République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Chadli Bendjedid
Faculté des Sciences et de la Technologie
Département d'Informatique

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة الشاذلي بن جديد
كلية العلوم و التكنولوجيا
قسم الاعلام الالي

جامعة الشاذلي بن جديد
UNIVERSITE CHADLI BENDJEDID

MEMOIRE

Présenté par

Mender Bilal

Pour l'obtention de diplôme de

MASTER

Filière : Informatique

Spécialité : Systèmes Informatiques Intelligents

Thème

Modèle de sujets pour l'indexation des tumeurs mammographiques (application de LSA et SIFT)

Soutenu le : 29/06/2022 Devant le Jury composé de :

Qualité	Nom et Prénom	Grade	Université
Président	Mr. Betouil Ali-Abdelatif	MCB	Chadli Bendjedid El-Tarf
Rapporteur	**Mr. Touahri Djamel Eddine**	**MAA**	**Chadli Bendjedid El-Tarf**
Examinateur	Mme. Bougarne Imene	MCB	Chadli Bendjedid El-Tarf

Année Universitaire : 2021/2022

Remerciements

Je remercie d'abord le bon Dieu, le tout puissant de nous avoir donné la force, la puissance et la volonté, pour atteindre notre but, symbolisé par ce modeste travail.

Mes premiers remerciements à mon encadreur Dr.Touahri Djamel Eddine pour les considérables qu'il a fourni afin de m'aider et m'éclairer de ces précieux conseils et dont l'amabilité et la patiente exemplaire m'ont aidés à mener à bien mon travail, et surtout pour leur gentillesse.

Je remercie aussi les membres de jury d'avoir accepté d'examiner notre travail.

Enfin, je souhaite associer à ces remerciements ma famille et tous ceux que je connais de près ou de loin.

Dédicace

Je dédie ce travail à :

Mes parents.

A mon frère et ma petite sœur, je te souhaite une bonne continuation.

A ma fiancée.

A tous mes amis et tous ceux que j'aime.

Table des matières

Remerciements **2**
Dédicace **3**
Table des matières **4**
Liste des figures **6**
Liste des tableaux **8**
Liste des acronymes **9**

Introduction Générale **10**
1. Contexte du projet et problématique 10
2. Motivations 10
3. Objectifs 11
4. Contenu du mémoire 11

Chapitre 1 : Etat de l'Art **12**
1. Introduction 12
2. Section 1 : Système de recherche d'image par le contenu CBIR 13
    A. Historique 13
    B. Architecture générale d'un CBIR 13
    C. Composants d'un CBIR 14
    D. Domaines d'application d'un CBIR 17
3. Section 2 : Les modèles de sujets 18
    A. Introduction 18
    B. Les modèles de sujets 18
4. Section 3 : Indexation d'image médicale 25
    A. L'indexation par le contenu d'un document image 25
    B. Les caractéristiques locales 27
    C. Les caractéristiques globales 29
    D. Les systèmes de recherche par le contenu adaptés au domaine médical 35
    E. L'apprentissage automatique 38
    F. Critères d'évaluation 41
5. Conclusion 42

Chapitre 2 : Conception **43**
1. Introduction 43

2. Représentation de notre système 43
A. Architecture de notre système 43
B. Description des différentes étapes de notre Système d'indexation 44
C. Reconnaissance de type de tumeur par l'application de KNN 49
3. Conclusion 50

Chapitre 3 : Implémentation 51
1. Introduction 51
2. Le langage de programmation et bibliothèques 51
A. Python 51
B. Bibliothèques utilisées 51
3. L'environnement de développement 52
A. NetBeans 52
B. Matlab 53
C. Google Colab 54
4. Détails de l'application 55
A. La base d'images utilisée 55
B. Extraction des caractéristiques 55
C. Construction des mots visuels 56
D. Application du modèle LSA 57
5. Discutions et évaluations 60
A. Comparaison de nos résultats avec du travail similaire 62
B. Evaluations des résultats de la classification par KNN 63
6. Conclusion 64

Conclusion et Perspectives 65
Références 67
A. Références Bibliographiques 67
B. Références Web (Techniques) 69
Annexe A 70
Annexe B 72
Annexe C 75

Liste des figures

Figure 1. Un graphique montrant le nombre de publications contenant les mots "Image retrieval" entre 2005 et 2016 12
Figure 2. Architecture générale d'un système de recherche d'images par le contenu. 13
Figure 3. Des images issues de la base IRMA 15
Figure 4. Exemples de formes de requêtes. 16
Figure 5. Diagramme à blocs de différentes applications du CBIR 17
Figure 6. Matrice termes-documents ; chaque élément de la matrice spécifie la fréquence d'apparition du mot de la ligne dans le document de la colonne 18
Figure 7. Factorisation de la matrice pour le modèle LSA 21
Figure 8. La représentation graphique du modèle pLSA 22
Figure 9. Modèle graphique du LDA 23
Figure 10. La représentation graphique du modèle CTM 24
Figure 11. La représentation graphique du modèle PAM 24
Figure 12. Résultats de l'interrogation pour la base de données des tumeurs du cerveau (une ligne par signature) 26
Figure 13. Exemple de régions d'intérêts choisies pour l'extraction des caractéristiques locales 27
Figure 14. Les phases d'extraction des caractéristiques locales 27
Figure 15. Détection des extrêmes par comparaison du pixel d'intérêt avec ses voisins du niveau courant ainsi que les niveaux adjacents 28
Figure 16. 2x2 vecteurs descripteurs calculés à partir d'un échantillon de 8x8 29
Figure 17. Une image et les points de caractéristiques locales extraites 29
Figure 18. Exemple de l'histogramme d'une image en niveaux de gris 30
Figure 19. Architecture détaillée de notre système 44
Figure 20. Préparation des images de la base 45
Figure 21. Différentes étapes de la description d'un point clé avec l'algorithme SIFT. Le point à décrire est représenté en rouge. Le cercle sur la figure du milieu illustre la gaussienne utilisée pour pondérer les amplitudes du gradient avant de construire l'histogramme final 46
Figure 22. Illustration du principe de construction de mots visuels 48
Figure 23. Illustration des composants de packadge Jama 52
Figure 24. La page principale de NetBeans IDE 13 53
Figure 25. Interface de Matlab R2013 54
Figure 26. Exemple des images de la base d'apprentissage 55
Figure 27. Un fragment de vecteur descripteur SIFT d'une image de la base 55

Figure 28. Un fragment de vecteur SIFT (positions de frames) d'une image de la base.....56
Figure 29. Un fragment de texte généré à la base des caractéristiques visuelles locales. Les mots sont des mots visuels 56
Figure 30. Un échantillon de la matrice co-occurrence « mot-visuel-image » de la base d'apprentissage.....57
Figure 31. Lecture des images d'apprentissage 58
Figure 32. Lecture des images de test.....58
Figure 33. Un échantillon des mots visuels dans une image 59
Figure 34. Calcul de la distance euclidienne entre l'image de test et les images de la base.....59
Figure 35. Montre le numéro et la position des images similaires pour la première image requête .60
Figure 36. Graphes des résultats de MAP par catégories de requêtes 61
Figure 37. Histogramme des Résultats globales de l'évaluation du modèle de classification KNN.64
Figure 38. Une partie du code présente les paramètres de la fonction SIFT 72
Figure 39. Construction de l'espace des échelles; l'image originale (a) et les octaves (b) 73
Figure 40. Une partie du code de la construction de l'espace des échelles et détection des extrémas 73
Figure 41. Création de l'histogramme des gradients 74
Figure 42. Une partie du code du résultat de descripteur SIFT 74
Figure 43. A gauche : Une partie du code pour calculer tf-idf des images de train, à droite : le résultat - matrice co-occurrence (lignes : 20 mots et colonne : 140 image) pour la base train 75
Figure 44. Une partie du code de la méthode LSA et la méthode distance euclidienne.....76
Figure 45. Une partie du code pour l'évaluation du MAP de LSA 76

Liste des tableaux

Tableau 1. Table de calcul des angles de l'ellipse équivalente à une région 32
Tableau 2. Différents types d'images et systèmes utilisant ces images 38
Tableau 3. Critères de distinction entre tumeurs bénignes et tumeurs malignes 45
Tableau 4. Résultats de l'évaluation du MAP de LSA avec les caractéristiques locales SIFT pour l'indexation des tumeurs mammaires 61
Tableau 5. Comparaison de nos résultats et résultats du travail [1] 62
Tableau 6. Résultats de l'évaluation de la reconnaissance de type de tumeur 63

Liste des acronymes

CBIR	Content-Based Image Retrieval; *recherche d'images Basée sur le Contenu*
QBIC	Query By Image Content; *requête par le contenu visuelle d'image exemple*
CIRES	Content Based Image Retrieval System; *système de recherche d'images basé sur le contenu*
CQA	Customized Queries Approach ; *approche des requêtes personnalisées*
QDMFF	Query Independent Multiview Features Fusion
LSA	Latent Semantic Analysis; *L'analyse sémantique latente*
PLSA	Probabiliste Latent Semantic Analysis ; *l'analyse sémantique latente probabiliste*
LDA	Allocation Latent Dirichlet; *modèle allocation de Dirichlet latente*
CTM	Correlated Topic Model*; modèle de sujets corrélée*
PAM	Pachinko Allocation Model; *modèle d'allocation de pachinko*
SVD	Singular Values *Decomposition; décomposition en valeurs singulières*
DoG	Difference of Gaussian ; *différence des Gaussiennes*
SIFT	Salient Invariant Feature Transform ; *transformation de caractéristiques visuelles invariante à l'échelle*
SURF	Speed Up Robuste Features ; *caractéristiques robustes accélérées*
IRM	Magnetic Resonance Imaging*; l'imagerie par résonance magnétique*
TEP	Positron Emission Tomography*; tomographie permission de positons*
US	Ultrasonic Imaging*; l'imagerie ultrasonore*
MDR	Searching Multimedia Documents*; recherche de documents multimédia*
AIA	Automatic Image Annotation*; l'annotation automatique d'images*
KNN	k Nearest Neighbor ; *méthode des k plus proches voisins*
LLSF	Linear Least Square Fit ; *la méthode des* moindres carrés
NB	Naive Bayes ; *classification naïve bayésienne*
SVM	Supports Vectors Machines ; *machines à vecteur support*
MAP	Mean Average Precision ; *moyenne des précisions moyennes*
BOVW	Bag Of Visual Words ; Sac de mots visuels
EDI	Integrated Development Environment*; un environnement de développement intégré*
CDDL	Common Development and Distribution License; *licence commune de développement et de distribution*
MATLAB	Matrix Laboratory; *laboratoire de matrice*
MIAS	Mammographic Image Analysis Society *; la société d'analyse des images de mammographie*

Introduction Générale

Le domaine de l'image numérique est un domaine en pleine expansion, il est devenu le cœur de tous les secteurs d'activités dans le monde médical, géographique,…etc. Pour gérer et utiliser efficacement les bases d'images, un système d'indexation et de recherche d'images est nécessaire. C'est pourquoi le sujet de la recherche d'images devient un sujet très actif dans la communauté internationale depuis plus d'une dizaine d'années.

Les premiers systèmes étaient basés sur la recherche par mot-clés, ces systèmes ont montré quelques limites à cause de la subjectivité des mots-clés attribués. Ces limites ont conduit à la naissance des systèmes d'indexation et de recherche d'images par le contenu physique de l'image (CBIR, en anglais Content-Based Image Retrieval).

Le contenu d'une image possède des caractéristiques permettant de la résumer par des métriques mathématiques appelés descripteurs, ces descripteurs sont fondés sur des caractéristiques visuelles comme la couleur, la texture, la forme…etc. Ces caractéristiques, dites de bas niveau, peuvent être calculées globalement sur l'image (descripteur global), comme ils peuvent être calculés au niveau local.

1. Contexte du projet et problématique

La mammographie constitue le moyen d'investigation le plus utilisé dans le diagnostic des lésions mammaires. Cependant, Les techniques existantes peuvent être insuffisantes pour montrer les structures du sein et faire apparaître les anomalies présentes et le médecin peut faire appel à d'autres modalités d'imagerie telle que l'imagerie IRM.

Une des difficultés majeures qui se pose dans le domaine de la recherche des images par le contenu visuel est le ***fossé sémantique*** existant entre une image et son sens. C'est-à-dire, à partir d'une image, retrouver ce qu'elle cherche à exprimer.

2. Motivations

Afin de traiter cette problématique, et dans le but d'aider les médecins et faciliter la tâche de diagnostique, nous nous proposons un système d'indexation d'images médicale.

3. Objectifs

Nous nous concentrons dans ce travail sur l'application des modèles de sujets dans le contexte de l'indexation et recherche d'image. Le défi consiste d'essayer à attribuer aux caractéristiques d'une image médicale des concepts sémantiques.

Pour cette raison nous proposons un système d'indexation qui s'exécute en trois étapes. La première est l'extraction des descripteurs SIFT (caractéristique locale) pour chaque image, la deuxième est la construction des mots visuels et la troisième c'est l'application de modèle de sujet LSA pour obtenir l'index de l'image.

Pour enrichir notre travail nous présentons une deuxième partie qui consistant à retrouver le type de tumeur de la requête en utilisant l'algorithme de classification KNN (K-nearest Neighbor) sur les résultats de la recherche.

4. Contenu du mémoire

Ce mémoire est constitué de trois chapitres principaux ainsi qu'une introduction et une conclusion générale.

- Le premier chapitre est consacré à l'état de l'art qui regrouper trois sections ; dans la première section, on va définir et présenter l'architecture générale ainsi que les domaines d'application d'un système d'indexation et recherche d'image par contenu (CBIR), puis dans la deuxième section, on va parler sur les différents types de modèles de sujets dans le contexte de modélisation de document. Dans la section troisième, on va entamer le sujet en analysant les travaux connexes du domaine et l'indexation d'image dans le domaine médical ainsi que quelques systèmes de recherche et d'indexation existants pour ce type d'image.
- Dans le chapitre deuxième nous décrivons la conception et l'architecture générale de notre système.
- En ce qui concerne le dernier chapitre, on va aborder l'implémentation et les détails techniques matériels, la plate-forme et les détails de programmation et l'interface de système et nous mettons en œuvre les difficultés rencontrées pour la mise du point du système.

1. Introduction

Le développement rapide dans le domaine des caméras numériques et des outils d'acquisition, de transmission et de stockage d'images présente des défis majeurs pour les utilisateurs. Ce défi se résume à la nécessité d'accéder et d'utiliser efficacement la grande quantité d'images disponibles.

Diverses tâches dans les domaines de l'indexation et de la recherche d'images assurent l'utilisation optimale des bases d'images et la fourniture d'outils permettant un traitement satisfaisant pour l'utilisateur.

La recherche sur l'indexation d'images a été fortement dynamisée grâce aux différentes techniques et approches qui ont été discutées et proposées depuis les années 1980. La figure 1 montre la variation annuelle du nombre de publications contenant le terme «Image Retrieval » de 2005 à 2016.

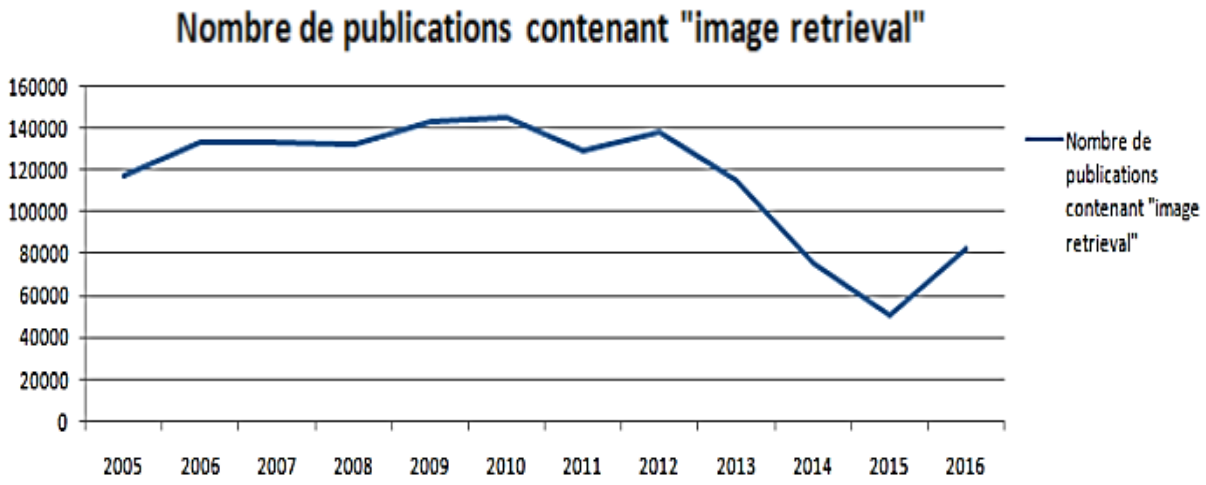


Figure 1. Un graphique montrant le nombre de publications contenant les mots "Image retrieval" entre 2005 et 2016 [1]

Les systèmes de recherche d'images basés sur le contenu (CBIRS) sont au centre de ce domaine de recherche. Ces systèmes fournissent aux utilisateurs des outils efficaces pour effectuer des recherches. Étant donné que la recherche en indexation se concentre sur les performances de ces systèmes, donc l'objectif est de développer des systèmes puissants qui répondent efficacement aux besoins des différents utilisateurs.

Nous allons présenter dans la section suivante les systèmes de recherche d'images à base de contenu (CBIRs).

2. Section 1 : Système de recherche d'image par le contenu CBIR

A. Historique

Les systèmes d'indexation et recherche d'images par le contenu (CBIR) permettent de rechercher les images d'une base d'images en fonction de leurs caractéristiques visuelles.

Le premier prototype de système CBIR a été proposé en 1970 et ce système a attiré l'attention de beaucoup de chercheurs. Quelques systèmes deviennent des systèmes commerciaux tels que QBIC (Query By Image Content), CIRES (Content Based Image Retrieval System).

L'expression "recherche d'images par le contenu" remonte aux travaux de Kato en 1992. Son système, ART MUSEUM, permet de retrouver des images d'art par couleurs et contours. Le terme s'est étendu par la suite à tout procédé permettant de rechercher des images selon des traits, pouvant être de type « signal », comme la couleur et la forme, mais également symboliques. **[2]**

B. Architecture générale d'un CBIR

Les systèmes de recherche d'images représentent un cas spécial des systèmes de recherche d'information. Les images sont l'information que l'utilisateur cherche à récupérer de la base en utilisant ces systèmes. L'architecture de base de CBIR est illustrée par la figure 2.

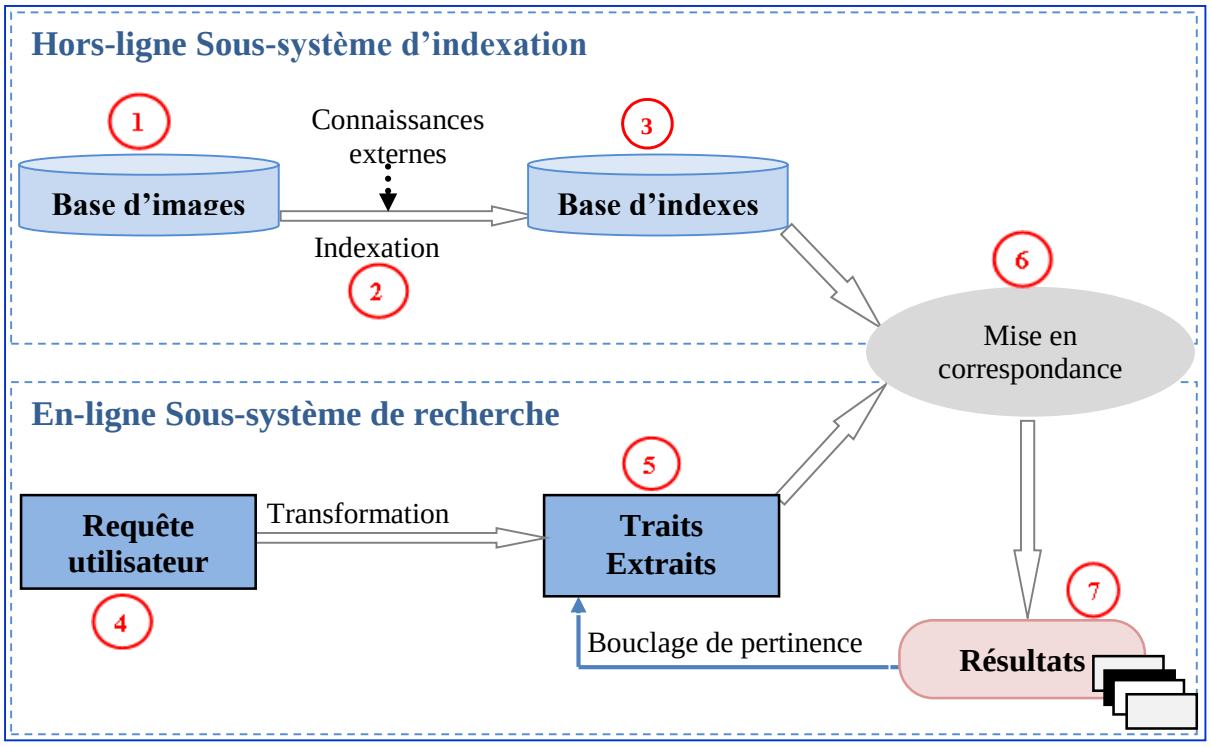


Figure 2. Architecture générale d'un système de recherche d'images par le contenu.

Les systèmes de recherche d'images basés sur le contenu ont deux aspects indissociables : l'indexation et la recherche.

La phase hors ligne est la phase d'indexation dans laquelle les images trouvées dans la base de données sont indexées à l'aide d'un ensemble de caractéristiques typiquement extraites de ces images. La représentation du contenu de cette image à l'aide de l'ensemble de caractéristiques extrait réduit la taille utilisée dans le processus de recherche. La pertinence de ces caractéristiques est essentielle à la bonne représentation du contenu des images dans la base.

La phase de recherche d'images est la phase en ligne. Les requêtes des utilisateurs doivent être traitées dans l'ordre et traduites en une représentation similaire aux caractéristiques de l'image dans la base.

C. Composants d'un CBIR

Voici une brève description des similitudes dans la plupart des étapes, telles que le traitement des bases d'images, les requêtes, la mise en correspondance et la présentation des résultats.

Premièrement (2) le descripteur est calculé à partir de chaque image de la collection (1), il peut être essentiellement un signal et/ou symbolique (vocabulaire d'indexation). Les données extraites (qui représentent maintenant le contenu de l'image du point de vue du système) forment une base d'index (3). La requête de l'utilisateur (4) est ensuite transformée pour être comparable à la requête indexée (5). La mise en correspondance entre la requête transformée et la base d'index (6) permet de générer le résultat de la requête (7). Le système peut également comprendre des composants liés à la personnalisation, comme l'extraction, le stockage et l'utilisation d'un profil d'utilisateur.

1. La base d'image

La collection d'images est la donnée principale du système. Les bases de données d'images diffèrent principalement par leur taille. La plupart des systèmes sont conçus pour des bases de données contenant des centaines ou des milliers d'images. La base d'images standards Wang est illustrée dans la figure 3.

Le type d'image a un effet significatif sur la conception globale du système, en particulier sur les descripteurs de bas niveau calculés. D'une manière générale, plus la variation dans et entre les images est grande, plus le système doit être riche et précis (et plus le problème d'indexation/recherche de ces images est difficile).

Figure 3. Des images issues de la base IRMA [W1]

2. L'indexation

Où les images trouvées dans les bases sont indexées en utilisant un ensemble de caractéristique généralement extraites à partir de ces images. Cette représentation du contenu de l'image en utilisant l'ensemble de caractéristiques extraites permet la réduction de la dimension utilisée durant la recherche, la pertinence de ces caractéristiques est primordiale pour assurer la bonne représentation du contenu des images de la base. Donc, l'indexation est l'ensemble des processus aboutissant à la construction d'un index de l'image.

Une fois les images indexées, ils peuvent être recherchées avec les modèles classiques de recherche tels que le modèle booléen, le modèle vectoriel et le modèle probabiliste.

3. La gestion des index

Cela affecte la gestion des index d'images (stockage et accès). Lorsque vous travaillez avec de grandes bases de données, la maintenance des index est une préoccupation majeure pour les collections de taille modeste. Le moyen le plus simple de stocker un index consiste à le stocker sous forme de liste séquentielle dans la mémoire ou dans un fichier. Cependant, plus le nombre d'images augmente, plus le temps d'accès aux images augmente linéairement, et les index s'organisent hiérarchiquement sous forme d'arbres (organisés par descripteurs) ou de tableaux pour accélérer l'accès aux informations nécessaires.

4. Les requêtes

Les requêtes peuvent prendre plusieurs formes dans un système d'indexation et recherche des images : soit une requête par mots clés, soit une requête par esquisse, soit une requête par exemple. La figure 4 montre des exemples de requêtes.

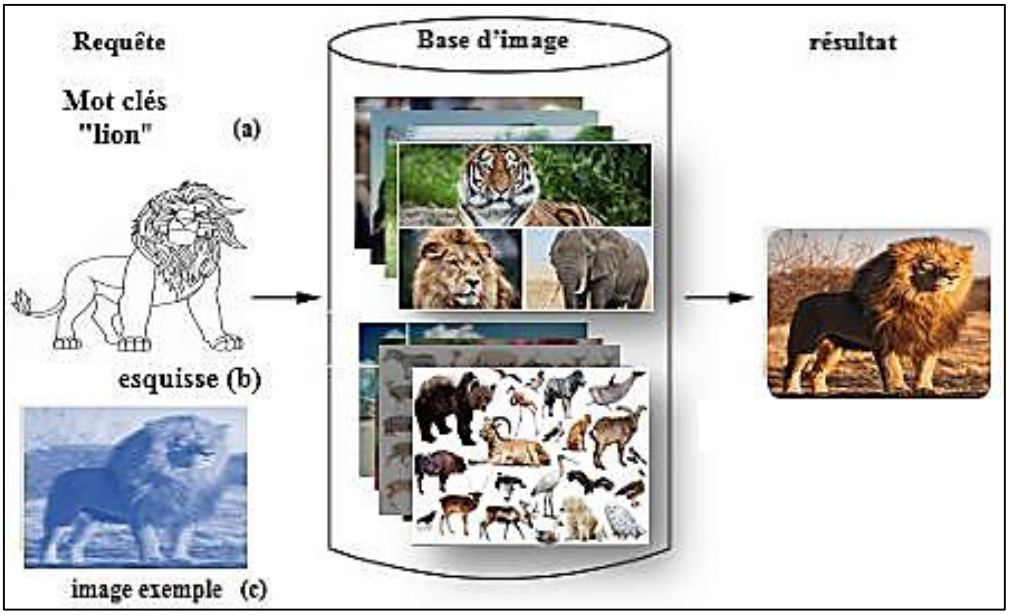


Figure 4. Exemples de formes de requêtes.

5. Analyse de la requête

Le but de cette étape est de transformer la requête de l'utilisateur pour qu'elle corresponde à l'index de la base d'images. Par conséquent, vous extrairez généralement le même type de descripteur que celui extrait de la base d'images lors de l'indexation.

6. Mise en correspondance requête / base

Estimez à quel point une image (son index) remplit une requête particulière. Dans le cadre de la recherche d'images, cela se résume souvent au calcul de la similarité entre les caractéristiques extraites de la requête et les caractéristiques de chaque image de la base. Cela donne généralement une valeur de correspondance qui caractérise la pertinence de l'image (du point de vue du système) par rapport à la requête.

Il existe plusieurs mesures de similarité utilisées pour définir la distance entre deux descripteurs d'image, comme la distance Euclidienne, la distance Cosine…etc.

7. La présentation des résultats

Dans la grande majorité des systèmes, le résultat d'une requête est présenté sous la forme d'une liste d'images (réduites à des vignettes) ordonnées par pertinence décroissante. Cette présentation peut prendre d'autres formats.

La présentation des résultats est une option d'interaction qui affine souvent la requête en indiquant des résultats pertinents et ceux qui ne le sont pas (bouclage de pertinence), ce qui permet de reformuler automatiquement la requête.

8. La phase de bouclage de pertinence

Eventuellement, on peut utiliser une phase de bouclage dans le système. Cette phase est le résultat de l'interaction des utilisateurs à travers plusieurs session avec les résultats retournés par le système, l'utilisateur participe à l'évaluation de ces résultats en jugeant leurs pertinence vis-à-vis ses requête.

Cette participation de l'utilisateur aide le système à ajuster les paramètres internes utilisés dans l'indexation.

D. Domaines d'application d'un CBIR

Le système de recherche d'images CBIR peut être associé à un grand nombre d'applications du monde réel et les principales applications sont orientées vers les types suivants : applications médicales, recherche d'images de télédétection, recherche d'images naturelles, applications médico-légales, applications de sécurité, applications commerciales et les applications diverses (voir la figure5).

- **Application médicale**

L'utilisation du CBIR peut donner lieu à des services puissants qui peuvent bénéficier aux systèmes d'information biomédicaux. Trois grands domaines peuvent profiter instantanément des techniques CBIR : l'enseignement, la recherche et le diagnostic. Plusieurs applications a été développé dans ce domaines tels que I-Browse, Customized Queries Approach (CQA) et Query Independent Multiview Features Fusion (QDMFF). **[3]**

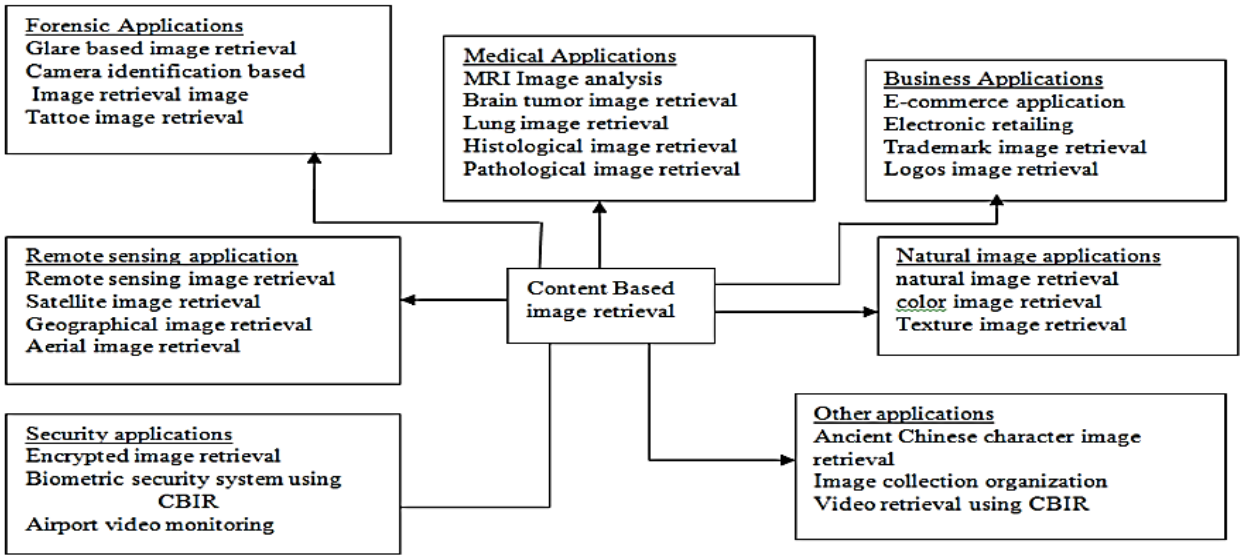


Figure 5. Diagramme à blocs de différentes applications du CBIR [3]

Dans la prochaine section, on va présenter les différents types de modèles de sujets dans le contexte de modélisation de document qui représentent un type particulier des modèles probabilistes.

3. Section 2 : Les modèles de sujets

A. Introduction

Les modèles de sujets sont utilisés depuis les années 90 en RI, ils ont été développées à l'origine pour les grandes collections de textes et ensuite adaptés aux images. Le premier modèle de sujet qu'on va présenter se nomme analyse de la sémantique latente LSA (*Latent Semantic Analysis*). Celle-ci **[4]** a été proposée comme approche vectorielle qui vise à découvrir une structure sémantique latente dans les collections de documents. Des extensions probabilistes ont été proposées pour ce modèle dans les deux dernières décennies.

Le modèle de sujet probabiliste était « *probabiliste Latent Semantic Analysis* » (PLSA), il introduit par Hofmann **[5]**. Chaque document est décrit par un mélange de sujets et à son tour, chaque sujet est caractérisé comme une distribution sur les mots dans un vocabulaire fini. Un nombre fixe de sujets est utilisé pour modéliser les documents dans la base de données. Le PLSA a ensuite été étendu par **[6]** pour un modèle entièrement génératif, appelé Allocation Latent Dirichlet (LDA). Un certain nombre d'extensions de ces modèles ont été proposées. Ils intègrent des structures hiérarchiques, modèle de sujets corrélée CTM **[7]** et Le modèle d'allocation de pachinko (PAM) **[8]**.

B. Les modèles de sujets

Les modèles de sujets sont classés en deux catégories : modèles à concepts déterministes et modèles à concepts probabilistes.

Tous ces modèles utilisent au départ une matrice de cooccurrence termes-documents. Etant donné que la base contient M documents $d_1, \dots, d_M$, le contenu des documents est issu d'un vocabulaire composé de N mots $w_1, \dots, w_N$. Les lignes de la matrice contiennent les mots de vocabulaire, et les colonnes contiennent les documents de la base. Chaque élément de la matrice représente la fréquence du mot de la ligne dans le document de la colonne, voir figure 6.

$$\begin{pmatrix} f(w_1,d_1) & \cdots & f(w_1,d_M) \\ \vdots & \ddots & \vdots \\ f(w_N,d_1) & \cdots & f(w_N,d_M) \end{pmatrix}$$

Figure 6. Matrice termes-documents ; chaque élément de la matrice spécifie la fréquence d'apparition du mot de la ligne dans le document de la colonne [1]

Le modèle conceptuel tente d'extraire de la matrice de cooccurrence un espace latent contenant des concepts cachés impliqués dans l'existence des mots dans le document. Ces concepts représentent les sujets (thèmes ou concepts) qu'ils contiennent les documents. Cet espace latent est utilisé pour

l'indexation et la recherche de documents au lieu d'utiliser la matrice initiale. L'utilisation de cet espace latent présente les avantages suivants :

- Le nombre de concepts est très faible par rapport au nombre de vocabulaires, ce qui réduit l'espace de recherche. Cette réduction accélère la recherche à la base.
- L'utilisation des concepts de l'espace latent permet de résoudre des problèmes tels que l'ambiguïté, les synonymes et l'absence.

B.1 Modèles à Concepts déterministes

Cette catégorie contient les modèles à concepts qui utilisent des techniques déterministes pour extraire les concepts à partir des documents. LSA est un exemple de ces techniques que nous allons présenter dans la section suivante et qu'on va utiliser dans notre travail.

❑ Latent Semantic Analysis (LSA)

LSA est une technique statistique automatique pour extraire et inférer des relations entre mots à partir de leur contexte. **[9]**

- **Principe**

Le sens d'un mot peut être défini statistiquement à partir de l'ensemble des contextes (phrases, paragraphes, textes) dans lesquels ce mot apparaît. Par exemple, le mot *autobus* sera souvent conjointement associé à *démarrer*, *route*, *gare routière*, et rarement à *fleur, barbecue*. Cependant, le contexte du mot n'est pas suffisant pour en définir le sens, car il ne dit rien sur les relations avec les mots qui n'apparaissent jamais ensemble. Par exemple, si les mots *autobus* et *autocar* n'apparaissent jamais ensemble, nous n'avons aucune information sur les liens sémantiques entre ces mots. Or *autocar* doit être considéré comme proche de *autobus* car tous les deux sont co-occurrents avec les mêmes mots. Ce sont donc des enchaînements de liens de cooccurrences à plusieurs niveaux qui permettent une représentation correcte du sens des mots. Pour résoudre cette difficulté, LSA construit une matrice de cooccurrences, constituée du nombre d'apparitions de chaque mot dans chaque contexte, sans tenir compte de leur ordre. Cette matrice est ensuite réduite à l'aide d'une décomposition en valeurs singulières (SVD) (généralisation de l'analyse factorielle) afin de capturer dans une certaine mesure les relations entre les mots et les documents et en espérant que les mots ayant un sens voisin (en particulier les synonymes) auront la même direction dans le nouveau sous-espace.

Dans le modèle vectoriel, chaque document est représenté sous la forme d'un vecteur (représentation « *sac de mots* »). On peut ainsi construire la matrice terme-document prenant en ligne les mots du lexique et en colonne les documents. Chaque case représente la fréquence d'un mot dans un document.

Un problème du modèle vectoriel est que les mots utilisés dans la requête ne sont pas forcément les mêmes que les mots utilisés dans les documents pertinents. En effet, des mots similaires peuvent avoir différents sens (*polysémie*) et différents mots peuvent avoir le même sens (*synonymie*). Pour résoudre ce problème, LSA utilise une SVD afin de prendre en compte le contexte des mots et de réduire les dimensions de l'espace.

Soit A la matrice terme-document et r son rang, par la méthode algébrique de décomposition en valeurs singulières, elle peut être écrite sous la forme du produit de trois matrices telles que :

$$A = W\sum D^t \qquad \begin{cases} A \in R^{n_W \times n_D} \\ W \in R^{n_W \times r} \\ \sum \in R^{r \times r} \\ D \in R^{n_D \times r} \end{cases} \tag{1}$$

où W et D sont des matrices orthonormales contenant les vecteurs singuliers gauche et droit de A, et $\sum$ est une matrice diagonale contenant les valeurs singulières de A :

$$\begin{cases} \Sigma = diag(\sigma_1, \sigma_2, \dots, \sigma_r) \quad avec\ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \\ WW^t = DD^t = I_r \\ rang(\Sigma) = \mathrm{r} \leq \min(n_W, n_D) \end{cases} \tag{2}$$

Pour diminuer le nombre de dimensions de l'espace, et si l'on suppose que les valeurs singulières de la matrice diagonale $\sum$ sont ordonnées, alors on peut trouver une bonne approximation $\bar{\mathrm{A}}$ de A en mettant à zéro les petites valeurs singulières de $\sum$ afin d'obtenir un espace réduit de dimension r^-choisie :

$$A \cong \bar{\mathrm{A}} = W^-\sum{}^- D_t^- \qquad \begin{cases} \bar{\mathrm{A}} \in R^{n_W \times n_D} \\ W^- \in R^{n_W \times r^-} \\ \sum{}^- \in R^{r^- \times r^-} \\ D^- \in R^{n_D \times r^-} \end{cases} \tag{3}$$

Où

$$\begin{cases} \sum{}^- = diag(\sigma_1, \sigma_2, \dots, \sigma_{r^-}) \quad avec \begin{cases} \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{r^-} \\ et \quad \sigma_{r^-+1} = \cdots = \sigma_r = 0 \end{cases} \\ W^-W_t^- = D^-D_t^- = I_{r^-} \\ rang(\sum{}^-) = r^- \leq rang(\sum{}^-) = \mathrm{r} \leq \min(n_W, n_D) \end{cases} \tag{4}$$

Cette opération réalise la projection de l'espace original vers un espace réduit à r^-dimensions. Il a été démontré que, sous certaines conditions, l'espace réduit capture dans une certaine mesure les relations sémantiques entre les mots du corpus. Le nombre de dimensions optimal de l'espace réduit pour la langue anglaise a été estimé empiriquement à 300 dimensions.

La SVD effectue un changement de base pour se placer suivant les axes de plus grande variation de la matrice A. De manière intuitive, on peut se représenter un mot comme un point dans un espace dont la dimension est le nombre de documents n_D. La matrice A donne les coordonnées des n_W mots. Ce nuage de points a des axes d'inertie qui sont précisément les axes de plus grande variation de A. En tronquant aux r^-premières valeurs singulières, on conserve les axes d'inertie suivant lesquels s'alignent le mieux les points du nuage. Ainsi on capture la structure la plus significative de la matrice. Il faut voir la décomposition en valeurs singulières comme une méthode qui réduit la dimension du problème et, surtout, qui permet de représenter mots et documents dans un même espace de dimension r^-. L'espace de dimension r^-s'interprète comme un espace de concepts. On ne peut pas vraiment espérer mettre un nom sur ces concepts. Mais ce n'est pas gênant : tout ce dont on a besoin est de savoir dans quelle mesure les différents concepts (abstraits) sont présents dans tel mot et tel document, de manière à comparer ceux-ci. Mathématiquement, puisque le mot et le document sont représentables dans un même espace, un simple calcul de la distance entre leurs représentants fournit une quantification de leur proximité. Au final, les documents renvoyés peuvent ne contenir aucun mot de la requête mais être pertinents.

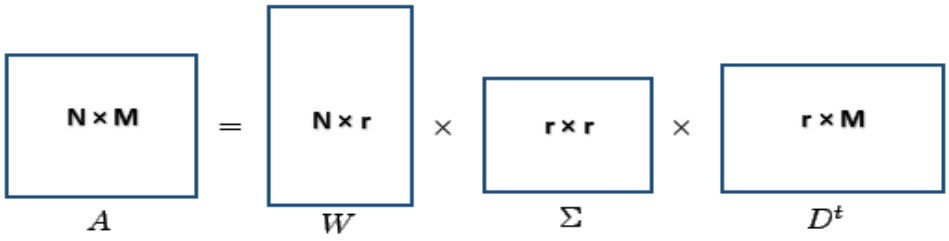


Figure 7. Factorisation de la matrice pour le modèle LSA

Pour comparer les documents dans l'espace réduit à un vecteur requête d_q, nous transformons tout d'abord le vecteur d_q en un pseudo document d_q^- dans l'espace réduit. Nous avons $X = W^- \sum^- D_t^-$ que nous pouvons dériver en $D^- = X^t W^- \sum^{-1}$. La vecteur ligne d_q^- dans l'espace réduit peut donc être obtenu par :

$$d_q^- = d_q^t W^- \sum^{-1} \qquad (5)$$

Il contient le contexte associé au document. Nous pouvons alors utiliser une mesure de similarité classique comme le cosinus pour calculer la distance entre d_q et chacun des documents dans l'espace réduit. De même, pour un mot W_q , nous pouvons obtenir sa représentation W_q^- dans l'espace réduit par :

$$W_q^- = \sum\nolimits_{-1}^{-1} D_{qt}^- W_q^t \qquad (6)$$

Pour mesurer la similarité entre le document i et le document j, il suffit de réaliser le produit scalaire entre les vecteurs lignes i et j de la matrice $D^{-}\sum^{-}$. Il est aussi possible de calculer les p mots les plus pertinents pour un document.

B.2 Les modèles à concepts probabilistes

La liste des modèles à concepts que nous allons présenté dans le reste de ce chapitre contient les modèles à concepts probabiliste qui diffèrent dans leur nature du modèle LSA présenté avant. Nous allons présenter dans cette section les modèles PLSA, LDA, CTM et PAM. Ces modèles représentent les documents par un modèle de mélange de concepts probabiliste. Les modèles à concepts probabilistes sont basés sur l'idée que les documents peuvent être modélisés par un mélange de concepts. Les modèles à concepts probabilistes sont des modèles génératifs spécifiant des procédures probabilistes permettant de générer les documents, en d'autre terme les règles probabilistes décrivant comment les mots peuvent être générés à la base des variables latentes (concepts). **[1]**

Des techniques d'inférence statistiques peuvent être utilisées pour inverser le processus et inférer les concepts qui ont été responsable de la génération des mots.

- **probabilistic Latent Semantic Analysis (PLSA)**

Le modèle pLSA est la variante probabiliste du modèle LSA. Le processus génératif du pLSA pour un document d_i est le suivant :

- Choisir un document avec une probabilité à priori $P(d_i)$.
- Pour chaque mot de la liste des mots du document d_i:
  - Choisir un concept latent avec la probabilité $P(z_n \backslash d_i)$.
  - Générer un mot avec une probabilité $P(w_n \backslash z_n)$.

La figure 8 donne le modèle graphique du PLSA. La probabilité du mot w_i dans le document d_i est présentée dans la formule suivante :

$$P(w_i \backslash d_i) = P(d_i) \sum_{k=1}^{k} P(w_j \backslash z_j = k) P(z_j = k \backslash d_i) \qquad (7)$$

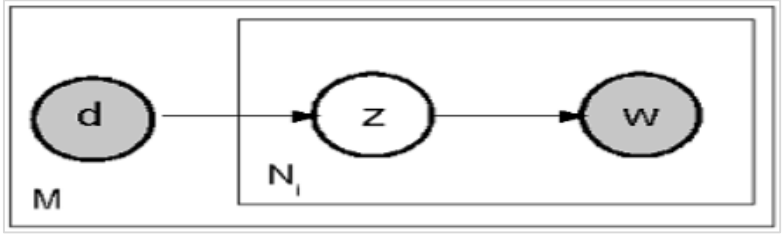


Figure 8. La représentation graphique du modèle pLSA [10]

❑ Latent Dirichlet Allocation (LDA)

Le modèle LDA est un modèle à concepts générative similaire au pLSA mais différents dans ses propriétés statistiques. Il spécifie aussi un modèle de mélange sur les concepts dont chaque concept est caractérisé par une distribution sur les mots, et chaque occurrence de mots dans un document est associée à un concept latent. Le LDA représente le mélange de concepts comme variable latente et pose un paramètre à priori de dirichlet sur celle-ci, la figure 9 montre le modèle graphique du LDA.

Le processus de génération des documents par le modèle LDA comporte les étapes suivantes :

- Choisir une variable aléatoire à K-dimensions $\theta_i \sim \mathrm{Dir}(\alpha)$, K est le nombre de concepts dans la base.
- Pour chaque mot de la listeN_i du document *i*
  - Choisir un concept $z_n \sim \mathrm{Multinomial}(\theta_i)$
  - Générer la valeur du w_n du $n^{ème}$ mot avec la probabilité conditionnelle $P(w_n \backslash z_n, \beta)$.

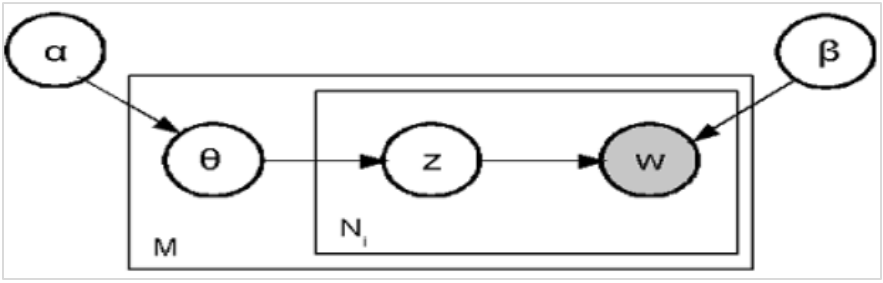


Figure 9. Modèle graphique du LDA [10]

La probabilité d'un document est donnée avec la formule suivante :

$$p(w_i \backslash \alpha, \beta) = \int P(\theta_i \backslash \alpha) \prod_{j=1}^{N_i} (\sum_{k=1}^{k} P(z_j = k \backslash \theta) P(w_j \backslash z_j = k, \beta)) d_\theta \qquad (8)$$

❑ Correlated Topic Model (CTM)

Le CTM est un modèle à concepts qui se diffère des autres modèles par ses propriétés statistiques. Le CTM choisit les proportions des concepts à partir d'une distribution normale logistique, la représentation graphique du modèle est illustrée dans la figure 10. Le processus génératif des documents pour le modèle CTM comporte les étapes suivantes :

- Choisir $\eta_i \ / \{\mu, \sum\} \sim N(\mu, \sum)$, avec μ comme vecteur de moyennes à K dimensions et $\sum$ une matrice de cooccurrence de taille K×K, Π une variable de mélange de concepts.
- Pour chacun des N_i mots dans le document *i* :
  - Choisir assignement des concepts à partir d'une loi multinomiale $f(\eta_i)$
  - Choisir un mot w_n à partir la probabilité multinomiale conditionnée avec les concepts z_n comme $p(w_n \backslash z_n, \beta)$.

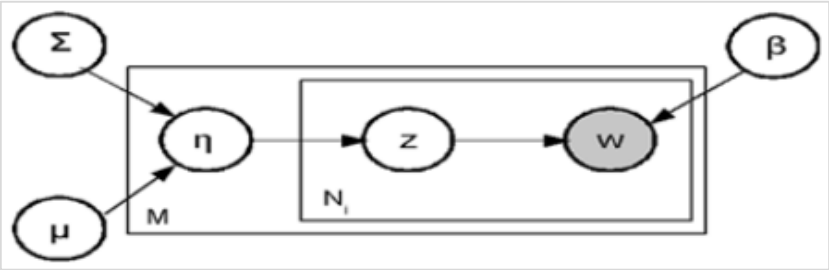


Figure 10. La représentation graphique du modèle CTM [10]

❑ Pachinko Allocation Model (PAM)

Le modèle PAM **[8]** diffère des autres modèles probabilistes par l'extraction de plusieurs niveaux des sujets. Ils s'intéressent non seulement par la capture de la corrélation entre les mots mais aussi entre les sujets eux même via l'extraction d'une couche supplémentaire des super-sujets. La figure 11 montre la représentation graphique du modèle PAM

Le processus génératif du modèle peut être décrit comme suit :

- Choisir $\theta_{t1}^{(d)}$, $\theta_{t2}^{(d)}$,…, $\theta_{ts}^{(d)}$ à partir de $g1(\alpha 1)$, $g2(\alpha 2)$,…, $gS(\alpha S)$ où $\theta_{ti}^{(d)}$ est une distribution multinomiale du sujet ti sur ses enfants.
- Pour chaque mot w dans le document ;
  - Echantillon de chemin du sujet Z_w de longueur L_w :< Z_{w1}, Z_{w2},…,Z_wL_w >, Z_{w1} est toujours la racine et Z_{w2} à Z_wL_w ont des nœuds dans le sujet T. Z_{wi} est un enfant de $Z_{w(i-1)}$ et il est échantillon en fonction de la distribution multinomiale $\theta_{Zw(i-1)}^{(d)}$.
  - Choisir mot de l'échantillon à partir de $\theta_{ZwLw)}^{(d)}$.

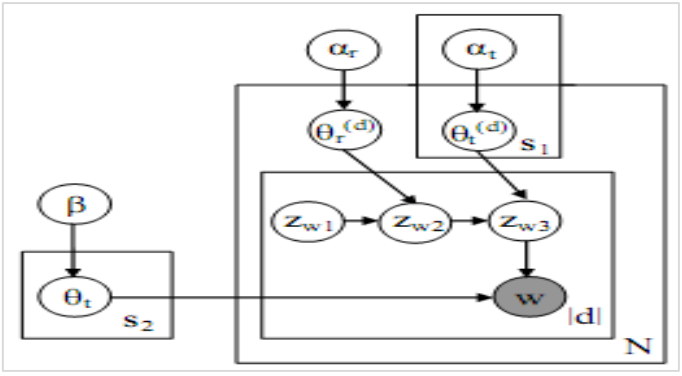


Figure 11. La représentation graphique du modèle PAM [8]

La section suivant sera dédié à l'étude détaillée sur l'indexation d'image médicale et décrire les travaux qui influencent fortement sur notre système d'indexation et de recherche d'images.

4. Section 3 : Indexation d'image médicale

A. L'indexation par le contenu d'un document image

A.1 Définitions et objectif

- Selon Marine Campedel l'indexation c'est l'exploitation d'une analyse fine du contenu du document visuel". **[W2]**
- Selon l'équipe de projet imedia **[11]** L'indexation par le contenu : C'est l'opération qui consiste à extraire d'un document (ici une image) des descripteurs visuels automatiques significatifs, compacts et structurés qui seront utilisés et comparés au moment de la recherche interactive.

Le but de l'indexation d'un document image est d'extraire et de représenter le contenu nécessaire et suffisant pour qu'il soit retrouvé par un utilisateur. Par conséquent, cette indexation est basée sur la représentation (supportée par le modèle) et le processus d'extraction. Pour éviter l'extraction d'informations non pertinentes dans un contexte particulier, chacun des éléments précédents doit en quelque sorte intégrer les besoins de l'utilisateur.

A.2 L'indexation en imagerie médicale

L'objectif est de développer des méthodes permettant de rechercher, dans des bases de données cliniques de référence, les cas semblables uniquement à partir du contenu numérique des images. L'image numérique est donnée en requête au système de recherche ou d'aide au diagnostic, et le système renvoie les cas clinique « renseignés » contenant des images similaires à l'image requête. Dans l'optique d'éviter tout biais, il faut développer des approches globales, qui ne reposent pas sur l'extraction de primitives particulières à des pathologies ou des organes. Ces approches permettent d'associer aux images de la base et aux images requêtes des signatures numériques. Deux points sont à considérer, conjointement ou séparément : d'une part la définition de la signature (index) d'une image, d'autre part les méthodes de comparaison des signatures.

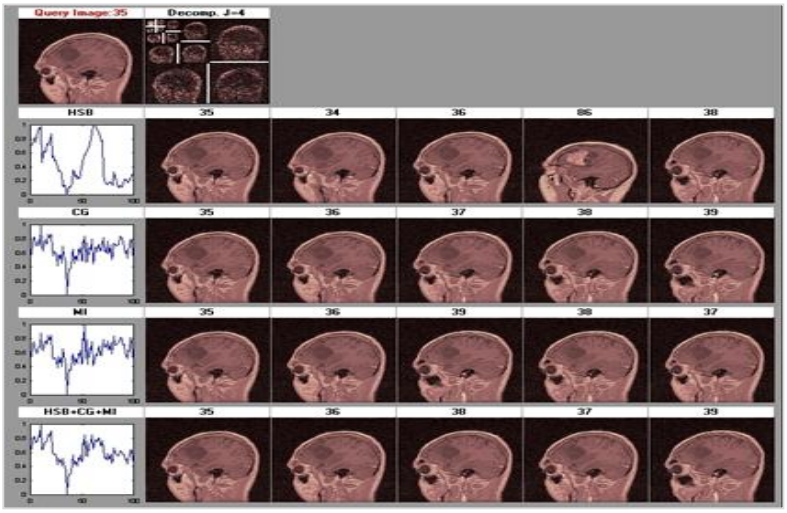


Figure 12. Résultats de l'interrogation pour la base de données des tumeurs du cerveau (une ligne par signature) [12]

A.3 Les différentes approches de l'indexation d'images

L'indexation d'images repose principalement sur une variété de fonctionnalités d'extraction et d'organisation pour représenter le contenu d'une image et est utilisée lors de la recherche d'images.

En général, il existe deux principales approches d'indexation : une approche textuelle basée sur l'indexation manuelle ou automatique du texte avec des requêtes basées sur des mots clés, et l'approche par le contenu basé sur le contenu visuel de l'image indexée lorsque la requête est soit une image ou croquis.

- **Indexation textuelle manuelle**

L'indexation manuelle du texte est effectuée par un opérateur humain. Il peut s'agir de simples utilisateurs qui tentent d'identifier et de catégoriser leurs collections personnelles, ou de documentalistes spécialisés dans l'indexation de collections d'images pour des organismes spécialisés tels que des agences de presse. Cette indexation a pour tâche de classer une image en l'associant à un groupe de mots décrivant une catégorie particulière. Décrit deux niveaux d'indexation manuelle ; le premier niveau concerne ce que l'indexeur affiche dans l'image, et le second niveau gère la signification de l'image. Le thésaurus facilite la tâche.

Les difficultés liées à ce type d'indexation peuvent être résumées dans les points suivants : la subjectivité et le volume important des données.

- **Indexation textuelle automatique**

L'indexation automatique des images est un processus nécessaire, même si l'image du consommateur a déjà été annotée manuellement. Exemple : Sur Internet, il s'agit d'un type d'indexation qui annote automatiquement les images sans intervention humaine.

❑ **Indexation par le contenu visuel**

Le principe de cette méthode est d'identifier l'image par son contenu (c'est-à-dire par les données de l'image elle-même, et non par le texte associé à l'image). Les caractéristiques de l'image doivent être extraites à l'avance pour l'indexation automatique des images.

On distingue usuellement les caractéristiques globales qui sont calculées sur toute l'image, par contre les caractéristiques locales sont calculées autour des points d'intérêt. La différence entre les caractéristiques globales et locales est taxonomiquement importante. Les caractéristiques locales sont distinctes en étant claires, robustes à l'occlusion (car il y a de nombreuses caractéristiques dans l'image ou la région) et ne nécessitant pas de segmentation. Ainsi, le descripteur local calculé pour chaque pixel de l'image ou région obtenue par segmentation et accumulé dans l'histogramme est une description globale de l'image ou région.

B. Les caractéristiques locales

Les caractéristiques de l'image sont calculées à des points spécifiques. La vue locale utilisée pour ce calcul contient un ensemble limité de pixels. Généralement, deux phases sont appliquées pour extraire les caractéristiques locales. La figure 13 montre un exemple d'imagettes sélectionnées pour calculer les caractéristiques locales de l'image.

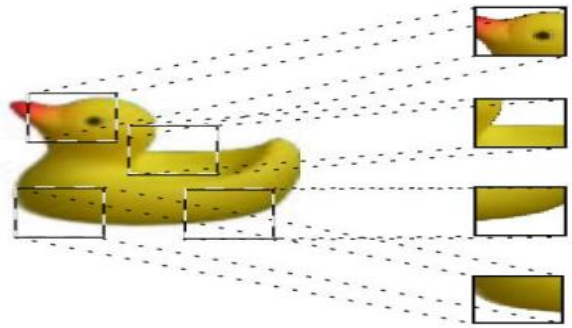


Figure 13. Exemple de régions d'intérêts choisies pour l'extraction des caractéristiques locales [9]

La première est la phase de détection des points locaux de l'image à laquelle est appliqué filtre détecteur de point à fort contraste. Un exemple d'un tel filtre est le détecteur DoG (Difference of Gaussian).

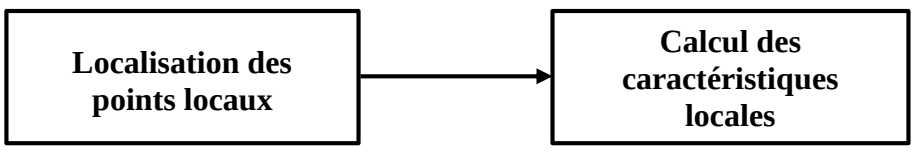


Figure 14. Les phases d'extraction des caractéristiques locales

SIFT (Salient Invariant Feature Transform) et SURF (Speed Up Robuste Features) sont deux méthodes d'extraction de caractéristiques locales des images. SIFT est calculé en présentant un histogramme dans la direction du gradient et SURF utilise l'approximation par ondelettes de Haar. Nous allons présenter la technique SIFT.

B.1 SIFT (Scale-invariant feature transform)

SIFT (scale-invariant feature conversion) développé par D. Lawe **[13]**, est très probablement le plus utilisé. Il a le grand avantage d'être invariant à la fois à la rotation et aux changements d'échelle. De plus, la densité des points de détection est élevée. Une particularité de cette procédure est le calcul combinatoire des points d'intérêt et des descripteurs associés. Les descripteurs sont des vecteurs qui caractérisent les voisins locaux aux points d'intérêt. Il caractérise le point d'intérêt par sa singularité.

SIFT utilise un détecteur DoG (Difference of Gaussian) pour identifier les points d'intérêt local dans une image. En commençant par la phase de détection du point d'intérêt, on passe à la phase de filtrage, stabilisant uniquement le point avec moins d'un certain niveau de bruit supplémentaire.

La location et l'échelle de l'image sont scannées pour identifier les points clés. Les emplacements et les échelles des points clés sont détectés comme l'extrême de la fonction $\mathrm{D}(\mathrm{x},\mathrm{y},\sigma)$ qui représente la différence gaussienne convolutionnée avec l'image $I(x,y)$ (voir la formule 1).

$$\mathrm{D}(\mathrm{x},\mathrm{y},\sigma) = \big(\mathrm{G}(\mathrm{x},\mathrm{y},\mathrm{k}\sigma) - \mathrm{G}(\mathrm{x},\mathrm{y},\sigma)\big) * \mathrm{I}(\mathrm{x},\mathrm{y}) \qquad (9)$$

Avec k un facteur constant multiplicatif et $\mathrm{G}(\mathrm{x},\mathrm{y},\sigma_\mathrm{i}) = \frac{1}{2\pi\sigma_\mathrm{i}^2}\mathrm{e}^{-(\mathrm{x}^2+\mathrm{y}^2)/2\sigma_\mathrm{i}^2}$ un noyau gaussien.

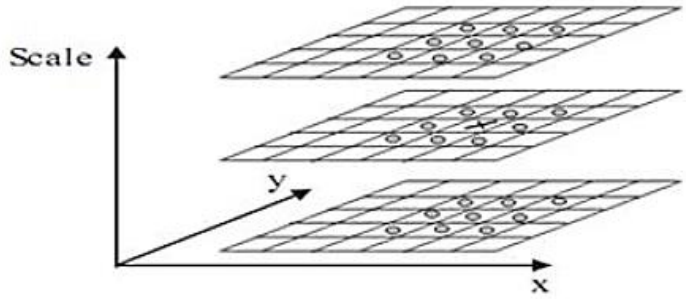


Figure 15. Détection des extrêmes par comparaison du pixel d'intérêt avec ses voisins du niveau courant ainsi que les niveaux adjacents [13]

Les extrêmes locaux sont détectés en comparant chaque pixel avec ses 26 voisins, 8 du niveau courant et 9 des deux niveaux inférieur et supérieur (voir figure 15). Le point est sélectionné s'il est le plus grand ou le plus petit de tous les points adjacents.

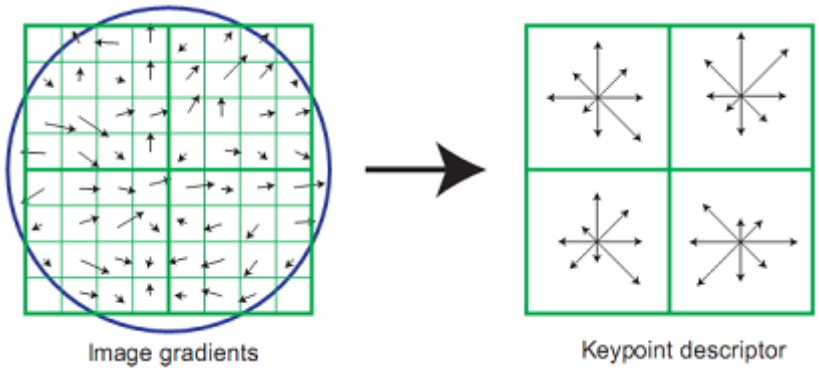


Figure 16. 2x2 vecteurs descripteurs calculés à partir d'un échantillon de 8x8 [13]

Le calcul des caractéristiques des points clés détectés par DoG est effectué à l'aide des techniques SIFT. Les points clés se voient attribuer une orientation, une échelle et un emplacement. L'échelle et l'emplacement sont extraits du détecteur DoG et attribuées à une ou plusieurs orientations en fonction de l'orientation principale du gradient de fenêtre entourant le point. Des histogrammes d'orientations sont calculés pour identifier les directions des gradients dominants.

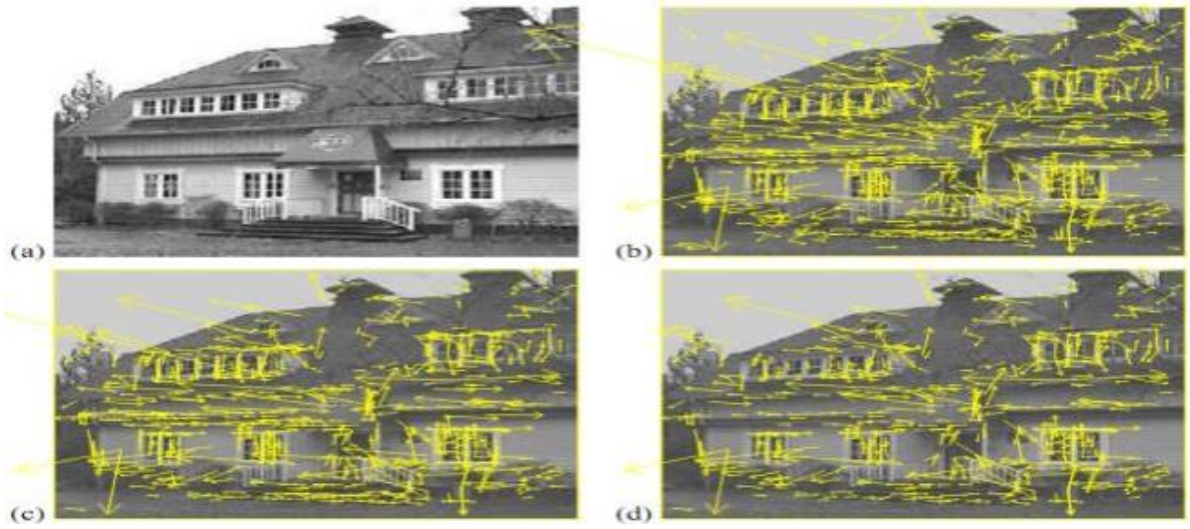


Figure 17. Une image et les points de caractéristiques locales extraites [13]

C. Les caractéristiques globales

Cette catégorie comprend essentiellement les caractéristiques de couleur, de texture et de forme.

C.1 La Couleur

Souvent, le premier descripteur utilisé pour la recherche d'image est la couleur. Plusieurs études ont déjà prouvé qu'il s'agit d'un descripteur efficace. Il existe différentes façons de caractériser la couleur,

comme l'histogramme ou le moment de couleur. Si on modifier l'espace de couleur avec la même manière, il peut révéler des informations différentes de l'image.

- **Les histogrammes**

Les histogrammes sont des indicateurs de répartition de niveaux de gris (ou de couleurs) dans une image. Ils sont très utilisés en recherche par le contenu car l'histogramme d'une image est presque invariant en rotation, translation et changement d'échelle de cette image.

À partir de l'histogramme, des attributs colorimétriques de l'image peuvent être extraits.

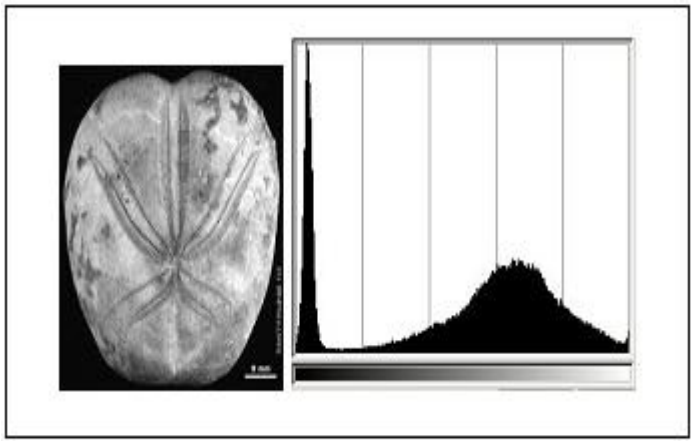


Figure 18. Exemple de l'histogramme d'une image en niveaux de gris [14]

Pour les images en couleurs, on utilise trois histogrammes, un par plan de bits de l'image (rouge, vert et bleu). De nombreux exemples d'attributs issus des histogrammes peuvent être calculés.

C.2 La Forme (en anglais : Shape)

La forme est un descripteur très important dans une base de données d'images. La forme fait référence à l'aspect général d'un objet, ses contours.

- **Les attributs géométriques de région**

Les attributs géométriques de forme permettent de distinguer les différents types de forme que peuvent prendre les objets d'une scène. Ils nécessitent une segmentation en région préalable de l'image. Ils sont ensuite calculés sur les différentes régions de l'image.

La surface relative (ou normalisée) d'une région R_k de l'image *I* est le nombre de pixels contenus dans cette région par rapport au nombre total de pixels de l'image :

$$S_k = \frac{card(R_k)}{hauteur(I) * largeur(I)} \qquad (10)$$

Le centre de masse des pixels de la région est défini par :

$$P = (P_i, P_j) = (\frac{\sum_{i \in R_k} i/card(R_k)}{largeur(I)}, \frac{\sum_{i \in R_k} i/card(R_k)}{hauteur(I)} \qquad (11)$$

La longueur du contour de la région est le nombre de pixels en bordure de la région :

$$S_k = card(contour(R_k)) \qquad (12)$$

La compacité traduit le regroupement des pixels de la région en zones homogènes et non trouées :

$$C_k = \frac{l_k^2}{S_K} \qquad (13)$$

Ces attributs très simples permettent d'obtenir des informations sur la géométrie des régions de l'image. Il existe d'autres attributs de forme, basés sur des statistiques sur les pixels des régions de l'image. **[14]**

- **Les moments géométriques**

Les moments géométriques permettent de décrire une forme à l'aide de propriétés statistiques. Ils sont simples à manipuler mais leur temps de calcul est très long.

Formule générale des moments :

$$m_{p,q} = \sum_{p=0}^{m} \sum_{q=0}^{n} x^p y^q f(x, y) \qquad (14)$$

L'ordre du moment est $p + q$. Le moment d'ordre 0 $m_{0,0}$ représente l'aire de la forme de l'objet.

Les deux moments d'ordre 1 $m_{0,1}$ et $m_{1,0}$, associés au moment d'ordre 0 $m_{0,0}$ permettent de calculer le centre de gravité de l'objet. Les coordonnées de ce centre sont :

$$x_c = \frac{m_{1,0}}{m_{0,0}}, \; y_c = \frac{m_{0,1}}{m_{0,0}} \qquad (15)$$

Il est possible de calculer à partir de ces moments l'ellipse équivalente à l'objet. Afin de calculer les axes de l'ellipse, il faut ramener les moments d'ordre 2 au centre de gravité :

$$\begin{cases} m_{2,0}^g = m_{2,0} - m_{0,0}\, x_C^2 \\ m_{1,1}^g = m_{1,1} - m_{0,0}\, x_c y_c \\ m_{0,2}^g = m_{0,2} - m_{0,0}\, y_C^2 \end{cases} \qquad (16)$$

Puis on détermine l'angle d'inclinaison de l'ellipse α.

$$\alpha = \frac{1}{2}\, arctan\, \frac{2m^g_{1,1}}{m^g_{2?0} - m^g_{0,2}} \qquad (17)$$

L'angle α est défini à π/2 près. La table 1 donne la valeur de l'angle en fonction du numérateur et du dénominateur de la dernière équation.

$m_{2,0}$-$m_{0,2}$	$m_{1,1}$	valeur	α
0	0		0
0	>0		$\frac{\pi}{4}$
0	<0		$-\frac{\pi}{4}$
>0	0		0
<0	0		$-\frac{\pi}{2}$
>0	>0	$\frac{1}{2} arctan \frac{2m^g_{1,1}}{m^g_{2,0} - m^g_{0,2}}$	$0< \alpha <\frac{\pi}{4}$
>0	<0	$\frac{1}{2} arctan \frac{2m^g_{1,1}}{m^g_{2,0} - m^g_{0,2}}$	$-\frac{\pi}{4}< \alpha <0$
<0	>0	$\frac{1}{2} arctan \frac{2m^g_{1,1}}{m^g_{2,0} - m^g_{0,2}} + \frac{\pi}{2}$	$\frac{\pi}{4}< \alpha <\frac{\pi}{2}$
<0	<0	$\frac{1}{2} arctan \frac{2m^g_{1,1}}{m^g_{2,0} - m^g_{0,2}} - \frac{\pi}{2}$	$-\frac{\pi}{2}< \alpha <-\frac{\pi}{4}$

Tableau 1. Table de calcul des angles de l'ellipse équivalente à une région

À partir des moments géométriques, Hu-Ming-Kuei a introduit sept invariants aux translations, rotations et changement d'échelle, appelés moments de Hu. **[14]**

$M_1 = \mu_{20} + \mu_{02}$

$M_2 = (\mu_{20} - \mu_{02})^2 + 4\mu^2_{11}$

$M_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$

$M_4 = (\mu 30 + \mu 12)2 + (\mu 21 + \mu 03)2$

$M_5 = (\mu_{30} - 3\mu_{12})\,(\mu_{30} + \mu_{12})\,[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2]$ (18)

$+ (3\mu_{21} - \mu_{03})\,(\mu_{21} + \mu_{03})\,[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$

$$M_6 = (\mu_{20} - \mu_{02})\,[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}\,(\mu_{30} + \mu_{12})(\mu_{03} + \mu_{21})$$

$$M_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})\,[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{21})\,(\mu_{12} + \mu_{03})\,[3(\mu_{30} + \mu_{12})^2 - (\mu_{12} + \mu_{03})^2]$$

C.3 La Texture

La texture peut être définie comme une forme visuelle avec des caractéristiques uniformités qui n'existent pas pour une seule intensité.

La texture est calculée pour une série de pixels. Plusieurs techniques ont été proposées pour extraire des caractéristiques de texture, qui sont généralement classées en deux catégories selon le domaine d'extraction. Ces deux grandes catégories sont l'extraction de caractéristiques de texture spatiale et l'extraction de caractéristiques de texture spectrale. La première approche extrait les caractéristiques en calculant les statistiques de pixels dans la région spatiale de l'image. La seconde approche transforme l'image en espace fréquentiel. C'est la source du calcul de la propriété.

Parmi les techniques d'extraction de caractéristiques ; les indices de Haralick, la transformée de Fourier discrète et les filtres de Gabor ont été largement utilisées.

- **La matrice de co-occurrence**

La texture d'une image peut être interprétée comme la régularité de l'apparition d'un couple de niveaux de gris en fonction d'une certaine distance dans l'image. La matrice de co-occurrence contient les fréquences spatiales relatives d'apparition des niveaux de gris selon quatre directions $\theta = 0, \theta = \frac{\pi}{2}, \theta = \frac{\pi}{4} et \theta = \frac{3\pi}{4}$. La matrice de co-occurrence est une matrice carrée n* n où n est le nombre de niveaux de gris de l'image.

On définit la matrice des fréquences relatives F par :

$$F(d,\theta) = (f(i,j\backslash d\ ,\theta)) \qquad (19)$$

Où f(i,j|d, θ) représente le nombre de fois où un couple de points séparés par la distance d dans la direction θ a présenté les niveaux de gris gi et gj . Pour obtenir la vraie fréquence relative, les éléments de la matrice doivent être normalisés en divisant par le nombre total de paires de points de base séparés par une distance *d* dans la direction *θ* dans toute l'image.

Les matrices de cooccurrences permettent de caractériser les textures présentes dans les images.

- **Les indices d'Haralick**

Dans son article "Textural features for image classification", Haralick **[15]** introduit quatorze attributs de texture extraits des matrices de cooccurrences. **(**Voir l'annexe A pour bien détailler ces attributs**)**.

- **La transformée de Fourier discrète**

De nombreuses méthodes d'extraction d'attributs de texture sont basées sur la transformée de Fourier. Cette dernière permet de passer du domaine spatial de l'image (coordonnées m et n) au domaine fréquentiel de l'image (coordonnées u et v). La transformée de Fourier discrète d'une séquence 2D correspondant au signal discret s[m,n], avec m et n entiers, $0 \leq m \leq M-1$, $0 \leq n \leq N-1$ est donnée par l'équation :

$$S(k,l) = \frac{1}{M}\frac{1}{N}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} s(m,n).e^{-2i\pi k\frac{m}{M}}.e^{-2i\pi l\frac{n}{N}} \qquad (20)$$

- **Les filtres de Gabor**

Le but du filtre de Gabor est de sélectionner un ensemble de caractéristiques fréquentielles pour chaque classe de texture dans le domaine de Fourier.

L'un des filtres les plus utilisés dans le domaine de la classification des images couleur texturées. En particulier, il est utilisé pour l'analyse de séquences d'images car il peut intégrer et relier des informations spatiales et temporelles. Cette propriété est utilisée pour détecter les objets en mouvement. Des études physiologiques montrent également que l'utilisation de ce type de filtre peut analogiser le travail de certains neurones du cortex visuel.

- **Les ondelettes**

Semblable au filtre de Gabor (en fait un cas particulier d'ondelettes), la transformée en ondelettes permet une représentation temps-fréquence. Il existe de nombreuses décompositions de ondelettes avec leurs propres caractéristiques uniques : transformées en ondelettes orthogonales, biorthogonales, transformées discrètes, continues, . . . etc.

Pour extraire les attributs de texture, on considère la transformée en ondelettes de l'image. Cette transformée est en fait une matrice de coefficients, de taille similaire à l'image initiale, dont on va extraire les attributs de texture. Ces transformations se font essentiellement pour les images en niveaux de gris.

D. Les systèmes de recherche par le contenu adaptés au domaine médical

D.1 Particularités de ces systèmes

Les particularités requises pour les systèmes CBIR adaptés au domaine médical proviennent, entre autres, des particularités des images médicales elles-mêmes. Il existe une grande variété de modalités d'acquisition d'images médicales, parmi lesquels l'imagerie par résonance magnétique (IRM), la tomographie permission de positons (TEP) et l'imagerie ultrasonore (US).

Les images fournies par ces différentes technologies sont très différentes en termes de résolution, contraste et rapport signal sur bruit. Elles sont très spécialisées et produisent des images porteuses d'informations différentes sur l'anatomie, la physiologie ou le métabolisme du patient. Les images médicales sont de plus des images d'intensité, qui portent moins d'information que les images couleur. Il arrive néanmoins que des images multimodales d'un même patient soient acquises (comme par exemple des images IRM et ultrasonores d'une même zone) mais ces images ne sont, la plupart du temps, pas recalées et nécessitent une procédure préalable de recalage qui s'avère complexe dans le cas de structures déformables. De plus, les images médicales peuvent être basse résolution et très bruitées. Elles sont de ce fait difficiles à analyser automatiquement pour en extraire des caractéristiques. En outre, comme mentionne dans **[16]**, la recherche d'images médicales doit la plupart du temps être effectuée selon des zones porteuses de pathologie précisément délimitées sur les images et difficilement détectables automatiquement dans le cas général.

Elle doit donc faire appel à une indexation locale des images, alors que la plupart des systèmes traditionnels se limitent à des caractéristiques globales. Ainsi, les systèmes CBIR adaptés au domaine médical nécessitent un haut niveau d'interprétation du contenu des images, ce qui reste aujourd'hui largement hors de portée des systèmes traditionnels.

Enfin, un niveau élevé de précision et de pertinence des requêtes effectuées sur ces systèmes est indispensable pour rendre ces systèmes dignes de confiance dans un contexte clinique. Par conséquent, comme le remarque Lehmann en, la plupart des systèmes CBIR dédies au domaine médical restent très spécifiques a une application et à un type d'images particuliers. **[16]**

D.2 Systèmes médicaux existants

- **Les techniques de raisonnement à partir de cas pour la recherche d'images médicales**

Plusieurs systèmes d'aide à la décision utilisent le principe RàPC (Raisonnement à Partir de Cas) dans le domaine de l'imagerie médicale.

- L'un des tous premiers systèmes de RàPC est CASEY qui est dédié au diagnostic cardiaque. On trouve aussi concernant PROTOS pour les problèmes d'audition.

- ISIS, par exemple permet une remémoration de cas d'interprétation de scanners, échographie ou IRM. Le système développé a été évalué et est en cours d'intégration dans un service de radiologie afin d'être utilisé en pratique quotidienne.
- Mac-Rad permet également la remémoration d'images morphologiques de référence (radiographie standard, scanners, RMN, et angiographie).
- ImageCreek associe un module de segmentation des images et un module de RàPC et met l'accent sur une interprétation à deux niveaux- segmentaire et globale – de l'image. Dans le domaine de la pathologie.
- IVY est dédié à l'aide au diagnostic des tumeurs pulmonaires. A partir de nombreux cas de départ, des cas « paradigmatiques » sont identifiés. Chacun d'entre eux est susceptible de représenter une classe. Le système propose un appariement, la phase d'adaptation n'est pas réalisée.
- La particularité de DIAGMED, développé dans le domaine de la pathologie rénale, est l'utilisation des logiques terminologiques pour la représentation des connaissances.
- CASIMIR non protocolaire est dédié au cancer du sein.
- MNAOMIA est dédié troubles du comportement alimentaire et enfin on retrouve des travaux concernant l'imagerie médicale IDEM (Image et Diagnostique par l'Exemple en Médecine).
- SRimCas «Système de Recherche d'Images Médicales par Cas » ce système permet, à partir de la description de cas à résoudre, de retrouver dans la base le(s) cas le(s) plus similaire(s) et également de visualiser, par une navigation qui utilise des liens de similarités, les cas «proches » du cas répondant à la requête. **[17]**

- **Les travaux sur les modèles à concepts dans le traitement d'image**

- **Yu Cao et al** ont présenté un système multi-madal de recherche des images médical en représentant l'image à l'aide des mots visuelle-textuel, Ces "mots" sont générés à partir des descripteurs visuels et l'information textuelle en utilisant le modèle PLSA. **[18]**
- **Spyridon Stathopoulos et al** ont présenté un système s'appuyant sur l'application de LSA sur plusieurs caractéristiques textuelles et visuelles de bas niveaux. **[19]**
- **Trong-Ton Pham et al** ont présenté un système pour étudier l'effet de l'analyse sémantique latente (LSA) sur deux tâches différentes : la recherche de documents multimédia (MDR) et l'annotation automatique d'images (AIA). La première s'agit de l'étude de l'influence du modèle LSA sur la recherche d'un nombre significatif de documents multimédias (par exemple, une collection de 20 000 images touristiques), au contraire la deuxième montre

comment différentes représentations d'images (basées sur les régions et sur les points clés) peuvent être combinées par LSA pour améliorer l'annotation automatique des images. **[20]**

- **Corina Văduva et al** ont présenté une approche basée sur le modèle probabiliste Allocation de Dirichlet latente pour l'analyse spatiale des images satellite. Les mots visuels spatiaux sont formés à partir de l'extraction de signatures spatiales invariantes qui décrivent les arrangements spatiaux dans la scène. **[21]**
- **Boulemden A et al** ont présenté un système de recherche d'images basé sur le contenu (CBIR) basé sur le modèle d'allocation pachinko (PAM) appliqués à deux modalités différentes de caractéristiques, les caractéristiques globales de l'image et les index textuels. **[22]**
- **Rasiwasia et Vasconcelos** ont présenté une technique pour la classification d'images basée sur le modèle d'allocation de Dirichlet latents. **[23]**

- **Systèmes divers**

Un nombre non négligeable de systèmes CBIR adaptés au domaine médical ont néanmoins été proposés.

- Chu et al présentent un système de recherche d'images dédie aux IRM du cerveau qui indexe les images essentiellement sur la forme de la région ventriculaire.
- Korn et al proposent un système de recherche de tumeurs dans les images mammographies.
- Comaniciu et al décrivent un système visant à aider les médecins au diagnostic des troubles lymphoproliferatifs du sang.
- Le système ASSERT, qui est dédié aux images HRCT du poumon et intègre des informations fournies par les médecins (comme des repères anatomiques et des régions porteuses de pathologie).
- Le système IRMA, qui propose une approche multi-étapes très complète pour la classification des images d'après leur modalité, leur angle de vue et la région anatomique à laquelle elles correspondent. **[24]**
- Le système de CervigramFinder a été développé pour étudier le cancer du col utérin, permet de calculer les caractéristiques locales d'une région définie par l'utilisateur.
- Spine pathologie et Image Retrieval System (SPIRS) est un système de recherche hybride basé sur le web, travaille en combinant les caractéristiques visuelles et l'information textuelle.
- Les systèmes SPIRS et IRMA ont été fusionnées pour former le système SPIRS-IRMA ; avec les fonctionnalités des deux. **[25]**

- SRIPCV « système d'indexation et de recherche d'Images Pulmonaires TDM par le contenu visuel » qui nous modélise le contenu visuel des images pulmonaires par un graphe (arbre) attribué. Ce dernier nous permet de bien représenter les différents aspects de similarité sur lesquels peut porter les requêtes des utilisateurs. **[26]**
- Système Indexation guidée par les connaissances en imagerie médicale, qui propose l'exploitation des deux sources de connaissances : une source théorique (la base de connaissances) et une source pratique (la base de cas). **[27]**
- Le système EMiner (modèles de mélange pour la recherche d'images par le contenu : applications aux pathologies ostéo-articulaire) est une nouvelle approche de recherche par boucle de pertinence est introduite. Elle est basée sur une classification semi-supervisée de la base des index utilisant les modèles de mélange et l'algorithme EM. **[28]**

Le tableau 2 présente un aperçu extrait des principaux systèmes CBIR médicaux existant et des types d'images auxquels ils sont dédiés. La variété des systèmes existant provient en grande partie de la grande diversité des types d'images médicales et de la spécificité des systèmes qui en découle. **[24]**

Images utilisées	Noms des systèmes
HRCT du poumon	ASSERT
TEP fonctionnelles	FICBDS
Rayons X de la moelle épinière	CBIR2
Images pathologiques	PathFinder, PathMaster
CT de la tête	MIMS
Mammographies	Tweed et al
Images issues de la biologie	BioImage
Dermatologiques	MELDOQ
Images variées	IRMA, KMed ,MedGIFT

Tableau 2. Différents types d'images et systèmes utilisant ces images

E. L'apprentissage automatique

L'apprentissage automatique (machine learning) fait référence au développement, à l'analyse et à l'implémentation de méthodes qui permettent à une machine (au sens large) d'évoluer grâce à un

processus d'apprentissage, et ainsi de remplir des tâches qu'il est difficile ou impossible de remplir par des moyens algorithmiques plus classiques.

Le premier stade de l'analyse est celui de la classification, qui vise à étiqueter chaque donnée en l'associant à une classe.

E.1 Types d'apprentissage

- **Apprentissage non supervisé**

En apprentissage non supervisé, vous apprenez à classer sans supervision. Au début du processus, il n'y a pas de définition ou de numéro des classes. C'est l'algorithme de classification qui détermine ces informations. De plus, les données d'entrée n'ont pas encore été catégorisées. C'est aussi à l'algorithme de découvrir des structures plus ou moins cachées dans les données elles-mêmes et de constituer des groupes d'individus aux caractéristiques communes.

Dans la littérature il existe plusieurs types d'algorithmes d'apprentissage non supervisé tels que les algorithmes de partitionnements et les algorithmes de classification hiérarchique :

- Le partitionnement : consiste à regrouper les données en fonction de leur degré de similarité. L'algorithme le plus connu de cette classe est K-means. Il s'agit d'un algorithme qui divise automatiquement le jeu de données en K clusters. Il consiste à sélectionner d'abord les k points qui représentent le centre du groupe formé, puis à associer les autres points au centre le plus proche. Cette affectation se fait en calculant la distance entre les points. Vous pouvez définir plusieurs distances telles que la distance euclidienne ou distance de Manhattan. Le raffinement est ensuite effectué en itérant jusqu'à l'étape de raffinement des groupes, en recalculant le centre du groupe après chaque itération et en réattribuant les points aux groupes. L'algorithme s'arrête quand tous les points ne bougent pas.
- La classification hiérarchique : Il existe deux types de classification hiérarchique: ascendante et descendante. La classification ascendante consiste à utiliser une matrice de similarité pour supposer une répartition plus fine vers un même groupe. Par conséquent, vous devez fusionner les groupes jusqu'à ce que vous ayez un seul groupe contenant tout le reste. Cette classification peut être représentée par un arbre hiérarchique ou un dendrogramme. La classification descendante est l'inverse de la classification ascendante. Par conséquent, le but est de décomposer un cluster unique en sous-groupes jusqu'à ce qu'un singleton soit obtenu.

- **Apprentissage supervisé**

Contrairement à l'apprentissage non supervisé, nous commençons avec un ensemble de classes pré-connues et définies. Il y a aussi la première sélection de données dont la classification est connue.

Ces données sont considérées comme indépendantes et distribuées de manière similaire. Ils nous aident à apprendre l'algorithme. La classification est effectuée par l'algorithme selon le modèle appris.

Il existe plusieurs algorithmes d'apprentissage supervisé, on peut présenter quelques-uns des plus connus parmi eux, il s'agit de KNN, LLSF, les réseaux de neurones et Naïve Bayes **[29]** :

- KNN (k nearest neighbor) : c'est une approche statistique de classification très connue. Il a été prouvé que c'est une des méthodes les plus performantes après des tests réalisés sur le corpus de données Reuters. Le principe de l'algorithme KNN est le suivant : étant donné un texte à classer, l'algorithme cherche les k voisins les plus proches parmi les documents utilisés au cours de la phase d'apprentissage, les catégories de ces k voisins les plus proches serviront à donner des poids aux catégories candidates de classification. C'est le degré de similarité entre le document test et le document voisin qui est utilisé comme poids de la catégorie de ce dernier, si plusieurs voisins partagent la même catégorie alors le poids attribué à cette catégorie est égal à la somme des degrés de similarité entre le document test et chacun des voisins appartenant à cette catégorie. Par cette méthode on peut obtenir une liste des poids attribués à chaque catégorie, le document test est classé dans une catégorie si le poids attribuée à celle-ci est supérieur à un seuil fixé à l'avance.

- LLSF (linear least square fit) : c'est une approche de mapping développée par Yang, il s'agit d'écrire les données d'apprentissage sous la forme de paires de vecteurs entrée/sortie, le vecteur d'entrée est composé des mots du texte accompagnés de leurs poids respectifs, alors que le vecteur de sortie est composé des différentes catégories avec leur poids binaires (1 si le texte appartient à une catégorie et 0 sinon), la résolution de l'équation suivante permet d'obtenir la matrice des coefficients de régression mot-catégorie.

$$F_{LS} = min_F \|FA - B\|^2 \qquad (21)$$

Où A et B sont deux matrices qui représentent les données d'apprentissage et dont les colonnes sont composées des paires des vecteurs entrée/sortie. La matrice solution FLS permet de donner pour tout texte un vecteur de catégories/poids. Comme pour KNN, un seuil est fixé et le document à classer appartient à toute catégorie ayant un poids supérieur au seuil fixé.

- Les réseaux de neurones : c'est une structure constituée de suite successive de couches de nœuds et qui permet de définir une fonction de transformation non linéaire des vecteurs d'entrées (composés dans le cas de classification des mots pondérés de leur poids) en vecteur de catégories. La disposition des neurones dans le réseau ainsi que le nombre de couches utilisées ont une influence sur le résultat de classification. Comparés aux autres méthodes de

classification par apprentissage supervisé, les réseaux de neurones ont l'inconvénient que le coût d'apprentissage est assez élevé.

- NB (Naive Bayes) : c'est une méthode de classification probabiliste. Elle consiste à utiliser les probabilités jointes des mots et des catégories pour estimer la probabilité d'une catégorie sachant un texte à classer. Le caractère « naïf » de cette approche est dû au fait que les mots sont considérés indépendants, c'est-à-dire que la probabilité conditionnelle d'un mot sachant une catégorie est supposée indépendante des probabilités conditionnelles des autres mots sachant la même catégorie, cette assomption rend NB très efficace par rapport aux autres approches bayésiennes. Plusieurs versions de NB sont proposées dans la littérature, le model mixte multinominal par exemple a permis d'avoir de bonnes performances.
- SVM « supports vectors machines » appelés aussi « maximum margin classifier» sont des techniques d'apprentissage supervisé basées sur la théorie de l'apprentissage statistique ou automatique. Les SVM sont relativement nouveaux, ils sont apparus en 1995 suite aux travaux de Vapnik. SVM traite d'un problème de classification bi classes.

F. Critères d'évaluation

Le domaine de l'indexation et recherche d'image par le contenu utilise les métriques d'évaluation de la recherche d'information. Les deux mesures les plus utilisées sont le rappel et la précision. Le rappel et la précision sont calculés selon les formules suivantes :

$$Precision = \frac{A}{B} \qquad (22)$$

$$Recall = \frac{A}{C} \qquad (23)$$

Avec :

A : le nombre d'images pertinentes retournées par le système.

B : le nombre total des images résultats.

C : le nombre d'images pertinentes dans la base.

Les systèmes sont souvent évalués en calculant le MAP (Mean Average Precision), cette mesure est calculée sur plusieurs requêtes et représente la moyenne arithmétique des différentes moyennes de précision de chaque requête. Le MAP permet d'évaluer la précision des réponses du système en tenant compte la position de la réponse pertinente dans l'ensemble des réponses.

Pour la phase de classification, il existe d'autres mesures telles que : F1-score, l'accuracy et macro-average.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (24)$$

L'accuracy permet de décrire la performance du modèle sur les individus positifs et négatifs de façon symétrique. Elle mesure le taux de prédictions correctes sur l'ensemble des individus :

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (25)$$

*O*ù **TP**: true positive, **TN**: true negative, **FP** : false positive, **FN** : false negative.

Le macro-moyen (macro-average) est calculé en utilisant la moyenne arithmétique (ou moyenne non pondérée) de tous les F1-scores par classe.

5. Conclusion

Dans ce chapitre, nous avons présenté dans la section 1 l'architecture générale et les domaines d'application d'un système d'indexation par image (CBIR), puis nous avons parlé dans la section 2 sur les différents types de modèles de sujets dans le contexte de modélisation de document. Nous avons donné en détail le modèle LSA qui est le modèle le plus utilisé dans ce contexte.

A la fin de ce chapitre, nous avons exposé l'indexation en imagerie médicale et les différentes caractéristiques d'image ainsi qu'on a présenté quelques systèmes d'indexation et de recherche d'image dans le domaine médical.

Les modèles de sujets ont été adaptés aux documents de types images et ils ont donné des résultats intéressants, surtout dans l'indexation et la recherche des images comme il était mentionné dans les travaux de Yu Cao et al **[18]**, Spyridon Stathopoulos et al **[19]**, Trong-Ton Pham et al **[20]**, Boulemden A et al **[22]**, Rasiwasia et Vasconcelos **[23]**.

Le chapitre suivant sera dédié à l'étude détaillé de la procédure de la conception de ce système d'indexation.

Chapitre 2 : Conception

1. Introduction

Après avoir abordé tous les notions et les concepts nécessaires, qui permettent de définir et de comprendre le fonctionnement d'un système « CBIR », ainsi que le schéma général de ce dernier, on peut maintenant présenter notre travail.

Ce projet concerne l'application du modèle de sujet LSA (Latent Semantic Analysis) pour l'indexation et la recherche des images de tumeurs mammaires. Le travail contient également une deuxième partie consistant à trouver le type de tumeur de la requête en utilisant l'algorithme de classification KNN (K-nearest Neighbor) sur les résultats de la recherche.

Afin de pouvoir réaliser le système proprement dit, il est nécessaire de :

- Proposer un modèle d'indexation par le contenu.
- Décrire les étapes à suivre afin de réaliser celui-ci.
- Appliquer de l'algorithme KNN sur les résultats de recherche pour classer les tumeurs.

2. Représentation de notre système

A. Architecture de notre système

Le modèle LSA constitue le cœur du module de l'indexation, il est appliqué sur des matrices de cooccurrence des mots visuels images. Ces matrices sont construites en utilisant les sacs de mots visuels.

Les caractéristiques locales sont calculées pour chaque image dans la base (voir partie 1 de la figure 19), à partir de ces caractéristiques extraites, le processus de construction des mots visuels est appliqué (BOVW pour bag of visual words, voir partie 2 de la figure 19). Le modèle LSA est par la suite appliquer sur la matrice de mot-visuels-images produite (voir partie 3 de la figure 19). L'application du modèle LSA sur la matrice de cooccurrences des mots visuels-images produit trois matrices. Il s'agit d'une matrice de document qui représente l'espace d'index de l'image d'apprentissage, une matrice de mots visuels et une matrice de valeurs singulières (voir la section 2 du chapitre 1).

Lors de la phase de la recherche, les caractéristiques locales sont calculées pour l'image de la requête (voir la partie 4 de la figure 19). La construction des mots visuels est par la suite appliquée sur ces caractéristiques extraites (voir la partie 5 de la figure 19). Cette construction utilisent les informations issues de la même phase exécutée pour les images de la base, les mots visuels de la requête vont être

construits en faisant référence aux mots visuels construits pour les images de la base (voir la partie 6 de la figure 19). Le modèle LSA permet par la suite de calculer un vecteur descripteur de la requête, ce calcul est effectué en utilisant les matrices W et E issues de l'application de LSA sur les images de la base (voir la partie 7 de la figure 19). La dernière étape consiste à mesurer la similarité entre le vecteur descripteur de la requête et ceux des images de la base. Ces derniers sont trouvés dans la matrice D. La liste des résultats est finalement affichée (voir la partie 8 de la figure 19).

Pour la phase de classification basée sur KNN, on a appliqué des prétraitements sur la base d'images puis donner à notre modèle pour l'apprentissage (voir les parties 9 et 11 de la figure 19). Les résultats de LSA seront triés et traités afin d'être prêt pour la phase de reconnaissance de type de tumeur (voir la partie 12 de la figure 19). La figure 19 montre le fonctionnement de notre système dans ses deux modes : mode apprentissage et mode recherche.

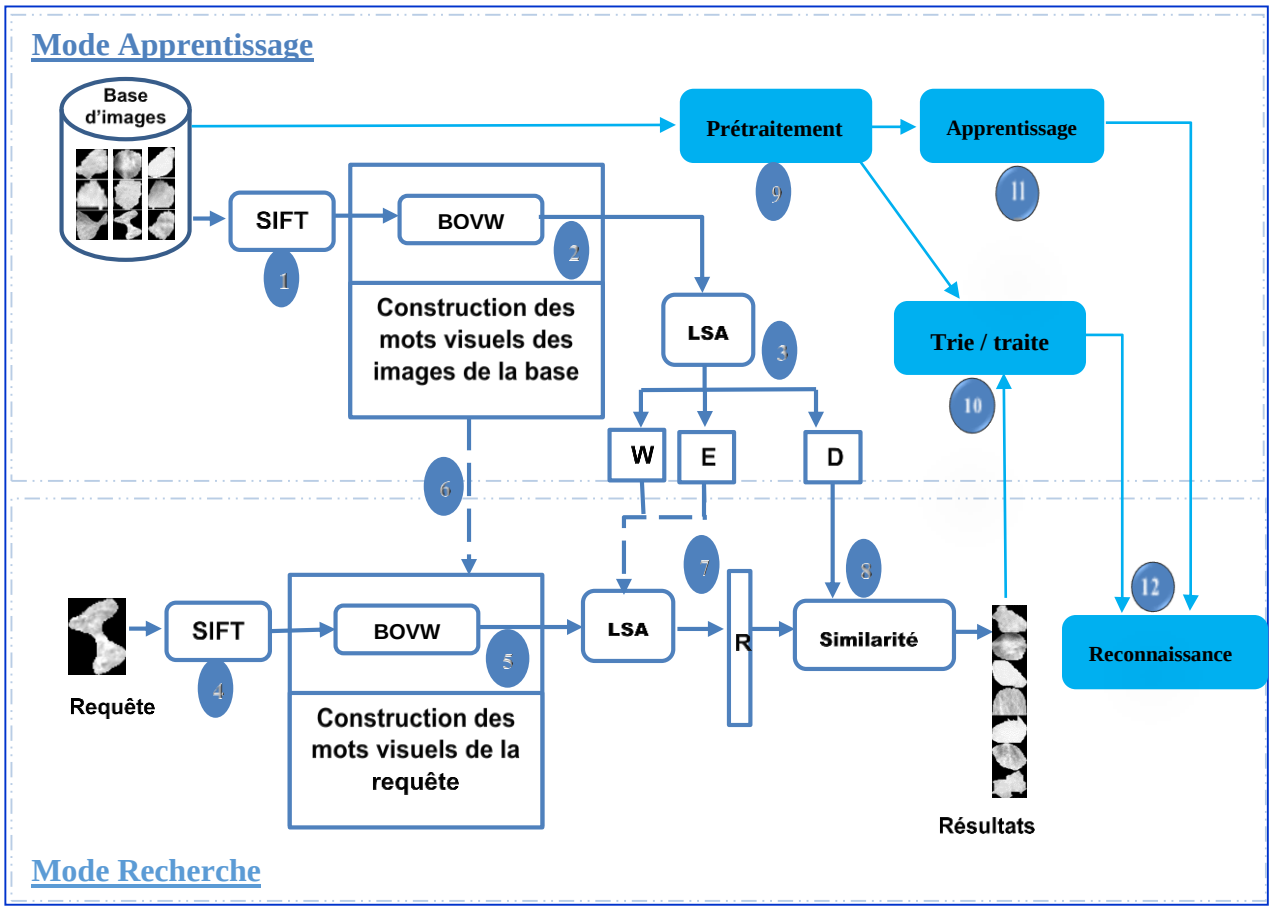


Figure 19. Architecture détaillée de notre système

B. Description des différentes étapes de notre Système d'indexation

Notre système d'indexation regroupe trois étapes principales : l'extraction des caractéristiques locales SIFT pour chaque image, la construction des mots visuels à partir de ces caractéristiques et à la fin l'application du LSA pour formant le vecteur descripteur, autrement dit l'index de l'image.

Avant de décrire le détail des étapes de notre système il est important de mentionner que les images de notre base sont obtenues manuellement après une détermination ou extraction de la zone de maladie.

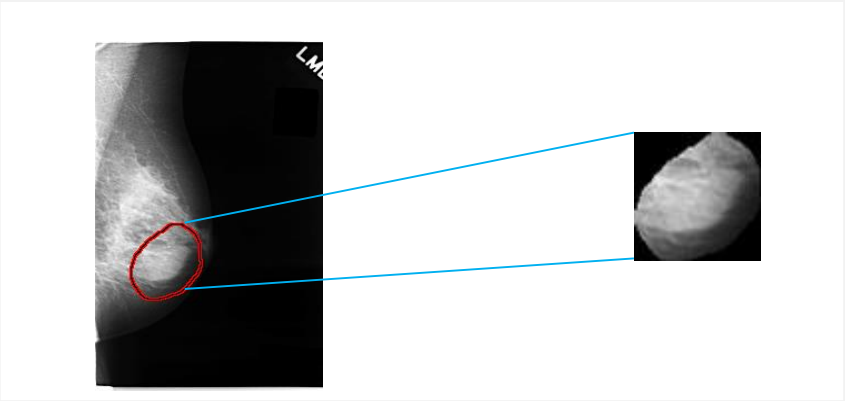


Figure 20. Préparation des images de la base

1. Extraction des caractéristiques

1.1 Critères de distinction entre tumeurs bénignes et tumeurs malignes

Tumeurs bénignes	Tumeurs malignes
Bien limitée	Mal limitée
Encapsulée	Non encapsulée
Histologiquement semblable au tissu d'origine	Plus ou moins semblable au tissu d'origine (dédifférenciation, différenciation aberrante)
Cellules régulières	Cellules irrégulières (cellules cancéreuses)
Croissance lente	Croissance rapide
Refoulement sans destruction des tissus voisins	Envahissement des tissus voisins
Pas de récidive locale après exérèse complète	Récidive possible après exérèse supposée totale
Pas de métastase	Métastase(s)

Tableau 3. Critères de distinction entre tumeurs bénignes et tumeurs malignes

Après la distinction entre les tumeurs bénignes et malignes, on peut maintenant projeter toutes ces informations sur le plan du traitement d'images numériques, pour que nous puissions les identifier automatiquement par le système. Donc chaque information peut se traduire par une ou plusieurs caractéristiques numériques.

1.2 Extraction des Caractéristiques SIFT

SIFT est un algorithme qui permet de détecter des points d'intérêt et d'extraire des caractéristiques distinctives de ces points pour la reconnaissance d'objet. SIFT est un vecteur de caractéristiques locales qui décrit un pixel et qui est robuste :

- À la translation ;
- Au changement d'échelle ;
- À la rotation ;
- Aux changements d'éclairage ;
- Aux projections affines ou 3D.

Dans l'algorithme SIFT proposé par Lowe **[13]** les caractéristiques locales sont décrites à l'aide d'un détecteur de type DoG (Difference of Gaussians). Ensuite chaque caractéristique "point clé" pour simplifier la lecture est décrite à l'aide d'un ensemble d'histogrammes des orientations comportant 8 intervalles. Pour ce faire, on définit une région de taille 16*16 autour du point clé. Cette région est ensuite divisée en 4 sous-régions de taille 4*4 dans lesquelles on calcule l'orientation et l'amplitude du gradient. À partir de ces informations on décrit le point par une concaténation de tous les histogrammes des 8 orientations du gradient dans chaque sous-région. L'histogramme de chaque sous-région de taille 4*4 est obtenu en faisant la somme des amplitudes du gradient en chaque point pondérée par une gaussienne centré sur le point clé et l'écart type est de 1.5 fois le facteur d'échelle du point clé. L'orientation du gradient détermine l'intervalle à incrémenter dans l'histogramme. Toutes ces différentes étapes de l'algorithme de calcul du descripteur SIFT sont illustrées par la figure 21.

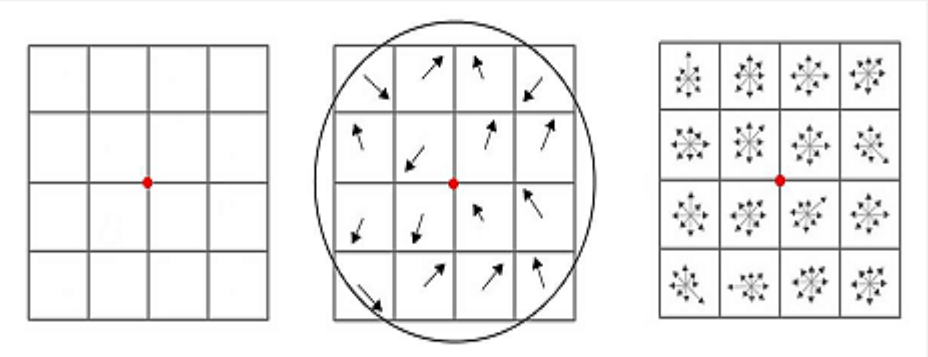


Figure 21. Différentes étapes de la description d'un point clé avec l'algorithme SIFT. Le point à décrire est représenté en rouge. Le cercle sur la figure du milieu illustre la gaussienne utilisée pour pondérer les amplitudes du gradient avant de construire l'histogramme final [1]

Le descripteur final est de taille 128 = 4*4*8 (grande dimension → obtenir une bonne capacité de discrimination).

2. Construction des Mots visuels

Pour la construction des mots visuels basés sur les caractéristiques locales SIFT, nous avons choisie 16 mots par image pour la première classification et 20 mots pour le vocabulaire général, réalise comme suit :

- Appliquer un clustering avec kmeans pour tous les points clés de chaque image. Les centroids sont considérés comme des mots visuels de l'image et la dimension du cluster comme sa fréquence.
- Définir le vocabulaire de la base, cette étapes consiste à un deuxième clustering appliqué sur les centroids de l'étape précédente. Les centroids du même cluster vont représenter le même mot dans toute la base (exemple word_5).

❑ Description de l'algorithme K moyenne

- Choisir *k* points qui représentent la position moyenne des partitions $m_1^{(1)},\ldots,m_k^{(1)}$ initiales (au hasard par exemple). **[W3]**
- Répéter jusqu'à convergence :
  - assigner chaque observation à la partition la plus proche.
    $S_i^{(t)} = \{X_j : \| X_j - m_i^{(t)} \| \leq \| X_j - m_{i^*}^{(t)} \| \forall\, i^* = 1, \ldots, k\}$
  - mettre à jour la moyenne de chaque cluster
    $m_i^{(t+1)} = \frac{1}{\left| S_i^{(t)} \right|} \sum_{X_j \in S_i^{(t)}} X_j$
- La convergence est atteinte quand il n'y a plus de changement.

Pour construire la matrice de co-occurrence « mot -visuel-image » qui correspond à notre modèle de sujets LSA, il est important de calculer la fréquence des mots visuels dans chaque image.

❑ Calcul de la fréquence

La figure 22 ci-dessous illustre le principe de construction de mots visuels.

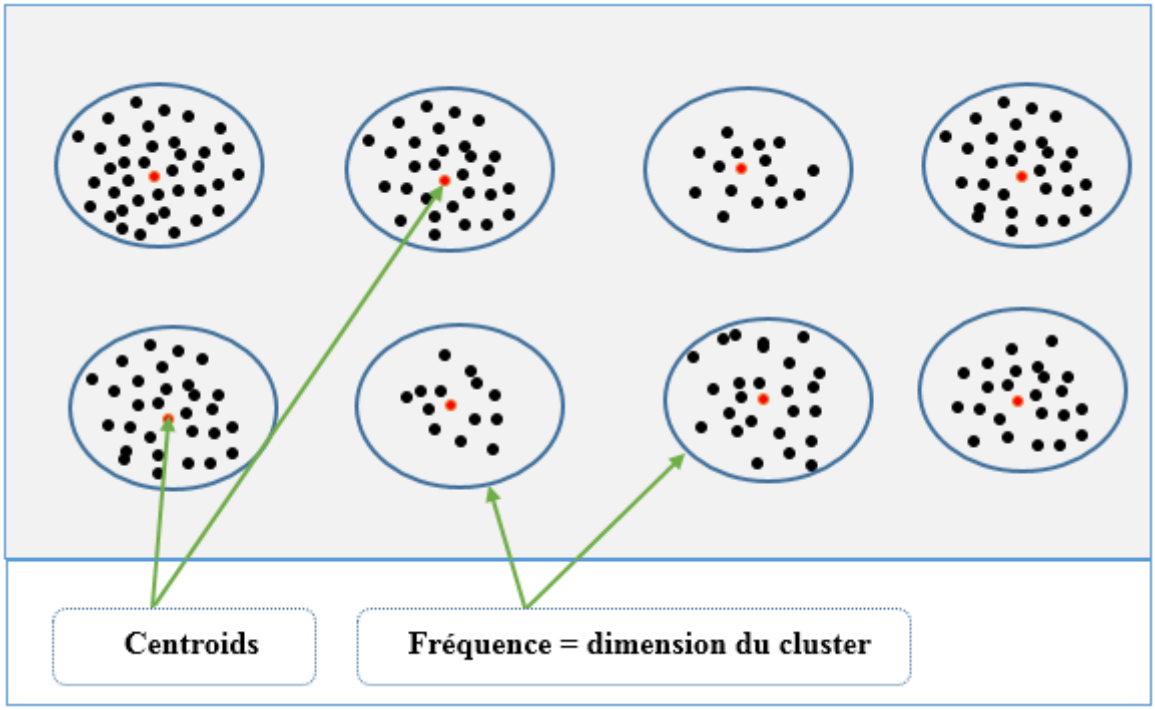


Figure 22. Illustration du principe de construction de mots visuels

3. Application du LSA

LSA permet d'extraire et d'inférer des relations entre les mots visuels. Dans cette étape, LSA construit la matrice de cooccurrences « mot-visuel-image », constituée des fréquences Fi de chaque mot visuel dans chaque image, sans tenir compte de leur ordre.

Cette matrice est ensuite réduite à l'aide d'une décomposition en valeurs singulières (SVD) afin de capturer dans une certaine mesure les relations entre les mots visuels et les images.

$$A = W\sum D^t \begin{cases} A \in \mathbb{R}^{n_W x n_D} \\ W \in \mathbb{R}^{n_W x r} \\ \sum \in \mathbb{R}^{r x r} \\ D \in \mathbb{R}^{n_D x r} \end{cases} \quad (26)$$

Où A est notre matrice de cooccurrences « mot-visuel-image », et W, D sont des matrices orthonormales contenant les vecteurs singuliers gauche et droit de A, et ∑ est une matrice diagonale contenant les valeurs singulières de A :

$$\begin{cases} \Sigma = diag(\sigma_1, \sigma_2, \ldots\ldots, \sigma_r) \\ \mathrm{WW^t = DD^t = I_r} \qquad avec\ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \\ \mathrm{rang(\Sigma) = r \leq \min(n_W, n_D)} \end{cases} \quad (27)$$

Pour diminuer le nombre de dimensions de l'espace, on peut trouver une bonne approximation $\tilde{A}$ de A en mettant à zéro les petites valeurs singulières de ∑ afin d'obtenir un espace réduit de dimension $\tilde{r}$ choisie :

$$A \cong \tilde{A} = \widetilde{W}\sum\widetilde{D}^{t} \begin{cases} \tilde{A} \in \mathbb{R}^{n_W x n_D} \\ \widetilde{W} \in \mathbb{R}^{n_W x \tilde{r}} \\ \widetilde{\sum} \in \mathbb{R}^{\tilde{r} x \tilde{r}} \\ \widetilde{D} \in \mathbb{R}^{n_D x \tilde{r}} \end{cases} \quad (28)$$

Nous obtiendrons à la fin une matrice réduite $\widetilde{D}^{t}$ où les colonnes représentent les vecteurs descripteurs des images de la base d'apprentissage.

Concernant l'image requête ; le vecteur d_q sera transformé en un pseudo vecteur $\tilde{d}_q$ dans l'espace réduit.

Remarque :

Pour calculer la similarité on a utilisé la distance euclidienne. La formule 29 montre le calcul de la distance euclidienne entre deux vecteurs i et j.

$$Dis(i,j) = \sqrt{\sum_{k=1}^{n}(X_{ik} - X_{jk})^2} \quad (29)$$

C. Reconnaissance de type de tumeur par l'application de KNN

La phase de l'indexation et la recherche des images est suivie par une phase de reconnaissance de type de tumeur mammaires de l'image requête. Cette reconnaissance est basée sur l'application de l'algorithme de classification de KNN (K nearest neighbors) (voir la section 3 du chapitre 1) sur l'ensemble des images résultats de la requête. La simplicité de l'algorithme KNN est avantageuse pour la rapidité. Le processus de classification comporte les étapes suivantes ;

- **Préparation de dataset**

Notre base d'images contient 140 images médicales mammographiques dont 70 pour la tumeur maligne et 70 pour le bénigne, pour atteindre à la classification il faut faire des traitements sur cette base comme suite :

- Toutes les images sont converties en trois dimensions 64 x 64 x 3 (où 3 désigne espace couleurs RVB).
- Puis convertir chaque image en vecteur de 12288 valeurs, ce dernier sera enregistré dans un matrice de dimension 140 ×12288.
- On a converti les informations de l'image en vecteur de 140 dimensions. les nombres de 1 à 70 représentants les tumeurs malignes prennent la valeur 1 et les tumeurs bénignes (71 à 140) prennent la valeur 0.

- Nous avons appliqué cette transformation sur la base pour adapter avec notre modèle de classification KNN parce que les algorithmes d'apprentissage automatique ne traitent que des valeurs numérique.

❑ **Phase d'apprentissage**

Après le prétraitement de la base, les données sont passées au modèle proposé pour apprentissage. Le modèle proposé basé sur KNN (K nearest neighbors).

❑ **Reconnaissance de type de tumeur**

Après l'apprentissage et l'enregistrement du modèle, nous avons fait une classification à l'aide des résultats de modèle LSA qui déjà triés et traités.

3. Conclusion

Dans ce chapitre, Nous avons détaillé les étapes de réalisation de notre système, en justifiant à chaque étape le choix de la méthode utilisé.

Dans le chapitre suivant, on va implémenter la conception de notre système d'indexation et de recherche d'image dans le domaine médical ainsi présenter les différents outils utilisés.

Chapitre 3 : Implémentation

1. Introduction

Le chapitre courant est consacré à l'implémentation de notre système, il concerne en général le développement et la réalisation du système, c'est à dire présenter le langage de programmation, la plateforme ou environnement de développement et les différents modules qui composent le système.

2. Le langage de programmation et bibliothèques

Notre application a été réalisée sur un PC portable de type Dell Inspiron 3520 i3, CPU 2.20 GHZ et 4.00 GB de RAM sous Windows 7 Pro. Les deux premières étapes de notre projet ont été réalisées sous Matlab R2013, la troisième étape qui concerne l'application de modèle LSA sous Apache NetBeans (c-à-d on a utilisé java comme langage de programmation).

Pour la phase de reconnaissance de type de tumeur par l'application de KNN on a utilisé le langage de programmation Python.

A. Python

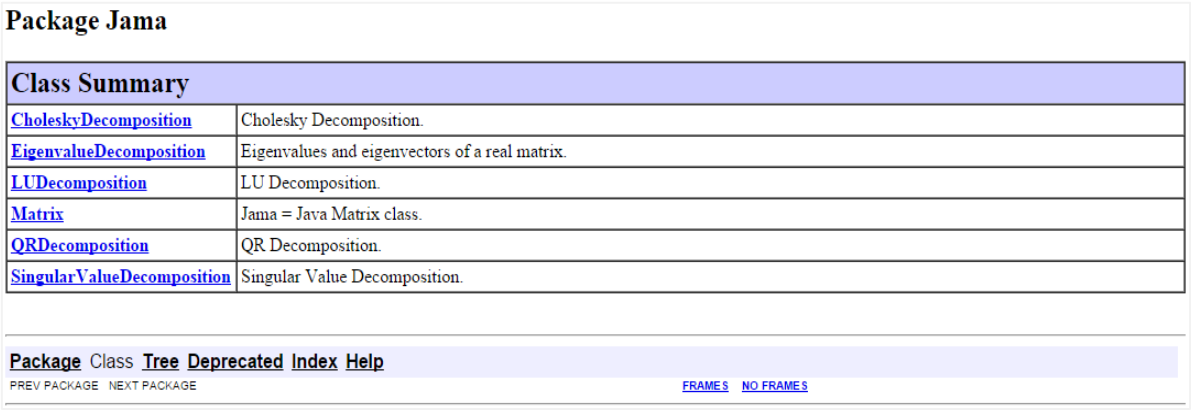
Python est le langage de programmation open source le plus employé par les informaticiens. Ce langage s'est propulsé en tête de la gestion d'infrastructure, d'analyse de données ou dans le domaine du développement de logiciels. En effet, parmi ses qualités, Python permet notamment aux développeurs de se concentrer sur ce qu'ils font plutôt que sur la manière dont ils le font. Il a libéré les développeurs des contraintes de formes qui occupaient leur temps avec les langages plus anciens. Ainsi, développer du code avec Python est plus rapide qu'avec d'autres langages. **[W4]**

B. Bibliothèques utilisées

- **Le packadge Jama**

JAMA est un paquet d'algèbre linéaire de base pour Java. Il fournit des classes de niveau utilisateur pour la construction et la manipulation réelle, matrices denses. Il est destiné à fournir une fonctionnalité suffisante pour des problèmes de routine, emballés d'une manière qui est naturel et compréhensible pour les non-experts. Il est destiné à servir de la classe de la matrice standard pour Java, et sera proposé en tant que telle à la Grande Forum Java, puis à Sun. Une implémentation de référence du domaine public simple a été développée par les MathWorks et NIST comme un homme de paille pour une telle classe. JAMA est composé de six classes Java : Matrix,

CholeskyDecomposition, LUDecomposition, QRDecomposition, SingularValueDecomposition et EigenvalueDecomposition. **[W5]**



Package Jama

Class Summary	
CholeskyDecomposition	Cholesky Decomposition.
EigenvalueDecomposition	Eigenvalues and eigenvectors of a real matrix.
LUDecomposition	LU Decomposition.
Matrix	Jama = Java Matrix class.
QRDecomposition	QR Decomposition.
SingularValueDecomposition	Singular Value Decomposition.

Package Class Tree Deprecated Index Help

PREV PACKAGE NEXT PACKAGE FRAMES NO FRAMES

Figure 23. Illustration des composants de packadge Jama [W5]

- **NumPy**

Nous avons utilisé cette bibliothèque pour adapter les types d'entrée selon la configuration du modèles utilisés, destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux. **[W6]**

- **Sklearn**

C'est l'une des bibliothèques d'apprentissage automatique les plus utiles en Python. La bibliothèque sklearn contient de nombreux outils efficaces pour l'apprentissage automatique et la modélisation statistique, notamment la classification, la régression, le clustering et la réduction de la dimensionnalité. **[W7]**

3. L'environnement de développement

A. NetBeans

Notre système a été implémenté dans NETBeans **[W8]** IDE version 13.

NetBeans est un environnement de développement intégré (EDI), placé en open source par Sun en juin 2000 sous licence CDDL (Common Development and Distribution License) et GPLv2. En plus de Java, NetBeans permet la prise en charge native de divers langages tels le C, le C++, le JavaScript, le XML, le Groovy, le PHP et le HTML, ou d'autres (dont Python et Ruby) par l'ajout de greffons. Il offre toutes les facilités d'un IDE moderne (éditeur avec coloration syntaxique, projets multi-langage, refactoring, éditeur graphique d'interfaces et de pages Web).

Compilé en Java, NetBeans est disponible sous Windows, Linux, Solaris (sur x86 et SPARC), Mac OS X ou sous une version indépendante des systèmes d'exploitation (requérant une machine virtuelle Java). Un environnement Java Development Kit JDK est requis pour les développements en Java.

NetBeans constitue par ailleurs une plateforme qui permet le développement d'applications spécifiques (bibliothèque Swing (Java)). L'IDE NetBeans s'appuie sur cette plateforme.

Vous pouvez télécharger l'IDE NetBeans à partir du site officiel : https://netbeans.apache.org/. Ci-dessous présente la page principale de NetBeans.

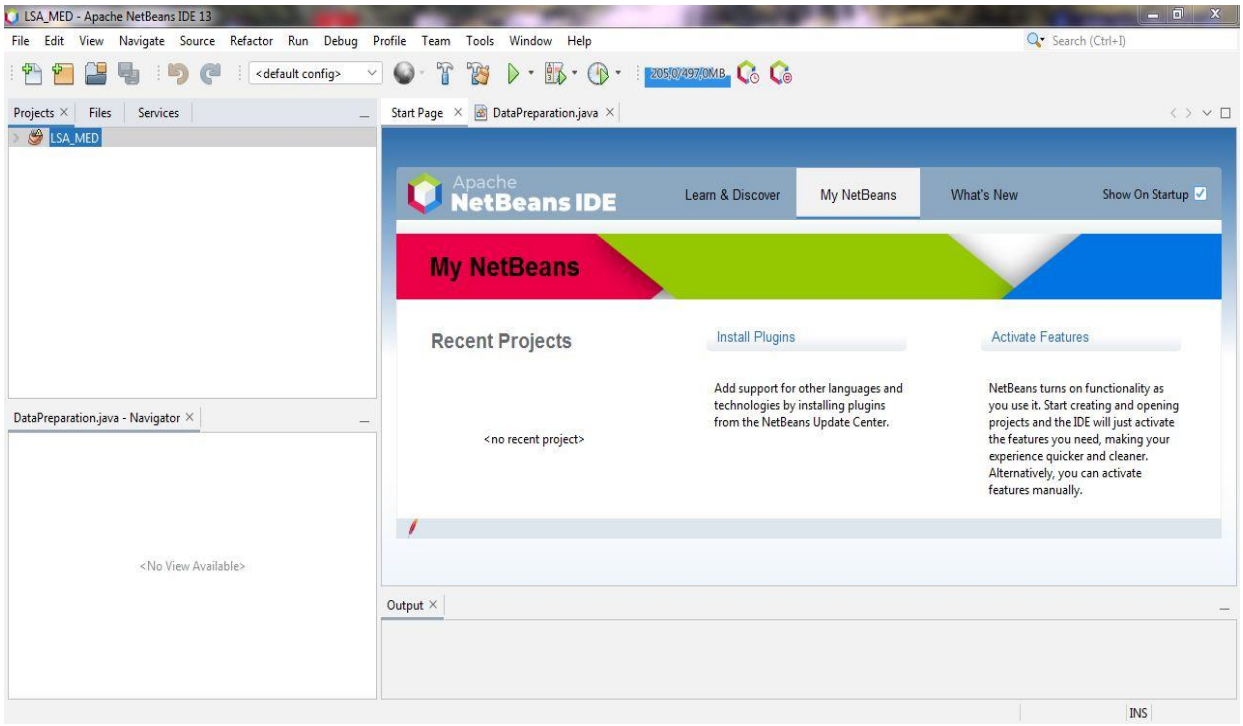


Figure 24. La page principale de NetBeans IDE 13

B. Matlab

MATLAB « matrix laboratory » est un langage de script émulé par un environnement de développement du même nom ; il est utilisé à des fins de calcul numérique. Développé par la société The MathWorks, MATLAB permet de manipuler des matrices, d'afficher des courbes et des données, de mettre en œuvre des algorithmes, de créer des interfaces utilisateurs, et peut s'interfacer avec d'autres langages comme le C, C++, Java, et Fortran. Les utilisateurs de MATLAB (environ 4 millions en 20193) sont de milieux très différents tels que l'ingénierie, les sciences et l'économie, dans un contexte aussi bien industriel que pour la recherche. Matlab peut s'utiliser seul ou bien avec des toolboxes (« boîte à outils ») **[W9]**. La figure 25 montre l'interface de Matlab R2013.

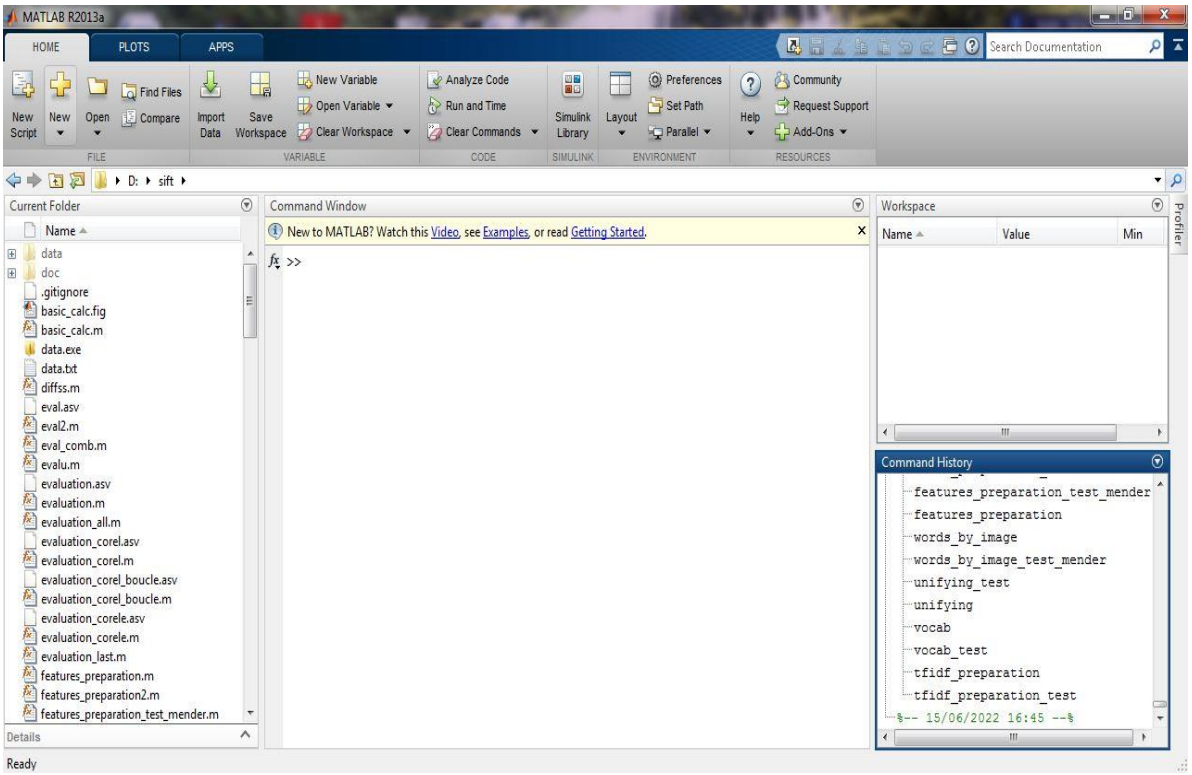


Figure 25. Interface de Matlab R2013

C. Google Colab

Google Colab ou également appelé Colaboratory, permet d'écrire et d'exécuter le code Python de son choix par le biais du navigateur. Offert par Google (gratuit), basé sur Jupyter Notebook et destiné à la formation et à la recherche dans l'apprentissage automatique.

Les principales fonctionnalités de Google Colab sont:

- ✓ Vous n'avez pas besoin de configuration supplémentaire. Tout Google Colab est en ligne, vous n'avez donc pas besoin de télécharger d'application ou de configurer quoi que ce soit sur votre ordinateur.
- ✓ Vous pouvez profiter des bibliothèques Python pour développer, analyser ou visualiser vos projets.
- ✓ Vous avez accès gratuitement au puissant matériel Google.
- ✓ Vous avez la possibilité d'enregistrer et de partager dans et depuis le cloud tous les notebooks que vous développez ou exécutez dans Colab. **[W10]**

4. Détails de l'application

A. La base d'images utilisée

Les images médicales mammographies utilisées dans nos expérimentations sont issues de la base **MIAS** (Mammographic Image Analysis Society), c'est une organisation de groupe de recherche du Royaume-uni qui s'intéresse à la compréhension et la génération des bases des images de mammographie numérique. Cette dernière se compose de deux collections, la première consiste en une base d'apprentissage et composé de 140 images de tumeurs mammaires dont 70 pour les tumeurs de type maligne et 70 pour les tumeurs de type bénigne. La deuxième collection représente la base de teste, elle est composée de 20 images requêtes dont 10 pour des tumeurs de type maligne et 10 pour des tumeurs de type bénigne. Chaque image dans cette base d'images a une taille de 64 × 64 pixels et de type PNG. Les images de ces deux collections contiennent seulement la partie de tumeurs qui a été extraite manuellement. La figure 26 montre un exemple d'une image de la base d'apprentissage utilisée dans les expérimentations. **[W11]**

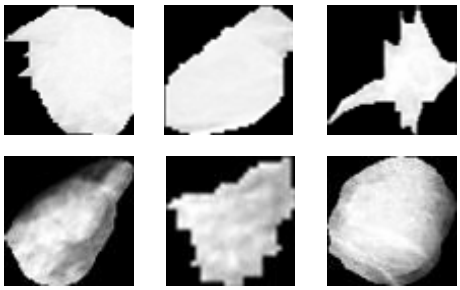


Figure 26. Exemple des images de la base d'apprentissage

B. Extraction des caractéristiques

L'extraction des caractéristiques visuelles locales est réalisée en utilisant le Matlab et se fait par le moyen de la technique SIFT (voir la section 3 du chapitre 1) c'est pour but de l'extraction des points clés et le calcule de leurs caractéristiques (voir **l'annexe B** pour la représentation d'une partie du code source matlab d'implémentation du technique SIFT).

```
0.007791463751345872900000000000000000000000000000000000000000000000
0.160157486796379090000000000000000000000000000000000000000000000000
0.040163964033126831000000000000000000000000000000000000000000000000
0.070191480219364166000000000000000000000000000000000000000000000000
0.000531515455804765220000000000000000000000000000000000000000000000
0.003890267573297023800000000000000000000000000000000000000000000000
0.045321989804506302000000000000000000000000000000000000000000000000
0.052782036364078522000000000000000000000000000000000000000000000000
0.096281841397285461000000000000000000000000000000000000000000000000
0.020111246034502983000000000000000000000000000000000000000000000000
0.006455824244767427400000000000000000000000000000000000000000000000
0.054186359047889709000000000000000000000000000000000000000000000000
0.226369306445121770000000000000000000000000000000000000000000000000
0.022643491625785828000000000000000000000000000000000000000000000000
```

Figure 27. Un fragment de vecteur descripteur SIFT d'une image de la base

Figure 28. Un fragment de vecteur SIFT (positions de frames) d'une image de la base

C. Construction des mots visuels

Pour la construction des mots visuels basés sur les caractéristiques locales SIFT, nous avons choisie 16 mots par image pour la première classification et 20 mots pour le vocabulaire général, réalise comme suit :

- Appliquer un clustering avec kmeans pour tous les points clés de chaque image. Les centroids sont considérés comme les mots visuels de l'image et la dimension du cluster comme sa fréquence.
- Définir le vocabulaire de la base, cette étapes consiste à un deuxième clustering appliqué sur les centroids de l'étape précédente. Les centroids du même cluster vont représenter le même mot dans toute la base.

Le résultat de ces étapes transforme les images en texte (textes visuels) et sauvegarder dans des fichiers texte dans le dossier D:/mender/tf_idf_prep et dans D:/mender/class/test/tf_idf_prep_test pour les images d'apprentissage et de test. La figure 29 montre la description textuelle construite pour l'image.

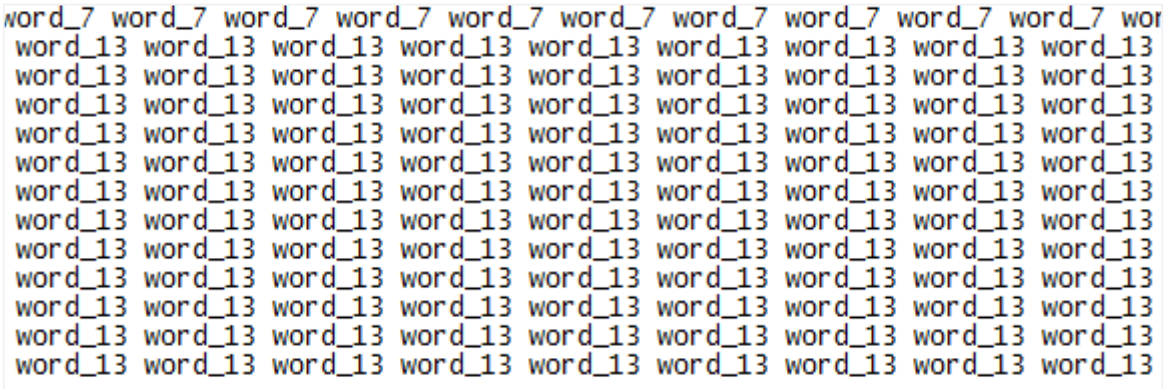
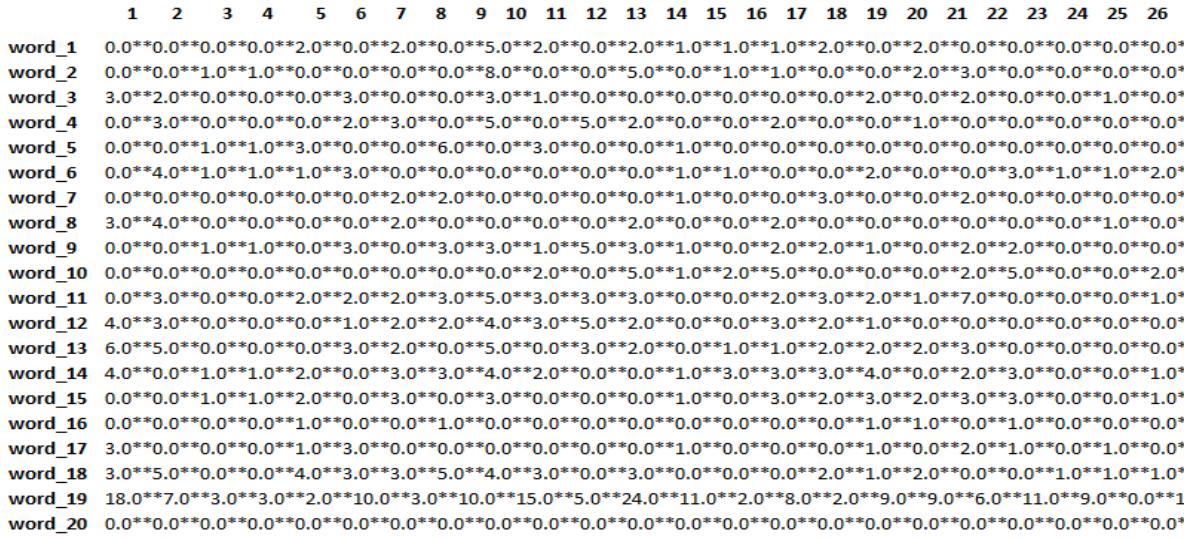


Figure 29. Un fragment de texte généré à la base des caractéristiques visuelles locales. Les mots sont des mots visuels

- Pour construire la matrice de co-occurrence « mot-visuel-image » qui correspond à notre modèle de sujets LSA, il est important de calculer la fréquence des mots visuels dans chaque image (voir chapitre 2 « figure 22 »). La figure 30 présente un échantillon de la matrice co-occurrence « mot-visuel-image » de la base d'apprentissage qui contient dans les colonnes 140 images et 20 mots visuels dans les lignes.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
word_1	0.0	0.0	0.0	0.0	2.0	0.0	2.0	0.0	5.0	2.0	0.0	2.0	1.0	1.0	1.0	2.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
word_2	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	8.0	0.0	0.0	5.0	0.0	1.0	1.0	0.0	0.0	2.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
word_3	3.0	2.0	0.0	0.0	0.0	3.0	0.0	0.0	3.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	2.0	0.0	0.0	1.0	0.0			
word_4	0.0	3.0	0.0	0.0	0.0	2.0	3.0	0.0	5.0	0.0	5.0	2.0	0.0	0.0	2.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0			
word_5	0.0	0.0	1.0	1.0	3.0	0.0	0.0	6.0	0.0	3.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			
word_6	0.0	4.0	1.0	1.0	1.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	2.0	0.0	0.0	3.0	1.0	1.0	2.0			
word_7	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	3.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0			
word_8	3.0	4.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0			
word_9	0.0	0.0	1.0	1.0	0.0	3.0	0.0	3.0	3.0	1.0	5.0	3.0	1.0	0.0	2.0	2.0	1.0	0.0	2.0	2.0	0.0	0.0	0.0			
word_10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	5.0	1.0	2.0	5.0	0.0	0.0	0.0	2.0	5.0	0.0	0.0	2.0			
word_11	0.0	3.0	0.0	0.0	2.0	2.0	2.0	3.0	5.0	3.0	3.0	3.0	0.0	0.0	2.0	3.0	2.0	1.0	7.0	0.0	0.0	0.0	1.0			
word_12	4.0	3.0	0.0	0.0	0.0	1.0	2.0	2.0	4.0	3.0	5.0	2.0	0.0	0.0	0.0	3.0	2.0	1.0	0.0	0.0	0.0	0.0	0.0			
word_13	6.0	5.0	0.0	0.0	0.0	3.0	2.0	0.0	5.0	0.0	0.0	3.0	2.0	0.0	0.0	1.0	1.0	2.0	2.0	2.0	3.0	0.0	0.0	0.0	0.0	0.0
word_14	4.0	0.0	1.0	1.0	2.0	0.0	0.0	3.0	3.0	4.0	2.0	0.0	0.0	0.0	1.0	3.0	3.0	3.0	4.0	0.0	0.0	2.0	3.0	0.0	0.0	1.0
word_15	0.0	0.0	1.0	1.0	2.0	0.0	0.0	3.0	0.0	3.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	3.0	2.0	3.0	0.0	2.0	3.0	3.0	0.0	0.0
word_16	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	1.0	0.0	0.0
word_17	3.0	0.0	0.0	0.0	0.0	1.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	2.0	1.0	0.0	0.0	1.0
word_18	3.0	5.0	0.0	0.0	0.0	4.0	3.0	3.0	5.0	4.0	3.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	2.0	1.0	0.0	2.0	0.0	0.0	0.0	1.0
word_19	18.0	7.0	3.0	3.0	0.0	2.0	10.0	3.0	10.0	0.0	15.0	5.0	0.0	24.0	0.0	11.0	0.0	2.0	8.0	0.0	2.0	9.0	0.0	9.0	0.0	6.0
word_20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Figure 30. Un échantillon de la matrice co-occurrence « mot-visuel-image » de la base d'apprentissage

D. Application du modèle LSA

Le paquet principal de notre système est le paquet « LSA_MED » qui contient la classe principale du système « LSA_MED ». La classe DataPreparation est la classe importante dans le système, elle contient toutes les méthodes nécessaires pour l'application du LSA (voir **l'annexe C** pour la représentation d'une partie du code source java d'implémentation du modèle LSA).

LSA est appliqué sous NetBeans IDE 13 et le packadge Jama a été utilisé.

Après l'exécution du projet LSA_MED nous obtiendrons les interfaces suivantes :

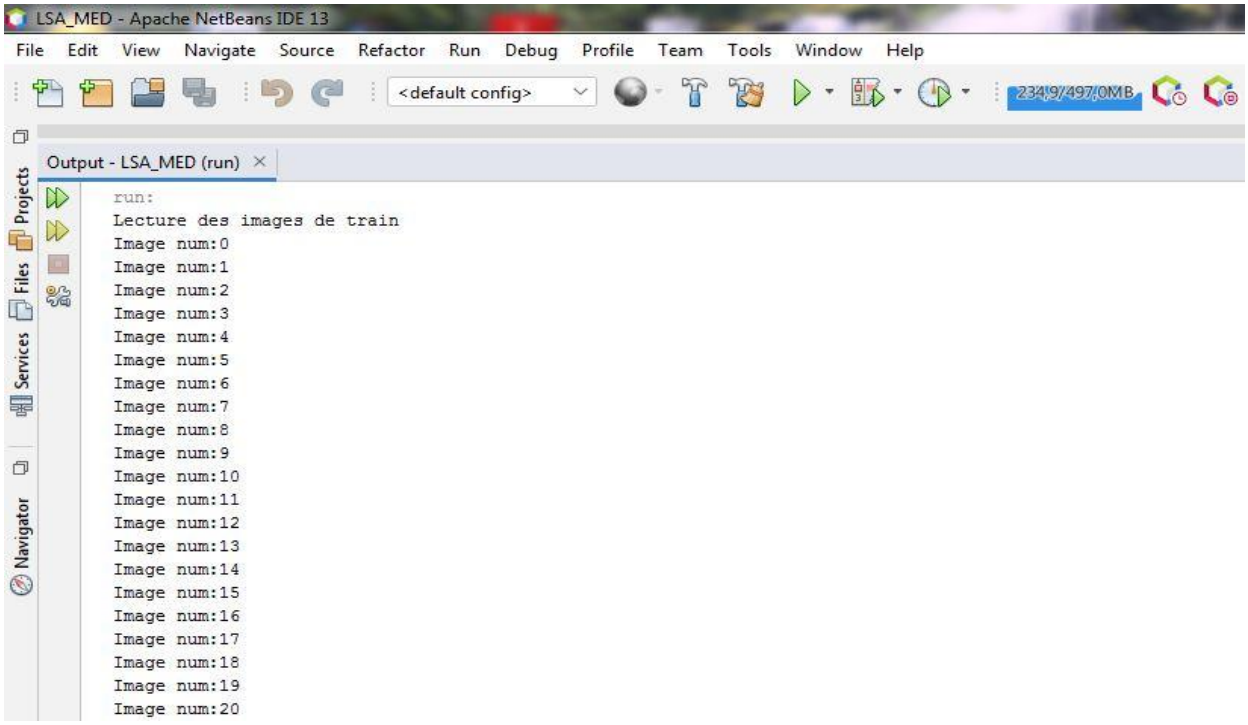


Figure 31. Lecture des images d'apprentissage

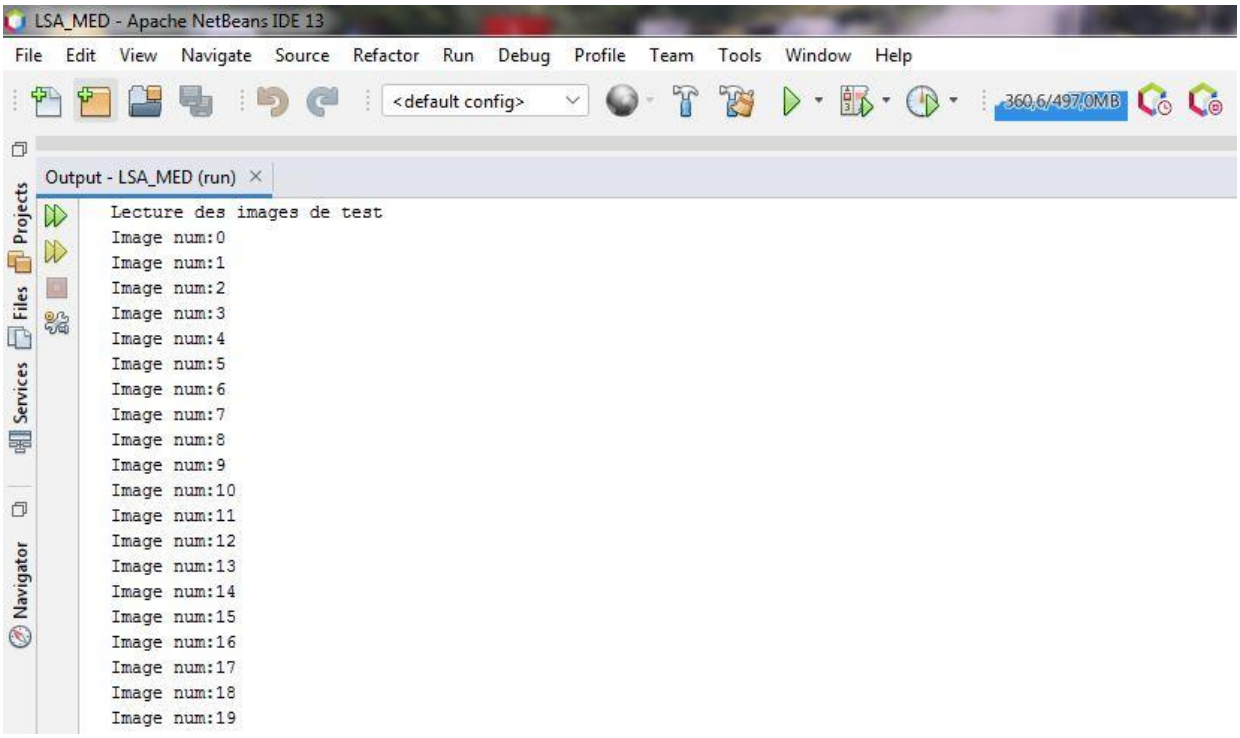


Figure 32. Lecture des images de test

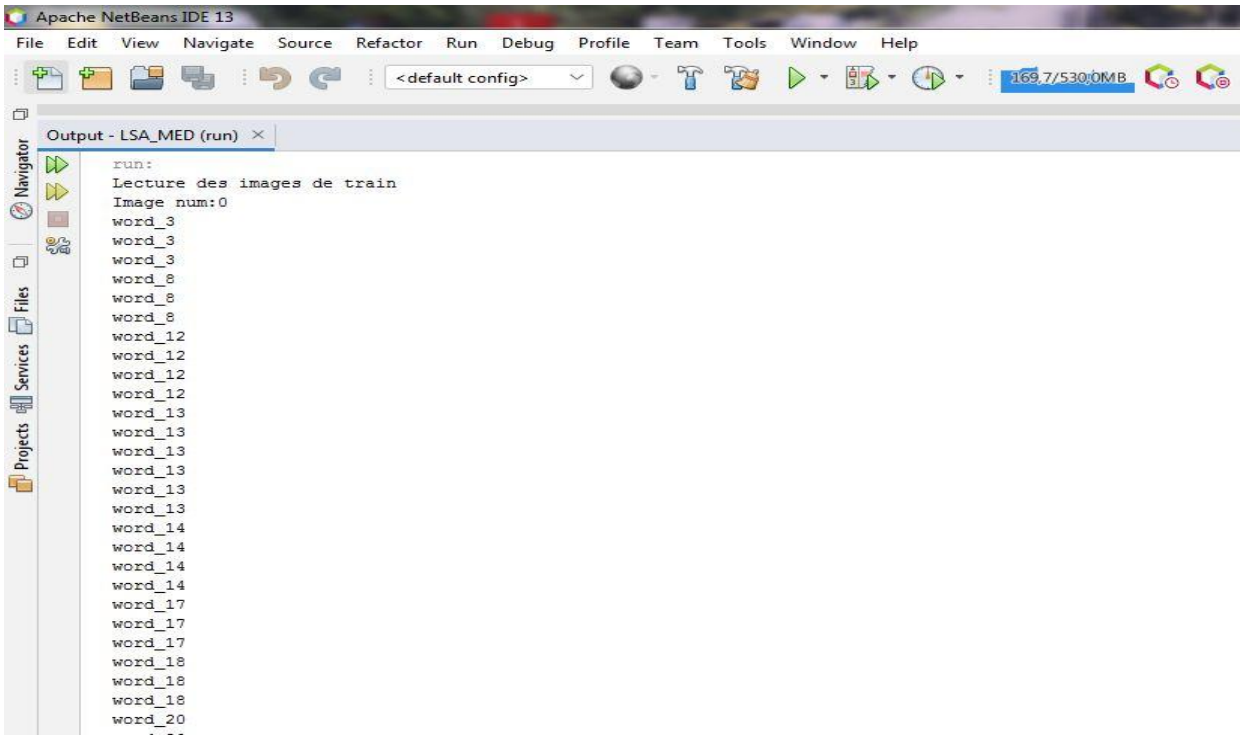


Figure 33. Un échantillon des mots visuels dans une image

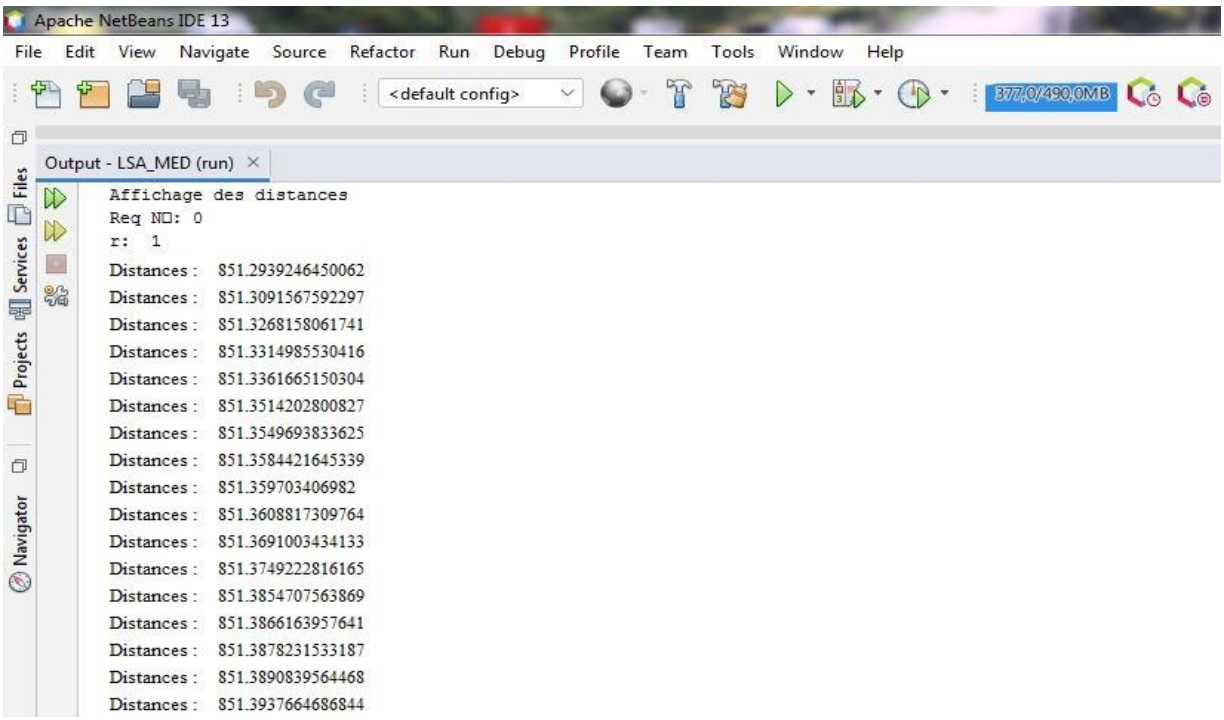


Figure 34. Calcul de la distance euclidienne entre l'image de test et les images de la base

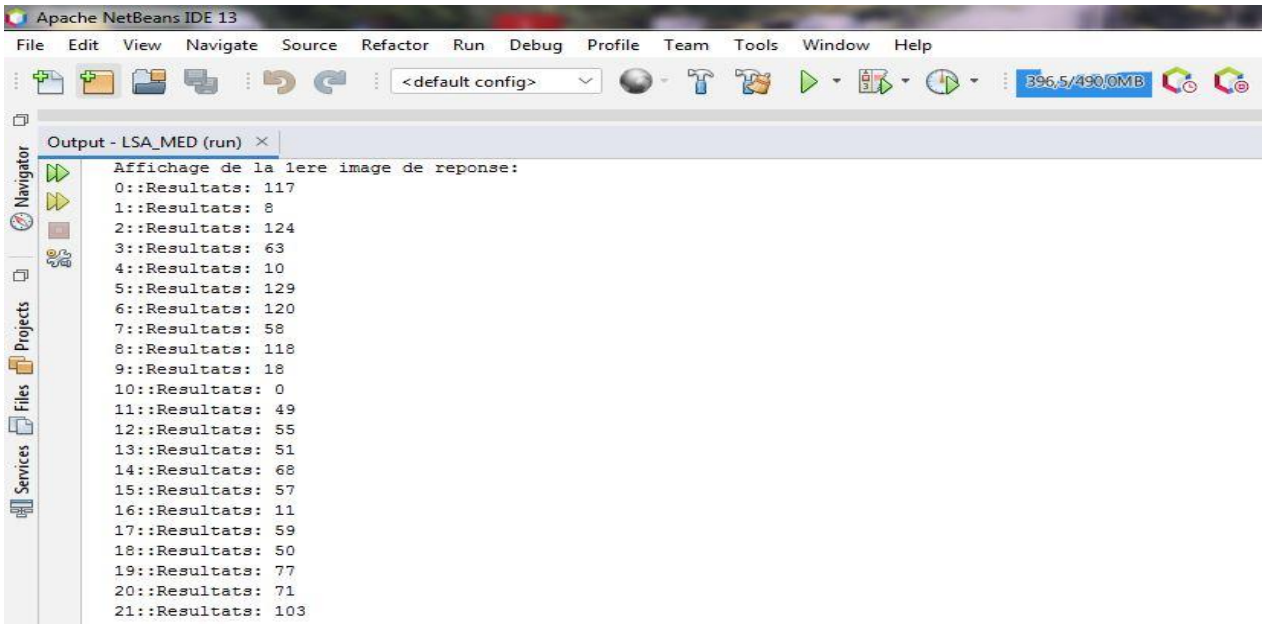


Figure 35. Montre le numéro et la position des images similaires pour la première image requête

5. Discutions et évaluations

Nous avons utilisé le MAP (Mean Average Précision) pour l'évaluation de l'indexation et la recherche d'images. Le MAP des requêtes de la base de teste est calculé pour le cas des caractéristiques locales de SIFT. Le tableau 4 montre les résultats de l'évaluation du MAP. Le nombre de requêtes utilisées est 20 requêtes dont les 10 premières sont des requêtes de tumeurs malignes et les 10 dernières sont des requêtes de tumeurs bénignes. La figure 36 montre le graphe des résultats des deux catégories des requêtes malignes et bénignes.

Le MAP permet d'évaluer la précision des réponses du système en tenant compte de la position de la réponse pertinente dans l'ensemble des réponses.

Requête malignes	Local SIFT	Requête bénignes	Local SIFT
Average Precisions de 1	0,5279	Average Precisions de 11	0,5100
Average Precisions de 2	0,5225	Average Precisions de 12	0,4867
Average Precisions de 3	0,5283	Average Precisions de 13	0,5270
Average Precisions de 4	0,5319	Average Precisions de 14	0,4670
Average Precisions de 5	0,5286	Average Precisions de 15	0,6124
Average Precisions de 6	0,5369	Average Precisions de 16	0,4704
Average Precisions de 7	0,5473	Average Precisions de 17	0,5164
Average Precisions de 8	0,5301	Average Precisions de 18	0,4810
Average Precisions de 9	0,5187	Average Precisions de 19	0,4967
Average Precisions de 10	0,4985	Average Precisions de 20	0,4679
Moyenne AP malignes	**0,5271**	**Moyenne AP bénignes**	**0,5036**
MAP malignes et bénignes	**0,5153**		

Tableau 4. Résultats de l'évaluation du MAP de LSA avec les caractéristiques locales SIFT pour l'indexation des tumeurs mammaires

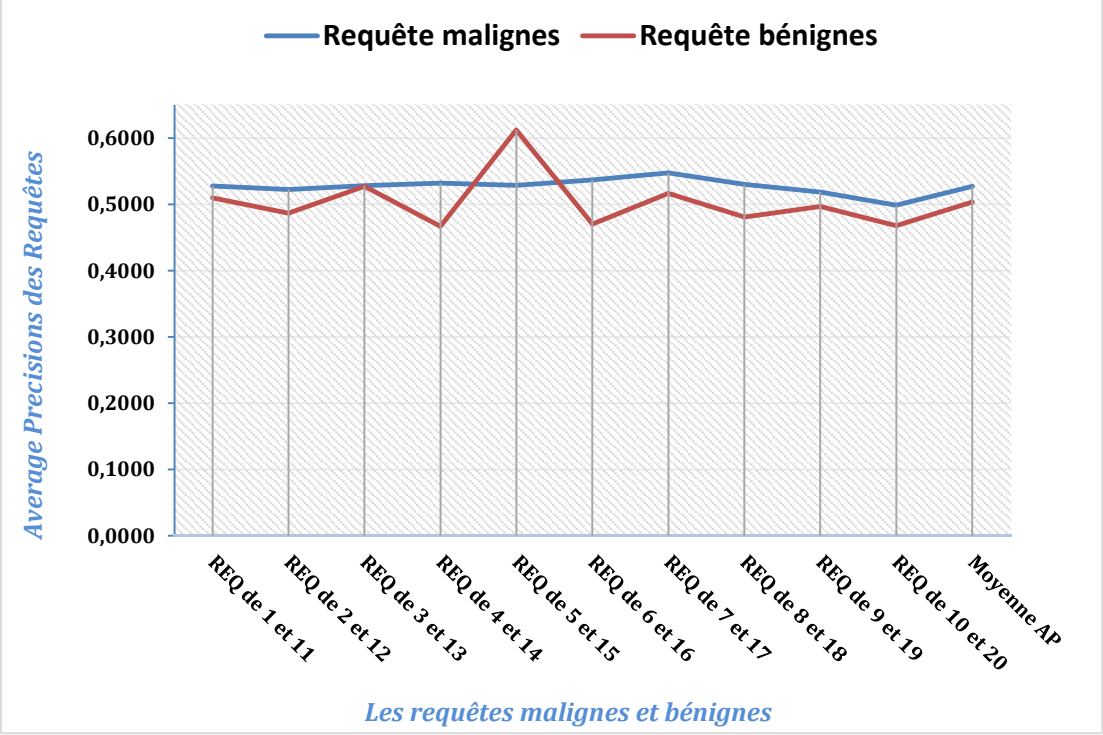


Figure 36. Graphes des résultats de MAP par catégories de requêtes

Le Map des requêtes malignes et bénignes est compris entre 47% à 61%. Le MAP globale de toutes les requêtes présentes 51,5%, c-à-d, qu'il représente des bons résultats. Ces résultats semblent encourageants pour les travaux futurs.

Le MAP des requêtes malignes pour LSA avec les caractéristiques locales SIFT est mieux de 2,35% par rapport aux requêtes bénignes.

A. Comparaison de nos résultats avec du travail similaire

Nous avons choisi de comparer nos résultats obtenus de notre travail avec d'autres résultats du travail dans **[1]**. Ces deux travaux ont utilisés le même modèle de sujet « LSA » ainsi que la même base des images « **MIAS** », le seul déférent est le type des caractéristiques à extraire. Le tableau 5 résume la comparaison.

	Notre travail	**[1]**
Modèle de sujet appliqué	LSA	LSA
Type de caractéristiques extraites	Caractéristiques locale SIFT	Caractéristiques globale Haralick
Base d'images utilisée	MIAS (160)	MIAS (160)
MAP des requêtes malignes	0,5271	0,4836
MAP des requêtes bénignes	0,5036	0,5230
MAP globale des malignes et bénignes	0,5153	0,5033

Tableau 5. Comparaison de nos résultats et résultats du travail [1]

Les résultats de l'application de LSA pour l'indexation et la recherche des tumeurs mammaires montrent que son utilisation est meilleure pour les tumeurs de type bénignes.

Le MAP globale de toutes les requêtes pour l'application de LSA avec les caractéristiques locales SIFT est mieux de 1,2% par rapport au résultat du MAP pour LSA avec les caractéristiques globales Haralick.

Le MAP des requêtes Malignes pour l'utilisation de LSA avec les caractéristiques locales SIFT est mieux de 4,35% par rapport aux requêtes malignes du LSA avec les caractéristiques globales

Haralick. Par contre, dans le cas de l'utilisation des caractéristiques de Haralick avec LSA, le MAP des requêtes bénignes est meilleur de 1,94%.

Les résultats restent proches et peuvent être considérés acceptable. L'amélioration des résultats constitue une perspective qui doit être abordé aussi avec le reste des expérimentations non réalisée.

B. Evaluations des résultats de la classification par KNN

Nous avons calculés la précision, le rappel, F1-score et l'accuracy de la classification des tumeurs pour l'évaluation de la reconnaissance par KNN. Nous avons utilisé plusieurs valeurs de k et on trouve des bons résultats lorsque k est égal à 5.

Les résultats de l'évaluation de la reconnaissance de type de tumeur sont illustrés par le tableau 6 et aussi par le graphe 37.

Requêtes	Precision	Recall	F1-score	Accuracy
Requête 1	0.64	0.40	0.49	0.59
Requête 2	0.67	0.46	0.54	0.61
Requête 3	0.67	0.46	0.54	0.61
Requête 4	0.65	0.44	0.53	0.61
Requête 5	0.68	0.43	0.53	0.61
Requête 6	0.68	0.43	0.53	0.61
Requête 7	0.65	0.43	0.52	0.60
Requête 8	0.67	0.34	0.45	0.59
Requête 9	0.67	0.46	0.54	0.61
Requête 10	0.65	0.42	0.51	0.59
Requête 11	0.72	0.66	0.69	0.70
Requête 12	0.73	0.65	0.69	0.69
Requête 13	0.76	0.67	0.71	0.70
Requête 14	0.67	0.65	0.66	0.67
Requête 15	0.61	0.61	0.61	0.66
Requête 16	0.68	0.68	0.68	0.69
Requête 17	0.75	0.67	0.71	0.71
Requête 18	0.72	0.64	0.68	0.69
Requête 19	0.72	0.66	0.69	0.70
Requête 20	0.73	0.67	0.70	0.70
Moyenne globale	**0.69**	**0.54**	**0.57**	**0.65**

Tableau 6. Résultats de l'évaluation de la reconnaissance de type de tumeur

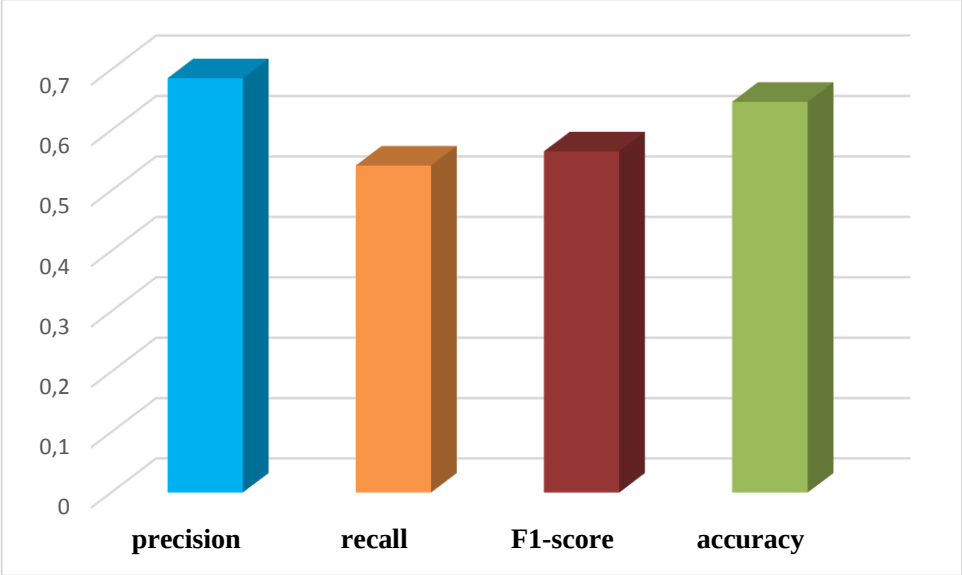


Figure 37. Histogramme des Résultats globales de l'évaluation du modèle de classification KNN

Les résultats de l'application de modèle de classification basé sur KNN montrent que son utilisation est meilleure pour les tumeurs de type bénignes.

Les résultats de l'utilisation de l'algorithme de KNN pour la reconnaissance de type de tumeurs montrent une variation dans l'accuracy compris entre 59% à 71%. L'accuracy maximum rencontré été à 71%. Tandis que l'accuracy minimum été à 59% avec K=5.

Nous jugeons l'accuracy et les autres métriques comme acceptable en vue la simplicité de l'algorithme de KNN. L'application d'un algorithme de classification plus sophistiqué peut donner des taux supérieurs. Nous préférons d'utiliser des algorithmes simple car ils permettent un passage à l'échelle plus efficace.

6. Conclusion

Nous résumons ce qui a été montré par les expérimentations comme suit :

- La combinaison des caractéristiques locales avec LSA donne des bons résultats.
- Il faut remarquer aussi que les expérimentations sont sensibles à quelques paramètres qui sont le nombre des mots visuels, les paramètres internes de modèle. Tous ces paramètres nécessitent une étude supplémentaire.
- LSA ne possède pas une interprétation probabiliste comme les autres modèles.
- Pour l'extraction des caractéristiques, il est préférable d'utiliser des caractéristiques locales et les caractéristiques globales pour enrichir la représentation de l'image.
- La classification par l'algorithme KNN (K-nearest Neighbor) donne des bons résultats.

Conclusion et Perspectives

L'évolution de la technologie a touché plusieurs secteurs dont le secteur médical, d'où l'application d'appareils d'acquisition d'image qui produisent un nombre important d'images chaque année. Cependant l'accès rapide à ces bases d'images énormes nécessite des systèmes d'indexation efficaces.

Plusieurs systèmes d'indexation et de recherche d'images par le contenu visuel ont été présentés dans la littérature pour répondre aux besoins d'utilisateurs et ainsi récupérer les images similaires à partir des bases.

Pour cette raison, on a présenté un système d'indexation d'image médicale par le contenu basé sur les modèles de sujets pour attribuer aux caractéristiques d'une image médicale des concepts sémantiques afin de réduire le fossé sémantique entre l'image médicale et son sens.

Notre approche est basée sur un modèle d'indexation appelée LSA (Latent Semantic Analysis). LSA appartient à la famille des modèles de sujets, cette famille initialement proposée pour la modélisation des connaissances dans les grands corpus de texte. L'avantage de cette famille de modèle réside dans sa capacité à capturer la corrélation des mots dans ces corpus.

Pour construire l'index de l'image, on a passé par plusieurs étapes : la première est la construction des descripteurs SIFT (Scale-invariant feature transform) pour chaque image, la deuxième est la construction des mots visuels à partir des descripteurs des images obtenus dans l'étape précédente c'est l'application de modèle de sujets LSA pour obtenir le descripteur de l'image autrement dit « l'index de l'image ». La dernière étape consiste à faire une classification par KNN pour la reconnaissance de type de tumeur.

Il est possible d'affirmer que le travail mené dans le cadre de ce mémoire a pour ambition de fournir un modèle d'indexation d'images par le contenu permettant d'améliorer la performance des systèmes CBIRs.

Pour l'évaluation de nos expérimentations en utilisent le MAP, qui a été calculé sur 20 requêtes. Ce résultat est compris entre 47% à 61%, c-à-d, qu'il représente des bons résultats, ainsi que les résultats de la tumeur maligne sont meilleurs que la bénigne.

Cependant, ce travail ouvre des perspectives, et on peut dire que beaucoup d'axes de recherche seront poursuivies à court et à long terme.

Nous donnons un résumé sur ces perspectives dans les points suivant :

- Réaliser l'étape d'apprentissage avec un corpus important d'images de différentes pathologies puis faire une classification avec d'autre méthode de classification exemple CNN.
- Appliquer LSA sur une combinaison des caractéristiques locales et globales de l'image pour améliorer la qualité d'indexation.
- Appliquer d'autres modèles de sujets.
- Utiliser le format DICOM car il est riche en informations.
- Utiliser d'autre méthode de classification dans l'étape de construction des mots visuelles.

A. Références Bibliographiques

[1] Mender, B.,Tlili, Y. (2016). Indexation et recherche des tumeurs de mammographie par LSA et caractéristiques globales de texture, Mémoire de Master, Université Badji-Mokhtar Annaba.

[2] Denaib, A., & Kouhili, M. (2015). La recherche et l'indexation des images par une hybridation d'une loi puissance et la méthode K-means (Doctoral dissertation, Université Ahmed Draïa-Adrar).

[3] D. Latha,Dr. Y. Jacob Vetha Raj.(2018). Different Types of CBIR Applications: A Survey. International Journal for Research in Engineering Application & Management (IJREAM). Department of PG Computer Science, NMCC, Tamilnadu, India

[4] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. Journal of the American society for information science, 41(6), 391-407.

[5] Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. Machine learning, 42(1), 177-196.

[6] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

[7] Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. The annals of applied statistics, 1(1), 17-35.

[8] Li, W., & McCallum, A. (2006, June). Pachinko allocation: DAG-structured mixture models of topic correlations. In Proceedings of the 23rd international conference on Machine learning (pp. 577-584).

[9] Tollari, S. (2006). Indexation et recherche d'images par fusion d'informations textuelles et visuelles (Doctoral dissertation, Toulon).

[10] Hörster, E., Lienhart, R., Effelsberg, W., & Möller, B. (2009). Topic models for image retrieval on large-scale databases. ACM Sigmultimedia Records, 1(4), 15-16.

[11] Projet imédia. (2002) « Images et Multimédia : Recherche et Navigation ».Institut National de Recherche en Informatique et en Automatique.

[12] Ordóñez, J. R., Cazuguel, G., Puentes, J., Solaiman, B., & Roux, C. (2003). Indexation d'images médicales basée sur les informations spectrale et spatiale extraites de JPEG-2000. Actes de CORESA'03, Lyon.

[13] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.

[14] Landre, J. (2005). Analyse multirésolution pour la recherche et l'indexation d'images par le contenu dans les bases de données images-Application à la base d'images paléontologique Trans' Tyfipal (Doctoral dissertation, Université de Bourgogne).

[15] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. IEEE Transactions on systems, man, and cybernetics, (6), 610-621.

[16] Glatard, T. (2004). Indexation d'images médicales basée sur le contenu: applicationa la recherche eta la segmentation d'images (Doctoral dissertation, Master's thesis, Ecole Doctorale EEA, Lyon, France).

[17] MEFTAH, M. S. Un serveur dédié à la recherche d'information médicales basé sur le raisonnement à partir de cas (Doctoral dissertation, Université de Ouargla-Kasdi Merbah).

[18] Cao, Y., Li, Y., Müller, H., Kahn Jr, C. E., & Munson, E. (2011). Multi–modal medical image retrieval. In SPIE Medical Imaging.

[19] Stathopoulos, S., Lourentzou, I., Kyriakopoulou, A., & Kalamboukis, T. (2013). IPL at CLEF 2013 Medical Retrieval Task. In CLEF (Working Notes).

[20] Pham, T. T., Maillot, N. E., Lim, J. H., & Chevallet, J. P. (2007, November). Latent semantic fusion model for image retrieval and annotation. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management (pp. 439-444).

[21] Văduva, C., Gavăt, I., & Datcu, M. (2012). Latent Dirichlet allocation for spatial analysis of satellite images. IEEE Transactions on Geoscience and Remote sensing, 51(5), 2770-2786.

[22] Boulemden, A., Tlili, Y., & Jalab, H. A. (2018). Content-based image retrieval with pachinko allocation model and a combination of colour, texture and text features. International Journal of Computational Vision and Robotics, 8(2), 122-139.

[23] Rasiwasia, N., & Vasconcelos, N. (2013). Latent dirichlet allocation models for image classification. IEEE transactions on pattern analysis and machine intelligence, 35(11), 2665-2679.

[24] Glatard, T. (2004). Indexation d'images médicales basée sur le contenu: applicationa la recherche eta la segmentation d'images (Doctoral dissertation, Master's thesis, Ecole Doctorale EEA, Lyon, France).

[25] Amaral, I. F. A. (2010). Content-based image retrieval for medical applications. Faculty of Science, University of Porto.

[26] Belhassen, M., Kalti, K., & Ayeb, B. Indexation des textures des images pulmonaires TDM en vue d'une recherche par le contenu.

[27] Al Sun, M. H. (2012). Indexation guidée par les connaissances en imagerie médicale (Doctoral dissertation, Télécom Bretagne, Université de Bretagne Occidentale).

[28] Najjar, M. (2004). Modèles de mélange pour la recherche d'images par le contenu: Applications aux pathologies ostéo-articulaires (Doctoral dissertation, Compiègne).

[29] Yang, Y., & Liu, X. (1999, August). A re-examination of text categorization methods. In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 42-49).

B. Références Web (Techniques)

[W1] Benchmark databases for CBIR (consulté le 8 Mai, 2022), http://savvash.blogspot.com/2008/12/benchmark-databases-for-cbir.html

[W2] Marine Campedel, Indexation des images (consulté le 1 Avril, 2022), https://muhaz.org/indexation-des-images-marine-campedel.html

[W3] Documentation K-moyennes, (consulté le 15 Mai, 2022), https://fr.wikipedia.org/wiki/K-moyennes

[W4] Documentation Python, (consulté le 18 Juin, 2022), https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1445304-python-definition-et-utilisation-de-ce-langage-informatique/

[W5] Joe Hicklin et al, *JAMA* : A Java Matrix Package, (consulté le 15 Mai, 2022), https://math.nist.gov/javanumerics/jama/

[W6] Documentation NumPy, (consulté le 18 Juin, 2022), https://numpy.org/

[W7] Documentation Scikit-learn(sklearn) in Python, (consulté le 18 Juin, 2022), https://www.analyticsvidhya.com/blog/2015/01/scikit-learn-python-machine-learning-tool/

[W8] Documentation NetBeans, (consulté le 15 Mai, 2022), https://fr.wikipedia.org/wiki/NetBeans

[W9] Documentation MATLAB, (consulté le 15 Mai, 2022), https://fr.wikipedia.org/wiki/MATLAB

[W10] Documentation Google Colaboratory, (consulté le 18 Juin, 2022), https://stepbystepinternet.com/fr/google-colaboratory-de-quoi-sagit-il-a-quoi-cela-sert-il-et-comment-pouvons-nous-profiter-de-colab/

[W11] The mini-MIAS database of mammograms 5, (consulté le 02 janvier, 2022), http://peipa.essex.ac.uk/info/Database/mias/

A. Les indices d'Haralick

Dans son article "Textural features for image classification", Haralick introduit quatorze attributs de texture extraits des matrices de cooccurrences. Ces attributs sont les suivants **[15]**:

A.1 Second moment angulaire :

$$f_1 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{p(i,j)\}^2 \quad (30)$$

Le second moment angulaire (ou énergie) mesure l'homogénéité de l'image. Plus cette valeur est faible, moins l'image est uniforme et dans ce cas, il existe beaucoup de transitions de couleurs.

A.2 Contraste :

$$f_2 = \sum_{n=0}^{N-1} n^2 \left\{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{p(i,j)\}\right\} \quad avec \quad |i-j| = n \quad (31)$$

La valeur en est d'autant plus élevée que la texture présente un fort contraste. Ce paramètre est fortement non corrélé à l'énergie.

A.3 Corrélation :

$$f_3 = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (ij) P(i,j) - u_x u_y}{\sigma_x \sigma_y} \quad (32)$$

Où σ_x, σ_y, u_x et u_y sont les moyennes et les standard déviations de P_x et P_y .

La corrélation mesure la dépendance linéaire (relativement à (i,j)) des niveaux de gris de l'image. La corrélation n'est corrélée ni à l'énergie, ni à l'entropie.

A.4 Variance :

$$f_4 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-\mu)^2 [P](i,j) \quad (33)$$

La variance mesure l'hétérogénéité de la texture. Elle augmente lorsque les niveaux de gris différent de leur moyenne. La variance est indépendante du contraste.

A.5 Moment différentiel inverse :

$$f_5 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1+(i-j)^2} P(i,j) \quad (34)$$

A.6 Moyenne des sommes :

$$f_6 = \sum_{i=2}^{2N_g} i \, P_{x+y}(i) \tag{35}$$

A.7 Variance des sommes :

$$f_7 = \sum_{i=2}^{2N_g} (1 - f_8)^2 P_{x+y}(i) \tag{36}$$

A.8 Entropie des sommes :

$$f_8 = -\sum_{\boldsymbol{i=2}}^{2N_g} P_{x+y}(\boldsymbol{i}) \log\{P_{x+y}(\boldsymbol{i})\} \tag{37}$$

A.9 Entropie :

$$f_9 = -\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j) \log\{P(i,j)\} \tag{38}$$

Ce paramètre mesure le désordre dans l'image. Contrairement à l'énergie, l'entropie atteint de fortes valeurs lorsque la texture est complètement aléatoire (sans structure apparente). Elle est fortement corrélée (par l'inverse) à l'énergie.

A.10 Variance des différences :

$$f_{10} = \sum_{i=0}^{N-1} (l - f_{11})^2 P_{x+y}(i) \tag{39}$$

A.11 Entropie des différences :

$$f_{11} = -\sum_{\boldsymbol{l=0}}^{\boldsymbol{N-1}} P_{x-y}(\boldsymbol{i}) \, \mathbf{log}\{P_{x-y}(\boldsymbol{i})\} \tag{40}$$

A.12 Information sur la corrélation :

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

$$HXY = -\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j) \log\{P(i,j)\} \tag{41}$$

Où HX et HY sont les entropies de Px et Py.

$$HXY1 = -\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j) \log\{P_x(i) P_y(j)\}$$

A.13 Information sur la corrélation :

$$f_{13} = (1 - exp[-2.0(HXY2 - HXY)])^{\frac{1}{2}} \tag{42}$$

Où $HXY2 = -\sum_{\boldsymbol{i=0}}^{\boldsymbol{N-1}} \sum_{\boldsymbol{j=0}}^{\boldsymbol{N-1}} P_x(\boldsymbol{i}) P_y(j) \log\{P_x(i) P_y(j)\}$

A.14 Coefficient de corrélation maximal :

$$f_{14} = \left(1^{ère} plus\ grand\ valeur\ propre\ de\ Q\right)^{\frac{1}{2}} \tag{43}$$

Les attributs de $f_6 \ldots . f_{14}$ apportent des informations supplémentaires sur les degrés d'homogénéité et de complexité de l'image, ainsi que sur la corrélation.

B. L'implémentation du SIFT

Cette annexe est dédiée à la représentation d'une partie du code source matlab d'implémentation du technique SIFT (Scale-invariant feature transform) pour l'extraction des caractéristiques locale d'images. Cette implémentation a été réalisée par Andrea Vedaldi et qui est compatible avec celle de D. Lowe **[13]**.

La fonction SIFT a de nombreux paramètres pour l'initialisation. Ces derniers peuvent être des paramètres qui décrivent l'espace de l'échelle, ainsi que des paramètres de détecteur et de descripteur. Les valeurs par défaut ont été choisies pour émuler l'implémentation originale de Lowe. La figure 38 présente ses paramètres.

```
%                                                     Check the arguments
% --------------------------------------------------------------------
if(nargin < 1)
  error('At least one argument is required.') ;
end
[M,N,C] = size(I) ;

% Lowe's equivalents choices

S      =  3 ;                          % NumLevels; Number of scale levels within each octave.
omin   = -1 ;                          % Index of the first octave
O      = floor(log2(min(M,N)))-omin-3 ; % Numoctaves; Number of octaves of the Gaussian scale space.
sigma0 = 1.6*2^(1/S) ;                 % Base smoothing [pixels], smooth lev. -1 at 1.6
sigman = 0.5 ;                         % Nominal smoothing of the input image.
thresh = 0.04 / S / 2 ;                % Maxima of the DOG scale space
r      = 10 ;                          % Localization threshold [>= 0, (10)]
NBP    = 4 ;                           % Number of spatial bins
NBO    = 8 ;                           % Number of orientation bins
magnif = 3.0 ;                         % Descriptor window magnification

% Parese input
compute_descriptor = 0 ;
discard_boundary_points = 1 ;
verb = 0 ;

for k=1:2:length(varargin)
  switch lower(varargin{k})

    case 'numoctaves'
      O = varargin{k+1} ;
```

Figure 38. Une partie du code présente les paramètres de la fonction SIFT

La fonction SIFT commence par générer l'espace d'échelles suivi de DoG (Difference of Gaussians) sur l'image pour former les octaves (Voir la figure 39).

Pour chaque octave, les différences entre les images gaussiennes (DoG) sont calculées pour localiser les extrema (ou bien des points SIFT) du DoG. Ensuite tous les points SIFT détectés sont filtrer pour éliminer les points les moins contrastés ainsi que les points situés sur une arrête.

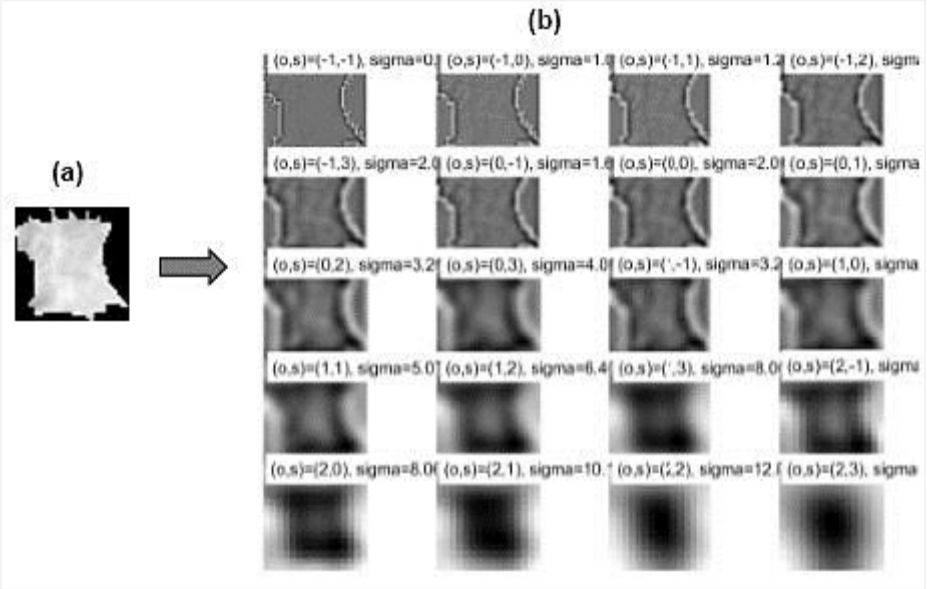


Figure 39. Construction de l'espace des échelles; l'image originale (a) et les octaves (b)

```
% Local maxima of the DOG octave
% The 80% tricks discards early very weak points before refinement.
idx = siftlocalmax(  dogss.octave{o}, 0.8*thresh  ) ;
idx = [idx , siftlocalmax( - dogss.octave{o}, 0.8*thresh)] ;

K=length(idx) ;
[i,j,s] = ind2sub( size( dogss.octave{o} ), idx ) ;
y=i-1 ;
x=j-1 ;
s=s-1+dogss.smin ;
oframes = [x(:)';y(:)';s(:)'] ;

if verb > 0
  fprintf('SIFT: %d initial points (%.3f s)\n', ...
    size(oframes, 2), toc) ;
  tic ;
end

% Remove points too close to the boundary
if discard_boundary_points
  % radius = maginf * sigma * NBP / 2
  % sigma = sigma0 * 2^s/S

  rad = magnif * gss.sigma0 * 2.^(oframes(3,:)/gss.S) * NBP / 2 ;
  sel=find(...
    oframes(1,:)-rad >= 1                          & ...
    oframes(1,:)+rad <= size(gss.octave{o},2) & ...
    oframes(2,:)-rad >= 1                          & ...
    oframes(2,:)+rad <= size(gss.octave{o},1)      ) ;
```

Figure 40. Une partie du code de la construction de l'espace des échelles et détection des extrémas

A la fin, des orientations des gradients des pixels voisins du point SIFT détecté sera calculées. Ces orientations sont rangées dans un histogramme comme l'illustre la figure 41.

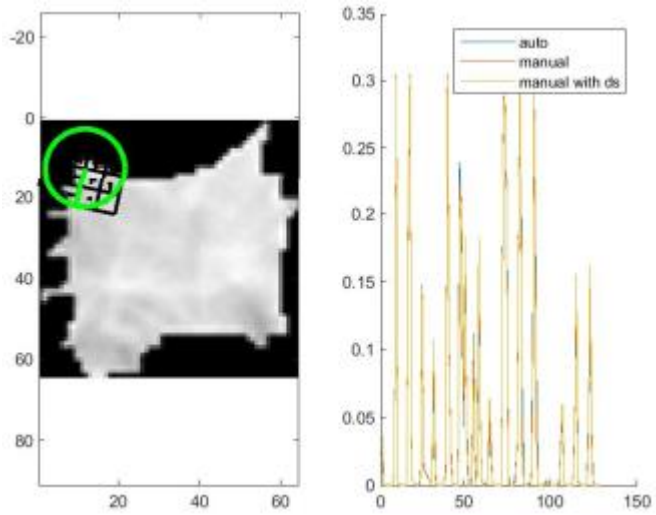


Figure 41. Création de l'histogramme des gradients

La fonction SIFT renvoie comme résultat des matrices de frames de taille 4× K contenant les frames SIFT et des matrices de descripteurs 128 × K contenant les descripteurs de ces frames.

```
% Compute the oritentations
oframes = siftormx(...
  oframes, ...
  gss.octave{o}, ...
  gss.S, ...
  gss.smin, ...
  gss.sigma0 ) ;

% Store frames
x     = 2^(o-1+gss.omin) * oframes(1,:) ;
y     = 2^(o-1+gss.omin) * oframes(2,:) ;
sigma = 2^(o-1+gss.omin) * gss.sigma0 * 2.^(oframes(3,:)/gss.S) ;
frames = [frames, [x(:)' ; y(:)' ; sigma(:)' ; oframes(4,:)] ] ;


% Descriptors
if nargout > 1
  if verb > 0
    fprintf('SIFT: computing descriptors...') ;
    tic ;
  end

  sh = siftdescriptor(...
    gss.octave{o}, ...
    oframes, ...
    gss.sigma0, ...
    gss.S, ...
    gss.smin, ...
    'Magnif', magnif, ...
    'NumSpatialBins', NBP, ...
    'NumOrientBins', NBO) ;

  descriptors = [descriptors, sh] ;

  if verb > 0, fprintf('done (%.3f s)\n',toc) ; end
```

Figure 42. Une partie du code du résultat de descripteur SIFT

C. L'implémentation du modèle LSA

L'annexe C représente une partie du code source java d'implémentation du LSA (Latent Semantic Analysis) pour obtenir les index des images tumeurs.

La classe « DataPreparation » contient toutes les méthodes nécessaires pour l'application du LSA, tel que ; analyse (lire les mots visuelles pour chaque image), tf_idf_prep (calcule de fréquence), Appliquer_LSA (utilise SVD), euclid (calcule les distance), tri (trier les distances) et evaluationMAP (calcule précision pour chaque requête).

Dans un premier temps, elle lit et calcule la fréquence (TF-IDF) des mots visuelle pour les images de train et test, afin de pouvoir construire les deux matrices co-occurrence « mot-visuel-image ».

```
76      public void tf_idf_prep()
77      {
78      System.out.println("Lecture des images de train");
79
80      for (int k=0;k<140;k++)
81          {
82          int f=k+1;
83          String chemin="d:/mender/tf_idf_prep/"+f+".txt";
84          ReadFile rf=new ReadFile();
85          String text =rf.reading(chemin);
86          System.out.println("Image num:"+k);
87          double [] vecteur=this.analyse(text);
88          for (int i=0;i<20;i++) this.term_doc[i][k]=vecteur[i];
89          }
90
91
92      for (int i=0;i<20;i++)
93          for (int j=0;j<140;j++)
94          {
```

```
0.0**0.0**0.0**0.0**0.0**0.0**3.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**2.0
3.0**3.0**1.0**1.0**1.0**0.0**2.0**2.0**2.0**2.0**3.0**0.0**1.0**0.0**2.0
0.0**0.0**1.0**1.0**0.0**3.0**0.0**2.0**0.0**2.0**2.0**0.0**0.0**1.0**0.0
0.0**0.0**1.0**1.0**2.0**0.0**0.0**1.0**0.0**0.0**0.0**3.0**0.0**1.0**3.0
4.0**4.0**1.0**1.0**0.0**4.0**0.0**1.0**4.0**1.0**4.0**0.0**1.0**0.0**0.0
0.0**5.0**0.0**0.0**0.0**1.0**1.0**2.0**5.0**1.0**5.0**2.0**0.0**0.0**1.0
4.0**4.0**0.0**0.0**0.0**3.0**2.0**2.0**5.0**0.0**5.0**3.0**0.0**2.0**1.0
4.0**0.0**1.0**1.0**2.0**0.0**4.0**2.0**3.0**1.0**4.0**0.0**1.0**0.0**3.0
5.0**0.0**1.0**1.0**2.0**0.0**4.0**2.0**4.0**2.0**0.0**0.0**1.0**2.0**3.0
5.0**0.0**1.0**1.0**0.0**2.0**0.0**0.0**0.0**1.0**5.0**1.0**1.0**0.0**0.0
2.0**4.0**0.0**0.0**0.0**0.0**2.0**0.0**5.0**0.0**0.0**4.0**0.0**0.0**0.0
0.0**0.0**1.0**1.0**0.0**0.0**1.0**5.0**4.0**0.0**0.0**5.0**0.0**1.0**0.0
0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**5.0**1.0**0.0**1.0**1.0
0.0**4.0**0.0**0.0**2.0**0.0**3.0**3.0**5.0**2.0**3.0**5.0**0.0**0.0**2.0
0.0**3.0**1.0**1.0**0.0**2.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**2.0**0.0
5.0**4.0**0.0**0.0**2.0**3.0**1.0**0.0**6.0**2.0**0.0**2.0**0.0**1.0**2.0
2.0**0.0**0.0**0.0**0.0**1.0**0.0**1.0**0.0**0.0**2.0**6.0**1.0**1.0**0.0
0.0**7.0**0.0**0.0**1.0**3.0**0.0**5.0**5.0**0.0**0.0**3.0**0.0**0.0**0.0
11.0**1.0**3.0**3.0**6.0**8.0**1.0**9.0**9.0**11.0**7.0**8.0**4.0**6.0**5
0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0**0.0
Lecture des images de test
Image num:0
```

Figure 43. A gauche : Une partie du code pour calculer tf-idf des images de train, à droite : le résultat - matrice co-occurrence (lignes : 20 mots et colonne : 140 image) pour la base train

LSA applique SVD sur le deux matrices co-occurrence afin de prendre en compte le contexte des mots visuelles et de réduire les dimensions de l'espace selon la formule mathématique (1), le SVD décompose chacune de ces deux matrices on trois matrices ; une matrice des documents (U) qui représente l'espace de l'indexation des images d'apprentissage, une matrice de mots visuels (V) et une matrice des valeurs singulières (S).

Après, la distance euclidienne sera calculé entre chaque image de requête et les images de train (entre les deux matrices transpose (U).

```
118         this.Appliquer_LSA();
119
120
121     }
122
123     public double euclid(double [] req, double [] im)
124         {
125         double dist=0.0;
126         for (int i=0;i<req.length;i++)
127         dist=dist+((im[i]-req[i])*(im[i]-req[i]));
128         dist=Math.sqrt(dist);
129         return dist;
130         }
131     public void Appliquer_LSA()
132         {
133
134             int [][] resultats=new int [20][140];
135             double [][] distances=new double[20][140];
136
137              Matrix train=new Matrix(this.term_doc);
138              Matrix test=new Matrix(this.term_doc_test);
139              Matrix s=train.transpose();
140              Matrix t=test.transpose();
141       SingularValueDecomposition svd=new SingularValueDecomposition(s);
142
143              Matrix U=svd.getU();//docs
144              Matrix S=svd.getS();
145              Matrix V=svd.getV();//mots
146            System.out.println("Affichage des distances");
```

Figure 44. Une partie du code de la méthode LSA et la méthode distance euclidienne

À la fin, le MAP est appliqué pour évaluer la précision des réponses du système en tenant compte de la position de la réponse pertinente dans l'ensemble des réponses.

```
196     public void evaluationMAP (int [][] resultat)
197     {
198         double [] precisions = new double [resultat[0].length];
199         double [] averagesP=new double[resultat.length];
200         double map;int i;
201         //*****************************************************
202        int nb_tr;double sumAP=0.0;
203        for (int k=0;k<averagesP.length;k++)
204        {
205            nb_tr=0;double sum=0.0;
206         for (i=0;i<precisions.length;i++)
207         {
208            if (resultat[k][i]<70 && k<10)
209            {
210            nb_tr++;
211
212                    }
213
214            if (resultat[k][i]>=70 && k>=10)
215            {
216            nb_tr++;
217
218            }
219            precisions[i]=(double) nb_tr/(i+1);
220            sum=sum+precisions[i];
221         if (nb_tr==70) break;
222         }
223         averagesP[k]=sum/i;
```

Figure 45. Une partie du code pour l'évaluation du MAP de LSA

Résumé

De nos jours, les systèmes médicaux produisent une grande quantité de données images qui sont stockées dans des bases de données, l'accès rapide à ces bases énormes nécessite des algorithmes d'indexation efficaces. L'indexation des images médicales est devenu pour les applications cliniques un outil essentiel, parce qu'elle apporte une aide efficace aussi bien en diagnostic qu'au suivi thérapeutique. Les systèmes de recherche d'images par contenu (en anglais Content Based Image Retrievel) sont l'une des solutions possibles pour gérer efficacement ces bases.

Nous avons concentré dans ce travail sur l'application des modèles de sujets (plus particulièrement le modèle Latent Semantic Analysis « LSA ») dans le contexte de l'indexation et recherche d'image médicale. Le défi consiste à attribuer aux caractéristiques d'une image des concepts sémantiques.

Les résultats de cette étude a permet de confirmer l'utilité de ce modèle comme modèle d'indexation.

Mots clés : L'indexation des images médicales, Les systèmes de recherche d'images par contenu, modèles de sujets, Latent Semantic Analysis.

Abstract

Today, medical systems produce a large amount of image data that is stored in databases. Fast access to these huge databases requires efficient indexing algorithms. Medical image indexing has become an essential tool for clinical applications because it provides effective support for both diagnosis and therapeutic follow-up. Content Based Image Retrieval systems are one of the possible solutions to manage these databases efficiently.

In this work, we focus on the application of topic models (more specifically the Latent Semantic Analysis « LSA » model) in the context of medical image indexing and retrieval. The challenge is to assign semantic concepts to the features of an image.
The results of this study confirmed the usefulness of this model as an indexing model.

Keywords: Medical image indexing, Content-based image retrieval systems, topic models, Latent Semantic Analysis.

ملخّص

في الوقت الحاضر، تنتج الأنظمة الطبية كمية كبيرة من بيانات الصور المخزنة في قواعد البيانات، ويتطلب الوصول السريع إلى هذه القواعد الضخمة خوارزميات فهرسة فعالة. لقد أصبحت فهرسة الصور الطبية للتطبيقات السريرية أداة أساسية، لأنها توفر مساعدة فعالة في التشخيص والمراقبة العلاجية، كما تعد أنظمة استرجاع الصور استنادا على المحتوى أحد الحلول الممكنة لإدارة قواعد البيانات هذه بكفاءة.
لقد ركزنا في هذا العمل على تطبيق نماذج المواضيع (وبشكل أكثر تحديدًا نموذج التحليل الدلالي الكامن "LSA") في مجال فهرسة واسترجاع الصور الطبية.

ويتمثل التحدي الآن في تعيين المفاهيم الدلالية إلى خصائص الصورة.

أكدت نتائج هذه الدراسة فائدة هذا النموذج كنموذج فهرسة.

الكلمات المفتاحية: فهرسة الصور الطبية، أنظمة البحث عن الصور استنادا على المحتوى، نماذج المواضيع، التحليل الدلالي الكامن.