
4.2.3 Interface graphique	68
4.3 Approche de 'BreastBioSeg'	69
4.3.1 Base d'images.....	70
4.3.2 Méthode du Fuzzy c means (FCM)	71
4.3.3 Clustering classique optimisé par les AG.....	75
4.3.4 Approche des K-means basée sur l'optimisation par la DPSO	84
4.4 Discussion générale	89
4.5 Conclusion	90

Table des figures

Figure 1 : Anatomie du sein.....	4
Figure 2 : image d'une tumeur maligne dans une masse.....	5
Figure 3:Sein gras et dense.....	8
Figure 4:l'appareil de mammographie	9
Figure 5: Différents types des microcalcifications.....	11
Figure 6 : Masse bénigne.....	14
Figure 7: Masse maligne.....	14
Figure 8: Formes possibles d'une masse	15
Figure 9 :Différents contours possibles d'une masse.....	16
Figure 10:Différents types de densités mammaires	17
Figure 11 :: Exemple d'un clustering hiérarchique.....	26
Figure 12 :Exemple d'un clustering de partitionnement	27
Figure 13: Fonction d'appartenance dans k means / fuzzy c means	35
Figure 14: Changement de la fonction objectif pour un problème de minimisation.....	36
Figure 15: Paradigme de l'algorithme génétique.....	39
Figure 16: Comportement des oiseaux en formation en V.....	43
Figure 17: Schéma de principe du déplacement d'une particule.....	44
Figure 18: Comparaison d'un ensemble classique et d'un ensemble flou.....	50
Figure 19 : Fonction caractéristique [38]	51
Figure 20: Fonction d'appartenance	51
Figure 21 : Organigramme de l'approche FCM	53
Figure 22: Principales étapes des AG	54
Figure 23: Approche adoptée de GAK-means	58
Figure 24: Icône de matlab R2019b.....	67
Figure 25: Interface graphique de MATLAB	68
Figure 26 : Synoptique de BreastBioSeg.....	70
Figure 27: Exemples de masses bénigne et malignes de la base MIAS.....	71
Figure 28: Résultats du premier test appliqué sur l'image mdb184	72
Figure 29: Résultats du second test appliqué sur l'image mdb184	72
Figure 30: Résultats du troisième test appliqué sur l'image mdb184.....	72
Figure 31: Résultats du premier test appliqué sur l'image mdb002	73
Figure 32 : Résultats du second test appliqué sur l'image mdb002.....	73
Figure 33: Résultats du troisième test appliqué sur l'image mdb002.....	74
Figure 34: Résultats du premier test appliqué sur l'image mdb001	75
Figure 35: Evolution des clusters vers les meilleurs pour le premier test de l'image mdb001	76
Figure 36: Résultats du second test appliqué sur l'image mdb001	77
Figure 37: Evolution des clusters vers les meilleurs pour le second test de l'image mdb001	77
Figure 38: Résultats du troisième test appliqué sur mdb001	78
Figure 39: Evolution des clusters vers les meilleurs pour le troisième test de l'image mdb001	78
Figure 40: Résultats du troisième test appliqué sur mdb002	79
Figure 41: Evolution des clusters vers les meilleurs pour le premier test de l'image mdb002	79
Figure 42: Résultats du second test appliqué sur mdb002	80

Figure 43: Evolution des clusters vers les meilleurs pour le second test de l'image mdb002	80
Figure 44: Résultats du troisième test appliqué sur mdb002	81
Figure 45: Evolution des clusters vers les meilleurs pour le troisième test de l'image mdb002	81
Figure 46: Résultats du premier test appliqué sur mdb184	82
Figure 47: Evolution des clusters vers les meilleurs pour le premier test de l'image mdb184	82
Figure 48: Résultats du second test appliqué sur mdb184	83
Figure 49: Evolution des clusters vers les meilleurs pour le second test de l'image mdb184	83
Figure 50 : Résultats du premier test appliqué sur mdb001	85
Figure 51 : Résultats du second test appliqué sur mdb001	85
Figure 52 : Résultats du troisième test appliqué sur mdb001	85
Figure 53: Résultats du premier test appliqué sur mdb002	86
Figure 54: Résultats du second test appliqué sur mdb002	86
Figure 55: Résultats du troisième test appliqué sur mdb002	86
Figure 56: Résultats du premier test appliqué sur mdb184	87
Figure 57: Résultats du second test appliqué sur mdb184	87
Tableau 1.1 : Critères de distinction entre tumeurs bénignes/ malignes.	6

Introduction générale

Le cancer du sein est un problème de santé publique majeur. En effet, il s'agit d'un fameux cancer, menaçant la vie de la plupart des femmes. Ainsi, environ une femme sur dix, sera touchée par cette maladie au cours de sa vie [1]. En Algérie, l'incidence de cette lourde pathologie reste croissante et, il existe très peu, de programmes de dépistage organisé. Cependant, la réduction du taux de mortalité due à ce type de cancer et la favorisation des chances de guérison ne peuvent être atteintes, que si la tumeur a été prise en charge dès les premiers stades de son apparition. Dans ce cas, la mammographie reste la technique de référence incontournable pour l'exploration mammaire et la plus efficace, en termes de surveillance et de détection précoce du cancer du sein.

Afin d'assurer un dépistage précoce d'une telle tumeur, les radiologues ont été amenés à augmenter la fréquence des mammographies, en particulier pour la tranche d'âge la plus concernée. Cela a engendré une augmentation exponentielle, du nombre de mammographies réalisées. Par conséquent, l'interprétation est devenue difficile à gérer par les radiologues.

Notre travail se situe dans ce contexte, où l'exploration principale, se rapporte à la segmentation des masses dans les images mammographiques, dans le but d'extraire les zones suspectes pour faciliter le diagnostic. La question n'est pas de remplacer le médecin, mais de lui proposer des outils performants qui l'aident dans son travail d'analyse d'images mammographiques. Il est alors, impératif que l'opération de la segmentation, soit précise, pour faciliter les tâches du haut niveau et, obtenir un diagnostic fiable. Nous proposons dans ce cas, l'étude et le développement de plusieurs approches de segmentation.

Les objectifs fixés, sont alors consacrés à l'exploration des méthodes efficaces de la segmentation par classification ; suivant la méthode fuzzy c-means standard (FCM), la technique des k-means optimisée par les algorithmes génétiques, d'une part et par l'algorithme d'optimisation par essaim particulaire (DPSO), d'autre part.

Ce mémoire est structuré autour de quatre chapitres, qui permettent de présenter les différents aspects de notre travail. Pour ce faire, nous avons suivi les étapes décrites par les chapitres suivants :

- Le premier chapitre introduit le contexte médical de la mammographie.
- Le second chapitre est consacré à l'étude de la segmentation par classification et au domaine d'optimisation. Nous dressons notamment un état de l'art à propos de quelques méta-heuristiques bio-inspirées appliquées, au problème de la segmentation d'image.
- Le troisième chapitre présente la technique de la segmentation basée sur la logique floue suivant l'algorithme « FCM », ainsi que l'approche proposée pour améliorer l'algorithme k-means suivant l'algorithme génétique et la DPSO.
- Le dernier chapitre, présente l'environnement de travail, la démarche suivie pour l'implémentation des techniques du clustering optimisées, ainsi que les résultats obtenus.

Enfin, ce modeste travail est terminé par une conclusion générale, incluant des remarques importantes et des perspectives.

Chapitre 01 : contexte mammographique

Chapitre 1 : contexte mammographique

1.1 Introduction

L'objectif de ce chapitre, est d'étudier le contexte médical. A cet effet, nous décrivons l'architecture de la glande mammaire et mettons en relief, la technique d'imagerie mammaire par excellence, la mammographie radiologique. L'étude des pathologies majeures, sera abordée suivant les masses et les calcifications. Une analyse approfondie porte sur le type d'anomalie mammaire « les masses », pour situer le contexte de notre application.

Chapitre 01 : contexte mammographique

1.2 Anatomie du sein

Le sein est composé d'une glande mammaire, de fibres de soutien (ligaments de Cooper) et de graisse (tissu adipeux) ; le tout est recouvert par la peau (figure 1.1). La quantité de chacune de ses composantes peut varier d'une femme à l'autre.

Le sein est situé par-dessus le muscle pectoral. On trouve également dans le sein des nerfs, des vaisseaux sanguins et lymphatiques. La glande mammaire est divisée en 15 à 20 sections qu'on appelle lobes, composés de lobules. Ceux-ci sont reliés à des canaux qui se rendent sous le mamelon (situé au centre du sein). Il existe également des chaînes de ganglions lymphatiques, qui filtrent les microbes et protègent le corps contre l'infection et la maladie.

Le cancer du sein peut se développer tant au niveau d'un canal galactophore que d'un lobule et il peut également se retrouver au niveau des ganglions lymphatiques [2].

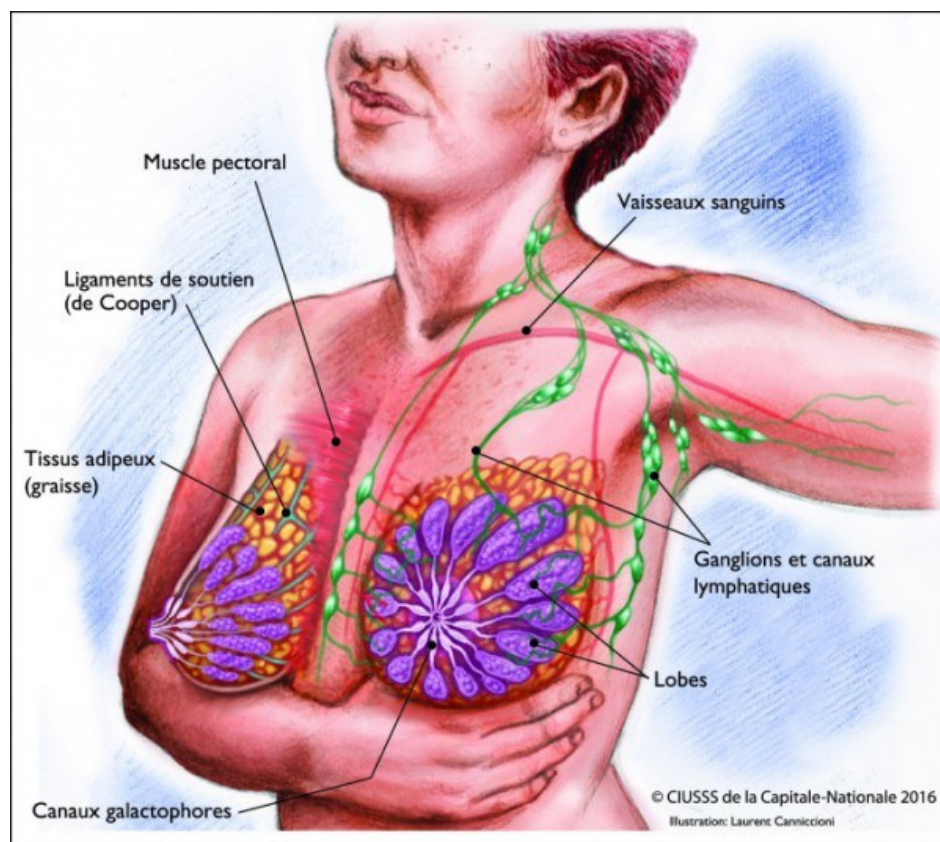


Figure 1 : Anatomie du sein

Chapitre 01 : contexte mammographique

1.3 Introduction au cancer du sein

Le cancer est une maladie qui affecte la cellule. En effet, la cellule est l'unité de base de la vie. Il en existe dans le corps plus de deux cents types différents ; qui ont tous un rôle spécifique : intervenant dans les muscles, les nerfs, les cellules osseuses, etc.

Une cellule cancéreuse est une cellule qui a changé. Habituellement, ces changements sont fixés par le corps. Cependant, lorsqu'une cellule devient cancéreuse, elle perd sa capacité à se réparer. Ensuite, elle commence à se multiplier et finit par former une masse, qu'on appelle tumeur maligne [3].

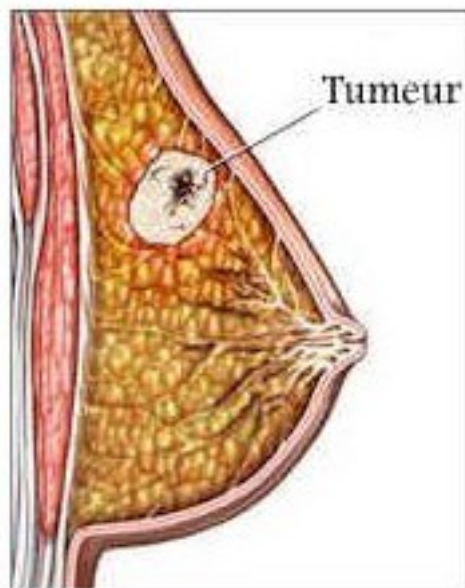


Figure 2 : image d'une tumeur maligne dans une masse

➤ Définition d'une tumeur

Le mot "tumeur" est un terme générique, correspondant au développement d'un tissu nouvellement formé au sein d'un tissu normal. La tumeur est causée par une anomalie du développement cellulaire [4].

Il existe deux principales classes de tumeurs :

Chapitre 01 : contexte mammographique

- les tumeurs bénignes ;
- les tumeurs malignes ou cancers.

Contrairement aux idées reçues, le terme de tumeur est utilisé pour des pathologies cancéreuses, mais également pour d'autres productions à caractère bénin. Une tumeur maligne, est un amas de cellules cancéreuses. Une tumeur bénigne n'est pas un cancer, contrairement à une tumeur maligne. Elle se développe lentement, localement, sans produire de métastases, et ne récidive pas si elle est enlevée complètement. Un kyste peut être considéré comme une tumeur bénigne [4].

Le tableau ci-dessous exprime les critères de distinction entre les tumeurs bénignes et malignes [5].

Tumeur bénigne	Tumeurs maligne
Bien limitée.	Mal limitée.
Encapsulée.	Non encapsulée.
Histologiquement semblable au tissu d'origine.	Plus ou moins semblable au tissu d'origine.
Cellules régulières.	Cellules irrégulières (cellules cancéreuses).
Croissance lente.	Croissance rapide.
Refoulement sans destruction des tissus voisins.	Envahissement des tissus voisins.
Pas de récurrence locale après exérèse complète.	Récurrence possible après exérèse supposée totale.
Pas de métastase.	Métastase(s).

Tableau 1:: Critères de distinction entre tumeurs bénignes/ malignes

Chapitre 01 : contexte mammographique

1.4 Imagerie mammographique

1.4.1 Mammographie

Une mammographie [6] est une radiographie des seins. Elle permet d'obtenir des images de l'intérieur du sein, à l'aide de rayons X et de détecter ainsi d'éventuelles anomalies au niveau des tissus, notamment au niveau de la glande mammaire.

Elle est pratiquée par un radiologue. En fait deux radiographies par sein sont réalisées, une de face et une en oblique, ce qui permet de comparer les deux côtés de chaque sein.

Une mammographie peut être réalisée soit dans le cadre d'un dépistage du cancer du sein (mammographie de dépistage), soit en présence de symptômes (mammographie de diagnostic).

- **Mammographie de dépistage**

La mammographie est l'examen de référence de dépistage, du cancer du sein. Elle permet de mettre en évidence des cancers de petite taille, à un stade précoce, avant l'apparition de symptômes.

- **Mammographie de diagnostic**

Une mammographie est réalisée lorsqu'une patiente présente des symptômes, pouvant faire évoquer un cancer du sein comme un écoulement du mamelon, une rougeur de la peau, une « boule » dans le sein et d'autres symptômes.

Seule la mammographie est capable de détecter des micro calcifications, de minuscules dépôts de calcium dans le sein qui peuvent révéler un début de cancer. Ces lésions très petites (moins de 1 mm) sont difficiles à repérer à la palpation.

D'autres anomalies sont détectées aussi en mammographie (asymétrie de densité, distorsion architecturale...). Mais certains seins sont moins lisibles que d'autres, comme Les seins denses qui comportent peu de graisse (adolescente, jeune femme, allaitement), donnent une image plus difficile à interpréter. Dans ce cas La mammographie doit alors être complétée, par une échographie. À l'inverse, les seins riches en graisse donnent une image contrastée, qui fait ressortir les anomalies (figure 1.3).

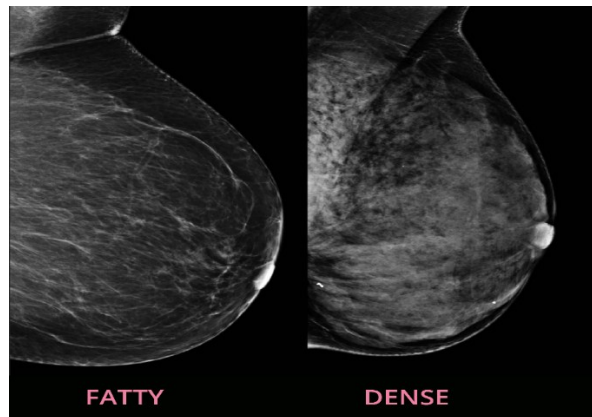


Figure 3: Sein gras et dense

1.4.2 Âge de pratique d'une mammographie

Les mammographies sont proposées à toutes les femmes entre 50 et 74 ans, sans symptômes et ne présentant aucun facteur de risque spécifique de développer un cancer du sein, autre que leur âge. En effet, les femmes âgées de 50 à 74 ans sont plus susceptibles de développer un cancer du sein [7].

Avant 50 ans, seules les femmes à risque sont invitées à subir des mammographies. Les femmes présentant des facteurs de risque (antécédents familiaux, antécédents personnels, prédisposition génétique, etc.) doivent faire un dépistage individuel du cancer du sein dès 40 ans. Il permet d'identifier les femmes effectivement porteuses d'une anomalie génétique qui augmente la probabilité de développer un cancer du sein. Si tel est le cas, un dépistage particulier ou un traitement préventif, peut leur être proposé.

1.4.3 Constitution de la salle mammographie

L'appareil de mammographie se compose d'un tube à rayons X et, d'un système de compression du sein. Le fait de comprimer les tissus du sein aide à rendre les clichés mammaires plus clairs et, ainsi de mieux voir les tissus internes du sein. La compression mammaire permet également d'émettre moins de radiations [8, 9].

Mammographie

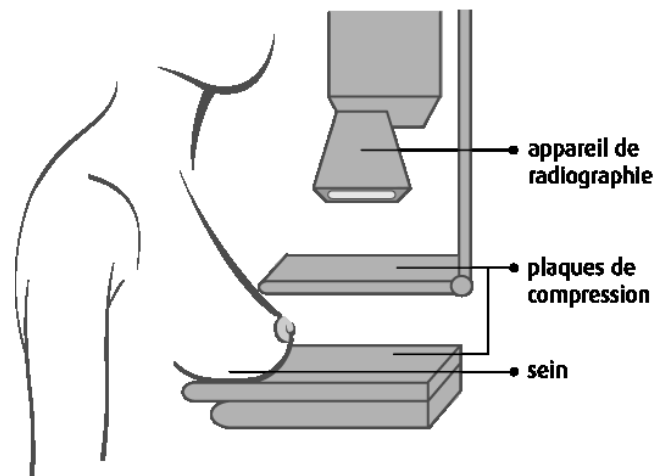


Figure 4:l'appareil de mammographie

1.4.4 Déroulement de la mammographie

Pendant l'examen le sein est placé dans la mammographie et, progressivement comprimé. Après avoir vérifié sa bonne position, le cliché radiographique est pris. La patiente ne doit pas bouger et sa respiration ne doit pas être bloquée. Le sein est décomprimé et l'opération est recommencée, pour la deuxième glande mammaire.

La mammographie de base dure en moyenne de 10 à 15 minutes. Pour obtenir une analyse de bonne qualité, chaque sein est successivement comprimé entre deux plaques pour bien l'étaler et éviter les superpositions des images. Il existe aussi une sécurité qui limite la pression maximale du sein. Ainsi, plusieurs clichés ou incidences (habituellement 2 par seins) sont réalisés sous différents angles [8].

1.4.5 Pathologies décelées par la mammographie

Les mammographies pourraient également détecter les maladies cardiovasculaires. La présence de microcalcifications dans les seins indiquerait leur présence, dans les artères également [10]. En effet, la détection de microcalcifications dans la poitrine, pourrait indiquer la présence d'autres petits dépôts de calcium dans les artères, ce qui augmente le risque cardiaque.

Chapitre 01 : contexte mammographique

Cette étude montre que la mammographie réalisée dès 50 ans pourrait fournir l'occasion d'identifier les femmes à risque cardiaque qui, normalement, n'auraient pas été soumises à un dépistage cardiovasculaire".

Chapitre 01 : contexte mammographique

1.5 Anomalies mammaires

Parmi les anomalies [11, 12], il existe les calcifications et les masses.

1.5.1 Calcifications

Les calcifications mammaires (figure 1.5) sont assez courantes et, la plupart ne sont pas associées au cancer. Afin de s'en assurer, le radiologue étudie leur taille, leur forme et leur disposition à l'aide, d'une mammographie sur laquelle, elles apparaissent souvent sous forme de petits points blancs. Certaines de leurs caractéristiques, comme une forme irrégulière ou certains regroupements, peuvent être suspects. On cite à cet effet ;

- Les macrocalcifications : ce sont des gros dépôts calcaires dans le sein qui sont associés à des affections bénignes.
- Les microcalcifications : ce sont des dépôts calcaires minuscules (<0.5 mm) qui ressemblent à des grains de sable dans le sein, facilement vue à la mammographie.

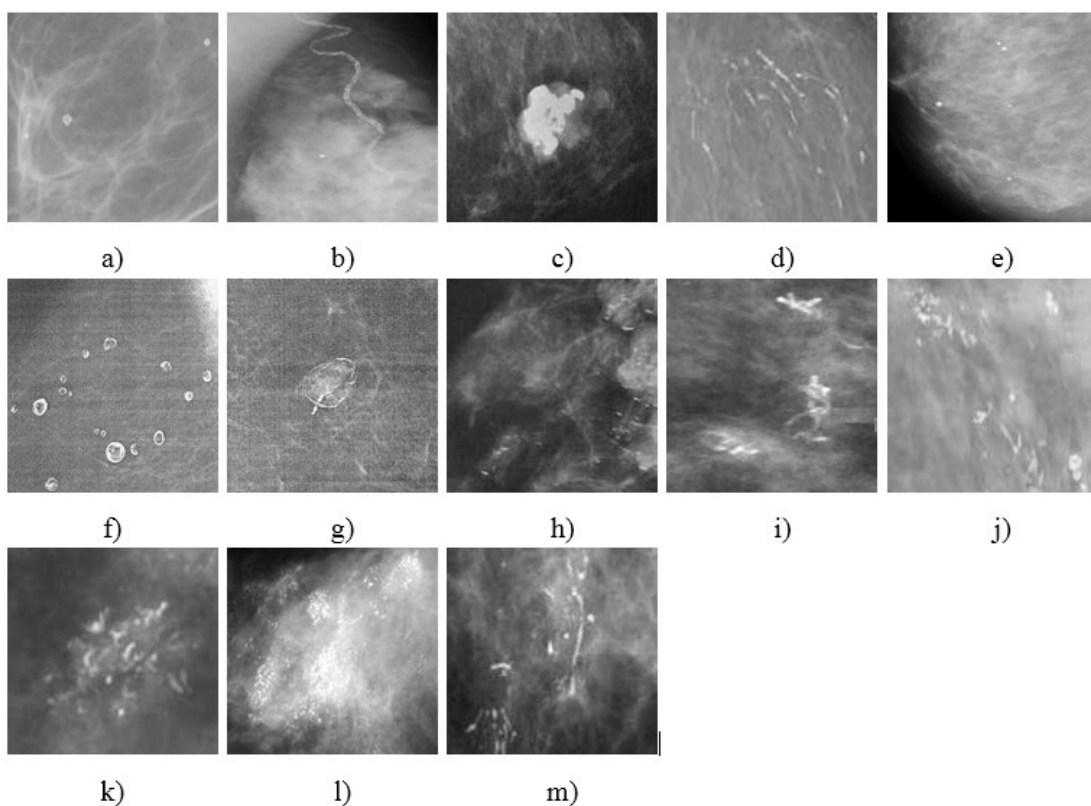


Figure 5: Différents types des microcalcifications

Chapitre 01 : contexte mammographique

- a) Microcalcifications cutanées ou dermiques : elles présentent typiquement un centre clair.
- b) Microcalcifications vasculaires : en rail ou linéaire, elles sont associées à des structures tubulaires.
- c) Microcalcifications grossières ou coralliformes : elles sont de grandes tailles diamètre supérieure à 2-3 mm.
- d) Microcalcifications en bâtonnet : elles sont généralement associées à une ectasie canalaire et sont dirigées vers le mamelon. Elles mesurent habituellement plus d'un mm de large et peuvent présenter un centre clair, si un dépôt calcique se fait dans la paroi du canal.
- e) Microcalcifications rondes : elles ont généralement une forme ronde et peuvent être de tailles variables. Lorsqu'elles mesurent moins de 0.5 mm, elles sont dites punctiformes ou pulvérulentes.
- f) Microcalcifications à centre clairs : leur taille peut s'étendre de 1 mm à plus de 1 cm. elles sont rondes ou ovales, à surface lisse et à centre clair.
- g) Microcalcifications en coquille d'œuf ou pariétale : elles sont très fines apparaissant comme des dépôts calciques, sur la surface d'une sphère. Elles mesurent généralement moins de 1 mm.
- a) Microcalcifications à type de lait calcique : elles sont sédimentées dans le fond de kystes. L'incidence de profil permet de démontrer leurs formes caractéristiques : semi-lunaires, en croissants, curvilignes ou linéaires.
- a) Microcalcifications de suture : elles correspondent à des dépôts calciques sur du matériel de suture. Ces microcalcifications sont typiquement linéaires ou tubulaires et présentent fréquemment des nœuds. Ces microcalcifications sont typiquement linéaires ou tubulaires.

Chapitre 01 : contexte mammographique

- b) Microcalcifications dystrophiques : elles mesurent habituellement plus de 0.5 mm de diamètre et, sont de formes irrégulières. Elles présentent parfois un centre clair. Ces microcalcifications sont souvent rencontrées, après un traumatisme mammaire.
- c) Microcalcifications amorphes ou indistinctes : elles sont souvent plus en moins rondes ou, en forme de flocons. Elles sont de petites tailles et généralement à contour vague, sans forme spécifique.
- d) Microcalcifications fines ou polymorphes : elles sont habituellement mieux visibles que les microcalcifications amorphes. Elles sont irrégulières de taille et de forme variables, mesurant généralement moins de 0.5 mm de diamètre.
- e) Microcalcifications linéaires et ramifiées : elles mesurent moins de 0.5 mm d'épaisseur. Elles sont irrégulières et de forme parfois linéaire, ou curviligne généralement discontinues, ou coudées.

1.5.2 Masses

Une masse (figures 1.6 et 1.7) est une opacité importante, occupant un espace dans le sein et vue, comme une tache blanche sur l'image mammographique. Il peut s'agir d'un kyste (collection liquidienne non cancéreuse) ou d'une liaison solide, qui peut correspondre à un cancer du sein.

Une opacité traduit une anomalie du tissu conjonctif ou épithéliale. Elle sera donc aisément visible au niveau d'une zone grasseuse et difficilement perceptible, dans un tissu dense.

Chapitre 01 : contexte mammographique



Figure 6 : Masse bénigne

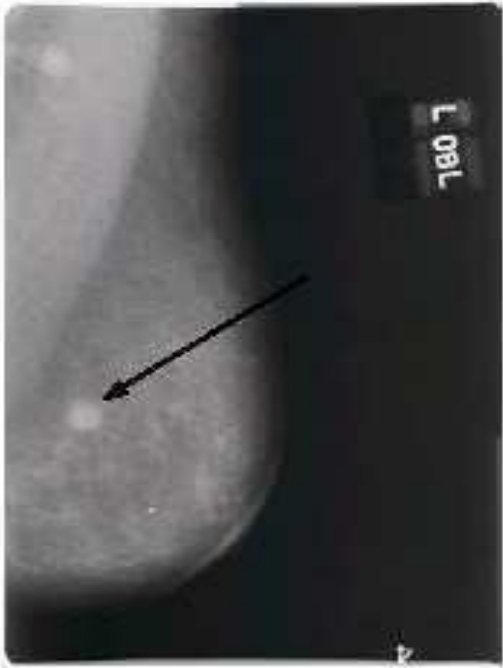


Figure 7: Masse maligne

Chapitre 01 : contexte mammographique

Différents attributs permettent aux médecins, de décrire la masse en vue de déterminer leur nature : sa forme, son contour et sa densité.

- **La forme**

Selon la description du BIRADS [11], les masses mammaires peuvent avoir la forme ronde, ovale, lobulée ou irrégulière (figure 1.8).

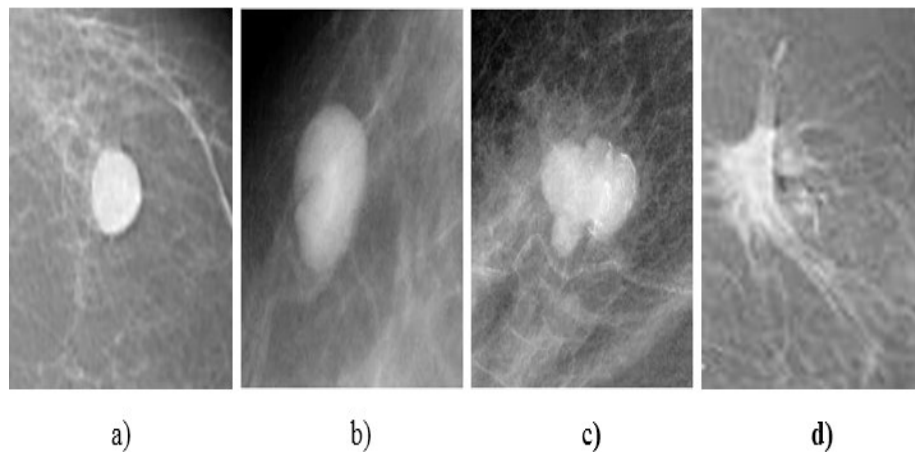


Figure 8: Formes possibles d'une masse

- a) Forme ronde : il s'agit d'une masse sphérique, circulaire ou globuleuse.
- b) Forme ovale : elle présente une forme elliptique (ou forme d'œuf).
- c) Forme lobulée : la forme de la masse présente une légère ondulation.
- d) Forme irrégulière : cette appellation est réservée aux masses dont la forme, est aléatoire et ne peut être caractérisée par les termes cités ci-dessus.

Chapitre 01 : contexte mammographique

- **Le contour**

Le contour des masses mammaires (figure 1.9) est soit : circonscrit, micro lobulé, masqué, indistinct ou spiculé.

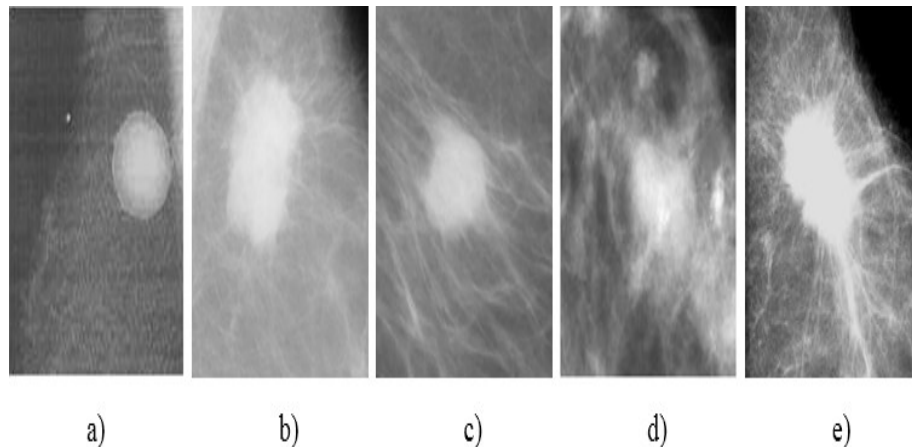


Figure 9 :Différents contours possibles d'une masse

Circonscrit b) microlobulé c) masqué d) indistinct e) spiculé

a) Circonscrit : Il s'agit d'une transition brusque entre la lésion et le tissu environnant. Le contour est alors net et bien défini. Pour qu'une masse soit qualifiée de circonscrite, il faut qu'au moins 75% de son contour, soit nettement délimité.

b) Microlobulé : Dans ce cas, de courtes dentelures du contour créent de petites ondulations.

c) Masqué : Un contour masqué est un contour qui est caché, par le tissu normal adjacent.

d) Indistinct. Dans ce cas, le contour est mal défini. Ce caractère indistinct (le contraire de circonscrit) peut correspondre à une infiltration.

e) Spiculé. La masse est caractérisée par des lignes radiaires, prenant naissance sur le contour de la masse. Ces lignes radiaires sont appelées les spicules.

Chapitre 01 : contexte mammographique

- **Densité mammaire**

Le sein est constitué de différents types de tissus : le tissu glandulaire et conjonctif ainsi que la graisse. La densité des seins décrit la quantité relative des différents tissus, dans le sein. Les seins denses présentent moins de tissu graisseux, plus de tissu conjonctif et glandulaire.

La recherche montre que le risque de présenter un cancer du sein, est 4 à 6 fois plus élevé, chez les femmes qui ont des seins denses. En effet, le tissu mammaire change au fur et à mesure que nous vieillissons, il devient généralement moins dense avec l'âge.

Le risque d'avoir un cancer en raison du caractère hétérogène du tissu fibre-glandulaire augmente, lors du passage de la catégorie 1 à 4 (figure 1.10).

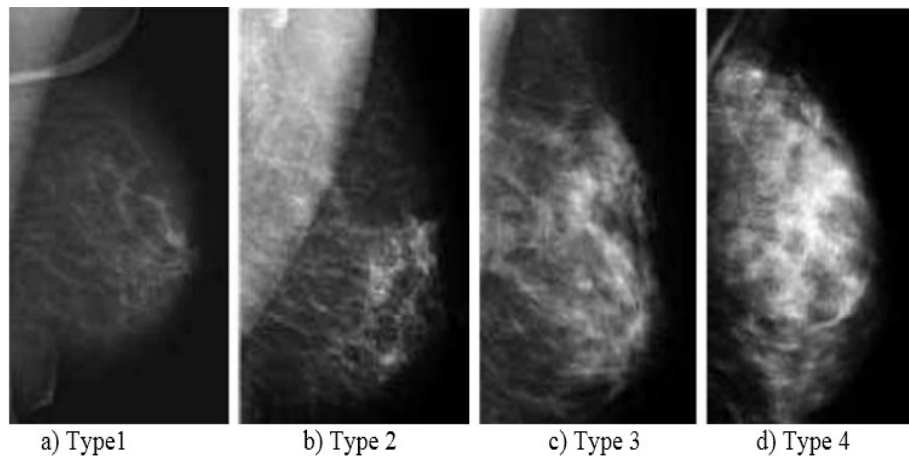


Figure 10: Différents types de densités mammaires

- ✓ Type 1 : le sein est presque entièrement graisseux : sein clair.
- ✓ Type 2 : il y a des opacités fibroglandulaires.
- ✓ Type 3 : le tissu mammaire est dense et hétérogène.
- ✓ Type 4 : le tissu mammaire est extrêmement dense.

Chapitre 01 : contexte mammographique

1.6 Classification ACR

Le système de l'American College of Radiology (ACR) utilise la classification BIRADS, pour classer les mammographies en 6 catégories [13].

- ACR 0 : classification d'attente. Des investigations complémentaires sont nécessaires.
- ACR 1 : mammographie normale.
- ACR 2 : il existe des anomalies bénignes (c'est-à-dire sans gravité) qui ne nécessitent ni surveillance, ni examen complémentaire.
- ACR 3 : il existe une anomalie probablement bénigne, pour laquelle une surveillance à court terme (3 ou 6 mois) est conseillée.
- ACR 4 : il existe une anomalie indéterminée ou suspecte.
- ACR 5 : il existe une anomalie évocatrice d'un cancer.

En cas d'images ACR 4 ou ACR 5, des prélèvements par biopsie percutanée sont nécessaires. Les mammographies classées 1 et 2 sont revues, en deuxième lecture par un second radiologue.

Chapitre 01 : contexte mammographique

1.7 Avantages et les inconvénients de la mammographie

La mammographie présente des avantages comme des inconvénients [14].

1.7.1 Avantages

Les avantages pouvant être générés par la mammographie, sont cités ci-dessous :

✓ *La réduction de la mortalité*

La réduction de la mortalité est la raison d'être, d'un programme de dépistage. Les données statistiques après 10 ans de fonctionnement du programme, indiquent une réduction de mortalité de l'ordre de 35 % chez les femmes participant au dépistage du cancer du sein.

✓ *Dépistage précoce*

Un cancer du sein peut être découvert plus tôt grâce au dépistage ce qui est un facteur important pour contribuer à la diminution de la mortalité. Les cancers trouvés au dépistage sont généralement plus petits que ceux trouvés à l'examen physique.

✓ *Chirurgie conservatrice*

Le dépistage précoce du cancer du sein, permet d'avoir une chirurgie moins étendue, permettant de conserver le sein. Il réduit aussi, le risque d'avoir de la chimiothérapie puisque le cancer mis en évidence, étant plus petit.

✓ *Dépistage un gain de temps sur la maladie*

Le dépistage permet de repérer des cancers de plus petite taille et moins évolués, avant qu'ils ne soient palpables. Ceci engendre :

- De réelles chances de guérison.
- Des traitements moins lourds et moins agressifs, avec moins de séquelles (le sein peut être conservé plutôt que retiré, par exemple).
- Un suivi régulier.
- Un examen pris en charge à 100% par l'assurance maladie.

Chapitre 01 : contexte mammographique

1.7.2 Inconvénients

Les inconvénients se rapportent à ne pas détecter uniquement, les vrais positifs. Des faux négatifs et positifs, peuvent apparaître lors d'une mammographie.

✓ *Les faux négatifs*

La mammographie n'est pas infaillible. Certains cancers, peuvent passer inaperçus. 10% des signes perçus, sont des faux négatifs.

✓ *Les faux positifs*

La mammographie demandant une investigation supplémentaire qui ne révélera pas de cancers considérés, comme des faux positifs (10 % des dépistages).

Les résultats faussement positifs ont plusieurs conséquences : de l'anxiété, une augmentation du nombre d'examens additionnels et requis, ainsi qu'une diminution de la participation des femmes aux dépistages subséquents.

✓ *L'exposition aux rayons x*

Comme toute radiographie, la mammographie expose la patiente, à des rayons X. Ceux-ci, s'ils sont répétés, peuvent conduire à l'apparition d'un cancer, que l'on appelle cancer radio-induits. C'est l'une des raisons pour lesquelles, le dépistage est recommandé uniquement tous les deux ans, à partir de 50 ans si la femme n'a pas de de symptôme ou, de facteurs de risque. Par ailleurs, après 50 ans, la composition des seins se modifie et les doses de rayons nécessaires à la mammographie, sont plus faibles. Cependant Etant utilisés à de très faibles doses lors de la mammographie, l'examen est considéré sans danger.

Chapitre 01 : contexte mammographique

1.8 Conclusion

Dans ce chapitre nous avons abordé des notions générales sur le cancer du sein et la mammographie. Après avoir présenté de façon concise l'anatomie du sein, l'objectif qui sous-tend la présentation de ces éléments, c'est de mettre en lumière une technique de détection d'un cancer ainsi qu'une sensibilisation au dépistage précoce, car vaut mieux prévenir que guérir.

Dans le prochain chapitre nous effectuons un parcours bibliographique sur les concepts liés à la segmentation par classification ainsi qu'aux techniques d'optimisation, pour l'obtention des meilleures performances possibles de détection des masses.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

Chapitre 2 : Techniques de la segmentation par classification et les algorithmes bio-inspirés

2.1 Introduction

L'imagerie médicale est l'un des domaines de la médecine qui a le plus progressé, ces 20 dernières années. Elle est le procédé par lequel, un médecin peut examiner l'intérieur du corps d'un patient sans l'opérer ; en créant une image visuelle compréhensible d'une information à caractère médical. L'image produite doit être partitionnée en régions connexe et homogènes, pour la représentation de l'information pertinente. Cette étape est très importante dans tout le processus d'analyse d'image.

Nous abordons dans ce chapitre, les techniques de la segmentation par classification, pour étudier par la suite les métaheuristiques bio-inspirées et, leurs applications à la segmentation d'images.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.2 Introduction à la segmentation

La segmentation est un traitement de bas-niveau, qui consiste à effectuer une partition de l'image en régions homogènes, par rapport à un ou plusieurs critères. Les régions obtenues se distinguent les unes des autres, par des différences significatives selon ces mêmes critères [15,16].

La segmentation consiste à extraire des lignes ou des régions. Le choix d'une technique de la segmentation est lié à plusieurs facteurs comme :

- la nature de l'image ;
 - éclairage non homogène, reflets,
 - présence de bruit, de zones texturées,
 - contours flous, en partie occultés.
- Les opérations situées en aval de la segmentation ;
 - reconnaissance des formes, interprétation,
 - diagnostic, contrôle qualité.
- Les primitives à extraire ;
 - contours, segments de droite, angles,
 - régions, formes,
 - textures.
- Les contraintes d'exploitation ;
 - complexité algorithmique, fonctionnement en temps réel,
 - taille de la mémoire disponible en machine.

On regroupe généralement les algorithmes de segmentation [15, 16] en trois grandes classes :

- la segmentation en régions ;
- la segmentation en contours ;
- la segmentation hybride.

L'objectif de la détection de contours, est la détermination des points de l'image, séparant des zones de caractéristiques différentes.

De manière duale, le but de la segmentation en régions, est de partitionner l'image en ensembles de points connexes, possédant des propriétés d'homogénéité. Dans ce cadre, les

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

méthodes existantes peuvent être rangées en deux classes, les procédés de « classification », et ceux de la croissance de régions ». La différence principale entre ces deux types d'approches, réside dans la stratégie d'utilisation de l'espace des luminances et des relations spatiales existant entre les pixels. Les méthodes de classification déterminent d'abord une partition de l'espace des luminances et, se servent ensuite des relations de connexité pour déterminer les régions.

La segmentation idéale, n'existe pas. La bonne technique, est celle qui permet d'aboutir à une interprétation efficace.

Les méthodes fixées par les objectifs de ce travail, se rapportent à la segmentation par classification.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.3 Segmentation par classification

La segmentation par classification permet une partition de données, en regroupant les pixels formant l'image à classifier, de telle sorte que les pixels appartenant à une classe, soient plus similaires entre eux que ceux des classes différentes.

Chaque entité de la scène traitée, se voit attribuer une étiquette indiquant son appartenance à une classe particulière. L'entité utilisée caractérise généralement un pixel ou un ensemble de pixel, tandis que l'étiquette constitue un thème choisi par l'utilisateur [16].

2.3.1 Introduction au Clustering

2.3.1.1 Définition

Le clustering est une étude non supervisée, visant à organiser un ensemble d'objets (physiques ou abstraits) en groupes ou clusters, de façon à avoir des objets similaires groupés et, les objets différents organisés dans des groupes différents. C'est l'une des étapes les plus importantes de l'analyse exploratoire des données [17].

2.3.1.2 Similarité entre les objets

La qualité d'un regroupement dépend donc de la mesure de similarité utilisée, par la méthode et de son implémentation [18]. Parmi ces mesures [18], nous pouvons citer le calcul des distances suivantes :

1. Distance de Minkowski

Elle porte le nom du mathématicien allemand Hermann Minkowski. Sa formule est donnée par l'équation 2.1.

$$d(i,j) = \sqrt[q]{(|x_{i1} - x_{j1}|^q + |x_{i2} - x_{j2}|^q + \dots + |x_{ip} - x_{jp}|^q)} \quad 2.1$$

Où $i = (x_{i1}, x_{i2}, \dots, x_{ip})$ et $j = (x_{j1}, x_{j2}, \dots, x_{jp})$ sont deux objets p -dimensionnels et q un entier positif.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2. Distance de Manhattan

Dans ce cas, si $q = 1$, d est la distance de Manhattan. Elle est représentée par l'équation 2.2.

$$d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}| \quad 2.2$$

3. Distance euclidienne

Dans ce cas, si $q = 2$, d est la distance Euclidienne. En effet, la distance euclidienne représente la distance entre deux points dans l'espace euclidien (équation 2.3). Avec cette distance, l'espace euclidien devient un espace métrique.

$$d(i, j) = \sqrt{(|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2)} \quad 2.3$$

2.3.2 Méthodologies du clustering

Deux méthodologies générales du clustering, sont données ; les algorithmes hiérarchiques et les algorithmes de partitionnement (k- means, FCM...) [19, 20, 21, 22].

2.3.2.1 Clustering hiérarchique

Dans un clustering hiérarchique [19], un cluster peut être divisé en sous clusters L'ensemble des clusters étant généralement, représenté par un arbre (figure 2.1).

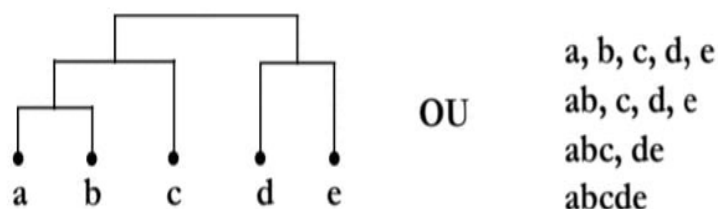


Figure 11 :: Exemple d'un clustering hiérarchique

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

La méthode hiérarchique consiste à effectuer, une suite de partitions emboîtées. Il existe deux types d'approches de clustering hiérarchique : les approches par agglomération et les approches par division [20].

2.3.2. 2 Partitionnement

Le partitionnement de données (ou data clustering en anglais), est une méthode en analyse des données [21]. Elle vise à diviser un ensemble de données en différents « paquets » homogènes, en ce sens que les données de chaque sous-ensemble, partagent des caractéristiques communes (figure 2.2).

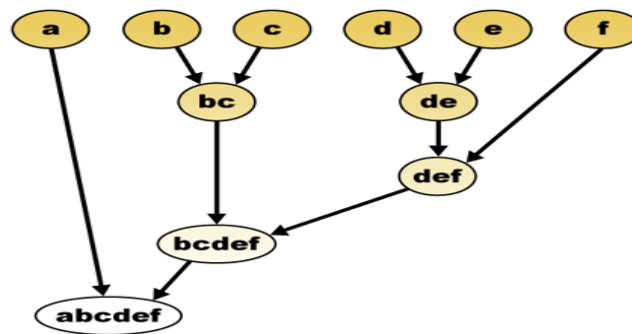


Figure 12 :Exemple d'un clustering de partitionnement

Le but principal de cette méthode, est de partitionner un ensemble X de N objets $\{x_1, x_2, \dots, x_N\}$ en k classes (k fixé par l'utilisateur). Dans cette méthodologie, même pour un nombre de classes et de pixels très réduits, le nombre de combinaison explose considérablement. Il n'est pas possible alors, d'énumérer complètement les partitions de l'image pour déterminer exactement la valeur optimale du critère de partitionnement. Une solution possible est d'utiliser des heuristiques, qui permettent de calculer une valeur approchée de la valeur optimale. Par exemple :

- K-means : chaque cluster est représenté par son centre de gravité.
- Fuzzy C-means : c'est une version floue du K-means.

Il a été montré que c'est un problème d'optimisation difficile. Pour trouver la partition qui minimise au mieux le critère de partitionnement, il faut énumérer toutes les partitions possibles de N pixels en K classes et choisir parmi elles, celle pour qui ce critère est minimal.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

Le nombre des partitions possibles noté $S(N, K)$ est donné par la formule suivante :

$$S(N, K) = \frac{1}{k!} \sum_{i=1}^k (-1)^{k-i} \binom{k}{i} i^N \quad 2.4$$

La plupart des heuristiques du partitionnement minimisent l'inertie intraclasse, de la partition I_{intra} (équation 2.5) où bien maximisent l'inertie interclasse de la partition I_{inter} (équation 2.6) suivant un nombre de classes k fixé. L'inertie totale I est définie comme suit :

$$I_{intra} = \sum_{j=1}^k \sum_{i=1}^{|c_j|} d^2(x_i^j, \mu_j) \quad 2.5$$

$$I_{inter} = \sum_{j=1}^k |c_j| d^2(\mu_j, \mu) \quad 2.6$$

D'où:

$$I_{intra} + I_{inter} = I \quad 2.7$$

Avec :

x_i^j : est le vecteur caractéristique du i ème pixel de la classe c_j .

μ_j : est le vecteur de caractéristiques moyen de la classe c_j .

c_j : est la taille de la classe c_j .

μ : est le vecteur moyen de caractéristiques de l'image à segmenter.

$d(.)$ est une mesure de distance (la plus connue est la distance euclidienne).

Les méthodes de classification par partitionnement, dépendent d'un processus itératif qui converge vers une partition qui optimise un critère prédéfini. Nous décrivons dans les sections suivantes, l'approche des K-means ainsi que celle des FCM.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.4 Clustering classique

Le terme « *k*-means » a été utilisé pour la première fois par Ames MacQueen en 1967, L'algorithme classique a été proposé par Stuart Lloyd en 1957 et mis en œuvre sous sa forme actuelle par E. Forgy en 1965 [22].

La méthode des *k*-means est un outil de classification non supervisée, qui permet de répartir un ensemble de données en *K* classes homogènes [23, 24]. Dans ce cas, l'intérêt généralement est de partitionner l'espace en classes concentrées et isolées les unes des autres. Dans cette optique, l'algorithme des *k*-means vise à minimiser la variance intra-classe, qui se traduit par la minimisation de la fonction objectif suivante :

$$J = \sum_{k=1}^C \sum_{i=1}^N u_{ki} \|o_i - z_k\|^2 \quad 2.8$$

Dans cet algorithme, un pixel o_i est attribué à la classe la plus proche parmi les *C* clusters.

Avec :

- c : Le nombre de classes (clusters)
- N : Une image avec *N* pixels $o_i = \{o_1, o_2, \dots, o_N\}$
- o_i : représente le vecteur caractéristique du pixel *i*
- z_k : Le *k*-ième centre de classe
- $\| \cdot \|^2$ désigne la distance euclidienne

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

Les solutions au problème s'écrivent :

$$\mathbf{u}_{ki} = \begin{cases} 1 & \text{si } \|o_i - Z_k\|^2 < \|o_i - Z_j\|^2 \forall k \neq j \\ 0 & \text{sinon} \end{cases} \quad 2.9$$

$$\mathbf{Z}_k = \frac{\sum_{i=1}^N \mathbf{u}_{ki} o_i}{\sum_{i=1}^N \mathbf{u}_{ki}} \quad 2.10$$

2.4.1 Etapes du K-means

En fixant le nombre de classes C et le critère d'arrêt ε l'approche des K-means suit les étapes ci-dessous :

1. Initialisation aléatoire du vecteur des centroïdes Z (centres) :

$$\mathbf{Z}_k = [\mathbf{z}_k^1, \mathbf{z}_k^2, \dots, \mathbf{z}_k^j, \dots, \mathbf{z}_k^p], \quad 1 < k < c$$

2. Calculer la distance de chaque pixel o_i aux centroïdes

3. Affecter chaque pixel à la classe la plus proche (suivant la distance minimale) (équation 2.9)

4. Mettre à jour le vecteur des centroïdes (centre de gravité) (équation 2.10)

5. Itérer les étapes 2 et 3. Jusqu'à $J^{(t)} - J^{(t+1)} < \varepsilon$; jusqu'à stabilité des centroïdes (minimisation de l'équation 2.8).

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.4.2 Fonctionnalités du k-means

Cette section introduit la paramétrisation des k-means, ainsi que ses avantages et ses inconvénients.

➤ **Choix du nombre k de clusters**

L'algorithme classique des k -means laisse un paramètre libre : le nombre de clusters. Ce qui dans le cas de la segmentation d'images, correspond au nombre d'intensités (pixel $(i, j, \text{intensité})$) utilisées pour représenter l'image. En général, k est choisi expérimentalement en spécifiant une valeur, qui réduit l'énergie.

➤ **Avantages de l'algorithme [25]**

Les k-means présentent l'avantage, d'avoir un faible coût de calcul et d'être facile à implémenter.

➤ **Inconvénients de l'algorithme [26]**

Un inconvénient possible des k-moyennes pour le partitionnement, est que les clusters dépendent de l'initialisation et de la distance choisie.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.5 Approche floue du clustering

La partition floue est basée sur la théorie des ensembles flous défini par Zadeh [27, 28], qui décrit la manière dont sont structurés les ensembles flous et, quelles sont les opérations permises sur ces ensembles. En traitement d'images, les méthodes de classification floue modélisent le degré d'appartenance d'un pixel à chaque classe, sous l'hypothèse que ce pixel puisse ne pas appartenir à une classe unique. Chacun des N pixels appartient à chacune des C classes avec un coefficient d'appartenance u .

L'ensemble des degrés d'appartenance est stocké dans la matrice de la C-partition floue U (équation 2.11).

$$U_{n \times c} = [u_{ij}]_{i=1 \dots N, j=1 \dots C} \quad 2.11$$

Où :

u_{ij} est le degré d'appartenance de pixel i à la classe j .

Cette propriété est exploitée en traitement d'images et plus précisément, en classification où les classes, appelées aussi régions, sont représentées par des ensembles flous. Cela est fort utile, lorsque les régions ne peuvent pas être définies de manière nette et précise.

Leur manipulation en gardant le caractère flou, permet de traiter des données imprécises, incertaines et/ou redondantes d'une manière plus flexible.

Pour calculer u_{ij} on utilise cette équation :

$$u_{ij} = \left(\sum_{k=1}^c \left(\left\| \frac{x_i - v_j}{x_i - v_k} \right\| \right)^{2/(m-1)} \right)^{-1} \quad 2.12$$

Avec : $\forall i = 1, \dots, n$ et $\forall j = 1, \dots, c$

Le paramètre m « fuzzifier » ($m \geq 1$) contrôle le degré de « fuzzification » du processus. Plus il est élevé, plus les appartenances aux classes sont lissées, plus m est grand plus la partition sera floue.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

L'objectif de l'algorithme *FCM* est de trouver U qui minimise la fonction 'objectif' suivante :

$$J_{FCM}(U, V) = \sum_{i=1}^n \sum_{j=1}^c (\mu_{ij})^m * \|x_i - v_j\|^2 \quad 2.13$$

Où :

- μ_{ij} est le degré d'appartenance flou.
- $\|x_i - v_j\|^2$ est la distance euclidienne entre le nœud c et le centre du cluster n .
- c : est le nombre des classes.
- n : est le nombre de pixels.
- m : est un nombre réel > 1 .

2.5.1 Fonctionnement du FCM

L'algorithme des c -moyens flous (fuzzy c -means) est un algorithme de classification floue fondé sur l'optimisation d'un critère quadratique de classification, où chaque classe est représentée par son centre de gravité. L'algorithme nécessite de connaître le nombre de classes au préalable et génère les classes par un processus itératif en minimisant une fonction 'objectif' [29].

L'algorithme FCM peut être exécuté par les étapes suivantes après avoir fixé le nombre de classes, le paramètre de flou m et le critère d'arrêt ε .

1. Initialiser la position des centres décrits par :

$$V = [v_1, v_2 \dots v_c]$$

2. Mettre à jour la matrice μ_{ij} . Le réajustement de la matrice d'appartenance en fonction de la position des centres (la partition floue initiale U) se fait selon l'équation (2.12).

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

Cette étape permet de calculer pour chaque pixel, ses degrés d'appartenances aux classes, avec la comparaison entre sa distance à v_j et sa distance aux autres barycentres conditionnels (v_k , $k= 1...c$).

3. Calculer les barycentres (centre de gravité) des différentes régions, en tenant compte des degrés d'appartenance et en utilisant cette formule :

$$v_j = \frac{\sum_{i=1}^N (u_{ij})^m x_i}{\sum_{i=1}^N (u_{ij})^m} \text{ avec } 1 < j < c \quad // \text{ mise à jour du vecteur des barycentres}$$

4. Calcul des distances entre les pixels traitants et les barycentres puis Affecter ces pixels aux classes en considérant la distance minimale.
5. Calculer la fonction 'objectif' par la relation (2.13).
6. Convergence du FCM. Autrement dit, la stabilité est réalisée, lorsque la matrice d'appartenance aux classes des individus Ω n'est plus modifiée. Sinon retourner en 2) jusqu'à ce que la valeur J minimale soit atteinte.

Cet algorithme utilise un critère de minimisation des distances intra-classes et de maximisation des distances interclasses, en tenant compte des degrés d'appartenance et en optimisant la fonction J.

➤ Avantages et inconvénients du FCM

L'approche des FCM, est une méthode de segmentation par clustering, qui est convergente.

La plupart des méthodes de classification utilisées, y compris le FCM, souffrent de trois défauts :

- ✓ sensibilité moindre au bruit ;
- ✓ le nombre de clusters doit être fourni à l'avance ;
- ✓ complexité du calcul.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.5.2 Différence entre les k means et le fuzzy c means

La différence entre les méthodes fuzzy c means et k means [30] dans la fonction d'appartenance d'un nuage de points à deux clusters, peut être résumée dans l'exemple de la figure 2.3.

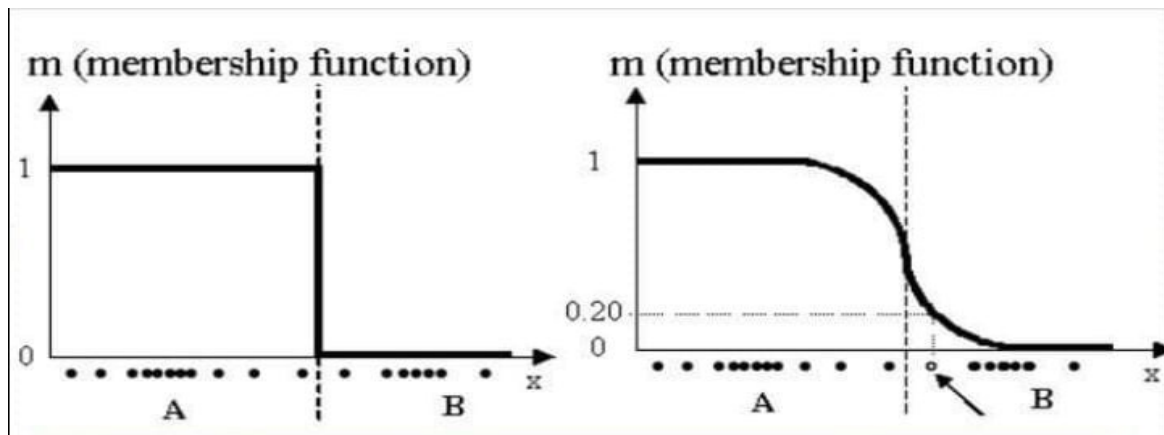


Figure 13: Fonction d'appartenance dans k means / fuzzy c means

La figure de gauche représente la fonction d'appartenance pour les k-means et celle de la droite, pour l'algorithme fuzzy c means. L'axe vertical montre la probabilité d'appartenance. Dans les k means un objet ne peut pas appartenir à deux clusters simultanément, ce qui explique la discrimination binaire entre les clusters. Mais dans les FCM il est possible qu'un objet, appartienne à deux ou plusieurs clusters selon des pourcentages différents. En effet, dans ce cas, les données sont liées à chaque cluster, par le biais d'une fonction d'appartenance, ce qui représente le comportement flou de cet algorithme.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.6 Introduction aux techniques bio inspirées

2.6.1 Introduction à l'optimisation

L'optimisation [23, 31] d'une fonction $f(s)$ où $s \in S$ (S est un ensemble de solution), consiste à trouver une solution $s_0 \in S$, dit une solution optimale, telle que $\forall s \in S$:

- Un problème de minimisation $f(s_0) \leq f(s)$.
- Un problème de maximisation $f(s_0) \geq f(s)$.

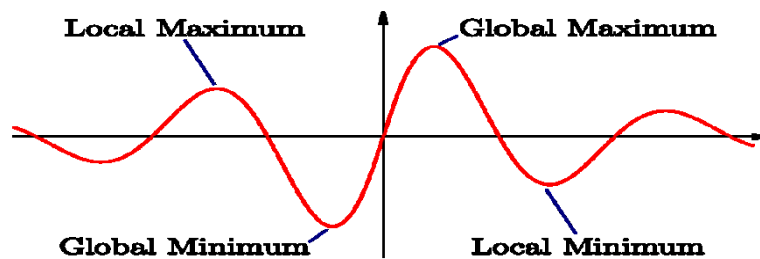


Figure 14: Changement de la fonction objectif pour un problème de minimisation

f est la fonction 'objectif' ou fonction fitness, est une mesure quantitative de la qualité des solutions. Chaque problème de minimisation peut être transformé, en un problème de maximisation et inversement.

Il existe deux types de minimum :

- Minimum local : la solution s est un minimum local dans un voisinage N si :

$$\forall s' \in N(s), f(s) \leq f(s')$$

- Minimum global : la solution est un minimum global, si :

$$\forall s' \in S, f(s) \leq f(s')$$

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.6.2 Catégories des méthodes d'optimisation

Un problème d'optimisation est un ensemble de variables, contraintes et une fonction 'objectif' qui doit être minimisée/maximisée. En effet, une fonction 'objectif' est utilisée pour déterminer la meilleure solution, à un problème d'optimisation. Le but du problème d'optimisation, est alors de minimiser ou de maximiser cette fonction jusqu'à l'optimum, par différents procédés [31]. Chaque problème d'optimisation possède un espace de recherche, qui représente l'ensemble des solutions possibles.

➤ **Les méthodes d'optimisation**

Il existe deux méthodes d'optimisation déterministes et non-déterministes.

➤ **Méthodes déterministes**

Ces méthodes sont généralement approuvées, lorsque la fonction 'objectif' du problème, présente un certain nombre de caractéristiques, telles que la convexité, la continuité ou le contraste.

➤ **Méthodes non-déterministes (méta heuristiques)**

Ces méthodes sont limitées. Par exemple, on y observe le manque de convexité, la non-dérivabilité, la discontinuité ou même, la difficulté à déterminer avec précision la fonction cible.

Pour résoudre ces problèmes, il est nécessaire de recourir à des méthodes approximatives également appelées méthodes stochastiques (aléatoires) qui utilisent des stratégies, afin de fournir des résultats encourageants.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

➤ **Paradigme des techniques bio inspirées**

La recherche d'algorithmes bio-inspirés [23] consiste à étudier et à imiter des phénomènes, des fonctions et des modèles de différents organismes vivants ; afin de développer de nouvelles solutions dans de nombreux domaines d'application (industrie, informatique, médecine, etc.). Cette source d'inspiration permet de manipuler des données inexacts ou incomplètes, pour atteindre un plus large éventail de solutions.

- **Définition**

La bio-inspiration [32] est un changement de paradigme, qui conduit des concepteurs à s'inspirer de la nature pour développer de nouveaux systèmes. La bio-inspiration s'appuie souvent sur le bio mimétisme, elle peut s'inspirer du monde des plantes, des animaux, des champignons ou des bactéries et virus.

Elle a déjà contribué à des applications dans divers domaines tels que l'aéronautique, la pharmacie, la marine, la médecine, la chimie verte, les matériaux composites, la robotique, l'intelligence artificielle et les nanotechnologies. Il existe plusieurs techniques de bio inspiré, nous en étudions celles qui nous intéressent dans ce mémoire.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.7 Algorithmes génétiques

Les algorithmes génétiques (AG) sont des algorithmes de recherche, basés sur les mécanismes de la sélection naturelle et de la génétique. Ils appartiennent à la famille des algorithmes évolutionnistes.

Ils combinent une stratégie de survie des plus forts, avec un échange d'information aléatoire mais structurée. Leur but est d'obtenir une solution approchée à un problème d'optimisation, pour le résoudre en un temps raisonnable. Cet algorithme est développé par John Holland [33].

2.7.1 Fonctionnement

Les algorithmes génétiques (figure 2.5) utilisent la théorie de Darwin [34] sur l'évolution des espèces. Elle repose sur trois principes :

- Le principe de variation : chaque individu au sein d'une population, est unique.
- Principe d'adaptation : les individus mieux adaptés à leur environnement, atteignent plus facilement l'âge adulte. Ainsi, ceux qui ont une meilleure capacité à survivre pourront donc, se reproduire davantage.
- Le principe d'hérédité : les caractéristiques des individus doivent être héréditaires, pour être transmises à leur descendance. Ce mécanisme permettra de faire évoluer l'espèce, pour partager les propriétés bénéfiques de sa survie.

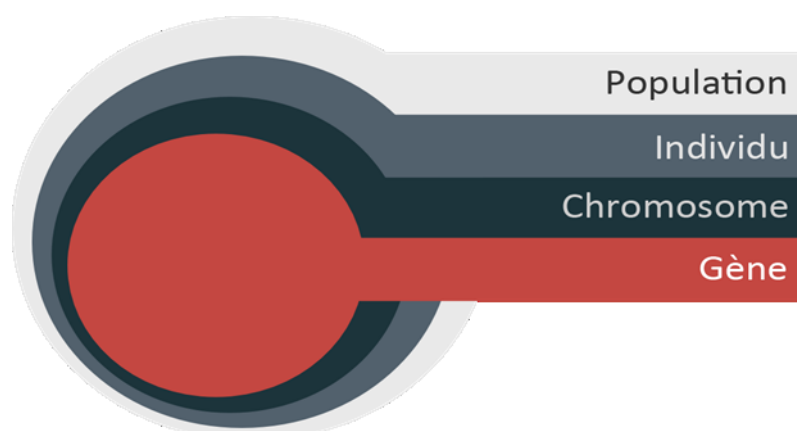


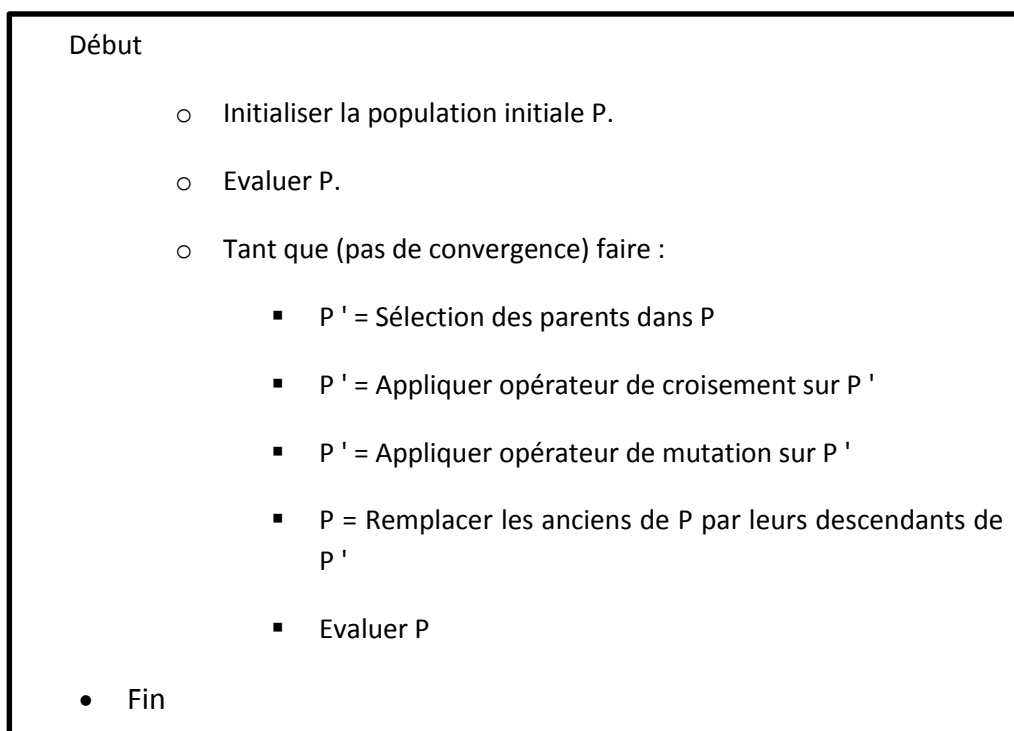
Figure 15: Paradigme de l'algorithme génétique

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

Un algorithme génétique générique [35] comprend les étapes suivantes :

Telles que :

- la population est l'ensemble des solutions envisageables ;
- l'individu représente une solution ;
- le chromosome est une composante de la solution ;
- le gène est une caractéristique, une particularité de l'individu.



2.7.2 Opérateurs d'évolution

Il y a trois opérateurs d'évolution dans les algorithmes génétiques.

- La sélection : choix des individus les mieux adaptés.
- Le croisement : Mélange par la reproduction des particularités des individus choisis.
- La mutation : Altération aléatoire des particularités d'un individu.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

➤ Sélection

La sélection consiste à choisir les individus les mieux adaptés, afin d'avoir une population qui converge vers l'optimum global.

Il existe plusieurs techniques de sélection dont les principales utilisées :

- sélection par rang : elle se concentre toujours sur la sélection des individus, qui possédant les meilleurs scores d'adaptation.
- Probabilité de sélection proportionnelle à l'adaptation : il s'agit de la technique de la roulette ou roue de la fortune, où pour chaque individu, la probabilité d'être sélectionné est proportionnelle à son adaptation au problème.
- Sélection par tournoi : l'utilisation de la sélection proportionnelle sur des paires d'individus, choisit parmi ces paires, l'individu qui a le meilleur score d'adaptation.
- Sélection uniforme : la sélection se fait aléatoirement, uniformément et sans intervention de la valeur d'adaptation.

➤ Croisement

Le croisement, ou enjambement, 'crossing-over', est le résultat obtenu lorsque deux chromosomes partagent leurs particularités. Il permet le brassage génétique de la population et l'application du principe d'hérédité, de la théorie de Darwin.

Il existe deux méthodes de croisement : simple ou double enjambement.

- Le croisement simple : consiste à combiner les particularités de deux individus, à partir d'un pivot, afin d'avoir un ou deux enfants.
- Le croisement double : est basé sur le même principe, sauf qu'il y a deux pivots.

➤ Mutation

La mutation consiste à altérer un gène dans un chromosome, selon un facteur de mutation. Ce facteur est la probabilité qu'une mutation, soit effectuée sur un individu. Cet opérateur est l'application du principe de variation, de la théorie de Darwin qui permet,

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

par la même occasion, d'éviter une convergence prématurée de l'algorithme vers un extremum local.

2.7.3 Avantages et inconvénients des algorithmes génétiques

Les algorithmes génétiques présentent des avantages comme des inconvénients [36].

➤ Avantages

- Les AG fournissent d'excellentes performances à de faibles coûts.
- Les AG permettent d'explorer des domaines possédant de très nombreuses solutions.

➤ Inconvénients

- il n'y a pas de garantie quant à l'obtention de la solution optimale, au problème posé en un temps fini.
- L'utilisation de ces algorithmes, est souvent coûteuse en temps de calculs.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.8 Optimisation par essaim de particules (PSO)

L'Optimisation par Essaim de Particules, connue sous le nom de 'Particle Swarm Optimization' (PSO), est un algorithme évolutionnaire. Il a été proposé par Russel Eberhart (ingénieur en électricité) et James Kennedy (socio psychologue) en 1995 [23, 37].

Cette méthode trouve sa source dans les observations faites, lors des simulations informatiques de vols groupés d'oiseaux et de bancs de poissons. Autrement-dit, elle s'inspire fortement de l'observation des relations grégaires d'oiseaux migrateurs, qui pour parcourir de « longues distances » (migration, quête de nourriture, parades aériennes, etc.), doivent optimiser leurs déplacements en termes d'énergie dépensée, de temps, (etc.), comme par exemple la formation en forme de V présentée dans la figure 2.6.



Figure 16: Comportement des oiseaux en formation en V

Le mouvement de la PSO est complexe. Sa dynamique est soumise à des règles et des facteurs très spécifiques qu'il s'agit de cerner :

- Chaque individu dispose d'une certaine intelligence « limitée » (qui lui permet de prendre une décision).
- L'individu doit connaître sa position locale et disposer d'information locale de chaque individu, se trouvant dans son voisinage.
- Obéir à ces trois règles simples, « rester proche des autres individus », « aller dans une même direction » ou « voler à la même vitesse ».

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.8.1 Mode de fonctionnement de l’algorithme PSO

La population dans l’algorithme PSO est nommée essaim. Chaque individu du groupe est dit particule, où le déplacement d’une particule (figure 2.7), est influencé par les trois composantes suivantes :

1. Une composante d’inertie : la particule tend à suivre sa direction courante de déplacement.
2. Une composante cognitive : la particule tend à se diriger vers le meilleur site, par lequel elle est déjà passée.
3. Une composante sociale : la particule tend à se fier à l’expérience de ses congénères et, ainsi, à se diriger vers le meilleur site déjà atteint par ses voisins.

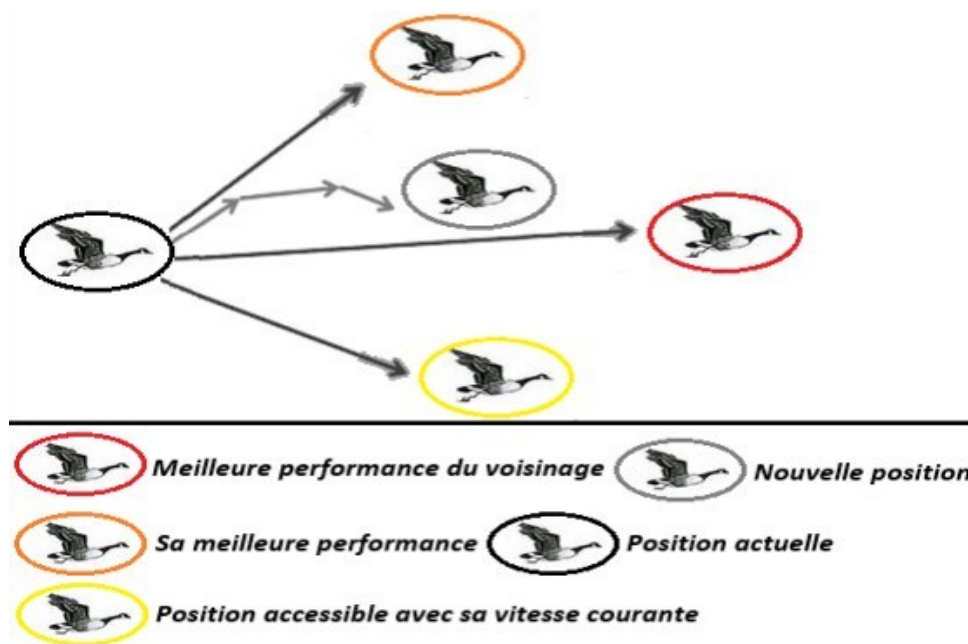


Figure 17: Schéma de principe du déplacement d’une particule

L’algorithme est résumé par les étapes suivantes :

- Dans l’espace de recherche de dimension D , une particule i de l’essaim est représentée par son vecteur position et par son vecteur vitesse ; formulés ainsi :

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

$$x_i = (x_{i1}, x_{i2} \dots x_{iD}) \quad 2.14$$

$$v_i = (v_{i1}, v_{i2} \dots v_{iD}) \quad 2.15$$

- La qualité de sa position est déterminée par la valeur de la fonction objectif, en ce point. Cette particule garde en mémoire la meilleure position, par laquelle elle est déjà passée, qui est notée $Pbest_i = (pbest_{i1}, pbest_{i2}, \dots, pbest_{iD})$.
- La meilleure position atteinte par les particules de l'essaim est notée $\vec{Gbest} = (gbest_1, gbest_2, \dots, gbest_D)$.

La notion Gbest (globale best) est calculée sur la version PSO globale, où toutes les particules de l'essaim sont issues de la particule i .

Etape1 : Initialiser aléatoirement N particules : position et vitesse. Par la suite à chaque itération, les particules se déplacent en fusionnant les trois composantes :

- $wv_{i,j}^t$: correspond à la composante d'inertie du déplacement, où le paramètre w contrôle l'influence de la direction de déplacement sur le déplacement futur.
- $c_1 r_{1,i,j}^t [pbest_{i,j}^t - x_{i,j}^t]$: correspond à la composante cognitive du déplacement, où le paramètre c_1 contrôle le comportement cognitif de la particule.
- $c_2 r_{2,i,j}^t [gbest_j^t - x_{i,j}^t]$: correspond à la composante sociale du déplacement, où le paramètre c_2 contrôle l'aptitude sociale de la particule.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

La combinaison de ces trois composantes, permet tout d'abord de déterminer la vitesse à laquelle la position d'une particule peut être calculée.

$$v_{i,j}^{t+1} = wv_{i,j}^t + c_1r_{1i,j}^t [pbest_{i,j}^t - x_{i,j}^t] + c_2r_{2i,j}^t [gbest_j^t - x_{i,j}^t], j \in \{1,2, \dots, D\} \quad 2.16$$

$$x_{i,j}^{t+1} = x_{i,j}^t + v_{i,j}^{t+1}, j \in \{1,2, \dots, D\} \quad 2.17$$

Avec :

W est une constante, appelée coefficient/pourcentage d'inertie.

r1, r2 sont deux nombres aléatoires tirés uniformément dans [0, 1], et ce à chaque itération t et pour chaque dimension j.

c1, c2 sont deux constantes, appelées coefficients/pourcentage d'accélération.

Etape2 : les nouvelles positions des particules sont évaluées et les deux vecteurs Pbesti et Gbest sont réindexés Selon les équations (2.18) si une fonction 'objectif' est minimisée alors elle est indiquée par la relation (2.19) dans PSOG.

$$Pbest_i(t+1) = \begin{cases} Pbest_i(t), & \text{si } f(x_i(t+1)) \geq Pbest_i(t) \\ x_i(t+1), & \text{sinon} \end{cases} \quad 2.18$$

Telle que f est la fonction 'objectif'.

Etape3 : pour chaque particule i, $pbest_i = x_i$.

Etape4 : calculer $Gbest$:

$$Gbest(t+1) = \arg \min_{pbest_i} f(pbest_i(t+1)), 1 \leq i \leq N \quad 2.19$$

Etape 5 : tant que le critère d'arrêt n'est pas satisfait faire :

- Déplacer les particules selon (2.16) et (2.17).
- Evaluer les positions des particules.
- Mettre à jour $pbest_i$ et Gbest selon (2.18) et (2.19).

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.8.2 Avantages et inconvénients de la PSO

Les algorithmes par essaim particulaire, présente des avantages, comme des inconvénients [37].

➤ **Avantages**

- Rapidité et robustesse.
- Implémentation simple, avec peu de paramètres.
- Résolution de différents problèmes.

➤ **Inconvénients**

- Problème de la convergence prématurée qui conduit à un optimum local.
- Problème de l'influence du choix des paramètres, sur la performance de l'algorithme.

Il existe d'autres algorithmes bio-inspirés [40] telle que :

- ❖ L'algorithme de colonies de fourmis ou ACO : il est Inspiré du comportement de recherche de nourriture des fourmis.
- ❖ L'Algorithme d'optimisation de colonie d'abeilles artificielle (ABC) : il est inspiré du comportement de recherche de nourriture intelligente par essaim d'abeilles.
- ❖ Le Système immunitaire artificiel (AIS) : qui est inspirée par les principes et le fonctionnement du système Immunitaire naturel (SIN).
- ❖ L'Algorithme du recuit simulé (SA) : qui vient de l'analogie avec la métallurgie où la méthode, pour atteindre des états de basse énergie d'un métal, quand on le chauffe jusqu'à des températures élevées, avant de le laisser refroidir lentement.

Chapitre 02 : techniques de la segmentation par classification et les algorithmes bio-inspirés

2.9 Conclusion

Nous avons présenté dans ce chapitre, deux grandes parties : la segmentation par clustering et les techniques bio inspirées, qui ont rencontré un vif succès grâce à leur simplicité, leur adaptabilité et leur capacité, de trouver des solutions presque optimales.

Nous avons particulièrement étudié les algorithmes génétiques ainsi que, les essaims particuliers et leur exploitation en segmentation d'images.

Dans le prochain chapitre, nous allons se concentrer sur l'étude détaillée, des approches du clustering optimisé par les techniques bio inspirées, que nous appliquerons pour la détection de masses en mammographie.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essais de particules

Chapitre 3 : Clustering optimisé par les algorithmes génétiques et les essais de particules

3.1 Introduction

Face au nombre croissant d'examens d'imagerie médicale, les logiciels de traitement d'images jouent un rôle majeur, afin de soutenir la tâche des praticiens. L'un des processus primordiaux pour tout traitement d'image est évidemment la segmentation.

Nous commençons ce chapitre par introduire la logique floue pour comprendre particulièrement le principe de l'approche FCM.

Une attention particulière, sera accordée à une méthode de classification des données utilisant un algorithme génétique et un algorithme K-Means, avec l'optimisation du centre de cluster initial, afin d'améliorer encore la qualité des clusters formés et, d'accélérer le processus de recherche.

Nous proposons également une nouvelle méthode de segmentation, basée sur le clustering optimisé par l'algorithme DPSO.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

3.2 Logique floue

L'idée de la logique floue [38] est de capturer l'imprécision de la pensée humaine et l'exprimer avec des outils mathématiques, appropriés. La théorie des ensembles flous, traite de la subjectivité et de l'incertitude ou de l'imprécision.

3.2.1 Notions d'ensemble et de sous-ensemble flou

La théorie mathématique des sous-ensembles flous est une extension de la théorie des ensembles classiques, pour la prise en compte des sous-ensembles définis de façon imprécise. L'exemple dans la figure ci-dessous montre que les contours de chaque ensemble flou ne sont pas « nets », mais « flous » ou « graduels ».

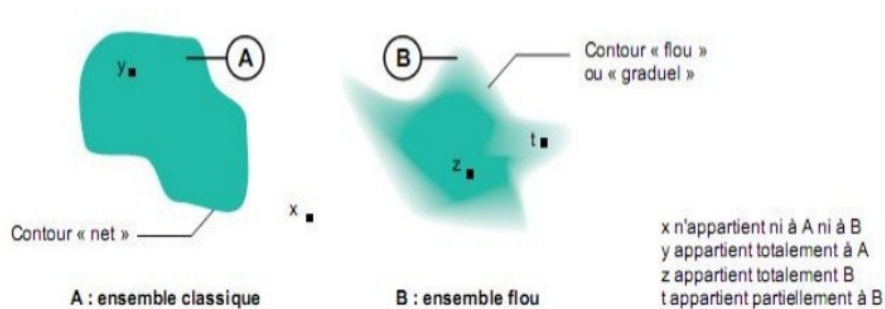


Figure 18: Comparaison d'un ensemble classique et d'un ensemble flou

La notion d'ensemble flou vient de l'observation que « très souvent, les classes d'objets rencontrés dans le monde physique, ne possèdent pas de critères d'appartenance bien définis » [38].

Mathématiquement, un ensemble flou B est défini sur un univers de discours U, par une fonction d'appartenance $\mu_B(x)$ qui prend ces valeurs dans l'intervalle [0,1]. Cette fonction donne le degré d'appartenance de chaque élément x à B. En fait, un élément de l'univers de discours peut être un membre entier ou partiel de l'ensemble B : $\mu_B(x) = \text{Degré}(x \in B)$.

3.2.2 Valeur d'appartenance

La valeur d'appartenance est le degré de compatibilité d'un élément, avec le concept qui est représenté par un ensemble flou.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

L'univers du discours $U \{0,1\}$ est l'ensemble référentiel, qui contient tous les éléments qui sont en relation avec le contexte donné, tel que 1 signifie que x appartient à B , et 0 si x n'appartient pas à B .

3.2.3 Fonctions d'appartenance

Un ensemble flou est défini par sa « fonction d'appartenance », qui correspond au concept de la « fonction caractéristique » en logique classique. Supposons afin de définir un ensemble des personnes de « taille moyenne ». En logique classique, par exemple les personnes de taille moyenne sont celles, dont la taille varie entre 1,60 m et 1,80 m. La fonction caractéristique de l'ensemble (figure 3.2) donne « 0 » pour les tailles hors de l'intervalle [1,60 m ; 1,80 m] et « 1 » dans cet intervalle.

L'ensemble flou des personnes de « taille moyenne » sera déterminé par une « fonction d'appartenance » qui diffère d'une fonction caractéristique, par le fait qu'elle peut prendre n'importe quelle valeur dans l'intervalle [0, 1]. A chaque taille possible, correspondra un «degré d'appartenance » à l'ensemble flou des « tailles moyennes » (figure 3.3), compris entre 0 et 1.

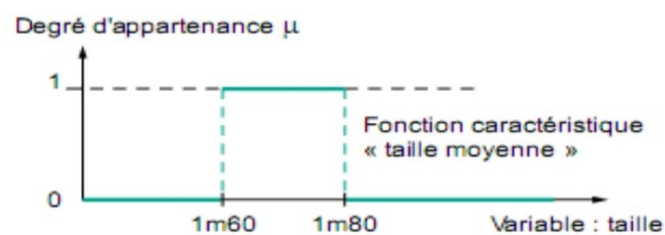


Figure 19 : Fonction caractéristique [38]

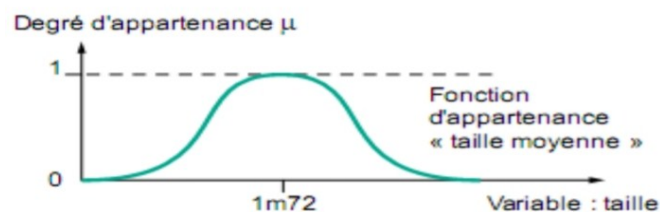


Figure 20: Fonction d'appartenance

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

3.3 Organigramme de l'approche Fuzzy C-Means

L'algorithme FCM, est basé sur la minimisation de la fonction objectif donnée dans l'équation 2.13. Dans cette approche, le pixel x_i n'est plus assigné à une unique classe, mais à plusieurs par l'intermédiaire de degrés d'appartenance $u_{ij} \in [0, 1]$ du pixel x_i à la classe j , plus l'appartenance est forte, plus le degré est proche de 1, et inversement [22].

On se rend bien compte qu'en fonction de la distance par rapport aux barycentres conditionnels, le degré d'appartenance des individus aux classes peut différer. En effet, l'algorithme FCM effectue une optimisation itérative en évaluant de façon approximative les minimums d'une fonction d'erreur.

Il existe toute une famille de fonctions d'erreur associées à cet algorithme qui se distinguent par des valeurs différentes prises par un paramètre réglable m , dite indice de flou "fuzzy index" « indice de flou » qui est utilisé pour déterminer le degré de flou de la partition obtenue et pour contrôler le flou d'appartenance de chaque donnée dans la plage $m \in [1, \infty]$.

Si $m=1$ cela conduit à un retour aux K-Means « crisp » ($u_{ij}=0$ ou 1) est c'est un partitionnement en associant chaque individu à la classe maximisant le degré d'appartenance. D'une manière générale, $m = 2$ dans les applications de FCM [22].

À chaque itération, la distance entre chaque pixel et le cluster est calculée. Si la distance est inférieure à la valeur seuil, la fonction 'objectif' est déterminée. Sinon les valeurs d'appartenance et du centroïde des clusters sont mis à jour. La segmentation est effectuée en fonction de la valeur de seuil calculée.

Les étapes expliquées dans l'organigramme suivant.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

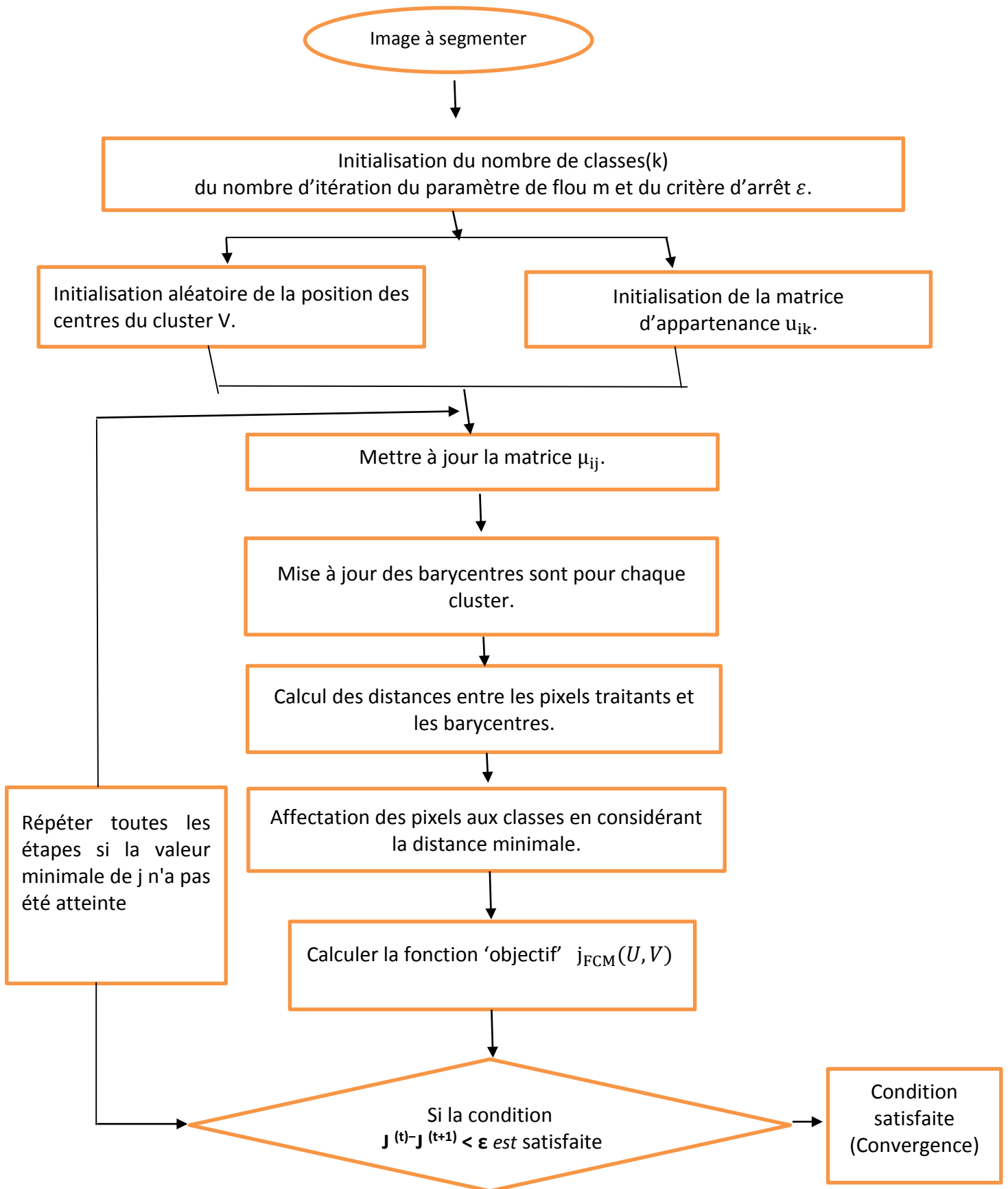


Figure 21 : Organigramme de l'approche FCM

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

3.4 Clustering optimisé par les AG

3.4.1 Principe général des AG

L'algorithme génétique commence au début, par choisir les individus autorisés à se reproduire, à partir d'une population de N individus [33, 34]. Ensuite, les individus choisis se croisent, de façon à obtenir une population d'enfants, tout en mutant aléatoirement certains gènes. La performance des individus est évaluée, grâce à la fonction fitness, qui permet de sélectionner les individus autorisés à survivre. Le processus se répète jusqu'à un critère d'arrêt, par exemple un certain nombre d'itérations est atteint ou aucune amélioration de la performance des individus. Les principes de bases de cet algorithme sont expliqués par le schéma de fonctionnement suivant :

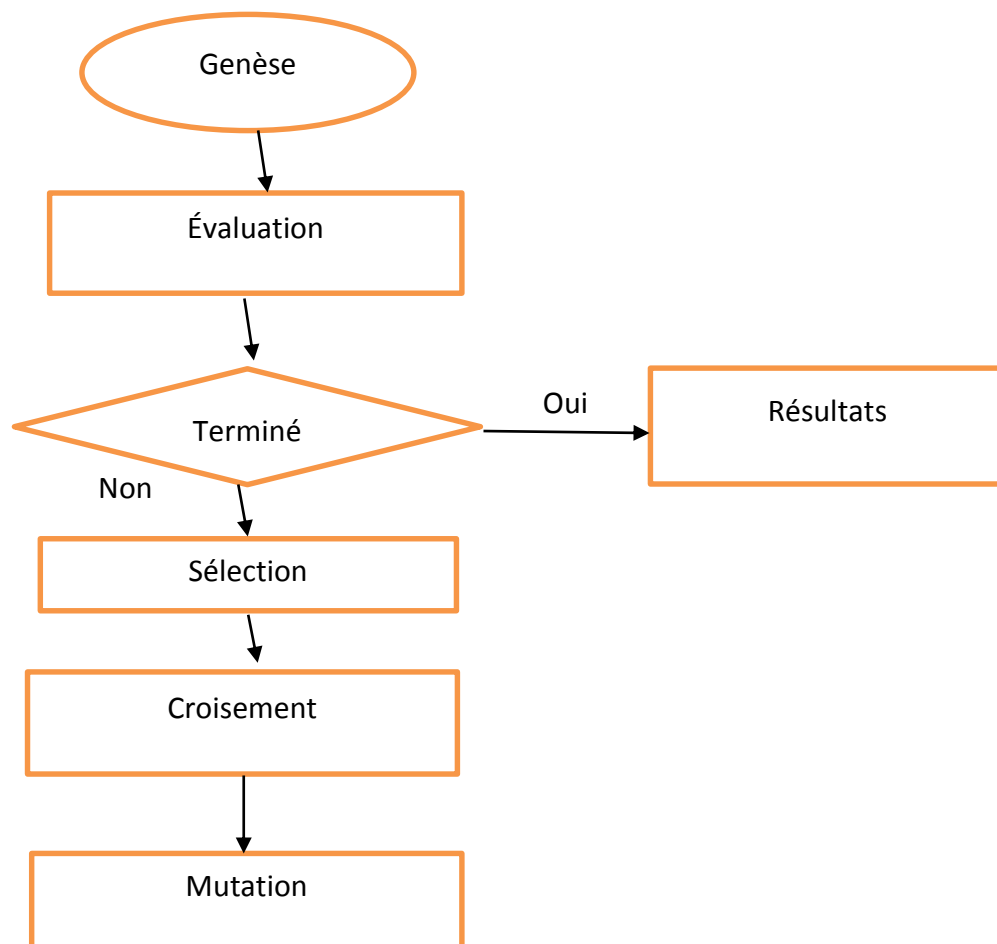


Figure 22: Principales étapes des AG

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

- La genèse est une étape pour créer une population aléatoire. C'est le point de départ des AG.
- L'évaluation est l'analyse des individus, pour analyser si une solution est disponible. Pour cela, La fonction coût doit être utilisée, pour déterminer le degré d'adaptation des individus lors du processus de sélection.
- L'utilisation d'une boucle est nécessaire, tant que l'évaluation estime que la solution n'est pas optimale.

3.4.2 Approche des K-means optimisée par les AG

L'algorithme de clustering K-means a des larges applications pour l'exploration de données et de documents, le traitement d'images numériques et différents domaines de l'ingénierie [39]. De plus, l'algorithme de clustering se compose d'étapes simples et a une convergence rapide. Cependant, son principal inconvénient, est la sélection initiale du centre de gravité, lors de la classification d'une image [34, 41].

La sélection initiale du centre de gravité, détermine les résultats et les performances cohérents de l'algorithme de clustering. Les bons barycentres des clusters initiaux, conduisent à de bons résultats de clustering et inversement.

Afin d'obtenir les performances cohérentes du clustering K-means, améliorer encore la qualité des clusters formés et accélérer le processus de la segmentation, il est proposé dans le cas de notre travail, une sélection initiale des barycentres des clusters des k-means par l'algorithme d'optimisation génétique. En effet, les AG adopte des opérations de recherche efficaces (sélection, croisement et mutation) pour la détermination des minima globaux, lors de la sélection des barycentres des clusters (figure 3.6).

3.4.2.1 Population génétique

Un ensemble de solutions ou de chromosomes, est appelé population. Chaque chromosome a un nombre N de gènes (variables) qui représentent les centres de gravité des clusters ; qui sont initialisés de manière aléatoire entre 0 (limite inférieure) et 1 (valeur de limite supérieure).

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

3.4.2.2 Fonction 'objectif'

La fonction fitness est la suivante : $f_i = (N - \sum_{k=1}^K a_i^k) / N$ $1 \leq i \leq l$

Dans cette formule, f_i est la fonction fitness du gène i ,

3.2

a_i^k est le nombre d'erreurs que le gène i catégorise dans les classes de cluster K ,

N est le nombre d'échantillons.

La fonction fitness doit être normalisée selon l'équation 3.3, pour que les données traitées aient la même dynamique.

$$\hat{f}_i = \frac{f_i - f^{max}}{f^{max} - f^{min}} \quad 1 \leq i \leq l \quad 3.3$$

\hat{f}_i engendre les valeurs de fitness normalisées du gène i , f^{max} , f^{min} sont le maximum et le minimum de la fitness dans la population.

L'AG détermine la valeur de fitness d'un chromosome, en évaluant une fonction 'objectif'. Il a pour fonction 'objectif', la minimisation de la somme de la différence intra-clusters, afin d'obtenir les meilleures valeurs initiales, du centre de gravité de cluster.

3.4.2.3 Fonctionnement

Cette partie conçoit la méthode de la segmentation des données mammaires, basée sur l'algorithme K-means. Afin de classer les données plus efficacement, le clustering initial est optimisé et les données sont nettoyées avec l'algorithme génétique. Ainsi, l'algorithme AG, est utilisé pour éliminer les attributs inutiles des données.

La méthode du clustering est utilisée pour détecter les valeurs aberrantes par mise en cluster. Les données similaires sont organisées en groupes ou cluster et, les données en hors du cluster, sont considérées comme des valeurs aberrantes.

Pour les données du clustering $S_i = \{\hat{S}_1, \hat{S}_2, \hat{S}_3, \dots, \hat{S}_N\}$ 3.1

La méthode de nettoyage des données basée sur l'algorithme génétique est la suivante :

Chapitre 03 : clusturing optimisé par les algorithmes génétiques et les essaims de particules

Étape 1 : Initialisation de la population

La méthode du codage et la fonction fitness, sont définies dans cette étape. La séquence des gènes est représentée par le codage 01, dans laquelle 0 signifie que l'individu n'a pas la particularité et, le chiffre 1 représente l'individu avec la particularité. La population initiale est une séquence 01 de 50 nombres générés aléatoirement (1,0, ... 1,0).

Étape 2 : Réglage de la fonction fitness.

La sélection des caractéristiques à laquelle chaque gène correspond, produit le résultat de l'algorithme K-means.

La valeur de similitude entre le résultat et l'ensemble des échantillons, est la valeur adaptative du gène. Le nombre d'individus dans la population est l.

Étape 3 : Sélection des individus en fonction de leur fitness et de l'algorithme de sélection proportionnelle

Ce processus se détermine selon les opérations ci-dessous ;

a) Calcul de la fonction fitness relative à tous les individus de la population. Dans ce cas, cette fonction représente la probabilité de sélectionner l'individu. Elle s'exprime comme le quotient de la fonction fitness d'un seul individu divisé par la somme de la fonction fitness de la population (équation 3.4).

$$p_i = \frac{\hat{f}_i}{\sum_{i=1}^l \hat{f}_i} \quad 3.4$$

b) Attribution de la surface de roulette correspondante, en fonction de la fonction fitness relative de chaque individu, puis une rotation de la roulette est réalisée. Plus la fonction fitness est élevée, plus la probabilité que le gène soit sélectionné, est importante.

Étape 4 : opération de croisement

Elle est l'opération de croisement des gènes. Elle représente la sélection au hasard des positions x et y de deux gènes, par un opérateur de croisement, à deux points. Il faut par la suite, échanger les séquences entre x et y des deux gènes.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

Étape 5 : Opération de mutation

Elle s'effectue par la sélection au hasard, de la position du gène en inversant la séquence 01 après la position sélectionnée.

Étape 6 : Sortie de la nouvelle population et conservation des résultats optimaux actuels

Si le nombre d'itérations a atteint le nombre d'itérations spécifié, l'algorithme cesse de s'exécuter on retourne à l'étape 2.

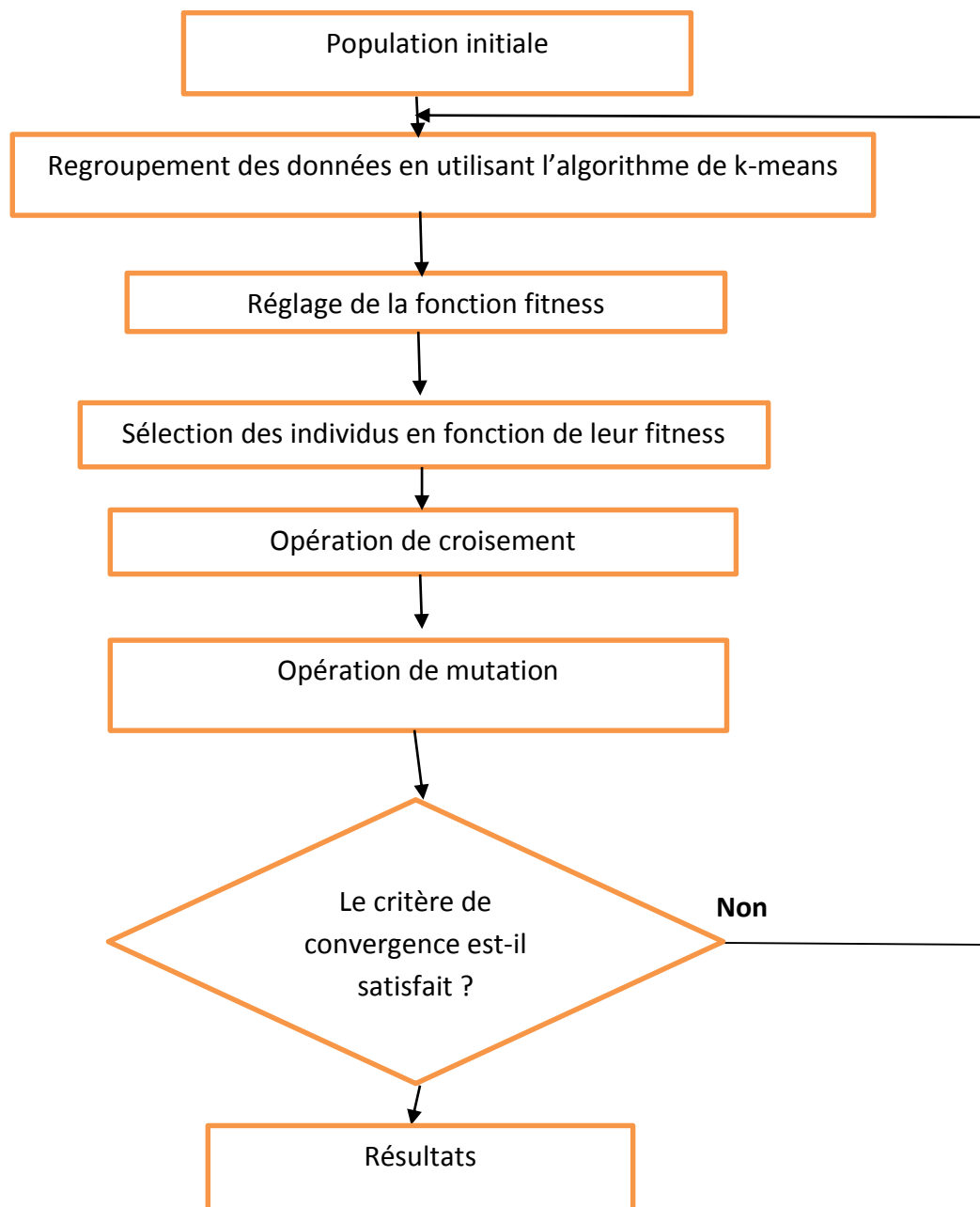


Figure 23: Approche adoptée de GAK-means

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

➤ Déroulement global du GAK-means

En fixant N_c : le nombre des clusters et le T le nombre maximum d'itérations, nous résumons ci-dessous, le déroulement global de l'approche du clustering optimisé par les k-means.

- 1) Initialiser chaque chromosome pour contenir k centres de gravité choisis au hasard dans l'ensemble de données.
- 2) Pour $t = 0$ à T
 - a) pour chaque chromosome i :
 - i) Calculer la distance euclidienne entre les centres de gravité et les pixels de l'image.
 - ii) affecter chaque pixel de données au cluster avec le centre de gravité le plus proche.
 - iii) Recalculer les centres de gravité de cluster du chromosome i comme la moyenne de leurs objets de données.
 - iiii) Calculer la fitness du chromosome i .
 - b) Créer la nouvelle génération de chromosomes en utilisant la sélection, le croisement et la mutation.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

3.5 Clustering k means optimisé par la DPSO

Cette partie vise à effectuer une segmentation des tumeurs mammaires sur des images IM, en utilisant une méthode de clustering améliorée basée sur l'algorithme K-Means. L'approche proposée traite de certaines limitations de l'algorithme K-Means standard, telles que l'initialisation aléatoire des centres de gravité des clusters. L'idée principale consiste à combiner la technique darwinienne d'optimisation des essaims de particules (DPSO), et l'algorithme K-Means. La technique DPSO [40] est appliquée aux images médicales IM pour l'initialisation des centroïdes de clusters.

3.5.1 Introduction à la DPSO

Les techniques de segmentation à plusieurs niveaux, offrent un moyen efficace d'effectuer une analyse d'image. Cependant, la sélection automatique d'un cluster optimal robuste est restée un défi dans la segmentation d'image.

La DPSO est utilisée afin de résoudre ce problème, pour la délimitation des valeurs de seuil à plusieurs niveaux, ou pour l'optimisation des clusters dans les K-means. Un problème général avec la PSO et d'autres algorithmes d'optimisation, est celui d'être piégé dans un optimum local de telle sorte, qu'il puisse bien fonctionner sur un problème mais échouer sur un autre.

Afin de surmonter cet inconvénient, l'optimisation des essaims de particules de Darwin (DPSO) a été formulée par Tillett et al en 2005 [40] à la recherche d'un meilleur modèle de sélection naturelle, à l'aide de l'algorithme PSO. Dans cet algorithme, de nombreux essaims, peuvent être des solutions. Chaque essaim fonctionne individuellement comme un algorithme PSO ordinaire, dans lequel la sélection naturelle (principe darwinien de survie du plus apte) est utilisée pour améliorer la capacité, à s'échapper des optima locaux.

Malgré les similitudes entre la PSO et les AG, comme la population générée aléatoirement, l'évaluation de la fonction de fitness, la mise à jour de la population ; la PSO n'utilisant pas d'opérateurs génétiques tels que le croisement et la mutation, n'est donc pas considérée comme une technique évolutive.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

Dans le cadre de cette étude, l'optimisation par la (DPSO) étend le k-means, pour déterminer si la sélection naturelle, peut améliorer la capacité du clustering, à échapper aux optima locaux.

3.5.1.1 Fonctionnement de l'algorithme (DPSO)

L'idée est d'exécuter plusieurs algorithmes PSO parallèles simultanés [40, 41, 42] ; chacun d'eux, étant un essaim différent, sur le même problème de test où, un simple mécanisme de sélection est appliqué. Lorsque la recherche tend vers un optimum local, le traitement dans cette zone est simplement rejeté et, une autre zone est recherchée à la place.

Dans cette approche, à chaque étape, les essaims qui s'améliorent, sont récompensés (prolongent la vie des particules ou engendrent un nouveau descendant). Les essaims qui stagnent, sont punis (réduisent la durée de vie des essaims ou suppriment des particules).

Pour analyser l'état général de chaque essaim, la fonction fitness de toutes les particules, est évaluée. Le voisinage et les meilleures positions individuelles de chacune des particules, sont mis à jour. Ainsi ;

- Si une nouvelle solution globale est trouvée, une nouvelle particule est générée.
- Une particule est supprimée, si l'essaim ne parvient pas à trouver un état d'ajustement dans un nombre défini d'étapes.

Certaines règles simples sont suivies pour supprimer un essaim, supprimer des particules, engendrer un nouvel essaim et une nouvelle particule :

- lorsque la population de l'essaim tombe en dessous d'une limite minimale, l'essaim est supprimé ;
- la particule la moins performante de l'essaim, est supprimée lorsqu'elle atteint un nombre maximal d'étapes (compteur de recherche SC_c^{max}) sans améliorer la fonction de fitness.

Après la suppression de la particule, au lieu d'être mis à zéro, le compteur est remis à une valeur approchant le nombre de clusters, selon :

Chapitre 03 : clusturing optimisé par les algorithmes génétiques et les essaims de particules

$$SC_C(N_{kill}) = SC_C^{max} \left[1 - \frac{1}{N_{kill} + 1} \right] \quad 3.5$$

Où N_{kill} est le nombre de particules supprimées de l'essaim sur une période au cours de laquelle, il n'y a pas eu d'amélioration de la fonction fitness.

Pour engendrer un nouvel essaim, un essaim ne doit avoir aucune particule jamais supprimée et le nombre maximal d'essaims, ne doit pas être dépassé. Pourtant, le nouvel essaim n'est créé qu'avec une probabilité de $p = f / NS$, avec f un nombre aléatoire dans $[0,1]$ et NS le nombre d'essaims. Ce facteur évite la création de nouveaux essaims, lorsqu'il existe un grand nombre.

L'essaim parent n'est pas affecté et la moitié des particules du parent, est sélectionnée au hasard pour l'essaim enfant. La moitié des particules d'un membre aléatoire de la collection d'essaims, est également sélectionnée. Si le nombre initial de population de l'essaim n'est pas obtenu, le reste des particules est initialisé au hasard et ajouté au nouvel essaim.

Une particule est engendrée, chaque fois qu'un essaim atteint un nouveau record et que la population maximale définie d'un essaim n'a pas été atteinte. Comme la PSO, quelques paramètres doivent également être ajustés pour exécuter l'algorithme efficacement :

- population initiale d'essaims ;
- population d'essaims maximale et minimale ;
- nombre initial d'essaims ;
- nombre maximum et minimum d'essaims ;
- seuil de stagnation.

Les étapes de l'algorithme DPSO se résument comme suit :

- ✓ Boucle principale du programme. Pour chaque essaim de la collection :
 - Faire évoluer l'essaim
 - Laisser l'essaim se reproduire
 - Supprimer les essaims ayant échoué
- ✓ Faire évoluer l'algorithme d'essaim. Pour chaque particule de l'essaim :
 - Mise à jour de la fonction fitness

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

- Mise à jour du meilleur des particules
 - Déplacer la particule,
 - Si l'essaim va mieux
- ✓ Essaim de récompenses. Particule d'apparition : prolonger la vie de l'essaim ;
- Si l'essaim ne s'est pas amélioré
- ✓ Punir l'essaim. Supprimer éventuellement la particule : réduire la durée de vie de l'essaim.

3.5.1.2 Algorithme des K-Means optimisé avec la DPSO

Le K-means dépend fortement du nombre de cluster initial et tombe facilement dans l'optimum local. En conséquence, il est souvent difficile d'obtenir des effets visuels satisfaisants. Une approche proposée est d'améliorer la capacité de recherche globale du clustering k-means avec l'algorithme DPSO. Les étapes de cette méthode sont :

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

1. Initialisation de la position des particules.
2. Regroupement des données en utilisant l'algorithme de k-means.
3. Pour chaque particule :
 - i. Calculer la distance euclidienne entre les centres de gravité et les pixels traitants.
 - ii. Affecter chaque pixel au cluster, avec le centre de gravité le plus proche.
 - iii. Recalculer les centres de gravité des clusters, de la particule comme la moyenne de leurs objets de données.
4. Calculer la valeur de la fonction fitness de chaque particule.
5. Tant que (condition d'arrêt) faire
6. Calculer p_{best} et G_{best} en fonction de la valeur de fitness des particules.
7. Calculer le vecteur de vitesse des particules.
8. Mettre à jour la position des particules.
9. Evaluer la valeur de fitness (fonction objective de l'essaim).
10. Si $\{fitness\}_{new} > \{fitness\}_{old}$ alors
11. Engendrer une nouvelle particule.
12. Sinon
13. Mettre à jour le compteur à 1.
14. Si $C_C = C_C^{max}$ **alors**
15. Supprimer une particule
16. Incrémenter N_{kill} par 1.
17. Fin
18. Fin
19. Si $N_{kill} = 0$ et $N_s < N_S^{max}$ alors
20. Engendrer une nouvelle particule.
21. Fin
22. Fin
23. Trouver la solution optimale.

Chapitre 03 : clustering optimisé par les algorithmes génétiques et les essaims de particules

3.8 Conclusion

Dans ce chapitre, nous avons présenté une méthode pour la segmentation des images, optimisée par les algorithmes bio inspirés.

L'algorithme des c moyennes floues est proche de l'algorithme des k-moyennes. Il présente l'avantage d'une sensibilité à l'initialisation plus faible.

Malgré la diversité des approches de la segmentation, les chercheurs ont conclu à l'inexistence d'une approche idéale, applicable à toutes les images. Dans ce cas, nous avons considéré cette problématique du point de vue d'un problème d'optimisation. C'est dans ce contexte, que sont introduites dans ce mémoire, l'étude de deux types de métaheuristiques. Les algorithmes génétiques interviennent afin de, nettoyer et optimiser les données des k-means. La DPSO est utilisée pour déterminer n-1 clusters optimaux, pour le partitionnement de l'image en améliorant les performances des k-means.

Le chapitre suivant présente l'environnement de travail, le traitement effectué sur des images mammographies réelles, ainsi qu'une discussion sur les résultats obtenus.

Chapitre 4

Mise en œuvre et résultats

4.1 Introduction

Le but de ce chapitre, est d'exposer l'environnement du travail, les approches adoptées, ainsi que les traitements effectués sur des images mammographiques réelles. L'objectif est d'extraire les masses à partir de la base référentielle MIAS [43].

4.2 Environnement de travail

Notre système a été développé, sous environnement Windows 8, à l'aide de l'outil de programmation Matlab de la version 2019b.

4.2.1 Matériel utilisé

L'implémentation de nos codes, a été réalisée grâce à un PC portable, ayant un processeur Intel Inside core i3, d'une capacité de mémoire vive « RAM de 8 Go ».

4.2.2 Langage utilisé

Le nom MATLAB vient de l'anglais Matrix Laboratory. MATLAB peut être considéré comme un laboratoire de manipulation, de matrices. La plupart des fonctions définies dans MATLAB, sont destinées pour des grandeurs matricielles et par extension, pour des données tabulées [44].

MATLAB comprend de nombreuses fonctions, de traitements de données, d'affichage, de tracés de courbes, de résolution de systèmes et d'algorithmes de calculs numériques au sens large du terme. Il s'agit d'un outil puissant, qui permet la résolution de nombreux

Chapitre 04 : Mise en œuvre et résultats

problèmes, en beaucoup moins de temps. On peut se rendre compte de la puissance du logiciel, en lançant la commande *demo*.

Les compléments d'information peuvent être facilement obtenus, grâce à l'aide en ligne. Matlab (figure 4.1) s'élargit sur la librairie des fonctions mathématiques, l'environnement graphique, ainsi que sur une interface de développement.

On le trouve dans les applications de :

- développement d'algorithmes et la programmation ;
- modélisation et simulation ;
- analyse et visualisation de données ;
- création de graphiques scientifiques ;
- création d'application avec interfaces utilisateurs ;
- l'interfaçage avec d'autres langages (**Fortran, C++**) ;
- le développement d'interface utilisateur (**GUI**) ;
- L'utilisation des boîtes à outils (Toolboxes) spécialisées.



Figure 24: Icône de matlab R2019b

Chapitre 04 : Mise en œuvre et résultats

4.2.3 Interface graphique

Le fonctionnement ergonomique de MATLAB, facilite grandement le traitement du notre sujet. Que ce soit la segmentation d'une image, son affichage, ou même l'implémentation d'un algorithme évolutif, toutes ces tâches sont réalisables à l'aide de MATLAB.

La figure 4.2 représente les fenêtres que nous avons utilisées, dans notre travail telles que :

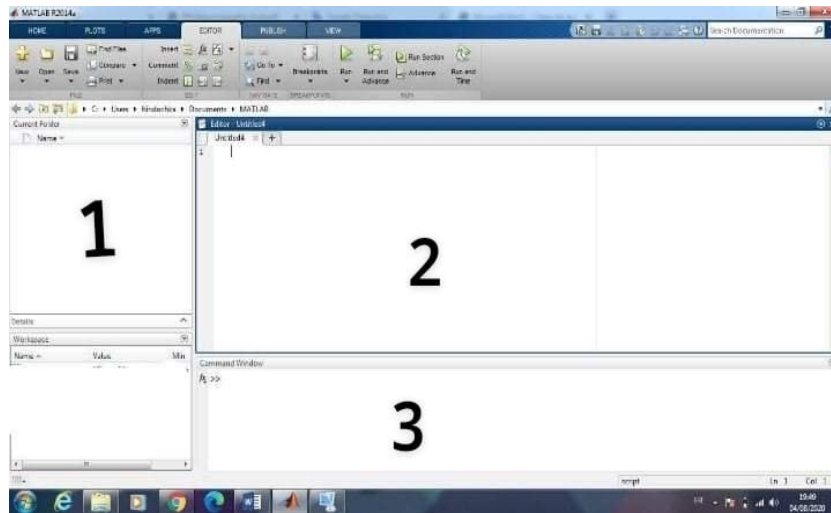


Figure 25: Interface graphique de MATLAB

- 1. Workspace** : il s'agit d'un espace de travail qui contient des variables créées ou importées dans MATLAB, à partir de fichiers de données.
- 2. Editeur** : c'est un éditeur de texte qui permet de saisir et modifier le script.
- 3. Command Window** : Il s'agit de la fenêtre principale, qui apparaît dans la partie centrale de l'interface. C'est dans cette fenêtre, que les commandes seront saisies.

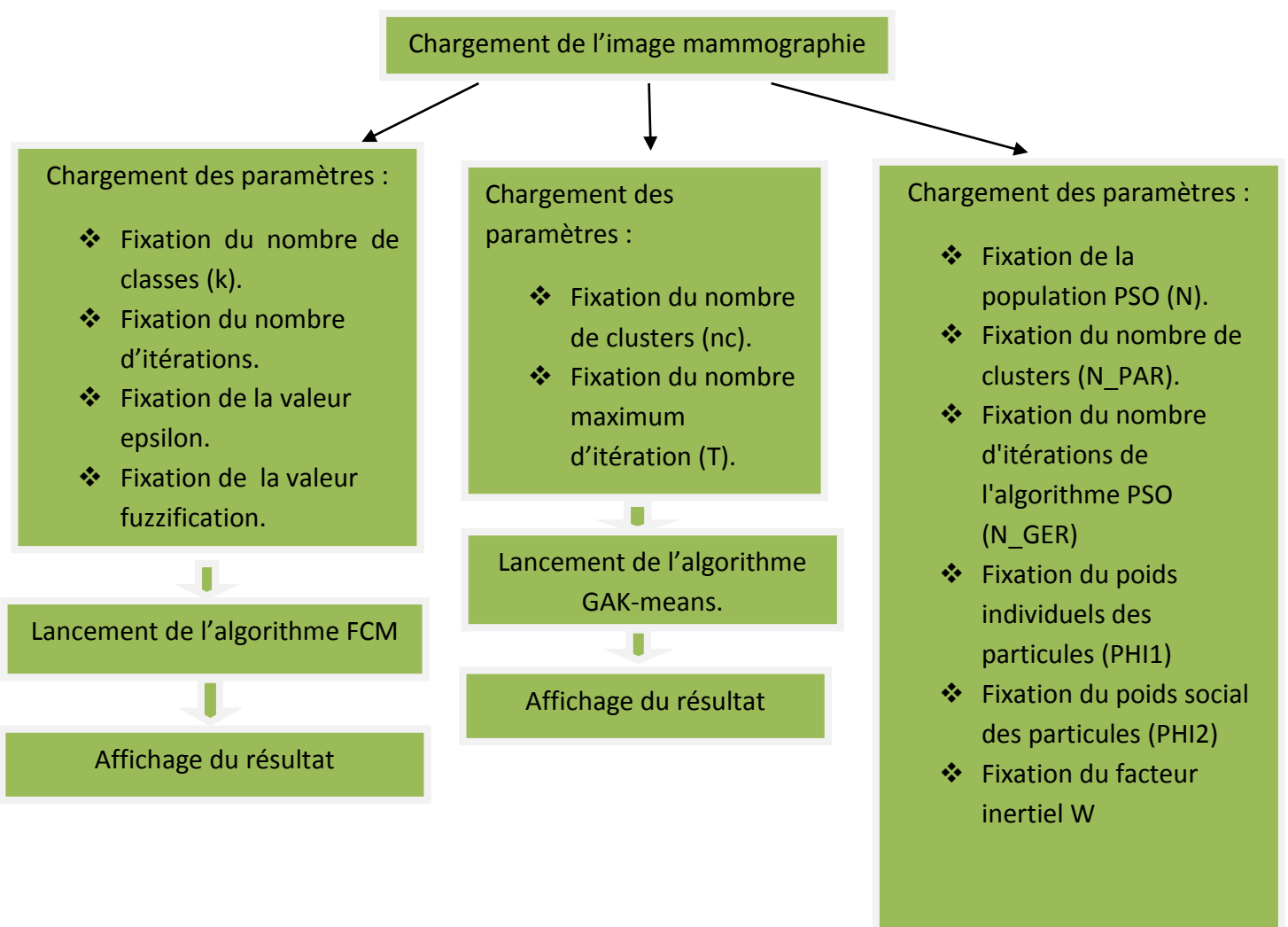
Chapitre 04 : Mise en œuvre et résultats

4.3 Approche de 'BreastBioSeg'

Le système 'BreastBioSeg' (figure 4.3) permet de détecter les masses malignes et bénignes dans les images mammographiques, suivant des méthodes de segmentation et d'optimisation. Notre travail a été consacré dans la première partie, à la segmentation d'image par la méthode de classification non supervisée Fuzzy C-Means ; qui est précédée par le choix du nombre de classes et le nombre maximum d'itérations ainsi que, le choix de la valeur de Fuzzification et la valeur du critère d'arrêt epsilon.

La seconde partie, est une combinaison de l'approche des k-means optimisée par les algorithmes génétiques.

La troisième partie, est dédiée à la segmentation par le clustering classique optimisé, par la DPSO. Une discussion sur les paramètres adoptés pour la segmentation d'une image mammographique clôture la partie expérimentale.



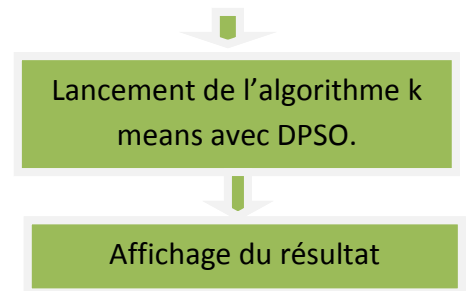


Figure 26 : Synoptique de BreastBioSeg

4.3.1 Base d'images

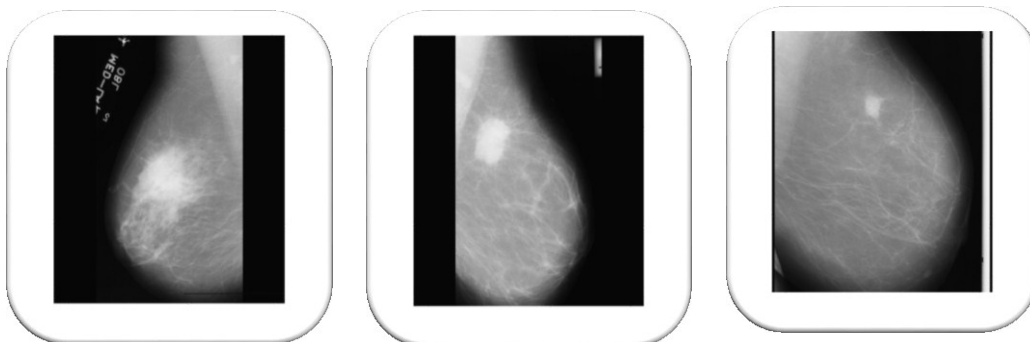
L'approche BreastBioSeg est appliquée sur les données mammographiques réelles (figure 4.4), de la base d'images Mias (Mammography Image Analysis Society) [43].

MIAS est une organisation des groupes de recherches sur les mammographies du royaume uni.

La base de données MIAS, comprend 322 images numériques, où 207 sont des cas sains, 26 représentent des calcifications et 89 sont des masses. La résolution spatiale de chaque image, est de 1024*1024 pixels [45].

Toutes les étiquettes des images mammographiques sont au format suivant : mdb XXXB, où :

- XXX: représente le nombre de l'image, allant de 001 à 322;
- B représente le côté du sein, prenant la valeur «L» pour le sein gauche ou «R» pour le sein droit (gauche ou droit).



Chapitre 04 : Mise en œuvre et résultats

Figure 27: Exemples de masses bénigne et malignes de la base MIAS

4.3.2 Méthode du Fuzzy c means (FCM)

Les algorithmes flous modélisent l'incertitude et l'imprécision dans la séparation des différentes régions, en optimisant une fonction coût tout en donnant un certain degré d'appartenance à chaque classe et ce, pour chaque pixel. Dans la première partie de notre travail l'objectif est la segmentation d'image, par la méthode de classification non supervisée Fuzzy C-Means. Le Principe de base de cette méthode est d'introduire un caractère flou d'appartenance aux classes k de dimension $(n \times K)$ (n nombre des pixels, K nombre de classes).

4.3.2.1 Chargement des paramètres du FCM

Nous avons défini nos paramètres d'entrée qui sont :

- le nombre de classes ;
- le nombre d'itérations ;
- la valeur Epsilon ;
- la valeur du degré de fuzzification.

Les valeurs des paramètres de la segmentation sont des valeurs expérimentales. Par conséquent, elles varient en fonction de la nature des images. Pour chaque image, différentes valeurs peuvent être fixées, en fonction des expériences réalisées.

4.3.2.2 Résultats de la segmentation FCM

Des tests ont été effectués dans différentes images mammographiques, qui présentent des masses malignes et bénignes. Les figures ci-dessous montrent les résultats obtenus pour quelques cas. Le nombre de clusters, a été fixé à 3, 4 et 5.

Les figure 4.5, 4.6 et 4.7 présentent un partitionnement d'une masse maligne, considérée dans l'image mdb184.

Chapitre 04 : Mise en œuvre et résultats



Figure 28: Résultats du premier test appliqué sur l'image mdb184

Pour ce premier test, le nombre de classes est fixé à 3, le nombre d'itérations à 15. Le clustering est obtenu pour une précision de 0.0001.



Figure 29: Résultats du second test appliqué sur l'image mdb184

Le clustering du second test, a été obtenu pour un nombre d'itérations 15, un nombre de classes de 4 et une précision de 0.0001.



Figure 30: Résultats du troisième test appliqué sur l'image mdb184

Chapitre 04 : Mise en œuvre et résultats

Le clustering du troisième test, a été obtenu pour un nombre d'itérations 15, un nombre de classes de 5 et une précision de 0.0001.

Les figures 4.8, 4.9 et 4.10, illustrent les résultats du clustering flou sur une masse bénigne, indiquée dans l'image mdb002.

Le premier test, est obtenu pour un nombre de classes de 3. Le second test, est réalisé pour un nombre de classes de 4 et le dernier, pour un nombre de clusters de 5.



Figure 31: Résultats du premier test appliqué sur l'image mdb002

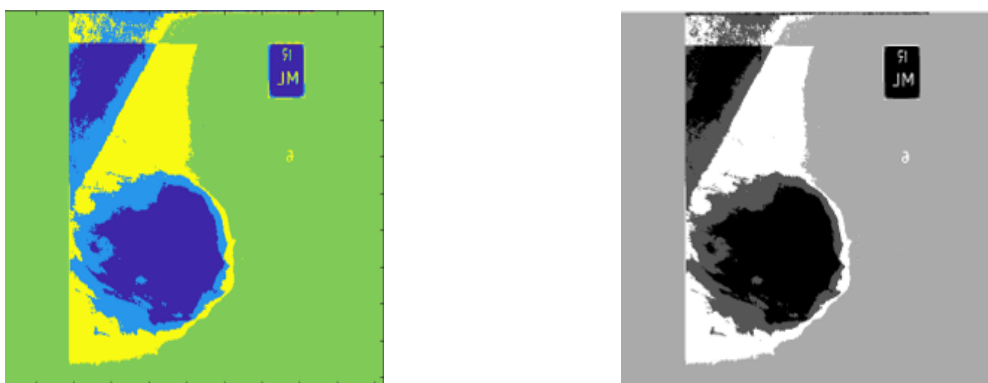


Figure 32 : Résultats du second test appliqué sur l'image mdb002



Figure 33: Résultats du troisième test appliqué sur l'image mdb002

4.3.2.3 Discussion

Le traitement engendré par le clustering flou, aboutit à des résultats satisfaisants. Les données ont été affichées en couleurs, pour la distinction des différentes structures de l'image, ainsi qu'en niveaux de gris.

Une simple visualisation sur les résultats obtenus, montre clairement que les classes obtenues sont disjointes et plus ou moins homogènes. Ceci est dû au fait de l'introduction de la logique floue dont l'apport a été bien traduit, par la qualité de la segmentation obtenue.

Les tests ont été effectués avec un nombre de clusters valant 3, 4, et 5. Nous avons constaté que plus le nombre de classes est élevé, plus le niveau de gris est commun avec d'autres objets dans l'image, tel que dans le test 3 lorsque le nombre de cluster était égale à 5. De plus, lorsque nous augmentons le nombre de clusters, nous avons constaté que les contours des images forment de nouveaux clusters, mais la qualité se dégrade.

L'image a été segmentée facilement ce qui montre que les objets à segmenter, doivent avoir une couleur suffisamment uniforme, pour que la détection soit envisageable. Nous avons pu constater aussi lorsque les contours d'image ne sont pas bien définis, la segmentation devient difficile.

Nous avons noté également que cette méthode de segmentation, prend un temps de calcul très long.

4.3.3 Clustering classique optimisé par les AG

L'algorithme des K-Means est considéré comme l'un des algorithmes, largement utilisés dans le clustering. Cependant, il a encore quelques problèmes. L'un d'eux, est dans la phase d'initialisation qui se fait, de manière aléatoire. Un autre problème produit par ce clustering, est qu'il converge vers les minima locaux. Les algorithmes génétiques sont l'un des algorithmes évolutifs inspirés de la nature et utilisés dans le domaine du clustering, pour régler le problème d'initialisation de centre de cluster.

4.3.3.1 Choix des paramètres

Les paramètres à fixer sont considérés par ; le nombre de classes (nc), le nombre d'itérations (T), un nombre de générations de 100 et une population de taille 20. La sélection se fait via la méthode de la roulette.

4.3.3.2 Résultats de la segmentation par GAK-means

Nous présentons les résultats obtenus par la méthode proposée, pour montrer l'effet des AG, sur le problème du clustering.

Les figures ci-dessous, illustrent quelques exemples sur l'approche du clustering, optimisée par les AG.

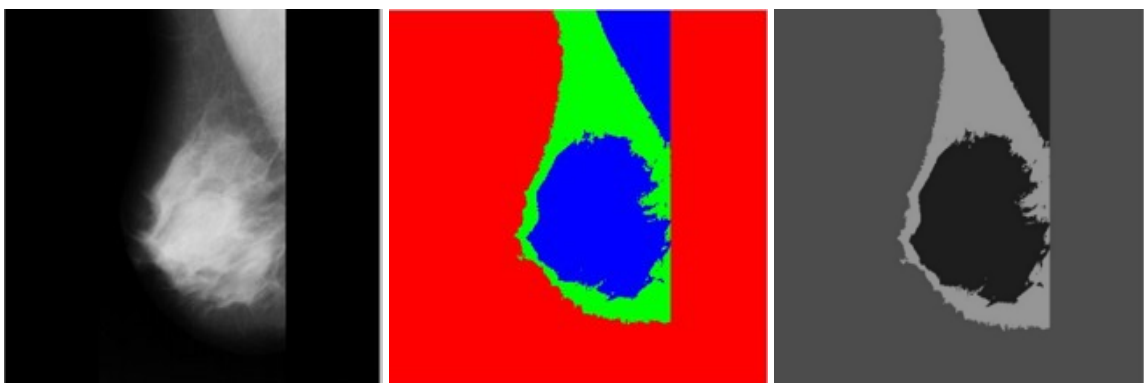


Figure 34: Résultats du premier test appliqué sur l'image mdb001

Pour ce test, le nombre de cluster est de 3 et le nombre d'itérations est de 50. Les meilleurs clusters, ont été obtenus, pour une fonction fitness de 272 (figure 34).

Chapitre 04 : Mise en œuvre et résultats

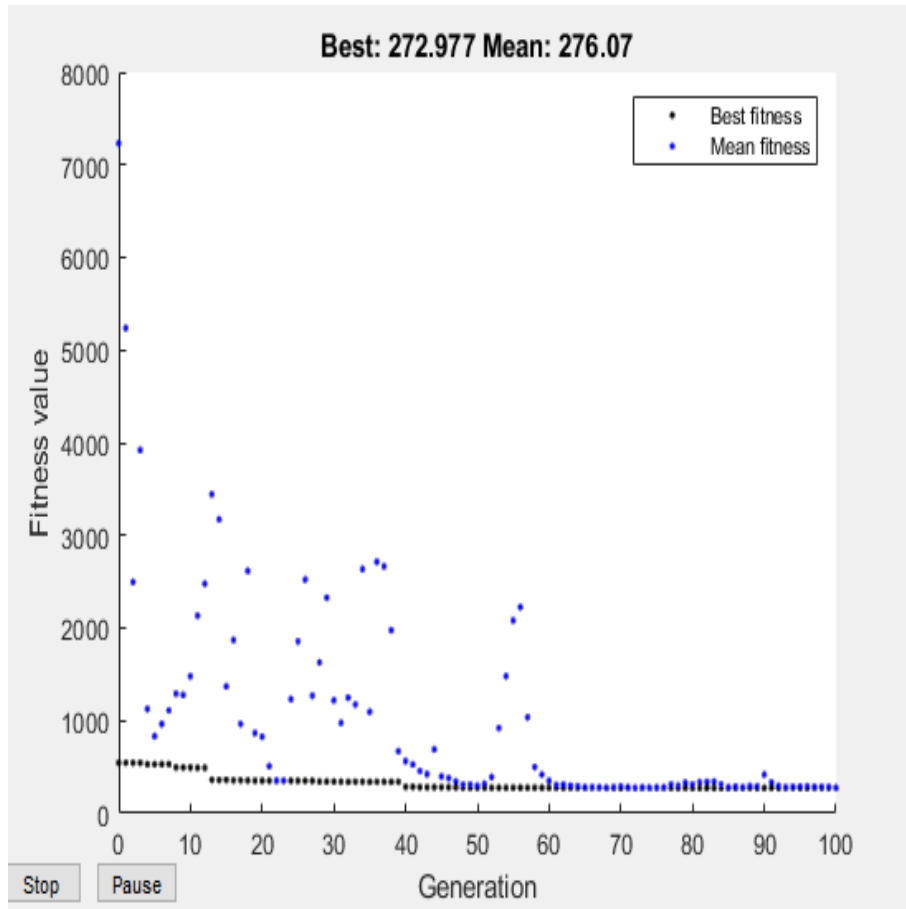


Figure 35: Evolution des clusters vers les meilleurs pour le premier test de l'image mdb001
Le second test engendre les résultats pour la détection de la masse bénigne de mdb001, suivant un nombre de clusters de 4 et un nombre d'itérations de 50. Les meilleurs clusters, ont été obtenus, pour une fonction fitness de 147 (figure 35).



Figure 36: Résultats du second test appliqué sur l'image mdb001

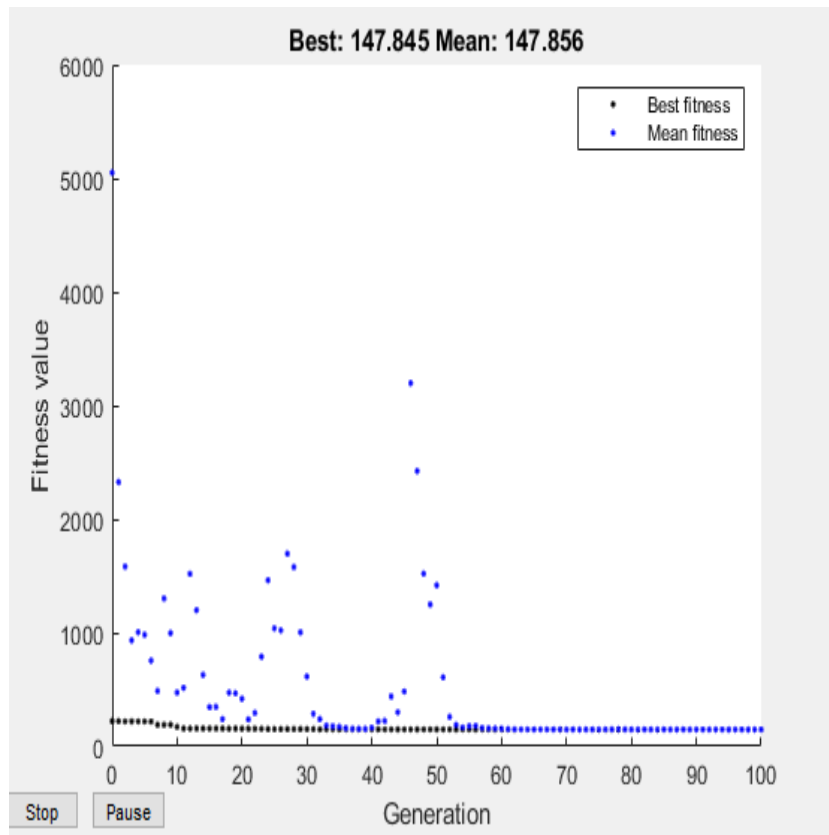


Figure 37: Evolution des clusters vers les meilleurs pour le second test de l'image mdb001

Le troisième test engendre les résultats pour la détection de la masse bénigne de mdb001, suivant un nombre de clusters de 5 et, un nombre d'itérations de 50. Les meilleurs clusters, ont été obtenus, pour une fonction fitness de 114 (figure 37).

Chapitre 04 : Mise en œuvre et résultats

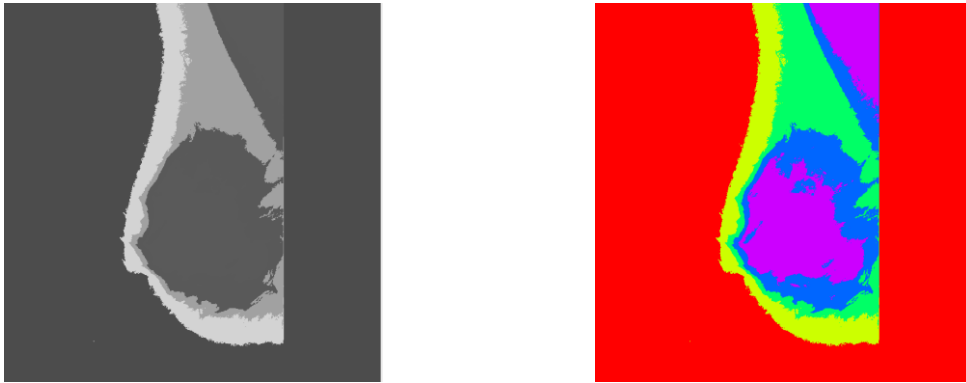


Figure 38: Résultats du troisième test appliqué sur mdb001

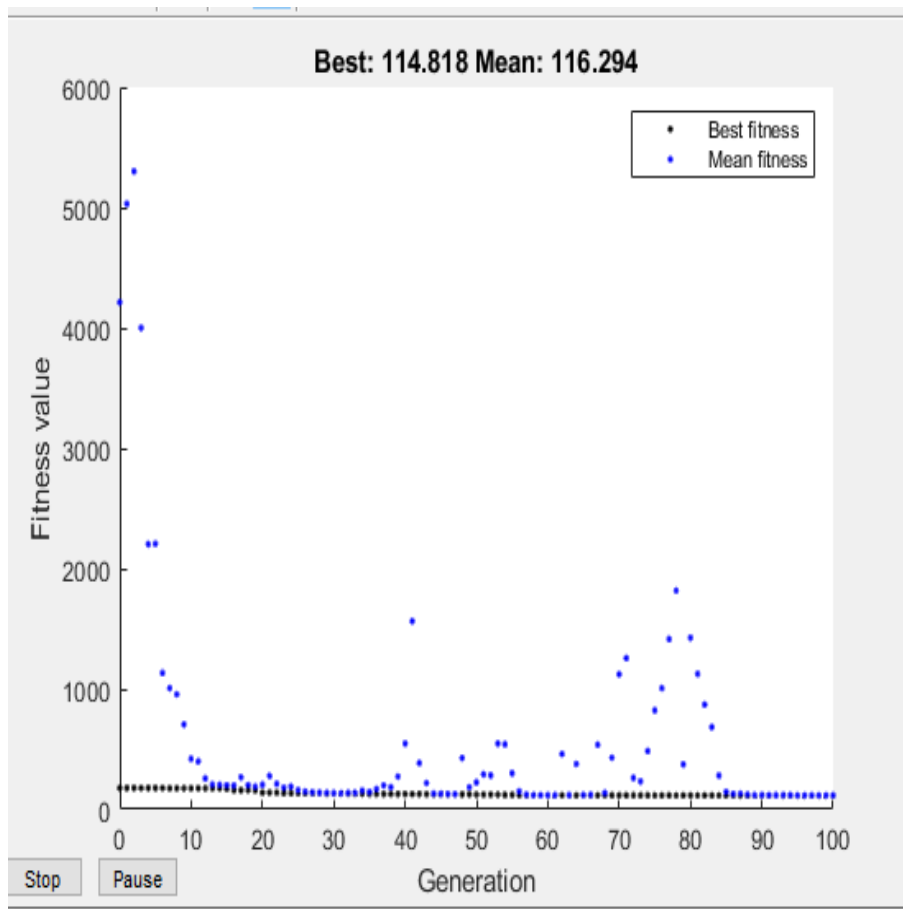


Figure 39: Evolution des clusters vers les meilleurs pour le troisième test de l'image mdb001

Les tests qui suivent, sont appliqués sur une seconde image, qui présente une masse bénigne, présentée sur l'image mdb002.

Chapitre 04 : Mise en œuvre et résultats

➤ 1^{er} test : nombre de clusters=3

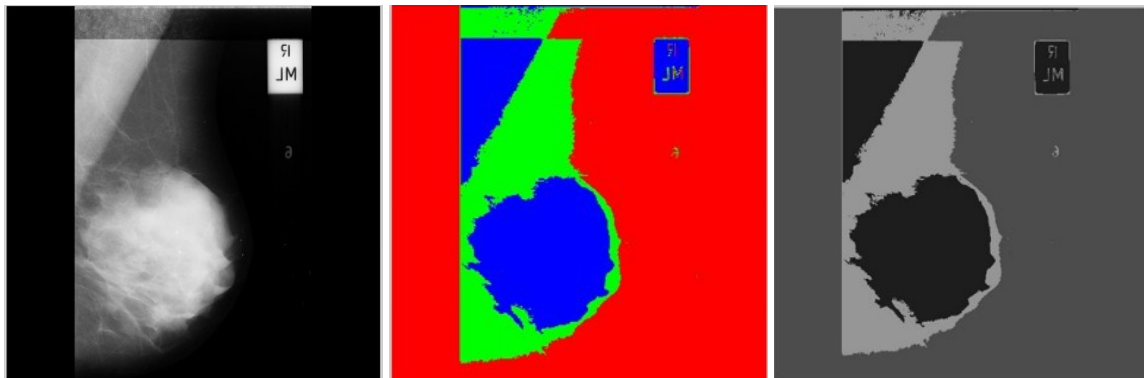


Figure 40: Résultats du troisième test appliqué sur mdb002

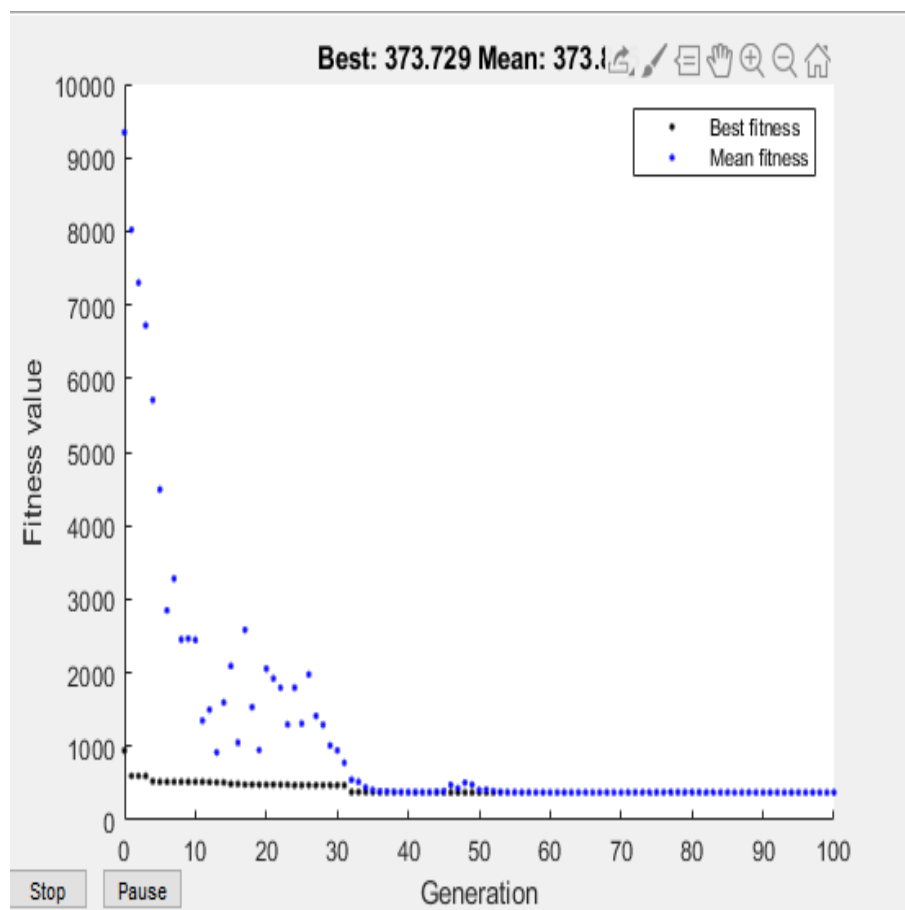


Figure 41: Evolution des clusters vers les meilleurs pour le premier test de l'image mdb002

Les meilleurs clusters, ont été obtenus, pour une fonction fitness de 372 (figure 41).

Chapitre 04 : Mise en œuvre et résultats

➤ **2nd test : nombre de clusters=4**

L'image segmentée, est observée sur la figure 4.19. Les meilleurs clusters, ont été obtenus, pour une fonction fitness de 235 (figure 4.20).

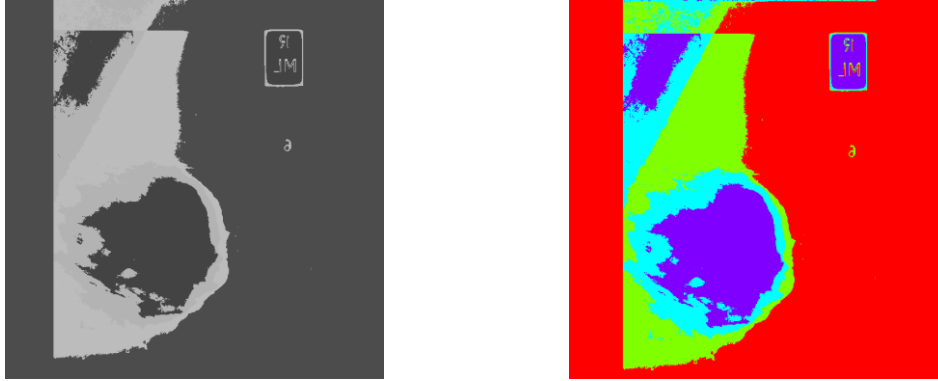


Figure 42: Résultats du second test appliqué sur mdb002

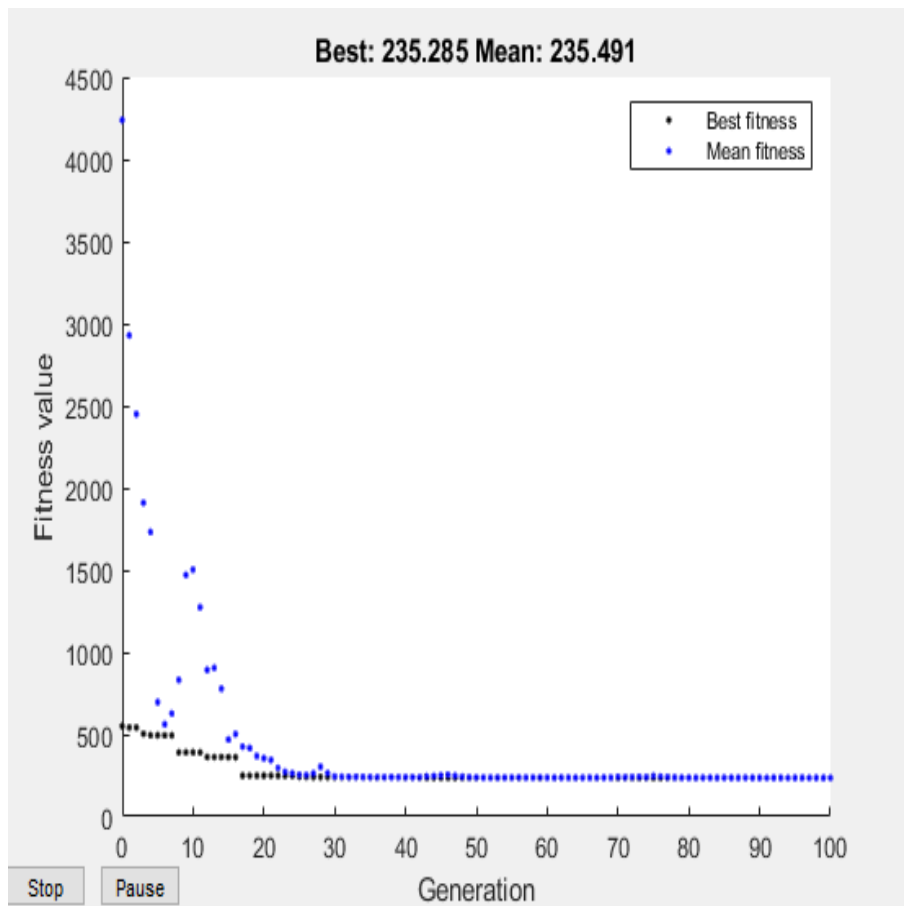


Figure 43: Evolution des clusters vers les meilleurs pour le second test de l'image mdb002

Chapitre 04 : Mise en œuvre et résultats

➤ 3^{ème} test : nombre de clusters=5

L'image segmentée, est observée sur la figure 4.21. Les meilleurs clusters, ont été obtenus, pour une fonction fitness de 176 (figure 4.22).

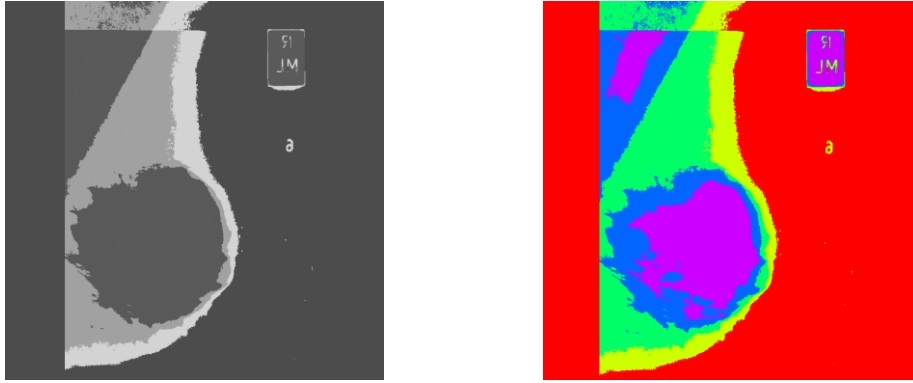


Figure 44: Résultats du troisième test appliqué sur mdb002

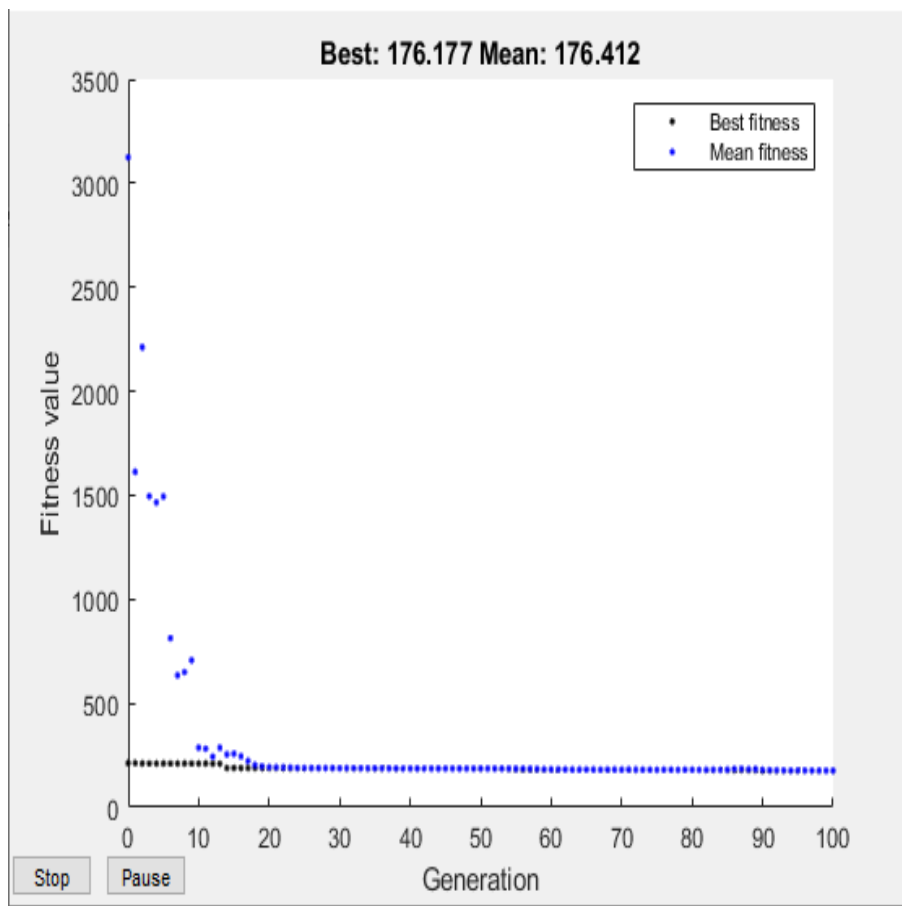


Figure 45: Evolution des clusters vers les meilleurs pour le troisième test de l'image mdb002

Chapitre 04 : Mise en œuvre et résultats

➤ Cas d'un traitement d'une masse maligne

Nous appliquons dans le cas de ce traitement, la forme optimisée du clustering, sur l'image mdb184.

➤ Premier test où le nombre de clusters=4

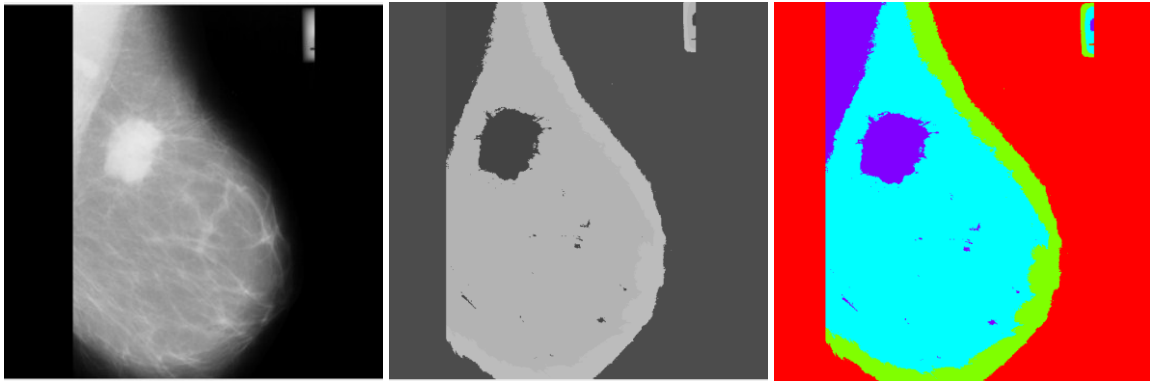


Figure 46: Résultats du premier test appliqué sur mdb184

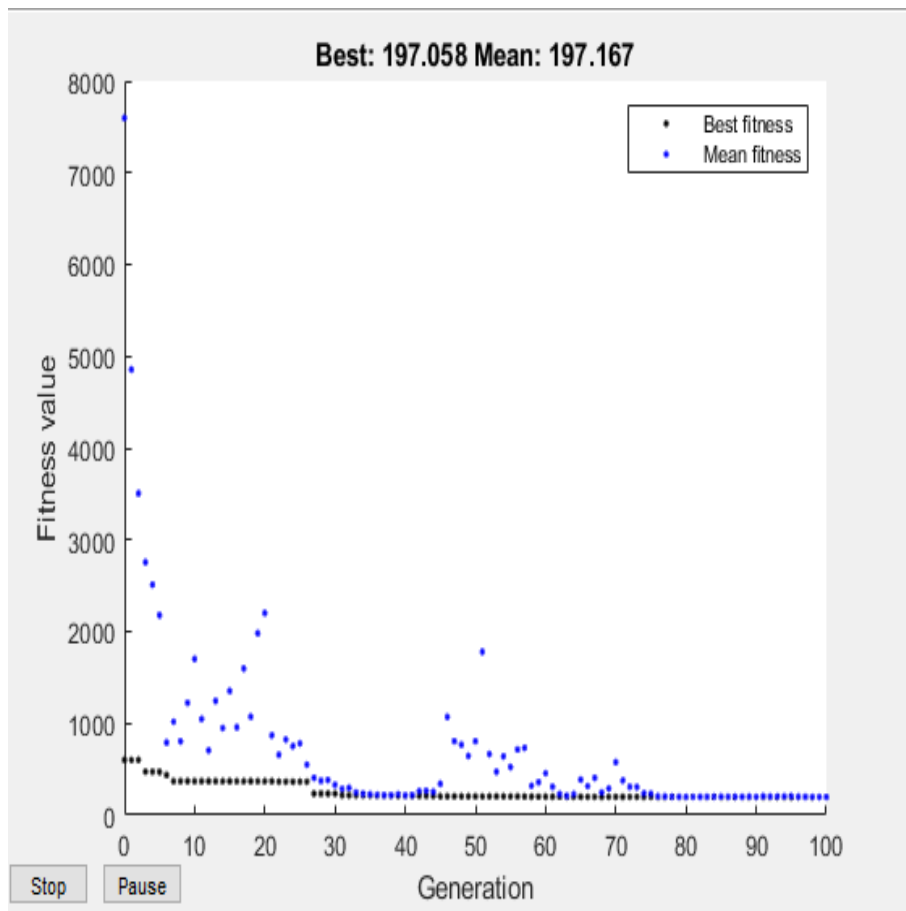


Figure 47: Evolution des clusters vers les meilleurs pour le premier test de l'image mdb184

Chapitre 04 : Mise en œuvre et résultats

Dans le cas de cette expérimentation, la masse a été détectée pour la fonction fitness, de 197 (figure 47).

➤ **Second test où le nombre de clusters=5**

Les résultats sont illustrés par les figures 4.25 et 4.26.

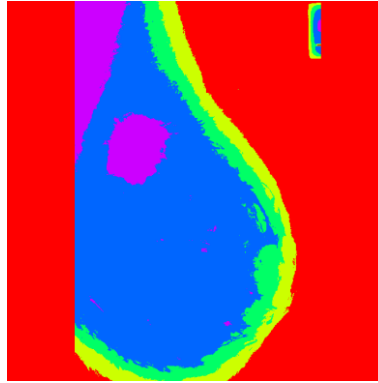


Figure 48: Résultats du second test appliqué sur mdb184

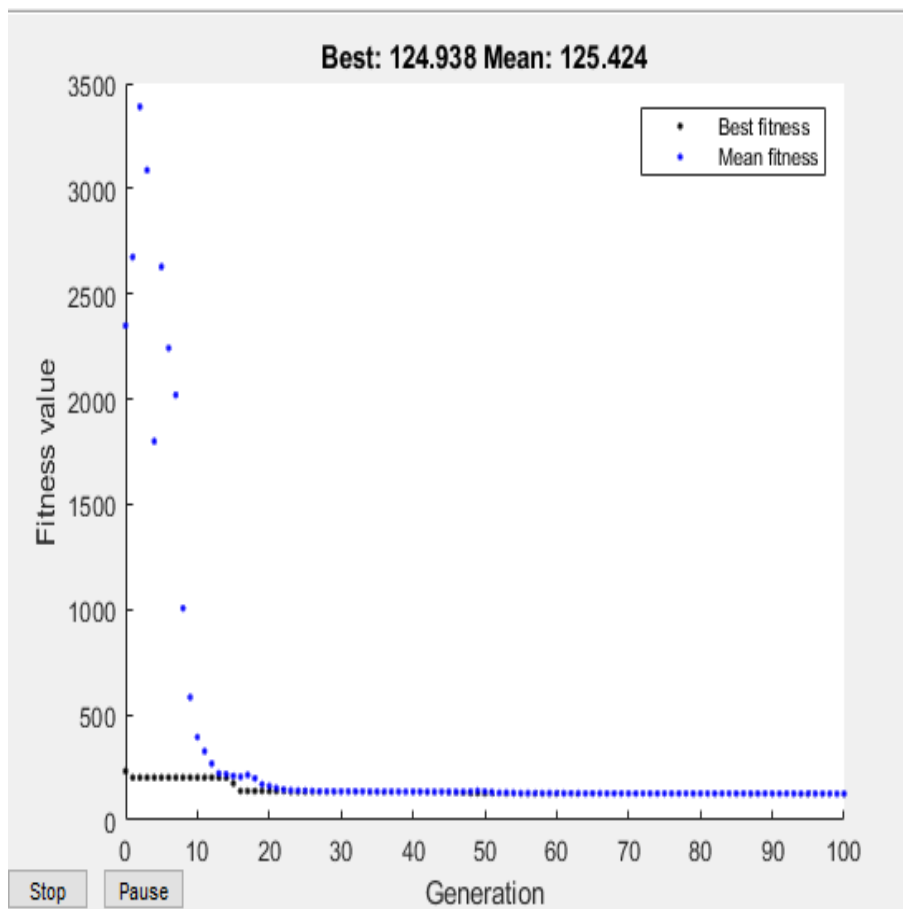


Figure 49: Evolution des clusters vers les meilleurs pour le second test de l'image mdb184

Chapitre 04 : Mise en œuvre et résultats

Le meilleur clustering est obtenu, pour une fonction fitness de 124.

4.3.3.3 Discussion

Après une série des tests suivant le changement du nombre de classes, nous avons remarqué que lorsque le nombre de classes augmente, les contours des images deviennent plus visibles. A cet effet, la zone d'intérêt détaillée, est plus perceptible pour un nombre de clusters de 5.

Le temps d'exécution de cette méthode, est plus rapide que pour l'approche du clustering flou.

4.3.4 Approche des K-means basée sur l'optimisation par la DPSO

Dans cette partie, un algorithme amélioré du clustering de k-moyennes, optimisé par la DPSO est proposé. Le but est de résoudre le problème d'initialisation des centres de clusters.

4.3.4.1 Chargement des paramètres

Certains paramètres de l'approche doivent être fixés, avant le traitement des données.

- La taille de la population PSO (N) est fixée à 50.
- Le nombre de clusters (N_PAR), fixé à 3, 4 et 5.
- Le nombre d'itérations de l'algorithme PSO (N_GER) est fixé à 150.
- Le poids individuel des particules (PHI1).
- Le poids social des particules (PHI2).
- Le facteur inertiel W.

4.3.4.2 Résultats de la méthode k-means et la DPSO

Nous avons effectué plusieurs tests, sur plusieurs images mammographiques suivant des variations du nombre de clusters. Le but de ces expérimentations, est de voir l'influence de la méthode DPSO relativement aux autres méthodes.

L'application de la méthode du clustering, par la PSO engendre les résultats représentés par les figures ci-dessous :

Chapitre 04 : Mise en œuvre et résultats

➤ Cas de la masse bénigne mdb001

Nos premières expériences avec cette approche, se font sur une image présentant une masse bénigne où le premier test, considère un nombre de classes de 3, le second de 4 et le troisième de 5.



Figure 50 : Résultats du premier test appliqué sur mdb001



Figure 51 : Résultats du second test appliqué sur mdb001

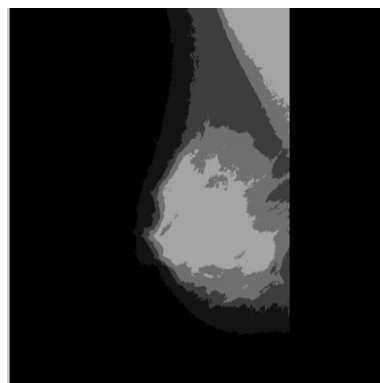


Figure 52 : Résultats du troisième test appliqué sur mdb001

Chapitre 04 : Mise en œuvre et résultats

➤ Cas de la masse bénigne mdb002

Nos expériences dans ce cas, se font sur une image présentant une masse bénigne où le premier test, considère un nombre de classes de 3, le second de 4 et le troisième de 5.



Figure 53: Résultats du premier test appliqué sur mdb002

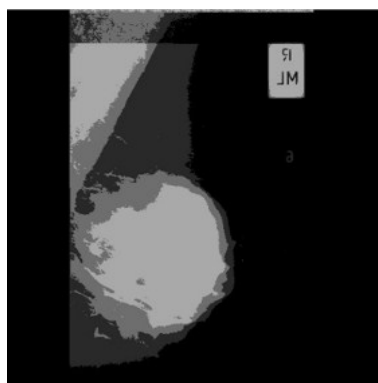


Figure 54: Résultats du second test appliqué sur mdb002

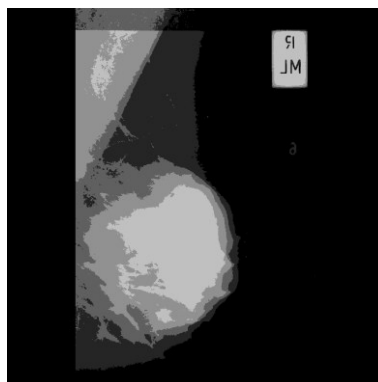


Figure 55: Résultats du troisième test appliqué sur mdb002

➤ **Cas de la masse maligne mdb184**

Nos expériences avec cette approche, se font sur une image présentant une masse maligne où le premier test, considère un nombre de classes de 4 et, le second de 5.

Les résultats obtenus, sont illustrés par les figures ci-dessous :

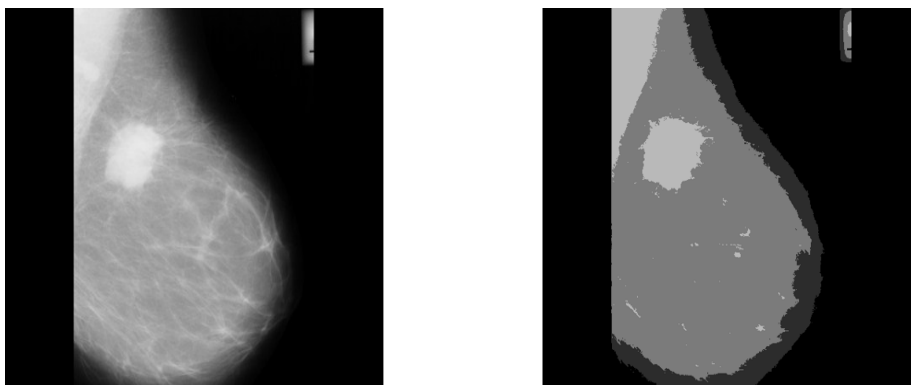


Figure 56: Résultats du premier test appliqué sur mdb184

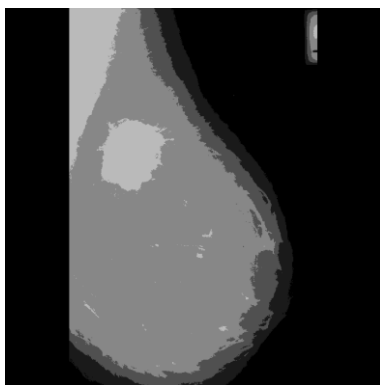


Figure 57: Résultats du second test appliqué sur mdb184

Chapitre 04 : Mise en œuvre et résultats

4.3.4.3 Discussion

Dans les essais obtenus sur Matlab, nous pouvons facilement remarquer, que la détection précise des régions d'intérêt, a une relation directe, avec le nombre de classes. En effet, plus le nombre de classes est élevée plus les données pertinentes de l'image mammographique, sont moins préservées.

Nous remarquons après l'analyse visuelle des images, que la méthode k-means optimisée par la DPSO, a engendré de très bons résultats, pour les images présentant des masses bénignes, avec un nombre des classes égal à 3. Les contours sont tous fermés et, très proches des limites des régions des masses. La masse maligne de l'image mdb184, est parfaitement extraite, pour un nombre de clusters de 4.

Comme le montrent ces images, la technique est bien adaptée à, l'imagerie mammographique.

Nous avons constaté également, lors des tests, que le temps de calcul pour la segmentation basée sur la DPSO était beaucoup plus court, que pour les autres méthodes. Cela s'est produit en raison du nombre de particules, dans la population de la DPSO constituée, de plusieurs essais plus petits. Nous pouvons conclure que la DPSO, est capable de trouver de meilleurs clusters en moins de temps.

4.4 Discussion générale

Les expériences considérées, par les techniques du clustering optimisées par les métaheuristiques, ont été réalisées sur différentes images prélevées de la base de données MIAS.

Les trois méthodes sont simples à comprendre et peuvent être applicables à divers types de données. En effet, nous avons constaté que les trois algorithmes nécessitent une connaissance préalable du nombre de clusters à déterminer et par leur caractère itératif, ils s'avèrent inefficaces lorsque le nombre de clusters devient important. L'algorithme des Fuzzy C-Means est bien adapté, à l'identification de classes et il est d'autant plus performant, lorsque nous connaissons le nombre de classes a priori.

A partir des résultats obtenus, nous pouvons juger que la méthode K-Means optimisée avec la DPOS est meilleure que les approches FCM et GAK-means, en matière de la qualité de la solution trouvée et, en temps d'exécution. Les résultats expérimentaux confirment que la méthode proposée avec la DPSO est efficace, avec une meilleure précision du clustering et un temps d'exécution très réduit.

Le choix d'une méthode est lié à plusieurs facteurs, dont l'ajustement des paramètres, qui gouverne le fonctionnement de l'algorithme mérite une attention particulière.

4.5 Conclusion

Dans ce chapitre, nous avons présenté des différentes méthodes de la segmentation d'image. Dans la partie expérimentale, nous avons exposé les résultats obtenus par la segmentation du Fuzzy c-means, du GAK-means et des k-means avec la DPSO. Ces résultats ont montré que la méthode Fuzzy C-Means est performante. Cependant, elle peut être toujours sujet à des imperfections, en introduisant d'autres critères tels que le degré d'homogénéité de la classe et, en testant d'autres types de distances.

Nous constatons que les algorithmes pré-initialisés comme l'algorithme GAK-means et k-means optimisé avec la DPSO fonctionnent bien et, engendrent des résultats significatifs et utiles en termes de temps de calcul ainsi qu'en recherche de bonnes configurations du clustering. En outre, il est plus significatif de sélectionner, à partir de chaque cluster, des centres importants.

Le résultat obtenu par l'optimisation de l'essaim de particules darwiniennes (DPSO) montre son efficacité pour résoudre le problème du clustering pour délimiter les classes et pour surmonter les inconvénients des méthodes évolutives, en termes de problème du blocage dans des minima locaux. La méthode DPSO est caractérisée par la rapidité du temps d'exécution plus que l'algorithme FCM et l'algorithme GAK-means.

Nous pouvons conclure que l'approche proposée du clustering avec la DPSO est robuste. Ce qui nous permet de confirmer, l'intérêt majeur des méta heuristiques, à améliorer la qualité de la segmentation.

Conclusion générale

Conclusion générale

L'initiation à la recherche réalisée dans le cadre de ce mémoire, s'inscrit dans les objectifs de l'équipe, "IMAD", du laboratoire LATSI, pour la conception de systèmes d'aide au diagnostic en imagerie médicale.

Dans l'analyse d'images médicales, la segmentation est sans aucun doute la tâche, qui mobilise le plus d'efforts, car elle nécessite une connaissance préalable sur le type, la qualité des images à traiter ainsi que, la structure à segmenter. Après avoir effectué un parcours bibliographique sur les méthodes de la segmentation par clustering et décelé, les avantages et les inconvénients de chacune, nous avons noté l'importance et la difficulté du problème de la segmentation d'images IM.

Au cours de ce mémoire, nous nous sommes intéressées à l'exploitation des méta heuristiques bio-inspirées, dans le domaine de la segmentation d'images mammographiques. L'objectif recherché était de reformuler le problème de la segmentation en un problème d'optimisation, en vue d'améliorer les résultats et aider le médecin, à caractériser les différentes pathologies du sein.

Dans ce mémoire nous avons développé des algorithmes de segmentation de masses, dans les images mammographiques. Ces approches sont basées sur le clustering flou FCM, la méthode des k-means optimisée par les algorithmes génétiques et, par la DPSO. Les objectifs fixés sont de :

- faciliter la détection des masses mammaires ;
- d'assister le radiologue, dans son interprétation ;
- réduire les erreurs de diagnostic.

L'algorithme fuzzy c means a été largement utilisé, en raison de son efficacité. Néanmoins, son principal inconvénient est lié à la forte dépendance, aux positions initiales des centres de classes, ce qui les conduit au blocage dans des minima locaux. Par conséquent, nous nous

Conclusion générale

concentrâmes, sur la mise en œuvre d'une approche, qui rend la segmentation plus efficace et plus robuste.

Il s'agit de l'algorithme k-moyenne hybridé d'une part avec les AG et d'autre part, avec la DSPO. Ces se caractérisent par leur simplicité et leur capacité d'adaptation, pour résoudre des problèmes d'optimisation. Nous avons utilisé ces approches, pour accélérer la convergence du clustering et améliorer la qualité de la segmentation, qui doit se traduire par une bonne initialisation du nombre de classes.

Les approches GAK-means et DSPO-K-means, ont détecté correctement les différents tissus mammaires, pour la majorité des images testées. En effet, divers essais ont démontré l'efficacité et la supériorité des méthodes K-means avec la DPSO par rapport au FCM et, au GAK-means. Les résultats obtenus, nous ont permis de nous assurer que l'utilisation d'algorithmes bio-inspirés permet de surmonter les problèmes rencontrés, par les méthodes classiques et d'améliorer significativement la qualité de la segmentation.

De manière plus subjective, nous avons pris plaisir à utiliser l'environnement de programmation MATLAB, et sommes persuadées que nous serons amenées, à retravailler sur ce langage.

A partir de ce travail, de nombreuses perspectives pourraient être envisagées dans le cadre de futurs travaux, nous pouvons citer les suivantes :

- Incorporer des informations spatiales dans la fonction 'objectif', pour rendre l'algorithme plus robuste à des niveaux de bruit élevés.
- Intégrer plusieurs sources d'informations, dans le cadre de la fusion des données.
- Exploiter d'autres techniques bio-inspirées et d'autres modes d'hybridation.
- Développer un système d'aide au diagnostic à travers les résultats de la segmentation générés par BreastBioSeg.

Enfin, une perspective intéressante reste l'adaptation d'algorithmes bio-inspirés à la segmentation de séquences d'images mammogr

Bibliographie

- [1] Organisation mondiale de la santé. Cancer du sein : prévention et lutte contre la maladie, <https://www.who.int/topics/cancer/breastcancer/fr/>, consulté en avril 2020.
- [2] J. Cardenas. Le sein : anatomie, maladies et examens, https://www.doctissimo.fr/html/sante/mag_2002/sem02/mag1004/dossier/sa_5966_anatomie_sein.htm, Doctissimo, septembre 2016, consulté en mars 2020.
- [3] Ablation d'un nodule du sein, <https://doctinet.info/chirurgie/gynecologie-obstetrique/160-ablation-d-un-nodule-du-sein.html>, 2016, consulté en mars 2020.
- [4] O. Redon. Tumeur maligne ou bénigne : définition, différence et symptômes, <https://sante.journaldesfemmes.fr/fiches-maladies/2524453-tumeur-maligne-ou-benigne-definition-difference-et-symptomes/2019>, consulté en mars 2020.
- [5] H. Chekkaf, I. Touil. Segmentation des masses dans les images mammographiques, Mémoire de Master en Informatique, Université Abou Bakr Belkaid, Tlemcen, 2012.
- [6] Entreprise Roche. Comprendre la mammographie, <https://www.roche.fr/fr/patients/info-patients-cancer/diagnostic-cancer/diagnostic-cancer-du-sein/depistage-mammographie.html>, consulté en avril 2020.
- [7] Entreprise Roche. Le dépistage du cancer du sein, <https://www.roche.fr/fr/patients/info-patients-cancer/diagnostic-cancer/diagnostic-cancer-du-sein/depistage-organise-cancer-sein.html#:~:text=Le%20d%C3%A9pistage%20du%20cancer%20du,un%20examen%20clinique%20des%20seins>, consulté en avril 2020.
- [8] S. Ronze. La mammographie : définition, âge, intérêt et déroulement de l'examen, <https://www.doctissimo.fr/html/sante/imagerie/mammographie.htm?fbclid=IwAR02PfhkMVSVM40EZYtJ9JGH7DCmOnH10KOyjxoBz9Ys7U2xnNxlqIDTK8>, juin 2016, consulté en mars 2020.
- [9] Société canadienne du cancer. Mammographie, <https://www.cancer.ca/fr-ca/cancer-information/diagnosis-and-treatment/tests-and-procedures/mammography/?region=qc>, consulté en mars 2020.

Bibliographie

- [10] C. Cordonnier. La mammographie ne détecte pas seulement le cancer du sein, <https://www.topsante.com/medecine/cancers/cancer-du-sein/la-mammographie-ne-detecte-pas-seulement-le-cancer-du-sein-610507>, mars 2016, consulté en mars 2020.
- [11] I. Cheikhrouhou, Description et classification des masses mammaires pour le diagnostic du cancer du sein, thèse de doctorat en informatique, Université d'Evry-Val d'Essonne, 2012.
- [12] C. Vachier. Extraction de caractéristiques, segmentation d'image et morphologie Mathématique. Thèse de doctorat en informatique, Ecole des Mines de Paris, Décembre 1995.
- [13] Mammographie, disponible sur <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-sein/Diagnostic/Mammographie>, consulté en avril 2020.
- [14] Programme québécois de dépistage du cancer du sein (PQDCS)-Région de la capitale-nationale, <http://www.depistagesein.ca/avantages-inconvenients/?fbclid=IwAR3HBdzNsHJrX1EC6bt3YSGawvU45L7l4j9daBMoxEaRVfaFYoh24UwnvDI#Xw77hqEzblU>, consulté en mars 2020.
- [15] J. P. Cocquerez & S. Philipp. Analyse d'images : filtrage et segmentation. Masson, 1995.
- [16] I. Bloch, Y. Gousseau, H. Maître, D. Matignon, B. Pesquet-Popescu, F. Schmitt, M. Sigelle, F. Tupin. Le traitement des images, Polycopié du cours ANIM Département TSI - Télécom-Paris, <http://www.ensta.fr/>, version 5.0, septembre 2004, consulté en avril 2020.
- [17] O. Alaoui Ismaili., V. Lemaire., A. Cornuéjols. Classification à base de clustering : ou comment décrire et prédire simultanément, Rencontre en Intelligence Artificielle (RJCIA 2015), <https://hal.inria.fr/hal-01178907/document>, Juin 2015, Rennes, France, consulté en avril 2020.
- [18] R. Eyraud. Classification non-supervisée : Regroupement (clustering), cours, Université Aix Marseille, http://pageperso.lif.univ-mrs.fr/~remi.eyraud/CAD/cours_4-clustering.pdf, consulté en avril 2020.
- [19] P. L. Gonzalez. Méthodes de classification, <http://maths.cnam.fr/IMG/pdf/Classification-2008-2.pdf>, consulté en avril 2020.
- [20] S. Hamidouche, T. Idjeraoui. Clustering : Approche par la théorie des jeux, mémoire de Master

Bibliographie

en Recherche Opérationnelle, Université Mira Abderrahmane de Béjaia, 2013.

[21] S. Ouadfel. Contributions à la Segmentation d'images basées sur la résolution collective par colonies de fourmis artificielles. Thèse de Doctorat en Informatique, Université Hadj Lakhdar de Batna, juillet 2006.

[22] R. RAKOTOMALALA. Méthode des centres mobiles, http://eric.univ-lyon2.fr/~ricco/cours/slides/classif_centres_mobiles.pdf, consulté en avril 2020.

[23] S. Moussa. Contribution à la segmentation d'image médicale par les algorithmes bio inspiré, thèse de doctorat en informatique, Université de Sétif, 2018.

[24] O. El Ganaoui, M. Perrot. Segmentation par la méthode des k-means, <http://www.tsi.enst.fr/pages/enseignement/ressources/beti/hyste-dyn/node2.html>, consulté en mai 2020.

[25] I. Hausmane. K-means : Définition & Avantages/Inconvénients, <https://le-datascientist.fr/k-means> 2019, consulté en mai 2020.

[26] N. Al-Amram. Algorithme des K-Moyennes, ENSA, Tetouan, <http://www.ieee.ma/uaesb/pdf/Algo-k-Moyennes.pdf>, consulté en avril 2020.

[27] C. M. Immaculate, S.V. Kasmir Raja. Improved Fuzzy C-Means Clusters With Ant Colony Optimization, International Journal of Computer Science & Emerging Technologies (E-ISSN: 2044-6004), Volume 1, N° 4, pages 2-3, December 2010.

[28] A Tutorial on Clustering Algorithms, http://home.deib.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html, consulté en mai 2020.

[29] M. Hachama, N. Bohoua, F. Olivier. Une segmentation grossière et rapide des images en couleurs, <http://www.tsi.enst.fr/pages/enseignement/ressources/mti/gric/schedule.htm>, consulté en mai 2020.

[30] M. Âya. Chapitre 2 : les méthodes de segmentation par clustering, https://www.academia.edu/24629026/Chapitre_2_memoire, consulté en mai 2020.

Bibliographie

- [31] L. Pennington. Solving Min-Max Problems Using Derivatives, <https://study.com/academy/lesson/solving-min-max-problems.html>, consulté en mai 2020.
- [32] J. Barolet. Bio-inspiration, <https://datafranca.org/wiki/Bio-inspiration>, 2020, consulté en juin 2020.
- [33] J. H. Holland. Adaptation in Natural and Artificial Systems, the MIT press, 1992.
- [34] Algorithme génétique, http://igm.univ-mlv.fr/~dr/XPOSE2013/tleroux_genetic_algorithm/fonctionnement.html#:~:text=Il%20y%20a%20trois%20op%C3%A9rateurs,des%20particularit%C3%A9s%20d'un%20individu, consulté en juin 2020.
- [35] H. Charif. Segmentation d'images par une approche biomimétique hybride, Magister en génie électrique et électrotechnique, Université de Boumerdes 2018.
- [36] G. Balmisse. Algorithme génétique, http://iacenter.free.fr/AG/index_ag.htm, consulté en mai 2020.
- [37] A. A Bestaoui. Gestion de spectres dans un réseau de radio cognitive en utilisant l'algorithme d'optimisation par essaim de particules, mémoire de Master en Réseaux et Systèmes Distribués (R.S.D), Université Abou Bakr Belkaid, Tlemcen, juin 2015.
- [38] M. H. Eddine Krim. Etude de l'algorithme FCM pour la segmentation des images satellites landsat 5, mémoire de Master en Modalisation Informatique des Connaissances et de Raisonnement « MICR ». Université Tahar Moulay, Saïda, Juin 2014.
- [39] H. Shi., M. Xu. A Data Classification Method Using Genetic Algorithm and K-Means Algorithm with Optimizing Initial Cluster Center, <https://ieeexplore.ieee.org/document/8542173>, 2018, consulté en juin 2020.
- [40] P. Ghamisi, M. S. Couceiro, J. A. Benediktsson, N. M.F. Ferreira. An efficient method for segmentation of images based on fractional calculus and natural selection, https://www.academia.edu/5853429/An_efficient_method_for_segmentation_of_images_based_on_fractional_calculus_and_natural_selection, 2012, consulté en juin 2020.

Bibliographie

[41] S. Chaabane, N. Haddache. Segmentation d'images mammographiques, mémoire de master en électronique, Université de Saida, 2017

[42] S. Benkhaled. Une approche d'optimisation par essaim de particules quantiques pour la sélection d'attributs dans la classification de données, mémoire de Master en Informatique, Université Abbès Laghrour, Khenchla, 2017.

[43] Medical Image/Video Engineering Group, MIAS: Mammographic Image Analysis Society, <http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html>, consulté en avril 2020.

[44] A brief introduction to MATLAB, <https://web.stanford.edu/class/cme001/handouts/matlab.pdf>, 2006, consulté en juin 2020.