

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE - MINISTERE DE
L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE -
ECOLE NATIONALE POLYTECHNIQUE.



M0033/05B

DEPARTEMENT D'ELECTRONIQUE
Option Signal et Communications

**MEMOIRE DE MAGISTER
EN ELECTRONIQUE**

المدرسة الوطنية المتعددة التقنيات
المكتبة — BIBLIOTHEQUE
Ecole Nationale Polytechnique

Présenté par :

DOUACHE Djamila

Ingénieur d'Etat en Electronique

Système Autonome de Diagnostic du Cancer du Col Utérin

Soutenu le 14/09/2005 à 10H

Devant le Jury d'examen composé de:

Président :	A. BELOUHRANI	Maître de Conférences	ENP
Rapporteur :	L. HAMAMI	Maître de Conférences	ENP
Examineurs :	A. ZERGUERRAS M. GUERTI N. LASSOUAOUI	Professeur Maître de Conférences Chargée de Recherche	ENP ENP CERIST
Invitée :	R. BELAHRACHE	Médecin	INSP

SEPTEMBRE 2005

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE – MINISTERE DE
L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE –
ECOLE NATIONALE POLYTECHNIQUE.



DEPARTEMENT D'ELECTRONIQUE
Option Signal et Communications

**MEMOIRE DE MAGISTER
EN ELECTRONIQUE**

Présentée par :

DOUACHE Djamila

Ingénieur d'Etat en Electronique



Systeme Autonome de Diagnostic du Cancer du Col Utérin

Soutenu le 14/09/2005 à 10H

Devant le Jury d'examen composé de:

Président :	A. BELOUHRANI	Maître de Conférences	ENP
Rapporteur :	L. HAMAMI	Maître de Conférences	ENP
Examineurs :	A. ZERGUERRAS	Professeur	ENP
	M. GUERTI	Maître de Conférences	ENP
	N. LASSOUAOUI	Chargée de Recherche	CERIST
Invitée :	R. BELAHRACHE	Médecin	INSP

SEPTEMBRE 2005

السرطان من إحدى الأمراض الخطيرة والشائعة في العالم في هذا القرن الأخير ويمثل من أكبر سبب لعدد الوفيات نقص الوسائل والمختصون في ميدان علم الخلايا أدى إلى تعطيل كشف تطعيم لهذا الأذى ولهذا أردنا مساعدة الأطباء في تشخيصه المرض وفي التحري بتحقيق نظام آلي يتركز على تقنيات معالجة الصور والذكاء الاصطناعي
مفاتيح : فرا كتال - تجزيء- الذكاء الاصطناعي - خلايا - السرطان

Résumé :

Durant ces dernières années, le rôle du traitement d'image dans le domaine biomédical a considérablement avancé; alors que ses ressources sont maintenant exploitées, pour aider les médecins dans leur diagnostic.

Nous avons proposé, la réalisation d'un système de diagnostic du cancer, basé sur une étude cytologique.

Notre travail, consiste à traiter et segmenter, une image de cellule prise à partir d'un prélèvement cervical, et de déterminer le type et la nature de chaque cellule. Ceci en utilisant l'approche multifractale pour la segmentation et les réseaux de neurones pour la reconnaissance.

Mots clés : Cellules - cancer - segmentation - réseau de neurones - multifractales

Abstract:

Cancer is the one of the most dangerous illness in the world. Unfortunately, high performances which permit its precocious detection don't exist currently; excepting the anatomic pathologic exam, named the tracking.

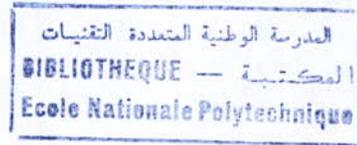
The tracking is possible by microscopic exam of a smear of cells sampled from an organ in order to check if it's an infected or not one.

The goal of our work is to conceive a system of tracking of cancer; it should recognize the malignity or not of cell, by a treatment of a picture of the cervical smears.

In that way, we create a system based on the image processing, it includes several modules as acquisition of the image, treatment, segmentation, analysis and recognition.

Keys words: Cancer - segmentation - fractal -neural network - smear cervical - traking .

Remerciements



Remerciements

En premier lieu, je tiens à adresser mes plus vifs remerciements à Mme HAMAMI, pour les conseils avisés, les connaissances dont elle m'a fait bénéficier, son suivi attentif et sa confiance qui m'a été très précieuse.

J'exprime ma profonde reconnaissance à Mr. A. BELOUHRANI, qui m'a fait l'honneur de présider le jury.

Je tiens aussi à exprimer mes remerciements et ma plus vive gratitude à Melle M. GUERTI, Melle N. LASSOUAOUI et Mr. A. ZERGUERRAS qui ont accepté de juger mon travail.

J'adresse mes plus vifs remerciements au Dr BELAHRACHE, qui m'a ouvert les portes de son laboratoire et a mis gracieusement à ma disposition les prélèvements biologiques.

Je remercie Mme DOUACHE qui a contribué activement à l'élaboration et la rédaction de ce travail, par ses conseils et son soutien.

Je remercie Mr DOUACHE pour le soutien moral mais aussi matériel qui a contribué à la concrétisation de ce projet.

Mes sincères remerciements vont à Melle DOUACHE Amina, qui a su me conseiller pour mettre en œuvre et coordonner ce travail.

Je remercie Melle BENDI Samah pour ses encouragements son aide et sa bonne humeur. Je remercie Mr AHMED-ZAID Salim pour les mots justes trouvés à chaque occasion. Un grand merci pour Mr MELLAK Driss pour avoir activement collaborer à la réalisation de ce travail.

Enfin, je remercie en particulier, toute personne ayant contribué de près ou de loin à l'élaboration de ce travail.

المدرسة الوطنية المتعددة التقنيات
المكتبة — BIBLIOTHEQUE
Ecole Nationale Polytechnique

Dédicaces

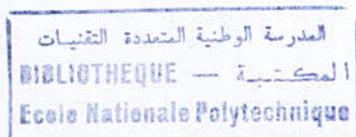
Dédicaces

*Je dédie ce travail à la mémoire de mes grands parents disparus trop tôt pour que je
puisse les connaître*

*A mes parents qui me sont très chers
A mon unique petite sœur
Ainsi qu'à toute ma famille*



Sommaire



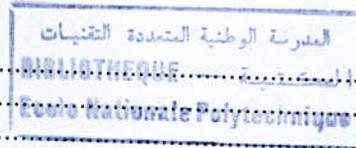
Sommaire

PROBLEMATIQUE	2
INTRODUCTION GENERALE	3
CHAPITRE 1 : CANCEROLOGIE ET CYTOLOGIE.....	7
INTRODUCTION :	7
1. QU'EST CE QUE LE CANCER ? :	8
2. LA CELLULE CANCEREUSE :	8
a) Les anomalies morphologiques :	8
a-1. Anomalies du noyau :	9
a-2. Anomalies du cytoplasme :	9
b) Les anomalies de la croissance et de la prolifération cellulaire :	10
c) Anomalies génétiques :	10
3. ETUDE SPECIFIQUE DE QUELQUES CANCERS :	11
3.1 Introduction :	11
3.2 Cancer du col de l'utérus :	11
3.2.1 Un peu d'anatomie :	11
3.2.2 Qu'est-ce que le cancer du col utérin? :	11
3.2.3 Qu'est-ce que la dysplasie du col utérin? :	12
3.2.4 Qu'est ce qu'un frottis du col utérin ? :	12
3.3 Morphologie des cellules normales :	13
3.3.1. La couche basale :	14
3.3.2. La couche para-basale :	14
3.3.3. La couche intermédiaire :	14
3.3.4. La couche superficielle :	15
4. Critères de malignité :	15
4.1 Perturbation du rapport nucléo-cytoplasmique en faveur du noyau :	15
4.2 Anisonucléose (Anisocariose) :	16
4.3 Difformité nucléaire (Polymorphisme nucléaire) :	16
4.4 Hyperchromatisme nucléaire :	16
CONCLUSION :	16
CHAPITRE 2 : VISION PAR ORDINATEUR	18
INTRODUCTION :	18
1. L'ACQUISITION :	19
1.1 Le scanner :	20
1.2 La caméra numérique :	20
1.3 Le microscope :	20
2. LE TRAITEMENT D'IMAGE :	21
2.1 Le prétraitement :	21
2.2. La segmentation :	22
2.2.1 Approche frontière :	22
2.2.2 Approche région :	22
3. La reconnaissance :	22
4. Les domaines d'application du traitement d'images :	23

CONCLUSION :	23
CHAPITRE 3 : SEGMENTATION.....	25
INTRODUCTION :	25
1. DEFINITION DE LA SEGMENTATION :	26
2.1 DEFINITION DES FRACTALES :	26
2.2. DIMENSION FRACTALE : [.....]	27
2.3. DIFFERENTS TYPES DE FRACTALES :	28
2.3.1. Fractales aléatoires :	28
2.3.2. Fractales déterministes :	28
2.3.3. Fractales à homothétie interne :	29
2.3.4. Multifractales :	29
2.4. FORMALISME MULTIFRACTAL :	29
a) L'exposant de singularité :	30
b) Le spectre de singularité :	30
c) Algorithme multifractal pour la segmentation:	30
3.1. DETECTION DE CONTOURS : APPROCHES GLOBALES.....	32
3.1.1 Les détecteurs de gradient par filtrage :	32
3.1.2 Les détecteurs de gradient par masques :	32
3.1.3 Les détecteurs de passage par zéro du laplacien :	32
3.1.4 Les détecteurs par masquage adapté :	33
3.2. LES CONTOURS ACTIFS = LES SNAKES :	33
3.2.1 Une approche physique des contours :	33
3.2.2 Mise en équations du principe du contour actif :	34
3.2.3 Energie totale du contour actif :	35
3.3. EVOLUTION TEMPORELLE DU CONTOUR ACTIF :	36
3.4. PROBLEMES LIES AUX CONTOURS ACTIFS :	37
CONCLUSION :	37
CHAPITRE 4 : LES RESEAUX DE NEURONES.....	39
INTRODUCTION :	39
1. LE NEURONE BIOLOGIQUE :	40
2. FONCTIONNEMENT DU NEURONE BIOLOGIQUE :	41
3. LE NEURONE FORMEL :	41
3.1. Interprétation mathématique :	42
3.2. Les Fonctions seuil de bases :	43
a) Fonction binaire à seuil :	43
b) Fonction sigmoïde :	43
4. PRESENTATION DU RNA :	44
4.1. Architecture d'un réseau de neurone artificiel :	45
a) Les réseaux "FEED-FORWARD" :	45
b) Les réseaux "FEED-BACK" :	46
4.2. Apprentissage :	47
a) Le mode supervisé :	48
b) Le mode non-supervisé (ou auto-organisationnel) :	48
c) Le mode hybride :	48
4.3. Règles d'apprentissage :	48
a) Règle de correction d'erreurs :	48
b) Apprentissage de Boltzmann :	48
c) Règles de Hebb :	49
d) Règle d'apprentissage par compétitions :	49
CONCLUSION :	49
CHAPITRE 5 : LE TRAVAIL EFFECTUE	52
INTRODUCTION :	52

Sommaire

1. L'ACQUISITION :	53
2. PRETRAITEMENT :	54
3. LA SEGMENTATION :	56
4. ANALYSE ET RECONNAISSANCE :	62
4.1. L'extraction des critères de malignité :	62
4.1.1 Le rapport nucléo-cytoplasmique :	62
4.1.2. L'anisocaryose.....	64
4.1.2.1. Calcul du centre de gravité de chaque noyau	64
4.1.2.2. Calcul de la taille moyenne du noyau.....	64
4.1.3. La difformité nucléaire.....	65
4.1.4. L'hyperchromatisme nucléaire.....	65
4.2. Classification des cellules en fonction du rapport nucléo-cytoplasmique.....	66
4.3. Les réseaux de neurones :	68
4.3.1. Le réseau Multicouche	68
4.3.1.1. Définition.....	68
4.3.1.2. Couche d'entrée.....	68
4.3.1.3. Couche cachée.....	68
4.3.1.4. Couche de sortie.....	69
5. DESCRIPTION DU LOGICIEL :	72
5.1. DESCRIPTION DE L'APPLICATION :	73
CONCLUSION :	79
CONCLUSION GENERALE.....	74
ANNEXE 1	V
ANNEXE 2	XIII
BIBLIOGRAPHIE	II



المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

Table des Figures

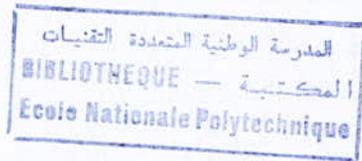


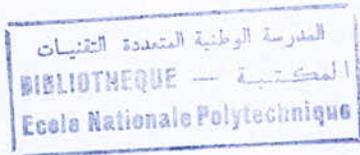
Table des Figures

FIGURE 1.1 : REPRESENTATION DE L'APPAREIL GENITAL FEMININ.....	11
FIGURE 1.2 : CELLULES MALPIGHIENNES - A : CELLULES SUPERFICIELLES (FLECHES) ; B : CELLULES INTERMEDIAIRES ; C : CELLULES PARABASALES ; D : CELLULES METAPLASIQUES.....	13
FIGURE 2.1 : SYSTEME DE VISION PAR ORDINATEUR.....	19
FIGURE 2.2 MICROSCOPE UTILISE.....	21
FIGURE 3.1 : MODELE DE FRACTALE IFS.....	29
FIGURE 3.2 : LE "FLOCON DE NEIGE" DE VON KOCK.....	29
FIGURE 3.3 : REPRESENTATION DU CONTOUR ACTIF.....	34
FIGURE 3.4 : ORGANIGRAMME POUR LA SEGMENTATION PAR CONTOUR ACTIF.....	36
FIGURE 3.5 : EXEMPLE D'APPLICATION DE L'ALGORITHME DE DETECTION DE CONTOUR PAR CONTOUR ACTIF.....	36
FIGURE 4.1 : LE NEURONE BIOLOGIQUE.....	40
FIGURE 4.2 LE NEURONE FORMEL.....	42
FIGURE 4.3 REPRESENTATION MATHEMATIQUES DU NEURONE FORMEL.....	42
FIGURE 4.4: FONCTION BINAIRE A SEUIL.....	43
FIGURE 4.5 : FONCTION EXPONENTIELLE.....	44
FIGURE 4.6 : FONCTION TANGENTIELLE.....	44
FIGURE 5.1 : DISPOSITIF D'ACQUISITION D'IMAGES.....	53
FIGURE 5.2 : IMAGE EXTRAITE DU MICROSCOPE.....	53
FIGURE 5.4 : IMAGE DE CELLULE (A) BRUTE (B) CONTRASTEE (C) GRISEE.....	55
FIGURE 5.5. (A) IMAGE AVEC DES ZONES INEXPLOITABLES (B) UTILISATION DE LA GOMME POUR LA SELECTION.....	56
FIGURE 5.5. (C) REMPLACEMENT DES ZONES INDESIREES PAR DES PIXELS BLANCS (D) IMAGE FILTREE.....	56
FIGURE 5.6 : IMAGE RESULTAT DE SEGMENTATION D'UNE CELLULE.....	60
FIGURE 5.7 : IMAGE NON SEGMENTABLE.....	60
FIGURE 5.8 : MISE EN EVIDENCE DES PARTICULES ISOLEES SUR UNE IMAGE FILTREE ET APRES LA SEGMENTATION DES NOYAUX.....	61

FIGURE 5.11 : IMAGES RESULTATS DE LA CLASSIFICATION PAR RAPPORT AU RNC 67	
FIGURE 5.12 : SCHEMA DU RESEAU DE NEURONES ADAPTE.....	69
FIGURE 5.13 : REPRESENTATION DE LA BASE DE DONNEE.....	71
FIGURE 5.14 : AFFICHAGE DU RESULTAT DE LA RECONNAISSANCE.....	72
FIGURE 5.15 : L'INTERFACE DU LOGICIEL	73
FIGURE 5.16 : BOITE DE DIALOGUE POUR OUVRIR.....	74
FIGURE 5.17: L'HISTOGRAMME DES NIVEAUX DE GRIS.....	75
FIGURE 5.18 : BOITE DE DIALOGUE POUR LA BINARISATION.....	75
FIGURE 5.19 : SEGMENTATION DES NOYAUX	77
FIGURE 5.20 : SEGMENTATION DU CYTOPLASME	77
FIGURE 5.21 : SEGMENTATION DES CELLULES	78
FIGURE 5.22 : RECAPITULATIF	78
FIGURE 5.23 : AFFICHAGE DE LA BOITE DE RESULTAT CALCUL DES CRITERES DE MALIGNITE.....	79

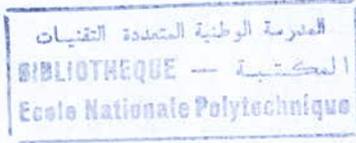
المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

Problématique

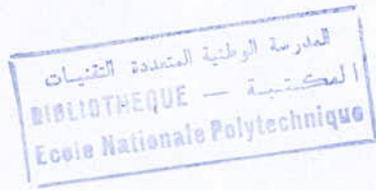


Problématique

Le cancer est l'une des maladies les plus dangereuses et les plus répandues dans le monde. Celui-ci est cause de nombreux décès et aucun moyen de prévention, tel que le vaccin, n'a encore été mis au point. En attendant, le seul moyen qui permette d'attaquer à temps cette maladie, et par conséquent de diminuer considérablement le taux de mortalité, est le dépistage (c'est un examen anatomo-pathologique ou au microscope, d'un fragment de tissu suspect). Ce dernier étant très fastidieux, des chercheurs ont pensé à automatiser ce travail. Cet automatisme consiste à mettre au point un système d'aide aux médecins en éliminant tous les cas normaux (par screening) et en ne présentant aux médecins que les cas anormaux par visualisation sur écran. Grâce à ce système, on pourra réaliser un dépistage précoce de la maladie, ce qui permettra un traitement plus efficace et augmentera, par conséquent, les chances de survie du patient. On se rend compte, toutefois, que l'examen des cellules prélevées pour le dépistage est particulièrement complexe et consommateur de temps, car il consiste à rechercher quelques cellules anormales parmi un amas de cellules normales. Le résultat de cet examen, ne repose que sur l'avis du médecin, ce qui augmente le risque d'erreurs. Dans le but de pallier cette difficulté, nous voulons réaliser un système qui permettra aux médecins de visualiser les cellules et d'avoir un outil d'aide à la décision en ce qui concerne la malignité ou non de ces cellules. La segmentation consiste en la séparation de la cellule en deux parties très significatives : noyau et cytoplasme. La classification exploite les résultats de la segmentation et son but est la séparation du tissu cellulaire composé de quatre couches de cellules (cas des frottis cervicaux), puis on établit une décision, à l'aide des réseaux de neurones.



Introduction Générale



Introduction Générale

Malgré des décennies de recherche et de progrès thérapeutiques, le cancer touche un individu sur trois et tue un quart des personnes qui en sont atteintes. Bien qu'il soit considéré comme une maladie des temps modernes, il existait déjà dans les siècles passés, mais à une échelle beaucoup plus réduite.

En 1900, le cancer ne causait que 4% des décès, on pourrait donc avoir l'impression qu'il y a eu, depuis le début du siècle, une véritable épidémie de cancers.

En fait, cette augmentation rapide de la mortalité par cancer est due à l'allongement de l'espérance de vie (45 ans en 1900, 78 ans en 1995)

Il faut rappeler que depuis 1939, la fréquence des cancers de l'estomac a été divisée par 4, celle de l'utérus par 2.

Par contre, celle des poumons a été multipliée par 4 ainsi que celle des mélanomes.

Ces évolutions spectaculaires reflètent les changements du mode de vie par :

- une alimentation plus saine (baisse des cancers de l'estomac)
- la pratique au frottis (baisse des cancers de l'utérus)
- l'augmentation de la consommation tabagique est à l'origine de l'accroissement du nombre des cancers du poumon.
- la mode des bains de soleil (élévation des mélanomes)

Heureusement que, pour lutter contre le cancer, il existe trois moyens :

- la prévention dite primaire qui s'attaque aux causes (tabac, alimentation trop riche...)
- Le dépistage
- Le traitement

En effet, le but du dépistage est de diagnostiquer et de traiter les cancers avant la dissémination métastatique qui va transformer une tumeur locale dont le traitement est aisé et efficace en une maladie généralisée dont la guérison reste exceptionnelle malgré les moyens modernes.

Introduction Générale

Néanmoins, on devrait pouvoir réduire la mortalité par cancer du col utérin, plus aisé à détecter à un stade précoce. Mais le dépistage ne sera efficace que s'il est doté de l'infrastructure nécessaire pour l'assurance de la qualité et du suivi des sujets dépistés.

Dans le but de palier à cette difficulté, on désire réaliser un système autonome pour le diagnostic du cancer, qui permettra aux médecins de visualiser les cellules et d'avoir un outil d'aide à la décision en ce qui concerne la malignité ou non des cellules.

Notre travail, s'intègre dans un projet de recherche dont le but est l'informatisation des opérations de routine des laboratoires d'anatomie cytologique, nous avons élaboré les différentes étapes de ce système.

Concernant le traitement des images des cellules, nous avons commencé par l'étape d'acquisition, réalisée à l'aide d'un microscope et une caméra numérique; suivie de l'étape de prétraitement qui consiste à éliminer les informations indésirables, en utilisant le filtrage et le rehaussement du contraste.

Pour la segmentation des images traitées, nous avons utilisé l'approche multifractale qui se base sur le critère d'homogénéité de l'image.

Dès que les objets seront bien déterminés, on établira une classification des cellules en leurs quatre couches essentielles (la superficielle, l'intermédiaire, la basale et la para-basale), par l'utilisation de l'algorithme de croissance de régions.

Ensuite, quand la classification sera réalisée et les critères de malignité calculés, on pourra déterminer son état (normal ou anormal)

Notre travail est divisé en 5 chapitres principaux:

Chapitre 1 : Nous présentons quelques notions, sur la cancérologie et la cytopathologie du col utérin en particulier pour mieux nous familiariser avec la structure du col.

Chapitre 2 : Nous allons définir les différentes étapes qui constituent un système de vision par ordinateur.

Chapitre 3 : Segmentation d'images nous abordons deux approche multi fractale et par contour actif . ce chapitre contiendra la définition de la méthode de segmentation d'images qui est une étape importante dans le processus de traitement d'images et qui permet, dans notre cas, la séparation de chaque cellule en son noyau et son cytoplasme.

Chapitre 4 : Dans ce chapitre, nous expliciterons les réseaux de neurones en donnant, leurs structures, leurs principales caractéristiques ainsi que les différents types existants et utilisés à ce jour.

Introduction Générale

Chapitre 5 : Ce chapitre contiendra notre travail effectué, ainsi que les résultats obtenus dans les différentes étapes. Et nous finirons en présentant notre application.

Et enfin, nous terminerons par une conclusion générale, contenant un résumé sur ce qui a été réalisé, ainsi que les éventuelles perspectives qui peuvent être apportées ultérieurement, sur notre travail pour une amélioration possible

Chapitre

1

Cancérologie et cytopathologie

Chapitre

1

Cancérologie et cytopathologie

Introduction :

Le cancer est une maladie très grave que l'on rencontre de plus en plus ces dernières décennies. De tous les cancéreux, une personne sur deux en meurt.

Le corps humain est très complexe. La cellule est sans aucun doute une des parties les plus compliquées du corps. Les chercheurs s'entendent même pour dire que la cellule humaine est en fait beaucoup plus complexe que toute la technologie qui a servi à conquérir la Lune. Chacune des quelques 50 000 milliards de cellules sont "programmées" pour mourir à un moment donné. Pourtant programmées pour mourir dès leur naissance, certaines cellules contredisent la nature. Elles se reproduisent alors, bien qu'elles soient dérégées. Les cellules se multiplient alors et créent une tumeur. Lorsque la tumeur se forme dans un sein, il y a des chances que celle-ci soit repérée. Elle peut alors être éliminée avant de produire des métastases. Une métastase se produit lorsqu'une cellule cancéreuse se sépare de la tumeur mère et se promène dans le corps et fonde une autre colonie. On assiste alors à un nouveau cancer. Généralement, lorsqu'il y a présence de métastases, le cancer devient très dangereux.

1. Qu'est ce que le cancer ? : [1]

Les cellules cancéreuses se propagent dans le corps à la suite d'une détérioration de la substance originelle des cellules. Cette détérioration peut être provoquée par un grand nombre de facteurs d'influence externes. Mais d'autres facteurs, tels que l'hérédité, les maladies préalables ou un âge très avancé, augmentent nettement le risque de souffrir d'un cancer.

Tout au long de notre vie, ces cellules anormales continuent de se développer. Mais en général, elles sont stoppées par nos défenses naturelles. Or, dans certains cas, les cellules anormales ne sont pas détruites. Elles se multiplient rapidement et forment des tumeurs.

Une tumeur située à un endroit précis du corps est rarement mortelle. Une petite tumeur ne cause encore aucune douleur. La diminution de la fonction normale de l'organe ne se produit souvent que lorsque la tumeur augmente en taille et qu'elle remplace le tissu environnant.

La maladie devient presque toujours mortelle lorsqu'elle se propage (métastases). Près de 90% des cancers avec issue mortelle sont imputés aux métastases, c'est-à-dire à la pénétration des cellules cancéreuses dans d'autres organes ou d'autres tissus.

Pour pénétrer plus facilement dans les organes, les cellules cancéreuses sécrètent des enzymes qui digèrent le tissu conjonctif environnant et leur ouvrent ainsi la voie vers d'autres organes.

2. La cellule cancéreuse : [2]

La cytologie est la science de la structure microscopique et moléculaire de la cellule.

La cellule cancéreuse possède de nombreuses caractéristiques la différenciant d'une cellule normale. Ces caractéristiques sont dues à la succession dans le clone tumoral d'anomalies moléculaires, d'origine génétique ("génétique somatique"), donnant au clone cellulaire tumoral un avantage sélectif et permettant sa croissance aux dépens des cellules normales. Ces caractéristiques sont de plusieurs ordres. On peut arbitrairement distinguer les anomalies morphologiques, les anomalies de la croissance et de la prolifération cellulaire, et les anomalies génétiques.

a) Les anomalies morphologiques :

L'appréciation des anomalies morphologiques cellulaires constitue un critère majeur du diagnostic des cancers (cytologie des cancers). Ces anomalies ne sont que la conséquence phénotypique de multiples anomalies moléculaires : anomalies des molécules d'adhérence, du cytosquelette, des organites intracellulaires, anomalies de la chromatine, anomalies chromosomiques (ploïdie).

L'étude morphologique peut porter sur des prélèvements cytologiques comportant des cellules indépendantes ou en amas, étalées sur une lame. Cette technique permet l'analyse

optimale des détails cytonucléaires. Mais l'examen histologique est obligatoire pour le diagnostic du cancer : il permet en particulier d'apprécier les anomalies architecturales qui échappent à la cytologie.

a-1. Anomalies du noyau :

Des anomalies encore mal connues de la conformation de la chromatine, les anomalies de la chromatine, en particulier la ploïdie, sont à l'origine de multiples anomalies nucléaires :

- Les noyaux sont globalement augmentés de taille. Cette augmentation provoque l'augmentation du rapport nucléo-cytoplasmique.
- Les noyaux sont de taille inégale d'une cellule à l'autre : anisocaryose.
- Les noyaux sont foncés, denses : hyperchromatisme.
- La chromatine est anormalement répartie, en mottes.
- La membrane nucléaire est épaissie, les contours nucléaires sont irréguliers.
- Une cellule tumorale peut avoir plusieurs noyaux : multinucléation.
- Les nucléoles sont multiples, volumineux, irréguliers.
- Le nombre de mitoses est augmenté. Ces mitoses sont fréquemment anormales (tri- ou tétrapolaires).

a-2. Anomalies du cytoplasme :

- Le cytoplasme est moins abondant, ce qui contribue aussi à l'augmentation du rapport nucléo-cytoplasmique.
- La taille des cellules est variable (anisocytose).
- Le cytoplasme est basophile (par augmentation de son contenu en acides nucléiques).
- Il peut contenir des inclusions (par exemple des amas de filaments intermédiaires) et des vacuoles (de lipides ou de glycogène).

Cependant, des cellules non cancéreuses peuvent présenter des anomalies identiques au cours des processus inflammatoires, des viroses, après irradiation... De plus, aucun de ces caractères morphologiques n'est constant; une cellule cancéreuse peut avoir une morphologie normale.

Ces anomalies cytologiques permettent l'identification de cellules cancéreuses (dépistage du cancer) au sein de frottis (frottis cervico-vaginaux), de produits d'expectoration (cytologie des crachats), de liquides biologiques (épanchement pleural, péricardique, péritonéal, liquide céphalo-rachidien), de ponctions de ganglions lymphatiques ou d'organes profonds (guidées par échographie ou par tomographie). Le cytodagnostic, cependant, n'a qu'une valeur d'orientation. Avant toute sanction thérapeutique, il doit être confirmé par l'étude histologique d'une biopsie.

Pour un même type de tumeur, l'intensité de ces anomalies est variable et permet de définir des grades de malignité des tumeurs en fonction de l'importance des altérations

morphologiques des cellules du clone tumoral, en particulier nucléaires. Cette "gradation" (grading) des tumeurs permet, pour un type tumoral donné, de définir des tumeurs de faible malignité (stade primaire) et des tumeurs de haute malignité (stade avancé). Dans de nombreux modèles tumoraux, un haut grade est corrélé à un mauvais pronostic ("grade histopronostique").

b) Les anomalies de la croissance et de la prolifération cellulaire :

L'homéostasie tissulaire est liée à un équilibre entre la prolifération (nombre de cellules en mitoses) et la mort des cellules (nombre de cellules en apoptose). La caractéristique principale des cellules tumorales est la mauvaise régulation de cet équilibre. Par exemple, les cellules tumorales peuvent croître en l'absence de stimuli mitogènes. Dans un organisme normal, les cellules issues d'une cellule souche se différencient grâce à l'expression séquentielle d'un ensemble de gènes permettant à la cellule d'exprimer son phénotype (phénotype mature) et d'assurer sa fonction, parfois hautement spécialisée (différenciation terminale). Au sein d'une tumeur, les cellules présentent des troubles de la différenciation. Elles peuvent être bien, moyennement ou mal différenciées. En l'absence de tout critère de différenciation, on dit que les cellules sont indifférenciées. La tumeur est dite alors anaplasique

c) Anomalies génétiques :

L'hypothèse selon laquelle le cancer aurait pour origine des anomalies du matériel génétique a été formulée dès le début du siècle par Boveri. Depuis les années 60, les progrès réalisés dans les techniques de culture des cellules tumorales ont permis l'obtention de mitoses in vitro et la réalisation de caryotypes tumoraux. Ces caryotypes ont montré la grande fréquence de remaniements chromosomiques dans les cellules tumorales. Il peut s'agir de remaniements numériques (monosomie, trisomie, tétrasomie) ou structuraux (translocations, isochromosomes, chromosomes en anneaux). On observe fréquemment des anomalies de la ploïdie ("aneuploïdie") où le génome tumoral ne contient pas 46 (2n) chromosomes mais 23 (haploïdie, n), ou 69 (triploïdie, 3n) ou 92 (tétraploïdie, 4n) chromosomes.

Initialement considérées comme des événements secondaires ("instabilité" génétique du processus tumoral), ces anomalies sont maintenant au coeur de la recherche oncoogénèse.

3. Etude spécifique de quelques cancers :

3.1 Introduction :

Des études de cancer relatif à quelques organes ont été entreprises à savoir les cas du poumon, du colon et du foie que nous avons préféré mettre en annexe [Annexe 1]. L'étude du cas du cancer du col utérin sera détaillée dans ce chapitre. Les raisons étant uniquement que pour cet organe nous disposons d'une base de données conséquente. Le but de notre projet étant de concevoir un système pouvant diagnostiquer un certain nombre de cancers issus de différents organes.

3.2 Cancer du col de l'utérus :

3.2.1 Un peu d'anatomie : [3]

L'utérus est un organe creux, situé dans le bassin de la femme, et dont la fonction est d'héberger le fœtus pendant la grossesse. L'utérus s'abouche dans le fond du vagin, et la portion d'utérus présente dans le vagin porte le nom de col de l'utérus. (Figure 1.1)

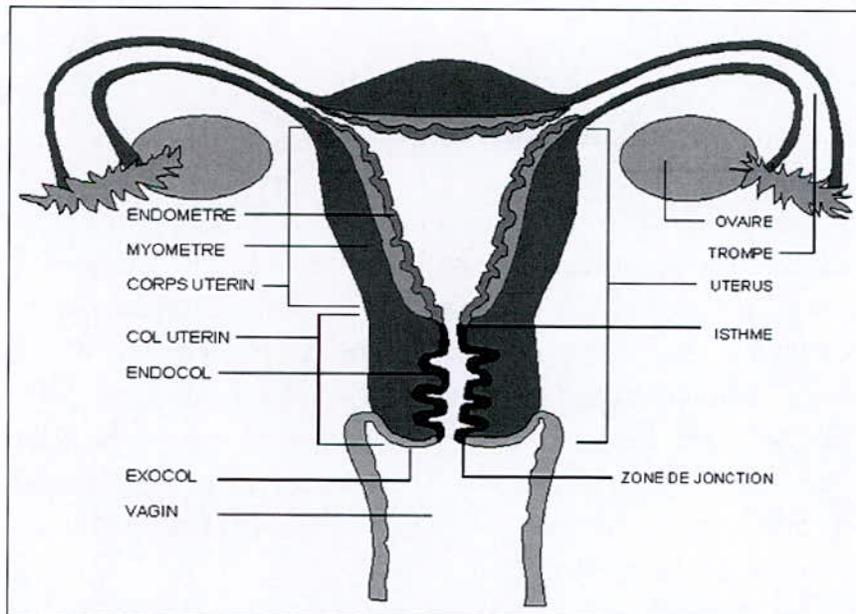


Figure 1.1 : Représentation de l'appareil génital féminin

3.2.2 Qu'est-ce que le cancer du col utérin? : [4]

Lorsqu'il ne faisait l'objet d'aucun dépistage, le cancer du col de l'utérus était le plus fréquent des cancers féminins. Ce cancer débute le plus souvent à la jonction entre la muqueuse qui recouvre le col de l'utérus, et celle qui tapisse l'intérieur de l'utérus (les deux muqueuses sont différentes, un peu comme la surface des lèvres diffère de la muqueuse de

l'intérieur de la bouche). Non dépisté, c'est un cancer très grave. Heureusement, le dépistage par frottis est extrêmement performant.

Les cellules qui recouvrent la surface du col utérin constituent le point d'origine de l'épithélioma, la forme la plus fréquente de cancer du col utérin. L'autre forme, plus rare, est appelée adénocarcinome et apparaît initialement dans les tissus glandulaires qui sécrètent la muqueuse.

3.2.3 Qu'est-ce que la dysplasie du col utérin? : [3]

Le terme dysplasie signifie que les cellules du col utérin ont subi des changements anormaux. Ces cellules anormales ne sont pas cancéreuses, mais elles peuvent le devenir si la dysplasie n'est pas traitée. La dysplasie du col utérin est également connue sous les noms suivants :

- lésions intra-épithéliales squameuses (LIS);
- néoplasie cervicale intra-épithéliale (CIN);
- changements précancéreux.

Toutes les femmes devraient subir régulièrement un test Papanicolou (frottis cervical) qui permet de déceler des changements au niveau du col utérin avant que les cellules anormales deviennent cancéreuses.

3.2.4 Qu'est ce qu'un frottis du col utérin ? : [1]

Le frottis du col de l'utérus est pratiqué au cours d'un examen gynécologique. A l'aide d'une petite spatule de bois, ou d'un très petit écouvillon, il gratte délicatement la surface du col, en pénétrant légèrement à l'intérieur de l'orifice du col (pour atteindre à coup sûr la zone de la jonction entre les deux muqueuses). Il récupère ainsi quelques cellules, qu'il étale sur une lame de verre. Ce prélèvement (ou frottis) est « fixé » sur la lame à l'aide d'un produit chimique en aérosol. Une fois fixé, le prélèvement est stable, et peut être examiné au microscope. Certains médecins préfèrent utiliser un milieu de conservation liquide (l'écouvillon est alors plongé dans un petit flacon spécial, l'étalement n'est pas fait par le médecin, mais le principe est le même).

3.3 Morphologie des cellules normales : [5]

La partie touchée par le cancer dans le col est la zone de jonction dont l'élément de base est l'épithélium.

L'épithélium est un ensemble de cellules (tissu) serrées les unes contre les autres recouvrant la surface de la muqueuse. Il est innervé mais ne comporte pas de vaisseaux sanguins. Il est composé de quatre couches de cellules qui sont :

- 1 – Une couche basale de cellules cubiques, avec des **mitoses**.
- 2 – Une couche para-basale de cellules **polyédriques**.
- 3 – Une couche intermédiaire de cellules qui tendent à s'aplatir.
- 4 – Une couche superficielle de cellules aplaties (pavimenteuses).

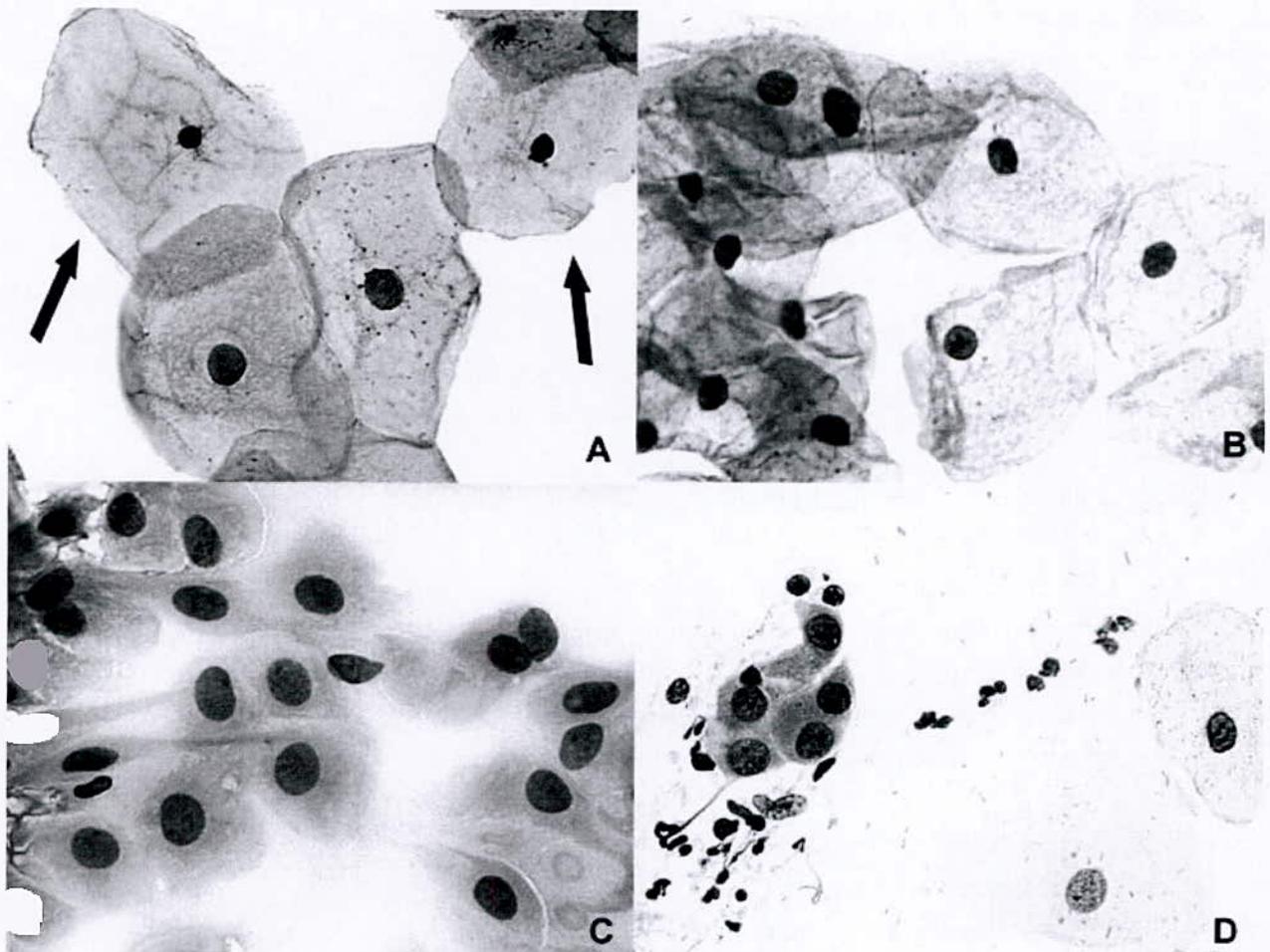


Figure 1.2 : Cellules malpighiennes - A : cellules superficielles (flèches) ; B : cellules intermédiaires ; C : cellules parabasales ; D : cellules métaplasiques.

3.3.1. La couche basale :

Elle est constituée de cellules nommées : (Figure 1.2.D) les basales à gros noyaux (BGN) mesurant environ 13 à 20 microns de diamètre, elles apparaissent rondes lorsqu'elles sont isolées les unes des autres, leur noyau est central relativement gros mesurant 9 à 10 microns de diamètre. Le cytoplasme est **cyanophile**, coloré en bleu vert avec les techniques cytologiques.

Ces cellules n'apparaissent pas normalement chez les femmes en activité génitale dans les frottis. On les rencontre quelques jours après la période pré-pubertaire, à la ménopause et chez les femmes en activité génitale en cas d'**érosion** épithéliale, inflammatoire, traumatique, etc. Leur **rapport noyau/cytoplasme** est compris entre 0.25 et 0.37.

3.3.2. La couche para-basale :

On les retrouve dans les frottis **atrophiques** et subatrophiques sous forme de cellules groupées et isolées (Figure 1.2.C). Elles contiennent la majorité des cellules exfoliées chez la femme ménopausée.

Elle est formée de cellules dites: les basales à petits noyaux (BPN), ces cellules sont plus volumineuses que les précédentes, elles sont rondes et mesurent 15 à 25 microns de diamètre, lorsqu'elles sont isolées. Ces cellules ont également un cytoplasme cyanophile et leur noyau central mesure environ 8 à 9 microns de diamètre.

Le rapport noyau/cytoplasme est compris entre 0.15 et 0.25.

3.3.3. La couche intermédiaire :

Ces cellules proviennent de la partie histologiquement appelée intermédiaire de l'épithélium malpighien située au dessus des cellules basales et aussi de la zone histologiquement qualifiée « pavimenteuse » (Figure 1.2.B).

On distingue deux sous variétés cellulaires dans ce groupe : les vraies intermédiaires et les grandes intermédiaires.

Les petites cellules intermédiaires (vraies intermédiaires) : Ces cellules sur les frottis, possèdent un cytoplasme plus ou moins ovale avec un grand axe mesurant de 20 à 35 microns. Le cytoplasme est toujours cyanophile avec parfois une zone **périnucléaire** plus ou moins importante, d'aspect plus clair, parfois jaunâtre, due à la présence de **glycogène**. En cas de grande richesse en glycogène, on observe une mince couronne exo-plasmique, cyanophile, plus foncée ; l'ensemble caractérise les cellules « naviculaires ». Le noyau est excentré, toujours bien structuré ; il mesure de 6 à 9 microns de diamètre.

Les grandes cellules intermédiaires : sont des éléments mesurant 35 à 60 microns, leur aspect est polygonal, elles sont généralement bien étalées sur la préparation. Ce sont réellement des cellules pavimenteuses, c'est-à-dire aplaties. Elles proviennent des **assises** histologiquement dites « pavimenteuses », leur noyau n'est pas **pycnotique**, il est bien structuré et mesure de 6 à 8 microns de diamètre.

Leur abondant cytoplasme est cyanophile, mais il peut apparaître éosinophile, dans certains cas, notamment lors d'un processus parakératosique. Ces derniers éléments constituent ce que l'on appelle la « fausse éosinophile » qui n'a aucune signification hormonale.

On rencontre des cellules intermédiaires à la période pré-pubertaire, à la ménopause et chez les femmes en activité génitale, surtout dans la deuxième partie du cycle. Pendant la gestation, leur nombre augmente, elles prennent un aspect allongé et leurs bords s'épaississent

Leur rapport nucléo-cytoplasmique est compris entre 0.04 et 0.15.

3.3.4. La couche superficielle :

Ce sont de grands éléments cellulaires. Leur masse cytoplasmique mesure 35 à 60 microns, quelquefois plus (Figure 1.2.A).

Ce sont des cellules pavimenteuses provenant des assises les plus haut situées de l'épithélium malpighien. Leur noyau est petit, central, très centré, sans structure chromatiniennne visible et prennent fortement les colorants nucléaires ; c'est un noyau pycnotique qui correspond à une dégénérescence cellulaire. Le diamètre, suivant la définition de Férin, doit mesurer 5 microns au maximum. Le cytoplasme de ces cellules superficielles peut être cyanophile ou **éosinophile**. Lorsqu'elles sont cyanophiles ces cellules ne se distinguent des grandes intermédiaires que par leur noyau pycnotique.

Leur rapport noyau/cytoplasme est compris entre 0.01 et 0.04. On les rencontre chez les femmes en activité génitale.

4. Critères de malignité : [3]

Le diagnostic de la malignité se fonde sur l'ensemble des particularités morphologiques atypiques de toutes les cellules du frottis.

4.1 Perturbation du rapport nucléo-cytoplasmique en faveur du noyau :

Les noyaux des cellules tumorales sont habituellement plus gros que ceux des cellules normales, mais sans que le cytoplasme présente une variation comparable. Les modifications du rapport nucléo-cytoplasmique constituent un des signes les plus valables de la malignité.

4.2 Anisonucléose (Anisocariose) :

La taille des noyaux dans les cellules tumorales isolées peut être très variable de l'un à l'autre. A cela s'oppose l'uniformité des noyaux des cellules normales parvenues au même degré de différenciation.

La constatation de noyaux de forte taille, aux côtés d'autres petits ou moyens, est souvent un indice de la provenance des cellules d'un tissu malin. Cette anisonucléose apparaît clairement si les cellules malignes se présentent en groupe ou en amas.

4.3 Difformité nucléaire (Polymorphisme nucléaire) :

Indépendamment des variations de la taille, il existe une diversité marquée de la forme des noyaux des cellules tumorales. Ils sont plus ou moins elliptiques, encochés, bizarrement déformés, et souvent en position excentrée. L'aspect fusiforme ou étiré est relativement fréquent. Le polymorphisme nucléaire apparaît mieux lorsque les cellules malignes sont groupées.

4.4 Hyperchromatisme nucléaire :

Les modifications de la structure nucléaire normale appartiennent toujours à l'image des cellules tumorales. Les noyaux hyperchromatiques apparaissent déjà par leur coloration sombre et massive. Du fait de la disparition de leur texture, ils se montrent entièrement bruns, sombres et noirs ou bien avec des amas irréguliers de chromatine. Le contour nucléaire est souvent nettement tranché et épais.

Conclusion :

Les mécanismes de développement des tumeurs sont complexes et mal compris. En général, un cancer ne peut pas être imputé à un seul agent ou à une seule cause. Il semble s'agir d'un processus aléatoire, dont la probabilité augmente par la contribution combinée de plusieurs facteurs. Nous avons constaté qu'il était indispensable de passer par cette étude, considérée comme un point de départ pour une bonne maîtrise de l'élément de base de notre sujet qui est le col utérin et sa tumoralité

Nous avons vu qu'il existait plusieurs critères de malignité permettant la détection des cellules malignes que nous allons considérer par la suite. Le rapport nucléo-cytoplasmique permettra particulièrement de faire une classification des cellules en leurs quatre couches.

Cependant, nous verrons dans le prochain chapitre, les notions et les concepts de base du traitement et de l'amélioration de l'image d'un système de vision par ordinateur.

Chapitre

2

Vision par ordinateur

Chapitre

2

Vision par ordinateur

Introduction

Une chaîne de traitement d'images se compose de trois parties fonctionnellement distinctes [6] :

- l'acquisition et la numérisation
- le traitement
- l'analyse et la décision : Reconnaissance

Notre travail qui constitue un système automatique suit le schéma fondamental donnant les différentes étapes d'un système de vision par ordinateur (Figure 2.1)

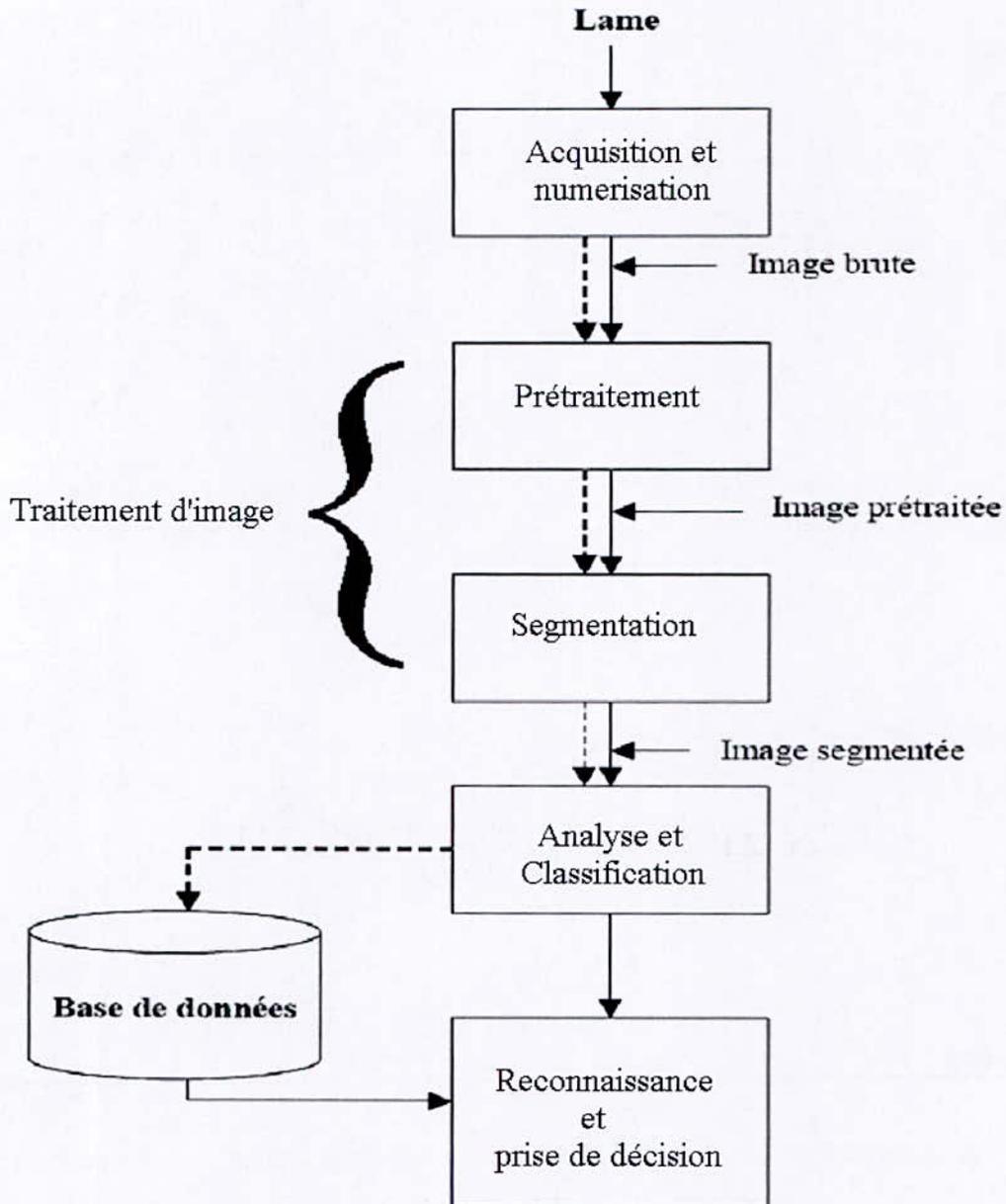


Figure 2.1 : Système de vision par ordinateur

1. L'acquisition : [7]

L'acquisition d'images constitue un des maillons essentiels de toute chaîne de conception et de production d'images. Pour pouvoir manipuler une image sur un système informatique, il est avant tout nécessaire de lui faire subir une transformation qui la rendra lisible et manipulable par ce système. Le passage de cet objet externe (l'image d'origine) à sa représentation interne (dans l'unité de traitement) se fait grâce à une procédure de

numérisation. Ces systèmes de saisie, dénommés optiques, peuvent être classés en deux catégories principales : les caméras numériques et les scanners.

A ce niveau, notons que le principe utilisé par le scanner est de plus en plus adapté aux domaines professionnels utilisant le traitement de l'image comme la télédétection, les arts graphiques, la médecine, etc.

Le développement technologique a permis l'apparition de nouveaux périphériques d'acquisition appelés cartes d'acquisition, qui fonctionnent à l'instar des caméras vidéo, grâce à un capteur C.C.D. (Charge Coupled Device). La carte d'acquisition reçoit les images de la camera, de la T.V. ou du scanner afin de les convertir en informations binaires qui seront stockées dans un fichier.

1.1 Le scanner :

Pour afficher une image sur l'écran d'un ordinateur, elle doit posséder un format numérique. Plusieurs méthodes permettent de numériser une image. L'utilisation du scanner est l'une de ces méthodes, il permet de « copier » l'image quelque soit son support (avec un logiciel de reconnaissance). Les types de scanners généralement rencontrés sont les scanners à plat et les scanners à rouleau.

1.2 La caméra numérique :

Ce type d'appareil n'a pas besoin de pellicule; les images sont stockées sous forme numérique dans une disquette, un disque dur ou même une puce. La qualité de l'image dépend de l'appareil utilisé.

Son principal avantage est le développement des photos instantané. En effet, elle peut être reliée à un PC par un câble ce qui permet de récupérer l'image de façon facile et pratique.

Un appareil numérique possède un viseur à cristaux liquides (mini écran à cristaux liquides affichant l'image). Ainsi on peut visualiser parfaitement ce qu'on est sur le point de photographier et effectuer les réglages nécessaires. Une fois la photo prise, le viseur permet de la consulter, de l'effacer ou de l'archiver.

1.3 Le microscope :

Le microscope est essentiellement constitué de trois systèmes optiques calés sur un même axe :

L'objectif : Système convergent de très faible distance focale, qui donne de l'objet une image réelle très agrandie.

L'oculaire : Système agissant à la façon d'une loupe, centré sur la première image, dont il donne à son tour une image virtuelle encore plus agrandie au niveau de l'œil de l'observateur.

Un condensateur : Destiné à concentrer, sur l'objet à observer, la lumière fournie par une source lumineuse. Cet objet est constitué par une couche mince de substance ou de cellules, étalée sur une lame de verre, ou une gouttelette liquide maintenue sur la lame par une très fine lamelle à faces parallèles et qu'on observe par transparence.

Le microscope peut être équipé d'une caméra numérique, qui nous permettra de prendre des photos à partir des lames et de les transférer vers notre ordinateur pour les traiter (Figure 2.2).



Figure 2.2 microscope utilisé

2. Le traitement d'image :

2.1 Le prétraitement : [8]

L'étape de prétraitement a pour but de faciliter la segmentation d'images. Elle doit lisser le bruit tout en sauvegardant les structures et les transitions entre les différentes régions de l'image.

L'image brute est une structure simple dans son organisation mais complexe dans son contenu. Ceci est dû à plusieurs facteurs, d'une part la grande quantité d'informations qui y réside, d'autre part le nombre important de processus indépendants de l'utilisateur qui tend à y introduire des distorsions indésirables, donc à la dégrader.

Cette étape consiste à :

- Rehausser le contraste.
- Filtrer l'image (diminuer les bruits).
- Effacer toutes les parties indésirables manuellement.

2.2. La segmentation : [9]

La phase prétraitement a consisté en une restauration et amélioration de l'image numérisée soit en une préparation de l'image à l'analyse.

La phase clé est celle de l'extraction des primitives, pour la description des aspects intéressants de l'image. Cette extraction nécessite une segmentation ou une division de l'image en ces éléments constitutants et pour ceci deux approches sont possibles :

- La première approche part du fait que dans la plupart des cas, les lignes de contours ou *frontière* apportent une information substantielle sur un objet. Ainsi plusieurs méthodes de détection de contour ont vu le jour.
- La seconde approche part de la constatation suivante : les contours seuls ne suffisent pas à caractériser un objet dans plusieurs cas, il est donc souvent nécessaire d'extraire non plus des contours, mais des régions entières ; une région est un ensemble connexe de points d'image (pixels) ayant des propriétés communes (intensité, texture,...) qui les différencient des pixels des régions voisines.

2.2.1 Approche frontière :

Dans une image numérique, les contours se situent entre les pixels appartenant à des régions ayant des intensités moyennes différentes ; il s'agit de contours de type « saut d'amplitude ». Un contour peut également correspondre à une variation locale d'intensité présentant un maximum ou un minimum. Donc les contours sont assimilés aux points de fort gradient ou de dérivée seconde nulle. (Point d'inflexion)

2.2.2 Approche région :

La segmentation en régions homogènes vise à segmenter l'image en se basant sur les propriétés intrinsèques de la région, le choix de ces propriétés détermine ce que l'on appelle le critère de la segmentation, elle consiste à rechercher et séparer des ensembles de pixels qui possèdent des propriétés communes. L'une des méthodes utilisée pour la segmentation en approche région est la segmentation multifractale.

3. La reconnaissance : [10]

A partir de l'image segmentée, il faudra reconnaître chaque objet. Et faire une classification de chaque objets suivant ses caractéristiques spécifiques.

Nous allons nous baser dans cette étape sur l'intelligence artificielle, par l'utilisation des réseaux de neurones, pour prendre la bonne décision. Le réseau de neurones nécessite un apprentissage, se basant sur les images dites prototypes, qui constitueront la base de données.

4. Les domaines d'application du traitement d'images : [11]

De nos jours, les ordinateurs accélèrent le processus de traitement d'un grand nombre de calculs et d'applications. Avec la parole, l'image constitue l'un des moyens les plus importants qu'utilise l'homme pour communiquer avec autrui. C'est un moyen de communication universel dont la richesse du contenu permet aux êtres humains de tout âge et de toute culture de se comprendre.

C'est aussi le moyen le plus efficace pour communiquer, chacun peut analyser l'image à sa manière, pour en dégager une impression et d'en extraire des informations précises.

De ce fait, le traitement d'images est l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes.

Les applications du traitement de l'image sont à leur tour très variées:

- Compression d'images.
- Amélioration d'images en vue d'un meilleur rendu visuel.
- Comptage et tri cellulaire en imagerie médicale.
- Identification de personnes : reconnaissance des empreintes digitales, des visages et de l'iris.
- Identification de documents (billets de banques) et reconnaissance de caractères.
- Télédétection.
- Indexation des images : organiser automatiquement de grandes bases d'images en fonction de leur contenu.

Conclusion :

Dans ce chapitre, nous avons mis en évidence les différentes étapes d'un système de vision par ordinateur. On remarque que l'étape primordiale est l'acquisition car une mauvaise acquisition donnera une image de basse qualité qui sera difficile à traiter, on passera à l'élimination des zones bruitées, la séparation des différents objets et finalement leurs classifications suivant les caractéristiques spécifiques.

Chapitre

3

Segmentation

Chapitre

3

Segmentation

Introduction

:

Une image est généralement composée de zones à niveaux de gris différents, correspondant aux différents objets de la scène. Il nous faut donc trouver un moyen de repérer ces derniers afin de pouvoir passer leurs caractéristiques à un éventuel processus de reconnaissance des formes. Pour cela, nous devons faire passer l'image par une étape très importante qui est la segmentation.

1. Définition de la segmentation : [9]

La segmentation est le processus de subdivision d'une image en parties qui la constituent. Elle consiste à effectuer une répartition de l'image en sous-ensembles R_i (régions homogènes) par rapport à un ou plusieurs critères. Chacun d'eux ayant les mêmes propriétés tel que :

- $\forall i, R_i \neq \Phi.$
- $\forall i, j : i \neq j, R_i \cap R_j = \Phi. \quad (1)$
- $I = \cup R_i / I : \text{image.}$

Soit R l'image entière et P le prédicat choisi. On cherche une partition de R , en N régions, telle que :

$$\begin{aligned}
 R_i \cap R_j &= \phi \quad \forall i, j, i \neq j \\
 P(R_i \cup R_j) &= \delta_{i,j} \\
 P(R_i) &= 1 \quad \forall i = 1, \dots, N
 \end{aligned}
 \quad (2)$$

La segmentation est l'élément le plus important dans le processus de traitement d'image, car c'est à ce moment précis que les objets ou autres entités sont extraits de l'image pour le reste du processus, à savoir l'extraction des caractéristiques et la reconnaissance.

Les algorithmes de segmentation sont, en général, basés sur une des deux propriétés des valeurs de niveau de gris : la discontinuité ou la similarité. Dans la première catégorie, on scinde l'image en se basant sur le changement de niveau de gris. On détecte, alors, les points isolés, les lignes et les contours. Dans la seconde catégorie, on se base sur la détection des régions homogènes.

L'approche Multifractale utilisée dans cette partie appartient à la deuxième catégorie.

2. Approche multifractale :

2.1 Définition des fractales : [12]

Apparues au XIXe siècle, les fractales n'acquirent un statut à part entière que dans les années 1970, grâce au mathématicien français Benoît Mandelbrot qui en fit l'objet d'une nouvelle discipline mathématique : la géométrie fractale. Le terme "fractale" (du latin fractus = brisé) fut d'ailleurs inventé par Mandelbrot pour désigner un type d'objets dont l'irrégularité les distingue des figures géométriques euclidiennes telles que la droite ou le cercle.

Une fractale est un objet géométrique qui possède une complexité intrinsèque, une irrégularité fondamentale qui se manifeste à toutes les échelles d'observation et surtout une propriété d'autosimilarité. Les objets fractals peuvent se définir comme des structures obtenues par l'itération d'un algorithme géométrique sur une figure. Pour construire des objets fractals, nous débutons avec un objet graphique quelconque (ligne, triangle, carré, cube, etc...). Par la suite, on définit une opération, ou une série d'opérations, qui ajouteront un élément de complexité à l'objet initial. Nous appliquons à l'infini les transformations choisies à l'objet de départ.

On définit donc une fractale comme une figure géométrique ou un objet naturel qui combine les caractéristiques suivantes :

- Ses parties ont la même structure que le tout, à ceci près qu'elles le sont à une échelle différente et peuvent être légèrement déformées
- Sa forme est extrêmement irrégulière ou fragmentée et le reste à toutes les échelles
- Elle contient des éléments discernables dans une large gamme d'échelle.

2.2. Dimension fractale : [13]

En géométrie euclidienne, les figures ont une dimension entière : 0 pour un point, 1 pour une ligne droite, 2 pour une courbe et 3 pour un volume. En revanche, la dimension d'une fractale peut prendre des valeurs qui ne sont pas des nombres entiers : la dimension fractale constitue une généralisation de la notion de dimension utilisée en géométrie euclidienne.

En géométrie ordinaire, on dit qu'une courbe est un espace à 1 dimension parce qu'on sait mesurer les longueurs le long de la courbe et qu'on peut repérer tout point de la courbe par un seul nombre, son abscisse, c'est à dire sa distance depuis le point de la courbe pris comme origine.

Pour les fractales, il est impossible de désigner un point initial d'une fractale, puisqu'on ne peut pas mesurer de longueur le long de la fractale. Une ligne fractale dessinée dans un plan appartient bien à ce plan, de dimension 2, mais on ne peut pas dire qu'elle soit de dimension 1.

En effet, les fractales ont des dimensions fractionnées ; par contre, on dispose du nombre D , nombre réel, qui est proche de 1 quand la fractale est "plutôt" lisse, qui augmente quand la fractale devient plus accidentée, et qui tend vers 2 quand la fractale remplit la portion de plan qui lui sert de support.

2.3. Différents types de fractales : [14]

2.3.1. Fractales aléatoires :

On définit des fractales aléatoires comme des fractales pour lesquelles le choix de l'opération appliquée à chaque itération suit une loi de probabilité. À partir d'un carré, on peut par exemple conserver deux quarts quelconques de ses quarts et obtenir ainsi une fractale de dimension 2. On peut encore généraliser la construction en variant aussi à chaque pas de l'itération le nombre de quarts conservés selon une loi de probabilité. On obtient alors une fractale aléatoire hétérogène; il semble que ce soit ces structures qu'on rencontre dans la nature, par exemple pour la distribution des galaxies.

2.3.2. Fractales déterministes :

Les fractales déterministes regroupent certainement les plus complexes et les plus spectaculaires fractales : elles combinent la science des mathématiques et la beauté de l'art.

Les exemples les plus fondamentaux sont l'ensemble de Mandelbrot et les ensembles de Julia (=sous-ensembles de l'ensemble de Mandelbrot).

Les fractales déterministes ne comprennent pas seulement les fractales obtenues par l'itération d'une équation en nombre complexe, mais aussi un autre sous catégorie appelée IFS.

Pour les fractales déterministes le processus de formation est beaucoup plus compliqué que celui des objets fractals. Elles sont développées par l'itération de polynômes en nombres complexes. Les ordinateurs sont indispensables à leur création. Il faut comprendre qu'il est difficile de dessiner des fractales à la main car nous sommes dans l'impossibilité de tracer tous les détails, à l'infini, avec précision. Les ordinateurs sont nécessaire pour de telles opérations car ils nous permettent de visualiser et de créer des fractales impressionnantes.



Figure 3.1 : Modèle de fractale IFS

2.3.3. Fractales à homothétie interne :

Les fractales à homothétie interne présentent en outre une propriété particulière : chacune de leur partie reproduit leur totalité. Elles présentent ainsi un aspect tout à fait identique quelle que soit l'échelle considérée : on parle également d'invariance par dilatation. La génération des fractales à homothétie interne est souvent basée sur un algorithme simple. Ainsi, si l'on examine un flocon de neige naturel à différentes échelles, on observe une structure à peu près identique. Cette propriété se traduit par une dimension fractale strictement comprise entre 1 et 2.

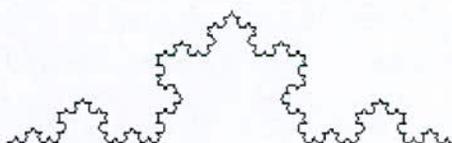


Figure 3.2 : le "flocon de neige" de Von Kock

2.3.4. Multifractales :

Une multifractale est une réunion de fractales de différentes dimensions d_F , munie d'un spectre continu ou discret $f(d_F)$ qui spécifie la mesure de chacune des fractales (ou pour être plus simple, une multifractale est une réunion de fractales de dimensions différentes ou égales).

2.4. Formalisme multifractal : [15]

Le terme multifractal et la notion de spectre de singularités ont été introduits pour la première fois par Fisch et Parisi.

a) L'exposant de singularité :

Soit μ une mesure quelconque et $\text{supp}\mu$ son support (Nous considérons le support d'une mesure comme étant l'ensemble sur lequel la mesure est concentrée).

On appelle exposant de singularité au point $x_0 \in \text{supp}\mu$, la limite :

$$\alpha(x_0) = \lim_{\epsilon \rightarrow 0} (\ln \mu(\beta_{x_0}(\epsilon)) / \ln \epsilon) \quad (3)$$

Où : $\beta_{x_0}(\epsilon)$ désigne une boule centrée en x_0 et de taille ϵ .

L'exposant de singularité en un point rend compte du degré local de régularité de la mesure considérée. La définition de $\alpha(x_0)$ indique simplement que de manière générale, on pourra écrire :

$$\mu(\beta_{x_0}(\epsilon)) \approx C \epsilon^{\alpha(x_0)} \quad \text{avec } \alpha \geq 0, \text{ quelconque} \quad (4)$$

b) Le spectre de singularité :

Le spectre $f(\alpha)$ des singularités associées à la mesure μ est la fonction qui, à tout α , associe la dimension fractale de l'ensemble de points x_0 tel que $\alpha(x_0) = \alpha$:

$$f(\alpha) = d_F(\{x_0 \in \text{Supp}\mu / \alpha(x_0) = \alpha\}) \quad (5)$$

Le spectre $f(\alpha)$ des singularités peut être introduit comme une quantité qui donne une caractérisation relativement précise du degré de régularité et d'homogénéité d'une mesure fractale.

c) Algorithme multifractal pour la segmentation:[16]

Dans cette méthode, nous adaptons le paramètre multifractal α à la segmentation des images. Pour cela, nous étudions le comportement des exposants de singularité locaux, c'est-à-dire calculés en chaque point de l'image, nous obtenons alors une image des exposants de singularité.

L'algorithme de segmentation est le suivant :

- Calculer les exposants de singularité en chaque point de l'image, et trouver la singularité minimale α_{\min} et la singularité maximale α_{\max} .
- Diviser le spectre de singularité en plusieurs parties égales, c'est-à-dire diviser l'intervalle $[\alpha_{\min}, \alpha_{\max}]$ en N sous intervalles.
- Pour chaque sous intervalle, trouver l'ensemble des points de l'image ayant l'exposant de singularité appartenant à l'intervalle, ce qui permet d'avoir N sous images, chacune correspondant à un intervalle iso-alpha. Sachant que pour un

Chapitre 3: Segmentation

intervalle $[\alpha_1, \alpha_2]$ donné, les pixels retenus seront ceux dont le niveau de gris sera compris dans l'intervalle $[p(1/L)^{\alpha_2}, p(1/L)^{\alpha_1}]$;

- Pour chaque image d'un intervalle iso-alpha, les points en noir sont ceux dont la valeur α est dans l'intervalle souhaité.

Dans cet algorithme, le calcul de l'exposant de singularité s'effectue comme suit, selon la formule:

$$m(l(x,y)) \approx (1/r)^{-\alpha(x,y)} \quad \text{quand } r \rightarrow 0, \quad (6)$$

Avec :

- $m(x,y)$: est la mesure associée au point (x,y) de niveau de gris $l(x,y)$.
- r : est la largeur de boîte entourant le point (x,y) .
- α : est l'exposant de Hölder au point (x,y) .

Ce qui nous permet d'écrire l'équation suivante :

$$\frac{p(x,y)}{p} = (r/L)^{\alpha(x,y)} \quad (7)$$

soit à l'échelle du pixel :

$$p(x,y) = p(1/L)^{\alpha(x,y)} \quad (8)$$

avec :

p : la somme des niveaux de gris de toute l'image.

L : la taille de l'image.

$P(x,y)$: le niveau de gris associé au pixel (x,y) .

3. Contours actifs :

Il y a eu divers efforts pour l'analyse d'images, en particulier de scènes avec des objets simples, par localisation des bords et des lignes dans l'image. Les bords doivent être rassemblés en segments de ligne, et ces segments de ligne doivent être regroupés pour former les contours d'objets. Généralement, de tels systèmes nécessitent un modèle du monde bien défini, permettant de regrouper ces segments en objets.

Les modèles classiques de détection de contour procèdent essentiellement par application d'un opérateur local de différentiation précédé d'un opérateur de filtrage.

Cette opération de filtrage peut soit être associée à un opérateur dont la réponse en fréquence est figée (comme dans le cas du masque de Sobel ou Prewitt [Annexe 2]) soit être associée à une opération de filtrage adaptable à l'aide d'un paramètre d'échelle comme c'est le cas avec les approches de Canny et dérivées. Ces opérateurs sont qualifiés d'opérateurs locaux ; car ils ne prennent pas en compte l'information globale sur le contour.

3.1. Détection de contours : approches globales

On présente dans cette partie du chapitre un certain nombre de modèles globaux de détecteurs de contours :[17]

3.1.1 Les détecteurs de gradient par filtrage :

Ces détecteurs reposent tous sur une recherche d'un extremum de la dérivée première (ou d'un passage par zéro d'une dérivée seconde), celle-ci étant calculée de diverses manières, mais généralement par un filtrage passe haut précédé d'un léger filtrage passe-bas pour s'affranchir des bruits.

3.1.2 Les détecteurs de gradient par masques :

A côté de ces approches très inspirées du traitement du signal, des filtres de dérivation plus empiriques ont été proposés à partir d'estimateurs locaux de l'image ou de ses dérivées. Ces estimés sont obtenus à l'aide de masques appliqués sur des fenêtres de 2 x 2 pixels ou 3 x 3 pixels (exceptionnellement, en cas d'images très bruitées, sur des fenêtres plus grandes) . On note que la somme des coefficients de ces filtres est nulle (fonction de transfert à la fréquence 0), et que les coefficients sont anti-symétriques.

3.1.3 Les détecteurs de passage par zéro du laplacien :

Cette méthode utilise le fait que le passage par zéro du laplacien permet de bien mettre en évidence les extremums de la dérivée. Le passage par zéro du laplacien correspond en effet bien au maximum du gradient dans la direction du gradient. Ces

méthodes tirent en outre profit du fait que les zéros de la dérivée seconde constituent un réseau de lignes fermées. Mais l'estimation de la dérivée seconde étant très sensible aux bruits, il convient de filtrer très fortement l'image avant d'en mesurer le laplacien.

3.1.4 Les détecteurs par masquage adapté :

Dans ces techniques, on recherche en chaque point de l'image la présence d'une configuration conforme à un gabarit appartenant à un dictionnaire de contours. Pour cela, on définit une distance entre un contour type et une fenêtre de l'image.

L'approche par masque adapté, permet de connaître aisément la direction du contour.

Dans une version améliorée de la détection des contours, un filtre de détection se compose de deux estimateurs de dérivées, l'un selon O_x , l'autre selon O_y . L'un de ces détecteurs se compose du produit de 2 fonctions :

- selon O_y c'est une fonction passe-bas, symétrique, dont l'étendue est fonction de l'importance du bruit et de la distance que l'on souhaite respecter entre deux contours ;
- selon O_x c'est une fonction passe-haut (anti-symétrique), souvent égale à la dérivée de la précédente.

3.2. Les contours actifs = les *snakes* :

Un contour actif est une courbe du plan qui, à partir d'un état initial donné, a la propriété de se déformer [18]. La technique des contours actifs consiste à déformer cette courbe afin de l'adapter au contour apparent d'un objet (dû à sa perception bidimensionnelle). Le contour actif se déforme, en accord avec l'optimisation d'un critère, pour converger vers un état final, qui sera la segmentation de l'image. L'image est alors séparée en deux zones correspondant à l'objet à l'intérieur du contour et le fond à l'extérieur (Figure 3.3).

L'intérêt majeur de cette approche est qu'elle exploite un contour déjà fermé et la connexité n'est plus à vérifier lorsque l'opération est terminée, contrairement aux détecteurs de bord traditionnels. C'est une opération qui permet d'obtenir les coordonnées du contour d'un objet, sa surface et sa position, contrairement aux filtres classiques de détection de contours qui sont des opérations traitant l'image globalement.

3.2.1 Une approche physique des contours :

Une approche très différente des méthodes antérieures de détection de contours a été proposée en 1987 par Kass, Witkins et Terzopoulos, appelée contours actifs ou *snakes*. Il s'agit d'une méthode semi-interactive dans laquelle l'opérateur place dans l'image, au voisinage de la forme à détecter, une ligne initiale de contour.

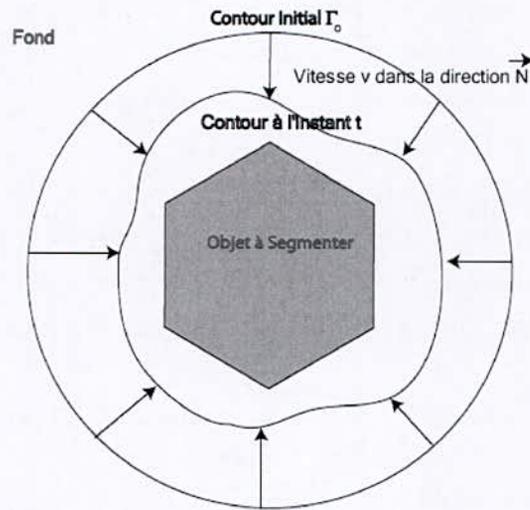


Figure 3.3 : Représentation du contour actif

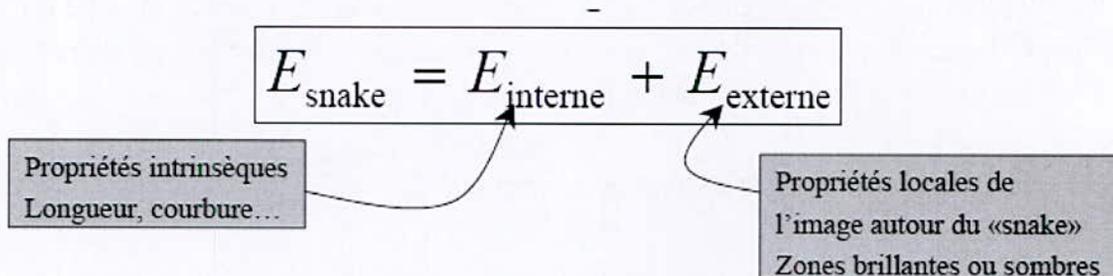
Cette ligne sera amenée à se déformer sous l'action de plusieurs forces :

- Une énergie propre, assimilée à l'énergie mécanique de torsion d'une ligne matérielle.
- Une énergie potentielle imposée par l'image qui vise à plaquer la courbe sur les contours.
- Une énergie externe, introduite par l'utilisateur pour traduire les contraintes spécifiques du problème qu'il se pose.

Sous ces énergies, le contour actif va évoluer pour rechercher la position d'énergie minimale, qui sera ainsi un compromis entre les diverses contraintes du problème.

3.2.2 Mise en équations du principe du contour actif :

Pour établir les expressions, on considère que notre snake est composé de n points V_i que l'on va définir de façon paramétrique en fonction de l'abscisse curviligne s . [18]



- **L'énergie interne :**

L'énergie interne permet de contrôler les propriétés du snake et régularise la solution au sens des problèmes mal posés. Le snake est considéré comme un matériau déformable

caractérisé par ses facultés d'allongement et de courbure. Elle va dépendre uniquement de la forme du snake. C'est une contrainte de régularité sur la courbe. Elle s'écrit:

$$E_{int} = (a(s)) \cdot ||(V_s(s))|| + b(s) \cdot ||(V_{ss}(s))||. \quad (9)$$

a et b sont les poids respectifs accordés à la dérivée première V_s et à la dérivée seconde V_{ss} . On ajustera a et b pour trouver un intermédiaire entre un contour souple mais pas forcément très régulier (et qui pourra alors se caler sur des coins et des angles aigus) et un contour très régulier qui suivra bien le contour sans s'accrocher sur des petites aspérités.

- **L'énergie potentielle liée à l'image :**

Elle caractérise les éléments vers lesquels on veut attirer le snake sur l'image. Pour notre problème, puisqu'il s'agit de détecter des contours, on va s'intéresser aux lignes de fort gradient:

$$E_{image} = -\text{gradient}(i). \quad (10)$$

Pour certains problèmes, on peut choisir d'attirer le snake vers des zones présentant d'autres caractéristiques: vers les zones les plus sombres par exemple.

- **L'énergie de "contrainte" :**

Elle est définie par l'utilisateur selon les spécificités du problème. On peut par exemple imposer une distance minimale ou maximale entre deux points consécutifs du contour actif. Elle va s'exprimer en fonction de $V(s)$ uniquement:

$$E_{cont} = E_{cont}(s) \quad (11)$$

3.2.3 Energie totale du contour actif :

L'énergie totale en un point du snake $V(s)$ va donc s'écrire:

$$E_{totale}(V(s), V_s(s), V_{ss}(s)) = E_{int}(V, V_s, V_{ss}) + E_{image}(V) + E_{cont}(V). \quad (12)$$

En minimisant cette somme pondérée, on parvient à une solution offrant un bon compromis entre les contraintes. [19]

Nous avons affaire à un nombre de points V_i fini. On exprime directement E_i : énergie au point V_i en fonction de tous les points V du snake en utilisant notamment:

$$V_s(s) = V_i - V_{i-1} \text{ et } V_{ss}(s) = V_{i+1} - 2V_i + V_{i-1} \quad (13)$$

Ce qui nous permet de discrétiser la formule de E_i : énergie au point V_i sous la forme:

$$E_i \text{ totale}(V_{i-1}, V_i, V_{i+1}) = E \text{ int}(V_{i-1}, V_i, V_{i+1}) + E \text{ image}(V_i) + E \text{ cont}(V_i). \quad (14)$$

Et donc l'énergie totale du snake vaut:

$$E \text{ totale}(V_{i-1}, V_i, V_{i+1}) = \text{somme}(E_i) = \text{somme}(E_i \text{ int}(V_{i-1}, V_i, V_{i+1}) + E_i \text{ image}(V_i) + E_i \text{ cont}(V_i)). \quad (15)$$

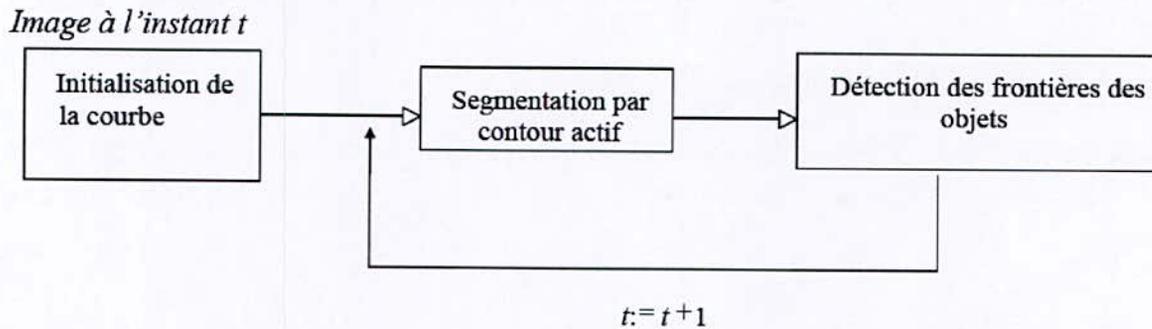


Figure 3.4 : Organigramme pour la segmentation par contour actif

3.3. Evolution temporelle du contour actif :

- Calculer l'énergie pour chaque point [19]
- Faire la liste des points par ordre d'énergie croissante (Figure 3.5)
- Faire évoluer le point avec l'énergie minimale
- Calculer l'énergie nouvelle pour ce point et organiser la liste
- si la distance entre deux points est trop grande, ajouter un point entre les deux

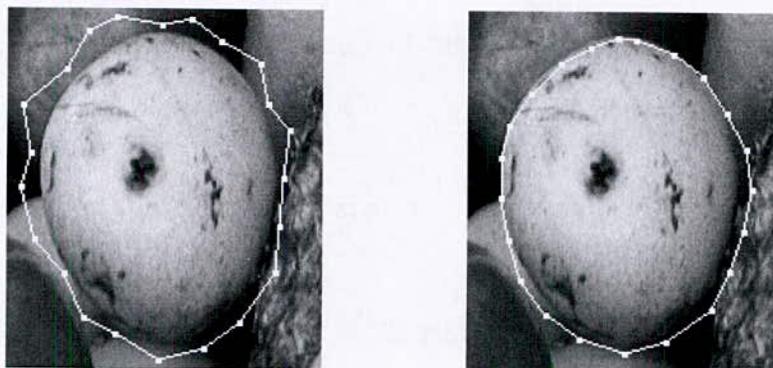


Figure 3.5 : Exemple d'application de l'algorithme de détection de contour par contour actif

3.4. Problèmes liés aux contours actifs :

- Le contour initial ne peut pas être sélectionné automatiquement
- Le contour initial doit être proche du contour final
- Le modèle n'est pas utilisable dans le cas de la présence de texture
- Le modèle peut être perturbé en présence de bruit
- La minimisation d'énergie demande l'inversion de matrices de grande taille à chaque itération

Conclusion :

Ce chapitre présente des notions mathématiques fractales et ces différents paramètres, tel que : la dimension fractale, l'exposant de singularité, ... ainsi qu'une présentation des contours actifs, une technique de segmentation devenue très efficace grâce à ses diverses évolutions.

L'utilisation de la dimension fractale locale permet en effet de détecter les brusques transitions présentes dans l'image pour chaque pixel.

Les algorithmes multifractals présentent l'avantage de la non nécessité de paramètres d'entrée, tel que : seuils, taille des masques, ... Ce qui augmente ses champs d'applications.

Notons, qu'avec l'algorithme multifractal, il est possible de détecter les contours dans des zones texturées, sachant que les exposants de singularité α mesurent la texture.

Chapitre

4

Les Réseaux de Neurones

Chapitre

4

Les Réseaux de Neurones

Introduction

Le terme « Réseau de neurones artificiel » fait rêver beaucoup d'ingénieurs soucieux d'exploiter au maximum les données dont il dispose.

Cet outil s'inspire essentiellement du processus de traitement de l'information effectué par le cerveau humain. [20]

Mais avant d'aboutir à ce système, deux principales étapes ont été franchies :

- L'approche algorithmique basée sur la connaissance.
- L'intelligence artificielle qui se révèle être un moyen de stockage de la connaissance sous forme explicite.

On pense que le système nerveux compte plus de 1000 milliards de neurones interconnectés. Bien que les neurones ne soient pas tous identiques, leur forme et certaines caractéristiques permettent de les répartir en quelques grandes classes. En effet, il est aussi important de savoir, que les neurones n'ont pas tous un comportement similaire en fonction de leur position dans le cerveau. Avant de rentrer plus en avant dans les détails, examinons un neurone.

1. Le Neurone Biologique : [21]

Le schéma de la figure 4.1 représente un neurone biologique :

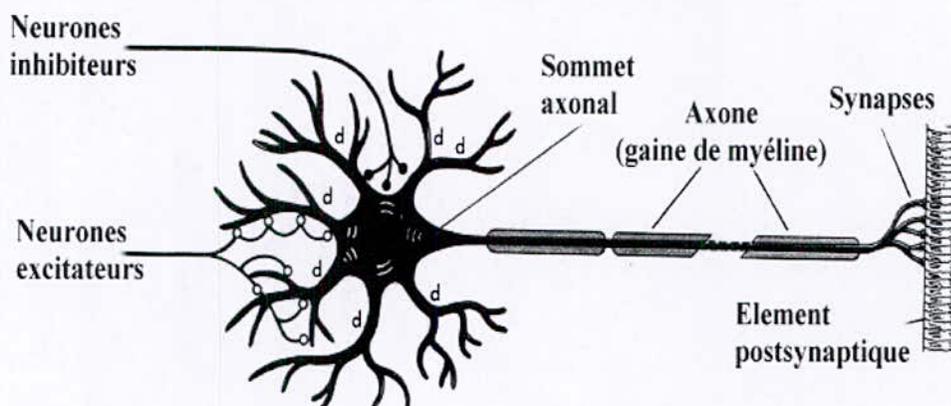


Figure 4.1 : Le neurone biologique

On peut le décomposer en trois régions principales:

Le corps cellulaire :

Il contient le noyau du neurone ainsi que la machine biochimique nécessaire à la synthèse des enzymes. Ce corps cellulaire de forme sphérique ou pyramidale contient aussi les autres molécules essentielles à la vie de la cellule. Sa taille est de quelques microns de diamètre.

Les dendrites :

Ce sont de fines extensions tubulaires qui se ramifient autour du neurone et forment une sorte de vaste arborescence. Les signaux envoyés au neurone sont captés par les dendrites. Leur taille est de quelques dizaines de microns de longueur.

L'axone:

C'est le long de l'axone que les signaux partent du neurone. Contrairement aux dendrites qui se ramifient autour du neurone, l'axone est plus long et se ramifie à son extrémité ou il se connecte aux dendrites des autres neurones. Sa taille peut varier entre quelques millimètres à plusieurs mètres.

Synapse :

Une synapse est une jonction entre deux neurones, et généralement entre l'axone d'un neurone et un dendrite d'un autre neurone (mais il existe aussi des synapses axo-axonales par exemple).

2. Fonctionnement du neurone biologique :

Du point de vue fonctionnel, il faut considérer le neurone comme une entité polarisée, c'est-à-dire que l'information ne se transmet que dans un seul sens : des dendrites vers l'axone.

Le neurone va donc recevoir des informations, venant d'autres neurones, grâce à ses dendrites. Il va ensuite y avoir sommation, au niveau du corps cellulaire, de toutes ces informations et via un potentiel d'action (un signal électrique) le résultat de l'analyse va transiter le long de l'axone jusqu'aux terminaisons synaptiques. A cet endroit, lors de l'arrivée du signal, des vésicules synaptiques vont venir fusionner avec la membrane cellulaire, ce qui va permettre la libération des neurotransmetteurs (médiateurs chimiques) dans la fente synaptique. Le signal électrique ne pouvant pas passer la synapse (dans la cas d'une synapse chimique), les neurotransmetteurs permettent donc le passage des informations, d'un neurone à un autre.

Les neurotransmetteurs excitent (neurotransmetteurs excitateurs) ou inhibent (neurotransmetteurs inhibiteurs) le neurone suivant et peuvent ainsi générer ou interdire la propagation d'un nouvel influx nerveux.

En effet, au niveau post-synaptique, sur la membrane dendritique, se trouvent des récepteurs pour les neurotransmetteurs. Suivant le type de neurotransmetteur et le type des récepteurs, l'excitabilité du neurone suivant va augmenter ou diminuer, ce qui fera se propager ou non l'information.

Les synapses possèdent une sorte de «mémoire» qui leur permet d'ajuster leur fonctionnement. En fonction de leur «histoire», c'est-à-dire de leur activation répétée ou non entre deux neurones, les connexions synaptiques vont donc se modifier.

Ainsi, la synapse va faciliter ou non le passage des influx nerveux. Cette plasticité est à l'origine des mécanismes d'apprentissage.

3. Le Neurone Formel : [22]

Le premier neurone formel est apparu en 1943. On le doit à Mac Culloch et Pitts. La figure 4.2 donne le schéma de leur modèle

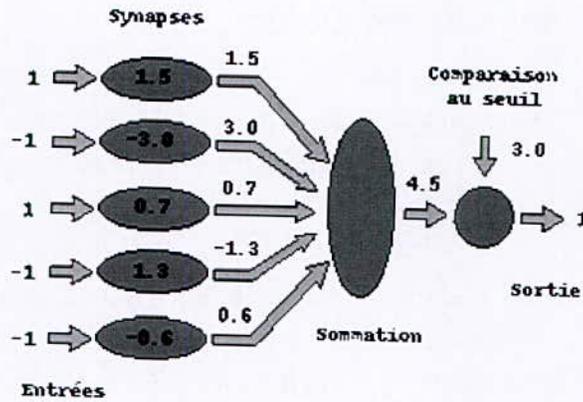


Figure 4.2 Le Neurone Formel

Le neurone formel est donc une modélisation mathématique qui reprend les principes du fonctionnement du neurone biologique, en particulier la sommation des entrées. Sachant qu'au niveau biologique, les synapses n'ont pas toutes la même «valeur» (les connexions entre les neurones étant plus ou moins fortes), les auteurs ont donc créé un algorithme qui pondère la somme de ses entrées par des poids synaptiques (coefficients de pondération). De plus, les 1 et les -1 en entrée sont là pour figurer une synapse excitatrice ou inhibitrice.

3.1. Interprétation mathématique :

D'un point de vue mathématique, le neurone formel peut être représenté de la manière suivante (figure 4.3):

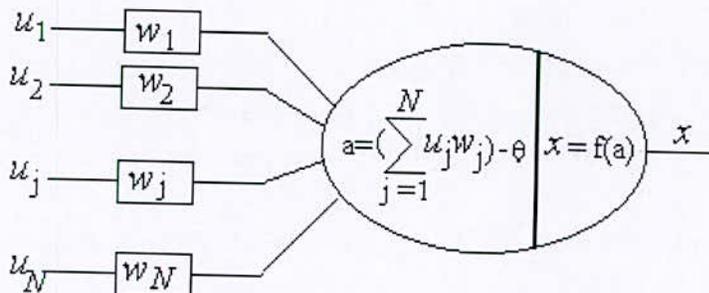


Figure 4.3 Représentation mathématiques du neurone formel

Pour un nombre compris entre $j (=1)$ et un nombre quelconque n , le neurone formel va calculer la somme de ses entrées (u_1, \dots, u_n), pondérées par les poids synaptiques (w_1, \dots, w_n), et la comparer à son seuil θ . Si le résultat est supérieur au seuil, alors la valeur renvoyée est 1, sinon la valeur renvoyée est 0.

D'où la formule (avec f = fonction seuil):

$$y = f\left(\sum_{j=1}^n w_j x_j - \theta\right) \quad (1)$$

3.2. Les Fonctions seuil de bases : [23]

La fonction de base ou de seuillage f définit l'état interne en fonction de ses entrées et des poids.

On peut citer deux exemples de fonctions.

a) Fonction binaire à seuil :

Ce type est caractérisé par deux états distincts : 0 = inactif, 1 = actif tel que le déclenchement de l'activité intervient si la somme des excitations dépasse un certain seuil propre au neurone.

Fonction de Heaviside : $f(x) = 1$ si $x >$ seuil

$$f(x) = 0 \text{ sinon} \quad (2)$$

Le modèle binaire a été perfectionné pour mieux servir la réalité.

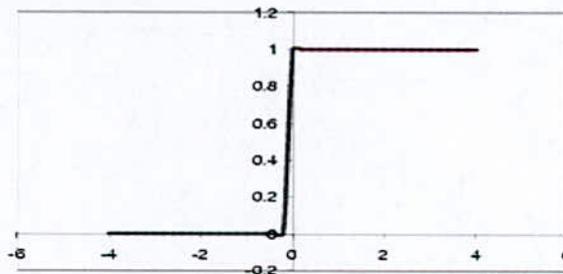


Figure 4.4: Fonction binaire à seuil

b) Fonction sigmoïde :

On dit qu'une fonction est sigmoïde lorsqu'elle est monotone et croissante, dérivable et bornée.

Ce type de fonction combine l'avantage de l'effet de seuil et de la dérivabilité.

On a plusieurs fonctions sigmoïdes fréquemment utilisées.

◆ la fonction exponentielle $f(x) = \frac{1}{1 + \text{Exp}(-x)}$

$$f'(x) = f(x)(1 - f(x)) \quad (3)$$

elle est la plus utilisée du fait que sa dérivée est simple à calculer

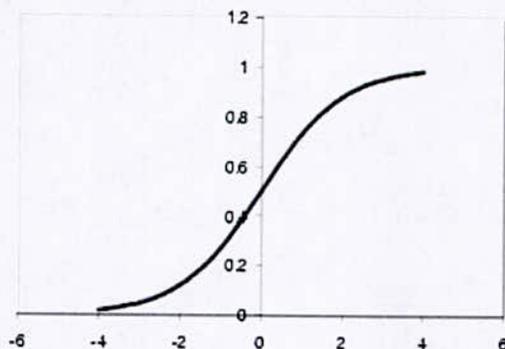


Figure 4.5 : fonction exponentielle

◆ la fonction tangentielle $f(x) = \tanh(x)$

$$f'(x) = 1 - f^2(x) \quad (4)$$

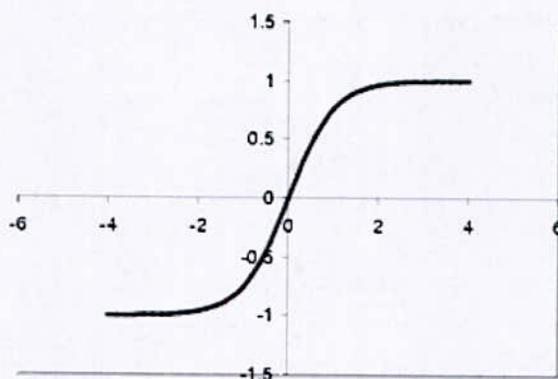


Figure 4.6 : fonction tangentielle

4. Présentation du RNA : [24]

Un RNA (Réseau de Neurones Artificiels) est un ensemble de neurones formels (d'unités de calcul simples, de noeuds processeurs) associés en couches (ou sous-groupes) et fonctionnant en parallèle.

Dans un réseau, chaque sous-groupe fait un traitement indépendant des autres et transmet le résultat de son analyse au sous-groupe suivant. L'information donnée au réseau

va donc se propager couche par couche, de la couche d'entrée à la couche de sortie, en passant soit par aucune, une ou plusieurs couches intermédiaires (dites couches cachées). Il est à noter qu'en fonction de l'algorithme d'apprentissage, il est aussi possible d'avoir une propagation de l'information à reculons ("back propagation"). Habituellement (excepté pour les couches d'entrée et de sortie), chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante.

Les RNA ont la capacité de stocker de la connaissance empirique et de la rendre disponible à l'usage. Les habiletés de traitement (et donc la connaissance) du réseau vont être stockées dans les poids synaptiques, obtenus par des processus d'adaptation ou d'apprentissage. En ce sens, les RNA ressemblent donc au cerveau car non seulement, la connaissance est acquise au travers d'un apprentissage mais de plus, cette connaissance est stockée dans les connexions entre les entités, soit dans les poids synaptiques.

4.1. Architecture d'un réseau de neurone artificiel :

On peut classer les RNA en deux grandes catégories:

a) Les réseaux "FEED-FORWARD" :

Appelés aussi "réseaux de type Perceptron", ce sont des réseaux dans lesquels l'information se propage de couche en couche sans retour en arrière possible.

- Le perceptron monocouche :

C'est historiquement le premier RNA, c'est le Perceptron de Rosenblatt. C'est un réseau simple, puisque il ne se compose que d'une couche d'entrée et d'une couche de sortie. Il est calqué, à la base, sur le système visuel et de ce fait a été conçu dans un but premier de reconnaissance des formes. Cependant, il peut aussi être utilisé pour faire de la classification et pour résoudre des opérations logiques simples (telle "ET" ou "OU"). Sa principale limite est qu'il ne peut résoudre que des problèmes linéairement séparables. Il suit généralement un apprentissage supervisé selon la règle de correction de l'erreur (ou selon la règle de Hebb).

- Le perceptron multicouches :

C'est une extension du précédent, avec une ou plusieurs couches cachées entre l'entrée et la sortie. Chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante (excepté pour les couches d'entrée et de sortie) et il n'y a pas de connexions entre les cellules d'une même couche. Les fonctions d'activation utilisées dans ce type de réseaux sont principalement les fonctions à seuil ou sigmoïdes. Il peut résoudre des problèmes non-linéairement séparables et des problèmes logiques plus compliqués, et notamment le fameux problème du XOR. Il suit aussi un apprentissage supervisé selon la règle de correction de l'erreur.

- Les réseaux à fonction radiale :

Ce sont les réseaux que l'on nomme aussi RBF ("Radial Basic Functions"). L'architecture est la même que pour les PMC cependant, les fonctions de base utilisées ici sont des fonctions Gaussiennes. Les RBF seront donc employés dans les mêmes types de problèmes que les PMC à savoir, en classification et en approximation de fonctions, particulièrement. L'apprentissage le plus utilisé pour les RBF est le mode hybride et les règles sont soit, la règle de correction de l'erreur soit, la règle d'apprentissage par compétition.

b) Les réseaux "FEED-BACK" :

Appelés aussi "réseaux récurrents", ce sont des réseaux dans lesquels il y a retour en arrière de l'information.

- Les cartes auto-organisatrices de Kohonen :

Ce sont des réseaux à apprentissage non-supervisé qui établissent une carte discrète, ordonnée topologiquement, en fonction des entrées. Le réseau forme ainsi une sorte de treillis dont chaque noeud est un neurone associé à un vecteur de poids. La correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins, vont être modifiés afin d'augmenter encore cette corrélation.

- Les réseaux de Hopfield :

Les réseaux de Hopfield sont des réseaux récurrents et entièrement connectés. Dans ce type de réseau, chaque neurone est connecté à chaque autre neurone et il n'y a aucune différenciation entre les neurones d'entrée et de sortie. Ils fonctionnent comme une mémoire associative non-linéaire et sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées. L'application principale des réseaux de Hopfield est l'entrepôt de connaissances mais aussi la résolution de problèmes d'optimisation. Le mode d'apprentissage utilisé ici est le mode non-supervisé.

- Les ART :

Les réseaux ART ("Adaptative Resonance Theorie") sont des réseaux à apprentissage par compétition. Le problème majeur qui se pose dans ce type de réseaux est le dilemme « stabilité/plasticité ». En effet, dans un apprentissage par compétition, rien ne garantit que les catégories formées vont rester stables. La seule possibilité, pour assurer la stabilité, serait que le coefficient d'apprentissage tende vers zéro, mais le réseau perdrait

alors sa plasticité. Les ART ont été conçus spécifiquement pour contourner ce problème. Dans ce genre de réseau, les vecteurs de poids ne seront adaptés que si l'entrée fournie est suffisamment proche, d'un prototype déjà connu par le réseau. On parlera alors de résonance. A l'inverse, si l'entrée s'éloigne trop des prototypes existants, une nouvelle catégorie va alors se créer, avec pour prototype, l'entrée qui a engendrée sa création. Il est à noter qu'il existe deux principaux types de réseaux ART : les ART-1 pour des entrées binaires et les ART-2 pour des entrées continues. Le mode d'apprentissage des ART peut être supervisé ou non.

- **Comparaison entre les différents réseaux de neurones :**

Modèle du Réseau \ Application	Rétro-propagation	Hopfield	Kohonen
Classification	Oui	Oui	Oui
Traitement d'image	Oui	Non	Oui
Aide à la décision	Oui	Non	Oui
Optimisation	Non	Oui	Oui

Vu le domaine de notre étude qui est « la reconnaissance des formes », et d'après le tableau ci-dessus, nous constatons que pour réaliser cette partie de l'application, nous avons le choix entre le modèle Multicouche ou celui de Kohonen.

4.2. Apprentissage :

Pour un RNA, l'apprentissage peut être considéré comme le problème de la mise à jour des poids des connexions au sein du réseau, afin de réussir la tâche qui lui est demandée. L'apprentissage est la caractéristique principale des RNA et il peut se faire de différentes manières et selon différentes règles.

a) Le mode supervisé :

Dans ce type d'apprentissage, le réseau s'adapte par comparaison entre le résultat qu'il a calculé, en fonction des entrées fournies, et la réponse attendue en sortie. Ainsi, le réseau va se modifier jusqu'à ce qu'il trouve la bonne sortie, c'est-à-dire celle attendue, correspondant à une entrée donnée.

- Le renforcement :

Le renforcement est en fait une sorte d'apprentissage supervisé et certains auteurs le classe d'ailleurs, dans la catégorie des modes supervisés. Dans cette approche le réseau doit apprendre la corrélation entrée/sortie via une estimation de son erreur, c'est-à-dire du rapport échec/succès. Le réseau va donc tendre à maximiser un index de performance qui lui est fourni, appelé signal de renforcement. Le système étant capable ici, de savoir si la réponse qu'il fournit est correcte ou non, mais il ne connaît pas la bonne réponse.

b) Le mode non-supervisé (ou auto-organisationnel) :

Dans ce cas, l'apprentissage est basé sur des probabilités. Le réseau va se modifier en fonction des régularités statistiques de l'entrée et établir des catégories, en attribuant et en optimisant une valeur de qualité, aux catégories reconnues.

c) Le mode hybride :

Le mode hybride reprend en fait les deux autres approches, puisque une partie des poids va être déterminée par apprentissage supervisé et l'autre partie par apprentissage non-supervisé.

4.3. Règles d'apprentissage :

a) Règle de correction d'erreurs :

Cette règle s'inscrit dans le paradigme d'apprentissage supervisé, c'est -à-dire dans le cas où l'on fournit au réseau une entrée et la sortie correspondante. Si on considère y comme étant la sortie calculée par le réseau, et d la sortie désirée, le principe de cette règle est d'utiliser l'erreur $(d-y)$, afin de modifier les connexions et de diminuer ainsi l'erreur globale du système. Le réseau va donc s'adapter jusqu'à ce que y soit égal à d .

Ce Principe est notamment utilisé dans le modèle du perceptron simple.

b) Apprentissage de Boltzmann :

Les réseaux de Boltzmann sont des réseaux symétriques récurrents. Ils possèdent deux sous-groupes de cellules, le premier étant relié à l'environnement (cellules dites visibles) et

le second ne l'étant pas (cellules dites cachées). Cette règle d'apprentissage est de type stochastique (= qui relève partiellement du hasard) et elle consiste à ajuster les poids des connexions, de telle sorte que l'état des cellules visibles satisfasse une distribution probabiliste souhaitée.

c) Règles de Hebb :

Comme je l'ai déjà dit dans l'historique, cette règle, basée sur des données biologiques, modélise le fait que si des neurones, de part et d'autre d'une synapse, sont activés de façon synchrone et répétée, la force de la connexion synaptique va aller croissant. Il est à noter ici que l'apprentissage est localisé, c'est-à-dire que la modification d'un poids synaptique w_{ij} ne dépend que de l'activation d'un neurone i et d'un autre neurone j .

d) Règle d'apprentissage par compétitions :

La particularité de cette règle, c'est qu'ici l'apprentissage ne concerne qu'un seul neurone. Le principe de cet apprentissage est de regrouper les données en catégories. Les patrons similaires vont donc être rangés dans une même classe, en se basant sur les corrélations des données, et seront représentés par un seul neurone, on parle de « winner-take-all ».

Dans un réseau à compétition simple, chaque neurone de sortie est connecté aux neurones de la couche d'entrée, aux autres cellules de la couche de sortie (connexions inhibitrices) et à elle-même (connexion excitatrice). La sortie va donc dépendre de la compétition entre les connexions inhibitrices et excitatrices.

Conclusion :

Au début de la cybernétique, de l'automatique et de l'informatique, l'Intelligence Artificielle (IA) classique et connexionniste se sont développées en parallèle.

Après l'article de Minsky et Papert en 1969 démontrant les faiblesses du Perceptron, les années 70 ont vu le désintérêt pour le modèle connexionniste et le succès de l'IA classique (cognitive) (manipulant des symboles logiques - comme les systèmes experts). Mais cette approche a aussi ses limites. D'abord sa lenteur : on est vite confronté à ce que l'on appelle l'explosion combinatoire. La faiblesse est aussi théorique ; la connaissance ne se réduit pas au langage. L'intelligence n'est pas qu'un raisonnement mathématique manipulant ces symboles. Une grande partie des experts travaillent de manière intuitive, en se fondant sur l'expérience accumulée. On a alors beaucoup de mal à faire ressortir les raisonnements qu'ils utilisent.

De plus, d'autres facultés humaines sont mal réalisées par l'IA classique, comme l'adaptation et la généralisation.

Chapitre 4: Les Réseaux de Neurones

Les réseaux neuromimétiques, bien qu'ils ne soient pas non plus une solution, ont aussi leur mot à dire : en particulier quand il n'existe pas de technologie ou d'algorithme classique mais des bases d'exemples, quand le système étudié est bruité, évolue dans le temps, etc. En 1982 Hopfield a remis à la mode l'approche connexionniste. Aujourd'hui, les nouvelles technologies du multimédia comme le traitement de l'image et du son demandent une rapidité de plus en plus grande.

Les ordinateurs séquentiels s'approchent de leur limite et l'on cherche à paralléliser les microprocesseurs et à trouver les algorithmes pouvant les utiliser à pleine puissance. Or les réseaux de neurones artificiels permettent, par leur conception même, cette parallélisation : ce sont un ensemble d'entités élémentaires travaillant en simultané.

Bien que l'on soit encore très loin des neuro-ordinateurs, d'ores et déjà les ordinateurs actuels pourraient s'équiper de modules spécialisés (cartes son, vidéo) s'inspirant de la technologie des réseaux neuromimétiques.

Les liens des réseaux de neurones avec les algorithmes génétiques (pour la modification des poids synaptique, par exemple), la théorie du chaos (modélisation d'un attracteur étrange), le besoin d'algorithmes parallèles, et l'association avec l'IA cognitive permettent d'augurer du développement de la technologie neuromimétique.

Chapitre

5

Le travail effectué

Chapitre

5

Le travail effectué

Introduction

:

Dans les chapitres précédents, nous avons présenté un aspect général de différentes étapes constituant notre travail qui se résument en :

L'acquisition et la numérisation pour l'obtention de l'élément de base à notre étude qui est l'image.

Les pré-traitements pour améliorer et filtrer l'image.

La segmentation pour extraire les cytoplasmes et les noyaux.

L'analyse et la reconnaissance pour donner une décision finale sur l'état de la cellule.

Dans ce chapitre, nous allons expliquer chaque étape, montrer quelques exemples d'images que nous avons traitées, puis nous allons interpréter les résultats obtenus en mettant en évidence les diversités des cas de figures.

1. L'acquisition : [7]

C'est la première étape d'un système de traitement d'image, l'image naturelle est une représentation bidimensionnelle d'objets tridimensionnels.

La numérisation peut se faire à l'aide de périphériques spécifiques tels que caméra numérique ou scanner. (Figure 5.1)

Pour ce travail, nous avons eu la chance d'avoir à notre disposition un microscope équipé d'une caméra numérique, ce qui nous a permis de prendre des photos des lames contenant des prélèvements cytologiques, et nous les avons visionné sur l'ordinateur directement.(Figure 5.2)

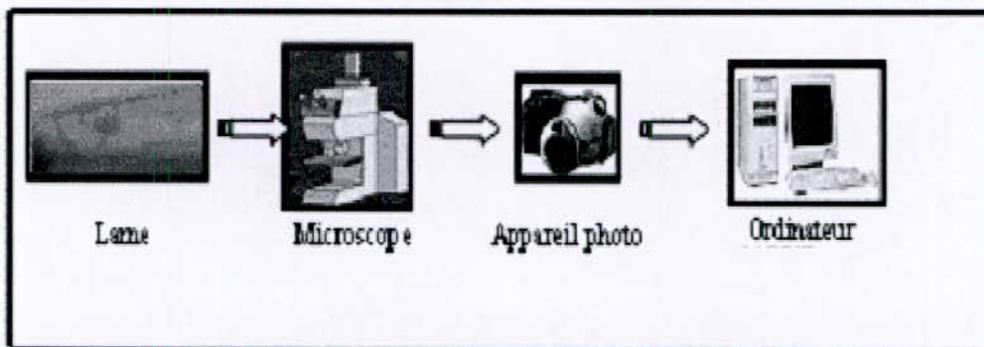


Figure 5.1 : Dispositif d'acquisition d'images

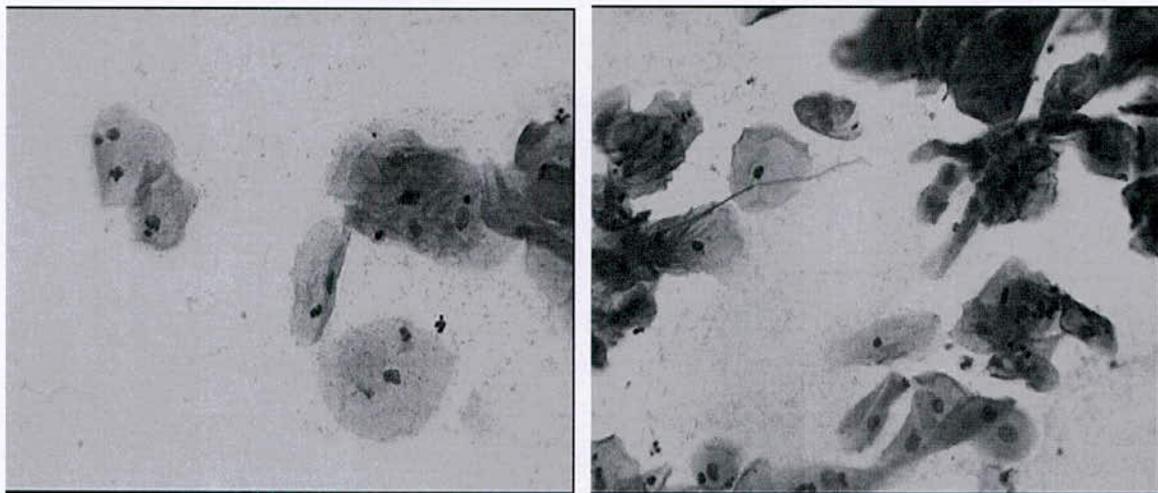


Figure 5.2 : Image extraite du microscope

Description du matériel utilisé :

Le microscope à permis la visualisation de l'image sur moniteur du PC (fonctionne sans carte d'acquisition) relié par câble USB.

Caractéristiques du microscope : Tourelle 5 objectifs orientée vers l'arrière. Eclairage halogène 20 W. Objectifs Achromatiques 4x, 10x, 40x, 100x à immersion. 2 oculaires grand champs WF 10x / 20. Condenseur d'Abbe ON : 1,25. Caméra couleur numérique intégrée. Capteur CCD 1/3". Balance des blancs automatiques.

2. Prétraitement :

L'image obtenue par l'acquisition est brute (Figure 6.4) ; elle peut contenir des informations indésirables ou inexploitable qui doivent être éliminées. Dans notre cas, et comme nous avons besoin des cellules complètes, c'est à dire avec des cytoplasmes et des noyaux complets. Ces informations indésirables se résument-en :

- Cytoplasme sans noyau,
- Noyau sans cytoplasme,
- Cellules tronquées (c'est à dire qu'une partie du cytoplasme ou du noyau a été tronquée lors de l'étape d'acquisition), (Figure 5.3)
- Cellules dont le cytoplasme est rabattu sur lui-même. Ceci se produit lorsque les tissus (les cellules) prélevés ont été mal fixés sur lame.

Cellule tronquée

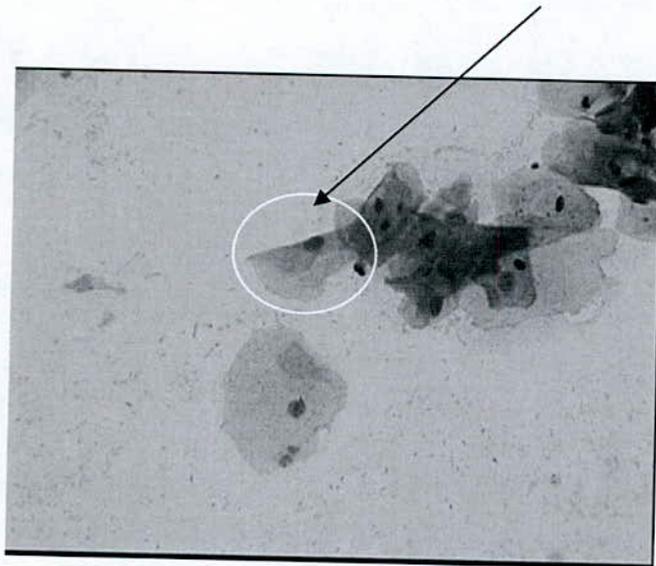


Figure 5.3 : Image de cellule inutilisable

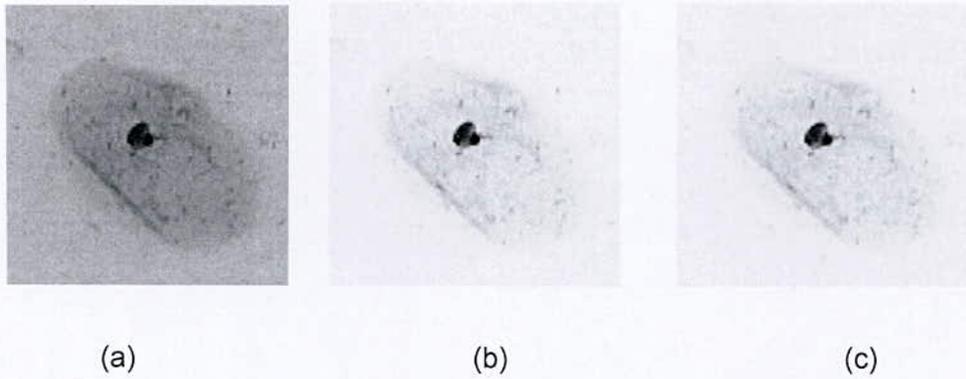


Figure 5.4 : Image de cellule (a) brute (b) contrastée (c) grisée

Pour éliminer toutes les zones contenant des points isolés, ou bien seulement des informations indésirables, on peut recourir à des logiciels confirmés tel que : PhotoShop ou bien PaintBrush, mais ceci risque de nous faire perdre beaucoup de temps, et en plus nous obliger à les installer partout, là où on doit utiliser notre application. Notre but étant d'alléger le travail des médecins, à cet effet, nous avons intégré une fonction de sélection de chaque zone non souhaitée à n'importe quel moment du traitement et de la remplacer par un rectangle blanc (ou ayant la structure du fond de l'image).

C'est une fonction « gomme » qui va permettre au médecin d'éliminer tout élément de l'image qu'il jugera indésirable. (Figure 5.5 a,b,c)

Ces deux zones représentent des

informations inexploitable

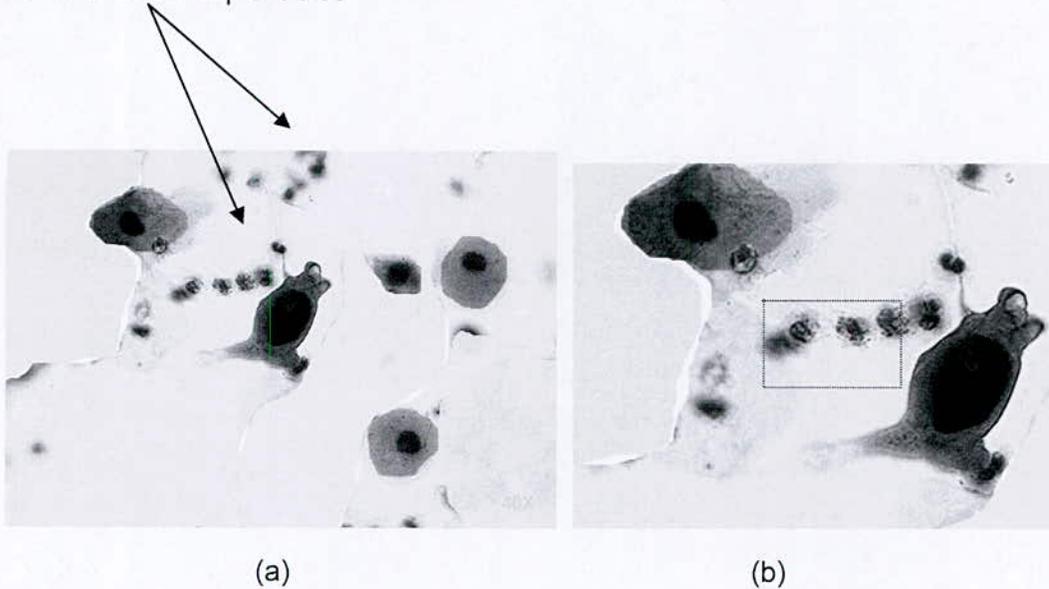


Figure 5.5. (a) Image avec des zones inexploitable (b) Utilisation de la gomme pour la sélection

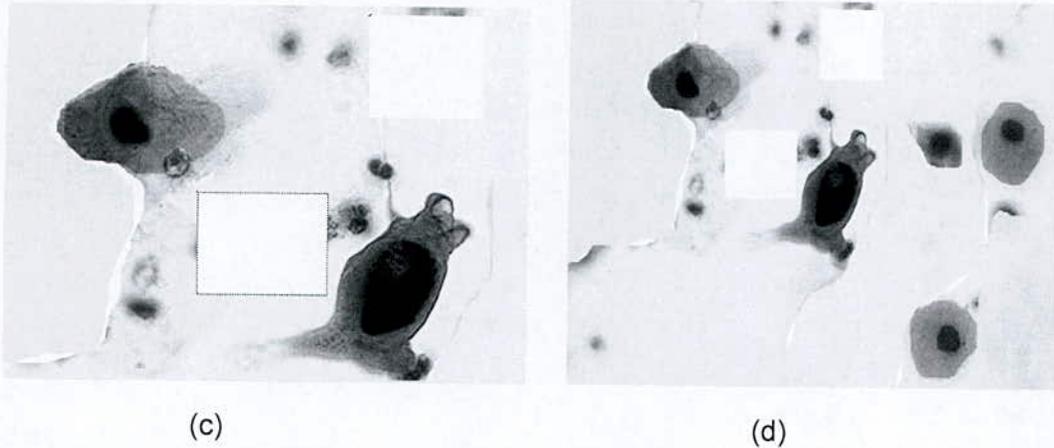


Figure 5.5. (c) remplacement des zones indésirées par des pixels blancs (d) Image filtrée

Pour les bruits, nous avons implémenté le filtre médian car en le comparant au filtre moyen ou au filtre de gauss il n'introduit pas de flou, et élimine le mieux les bruits impulsionnels. [Annexe 2] (Figure 5.5.d)

Les filtres morphologiques (Erosion et Dilatation) sont aussi utilisés, leur plus grand défaut est la déformation (changement de taille) de l'objet traité ce qui peut fausser les résultats obtenus sur ces derniers. Pour cela on favorise l'utilisation de l'ouverture ou de la fermeture qui restitue la taille de l'objet.

3. La segmentation :

La segmentation, dans le processus du traitement des images, est l'étape de subdivision de l'image en parties qui la constituent. Elle est considérée comme la phase la plus importante, car c'est à ce niveau que les objets et les autres entités, sont extraits. [2] D'après l'étude de la cellule cancéreuse et de ses critères de malignité, qui dépendent essentiellement de certaines caractéristiques du noyau et celles du cytoplasme de la cellule ; (rapport nucléo-cytoplasmique, anisocaryose, la difformité, l'hyper chromatisme nucléaire) nous avons considéré les objets de notre segmentation comme étant : les noyaux et les cytoplasmes.

- ✓ **L'objet noyau** : dans l'image, les noyaux apparaissent comme les régions les plus foncées.
- ✓ **L'objet cytoplasme** : apparaît comme des régions moins foncées que celles de l'objet noyau de l'image.

Chapitre 5 : Le travail effectué

La segmentation par la méthode multifractale, que nous avons choisie après une longue étude comparative [25].

L'approche multifractale se base sur les exposants de singularité des pixels . [26]

- Les régions les plus foncées correspondent aux exposants de singularité supérieurs,
- Les régions les plus claires correspondent aux exposants de singularité inférieurs.

Nous adaptons le paramètre multifractal Alpha à la segmentation des images. Pour cela, nous étudions le comportement des exposants de singularité, calculés en chaque point de l'image. Nous obtenons donc une image des exposants de singularité.

$$m(l(x,y)) \approx (1/r)^{-\alpha(x,y)} \quad \text{quand } r \rightarrow 0, \quad (1)$$

Où :

$m(x, y)$: est la mesure associée au point (x, y) de niveau de gris $l(x, y)$,

r : est la largeur de la boîte autour du point (x, y) ,

α : est l'exposant de Hôlder au point (x, y) .

Ceci est fait par l'estimation de la limite du rapport

$(\log(P_i) / \log(1/r))$ par une régression linéaire. Ce qui permet d'obtenir l'exposant de singularité de chaque pixel.

P_i : est la mesure d'un sous-ensemble(niveau de gris).

Donc l'exposant de singularité (Alpha), représente le degré d'homogénéité de l'image et est calculé comme étant l'estimation de la pente du nuage de points :

$$\{ \log(U(V(r))), \log(1/r) \} \quad (2)$$

avec :

$V(r)$: voisinage de taille r du pixel concerné,

$U(V(r))$: la mesure du voisinage du pixel concerné.

Le calcul de Alpha portera seulement sur le voisinage (d) du pixel ce qui évite le calcul de la régression, il en découle un gain en temps de calcul considérable. Alpha est obtenu par la relation suivante :

$$(P(x,y)/p) = (1/(d/L))^{-\text{Alpha}(x,y)} \quad (3)$$

Donc :

$$\log(P(x,y)/p) = -\text{Alpha}(x,y) * \log(1/(d/L)) \quad (4)$$

On trouve :

$$\text{Alpha}(x,y) = - (\log(P(x,y)/p) * (1/\log(1/(d/L)))) \quad (5)$$

$$\text{Alpha}(x,y) = - (\log(P(x,y)/p) * (1/\log(L/d))) \quad (6)$$

Si on pose $d = 1$

On obtient :

$$\text{Alpha}(x,y) = - (\log(P(x,y)/p) * (1/\log(L))) \quad (7)$$

Tel que :

$P(x,y)$: Le niveau de gris du pixel (x,y) ,

Chapitre 5 : Le travail effectué

p : La somme des niveaux de gris des pixels de l'image.

L : La taille de l'image.

L'algorithme de segmentation est le suivant : [16]

- Calculer les exposants de singularité en chaque point de l'image,
- Trouver la singularité minimale AlphaMin et la singularité maximale AlphaMax ,
- Diviser l'intervalle $[\text{AlphaMin}, \text{AlphaMax}]$ en plusieurs parties (N sous intervalles),
- Pour chaque sous intervalle, on obtient l'ensemble des points de l'image ayant l'exposant de singularité appartenant à l'intervalle, ce qui permet d'avoir N sous images, chacune correspondant à des intervalles iso-alpha $[\text{Alpha1}, \text{Alpha2}]$:

Tel que :

$\text{AlphaMin} < \text{Alpha1} < \text{AlphaMax}$,

$\text{AlphaMin} < \text{Alpha2} < \text{AlphaMax}$.

Sachant que pour l'intervalle $[\text{Alpha1}, \text{Alpha2}]$ donné, les pixels retenus seront ceux dont le niveau de gris sera compris dans l'intervalle $[P^*(1/L)\text{Alpha1}, P^*(1/L)\text{Alpha2}]$.

- Pour chaque image d'un intervalle iso-alpha, les points en noir sont ceux dont la valeur Alpha est dans l'intervalle souhaité.

Dans cet algorithme, le calcul de l'exposant de singularité est effectué selon la formule déjà vue (1) .

Ce qui permet d'écrire l'équation suivante :

$$(P(x,y)/p) = (r/1)^{\text{Alpha}(x,y)} \quad (8)$$

Soit à l'échelle du pixel

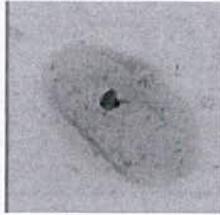
$$P(x,y) = p^*(1/L)^{\text{Alpha}(x,y)} \quad (9)$$

Avec :

p : La somme des niveaux de gris de l'image,

L : La taille de l'image

Nous avons obtenus différents résultats pour l'approche multifractale présentés ci-dessous



(a1)



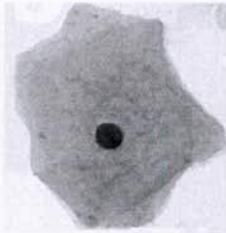
(b1)



(c1)



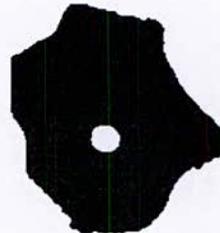
(d1)



(a2)



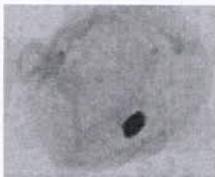
(b2)



(c2)



(d2)



(a3)



(b3)



(c3)



(d3)



(a4)



(b4)



(c4)



(d4)

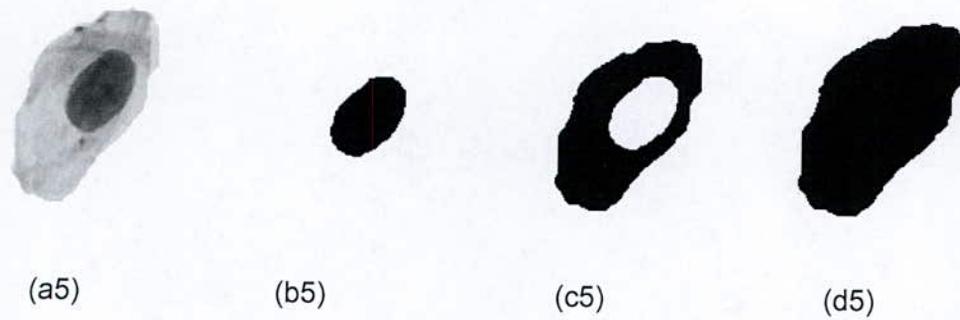


Figure 5.6 : Image résultat de segmentation d'une cellule
(a) image originale , (b) image noyau, (c) image cytoplasme , (d) image cellule

Nous avons procédé à la segmentation de plusieurs cellules. Nous avons obtenu des images de noyaux, de cytoplasmes et de cellules (Figure 5.6) sur lesquelles nous allons faire des calculs des différents critères de malignité cités ci-dessus. Nous avons obtenus un taux de segmentation égale à 80% de réussite , parmi les images jugées non segmentées nous avons l'exemple suivant :



Figure 5.7 : Image non segmentable

Interprétation des origines de certains problèmes rencontrés lors de la réalisation de notre application :

Parmi les problèmes rencontrés durant notre travail et essentiellement durant la segmentation, nous avons été confronté à certaines particules isolées du noyau (Figure 5.8).

Dans le but de faciliter et d'appuyer notre argumentation, nous allons définir au préalable certaines notions fondamentales de biologie cellulaire.

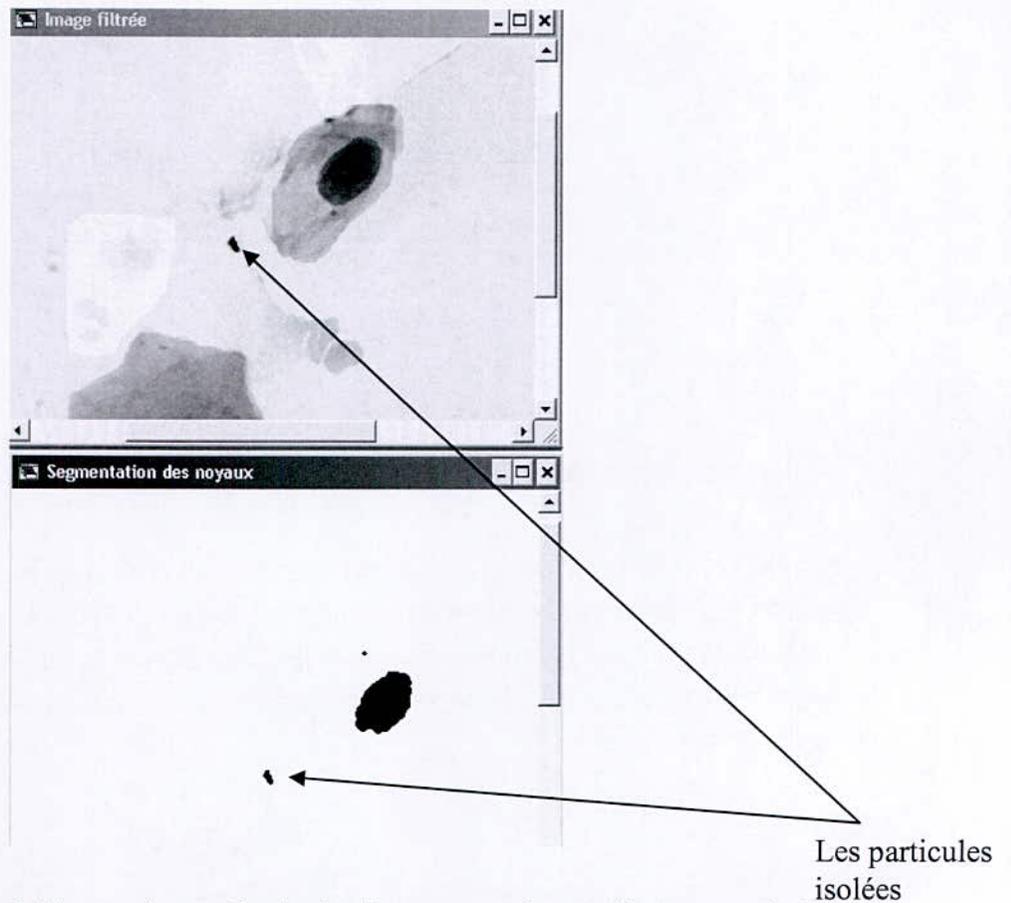
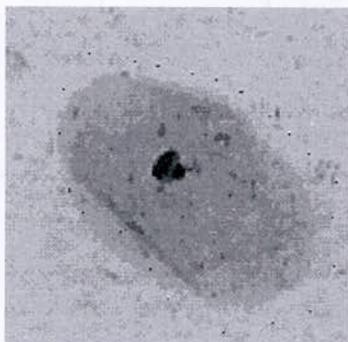


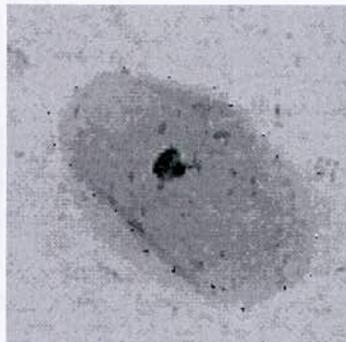
Figure 5.8 : Mise en évidence des particules isolées sur une image filtrée et après la segmentation des noyaux

Ces particules sont des constituant ou plus précisément certaines copies partielles du code génétique de l'ADN contenu dans le noyau et donc a l'échelle microscopique de même constitution que celui ci, peuvent fréquemment, surtout pendant les différentes phases de division cellulaire se retrouver dans le cytoplasme a l'extérieur du noyau et peuvent par conséquent être a l'origine des différentes imperfections que nous avons obtenus sur nos différents clichés.

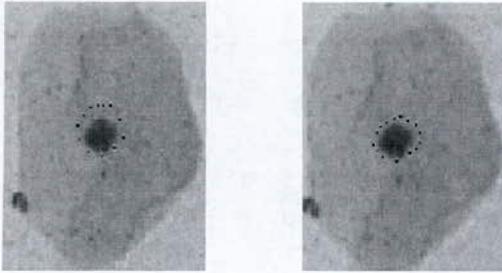
Pour les contours actifs nous avons obtenus les résultats suivants :



(d1)



(e1)



(d2)

(e2)

Figure 5.9 : Segmentation par contour actif (d) : contour initial, (e) : contour final

Cette méthode peut servir à extraire des zones comme le noyaux ou la cellule tout entière pour la segmentation ou le calcul de la surface des composants de l'image. (figure 5.8)

Une fois nos deux objets séparés , nous passons à la reconnaissance et ceci en appliquant l'algorithme d'étiquetage par partie connexe [9] pour récupérer sur l'image la cellule et calculer les critères de malignité.

4. Analyse et reconnaissance :

L'analyse des cellules permet d'identifier puis d'extraire les caractéristiques des cellules normales et anormales.

La reconnaissance quant à elle permet par l'utilisation de ces caractéristiques et en comparaison avec celles de la base de données, de prendre une décision sur l'état de la cellule.

Nous avons utilisé pour cette dernière étape les réseaux de neurones multicouches.

4.1. L'extraction des critères de malignité :

D'après les spécialistes, les quatre critères de malignité (cités précédemment) sont suffisants pour identifier la plupart des cellules cancéreuses. Elles sont les caractéristiques les plus fréquentes du cancer.

4.1.1 Le rapport nucléo-cytoplasmique :

C'est le rapport entre la surface prise par le noyau et celle qui constitue le cytoplasme. Dans la plupart des tissus cancéreux, les noyaux croissent de manière anarchique et envahissent presque toute la cellule. Cette évolution anormale des noyaux permet d'identifier la plupart des cellules cancéreuses.

Ce rapport permet ainsi dans le cas de frottis cervicaux, de classer les différents types de cellules (superficielles , intermédiaires, basales et parabasales)

Chapitre 5 : Le travail effectué

Pour réaliser la classification des cellules en fonction du rapport de surface noyau/cytoplasme, il faut calculer :

- Les surfaces des différents noyaux.
- Les surfaces des différents cytoplasmes.
- Le rapport noyau/cytoplasme correspondant à chaque cellule.

L'algorithme que nous avons utilisé est celui de « croissance de région », itératif et récursif, basé sur le principe suivant :

Pour chaque point de l'image $I(x,y)$ rencontré de niveau de gris zéro, on fait appel à la fonction **CalSurface** qui permet de calculer la surface d'une région connexe de points qui peut être noyau ou cytoplasme.

Le programme principal :

début

Pour chaque pixel $I(x,y)$

faire

Si ($I(x,y)=0$) alors début

NbrSurf := NnbrSurf + 1;

Compteur := 1;

S := **CalSurface**(x,y) ; // Appel à la fonction CalSurface

Sauvegarder le point de départ (x,y) dans une variable

Adrsurface ;

Surface := S ;

Fin ;

FSi ;

FPour ;

Fin.

La fonction CalSurface :

CalSurface(x,y)

Début

On met $I(x,y)$ à 255 afin de ne pas le réexaminer ;

Pour tout pixel $I(m,n)$ adjacent à $I(x,y)$

Faire

Si ($I(m,n) = 0$) Alors

Compteur := compteur + 1 ;

$I(m,n)$:= 255 ;

CalSurface(m,n) ;

FSi ;

Fait ;

Renvoyer Compteur ;

Fin.

Cet algorithme nous permet d'avoir :

1. Le nombre de cellules, qui n'est autre que le nombre de surface estimé (variable NbrSurf).
2. La surface de chaque cellule, celle-ci est emmagasinée dans le vecteur surface de taille NbrSurf.
3. AdrSurf qui contient l'adresse du premier point noir de chaque surface, il a un rôle très important à jouer dans l'étape de classification en classes cellulaires.

4.1.2. L'anisocaryose

L'anisocaryose ou l'inégalité de la taille des noyaux d'une cellule à l'autre (pour le même type de cellule) , est calculée à partir de l'image des noyaux, tel que pour chaque noyau on suit les étapes suivantes :

4.1.2.1. Calcul du centre de gravité de chaque noyau

Ce calcul peut se faire par l'une des deux méthodes :

- Soit par une suite d'érosions, tel que le centre de gravité est le dernier point avant la disparition du noyau.

- Soit par un calcul mathématique en suivant ces étapes :

On détecte le contour en appliquant l'un des opérateurs qui détecte le contour (Robert, Sobel, ...)

On applique au contour détecté l'algorithme d'amincissement pour obtenir un contour d'un pixel d'épaisseur, puis on extrait les coordonnées des points du contour $M_i(x_i, y_i)$.

On calcule enfin les coordonnées du centre de gravité données par :

$$G_x = \frac{\sum_{i=1}^N x_i}{N} \quad G_y = \frac{\sum_{i=1}^N y_i}{N} \quad (10)$$

Où : N est le nombre de points du contour.

4.1.2.2. Calcul de la taille moyenne du noyau

La taille moyenne des noyaux est la moyenne des huit distances caractéristiques calculées à partir du centre de gravité :

$$T_{mi} = \text{MOY}(d_{ij}) = 1/8 \sum_{j=1}^{j=8} d_{ij} \quad (11)$$

Après avoir appliqué ces deux opérations sur chaque noyau de l'image, on calcule la taille moyenne des noyaux :

$$T_m = \text{Moy}(T_{mi}) \quad (12)$$

$$\text{Finalement : Anisocaryose} = \text{Var}(T_{mi}) / T_m^2 \quad (13)$$

$$\text{Où : } \text{Var}(T_{mi}) = 1/nb \sum_{i=1}^{i=nb} (T_{mi} - T_m)^2$$

nb : Nombre de cellules dans l'image

4.1.3. La difformité nucléaire

Il est clair qu'un noyau de forme quasi circulaire présentera des distances caractéristiques proches, c'est à dire avec peu de variance de distance caractéristique, sur toutes les directions significatives. On peut traduire le caractère difforme du noyau par :

$$\text{Difformité Nucléaire} = \text{MOY}(DF_i) / T_m^2 \quad (14)$$

Tel que : la difformité du noyau est égale à la

variance des huit distances caractéristiques précédentes :

$$DF_i = \text{Var}(d_{ij}) \quad (15)$$

4.1.4. L'hyperchromatisme nucléaire

Rappelons que le cancer provoque un accroissement anormal de l'activité nucléaire ; ce qui se traduit dans l'image par des noyaux granuleux, présentant une texture particulière. Il s'agit donc ici, de travailler à nouveau à partir de l'image « noyaux » originale.

Dans une cellule cancéreuse, le noyau contient des sauts de luminance, d'où l'existence des zones claires. Pour détecter ces creux de luminance, on suit les étapes suivantes :

- Obtenir l'image des noyaux avec la luminance qu'ils présentent dans l'image originale.
- Effectuer une fermeture de l'image noyaux par un petit élément structurant.
- Effectuer la différence de luminance entre l'image noyaux originale et son image fermée: nous obtenons une image où ne sont présentées que les irrégularités de luminance.
- Calcul du paramètre « hyperchromatisme nucléaire » qui est la somme des luminances des pixels de cette image, divisée par le nombre de pixels de l'image noyaux

Nous avons obtenus les résultats illustrés ci-dessous (Figure 5.10)

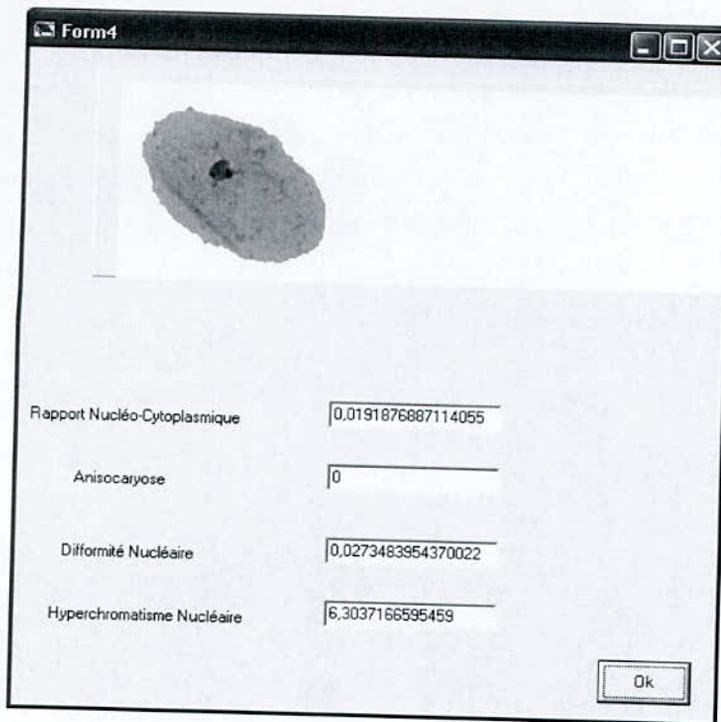


Figure 5.10 : calcul des critères de maligité

4.2. Classification des cellules en fonction du rapport nucléocytoplasmique

Précédemment, nous avons dit que ce qui permettait de différencier les types de cellules est le rapport noyau / cytoplasme. Rappelons que plus la cellule est jeune et se trouve dans une couche profonde, plus le rapport est élevé ; inversement, plus la cellule est vieille et se trouve dans une couche superficielle, plus le rapport est faible. [5]

Les intervalles du rapport nucléocytoplasmique, correspondant à chaque type de cellules, sont les suivants :

- Cellules basales : $0,25 \leq \text{rapport} < 0,37$.
- Cellules para-basales : $0,15 \leq \text{rapport} < 0,25$.
- Cellules intermédiaires : $0,04 \leq \text{rapport} < 0,15$.
- Cellules superficielles : $0,01 \leq \text{rapport} < 0,04$.

La classification est réalisée selon le principe suivant : pour chaque rapport nucléocytoplasmique qui appartient à un intervalle définissant le type ou la classe, nous récupérons son image originale. Et le résultat de chaque traitement se trouve dans une image qui correspond à la classe de la cellule. [28]

Pour réaliser cette opération, on exécutera les étapes suivantes pour chaque cellule :

Chapitre 5 : Le travail effectué

- Comparer le rapport nucléo-cytoplasmique aux différents intervalles.
- Après avoir déterminé l'intervalle qui convient, faire appel à la fonction qui correspond à cet intervalle et qui affectera la cellule à l'image adéquate. Il est important de signaler qu'on a quatre fonctions, chacune correspond à une catégorie de cellules, elles sont basées sur le même principe, la seule différence réside dans l'image résultat. (Figure 5.11)

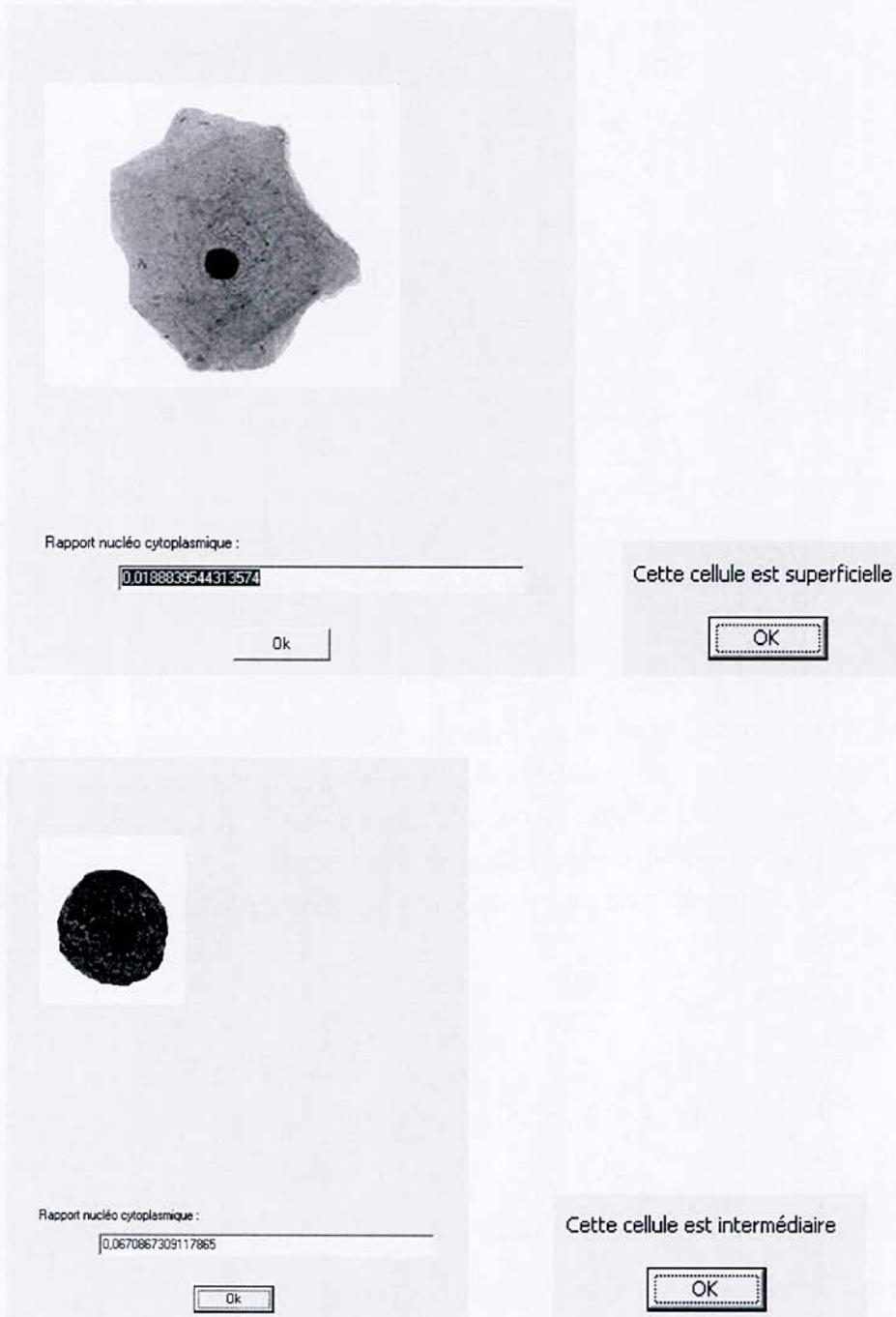


Figure 5.11 : Images résultats de la classification par rapport au RNC

4.3. Les réseaux de neurones :

Un réseau de neurones est constitué de cellules ou neurones. Ces cellules sont reliées entre elles par des connexions, ce qui leur permet de disposer d'un canal pour envoyer et recevoir des signaux en provenance d'autres cellules du réseau. Chacune de ces connexions reçoit une valeur: un poids (une pondération en quelque sorte). Les cellules disposent d'une entrée, qui permet de recevoir de l'information provenant d'autres cellules. Elles disposent aussi d'une fonction d'activation, qui est dans les cas les plus simples, une simple identité du résultat obtenu par l'entrée. Pour notre application nous avons opté pour le modèle multicouche, car ses avantages dans le module de reconnaissance sont plus frappants que ceux du modèle Kohonen.

Ces avantages se résument comme suit :

- La facilité de sa compréhension.
- Son apprentissage est supervisé, ce qui va nous permettre d'intervenir dans le choix des entrées et des sorties du réseau.
- Plusieurs réseaux multicouches utilisant la technique d'apprentissage ont été développés. Ces réseaux fournissent des résultats satisfaisants.

(voir chapitre 4.)

4.3.1. Le réseau Multicouche [21]

4.3.1.1. Définition

Cette architecture de réseau, appelée aussi « perceptron multicouche » (Multilayer perceptron), est un ensemble de neurones organisés en couches successives. Les éléments de traitement d'une même couche ne sont pas connectés entre eux, par contre ils sont connectés aux éléments du traitement des couches précédente et suivante.

Ce réseau est constitué de trois principales couches :

4.3.1.2. Couche d'entrée

Les neurones de cette couche n'effectuent aucun traitement sur les valeurs qui leur sont présentées, ils ne font que transmettre les informations qui leur parviennent de l'extérieur vers l'intérieur. Le nombre de neurones de cette couche dépend de l'application à laquelle répond le réseau de neurones.

4.3.1.3. Couche cachée

Elle contient une ou plusieurs couches cachées, dans chacune il y a plusieurs neurones (leur nombre est variable) le rôle des neurones de cette couche est de codifier de façon interne l'information présentée à l'entrée du réseau.

4.3.1.4. Couche de sortie

Un neurone de cette couche effectue un traitement sur les entrées provenant de la couche précédente (couche cachée) et délivre en sortie l'information utilisable par le système externe.

Le nombre de neurones de cette couche est déterminé par l'application.

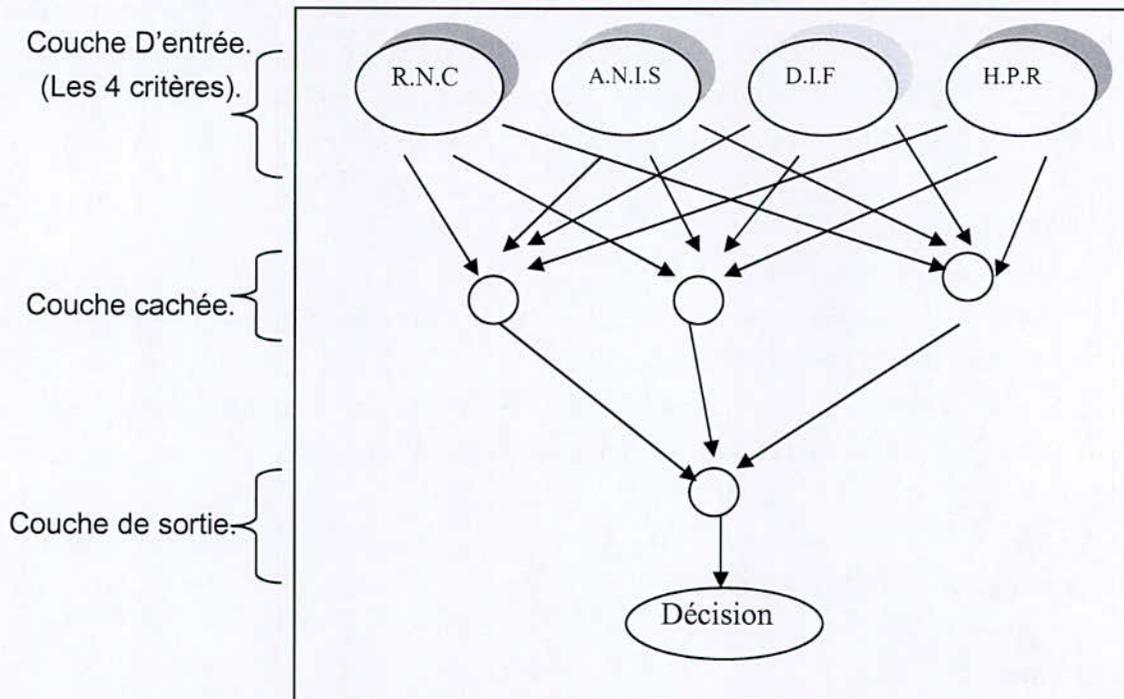


Figure 5.12 : Schéma du réseau de neurones adapté.

Structure du réseau :

Nous avons choisi le perceptron multicouche car par rapport au réseau de Kohonen qui donne lui aussi de bons résultats en classification, le fonctionnement du perceptron est plus simple à comprendre, son apprentissage est supervisé, ce qui va nous permettre d'intervenir dans le choix des entrées et des sorties du réseau.

Pour la couche d'entrée, nous avons mis 4 neurones qui correspondent aux 4 critères de malignité. La couche cachée possède 3 neurones, leurs poids sont une pondération des neurones de la couche d'entrée, tandis qu'en sortie un neurone va représenté la décision finale qui donnera l'état de la cellule : normale ou anormale. (Figure 5.12)

Le réseau va étudier chaque image introduite et la comparer aux images prototypes de la base de données pour déterminer son état.

La fonction de transfert :

Nous utilisons dans notre étude la fonction sigmoïde. Les caractéristiques de cette fonction sont :

Chapitre 5 : Le travail effectué

- Contenues sur l'ensemble des valeurs réelles.
- Ses valeurs sont réelles.
- Sa dérivée est toujours positive.
- Elle est bornée.

Les intervalles des valeurs de sortie les plus utilisés sont $[0,1]$ et $[-1,1]$; l'approche neuronale utilise fréquemment cette fonction comme fonction d'activation. Nous avons utilisé la fonction logistique comme fonction sigmoïde.

La fonction logistique est définie par :

$$F : \mathbb{R} \rightarrow]0,1[$$
$$F(x) = 1 / (1 + \exp (-x)).$$

L'erreur de sortie minimale du réseau de neurones est établie avec une base d'apprentissage. Ainsi les exemples de la base d'apprentissage seront très variés afin d'éviter le problème de sur-apprentissage.

- ◆ Nous avons fixé les poids de la couche d'entrée arbitrairement dans l'intervalle $[-1,+1]$, ces poids sont respectivement pour chaque neurone de la couche d'entrée : $\{0.25, -0.12, 0.365, 0.058\}$
- ◆ Chacune des entrées de la couche cachée est une sommation pondérée par les différents poids de cette dernière. On a donc trois séries de quatre poids qui sont affectés respectivement à chaque neurone de la couche cachée :

0.35	-0.48	0.14	-0.24
0.63	0.15	-0.45	0.35
0.25	0.45	0.69	-0.11

- ◆ De la même manière, nous avons trois séries de trois poids qui sont affectés respectivement à chaque neurone de la couche de sortie :

0.21	0.25	0.36
-0.54	0.236	0.147
0.25	0.14	0.26

- ◆ à la fin de l'apprentissage, nous avons trouvé les poids « couche cachée – couche de sortie » suivants pour chaque neurone :

Chapitre 5 : Le travail effectué

Sortie1: 1.139 1.179 1.289

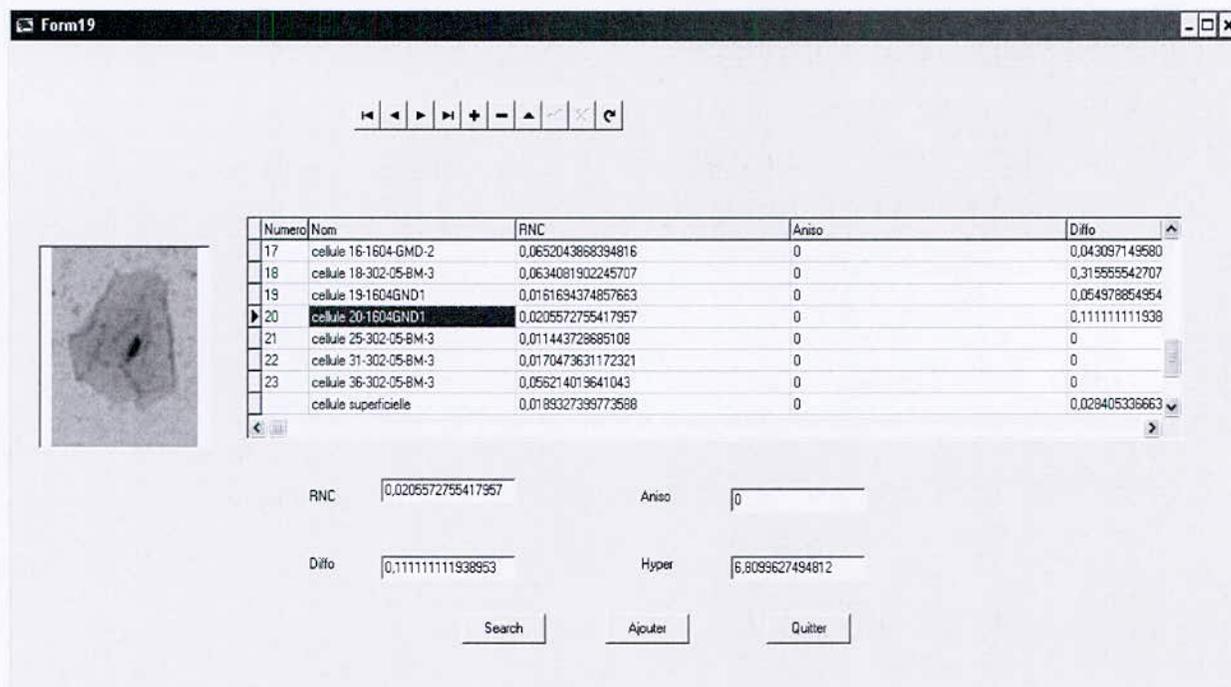
Sortie2: 0.770 1.546 1.457

Sortie3: 0.450 0.340 1.457

- ◆ Ces poids vont être stockés dans la base d'apprentissage dans le cas où l'image des cellules serait une image témoin (image pour laquelle on connaît la nature des cellules au préalable qui sont normales). L'observation ou la détection des anomalies dans le diagnostique. On compare les poids de la couche cachée et la couche sortie avec ceux qui sont stockés dans la base de donnée. Les résultats d'un diagnostique anatomo-cytologique en macroscopie engendrent à l'obtention de deux formes tissulaires ; Tissu anormal (masse) ou tissu normal (non masse).

Nous avons constitué une base de données à partir de lame fourni par l'hôpital de Belfort sur une vingtaine de lame nous avons pu récupérer une trentaine de cellules utilisables et bien déterminées telque :

Nous avons obtenu 25 cellules saines , normales et 5 anormales , non reconnues



The screenshot shows a software window titled "Form19" with a toolbar at the top. Below the toolbar is a table with the following data:

Numero	Nom	RNC	Aniso	Diffo
17	cellule 16-1604-GMD-2	0,0652043868394816	0	0,043097149580
18	cellule 18-302-05-BM-3	0,0634081902245707	0	0,315555542707
19	cellule 19-1604GND1	0,0151694374857663	0	0,054978854954
20	cellule 20-1604GND1	0,0205572755417957	0	0,111111111938
21	cellule 25-302-05-BM-3	0,011443728685108	0	0
22	cellule 31-302-05-BM-3	0,0170473631172321	0	0
23	cellule 36-302-05-BM-3	0,056214013641043	0	0
	cellule superficielle	0,0189327399773588	0	0,028405338663

Below the table, there are input fields for "RNC" (0,0205572755417957), "Aniso" (0), "Diffo" (0,111111111938953), and "Hyper" (6,8099627494812). At the bottom, there are buttons for "Search", "Ajouter", and "Quitter". To the left of the table, there is a small image of a cell.

Figure 5.13 : représentation de la base de donnée

Nous avons pu constituer un tableau contenant le nom de la cellule ainsi que l'image associée et tout les critères calculés. (Figure 5.13)

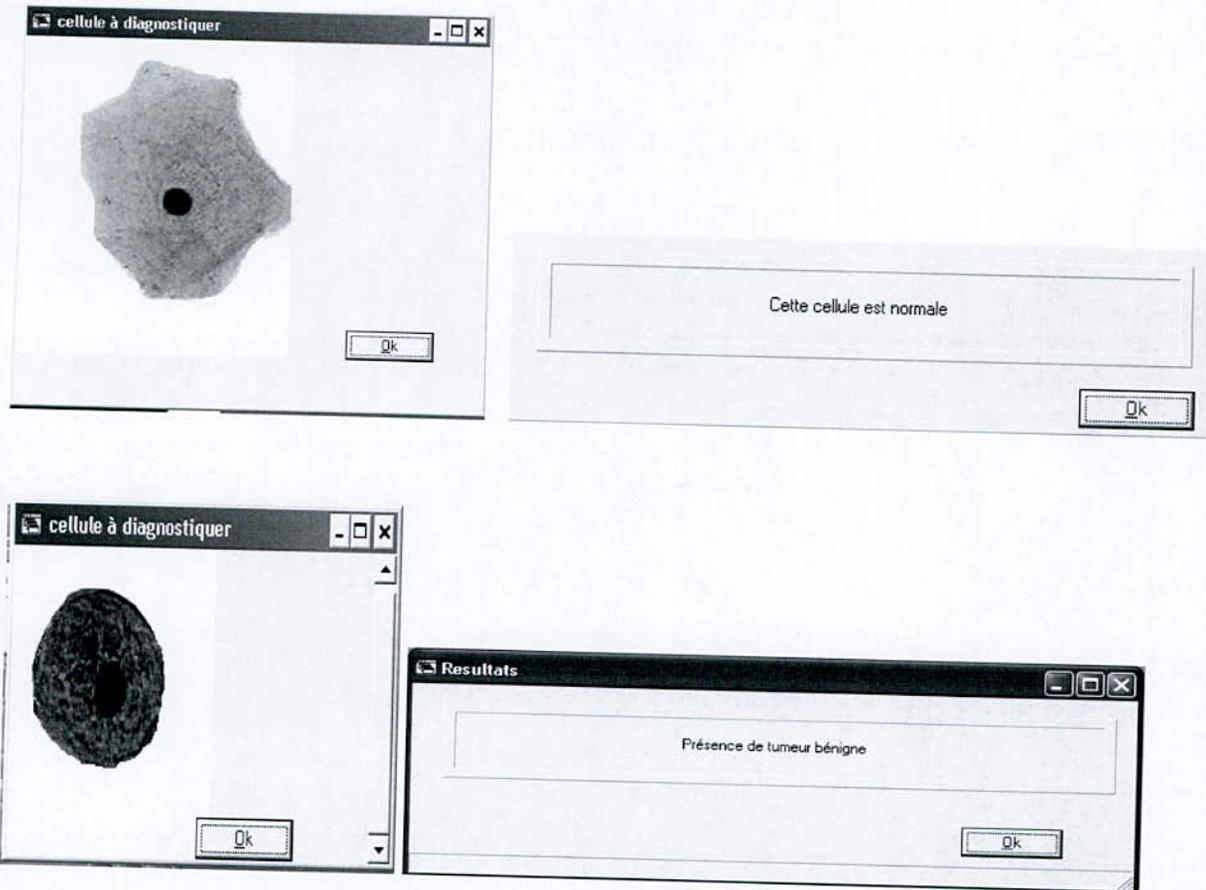


Figure 5.14 : Affichage du résultat de la reconnaissance

Nous avons obtenus un taux de 67% de cellules reconnues.

5. Description du logiciel :

Nous allons présenter le logiciel réalisé dans le cadre de notre projet de magister. Notre logiciel a été réalisé grâce à l'outil de programmation Builder C++ version 5 sous environnement Windows XP, il se nomme « DIAG-CANCER » : Diagnostic du cancer. Le choix du langage est justifié par la souplesse et la simplicité avec laquelle celui-ci supporte la programmation orientée objet, et par le besoin d'un environnement de programmation évolué, permettant une description rapide des algorithmes en POO (Programmation Orientée Objet).

Dans ce qui suit nous présenterons la fenêtre principale de l'application et les menus secondaires avec leurs différentes opérations.

5.1. Description de l'application :

L'interface du logiciel est présentée ci-dessous par la figure suivante :

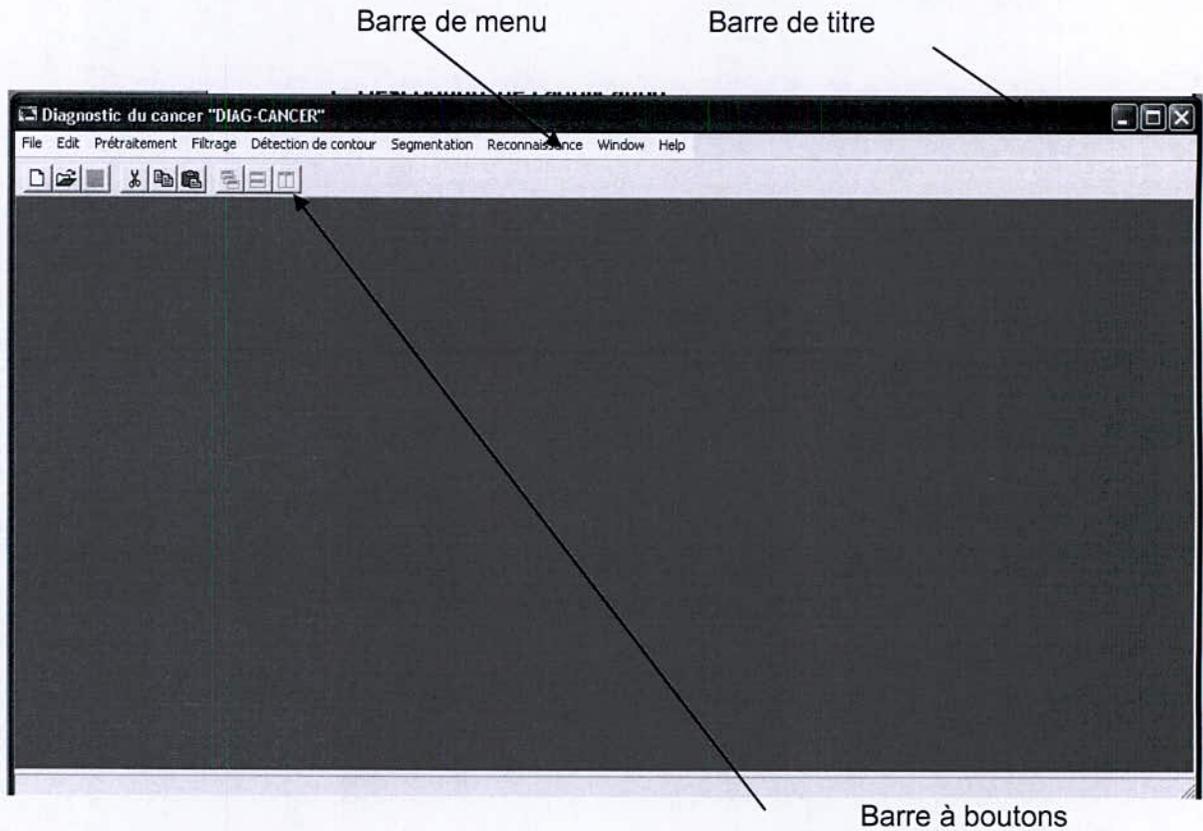
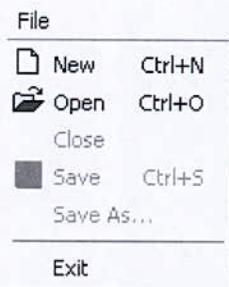


Figure 5.15 : L'interface du logiciel

✚ La barre des menus :

C'est une barre contenant des articles que l'on peut sélectionner à partir de la souris. L'activation d'un article fait appel soit à une fonction, soit à une boîte de dialogue.

❖ *Menu File :*



- **Commande Ouvrir (Open) :**

A l'exécution de cette commande une boîte de sélection apparaît permettant à l'utilisateur de choisir le nom du fichier image, d'extension (*.bmp), à ouvrir.

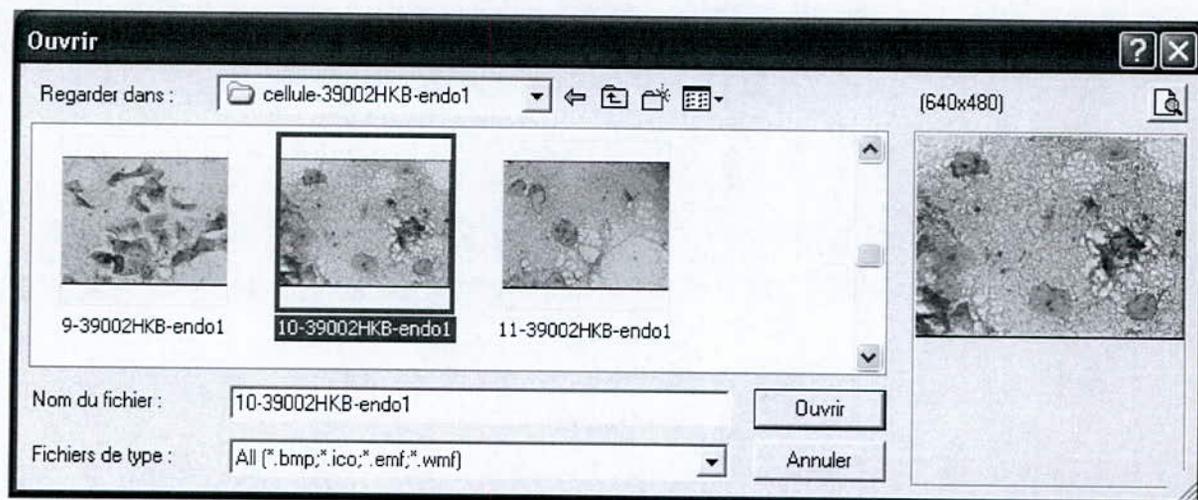
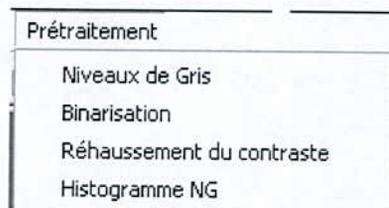


Figure 5.16 : Boite de dialogue pour ouvrir

- ❖ **Menu Prétraitement :**



- **Commande Niveaux de gris :**

Après avoir ouvert notre image, elle peut être en couleurs, on lui applique cette commande pour la rendre en niveaux de gris.

- **Commande HistogrammeNG :**

Cette commande calcule l'histogramme des niveaux de gris.

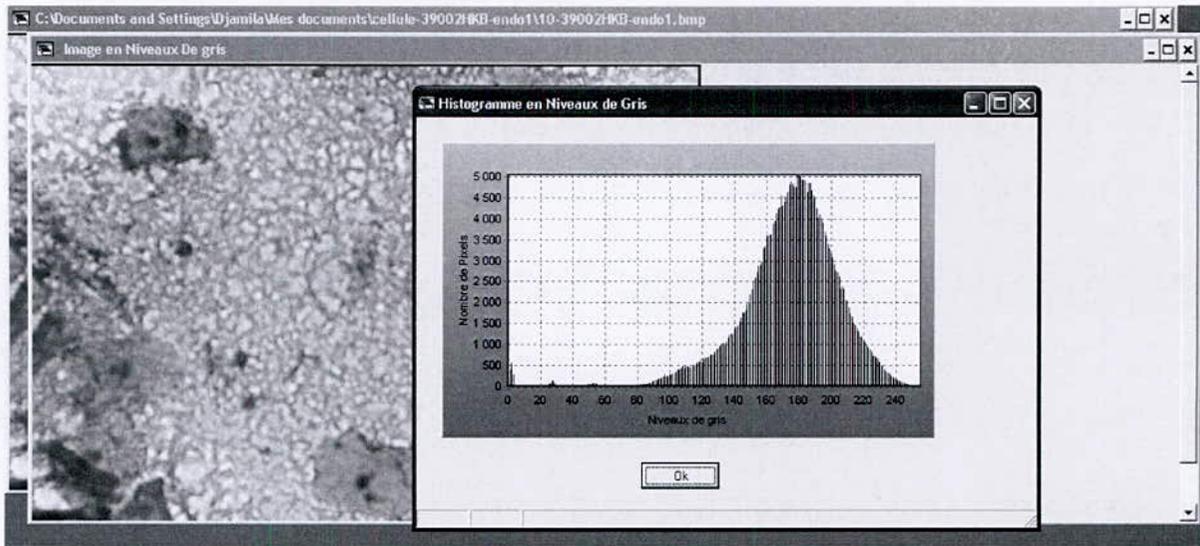


Figure 5.17: L'histogramme des niveaux de gris

▪ **Commande Binarisation :**

La binarisation est l'étape qui consiste à transformer l'image en seulement deux niveaux de gris le blanc et le noir et ceci se fait par un calcul automatique du seuil ou bien manuellement. Le choix de la méthode de calcul se fait grâce à la boîte de dialogue présentée ci-dessous :

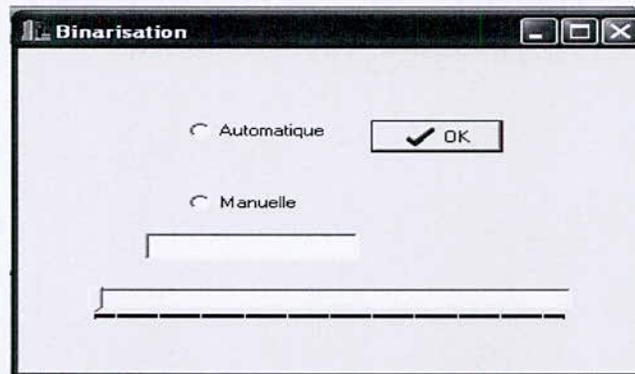


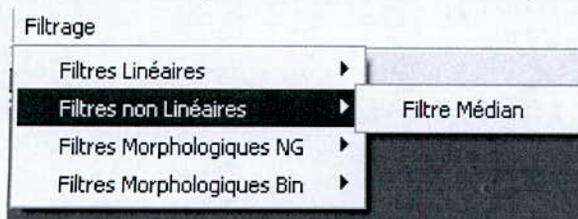
Figure 5.18 : boîte de dialogue pour la binarisation

❖ **Menu Filtrage:**

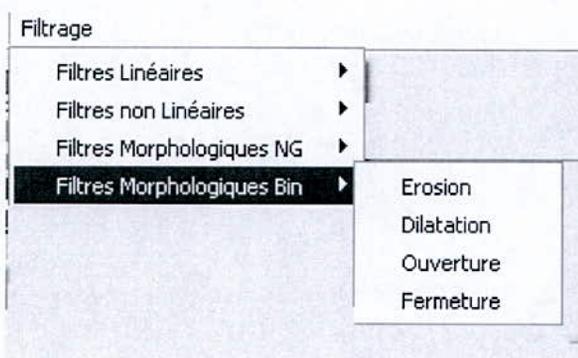
Ce menu se compose de deux sous menus déroulants : le filtre non linéaire, les filtres linéaires et les filtres morphologiques en niveaux de gris et en noir et blanc.

▪ Les filtres non linéaires :

Nous avons implémenté le filtre médian pour éliminer le bruit dans l'image.

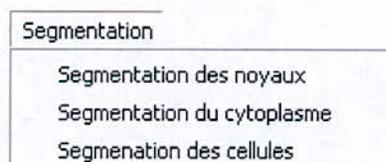


▪ Les Filtres morphologiques :



❖ **Menu Segmentation :**

Ce menu contient la phase de segmentation des noyaux, du cytoplasme et des cellules .



▪ La commande segmentation des noyaux :

Cette commande permet d'extraire les noyaux des cellules dans une image à part.

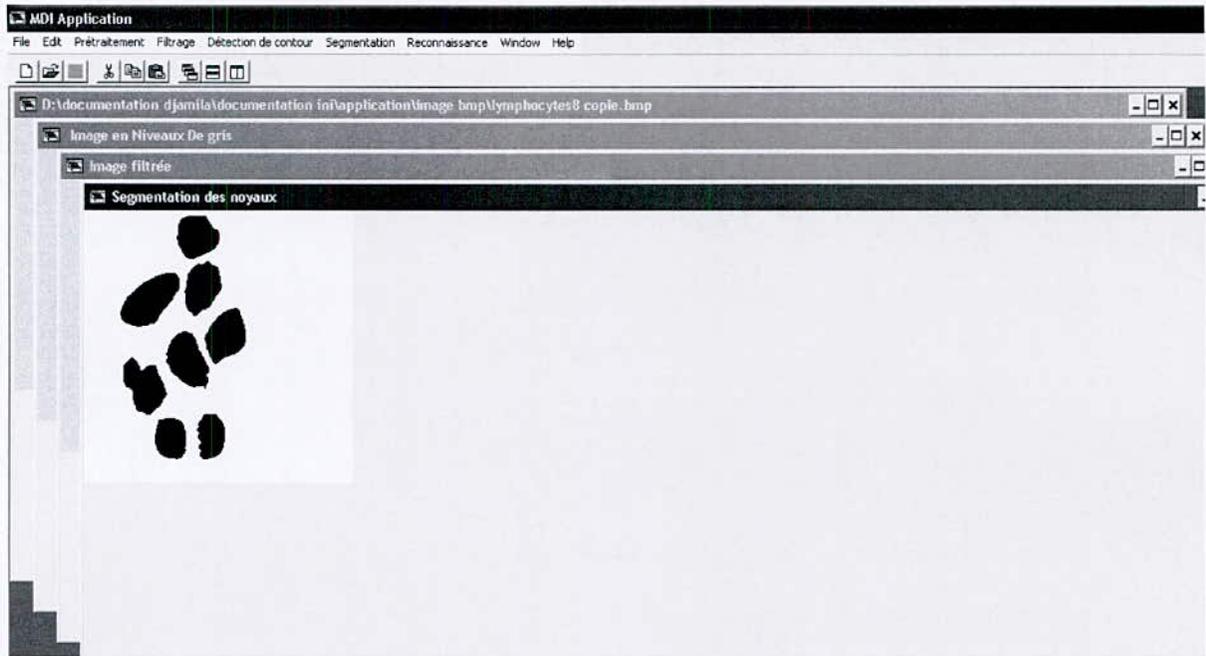


Figure 5.19 : Segmentation des noyaux

▪ La commande segmentation du cytoplasme :

Cette commande permet d'extraire le cytoplasme, ceci toujours par l'approche multifractale.

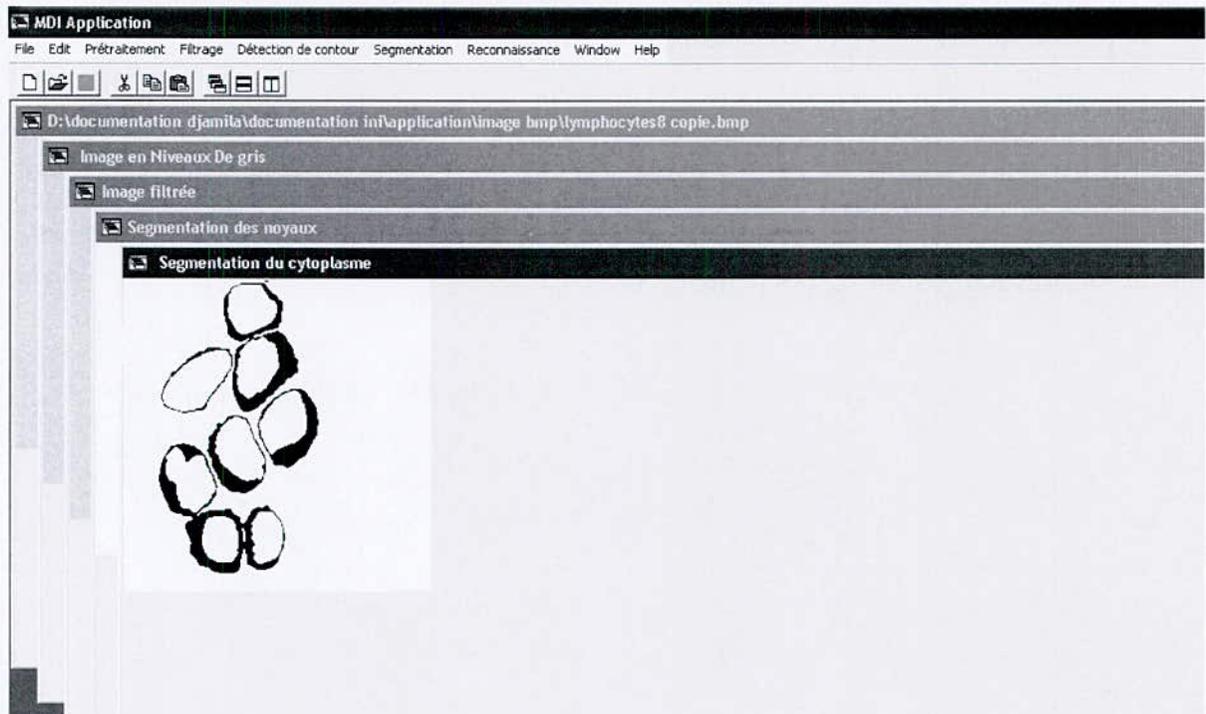


Figure 5.20 : Segmentation du cytoplasme

- La commande segmentation des cellules :
A l'aide de cette commande on obtient l'image de cellule

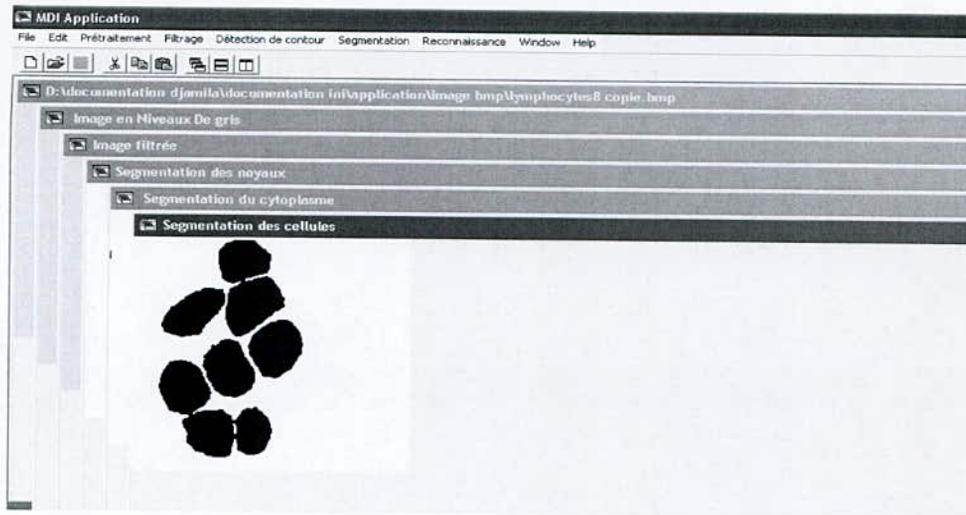


Figure 5.21 : Segmentation des cellules

Récapitulatif :



Figure 5.22 : Récapitulatif

❖ **Menu Reconnaissance :**

Ce menu contient l'étape finale de notre travail, on calcul les critères de malignité pour chaque cellule de l'image, on détermine le type de chaque cellule à l'aide du calcul du RNC et finalement on décide de l'état de la cellule .

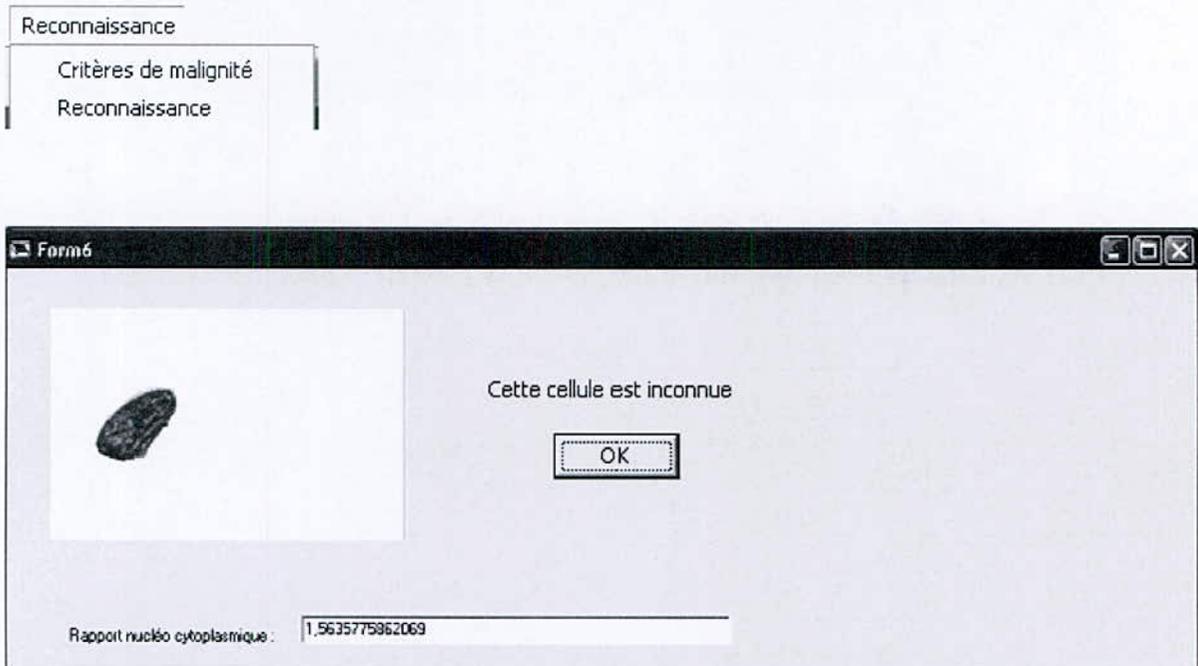


Figure 5.23 : Affichage de la boîte de résultat calcul des critères de malignité

Si l'on se réfère au rapport nucléo-cytoplasmique affiché sur l'image précédente, qui est supérieur à 1, cette cellule n'a pas été reconnue.

De ce fait, il y a une possibilité de présence de tumeur sur cette cellule, mais cette décision pourra être vérifiée après le calcul des critères de malignité restants.

Conclusion :

Ce chapitre a été consacré à la présentation des résultats d'application des différents algorithmes étudiés, ces algorithmes englobent le prétraitement, le filtrage, la segmentation, ainsi que la classification et la reconnaissance. nous avons fait une description de notre application, en donnant des exemples et des visualisations sur écran des résultats obtenus.

Notre but, étant de mettre en œuvre une application autonome et simple à l'utilisateur, nous avons essayé d'exploiter les fonctionnalités offertes par le Builder.

Conclusion Générale

Conclusion Générale

Le cancer est considéré comme l'une des maladies les plus dangereuses dans le monde, menant à une augmentation remarquable du taux de mortalité; ce qui rend son dépistage précoce un véritable défi.

Ce défi a poussé les chercheurs à exploiter l'outil informatique dans le domaine médical afin de diminuer le temps du dépistage ainsi que le risque d'erreurs éventuellement commis par le médecin.

Pour essayer d'aboutir aux résultats escomptés, une étude théorique sur la cytologie du col utérin a été faite, ce qui nous a permis de déterminer les critères de malignité qui sont: le rapport noyau-cytoplasmique, l'hyperchromatisme nucléaire, la difformité du noyau et l'anisocaryose.

Concernant le traitement des images des cellules, nous avons commencé par l'étape d'acquisition, réalisée à l'aide d'un microscope et une caméra numérique; suivie de l'étape de prétraitement qui consiste à éliminer les informations indésirables. Enfin, l'étape de filtrage qui permet l'élimination du bruit, pour laquelle nous avons opté pour le filtre médian que nous a donné de meilleurs résultats, que ceux obtenus par l'utilisation des filtres linéaires classiques.

Pour la segmentation des images traitées, nous avons utilisé l'approche multifractale qui se base sur les exposants de singularité, ainsi que les contours actifs. Les résultats de la segmentation sont : une image de noyau, une de cytoplasme et une autre de cellule.

Ces résultats sont utilisés comme entrées à l'étape de classification, dont le rôle est la **séparation des cellules en leurs quatre couches essentielles** (la superficielle, l'intermédiaire, la basale et la para-basale), par l'utilisation de l'algorithme de croissance de régions.

Conclusion Générale

Après l'étape de classification, nous avons présenté l'étape de reconnaissance qui permet de déterminer l'état de la cellule (normale ou anormale). Cette étape a été réalisée en utilisant les réseaux de neurones multicouches, et la base de connaissances déjà créée par le module d'apprentissage.

Pour illustrer tout ce qui a été étudié, nous avons présenté les résultats de l'application des différents algorithmes sur quelques images.

Enfin, nous espérons que notre étude répond au mieux aux besoins demandés; en prouvant, en quelque sorte, que l'application de l'outil informatique peut être très bénéfique pour les autres sciences.

Comme pour tout autre logiciel, le notre reste ouvert à d'éventuelles améliorations parmi lesquelles, nous pouvons citer:

- Traitement des images de cellules, en couleurs.
- Implémentation d'une méthode de segmentation hybride
- Généraliser notre application à d'autres organes.
- Donner un résultat plus précis quant au type d'anomalie en faisant une étude plus approfondie conjointement avec les médecins.

Bibliographie

Bibliographie

- [1] A. YAKER , Cancérologie générale, anatomie pathologique , OPU, 1985
- [2] J.M. VIGNAUD, Cours de Pathologie tumorale, 2004
- [3] M . MAILLET , Cytologie gynécologique normale et paranormale , PICCIN, 1991
- [4] N. DALY SCHVEITZER , abrégé de cancérologie clinique , MASSON, 1998
- [5] C .GOMPEL, Atlas de cytologie clinique, MALOINE S.A. éditeur, 1982
- [6] M.KUNT, traitement numérique des images , PPUR, 1993.
- [7] O.PAVIE, Photos, scans et retouches sous Windows 95/98 , CAMPUS , 1999.
- [8] J.J.TOUMAZET, Traitement de l'image sur micro-ordinateur, SYBEX, 1987.
- [9] J-P. COCQUEREZ, Analyse d'image : Filtrage et segmentation, MASSON,1995
- [10] M.KUNT, Reconnaissances des formes et des scènes, PPUR, 1993.
- [11] : Traitement d'images : <http://pczenith.univ-mlv.fr/jf/>
- [12] : B. MANDELBROT , Les objets fractals , Flammarion, 1995.
- [13] J F GOUYET, physique et structures fractales, MASSON, 1992
- [14] :Segmentation des images : <http://perso.magic.fr/ormery/fractales/fractal.html>
- [15] J.P. BERROIR , Application des mesures multifractales à l'analyse et la segmentation d'images médicales , Rapport d'INRIA, N°1345, pp : 979-987.

Annexes

- [16] N. LASSOUAOUI, Segmentation des images biomédicales par des approches fractales, Conception d'un logiciel de traitement et de segmentation des images , Thèse de Magister, Ecole Nationale Polytechnique d'Alger, 2000.
- [17] CAVARO-MENARD C, EA Signaux et Images en Biologie et Medecine : Traitement d'Images , Université d'Angers
- [18] C. CHESNAUD , techniques statistiques de segmentation par contour actif et mise en oeuvre rapide, Thèse de doctorat ,Marseille, 1999
- [19] E. HUEBER , Segmentation d'images par contour actif: implantation optique avec un corrélateur incohérent ombroscopique, Thèse de doctorat, Université de Haute-Alsace, 1992
- [20] G . DREYFUS, Réseaux de neurones : Méthodologie et applications, Eyrolles ,2002
- [21] J.F JODOUIN, Les réseaux de neurones : principes et définitions, Hermès, 1994.
- [22] J. HERAULT, Réseaux neuronaux et traitement du signal , Hèrmes ,1994
- [23] Quelques types de réseaux de neurones www.asi.insa-rouen.fr/pleray/ftp/rn.pdf
- [24] E. DAVALO , Des réseaux de neurones, Eyrolles 1991
- [25] N. LASSOUAOUI , Segmentation des images par différentes approches et optimisation avec les algorithmes génétiques. Thèse de doctorat en électronique, ENP, 2004
- [26] A. ARNEODO, Ondelettes, multifractales et turbulences de l'ADN aux croissances cristallines, Arts et Sciences, 1995
- [27] L.HAMAMI & N.LASSOUAOUI, Genetic algorithms and Multifractal segmentation of cervical cell images, ISSPA'03, Vol.2, p.1-4, Paris June 2003
- [28] N. LASSOUAOUI, L. HAMAMI, A. ZERGUERRAS, Segmentation and classification of biological cell images by a multifractal approach, International Journal of Intelligent systems, Vol°18, N°6, pp: 657-678, Wiley Edition, June 2003

Annexe

Annexe

1

Outils et notions fondamentales sur le traitement d'images

1. Caractéristiques principales des images numériques :

1.1 Pixel :

Abréviation de picture Element. Le pixel constitue le plus petit élément de l'image. Il est caractérisé par ses coordonnées dans l'image et par son intensité lumineuse (couleur ou niveau de gris).

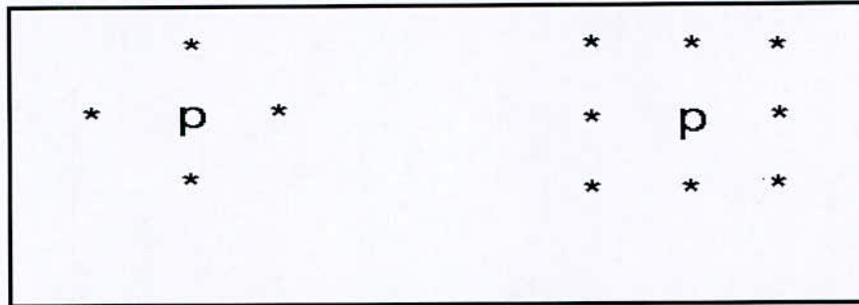
1.2 Le voisinage d'un pixel :

Le voisinage d'un pixel est composé de tous les pixels qui l'entourent immédiatement. Si p est un pixel d'une image D , alors le voisinage de p est le plus petit sous-ensemble de D qui contient p .

Annexes

Dans une image numérique, on distingue deux types de connexités relatives au voisinage utilisé : la 4-connexité et la 8-connexité.

Le voisinage d'un pixel (i,j) est dit 4-connexe s'il est formé des quatre pixels de coordonnées spatiales (i+1,j),(i,j+1),(i-1,j),(i,j-1), et il est dit 8-connexe s'il est formé des pixels de coordonnées spatiales (i+1,j),(i-1,j),(i,j+1),(i,j-1),(i+1,j-1),(i-1,j+1),(i-1,j-1).



Voisinage 4-connexes

Voisinage 8-connexes

Figure 1: Voisinage d'un pixel.

1.3 Le contour :

Il est défini comme étant la frontière entre les objets de l'image. Les pixels qui décrivent les contours représentent une brusque variation de couleur ou de niveau de gris par rapport aux pixels voisins.

1.4 Le contraste :

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images. Si L1 et L2 sont les degrés de luminosité respectivement de deux zones voisines A1 et A2 d'une image, le contraste C sera défini par

$$\text{le rapport : } C = \left| \frac{L1 - L2}{L1 + L2} \right|$$

1.5 Niveau de gris :

Dans une image en niveau de gris (nuances de gris), la couleur d'un pixel peut prendre des valeurs allant du noir au blanc, en passant par un nombre fini de niveaux intermédiaires. Dans une telle image les intensités du rouge, du vert, et du bleu de chaque pixel sont égales.

1.6 Bruit d'image :

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient soit du dispositif d'acquisition (caméra, scanner,...) soit de la scène elle-même (poussières, rayures,...).

1.7 Histogramme :

L'histogramme des niveaux de gris d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris dans l'image (figure II.4). En abscisse sont représentés les niveaux de gris de 0 à N (255 dans notre cas). En ordonnée est affiché le nombre de pixels affecté par chaque niveau de gris. L'histogramme permet de donner un grand nombre d'informations sur la distribution des niveaux de gris de l'image. Il permet notamment d'analyser entre quelles bornes est répartie la majorité des niveaux de gris dans le cas d'une image très claire ou très foncée. Comme illustré Figure II.4 :

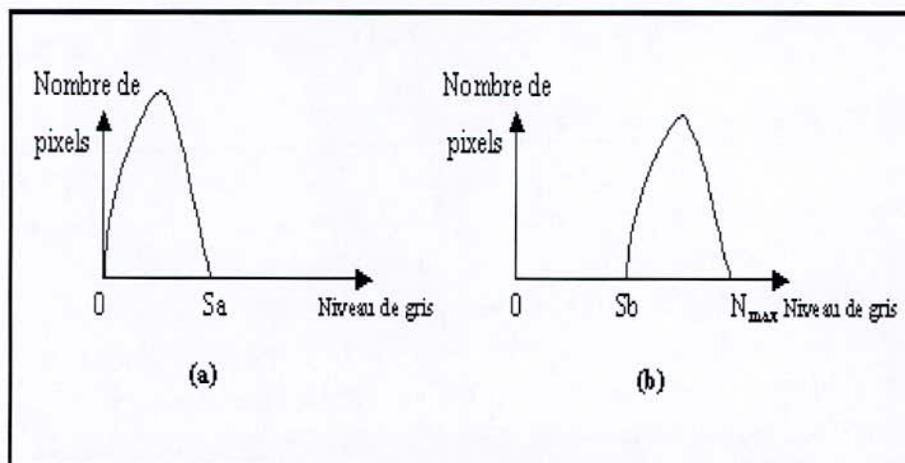


Figure 2 : Histogramme d'image. a : trop foncée, b : trop claire.

1.8 Modification d'histogramme :

Certaines images sont initialement trop claires, trop foncées, ou bien peu contrastées. Cela est dû respectivement au fait que les niveaux de gris de l'image sont tassés vers le haut de l'échelle, vers le bas, ou bien sont regroupés dans un intervalle étroit.

Ce défaut est très visible sur l'histogramme, le but est de redistribuer les niveaux de gris de l'image afin de leur faire occuper toute la bande de nuances possibles.

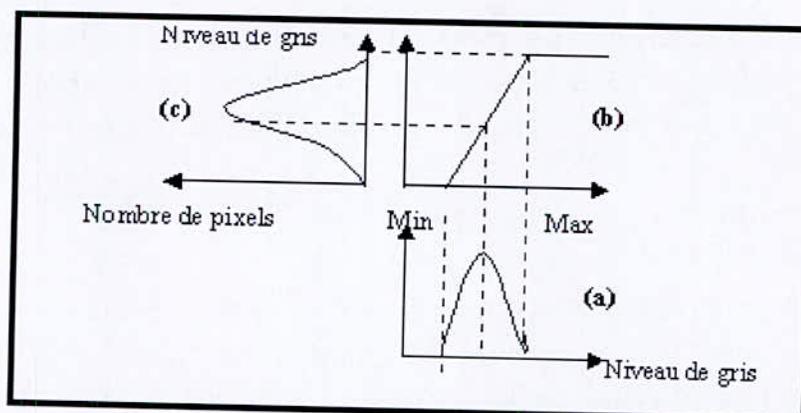


Figure 3 : Principe de la modification d'histogramme

a) Histogramme de l'image originale. b) Fonction de transformation.

c) Histogramme recadré.

Cette méthode est basée sur les transformations ponctuelles d'intensité. C'est à dire, à tout pixel d'intensité a_s on associe une intensité $a'_s = T(a_s)$.

2. Filtrage d'image

Nous avons vu qu'avec les méthodes de transformations ponctuelles chaque pixel était transformé indépendamment de la valeur de ses voisins.

Nous avons également défini le bruit comme un phénomène de brusque variation d'un pixel isolé par rapport à ses voisins. Or une image possède une certaine redondance spatiale : les pixels voisins ont généralement les mêmes caractéristiques. Il est donc nécessaire d'opérer des transformations qui pour chaque pixel tiennent compte de son voisinage. Les méthodes parmi les plus utilisées sont les techniques dites de *filtrage* de l'image.

On distingue : le filtrage linéaire, non linéaire, et morphologique.

2.1 Filtrage linéaire :

La transformation d'un pixel est le fruit d'une combinaison linéaire des pixels voisins. Parmi les filtres linéaires on définit, le filtre moyen :

▪ Filtre moyen :

Il consiste à considérer chaque pixel de l'image et d'en faire la moyenne avec les huit pixels qui lui sont voisins. Ceci a pour effet d'adoucir l'image en réduisant les fluctuations des niveaux de gris. Ce type de filtre peut être mis sous la forme d'un masque tel que celui-ci :

$$H = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \dots\dots\dots (1)$$

Ce masque est déplacé sur toute l'image, le pixel affecté par la transformation étant le pixel central du masque. Le facteur 1/9 est égal à la somme des coefficients du masque et sert à normaliser le filtre de manière à ce que celui-ci n'influe pas sur l'intensité globale de l'image. Si nous posons I_1 comme étant l'image de départ, I_f l'image résultat, toutes deux de taille $p \times p$, et H_1 le masque, nous avons alors pour chaque pixel de coordonnées x, y :

$$I_f(x,y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 H_1(i+1,j+1) \cdot I_1(x+i,y+j) \dots\dots\dots (2)$$

2.2 Filtrage non linéaire :

Les pixels voisins interviennent suivant une loi non linéaire. L'exemple le plus classique est celui du filtre médian de Tuckey.

▪ Filtre Médian :

Le filtrage médian consiste à remplacer un pixel par la valeur médiane de ses voisins.

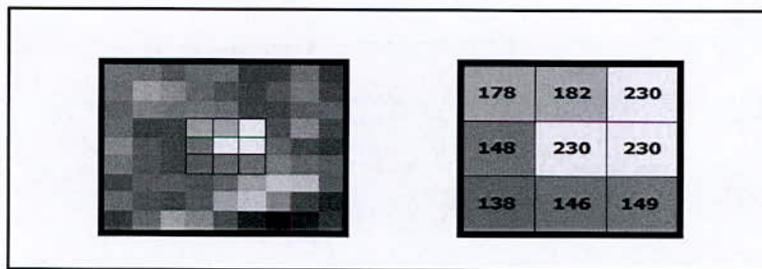


Figure 4 : a) Portion agrandie d'une image bruitée. b) niveau de gris des pixels voisins.

- La correction du pixel central de l'image de la figure II.6 se ferait donc de la façon suivante :
- Classer toutes les valeurs de luminance en ordre croissant : 138, 146, 148, 149, 178, 182, 230, 230, 230
 - Remplacer le pixel par la valeur médiane du classement. Dans ce cas 230 (intensité du pixel central) sera remplacée par 178

2.3 Opérateurs morphologiques :

Le filtrage morphologique est réalisé par une combinaison de traitements linéaires et non linéaires. En appliquant ce type de filtrage à des images binaires -noir et blanc- tel que chaque pixel est codé comme un élément logique, donc à valeur vraie ou fausse. Il est ainsi l'objet d'une transformation par une fonction logique.

La dilatation :

Elle consiste à dilater l'image. De ce fait, les points noirs isolés au milieu de parties blanches sont éliminés. Le processus est réalisé en balayant sur l'image une fenêtre carrée de taille $(N+1)*(N+1)$ et en effectuant pour chaque pixel de l'image le OU logique des $(N+1)^2-1$ voisins.

Les points noirs sont éliminés, mais la taille des objets présents dans l'image diminue.

L'érosion :

Le procédé est dual à la dilatation. On effectue cette fois le ET logique des $(N+1)^2-1$ voisins. De ce fait, les points blancs isolés au milieu de parties noires sont éliminés mais des pixels noirs sont ajoutés au contour des objets présents dans l'image.

L'ouverture :

L'ouverture est constituée par une opération d'érosion suivie d'une dilatation. Elle permet de retrouver la taille normale des objets de l'image.

La fermeture :

C'est une opération de dilatation suivie d'une érosion.

3. La binarisation

La binarisation consiste en l'obtention à partir d'une image à plusieurs niveaux de gris, une image à deux niveaux de gris (0 pour le noir et 255 pour le blanc). Ceci se fait en choisissant un seuil $S1$ pour lequel :

$$F(L(i,j)) = \begin{cases} \text{Noir si } L(i,j) > S1 \\ \text{Blanc si } L(i,j) < S1 \end{cases}$$

$L(i,j)$ est le niveau de gris du pixel (i,j)

Le choix du seuil S1 se fait avec l'utilisation de l'histogramme des niveaux de gris, il s'agit du minimum local entre deux pics.

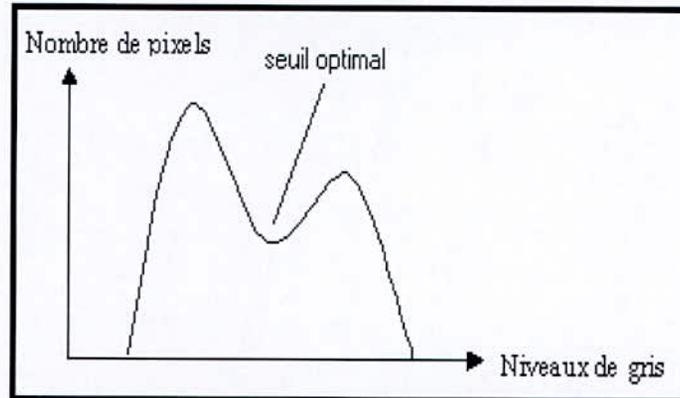


Figure 5 : Seuil de binarisation.

4. Le rehaussement du contraste

On considère les images à niveaux de gris compris entre m et M ($m=0$ et $M = 255$ dans le cas de nos images). On dit qu'une image est bien contrastée si les luminosités des pixels de l'image sont au mieux réparties sur l'intervalle des niveaux de gris $[m,M]$. Pour traiter le contraste d'une image, on applique uniformément à tous les pixels de celle-ci une transformation de niveaux de gris T_g . En d'autres termes, à partir d'une image I associant à tout pixel p le niveau de gris $g_e(i, j)$, on obtient une nouvelle image J dont le niveau de gris au pixel p est $J(p)=I(p)=g_s(i, j)$.

Le but des procédures de rehaussement du contraste est d'améliorer leur qualité visuelle qui facilite leur interprétation.

Les opérations de rehaussement de contraste peuvent être décrites de la façon suivante:

$$g_s(i, j) = T_g g_e(i, j)$$

$g_s(i, j)$ Le niveau de gris d'un pixel de l'image de sortie

$g_e(i, j)$ Le niveau de gris du même pixel de l'image d'entrée

T_g Fonction mathématique de transformation

5. Détecteur de contour

Un contour, dans la scène, correspond généralement aux changements des objets. Dans une image, on définit le contour comme une brusque variation de niveau de gris.

L'extraction de contour est réalisée en dérivant l'image une fois ; on obtient alors un gradient, ou bien deux fois, et on obtient un laplacien.

5.1. Opérateurs dérivatifs du premier ordre (Gradient) :

5.1.1. L'opérateur de Roberts :

Parmi les opérateurs de dérivation, le premier développé fut celui de Roberts en 1965, il n'est autre qu'une application directe de la formule d'une dérivée. Il se présente sous la forme suivante :

Pour chaque pixel $I(x,y)$ d'une image I , les dérivées en x et y sont :

$$A_x = I(x+1, y) - I(x, y).$$

$$A_y = I(x, y+1) - I(x, y).$$

Ce qui revient à convoluer l'image avec les masques :

$$H_x = (-1, +1) \quad \text{et} \quad H_y = \begin{pmatrix} -1 \\ +1 \end{pmatrix}$$

L'inconvénient de cet opérateur est sa sensibilité au bruit.

5.1.2. L'opérateur de Sobel et Prewitt :

Contrairement à l'opérateur de Roberts, il présente l'avantage d'être moins sensible au bruit. En effet, tout processus de dérivation d'un signal tend à accentuer le bruit présent dans ce signal ; cette sensibilité a été diminuée en effectuant une moyenne locale sur le domaine couvert par le masque.

$$\text{Opérateur Prewitt: } H_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}; \quad H_y = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\text{Opérateur Sobel: } H_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}; \quad H_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

5.2. Opérateurs dérivatifs du deuxième ordre :

5.2.1. Le Laplacien :

Comme son nom l'indique, il consiste à calculer la dérivée seconde de l'image. Il existe plusieurs masques du laplacien, tous très similaires tels que :

$$H = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} ; H = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} ; H = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

On obtient après l'application du laplacien sur l'image, un résultat où les points du contour correspondent aux valeurs nulles, bordées de deux extrémités. Le contour est ensuite déterminé par la détection des passages par zéro du laplacien de l'image.

Notons que le passage par zéro étant unique, le laplacien fournit directement des contours d'un pixel d'épaisseur. Le laplacien possède, cependant, un inconvénient majeur qui est sa grande sensibilité au bruit.

Annexe

2

Glossaire

Anatomo-pathologique : L'étude des modifications apportées par la maladie aux différents organes, tissus et cellules.

Anisocytose : Inégalité de taille des cellules, les unes par rapport aux autres. S'observe dans les dysplasies.

Anisocaryose : Inégalité de taille des noyaux des cellules, les uns par rapport aux autres. S'observe dans les dysplasies.

Basales (cellules basales) et parabasales : Cellules provenant des couches profondes de l'épithélium malpighien exocervical. En période d'activité génitale, en l'absence d'ulcération ou d'inflammation, elles ne sont pas recueillies sur le frottis, la spatule ne ramenant que les éléments cellulaires des couches superficielles et intermédiaires. Elles s'observent sur le frottis en cas d'atrophie chez la femme ménopausée. Il est alors fréquent d'observer des remaniements dystrophiques parfois inquiétants cytologiquement, justifiant un frottis de contrôle après estrogénothérapie locale qui aura souvent pour effet de faire croître la muqueuse et de la régulariser.

Bénigne (Tumeur bénigne): Ces tumeurs ne sont pas cancéreuses et les cellules dont elles sont formées ne se propagent pas ailleurs dans le corps.

Bibliographie

Cytologie : Partie de la biologie générale qui étudie la cellule vivante sous tous ses aspects.

Dysplasie : Etymologiquement signifie "construction perturbée". Il s'agit de lésions précurseurs du cancer infiltrant du col caractérisées par un trouble de la croissance et de la différenciation épithéliale associant des anomalies architecturales (appréciées en histologie) et cytologiques (visibles sur le frottis). Par définition, la lésion est cantonnée à l'épithélium (elle reste intra-épithéliale) et ne franchit pas la membrane basale.

Hyperchromatisme : Dans une cellule normale, le noyau possède une chromatine fine. Dans les dysplasies et les carcinomes, la chromatine se dispose en mottes irrégulières ; elle devient plus dense, hyperchromatique.

Intermédiaires (cellules) : Cellules provenant des couches moyennes de l'épithélium malpighien exocervical. Elles sont nombreuses en deuxième partie de cycle (imprégnation lutéale) et au cours de la grossesse (cellules naviculaires).

Lésion malpighienne intra-épithéliale : Terminologie cytologique utilisée pour désigner les anomalies épithéliales malpighiennes.

Mycose : Les infections à candida se traduisent par la présence de filaments segmentés et/ou de spores sur un fond plus ou moins inflammatoire. Il s'y associe parfois des remaniements dystrophiques avec dyskératose pouvant justifier des frottis de contrôle après traitement.

Normal : Un frottis normal est un frottis correctement prélevé, bien fixé, suffisamment cellulaire et dont le fond ne gêne pas l'interprétation. Un fond hémorragique ou inflammatoire peut masquer les cellules épithéliales.

Parabasales (cellules) : Trouble de maturation se traduisant par une kératinisation de la surface de l'épithélium avec persistance des noyaux. L'épithélium desquame sous forme de placards denses de cellules à noyau de petite taille, à cytoplasme acidophile. S'observe dans des circonstances variées, dystrophies, infections à HPV, dysplasies.

Polyédrique (Polyèdre) : En géométrie, solides dont les faces sont des polygones.

Pycnotique : petites, très petites.

Rapport nucléo-cytoplasmique : Rapport existant entre la taille du noyau et celle du cytoplasme. Il est augmenté dans les dysplasies.

Superficielles (cellules) : Cellules provenant de la couche superficielle de l'épithélium malpighien exocervical. Elles sont plus nombreuses en première partie de cycle (stimulation estrogénique).

Bibliographie

Tumeur : Augmentation pathologique du volume d'un tissu ou d'un organe, due à une multiplication des cellules.