

**THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE
FRANCHE-COMTÉ**

PRÉPARÉE À L'UNIVERSITÉ DE BOURGOGNE

École doctorale n°37

Sciences Pour l'Ingénieur et Microtechniques

Doctorat d'Informatique et Instrumentation de l'Image

par

ALEXANDRE KREBS

**Imagerie couleur et multispectrale pour la détection de lésions pré-cancéreuses dans
l'estomac**

Thèse présentée et soutenue à Dijon, le 11 Décembre 2019

Composition du Jury :

ELMOATAZ ABDERRAHIM	Professeur à Normandie Université, Caen	Rapporteur
ROMAIN OLIVIER	Professeur à l'Université de Cergy-Pontoise	Rapporteur
LAMARQUE DOMINIQUE	Professeur à l'Université de Versailles St-Quentin-en-Yvelines - Praticien Hospitalier à l'hôpital Ambroise Paré, Boulogne-Billancourt	Examineur
MOUSSATA DRIFFA	Professeure à l'Université de Tours - Praticienne Hospitalier au CHRU Hôpitaux de Tours - Trousseau	Examinatrice
RUAN SU	Professeure à l'Université de Rouen	Examinatrice
BENEZETH YANNICK	Maitre de conférences à l'Université de Bourgogne	Codirecteur de thèse
MARZANI FRANCK	Professeur à l'Université de Bourgogne	Directeur de thèse

Titre : Imagerie couleur et multispectrale pour la détection de lésions pré-cancéreuses dans l'estomac

Mots-clés : Endoscopie digestive, Imagerie NBI et multispectrale, Classification des lésions de l'estomac, Apprentissage supervisé, Apprentissage par transfert, Problèmes inverses et optimisation

Résumé :

Les travaux de thèse s'inscrivent dans le cadre du projet ANR EMMIE. Ce projet vise à développer un système multimodal innovant pour améliorer la détection de lésions inflammatoires dans l'estomac. Pour ce faire, un prototype a été mis en place pour pouvoir acquérir des images endoscopiques NBI et des images multispectrales lors de l'exploration antrale chez l'Homme. Le prototype est constitué d'une colonne endoscopique auquel nous ajoutons des caméras multispectrales.

Les deux modalités, les images NBI et les images multispectrales, sont traitées de manière indépendante. Les images endoscopiques NBI sont traitées avec des descripteurs issus du traitement d'images pour pouvoir détecter quatre pathologies : la gastrite active, la gastrite chronique, la métaplasie intestinale et l'atrophie. Nous avons également développé avec un gastroentérologue des descripteurs visuels. Ces descripteurs nous ont permis de classer avec succès des patches d'images NBI.

Les images NBI sont prétraitées en appliquant une "décomposition en images intrinsèques". Cette décomposition est faite grâce au modèle dichromatique. Nous avons développé deux méthodes pour inverser

ce modèle, respectivement formulées comme des optimisations quadratique et des déconvolutions. Nos algorithmes ont été testés sur trois bases de données différentes. Nous avons comparé nos méthodes avec celles de l'état de l'art et montré des résultats similaires ou meilleurs à ces dernières.

Comme travail préliminaire, nous avons également mené des campagnes d'acquisitions sur un modèle animal de souris. Des spectres de réflectance ont été acquis sur des estomacs réséqués de souris témoins et de souris avec inflammation. Nous avons mis en place une méthode de classification qui nous a permis d'obtenir de bons résultats de classification et d'identifier deux intervalles dans le proche infrarouge utiles à la discrimination.

Cette méthodologie est reprise pour traiter les spectres acquis *in vivo* dans l'estomac chez l'Homme avec le prototype d'acquisition. Le but étant de caractériser l'inflammation et de pouvoir classer les spectres par pathologie. Les résultats ainsi obtenus sont comparés aux résultats des biopsies et montrent une bonne capacité de reconnaissance de la gastrite chronique et de la métaplasie intestinale par rapport à des estomacs témoins.

Title: Imagerie couleur et multispectrale pour la détection de lésions pré-cancéreuses dans l'estomac

Keywords: Digestive endoscopy, Narrow Band Imaging and multispectral imaging, Classification of stomach lesions, Machine Learning, Transfer Learning, Inverse problem and optimization

Abstract:

The work presented in this manuscript is part of the ANR project EMMIE. This project aims to develop an innovative multimodal system for the detection of inflammatory lesions in the stomach. To this purpose, a prototype has been developed to be able to acquire NBI endoscopic images and multispectral images during human's antrum exploration. The prototype is made of a standard endoscope and multispectral images.

The prototype can acquire two types of data: NBI images and spectra. These two modalities are processed independently. Common image processing features are used to recognize four kind of diseases: active gastritis, chronic gastritis, metaplasia and atrophy. In addition, visual based features have been developed with an experienced endoscopist. These criteria were successfully used to classify endoscopic image patches.

NBI images are pre-processed with "intrinsic image decomposition". This decomposition is made by the

dichromatic model. We developed two methods to inverse this model, respectively with quadratic programming and deconvolutions. Our algorithms are tested on three datasets. Equivalent or better results are found compare to the state of the art.

Acquisitions are also made on mice model. Reflectance spectra have been acquired on resected stomach from control and inflamed mice. A pipeline was build to classify the two groups. The pipeline allowed us to obtain good classification results and to identify two interesting wavelengths range in the near infrared.

This pipeline is re-used to process *in vivo* spectra from human's stomach with the developed acquisition prototype. The aim is to characterize the inflammation and being able to classify spectra according to the disease. Obtained results are compared with biopsies and show good diagnostic prognosis for chronic gastritis and intestinal metaplasia.

REMERCIEMENTS

Messieurs Elmoataz et Romain pour avoir accepté d’être rapporteurs de ce manuscrit de thèse et pour leurs retours bienveillants.

Tous les membres du jury qui ont accepté de donner leurs expertises pour juger ce travail.

Franck, pour sa patience et sa présence malgré ses obligations. Il a su parfaitement écouter, motiver et proposer des solutions tout au long de l’encadrement de la thèse. Il était également très impliqué dans la relecture des articles et de ce manuscrit.

Yannick, pour son accompagnement, son point de vue scientifique et sa motivation. Il a suivi de près les avancées des travaux, et, au même titre que Franck, était toujours à l’écoute, a proposé des solutions et était très impliqué dans la relecture des articles et de ce manuscrit.

Dominique, pour son implication exceptionnelle dans le projet EMMIE. Il a passé du temps avec nous pour annoter les images endoscopiques et nous a apporté son expertise médicale indispensable à la réussite du projet.

Eliette et Vania, pour les expériences passionnantes que nous avons menées à l’institut Pasteur sur le petit animal et leur implication dans la rédaction des articles communs. En plus de leur expertise scientifique, leur accueil à l’institut Pasteur était toujours très chaleureux.

“ Je suis de ceux qui pensent que la science est d’une grande beauté. Un scientifique dans son laboratoire est non seulement un technicien : il est aussi un enfant placé devant des phénomènes naturels qui l’impressionnent comme des contes de fées. ”

Marie Curie, *Madame Curie*,

“ L’oiseau est un instrument fonctionnant selon les lois mathématiques ; il est au pouvoir de l’Homme de le reproduire avec tous ses mouvements, mais non avec autant de puissance, seulement du moins pour la faculté d’équilibre. ”

Léonard de Vinci, *Carnet*,

SOMMAIRE

1	Introduction et contexte	1
1.1	Motivations	2
1.2	Évolution des lésions inflammatoires de l'estomac chez l'Homme	2
1.3	Systèmes de diagnostic des lésions inflammatoires de l'estomac	5
1.4	Le projet ANR EMMIE	7
1.5	Objectifs de la thèse	9
1.6	Organisation du manuscrit	10
I	Traitement et analyse des images endoscopiques	11
2	Décomposition d'une image via le modèle dichromatique	13
2.1	Le modèle dichromatique pour la décomposition en images intrinsèques	14
2.2	Méthodes de décomposition existantes	15
2.3	Inversion du modèle par optimisation quadratique	16
2.3.1	Détecteur de pixels gris et mesure de similarité entre couleurs	16
2.3.2	Estimation de la composante shading	17
2.3.3	Estimation de la composante spéculaire	19
2.4	Inversion du modèle par déconvolution	20
2.4.1	Définitions et hypothèses	20
2.4.2	Perfectionnement du détecteur de pixels gris et de la mesure de similarité	22
2.4.3	Formulation par déconvolution	22
2.5	Résultats comparatifs et discussion	24
2.5.1	Résultats sur la base de données du MIT	26
2.5.2	Résultats sur la base de données SIID	29
2.5.3	Résultats sur la base de données CAVE	31
2.6	Conclusion et perspectives	31
3	Détection de lésions dans les images NBI	33
3.1	Traitement d'images pour la détection de lésions de l'estomac par endoscopie . .	34

3.2	Acquisition des données endoscopiques	35
3.3	Méthodologie	36
3.3.1	Classification par patches caractéristiques	36
3.3.2	Application du modèle dichromatique	40
3.3.3	Association des résultats à la mosaïque	43
3.4	Détermination de critères visuels	43
3.5	Conclusion et perspectives	48
II	Traitement et analyse des spectres chez la souris et chez l'Homme	51
4	Étude de la réflectance de l'estomac chez la souris	53
4.1	<i>Helicobacter pylori</i>	54
4.2	Protocole d'acquisition	55
4.3	Pipeline de classification des spectres	58
4.4	Résultats de classification	61
4.5	Conclusion et perspectives	63
5	Présentation et calibrage du prototype d'acquisition	65
5.1	Prototype d'acquisition	66
5.2	Calibrage spectral	68
5.2.1	Transformation linéaire pour le calibrage	69
5.2.2	Validation de l'estimation spectrale	70
5.3	Vers la simulation en lumière blanche et en NBI	73
5.3.1	Extension de la méthode pour la simulation d'image en vraie ou fausse couleur	73
5.3.2	Transformation en lumière blanche	73
5.3.3	Transformation en NBI	75
5.4	Conclusion et perspectives	76
6	Étude préliminaire de la réflectance chez l'Homme	79
6.1	Méthodes existantes d'analyse multispectrale	80
6.2	Acquisitions multispectrales chez l'Homme	81
6.2.1	Collecte des données	81
6.2.2	Première analyse des données	84
6.3	Classification des pathologies de l'estomac	85
6.4	Conclusion et perspectives	87

7 Conclusion et perspectives globales	89
7.1 Apports scientifiques	89
7.2 Perspectives du projet	90

INTRODUCTION ET CONTEXTE

Dans ce chapitre introductif, nous commençons par présenter les motivations qui nous ont conduit à nous intéresser aux lésions pré-cancéreuses de l'estomac. Nous détaillons ensuite les aspects médicaux propres à chaque stade pré-cancéreux et explorons les systèmes existants permettant de détecter et caractériser ces lésions. Enfin, nous présentons les solutions proposées dans le cadre du projet ANR EMMIE dans lequel s'inscrit la thèse. Pour mener à bien les recherches sur les lésions pré-cancéreuses, ce projet regroupe plusieurs partenaires ayant des expertises clinique, biologique et en traitement d'images.

Sommaire

1.1	Motivations	2
1.2	Évolution des lésions inflammatoires de l'estomac chez l'Homme	2
1.3	Systèmes de diagnostic des lésions inflammatoires de l'estomac	5
1.4	Le projet ANR EMMIE	7
1.5	Objectifs de la thèse	9
1.6	Organisation du manuscrit	10

1.1/ MOTIVATIONS

D'après l'Institut National du Cancer (INCa) et son rapport "Les cancers en France, l'essentiel des faits et chiffres, édition 2019" [135], on estime que 4272 décès (1478 femmes et 2794 hommes) sont dus au cancer de l'estomac en 2018. Ces chiffres le place en quatrième position des cancers les plus meurtriers après celui du poumon, du pancréas et du foie. En 2018, on dénombre également 6557 nouveaux cas (2293 femmes et 4264 hommes), ce qui place le cancer de l'estomac en cinquième position des cancers les plus incidents derrière le cancer du poumon, celui du pancréas, ceux de la zone "lèvre, cavité orale, pharynx" et celui du foie. Ces chiffres de la mortalité et de l'incidence du cancer de l'estomac font de lui un cancer avec un impact sociétal important.

De manière générale, le taux de mortalité, tout cancer confondu, tend à diminuer. Entre 2010 et 2018, on enregistre une baisse de 0,7% par an pour les femmes et de 2% pour les hommes. Ces chiffres sont encourageants mais dépendent fortement du type de cancer. La survie nette à 5 ans a bien augmenté pour certains cancers comme pour le cancer du sein ou celui de la prostate, notamment grâce à des diagnostics plus précoces. Malheureusement, la survie nette à 5 ans stagne pour le cancer de l'estomac. Pour la période 1989-1993 et la période 2005-2010, elle est dans les deux cas de l'ordre de 30%. Le pronostic de ce cancer se trouve donc dégradé et doit être amélioré car on enregistre trop peu de progrès. L'organisme INCa s'intéresse bien sûr aux cancers les plus fréquents mais également aux cancers, comme celui de l'estomac, pour lesquels on n'observe pas ou peu d'amélioration de la survie.

Pour améliorer le taux de survie net, on peut s'intéresser aux principaux facteurs de risque. Ceux-ci sont aujourd'hui connus. On peut notamment citer l'infection par la bactérie *Helicobacter pylori*, le tabagisme et une alimentation trop riche en sel et pauvre en fruits et légumes frais. On peut noter que le risque de développer un cancer de l'estomac est également augmenté en cas d'antécédents familiaux ou de prédisposition génétique. Certains de ces facteurs sont malheureusement inévitables. Il s'agit donc plutôt de détecter le plus précocement possible les signes d'un futur cancer. La compréhension de l'évolution de la maladie est alors importante ainsi que la caractérisation des tumeurs.

En ce sens, l'exploitation de bases de données avec des outils d'intelligence artificielle est bénéfique en permettant de "prédire" la survenue des cancers à des stades précoces. L'émergence de nouvelles technologies d'imagerie est également un atout pour améliorer la détection et la caractérisation des stades précoces du cancer de l'estomac car les technologies d'imagerie utilisées actuellement sont encore limitées et ne permettent pas une détection aisée des pathologies.

1.2/ ÉVOLUTION DES LÉSIONS INFLAMMATOIRES DE L'ESTOMAC CHEZ L'HOMME

L'estomac peut être divisé en trois grandes parties selon les différences histologiques de la muqueuse : le fundus, le corps et l'antrum (voir figure 1.1). Certaines parties de l'estomac sont plus susceptibles de développer des pathologies : 35% des pathologies sont retrouvées dans le fundus, 25% dans le corps et 40% dans l'antrum [36]. Ces différences pourraient être dues aux résidus permanents de suc gastrique dans l'antrum. Nous nous intéressons donc préférentiellement à cette zone qui est petite mais qui est aussi celle dans laquelle se concentrent les lésions inflammatoires.

1.2. ÉVOLUTION DES LÉSIONS INFLAMMATOIRES DE L'ESTOMAC CHEZ L'HOMME³

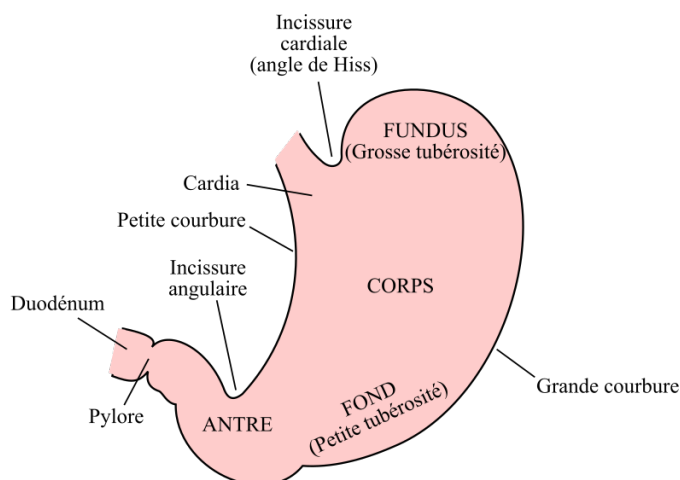


FIGURE 1.1 – Morphologie de l'estomac (issue de <https://www.passeportsante.net/fr/parties-corps/Fiche.aspx?doc=fundus>).

Le développement d'un cancer de l'estomac suit plusieurs stades connus et bien définis. Chacun de ces stades se caractérise par un changement dans la composition ou la structure de la paroi de l'estomac (muqueuse) que nous illustrons figure 1.2. Nous listons ci-après les caractéristiques principales des pathologies suivant l'ordre d'évolution vers le cancer de l'estomac.

- La gastrite aigüe ou active est souvent réversible et de courte durée. La gastrite aigüe est caractérisée par une infiltration de polynucléaires neutrophiles et de lymphocytes (mononucléaires) dans le chorion (que l'on appelle aussi *lamina propria* comme dans le schéma 1.2). La gastrite aigüe est associée à la présence de *Helicobacter pylori*. À l'endoscopie, on observe un œdème de la muqueuse et parfois des érosions.
- La gastrite chronique est une inflammation de la muqueuse de l'estomac. Elle est non-réversible. Biologiquement, elle correspond à l'infiltration du chorion uniquement par des lymphocytes. La densité des lymphocytes détermine le caractère visible ou non à l'endoscopie : la densité est en pratique souvent trop faible pour être identifiable. La gastrite chronique peut être vue comme une séquelle de l'infection à *Helicobacter pylori* et est le résultat du mode d'évolution de la gastrite aigüe.
- L'atrophie glandulaire de la muqueuse gastrique se définit par une diminution de l'épaisseur de la paroi. La densité des glandes diminue (voir schéma 1.3) et le revêtement de l'estomac s'amincit ce qui rend les micro-vaisseaux sanguins entre les glandes plus visibles (voir schéma 1.2). Elle s'observe le plus souvent chez les personnes âgées. À l'endoscopie, on visualise anormalement ces micro-vaisseaux.
- La métaplasie intestinale est l'apparition, au sein des glandes atrophiques de cellules que l'on trouve normalement dans l'intestin grêle ou le gros intestin (voir schéma 1.3). Le phénomène n'est, dans la majorité des cas, pas réversible. Il s'accompagne d'une augmentation de la micro-vascularisation capillaire qui est identifiable en endoscopie "Narrow Band Imaging" (NBI) par un aspect violacé de la muqueuse avec un épaissement et un allongement des orifices glandulaires (cryptes) qui apparaissent selon une trame blanchâtre.
- La métaplasie peut évoluer en dysplasie. Les cellules prolifèrent et se superposent de façon désordonnée. Ces cellules peuvent devenir cancéreuses et évoluer vers un adénocarcinome.

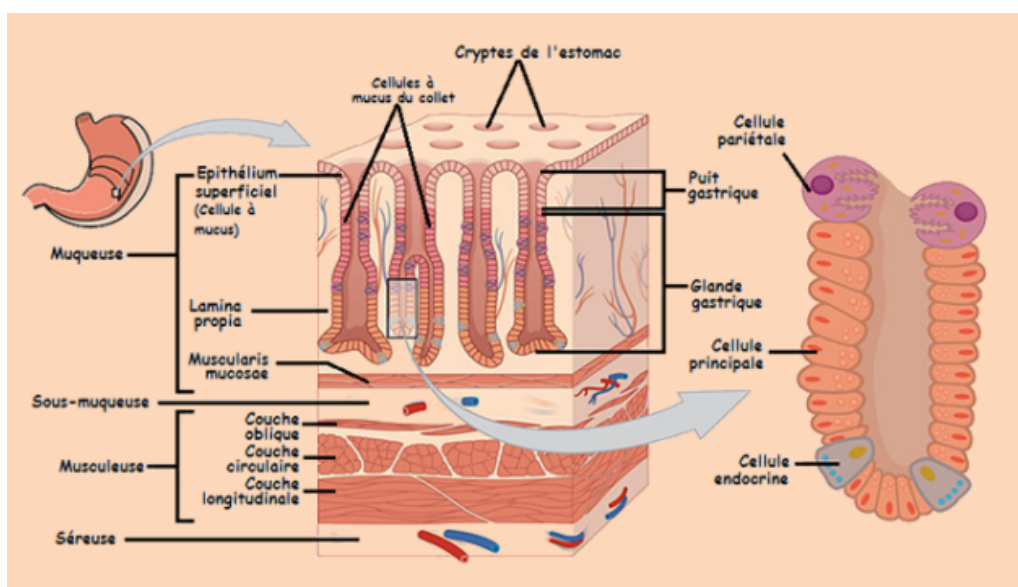
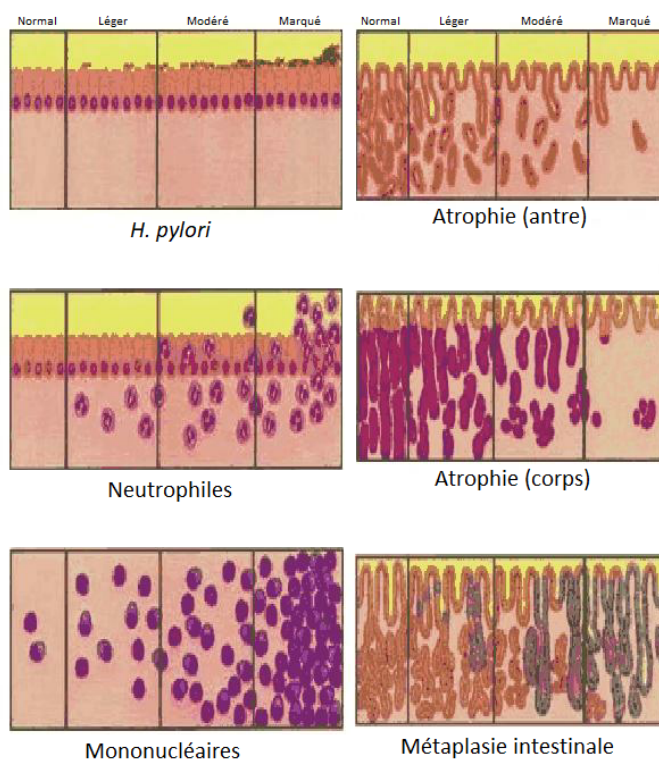


FIGURE 1.2 – Paroi de l'estomac

(issue de <http://ressources.unisciel.fr/physiologie/co/grain6b2.html>).FIGURE 1.3 – Score Sydney (issue de <https://www.hindawi.com/journals/grp/2013/393015/fig3/>).

La gravité des lésions est estimée par le score de Sydney [21] [107]. Il s'agit de quantifier la présence d'*Helicobacter pylori*, la densité des neutrophiles, la densité des lymphocytes (mononucléaires), l'atrophie et la présence de métaplasie intestinale comme le montre la figure 1.3. Pour

chacun des critères, une note entre 0 (normal) et 3 (marqué) est donnée. Il existe d'autres scores, le score "Operative Link on Gastritis Assessment" (OLGA) ou le score "Operative Link on Gastric Intestinal Metaplasia" (OLGIM) qui prennent en compte la localisation des prélèvements et différencie le score dans l'antrum du score dans le corps gastrique [114]. Dans notre cas, seul l'antrum est étudié, nous nous basons donc uniquement sur le score de Sydney.

1.3/ SYSTÈMES DE DIAGNOSTIC DES LÉSIONS INFLAMMATOIRES DE L'ESTOMAC

A ce jour, les cliniciens utilisent majoritairement l'endoscopie digestive haute, en lumière blanche ou en NBI pour diagnostiquer les pathologies de l'estomac. En pratique, il reste toujours difficile de reconnaître ces lésions inflammatoires notamment parce qu'elles sont pratiquement indécélables, en particulier avec l'examen en lumière blanche. Les lésions sont en effet souvent trop petites pour être vues au niveau macroscopique. Cela conduit l'endoscopiste à faire des prélèvements systématiques et non-orientés de la muqueuse que l'on appelle biopsies : on prélève des petits morceaux d'estomac que l'on analyse *a posteriori* en laboratoire pour établir le diagnostic du patient. Durant le prélèvement, des pinces sont insérées dans le canal opérateur de l'endoscope pour découper un bout de tissu. La localisation des prélèvements est aléatoire et peut ne pas refléter l'état général de la paroi gastrique. De plus, le procédé n'est pas reproductible, le résultat dépend fortement de l'expérience de l'opérateur.

Le NBI, créé par Olympus est un système constitué de deux ou trois filtres sensibles à des longueurs d'onde spécifiques, 415 et 540 nm [6] ou 415, 445 et 500 nm comme illustré par la figure 1.4. Ces filtres sont respectivement placés dans les canaux bleu, vert et rouge de l'image endoscopique créant ainsi une image en fausse couleur (dans le cas où l'on a seulement deux longueurs d'onde différentes, les canaux bleu et vert contiennent la même information). Biologiquement, ces longueurs d'onde correspondent aux pics d'absorption de l'hémoglobine. Les vaisseaux sanguins apparaissent ainsi en noir, ce qui permet de bien les visualiser. Le système a fait ses preuves en permettant notamment de diagnostiquer l'œsophage de Barrett [113] et de visualiser des textures caractéristiques de la gastrite chronique et de la métaplasie [37].

La principale alternative au NBI est le système "Fuji Intelligent Chromo Endoscopy" (FICE) créé par FUJI [2]. Le système FICE permet d'estimer une image multispectrale à partir d'une image endoscopique classique. Le système permet d'estimer des images aux longueurs d'onde entre 400 et 695 nm par pas de 5 nm, soit un total de soixante longueurs d'onde. On peut alors choisir d'afficher trois images parmi les soixante dans les trois canaux usuels rouge, vert et bleu. Dix configurations sont prédéfinies sur les endoscopes rendant ainsi possible une meilleure visualisation, notamment des vaisseaux sanguins et des zones anormales. Le système FICE s'est avéré utile pour différencier deux types de métaplasie (complète et incomplète) sur des patients atteints de gastrite chronique avec atrophie [35]. On peut noter que la différence majeure avec le NBI est que ce dernier se base sur une technologie optique, c'est à dire une amélioration matérielle par rapport à l'endoscopie classique alors que le système FICE propose une amélioration logicielle avec l'estimation d'une réponse spectrale [41].

Les constructeurs d'endoscopes (Fuji, Olympus ou Pentax) ont amélioré la résolution des images endoscopiques dans les dernières années permettant une meilleure visualisation des lésions [121]. Pourtant, s'il y a des améliorations des technologies d'un point de vue résolution, il n'y a pas ou peu d'améliorations du point de vue de la détection des lésions. Le NBI est limité à deux ou trois bandes et FICE ne permet que d'estimer une réponse spectrale à partir de trois bandes. Il existe de meilleurs systèmes utilisant le NBI avec grossissement (ME-NBI) permettant une meilleure discrimination des tissus [27]. Il s'agit d'un système NBI permettant un zoom optique. Uedo *et al.*

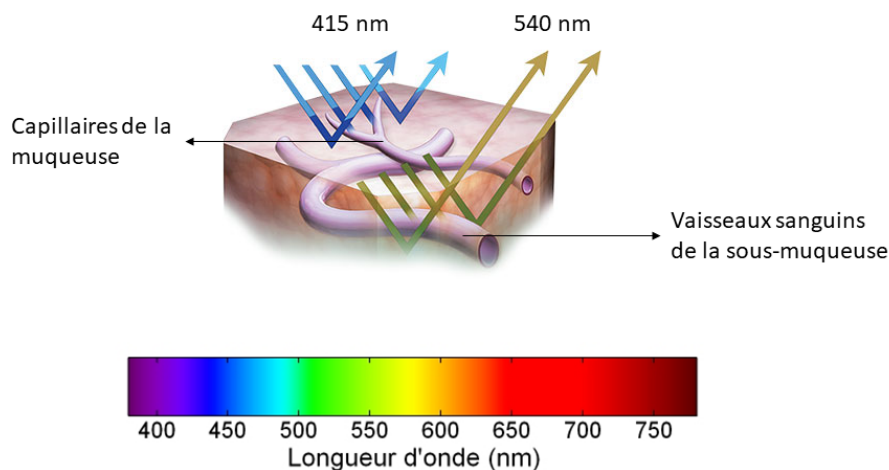


FIGURE 1.4 – Fonctionnement du NBI : deux filtres centrés sur les longueurs d’onde 415 et 540 nm sont placés devant la lumière d’un endoscope. La lumière filtrée pénètre le tissu et est absorbée par les vaisseaux sanguins et les capillaires (issue de <https://www.barmherzige-brueder.at/site/wien/home>).

ont travaillé avec ce système et ont corrélé de manière significative l’apparition de crêtes blanches et bleues à la surface de l’estomac avec la présence de métaplasie [34].

Ces deux modalités montrent que l’imagerie non-conventionnelle, en particulier l’utilisation d’informations spectrales, permet d’améliorer la visualisation des lésions dans l’estomac. Les recherches continuent dans ce sens. Gu *et al.* ont proposé un système multispectral à sept bandes entre 465 et 665 nm. Les bandes 645 nm, 545 nm et 495 nm ont été sélectionnées comme étant celles qui minimisaient le plus l’information mutuelle. Ces trois bandes ont ensuite été utilisées pour créer une image RGB rehaussée permettant ainsi une meilleure visualisation. Le système a été testé sur un cas pratique de détection d’ulcère dans l’estomac [106].

La visualisation est une première aide à la détection des pathologies mais l’interprétation est dans tous les cas laissée aux endoscopistes. Le diagnostic dépend alors de l’expérience du clinicien et n’est pas systématiquement corrélé aux résultats histologiques [81]. Il est alors important d’améliorer la détection par des outils de traitement du signal et de l’image adaptés aux modalités non-conventionnelles. Par exemple, Kiyotoki *et al.* ont utilisé une caméra hyperspectrale entre 400 et 800 nm pour caractériser des spectres de tumeurs de 14 patients réséquées [80]. Des travaux similaires ont été conduits par Ogihara *et al.*. Ils ont également utilisé une caméra hyperspectrale entre 400 et 800 nm [112] pour photographier des estomacs avec cancer réséqués. Ils ont observé une forte variabilité inter-patient qui les a conduit à normaliser l’ensemble des spectres en utilisant le spectre moyen des zones saines. Ces travaux sont très prometteurs mais restent cloisonnés à une utilisation *ex vivo*.

Des travaux *in vivo* utilisant des systèmes multispectraux existent pourtant aussi dans la littérature. En 2017, Hohmann *et al.* ont démontré l’intérêt des données multispectrales dans la détection d’adénocarcinomes dans l’estomac [122]. Leur système consiste en un endoscope standard couplé à une source de lumière modifiée qui permet d’extraire des données à six longueurs d’onde différentes entre 438 et 628 nm. Makhlof *et al.* ont élaboré un microendoscope confocal fonc-

tionnant en deux modes. L'un donnant une image en niveau de gris pour observer la réponse en fluorescence l'autre en mode "multispectral" pour obtenir une image multispectrale du signal de fluorescence [39]. Martinez *et al.* ont également proposé un système utilisant une roue à filtres, leur permettant d'acquérir des images à six longueurs d'onde différentes. Ces images doivent ensuite être recalées. Un classifieur de type SVM avec noyau gaussien a été entraîné avec des patches de tailles 5x5 pour détecter des zones pré-cancéreuses de l'estomac [86].

Dans la littérature, d'autres techniques ont été développées afin d'améliorer l'identification des pathologies par imagerie optique. Par exemple, on peut utiliser des agents de contraste pour améliorer la caractérisation de l'estomac, comme avec la chromoendoscopie [40]. Certains de ces agents sont absorbés par les cellules ou s'accumulent sur les pics et les creux de la muqueuse pour mettre en avant la structure des cellules. On visualise donc encore mieux à l'endoscopie les zones pathologiques. L'utilisation d'agents de contraste n'est pas forcément nécessaire, d'autres substances biologiques, nommées fluorophores peuvent être stimulées par la lumière. Ce phénomène est exploité par la technique d'autofluorescence : en excitant les fluorophores à une certaine longueur d'onde, ceux-ci ré-émettent de la lumière à une longueur d'onde plus élevée [97]. En l'occurrence, le collagène de la sous-muqueuse est le fluorophore le plus important mais on trouve également le phosphate de pyridoxal (PLP), la riboflavine, les phospholipides et les porphyrines I et II (PPS) [137]. Au contraire, les chromophores (l'hémoglobine par exemple) ne ré-émettent pas de lumière mais l'absorbent. Le résultat de l'influence des fluorophores et des chromophores est en général visualisé grâce à des caméras multispectrales [26].

Un aspect intéressant de l'acquisition de données hyperspectrales est la possibilité de remonter à des paramètres biologiques. Un spectre étant la somme des contributions de plusieurs agents biologiques, on peut réaliser une opération de "démixage", équivalent à un problème de séparation de sources. Par exemple, Grosberg *et al.* ont caractérisé les tissus gastriques grâce à un système de microscopie hyperspectrale à deux photons leur permettant d'obtenir des images hyperspectrales à l'échelle cellulaire. Ils ont réalisé un démixage sur les spectres en quatre catégories de tissu : l'épithélium, la *lamina propria*, le collagène et les tissus lymphatiques [59]. Bergholt *et al.* ont acquis des spectres via la spectroscopie Raman. Les spectres étaient également démixés pour trouver les contributions relatives de l'ADN, des protéines, des lipides, du sang et des glycoprotéines. Les contributions ont permis de reconnaître quatre types de tissu différents : les tissus normaux, la métaplasie intestinale, la dysplasie et les adénocarcinomes [73].

1.4/ LE PROJET ANR EMMIE

C'est donc de ce constat qu'il y a besoin de plus d'études *in vivo* et que les modalités basées sur des technologies multispectrales apportent plus d'information qu'est né le projet "Endoscopie MultiModale pour les lésions Inflammatoires de l'Estomac" abrégé EMMIE. Ce projet est financé par l'Agence Nationale de la Recherche (ANR) dans le cadre de programmes de recherche translationnelle en santé. Il consiste à développer un système innovant qui allie plusieurs modalités pour détecter et caractériser l'inflammation de la muqueuse de l'estomac. Le projet EMMIE ne s'intéresse pas directement au cancer comme la majorité des travaux cités précédemment mais aux lésions pré-cancéreuses. En effet, c'est en détectant au plus tôt les signes de lésions inflammatoires que l'on pourra améliorer la situation en augmentant significativement les chances de réversibilité des pathologies et en proposant des traitements adaptés.

Plusieurs études préliminaires ont démontré qu'il pouvait exister des différences entre les spectres de réflectance de tissus normaux et ceux de tissus inflammatoires. Pourtant, peu de systèmes exploitent cette différence. Ils contiennent un nombre réduit de bandes, sont au stade expérimental

ou ne sont utilisés que pour des acquisitions *ex vivo*. De plus, l'intérêt est souvent porté sur le cancer alors que la détection de stades précoces, avant le cancer, est plus complexe et plus rarement abordée.

La détection de zones inflammatoires peut être améliorée en utilisant des longueurs d'onde spécifiques qui montrent une différence significative entre les tissus normaux et ces lésions. Le développement d'un système multispectral et des méthodes d'analyse associées permettrait d'avoir une méthode quantitative reproductible pour caractériser les lésions inflammatoires. Dans le projet EMMIE, nous proposons donc de développer un prototype innovant permettant l'acquisition d'un flux vidéo multispectral et d'un flux vidéo endoscopique de manière simultanée. Pour ce faire, nous utilisons un fibroscope que l'on insère dans le canal opérateur d'un endoscope standard. Les changements de réflectance dus à l'inflammation sont caractérisés d'abord *ex vivo* sur un modèle animal de souris avant d'être caractérisés *in vivo* chez l'Homme. Chez ce dernier, la zone de l'estomac étudiée en priorité est la portion distale (antre) qui est la zone la plus à risque.

Le projet EMMIE est la continuité de plusieurs travaux sur les tissus épidermiques [67, 74] [52, 60, 78] [79] [131] [54] [82, 83, 87, 88] [99] et l'endoscopie multispectrale [86, 77, 84, 95] menés par les partenaires du projet EMMIE. Le projet regroupe exactement six partenaires se connaissant et ayant déjà eu des interactions.

- L'unité "Pathogénèse de *Helicobacter*" de l'institut Pasteur, dont les recherches sont orientées, entre autres, vers l'étude des virus, des bactéries, des champignons, des parasites et des maladies associées.
- L'unité INSERM 1173 "Infection et Inflammation Chronique" rattachée à l'université de Versailles Saint-Quentin-en-Yvelines. Cette unité travaille sur les réponses inflammatoires et infectieuses cellulaires sur différents modèles animal et chez l'Homme.
- Le service d'hépatogastroentérologie de l'hôpital Ambroise Paré de Boulogne-Billancourt qui apporte l'expertise clinique nécessaire au projet.
- Le laboratoire Pluridisciplinaire de Recherche en Ingénierie des Systèmes, Mécaniques et Énergétiques (PRISME) d'Orléans. Les thématiques de recherche du laboratoire incluent l'aérodynamique, l'automatique, la robotique et le traitement du signal et de l'image.
- Le Centre de Recherche en Automatique de Nancy (CRAN). Il s'agit d'une Unité Mixte de Recherche (UMR 7039) commune à l'université de Lorraine et au CNRS. Le CRAN accueille également des chercheurs de l'institut de Cancérologie de Lorraine. Les recherches du laboratoire concernent l'automatique mais aussi le traitement du signal et le génie informatique.
- Enfin, le laboratoire Imagerie et Vision Artificielle (ImViA), fondé en janvier 2019, est né de la restructuration de l'ancien laboratoire LE2I. Il s'agit d'une Équipe d'Accueil (EA) sous la tutelle de l'université de Bourgogne. Le laboratoire effectue ses recherches dans les domaines de la vision et de l'analyse d'images, notamment avec des applications médicales. Le projet est coordonné par ce laboratoire et je suis rattaché à celui-ci.

Ces partenaires ont des expertises différentes qui permettent de répondre aux objectifs du projet EMMIE : ils sont complémentaires à la fois sur des aspects biologique, clinique et sur le traitement d'images. Différentes tâches ont donc été décrites et réparties entre les partenaires selon leurs expertises comme le montre la figure 1.5. L'unité 1173 de l'INSERM et l'institut Pasteur travaillent sur l'exploration fonctionnelle et l'acquisition multimodale. L'objectif de cette tâche est d'acquérir des spectres de réflectance sur des estomacs de souris. Le traitement et l'analyse de ces spectres ont été faits en étroite collaboration avec le laboratoire ImViA et font partie du travail de thèse décrit dans ce manuscrit au chapitre 4. L'hôpital Ambroise Paré est responsable des acquisitions chez

l'Homme via le prototype décrit. Le prototype a été validé par le Comité de Protection des Personnes (CPP) du Sud-Est avec le numéro d'enregistrement 2019-038 B. L'élaboration du prototype et l'analyse des données acquises sont faites en partenariat avec le laboratoire ImViA et font partie du travail de thèse au travers des chapitres 3, 5 et 6. Le laboratoire PRISME s'occupe du calibrage spatial du prototype, permettant ainsi le recalage entre les images endoscopiques et les images multispectrales. Le CRAN a pour tâche de créer un logiciel capable de créer une mosaïque à partir des images endoscopiques. Il s'agit de reconstruire en 2D ou 2D^{1/2} l'estomac à partir des images endoscopiques. Le travail comprend l'estimation du flux optique entre des images successives et leur agencement en mosaïque, permettant la reconstruction logicielle de l'estomac. Enfin, le laboratoire ImViA travaille sur les aspects de traitements des données, en particulier, des traitements des spectres de réflectance et des images endoscopiques.

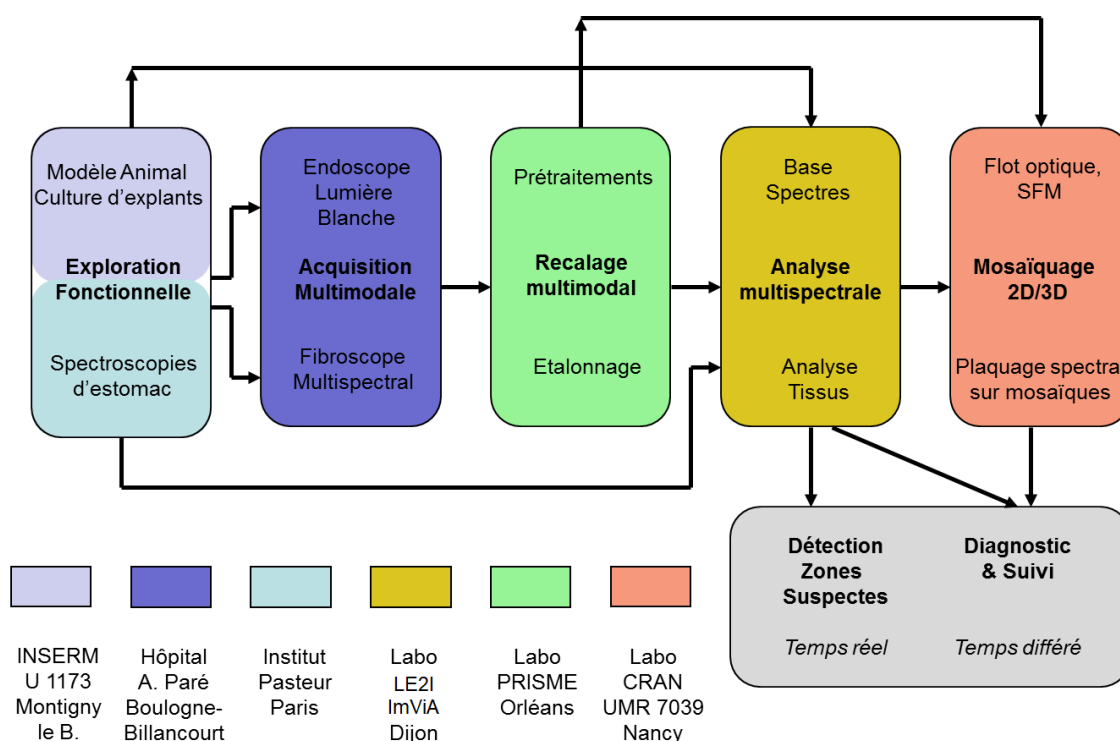


FIGURE 1.5 – Répartition des tâches entre les acteurs du projet comme initialement définie lors du dépôt de dossier scientifique.

1.5/ OBJECTIFS DE LA THÈSE

Les travaux de thèse s'imbriquent dans le projet EMMIE. Le premier objectif était de développer le prototype qui permet d'acquérir des images endoscopiques et des images multispectrales. Le prototype consiste en un endoscope classique utilisé dans les blocs opératoires auquel sont ajoutés des caméras multispectrales, un système optique et un ordinateur pour recueillir les acquisitions. Le système a besoin d'être calibré spatialement et spectralement : spatialement parce qu'il est nécessaire de connaître la transformation géométrique pour re-situer l'image multispectrale dans l'image endoscopique et spectralement parce que les caméras doivent être calibrées pour donner un spectre de réflectance.

Le cœur du projet EMMIE est la détection de l'inflammation chez l'Homme. Le deuxième objectif est donc l'analyse des deux modalités acquises, c'est-à-dire des images endoscopiques NBI et des spectres. Dans un premier temps, les images endoscopiques NBI sont étudiées. Nous traitons les images endoscopiques par patchs sur lesquels plusieurs types de descripteurs liés à la couleur, à la texture ou issus de l'apprentissage profond sont testés. Les descripteurs sont combinés avec un classifieur de type SVM pour évaluer leur capacité à discriminer les lésions inflammatoires. En parallèle de cela, nous avons développé avec un gastroentérologue expérimenté des critères visuels utiles pour classifier les patchs.

Pour mieux appréhender ces deux types de données, nous nous sommes intéressés aux méthodes de traitements que nous aurions à utiliser. D'une part, les images sont soumises aux gradients d'illumination qui peuvent rendre complexe l'interprétation des images. Nous avons donc développé une méthode basée sur un modèle mathématique, le modèle dichromatique qui permet de s'affranchir de ce problème. D'autre part, nous avons également étudié la réflectance chez la souris. Les données acquises chez la souris nous permettent d'identifier des longueurs d'onde utiles, d'extraire des descripteurs spectraux utiles à la détection de l'inflammation et de définir un processus robuste de classification que nous appliquons ensuite sur les spectres acquis chez l'Homme.

1.6/ ORGANISATION DU MANUSCRIT

Le manuscrit de thèse est organisé de la manière suivante : dans une première partie, nous exposons les méthodes de traitement d'images que nous avons utilisées pendant la thèse. En particulier, nous avons développé des algorithmes permettant d'estimer la réflectance à partir d'une seule image RGB ou multispectrale. Les algorithmes se basent sur le modèle dichromatique, qui permet de décomposer une image en ses composantes intrinsèques (ou invariants photométriques). Le modèle et les méthodes pour l'inverser sont présentées dans le chapitre 2. Dans un deuxième temps, nous détaillons la manière dont nous traitons les images endoscopiques NBI que les gastroendoscopistes sont habitués à voir. Celles-ci peuvent être sujettes aux gradients d'illumination ou aux tâches spéculaires. Grâce à l'inversion du modèle dichromatique, nous pouvons supprimer ces phénomènes. Nous travaillons alors sur ces images pour détecter plusieurs pathologies de l'estomac. Pour ce faire, nous utilisons des méthodes d'apprentissage supervisé et des méthodes d'apprentissage profond qui sont détaillées dans le chapitre 3.

Dans une deuxième partie, nous présentons l'ensemble des méthodes d'analyse de spectres. Une première approche consiste à étudier les spectres de réflectance diffuse chez la souris. Ces travaux sont décrits dans le chapitre 4. Nous comparons les spectres de la muqueuse gastrique provenant de souris témoins et de souris infectées par *Helicobacter pylori*, une bactérie provoquant une inflammation. Les méthodes ainsi développées sont ensuite ré-utilisées pour analyser les spectres chez l'Homme. Dans le chapitre 5, nous avons mis en place un prototype qui s'ajoute au système endoscopique existant dans les blocs opératoires. Ce prototype permet d'acquérir des images multispectrales en plus des images endoscopiques chez l'Homme. Les spectres issus des images multispectrales obtenues chez l'Homme sont analysés au chapitre 6 pour trouver les caractéristiques spectrales qui indiquent l'inflammation.

I

TRAITEMENT ET ANALYSE DES IMAGES ENDOSCOPIQUES

DÉCOMPOSITION D'UNE IMAGE VIA LE MODÈLE DICHROMATIQUE

Les images multispectrales sont, au même titre que les images RGB souvent sujettes aux gradients d'illumination. Il est nécessaire de pouvoir s'affranchir de ces variations pour retrouver la réflectance d'un matériau ou la vraie couleur d'une image et ainsi pouvoir mieux l'analyser. Pour cela, il existe un modèle mathématique, le modèle dichromatique qui permet de séparer les différentes composantes d'une image tels l'illuminant, la réflectance du matériau et la réflexion spéculaire.

Nous avons donc élaboré un moyen d'inverser le modèle dichromatique. Plus précisément, deux approches ont été proposées, l'une en trouvant les minima de fonctions de coût quadratiques et l'autre, en résolvant deux problèmes de déconvolution. Les deux approches sont comparées qualitativement et quantitativement aux algorithmes de l'état de l'art qui inversent également le modèle.

Sommaire

2.1	Le modèle dichromatique pour la décomposition en images intrinsèques . .	14
2.2	Méthodes de décomposition existantes	15
2.3	Inversion du modèle par optimisation quadratique	16
2.3.1	Détecteur de pixels gris et mesure de similarité entre couleurs	16
2.3.2	Estimation de la composante shading	17
2.3.3	Estimation de la composante spéculaire	19
2.4	Inversion du modèle par déconvolution	20
2.4.1	Définitions et hypothèses	20
2.4.2	Perfectionnement du détecteur de pixels gris et de la mesure de similarité	22
2.4.3	Formulation par déconvolution	22
2.5	Résultats comparatifs et discussion	24
2.5.1	Résultats sur la base de données du MIT	26
2.5.2	Résultats sur la base de données SIID	29
2.5.3	Résultats sur la base de données CAVE	31
2.6	Conclusion et perspectives	31

2.1/ LE MODÈLE DICHROMATIQUE POUR LA DÉCOMPOSITION EN IMAGES INTRINSÈQUES

La formation d'une image résulte de l'interaction de rayons lumineux à la surface des objets. Il y a principalement deux types de réflexion : la réflexion diffuse et la réflexion spéculaire. La réflexion diffuse correspond à un rayon réfléchi à la surface d'un objet dans toutes les directions alors que la réflexion spéculaire correspond à un rayon réfléchi dans une seule et même direction, symétrique de la direction incidente par rapport à la normale de l'objet, comme avec un miroir. La séparation de ces deux composantes est très utile en traitement d'images, notamment pour les problèmes de suppression du spéculaire [126] [125], d'estimation de la géométrie d'un objet [92], de la re-coloration [111], du remplacement de fond [96], de la segmentation [134] ou encore de la classification [70]. On appelle "images intrinsèques" ou "invariants photométriques" les résultantes d'une telle décomposition.

En 1985, Shafer *et al.* proposent le modèle de réflexion dichromatique pour modéliser les réflexions sur des objets non-lambertiens [14]. Il s'agit d'un cas particulier du modèle de réflectivité bidirectionnelle (Bidirectional Reflection Distribution Function, BRDF en anglais) dans lequel on fait l'hypothèse que la lumière est uniforme sur tout l'espace. Le modèle dichromatique relie une image à ses invariants photométriques par une combinaison linéaire. On l'exprime grâce à l'équation suivante :

$$I(u, \lambda) = L(\lambda)(g(u)S(u, \lambda) + k(u)) \quad (2.1)$$

dans laquelle l'indice u est la position d'un pixel et l'indice λ la longueur d'onde. L est le spectre de l'illuminant multiplié par la sensibilité du capteur. g est le facteur d'échelle dépendant uniquement de la géométrie. Par commodité, nous nommerons ce facteur par son équivalent anglais "shading". k désigne le coefficient spéculaire et S désigne la réflectance du matériau. Il est à noter que g et k sont des images en niveaux de gris alors que S est une image à plusieurs canaux, le même nombre que pour l'image de départ I . Une image est ainsi composée d'une partie diffuse $L(\lambda)g(u)S(u, \lambda)$ et d'une partie spéculaire $L(\lambda)k(u)$. En général, on souhaite supprimer la partie spéculaire, enlever l'influence de l'illuminant et la sensibilité du capteur. On souhaite garder le shading pour avoir des informations sur la forme de l'objet ou la réflectance du matériau pour avoir des informations sur la couleur de l'objet.

L'inversion de ce modèle décompose une image en images intrinsèques, *i.e.* les invariants photométriques. Malheureusement, la décomposition n'est pas aisée parce que le problème est mal conditionné : il y a un nombre infini de solutions. Pour contrer cette difficulté, les chercheurs utilisent soit plusieurs vues du même objet ou bien ajoutent des contraintes pour favoriser certaines solutions. Par exemple, Xie *et al.* utilisent des images stéréo et ajoutent des contraintes locales et globales ainsi que des contraintes de disparité entre les images gauche et droite pour pouvoir réaliser la décomposition en images intrinsèques [117]. Une contrainte que l'on retrouve souvent est une contrainte de lissage. En effet, les objets ont tendance à être lisses. Par conséquent, l'image g l'est aussi. Huynh *et al.* utilisent cette contrainte en introduisant un terme de régularisation qui pénalise les fortes discontinuités [51].

Dans les prochaines sections, nous présentons d'abord l'état de l'art des méthodes existantes pour inverser le modèle dichromatique puis nous présentons deux nouvelles méthodes, que nous avons développées pour l'inversion à partir d'une seule image RGB ou multispectrale. La première méthode est basée sur deux optimisations quadratiques sous contraintes et la deuxième présente le problème sous la forme de deux déconvolutions. Enfin nous montrerons les résultats qualitatifs et quantitatifs de ces deux méthodes comparés à ceux des méthodes l'état de l'art.

2.2/ MÉTHODES DE DÉCOMPOSITION EXISTANTES

L'inversion du modèle dichromatique fait partie des problèmes inverses sous-déterminés. La méthode principale pour réduire ces indéterminations consiste à réduire le nombre d'inconnues. En l'occurrence, la composante L est souvent connue ou peut être estimée au moyen des algorithmes classiques du monde gris [13], du patch blanc [25] ou des "bords gris" [15]. De nouveaux algorithmes, basés sur des techniques d'apprentissage profond, ont émergé pour estimer L . Ces algorithmes utilisent des réseaux de neurones convolutionnels (CNN) comme celui proposé par Bianco *et al.* [93] ou Barron *et al.* [91]. Fourure *et al.* ont également proposé une structure basée sur des réseaux mutuels [104]. Dans notre cas, nous supposons un illuminant blanc ou connu et nous omettons donc l'illuminant de l'équation 2.1 et considérons l'équation suivante :

$$R(u, \lambda) = \frac{I(u, \lambda)}{L(\lambda)} = g(u)S(u, \lambda) + k(u). \quad (2.2)$$

Le terme R est souvent nommé réflectance dans la littérature. Pour lever l'ambiguïté avec S nous utiliserons le terme "réflectance totale" pour le désigner. L'inversion de la nouvelle équation 2.2 est toujours sous-déterminée. Plusieurs méthodes sont présentes dans la littérature pour inverser le modèle à partir d'une image RGB ou d'une image multispectrale. En ce qui concerne les images RGB, les méthodes les plus récentes sont basées sur l'apprentissage profond. Dans ce cas, aucune contrainte n'est véritablement exposée ce qui signifie que les contraintes sont elles-mêmes apprises mais ne peuvent pas être explicitées facilement. Dans le cas multispectral, il est très difficile d'utiliser des méthodes d'apprentissage profond car il n'existe pas de bases de données de grande taille. De plus, comme le nombre de bandes change d'une caméra multispectrale à l'autre, les réseaux ne seraient pas suffisamment génériques comparés aux autres méthodes et devraient être ré-entraînés systématiquement.

Parmi les méthodes ne requérant pas d'apprentissage et qui s'appliquent aux images RGB, nous pouvons citer les travaux de Yang *et al.* [100]. Ceux-ci consistent à enlever le spéculaire par application itérative de filtrage guidé [76]. Seules quelques itérations semblent suffire, ce qui fait de l'algorithme une solution efficace d'un point de vue temps de calcul. Suo *et al.* ont également développé une méthode rapide de suppression du spéculaire basée sur les couleurs normalisées et une étape de regroupement des couleurs [115].

Les autres méthodes ne requérant pas d'apprentissage résolvent le problème de décomposition en images intrinsèques par des procédures d'optimisation de fonctions de coût basées sur le modèle dichromatique : par exemple Gu *et al.* ont proposé une résolution en trois phases basée sur des "Shapelets" [75]. Plus récemment Gu *et al.* ont proposé une méthode permettant d'inverser le modèle via deux optimisations quadratiques [105]. Akashi *et al.* ont proposé une approche utilisant la factorisation en matrices creuses non négatives (Sparse Non Negative Matrix Factorization, NNMF en anglais) [102]. Ces méthodes basées sur de l'optimisation ont l'avantage d'être applicables aussi bien aux images RGB qu'aux images multispectrales.

D'autres modèles que le modèle dichromatique ont été proposés pour la décomposition en images intrinsèques à partir d'une seule image RGB. On peut notamment citer le modèle SIRFS (Shape, Illumination and Reflectance From Shading) proposé par Barron *et al.* [92]. Dans ce modèle, la forme, la réflectance et l'illumination sont inconnues. La méthode proposée est basée sur des *a priori* sur les trois inconnues. Par exemple, les formes tendent à être lisses et isotropes. La réflectance tend à être parcimonieuse, c'est à dire que le nombre de couleurs utilisées pour décrire une scène tend à être petit. Les paramètres de ces *a priori* sont appris avec un sous-ensemble des bases de données d'images. La méthode demande donc une étape d'apprentissage.

Les méthodes plus récentes sont souvent basées sur l'apprentissage profond. Baslamisli *et al.*

ont proposé une architecture de type encodeur-décodeur pour estimer la réflectance et le shading à partir d'une image RGB [118]. L'entrée du réseau consiste en une paire d'images, l'image à décomposer et son gradient.

Il y a moins de travaux qui s'intéressent aux images multispectrales. Comme expliqué précédemment, les méthodes qui s'appliquent aux images multispectrales sont souvent basées sur la minimisation d'une fonction de coût. Par exemple Huang *et al.* utilisent le fait que les images se trouvent dans des sous-espaces de petites dimensions [128]. Leurs travaux n'utilisent pas le modèle dichromatique mais une extension de Retinex pour les images multispectrales appelée MIID. Chen *et al.* ont également travaillé sur les images multispectrales en utilisant des contraintes de parcimonie sur la réflectance S et des contraintes de lissage sur g [120].

En prenant en considération ces travaux de la littérature, nous proposons deux algorithmes, l'un basé sur de l'optimisation avec des fonctions de coût quadratiques et l'autre sur des déconvolutions. Ces algorithmes permettant d'obtenir la décomposition en images intrinsèques ne requièrent pas d'apprentissage et sont applicables en même temps aux images RGB et aux images multispectrales.

2.3/ INVERSION DU MODÈLE PAR OPTIMISATION QUADRATIQUE

Gu *et al.* ont utilisé des fonctions quadratiques notamment en raison de leur simplicité [105]. En effet, il existe une variété d'algorithmes d'optimisation pour trouver le minimum de fonctions quadratiques sous contraintes, entre autres les algorithmes du gradient conjugué, du gradient projeté, du lagrangien augmenté, du simplex ou l'algorithme des points intérieurs [33]. Gu *et al.* procèdent en deux étapes d'optimisation quadratique : dans un premier temps ils estiment g , puis dans un second temps, ils estiment k . S est alors facilement déduit par égalité du modèle. Nous nous inspirons de cette manière de faire et procédons également en deux étapes qui sont détaillées dans les sous-sections 2.3.2 et 2.3.3. La différence avec la méthode de Gu *et al.* est que nos deux étapes sont parfaitement indépendantes, il n'y a donc pas de propagation d'erreur et les deux étapes peuvent être facilement parallélisées.

La section est organisée en trois parties. Tout d'abord, nous définissons une nouvelle mesure de similarité dans la sous-section 2.3.1. Cette mesure de similarité est essentielle à la formulation de deux fonctions de coût quadratiques qui sont décrites dans les sous-sections 2.3.2 et 2.3.3.

2.3.1/ DÉTECTEUR DE PIXELS GRIS ET MESURE DE SIMILARITÉ ENTRE COULEURS

Comme le terme g est multiplicatif avec S dans l'équation 2.2, la mesure de similarité entre spectres doit être invariante à l'échelle. Nous avons donc décidé d'utiliser SAM (Spectral Angle Mapper en anglais) qui est justement invariant à l'échelle. SAM est défini comme l'arc cosinus du produit vectoriel entre deux vecteurs S_1 et S_2 divisé par leur norme euclidienne [66] :

$$SAM(S_1, S_2) = \arccos \left(\frac{\sum_{i=1}^{N_c} s_{1i} s_{2i}}{\sqrt{\sum_{i=1}^{N_c} s_{1i}^2} \sqrt{\sum_{i=1}^{N_c} s_{2i}^2}} \right) \quad (2.3)$$

où s_{1i} et s_{2i} sont respectivement les $i^{\text{ème}}$ composantes des vecteurs S_1 et S_2 et N_c le nombre d'éléments de ces vecteurs.

Pour S_1 et S_2 , deux vecteurs à valeurs positives, SAM retourne une valeur comprise entre 0 et $\frac{\pi}{2}$ 0 signifiant que S_1 et S_2 sont identiques à un facteur près.

En pratique, la mesure de similarité entre deux spectres S_1 et S_2 que nous utilisons est :

$$\psi_{S_1, S_2} = \exp\left(\frac{-SAM(S_1, S_2)^2}{2r^2}\right) \quad (2.4)$$

avec r le paramètre qui contrôle l'épaisseur de la gaussienne. Cette exponentielle négative nous permet d'obtenir des valeurs entre 0 et 1 et renforce la similarité lorsque les deux spectres sont proches.

De la même manière, nous pouvons déterminer si un pixel u est gris en comparant son spectre à un spectre blanc, c'est à dire, à un vecteur de 1. Nous noterons ce détecteur de pixel gris ζ_u :

$$\zeta_u = \psi_{u, 1}. \quad (2.5)$$

2.3.2/ ESTIMATION DE LA COMPOSANTE SHADING

La première étape de notre méthode consiste à estimer la composante shading g . Pour cela, nous allons étudier trois cas.

- Nous nous intéressons tout d'abord au cas où deux pixels voisins u et v appartiennent au même matériau. Dans ce cas, ils ont la même réflectance S *i.e.*

$$\forall \lambda, S(u, \lambda) = S(v, \lambda). \quad (2.6)$$

Il est intéressant de constater que le ratio de l'écart type de R selon l'axe spectral noté σ peut s'exprimer très simplement en fonction du ratio du shading g :

$$\alpha_{u,v} = \frac{\sigma_u}{\sigma_v} = \frac{g_u \hat{S}_u}{g_v \hat{S}_v} = \frac{g_u}{g_v} \quad (2.7)$$

où $\hat{S}_u$ désigne l'écart type de S selon l'axe spectral. Cette équation nous incite à écrire une fonction de coût comme un moindre carré de la différence entre g_u et $g_v \alpha_{u,v}$:

$$f_1(g) = \sum_u \sum_{v \in N(u)} \psi_{u,v} (g_u - g_v \alpha_{u,v})^2 \quad (2.8)$$

où $N(u)$ désigne le voisinage de u . Cette fonction minimise la distance entre g_u et $g_v \alpha_{u,v}$ seulement dans le cas où les pixels u et v ont la même réflectance et ce, grâce au terme $\psi_{u,v}$.

- Cette fonction de coût peut être complétée si l'on considère le cas où les pixels u et v n'appartiennent pas au même matériau. Nous utilisons alors le fait que l'image g tend à être lisse pour minimiser le carré de la différence entre g_u et g_v

$$f_2(g) = \sum_u \sum_{v \in N(u)} (1 - \psi_{u,v}) (g_u - g_v)^2. \quad (2.9)$$

Cette fois, la fonction minimise la distance entre g_u et g_v seulement dans le cas où les pixels u et v n'ont pas la même réflectance et ce, grâce au terme $1 - \psi_{u,v}$.

- Enfin, un troisième cas, plus complexe est à considérer. Il s'agit du cas où u et v sont des pixels gris. L'écart type est alors nul ou proche de zéro et le ratio $\alpha_{u,v}$ n'est pas calculable. De plus, si l'on se réfère à l'équation 2.2, un pixel gris est un pixel pour lequel $R(u, \lambda)$

est constant pour toutes les longueurs d'onde. $R(u, \lambda)$ ne peut alors pas être séparé en une composante diffuse et une composante spéculaire, car il ne dépend que de la position du pixel u et ne dépend pas de λ . En pratique, nous utilisons le ratio des moyennes μ à la place du ratio des écarts types :

$$\beta_{u,v} = \frac{\mu_u}{\mu_v}. \quad (2.10)$$

Cette affirmation correspond en fait au cas où la composante spéculaire serait nulle sur les objets gris, *i.e.* que $k = 0$. Cette considération nous fait ajouter une troisième partie à notre fonction de coût :

$$f_3(g) = \sum_u \sum_{v \in N(u)} \psi_{u,v} \zeta_u \zeta_v (g_u - g_v \beta_{u,v})^2. \quad (2.11)$$

f_1 et f_2 sont réécrites avec un facteur additionnel qui enlève les pixels gris :

$$f_1(g) = \sum_u \sum_{v \in N(u)} \psi_{u,v} \zeta_u \zeta_v (g_u - g_v \alpha_{u,v})^2, \quad (2.12)$$

$$f_2(g) = \sum_u \sum_{v \in N(u)} (1 - \psi_{u,v} \zeta_u \zeta_v - \psi_{u,v} \overline{\zeta_u \zeta_v}) (g_u - g_v)^2 \quad (2.13)$$

où $\overline{\zeta_u} = 1 - \zeta_u$.

Enfin, la fonction de coût est décrite par la somme des trois fonctions définies précédemment.

$$f_{shading}(g) = f_1(g) + f_2(g) + f_3(g). \quad (2.14)$$

On peut alors noter que, comme $f_{shading}$ est la somme de fonctions quadratiques, elle est donc elle-même quadratique et peut être écrite sous forme matricielle :

$$f_{shading}(g) = \mathbf{g}^t \times \mathbf{H}_{shading} \times \mathbf{g} \quad (2.15)$$

où $\mathbf{H}_{shading}$ est une matrice carrée $N_p \times N_p$ où N_p désigne le nombre de pixels dans l'image. Il se peut que $\mathbf{H}_{shading}$ ne soit pas symétrique. Il est préférable de travailler avec des matrices symétriques en utilisant $\frac{\mathbf{H}_{shading} + \mathbf{H}_{shading}^t}{2}$ à la place de $\mathbf{H}_{shading}$. Un avantage de cette formulation est que $\mathbf{H}_{shading}$ est une matrice creuse, c'est à dire une matrice dont seuls quelques coefficients sont non-nuls, puisque nous travaillons uniquement dans le voisinage de chaque pixel ; les interactions entre pixels plus éloignés ne sont pas considérées.

En considérant le modèle dichromatique, il pourrait y avoir n'importe quel facteur entre $\mathbf{g}$ et $\mathbf{S}$, *i.e.* si $\mathbf{S}^*$ et $\mathbf{g}^*$ sont solutions du problème inverse, alors $a\mathbf{S}^*$ et $\frac{\mathbf{g}^*}{a}$ sont également solutions du problème pour n'importe quel réel strictement positif a . Donc nous pouvons fixer la somme des éléments de $\mathbf{g}$ à une certaine constante c strictement positive sans biaiser le problème : nous écrivons alors la contrainte $\mathbf{g}^t \times \mathbf{1} = c$.

De plus, si l'on possède un masque binaire qui délimite l'objet pour valider ou invalider certains pixels, la contrainte devient $\mathbf{g}^t \times \mathbf{m} = c$ et $\mathbf{g}^t \times \overline{\mathbf{m}} = 0$ où $\mathbf{m}$ est un vecteur binaire et $\overline{\mathbf{m}}$ est l'inverse binaire de $\mathbf{m}$. Ces deux contraintes expriment le fait que la somme des pixels hors du masque est nulle et est égale à une constance c à l'intérieur du masque. $\mathbf{g}$ est alors finalement obtenu en résolvant

$$\left\{ \begin{array}{l} \mathbf{g}^* = \underset{\mathbf{g}}{\operatorname{argmin}} f_{shading}(\mathbf{g}) \\ \text{sous les contraintes} \\ \mathbf{g}^{*t} \times \mathbf{m} = c \\ \mathbf{g}^{*t} \times \overline{\mathbf{m}} = 0 \\ \mathbf{g}^* \geq 0. \end{array} \right. \quad (2.16)$$

2.3.3/ ESTIMATION DE LA COMPOSANTE SPÉCULAIRE

La deuxième étape de notre méthode est l'estimation de la composante spéculaire k . La méthode décrite dans cette sous-section est indépendante du calcul de g vu précédemment. Les deux optimisations sont donc parfaitement interchangeables et peuvent être parallélisées pour accélérer l'algorithme.

Pour commencer, nous définissons J le ratio entre R et son écart type σ :

$$J_u = \frac{R_u}{\sigma_u} \quad (2.17)$$

$$= \frac{g_u S_u}{\sigma_u} + \frac{k_u}{\sigma_u} \quad (2.18)$$

$$= \frac{S_u}{\hat{S}_u} + \frac{k_u}{\sigma_u}. \quad (2.19)$$

Nous définissons également un intermédiaire de calcul h comme :

$$h_u = \frac{k_u}{\sigma_u} - \overline{J_u} \quad (2.20)$$

où $\overline{J_u}$ est la moyenne dans l'axe spectral de J . Comme décrit dans l'équation 2.6, si deux pixels u et v appartiennent au même matériau, ils ont la même réflectance. Si de plus, on considère que k est lisse, alors la différence $h_u - h_v$ est nulle :

$$h_u - h_v = \frac{k_u}{\sigma_u} - \overline{J_u} - \frac{k_v}{\sigma_v} + \overline{J_v} \quad (2.21)$$

$$= \frac{k_u}{\sigma_u} - \frac{k_v}{\sigma_v} \quad (2.22)$$

$$= 0. \quad (2.23)$$

Comme précédemment, notre fonction de coût peut-être écrite comme le carré de la différence entre h_u et h_v pondérée par $\psi_{u,v}$ pour exprimer la similarité de matériau et $\overline{\zeta_u \zeta_v}$ pour enlever les pixels gris.

$$f_{\text{specular}}(h) = \sum_u \sum_{v \in N(u)} \psi_{u,v} \overline{\zeta_u \zeta_v} (h_u - h_v)^2. \quad (2.24)$$

Cette fonction de coût est également quadratique et peut être écrite sous la forme :

$$f_{\text{specular}}(h) = \mathbf{h}^t \times \mathbf{H}_{\text{specular}} \times \mathbf{h} \quad (2.25)$$

avec $\mathbf{H}_{\text{specular}}$ une matrice symétrique $N_p \times N_p$.

Pour des raisons purement physiques, k ne peut être négatif. De plus, si l'on regarde le minimum de R dans l'axe des longueurs d'onde, on s'aperçoit que k_u possède également une borne supérieure qui est $\min_{\lambda}(R_u)$:

$$\min_{\lambda}(R_u) = g_u \min_{\lambda}(S_u) + k_u. \quad (2.26)$$

Les contraintes physiques bornent k_u entre 0 et $\min_{\lambda}(R_u)$. Donc, en reprenant l'équation 2.20, h_u est également borné :

$$-\bar{J}_u \leq h_u \leq \frac{\min(R_u)}{\sigma_u} - \bar{J}_u. \quad (2.27)$$

h s'obtient alors en résolvant le système suivant :

$$\begin{cases} h^* = \underset{h}{\operatorname{argmin}} f_{\text{specular}}(h) \\ \text{sous la contrainte} \\ -\bar{J} \leq h^* \leq \frac{\min(R)}{\sigma} - \bar{J}. \end{cases} \quad (2.28)$$

Une fois que g et h ont été estimés, k est obtenu grâce à

$$k = h + \bar{J}\sigma \quad (2.29)$$

et S par

$$S = \frac{R - k}{g}. \quad (2.30)$$

Pour résumer, nous avons proposé une méthode pour estimer g et k en cherchant le minimum de fonctions quadratiques sous contraintes. S est ensuite déduit par égalité du modèle. Nous présentons les résultats de cette méthode dans la section 2.5 en même temps que les résultats de la seconde méthode présentée dans la section 2.4. Cette seconde méthode vise à améliorer la méthode par optimisation quadratique d'un point de vue stabilité numérique et vitesse de calcul.

2.4/ INVERSION DU MODÈLE PAR DÉCONVOLUTION

Dans cette section, nous proposons une deuxième méthode pour estimer les invariants photométriques. Cette nouvelle méthode se base sur les mêmes observations que la méthode précédente mais elle diffère dans le sens où l'on travaille avec le logarithme de l'image et que la méthode s'écrit sous la forme de deux déconvolutions.

2.4.1/ DÉFINITIONS ET HYPOTHÈSES

Nous avons déjà vu l'intérêt d'utiliser la moyenne μ et l'écart type σ de l'image R selon l'axe spectral. Ces deux images en niveaux de gris ont des propriétés intéressantes par rapport au modèle dichromatique. En particulier, σ se trouve être indépendant du spéculaire k .

$$\mu = g\bar{S} + k, \quad (2.31)$$

$$\sigma = g\hat{S}. \quad (2.32)$$

$\bar{S}$ et $\hat{S}$ sont respectivement la moyenne et l'écart type de S . Certains chercheurs ont proposé de travailler avec le logarithme des images. On peut notamment citer Barron *et al.* [92] et Robles-Kelly et Huynh [70]. Le logarithme transforme les multiplications en additions et, de ce fait, linéarise

le problème par rapport au shading et à la réflectance. Nous réécrivons donc les équations 2.31 et 2.32 en appliquant le logarithme :

$$\log(\mu) = \log(g\bar{S}) + \log\left(1 + \frac{k}{g\bar{S}}\right) = \log(g) + \log\left(\bar{S} + \frac{k}{g}\right), \quad (2.33)$$

$$\log(\sigma) = \log(g) + \log(\hat{S}). \quad (2.34)$$

Cette formulation nous mène à estimer indirectement g , S et k . En particulier, nous nous intéressons à $\log(\hat{S})$ et $\log(g\bar{S})$. L'estimation de ces deux images en niveaux de gris est suffisante pour inverser tout le modèle dichromatique et estimer g , S et k .

À partir des équations 2.33 et 2.34, nous considérons une nouvelle fois les trois cas exprimés dans la méthode précédente : soit les pixels voisins u et v appartiennent au même matériau (*i.e.* $S(u, \lambda) = S(v, \lambda)$), soit ils appartiennent à des matériaux différents, soit les pixels sont gris. Les trois points suivants détaillent chacun de ces trois cas en estimant les différences $\log(\hat{S})_u - \log(\hat{S})_v$ et $\log(g\bar{S})_u - \log(g\bar{S})_v$, c'est à dire, les images $\log(\hat{S})$ et $\log(g\bar{S})$ aux pixels u et v .

- Dans le cas où les pixels appartiennent au même matériau, S a la même valeur aux pixels u et v . Nous pouvons donc écrire

$$\log(\hat{S})_u - \log(\hat{S})_v = 0. \quad (2.35)$$

De la même manière,

$$\begin{aligned} \log(g\bar{S})_u - \log(g\bar{S})_v &= \log(g)_u - \log(g)_v \\ &= \log(\sigma)_u - \log(\sigma)_v. \end{aligned} \quad (2.36)$$

- Si u et v n'appartiennent pas au même matériau, alors, nous utilisons le fait que g tend à être lisse.

$$\begin{aligned} \log(\hat{S})_u - \log(\hat{S})_v &= \log(\sigma)_u - \log(\sigma)_v - \log(g)_u + \log(g)_v \\ &\simeq \log(\sigma)_u - \log(\sigma)_v. \end{aligned} \quad (2.37)$$

Et pour $\log(g\bar{S})$, nous considérons que l'influence du spéculaire est négligeable :

$$\log(g\bar{S})_u - \log(g\bar{S})_v \simeq \log(\mu)_u - \log(\mu)_v. \quad (2.38)$$

- Finalement, si u et v sont des pixels gris, il est mathématiquement impossible de déterminer g , S et k . Dans ce cas précis, nous considérons également que le spéculaire est négligeable :

$$\begin{aligned} \log(\hat{S})_u - \log(\hat{S})_v &= \log(\sigma)_u - \log(\sigma)_v - \log(g)_u + \log(g)_v \\ &\simeq \log(\sigma)_u - \log(\sigma)_v - \log(\mu)_u + \log(\mu)_v. \end{aligned} \quad (2.39)$$

De manière similaire :

$$\log(g\bar{S})_u - \log(g\bar{S})_v \simeq \log(\mu)_u - \log(\mu)_v. \quad (2.40)$$

Les différences $\log(\hat{S})_u - \log(\hat{S})_v$ et $\log(g\bar{S})_u - \log(g\bar{S})_v$ déclinées selon les trois cas sont utilisées dans la partie suivante pour formuler le problème sous forme de déconvolution.

2.4.2/ PERFECTIONNEMENT DU DÉTECTEUR DE PIXELS GRIS ET DE LA MESURE DE SIMILARITÉ

Comme nous l'avons vu dans la méthode précédente, il est indispensable de détecter les pixels gris et d'avoir une mesure de similarité entre couleur. Pour cela, nous utilisons l'angle spectral SAM déjà utilisé dans la méthode par optimisation quadratique.

Étant donnés deux pixels, une valeur entre 0 et $\frac{\pi}{2}$ est retournée. Dans la première méthode, nous utilisons une gaussienne pour obtenir des valeurs entre 0 et 1, 1 désignant une forte similitude entre les spectres. Ici, nous avons décidé d'utiliser une fonction sigmoïde à la place de la gaussienne. La sigmoïde permet de modéliser un seuillage plus ou moins doux :

$$\text{sig}_{\gamma,\beta}(x) = \frac{1}{1 + \exp^{-\gamma(x-\beta)}}, \quad (2.41)$$

β est la valeur du seuil entre 0 et $\frac{\pi}{2}$. Plus β est grand, plus le seuillage est permissif. Le paramètre gamma γ contrôle le comportement du seuillage : plus γ est grand, plus on se rapproche d'un seuillage dur, plus gamma est petit, plus le seuillage est doux.

Finalement l'indice ζ indique si un pixel est gris :

$$\zeta_u = \text{sig}_{\gamma,\beta}(\text{SAM}(R_u, \mathbf{1})), \quad (2.42)$$

où $\mathbf{1}$ est un vecteur de 1. De la même manière, la similarité entre deux couleurs peut être évaluée par :

$$\psi_{u,v} = \text{sig}_{\gamma,\beta}(\text{SAM}(R_u, R_v)). \quad (2.43)$$

2.4.3/ FORMULATION PAR DÉCONVOLUTION

En utilisant ces deux outils, nous pouvons écrire une somme pondérée des équations 2.35 (première ligne), 2.39 (deuxième ligne) et 2.37 (troisième ligne) telle que :

$$\begin{aligned} \log(\hat{S})_u - \log(\hat{S})_v &\simeq (0) [\psi_{u,v}(1 - \zeta_u)(1 - \zeta_v)] \\ &+ (\log(\sigma)_u - \log(\sigma)_v - \log(\mu)_u - \log(\mu)_v) [\psi_{u,v}\zeta_u\zeta_v] \\ &+ (\log(\sigma)_u - \log(\sigma)_v) [1 - \psi_{u,v}(1 - \zeta_u)(1 - \zeta_v) - \psi_{u,v}\zeta_u\zeta_v]. \end{aligned} \quad (2.44)$$

Les poids entre crochets ont été choisis de sorte que leur somme fasse 1.

De manière analogue, nous pouvons écrire une somme pondérée des équations 2.36 (première ligne), 2.38 et 2.40 (deuxième ligne) de façon à obtenir la relation pour $\log(g\bar{S})$:

$$\begin{aligned} \log(g\bar{S})_u - \log(g\bar{S})_v &\simeq (\log(\sigma)_u - \log(\sigma)_v) [\psi_{u,v}(1 - \zeta_u)(1 - \zeta_v)] \\ &+ (\log(\mu)_u - \log(\mu)_v) [1 - \psi_{u,v}(1 - \zeta_u)(1 - \zeta_v)]. \end{aligned} \quad (2.45)$$

Pour arriver à formuler le problème sous forme de déconvolution, il suffit de sommer les relations 2.44 et 2.45 pour tous les pixels $v = u + du$ du voisinage de u . La somme nous donne une estimation du laplacien de $\log(\hat{S})$ et de $\log(g\bar{S})$.

$$\begin{aligned}
8 \log(\hat{S})_u - \sum_{du \neq 0} \log(\hat{S})_{u+du} \simeq \\
\sum_{du \neq 0} \left((0) \psi_{u,u+du} (1 - \zeta_u) (1 - \zeta_{u+du}) \right. \\
+ (\log(\sigma)_u - \log(\sigma)_{u+du} - \log(\mu)_u + \log(\mu)_u + du) \psi_{u,v} \zeta_u \zeta_{u+du} \\
\left. + (\log(\sigma)_u - \log(\sigma)_{u+du}) (1 - \psi_{u,u+du} (1 - \zeta_u) (1 - \zeta_{u+du}) - \psi_{u,u+du} \zeta_u \zeta_{u+du}) \right)
\end{aligned} \quad (2.46)$$

que nous notons de manière plus simple

$$8 \log(\hat{S})_u - \sum_{du \neq 0} \log(\hat{S})_{u+du} \simeq \mathcal{A} \quad (2.47)$$

et

$$\begin{aligned}
8 \log(g\bar{S})_u - \sum_{du \neq 0} \log(g\bar{S})_{u+du} \simeq \\
\sum_{du \neq 0} \left((\log(\sigma)_u - \log(\sigma)_{u+du}) \psi_{u,u+du} (1 - \zeta_u) (1 - \zeta_{u+du}) \right. \\
\left. + (\log(\mu)_u - \log(\mu)_{u+du}) (1 - \psi_{u,u+du} (1 - \zeta_u) (1 - \zeta_{u+du})) \right)
\end{aligned} \quad (2.48)$$

que nous notons de manière plus simple

$$8 \log(g\bar{S})_u - \sum_{du \neq 0} \log(g\bar{S})_{u+du} \simeq \mathcal{B}. \quad (2.49)$$

Nous obtenons deux relations pour lesquelles l'inconnue est une image et dont nous avons une estimation de la convolution par le filtre linéaire $\mathcal{F}$:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$

Comme le filtre laplacien est le filtre identité moins un filtre moyennneur, nous pouvons envisager une forme plus générale et utiliser le filtre identité dont on soustrait un filtre passe-bas de la taille désirée. Ceci simplifie le problème à la résolution de deux déconvolutions pour estimer $\log(\hat{S})$ et $\log(g\bar{S})$:

$$\log(\hat{S}) * \mathcal{F} = \mathcal{A}, \quad (2.50)$$

$$\log(g\bar{S}) * \mathcal{F} = \mathcal{B}. \quad (2.51)$$

Les déconvolutions sont en soit aussi résolues par la méthode des moindres carrés (optimisation quadratique). À noter qu'il faut respecter les contraintes de positivité pour g , S et k . Le logarithme rend cette contrainte inutile pour $\log(\hat{S})$ mais doit toujours être considérée pour $\log(g\bar{S})$. En effet,

$$\log(\mu - \min_{\lambda} R(:, \lambda)) < \log(g\bar{S}) < \log(\mu). \quad (2.52)$$

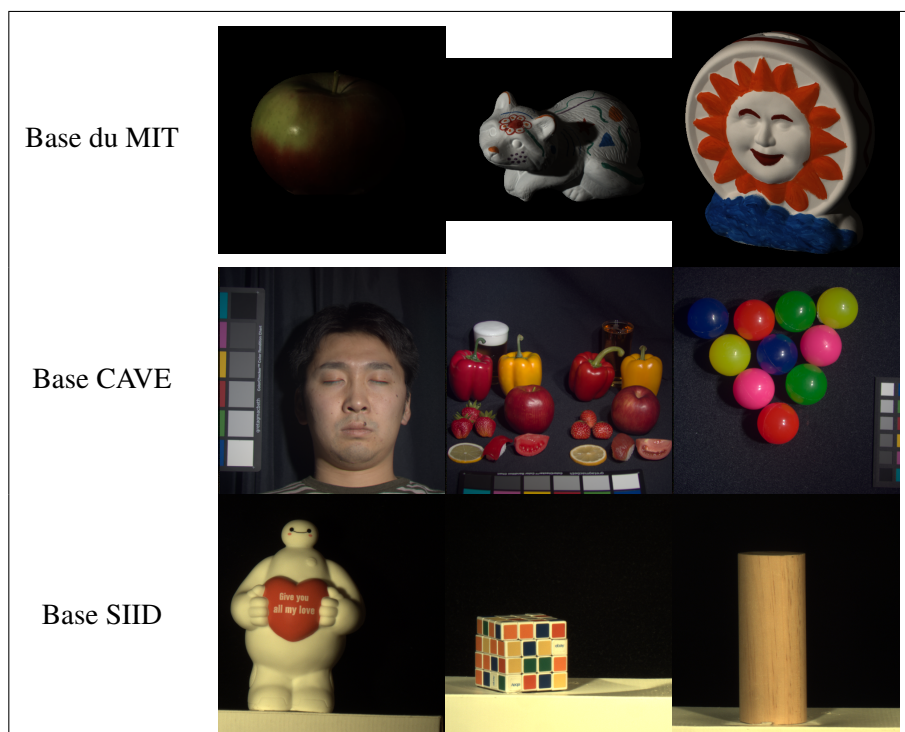
Les avantages de cette méthode sont que la décomposition est réduite à deux déconvolutions d'images en niveaux de gris indépendantes, ce qui la rend significativement plus rapide. De plus, les déconvolutions peuvent être parallélisées et la méthode est indépendante du nombre de longueurs d'onde de l'image initiale car seuls la moyenne et l'écart type dans l'axe spectral sont requis. De plus, la méthode a été construite pour contrer les limitations courantes, à savoir qu'elle est aussi bien applicable aux images multispectrales qu'aux images RGB et ne requiert pas d'apprentissage. Dans la section suivante, nous évaluons de manière quantitative et qualitative les résultats

de ces deux méthodes et les comparons à ceux de l'état de l'art. Trois bases de données différentes, une base d'images RGB et deux bases d'images multispectrales sont utilisées pour tester la robustesse de notre algorithme.

2.5/ RÉSULTATS COMPARATIFS ET DISCUSSION

Nous avons utilisé trois bases de données différentes pour tester nos algorithmes. La base du MIT (Massachusetts Institute of Technology), la base CAVE et la base SIID. La base de données du MIT est composée d'un ensemble de 20 images RGB et leur vérité terrain de décomposition en images intrinsèques [43]. CAVE est une base de données d'images multispectrales prises sous l'illuminant D65 mais il n'existe pas de vérité terrain [56]. Enfin, nous avons appliqué notre algorithme aux images de SIID (Spectral Intrinsic Ground-truth Dataset). Cette base d'images a été créée en 2017 par Chen *et al.* et est la toute première base de données multispectrales contenant une vérité terrain [120]. Les images ne contiennent pas de composante spéculaire. Les images ont été acquises sous deux sources de lumière différentes : des LEDs (Diodes ElectroLuminescentes) avec une température de 5500K et une source iode-tungstène (ITL). Des images provenant de ces trois bases de données se trouvent figure 2.1 pour illustrer la diversité des bases et des images.

TABLE 2.1 – Exemples d'images des trois bases de données : de gauche à droite et de haut en bas, une pomme, un raton-laveur et un soleil de la base du MIT ; un visage, de la nourriture réelle et factice et des balles de la base CAVE ; et Baymax, un cube et un cylindre de la base SIID.



Les deux méthodes proposées sont comparées à six autres méthodes de l'état de l'art. Parmi ces méthodes, trois ne requièrent pas d'apprentissage et trois autres sont dépendantes d'un apprentissage. Les trois méthodes n'étant pas basées sur un apprentissage sont notées de la manière suivante :

- *HUY* : Huynh *et al.* ont proposé une solution pour inverser le modèle dichromatique par optimisation. En l'occurrence, la fonction de coût est basée sur l'équation qui définit le modèle dichromatique et sur l'addition d'un terme de régularisation qui force le lissage de l'image g [51]. Nous utilisons l'implémentation donnée par les auteurs via la librairie Scyllarus [7].
- *GU* : Gu *et al.* ont utilisé des "shapelets". La réflectance et les coefficients spéculaires sont retrouvés par optimisation avec contraintes. Les fonctions de coûts sont sous la forme des moindres carrés [75]. Nous utilisons l'implémentation donnée par les auteurs via la librairie Scyllarus [7].
- *YAN* : Yang *et al.* se sont intéressés à la fraction de couleur diffuse dans une couleur. Cette fraction varie localement de manière très lisse. Ils ont utilisé cette observation pour créer une méthode basée sur un filtrage passe-bas qui préserve les bords, notamment en utilisant le filtrage bilatéral [100]. Nous avons ré-implémenté cette méthode par nos propres moyens en utilisant l'article original.

Les trois autres méthodes sont :

- *BAR* : Barron *et al.* ont proposé leur propre modèle nommé SIRFS. Les images intrinsèques que sont la réflectance, la profondeur et l'illumination de l'image sont retrouvées par optimisation de fonction de coût reflétant la réalité des caractéristiques des images intrinsèques. Par exemple, le nombre de couleurs différentes dans une image tend à être petit (contrainte de parcimonie) et les surfaces tendent à être lisses (contrainte de lissage) [92]. L'implémentation est fournie par l'auteur sur sa page personnelle [5].
- *NAR* : Narihira *et al.* ont proposé une approche basée sur de l'apprentissage profond [98]. Le réseau a été entraîné sur la base MPI Sintel [65, 72] et est capable de séparer une image RGB en son shading et sa réflectance. L'auteur a fourni une version Python de son implémentation en ligne [1].
- *BAS* : Baslamisli *et al.* ont également proposé un réseau profond. L'architecture employée est un encodeur-décodeur qui sépare l'image en sa réflectance et son shading. Ils ont testé deux structures différentes : IntrinsicNet pour lequel l'entrée du réseau est l'image à décomposer et RetiNet pour lequel les entrées du réseau sont l'image à décomposer et son gradient [118]. L'implémentation n'est pas disponible en ligne mais les images-résultats nous ont été gracieusement partagées par l'auteur.

Parmi ces méthodes, seules celles proposées par *GU*, *HUY* sont applicables aux images multispectrales (c'est-à-dire aux bases de données CAVE et SIID) et toutes les méthodes sont applicables à la base de données du MIT.

Pour comparer proprement les différentes méthodes, le choix des métriques est essentiel. Comme le shading g et la réflectance S sont des termes multiplicatifs dans l'équation simplifiée 2.2. Ils ne peuvent être estimés qu'à un scalaire près. Donc au lieu de calculer l'erreur quadratique moyenne (MSE), l'erreur entre g et sa vérité terrain g_{GT} (respectivement S et S_{GT}) est calculée comme le minimum de :

$$err = \min_{\alpha} \sum_u (\alpha g(u) - g_{GT}(u))^2 \quad (2.53)$$

pour tout scalaire positif α . De plus Barron *et al.* [92] et Grosse *et al.* [43] suggèrent qu'il est mieux de considérer une erreur quadratique local (LMSE). Le LMSE est la somme des MSE normalisés sur toutes les fenêtres de taille $w \times w$ par pas de $w/2$. En pratique, nous utilisons une fenêtre de taille $w = 20$ comme proposé dans l'article de Grosse *et al.* [43].

Pour le spéculaire k , le problème du facteur multiplicatif ne se pose pas. Nous utilisons donc un MSE classique. Enfin, pour les images RGB, nous pouvons également calculer le ΔE_{ab} sur

l'image de réflectance. Il s'agit de la distance euclidienne calculée dans l'espace $L^*a^*b^*$. Le ΔE_{ab} est généralement plus proche de la perception humaine que le MSE dans l'espace RGB.

2.5.1/ RÉSULTATS SUR LA BASE DE DONNÉES DU MIT

Dans leur article, Barron *et al.* ont travaillé sur deux versions différentes de la base de données du MIT, l'une sous illumination naturelle et l'autre sous illumination de laboratoire [92]. Les algorithmes ont été testés sur les deux versions de la base de données. Une des différences est qu'il n'y a pas de composante spéculaire dans les images sous illumination naturelle et donc, nous ne calculons pas de métrique sur k dans ce cas là.

Les résultats quantitatifs sont présentés dans le tableau 2.2. Pour chaque métrique, les erreurs les plus basses sont présentées en gras. Le tableau montre la supériorité de *NAR* pour l'estimation de la réflectance S avec un LMSE de $0,22 \times 10^{-4}$ et un ΔE_{ab} de 4,13. Pour autant, la deuxième méthode que nous proposons est plus performante pour l'estimation de g et k (un LMSE sur g de $0,48$ et un MSE sur k de $1,8 \times 10^5$). Les deux méthodes proposées occupent les deux premières places pour l'estimation de g ($0,48 \times 10^{-4}$ et $0,51 \times 10^{-4}$) et les deuxième et troisième places pour l'estimation de S ($0,27 \times 10^{-4}$ et $0,30 \times 10^{-4}$). Les résultats sont un peu moins bons avec l'illuminant naturel mais restent proches de l'état de l'art et rivalisent avec les meilleures méthodes qui utilisent un apprentissage.

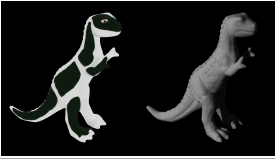
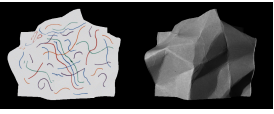
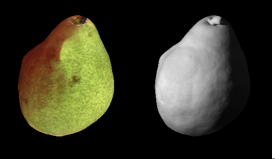
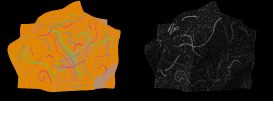
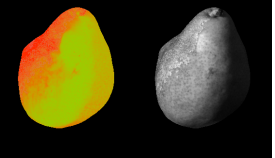
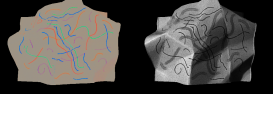
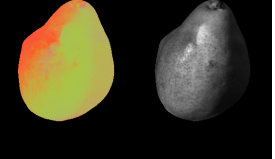
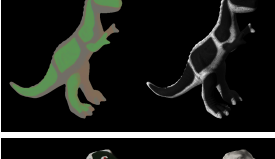
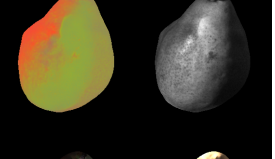
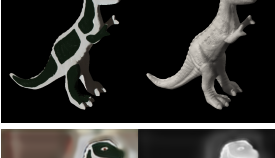
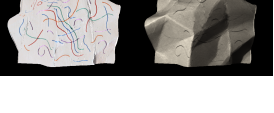
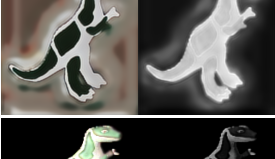
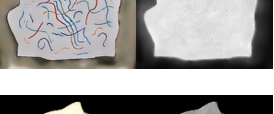
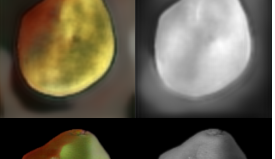
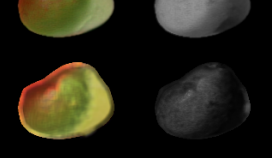
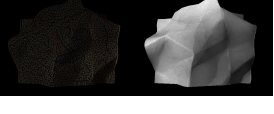
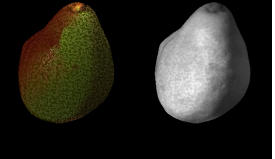
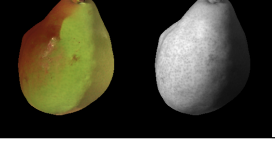
Il y a eu une nette amélioration des résultats entre la méthode par optimisation quadratique et la méthode par déconvolution. La deuxième méthode par déconvolution apporte de la stabilité numérique de par l'utilisation de l'espace logarithmique et de meilleurs seuillages grâce à la fonction sigmoïde présentée avec l'équation 2.41.

Les résultats quantitatifs peuvent s'expliquer visuellement en regardant l'ensemble des figures 2.3. Elles représentent le résultat de sortie de chaque méthode, la première ligne étant la vérité terrain. Pour certaines de ces méthodes, le shading est mal estimé notamment à cause des changements de couleurs brusques. Par exemple, le dinosaure est principalement composé de deux couleurs, le vert et le blanc. Pour *GU*, *YAN* et *BAS*, ces changements de couleurs se voient dans l'image de shading alors que ça ne devrait pas être le cas. Le shading est indépendant de la couleur des objets. Le même phénomène peut être observé sur l'image représentant une feuille de papier. Les lignes de couleur sur le papier ne devraient pas être visibles sur l'image de shading. De plus, certaines couleurs ne semblent pas bien rendues pour *HUY*, *GU* et *YAN*. Le vert et le blanc du dinosaure semblent faux, notamment le blanc qui apparaît orange ou gris. Nous notons également que les images produites par *BAS* n'ont pas la même taille que les autres. En effet, la plupart des réseaux profonds ne fonctionnent qu'avec des images de taille spécifique en entrée. Ceci est une autre limitation des réseaux profonds.

TABLE 2.2 – Résultats quantitatifs sur la base de données du MIT.

	Métrique	<i>HUY</i>	<i>GU</i>	<i>YAN</i>	<i>BAR</i>	<i>NAR</i>	BAS (IntrinsicNet)	BAS (RetiNet)	Méthode par optimisation quadratique	Méthode par déconvolution
MIT illuminant laboratoire	LMSE sur g ($\times 10^{-4}$)	1,86	6,02	2,30	0,65	1,04	1,17	1,38	0,51	0,48
	LMSE sur S ($\times 10^{-4}$)	0,50	2,11	0,55	0,56	0,22	0,43	0,54	0,30	0,27
	MSE on k ($\times 10^7$)	0,12	3,89	2,06	0,028	0,028	0,028	0,028	7,88	0,018
	ΔE_{ab} on S	8,35	17,3	9,03	6,83	4,13	7,21	8,03	13,6	7,04
MIT illuminant naturel	LMSE sur g ($\times 10^{-4}$)	2,06	3,88	2,33	0,20	0,68	1,06	1,44	0,48	0,53
	LMSE sur S ($\times 10^{-4}$)	0,71	1,77	0,73	0,33	0,22	0,6	0,72	2,97	0,35
	ΔE_{ab} sur S	10,3	15,6	10,9	4,36	4,03	9,30	9,62	13,2	9,12

TABLE 2.3 – Résultats sur trois images de la base du MIT : de gauche à droite, successivement la réflectance et le shading des images du dinosaure, du papier et de la poire.

Vérité Terrain			
<i>HUY</i>			
<i>GU</i>			
<i>YAN</i>			
<i>BAR</i>			
<i>NAR</i>			
BAS (IntrinsicNet)			
BAS (RetiNet)			
Méthode par optimisation quadratique			
Méthode par déconvolution			

2.5.2/ RÉSULTATS SUR LA BASE DE DONNÉES SIID

La base de données SIID contient les vérités terrain pour les images de réflectance et de shading. Les images de cette base ne présentent pas de composante spéculaire ; il n'est donc pas étudié ici. Dans la deuxième méthode proposée, nous pouvons choisir de n'effectuer que la première déconvolution donnée à l'équation 2.50. En l'absence de spéculaire, résoudre cette équation suffit car nous nous retrouvons dans le cas où $R = gS$. La méthode est donc simplifiée en l'absence de composante spéculaire.

Le tableau 2.4 rassemble les LMSE sur g et S pour toutes les méthodes. Les deux méthodes proposées sont au-dessus de l'état de l'art sur les quatre métriques et particulièrement pour l'estimation du shading. La méthode par déconvolution affiche les erreurs les plus basses : par exemple pour la base d'images sous ITL, nous avons obtenu un LMSE sur g de $0,08 \times 10^{-4}$ et un LMSE sur S de $0,25 \times 10^{-5}$. Sous les LEDs, nous avons réussi à obtenir un LMSE sur g de $0,09 \times 10^{-4}$ et un LMSE sur S de $0,51 \times 10^{-5}$.

TABLE 2.4 – Résultats sur la base SIID.

Illuminant	Métrique	HUY	GU	Méthode par optimisation quadratique	Méthode par déconvolution
SIID sous ITL	LMSE sur g ($\times 10^{-4}$)	0,32	0,41	0,09	0,08
	LMSE sur S ($\times 10^{-5}$)	0,27	0,27	0,27	0,25
SIID sous LEDs	LMSE sur g ($\times 10^{-4}$)	0,45	0,55	0,14	0,09
	LMSE sur S ($\times 10^{-5}$)	0,083	0,086	0,053	0,051

Comme pour la base du MIT, nous observons également les résultats d'un point de vue qualitatif. Malheureusement, il est impossible d'afficher une image multispectrale directement, c'est pour cela que nous appliquons une transformation linéaire pour retrouver les trois bandes RGB usuelles. Dans le tableau 2.5 se trouvent les images résultantes de toutes les méthodes appliquées à trois images : Baymax, un cube et un hibou. En observant l'image de Baymax, nous pouvons observer que tous les résultats sont affectés par les changements brusques de couleur, notamment entre le corps et le cœur de Baymax. On passe du blanc au rouge et c'est pour cela que le cœur apparaît plus sombre qu'il ne le devrait. Cette observation est moins marquée pour nos deux méthodes mais reste présente. Comme la surface du cœur est lisse, les lettres ne devraient pas apparaître non plus dans l'image de shading. Pour le cube, les méthodes semblent estimer une couleur correcte mais pourtant, l'image de shading, elle, ne l'est pas. Comme les faces du cube sont lisses, elles doivent avoir un shading uniforme comme le montre la vérité terrain. De la même manière, le contraste, autour des yeux du hibou, ne devraient pas apparaître dans l'image de shading.

TABLE 2.5 – Résultats sur trois images de la base SIID : de gauche à droite, successivement la réflectance et le shading des images de Baymax, du cube et d'un hibou.

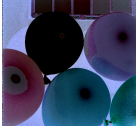
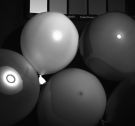
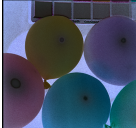
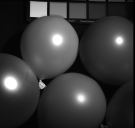
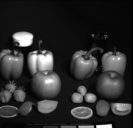
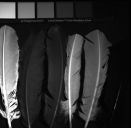
Vérité Terrain						
<i>HUY</i>						
<i>GU</i>						
Méthode par optimisation quadratique						
Méthode par déconvolution						

2.5.3/ RÉSULTATS SUR LA BASE DE DONNÉES CAVE

La base de données nommée CAVE ne contient pas de vérité terrain. Nous ne pouvons présenter ici qu'une analyse qualitative des résultats. Le tableau 2.6 présente les images résultantes de toutes les méthodes sur quelques exemples de la base CAVE.

L'image avec les ballons nous montre clairement des taches spéculaires. Même sans avoir de vérité terrain, nous savons que ces taches ne devraient pas apparaître ni sur l'image de shading, ni sur l'image de réflectance. Ces taches apparaissent pourtant pour *HUY* et *GU* et est présente, bien que de manière moins significative, sur les deux méthodes que nous avons proposées. Nous pouvons également observer que, sur l'image avec les plumes, la plume de gauche, qui devrait être blanche, apparaît noire pour *HUY* et *GU*. Nous faisons exactement l'observation inverse sur la plume de droite, qui devrait être noire et qui apparaît blanche pour *HUY* et *GU*. Ce problème est dû à une mauvaise estimation du shading. Toutes les plumes devraient à peu près avoir le même shading parce que la scène est plate, mais pour *HUY* et *GU*, ce n'est pas le cas.

TABLE 2.6 – Résultats sur trois images de la base CAVE : de gauche à droite, successivement la réflectance et le shading des images de ballons, de vrais et faux aliments et des plumes.

Image d'entrée			
<i>HUY</i>			
<i>GU</i>			
Méthode par optimisation quadratique			
Méthode par déconvolution			

2.6/ CONCLUSION ET PERSPECTIVES

En conclusion, nous avons proposé deux méthodes pour réaliser la décomposition en images intrinsèques, l'une par optimisation quadratique et son amélioration, écrite sous forme de déconvo-

lution. Les méthodes rivalisent avec l'état de l'art voire le dépassent. Elles font parties des rares méthodes applicables aux images RGB et multispectrales et qui, de surcroît, ne demandent pas d'apprentissage.

Néanmoins, les deux méthodes se basent sur des *a priori*, notamment l'affirmation que l'image du shading est lisse et qu'il n'y a pas de spéculaire sur les zones grises. Ces affirmations pourraient être allégées voire supprimées, notamment si l'on utilise plusieurs images montrant la même scène sous des angles différents. Nous serions alors dans le cas où l'indétermination serait allégée par l'addition de données. Une extension de ces travaux serait donc de modifier l'algorithme pour avoir l'appliquer, non pas à une image mais à une série d'image comme dans les vidéos endoscopiques acquises chez l'Homme.

Dans le chapitre suivant, nous appliquons la décomposition en images intrinsèques aux images endoscopiques acquises dans l'estomac. Cette étape peut être vue comme un pré-traitement des images qui peut faciliter la classification.

DÉTECTION DE LÉSIONS PRÉ-CANCÉREUSES DANS LES IMAGES ENDOSCOPIQUES NBI

Le système installé à l'hôpital Ambroise Paré nous permet d'enregistrer des images NBI et des images multispectrales lors d'une exploration *in vivo* de l'estomac. Dans ce chapitre, nous analysons uniquement les images NBI pour estimer leur potentiel de détection des lésions de l'estomac.

Nous avons ainsi, avec l'aide d'un gastroentérologue expérimenté, annoté des patchs caractéristiques de certaines lésions de l'estomac comme la gastrite chronique, la gastrite active, l'atrophie et la métaplasie intestinale. Ces patchs nous servent ensuite à entraîner un classifieur à reconnaître les différents types de lésions pré-cancéreuses.

En parallèle, nous avons également développé cinq critères visuels que le gastroentérologue utilise de manière "inconsciente" pour établir un diagnostic à partir d'images : il s'agit de la couleur de la muqueuse, de la régularité de la trame, de l'épaisseur de la trame, de la texture à l'intérieur de la trame et de la densité des vaisseaux sanguin.

Sommaire

3.1	Traitement d'images pour la détection de lésions de l'estomac par endoscopie	34
3.2	Acquisition des données endoscopiques	35
3.3	Méthodologie	36
3.3.1	Classification par patchs caractéristiques	36
3.3.2	Application du modèle dichromatique	40
3.3.3	Association des résultats à la mosaïque	43
3.4	Détermination de critères visuels	43
3.5	Conclusion et perspectives	48

3.1/ TRAITEMENT D'IMAGES POUR LA DÉTECTION DE LÉSIONS DE L'ESTOMAC PAR ENDOSCOPIE

Le traitement automatique d'images est souvent utilisé dans le domaine médical pour détecter un large nombre de pathologies. On peut distinguer les méthodes par apprentissage profond des méthodes plus conventionnelles de traitement d'images se basant sur des descripteurs de couleur ou de texture.

Dans les descripteurs de couleur, on trouve notamment les histogrammes de couleur. Ils sont souvent calculés dans l'espace HSV qui a l'avantage d'être un espace où les composantes sont décorréliées. Bashar *et al.* ont utilisé ces descripteurs de couleur ainsi que des descripteurs de texture pour sélectionner des images informatives provenant d'une capsule endoscopique, notamment pour supprimer les images contenant des bulles [38].

Les descripteurs de textures regroupent entre autres, les Local Binary Patterns (LBP) [48]. Li *et al.* ont pu détecter des ulcères dans des images de capsules endoscopiques grâce à une transformée en curvelet et le calcul de LBP dans l'espace YCbCr. Les descripteurs ainsi calculés sont utilisés pour entraîner un réseau de neurones à trois couches [45, 53].

Il existe bon nombre de variantes des Local Binary Patterns. Dans un contexte endoscopique, Hafner *et al.* ont développé les Local Color Vector Patterns operator (LCVP) [68] qui sont une extension couleur des LBP dans laquelle chaque pixel est considéré comme un vecteur. Les LBP sont alors calculés sur une distance entre vecteurs. Les LCVP ont ainsi permis de classer cinq types de motifs que l'on trouve sur la muqueuse du colon.

Iakovidis *et al.* ont testé différentes versions de la transformée en ondelettes pour la détection d'adénomes dans l'estomac [31]. On peut par exemple citer les WCS (Wavelet Correlation Signatures) qui prennent en considération les corrélations entre les trois canaux de l'image et les CWC (Color Wavelet Covariance) qui consistent à calculer la transformée en ondelettes puis de garder les informations statistiques de second ordre des coefficients des ondelettes au moyen des matrices de co-occurrence.

Les matrices de co-occurrence (GLCM) sont une autre catégorie de descripteurs de texture. Kanesaka *et al.* ont montré que l'on pouvait détecter et délimiter des cancers de l'estomac dans les images NBI avec grossissement [130] en calculant les 8 descripteurs proposés originellement dans [18] sur les matrices de co-occurrence. Shen *et al.* ont également utilisé les GLCM et les ont comparés aux LBP. Dans les deux cas, l'extraction de descripteurs est suivie par une transformée en contourlet et d'une classification visant à détecter les lésions de l'estomac [71].

De manière générale, ces méthodes permettent d'extraire des descripteurs liés à la perception visuelle humaine et sont ensuite associés à des classifieurs pour détecter diverses lésions de l'estomac [101]. Ce classifieur est souvent de type SVM (machine à vecteurs de support) [47].

À côté des techniques précédemment citées, nous trouvons également beaucoup de références utilisant l'apprentissage profond pour traiter des images endoscopiques. L'apprentissage profond consiste en un sous-ensemble de l'apprentissage automatique qui comprend des réseaux de neurones ayant des structures complexes et variées.

Les réseaux peuvent être entraînés pour une application spécifique mais ceci requiert en général une grande base de données et un long temps de calcul. C'est pour cela que certains chercheurs se penchent sur l'apprentissage par transfert. Il s'agit d'utiliser un réseau pré-existant, par exemple dédié à la reconnaissance d'objets, de figer partiellement ou en totalité les poids des neurones qui le composent et de le spécialiser ensuite pour l'application que l'on désire [46].

À titre d'exemple, Tajbakhsh *et al.* ont comparé des réseaux pré-entraînés et des réseaux ré-entraînés complètement sur trois applications d'imagerie médicale : en radiologie, cardiologie et gastroentérologie [116]. Ils ont ainsi montré que les réseaux pré-entraînés avaient une performance supérieure à celle des réseaux ré-entraînés complètement. Certains de ces réseaux pré-entraînés ont déjà démontré leur efficacité pour la détection de lésions dans des images NBI avec grossissement. Liu *et al.* ont ainsi utilisé les réseaux VGG-16, Inception V3 et InceptionResNet v2 et les ont spécialisés leur permettant ainsi de reconnaître des cancers gastriques en formation sur des images NBI avec grossissement avec une précision avoisinant les 98% [132]. Plus récemment, Wu *et al.* ont utilisé des réseaux profonds pré-entraînés, VGG-16 et ResNet-50, et les ont spécialisés pour détecter des lésions sur des images NBI [140]. Plus précisément, la dernière couche est remplacée pour s'adapter à leur problème et toutes les couches sont entraînées sur 3170 images de cancer gastrique et 5981 images bénignes. Ils ont pu atteindre de bons résultats de classification avec une précision de 92,5%, une sensibilité de 94% et une spécificité de 91%. Ces travaux sont malheureusement orientés sur la détection du cancer gastrique. Dans notre cas, nous préférons nous intéresser aux lésions pré-cancéreuses, c'est à dire aux lésions pouvant potentiellement se transformer en cancer avec le temps. Plus la détection est précoce, meilleures sont les chances de survie pour les patients.

De rares travaux s'intéressent aux lésions pré-cancéreuses, Itoh *et al.* ont travaillé sur des images en lumière blanche pour détecter les patients atteints de gastrite chronique due à *Helicobacter pylori*. Ils ont entraîné complètement un réseau de type CNN pour classifier les images endoscopiques [129]. Leenhardt *et al.* ont également entraîné un réseau de type CNN pour segmenter des images de capsules endoscopiques dans le but de détecter l'angiectasie gastrointestinale [136].

Les travaux que nous présentons diffèrent des travaux précédemment cités par le fait que nous nous intéressons à plusieurs types de lésions pré-cancéreuses et que nous utilisons le NBI et non le NBI avec grossissement ou la lumière blanche comme dans la majorité des travaux.

3.2/ ACQUISITION DES DONNÉES ENDOSCOPIQUES

En partenariat avec l'hôpital Ambroise Paré de Boulogne-Billancourt, nous avons pu collecter des vidéos endoscopiques NBI avec le consentement des patients. L'endoscopie en mode NBI est utilisée au quotidien dans le service d'hépatogastroentérologie, elle est préférée à l'endoscopie en lumière blanche car elle permet de visualiser beaucoup plus d'éléments dans l'estomac. Un total de 57 vidéos de patients différents a donc été collecté. Des biopsies ont été prélevées sur chacun de ces patients ce qui nous permet d'associer chaque patient à son résultat histologique. Le détail de la répartition des patients selon leurs résultats histologiques se trouve dans le tableau 3.1.

TABLE 3.1 – Répartition des données vidéos selon les grades histologiques.

Normal	Gastrite active	Gastrite chronique	Atrophie	Métaplasie Dysplasie
12	9	18	7	11

Il est important de mentionner ici que le résultat histologique est issu de plusieurs prélèvements endoscopiques, le grade final étant le grade le plus avancé sur tous les prélèvements. D'autre part, les sites des prélèvements étant localisés, les résultats des biopsies peuvent ne pas être représentatifs de plusieurs lésions concomitantes dans l'estomac. Ces remarques sont à garder en mémoire

pour la bonne interprétation des résultats de classification et seront reprises par la suite.

3.3/ MÉTHODOLOGIE

3.3.1/ CLASSIFICATION PAR PATCHS CARACTÉRISTIQUES

Dans une image endoscopique, il n'y a en réalité qu'une toute petite portion de l'image sur laquelle on peut percevoir les indices d'une pathologie. De plus, nous pouvons rencontrer plusieurs pathologies dans une même image. Il nous a donc semblé plus simple, dans un premier temps, de travailler sur des patchs de taille 112x112 caractéristiques des pathologies décrites précédemment. La taille des patchs a été choisie pour être suffisamment fine pour n'observer qu'une seule pathologie et suffisamment grande pour garder le contexte visuel nécessaire à la reconnaissance. Ces patchs ont été choisis avec un gastroentérologue expérimenté qui nous a ainsi appris à reconnaître les pathologies. Par exemple, la figure 3.1 montre une image endoscopique dont on extrait un patch 112x112 caractéristique de la métaplasie intestinale.

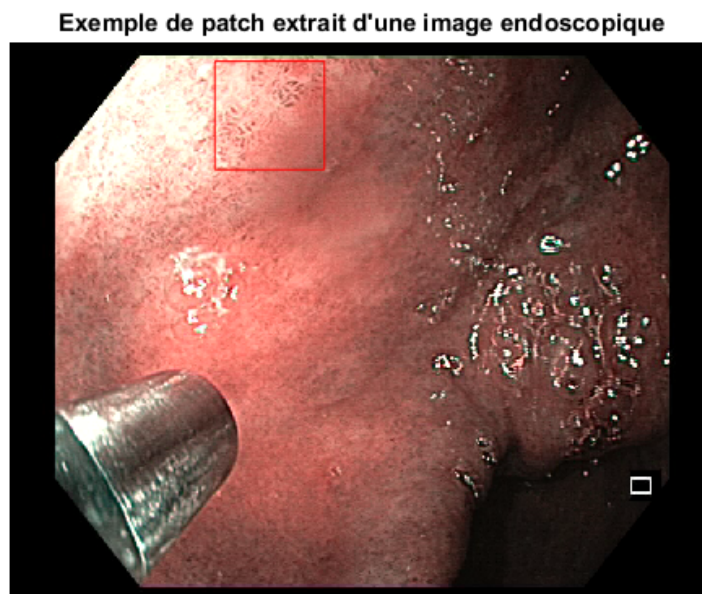


FIGURE 3.1 – Exemple de patch extrait d'une image endoscopique NBI

Sur la figure 3.2, nous donnons des exemples de patchs caractéristiques de chaque lésion.

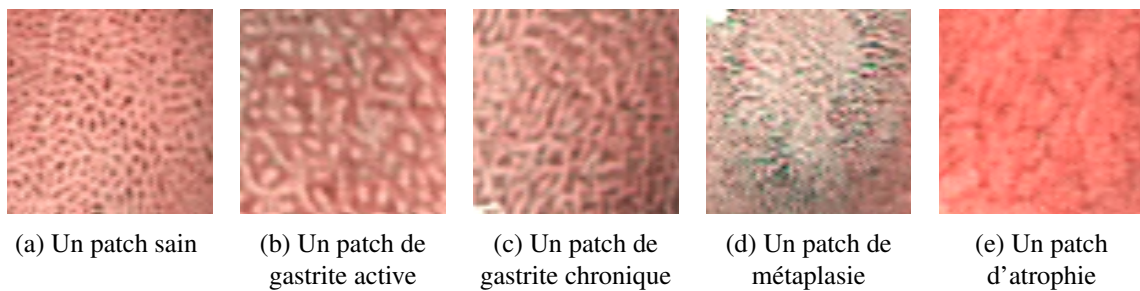


FIGURE 3.2 – Exemples de patches extraits des images NBI.

1. Une muqueuse normale se caractérise par un motif circulaire ou ovale régulier comme le montre la figure 3.2a.
2. La gastrite active peut se caractériser par la présence de points blancs à l'intérieur de la trame comme sur la figure 3.2b.
3. L'apparition de la gastrite chronique désorganise la trame comme on peut le voir sur l'image 3.2c.
4. De la même manière, la métaplasie intestinale se caractérise par une désorganisation de la trame et parfois l'apparition de mucus et de crêtes bleues [138] comme le montre la figure 3.2d.
5. Enfin, l'atrophie est reconnaissable lorsqu'on observe des petits vaisseaux sanguins qui ne sont pas visibles habituellement comme on le voit sur la figure 3.2e.

Nous présentons ici les patches les plus caractéristiques mais nous avons une forte variabilité au sein de chaque classe et les éléments cités ci-dessus ne sont pas toujours rencontrés et ne sont que des indices sur l'identification de la pathologie. Pour visualiser ce phénomène, nous présentons sur la figure 3.3 des cas plus difficiles que nous rencontrons dans la base de données. Le patch 3.3a, bien que normal, montre une trame désordonnée alors que l'on attendrait un motif régulier. Dans le patch 3.3b, nous observons très peu de points blancs caractéristiques de la gastrite active. Le patch 3.3c est très sombre pour représenter de la gastrite chronique. On peut facilement le confondre avec la métaplasie. Le patch 3.3d montre de la métaplasie intestinale mais nous observons pourtant une trame très régulière. Enfin, le dernier patch 3.3e d'atrophie ne montre que très peu de vaisseaux sanguins.

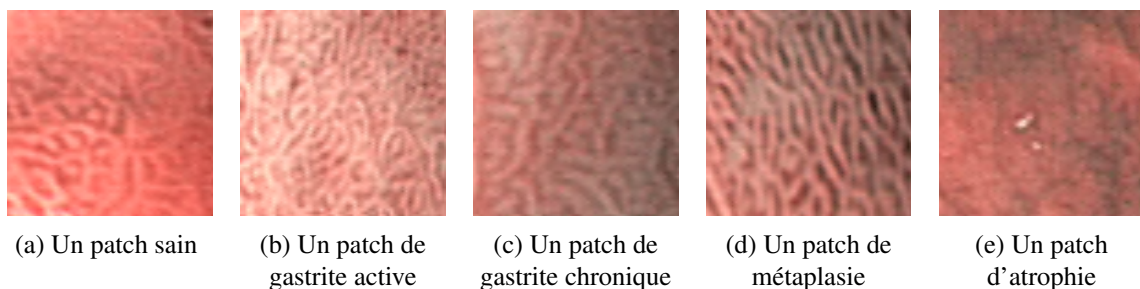


FIGURE 3.3 – Exemples de patches difficiles à évaluer.

Pour pouvoir être exploitables, un certain nombre d'opérations sont appliquées à chaque patch. Tout d'abord, les images sont dé-entrelacées. Ensuite, pour augmenter la taille de notre base d'apprentissage et dans le même temps, pour avoir le même nombre d'images par patient, nous appliquons des rotations, des réflexions (verticales ou horizontales) ou des zooms de manière aléatoire.

Augmenter la base de données permet généralement d'augmenter les performances de classification en rendant le classifieur invariant par les différentes transformations précédemment citées. Pour chaque patient, nous prenons donc 10 fois de suite un de ses patches de manière aléatoire et lui appliquons les transformations indiquées plus haut. Nous obtenons ainsi un total de 570 patches, soit 10 par patient. Les données sont équilibrées de manière à avoir le même nombre d'images par patient et non pas le même nombre d'images par pathologie. En effet, il n'est pas possible d'équilibrer en même temps les patients et les classes. De plus, nous avons observé une très forte variabilité inter-patient qui motive notre choix. Les classes sont dans tous les balancées au niveau du classifieur en appliquant des poids différents selon les classes.

À partir de ces patches, nous avons extrait différents types de descripteurs incluant des descripteurs classiques en traitement d'images et des descripteurs provenant de réseaux profonds :

Parmi les descripteurs classiques en traitement d'images, nous avons comparé :

- les motifs binaires locaux, Local Binary Patterns (LBP) en anglais [22]. Dans notre cas, nous les calculons sur les images transformées en images en niveaux de gris avec un rayon de 7 pixels et en les rendant invariant à la rotation.
- Les histogrammes de gradient orienté, Histograms of Oriented Gradients (HOG) en anglais [29]. Les HOG sont calculés avec des blocs de 2x2 cellules par blocs, 14x14 pixels par cellules et 9 orientations différentes. Dans le cas d'images couleur, le gradient le plus fort sur les trois canaux est utilisé.
- DAISY, un dérivé des descripteurs SIFT implémentés de manière à permettre une extraction dense et rapide [55]. les descripteurs sont obtenus avec 3 anneaux, un rayon de 15 pixels et des histogrammes de 8 intervalles et 8 orientations.
- Les matrices de co-occurrence, Gray Level Co-occurrence Matrix (GLCM) en anglais [12]. Six descripteurs sont calculés sur l'image transformée en image à 32 niveaux de gris : le contraste, la dissimilitude, l'homogénéité, l'énergie, la corrélation et le second moment angulaire. Ces descripteurs sont calculés sur quatre matrices de co-occurrence avec un rayon de 7 pixels et les angles 0 , $\frac{\pi}{4}$, $\frac{\pi}{2}$ et $\frac{3\pi}{4}$.
- Nous ajoutons également une combinaison de descripteurs de couleur et de texture comprenant les LBP, les moyennes et variances des trois canaux dans les espaces RGB, Lab et HSV ainsi que les histogrammes de couleurs des trois canaux RGB sur 42 intervalles.

L'ensemble des descripteurs classiques en traitement d'images ont été extraits grâce à la librairie Python Scikit-Image [90].

Les descripteurs provenant de réseaux pré-entraînés ont été extraits avec les librairies Python Pytorch et Torchvision. On peut notamment citer :

- AlexNet [85], un réseau profond qui a remporté le challenge de reconnaissance sur la base de données ImageNet en 2012 [42],
- DenseNet [123], un réseau profond densément connecté. Les résultats des couches sont reportés dans des couches plus profondes du réseau permettant de réduire les problèmes de disparition du gradient.
- SqueezeNet 1.1 [108], un réseau profond léger ayant les mêmes performances que AlexNet avec 50 fois moins de paramètres.
- VGG [89], des réseaux profonds, basé sur des blocs de couches de convolution, de MaxPooling et de couches complètement connectées pour la classification.

Nous avons ensuite entraîné un classifieur de type SVM avec noyau linéaire en utilisant une validation croisée par patient (Leave One Patient Out en anglais) : pour chaque patient, les patches issus de

ce patient sont mis de côté pour tester le classifieur et tous les autres patchs des autres patients sont utilisés pour entraîner le classifieur. Ce point est spécialement important. Les patchs d'un patient donné ont tendance à se ressembler beaucoup. Les résultats de classification seraient fortement biaisés sans la validation croisée par patient. Pourtant, un certain nombre d'articles scientifiques n'en font pas mention.

Le tableau 3.2 répertorie les résultats de classification obtenus pour chaque type de descripteur. Les chiffres présentés sont la précision, le rappel et le score F1 moyennés sur les cinq classes (Normal, gastrite active, gastrite chronique, atrophie et métaplasie intestinale). Nous remarquons que les descripteurs issus des réseaux de neurones sont légèrement supérieurs (entre 40% et 50%) par rapport aux descripteurs classiques (entre 25% et 35%).

TABLE 3.2 – Résultats de classification pour les différents descripteurs.

	Précision	Rappel	Score F1
LBP	0,41	0,44	0,40
HOG	0,35	0,28	0,31
Daisy	0,33	0,26	0,28
GLCM	0,34	0,40	0,34
Pixels vectorisés	0,28	0,29	0,26
LBP + histogramme couleur	0,51	0,44	0,45
Alexnet	0,42	0,41	0,41
Squeezenet 1,1	0,43	0,43	0,41
Densenet	0,49	0,47	0,47
VGG 11	0,52	0,49	0,48

Pour autant, les résultats ne sont pas aussi bons que ce que l'on aurait pu espérer. On peut alors s'intéresser en détail à la matrice de confusion pour mieux comprendre le phénomène. Le tableau 3.3 représente la matrice de confusion obtenue avec le réseau VGG11, celui qui semble donner les meilleurs résultats. Sur les deux premières lignes, nous observons que les patchs normaux et les patchs de gastrite active sont souvent confondus avec les patchs de gastrite chronique. Cette confusion peut venir du fait que nous avons beaucoup de patchs de gastrite chronique dans la base de données. La gastrite chronique est généralement bien reconnue. Les patchs de métaplasie sont souvent confondus avec les patchs de gastrite chronique, ce qui est attendu dans le sens où les motifs se ressemblent beaucoup : une trame épaisse et désorganisée. D'autant plus que la métaplasie se présente sous la forme de plaques au milieu de la gastrite chronique, on trouvera donc nécessairement des patchs de gastrite chronique chez des patients ayant de la métaplasie intestinale, cette confusion est acceptable d'un point de vue clinique. Enfin les patchs d'atrophie sont bien reconnus en général. En effet, le motif donné par l'atrophie est différent de tous les autres, il se caractérise par la disparition de la trame et l'apparition de petits vaisseaux sanguins.

Le tableau 3.4 donne la précision, le rappel et le score F1 pour les cinq classes. Ce tableau nous confirme que l'atrophie a une bonne précision de 72% et un bon rappel de 73%. C'est la classe qui est la mieux reconnue par le classifieur. La gastrite a également un bon rappel de 75%. Les résultats globaux sont donc à modérer. D'un côté, la précision moyenne n'est que de 52% mais il s'agit d'une classification à 5 classes, une classification aléatoire aurait des scores avoisinant les 20%. De plus, certaines confusions sont acceptables d'un point de vue clinique comme la confusion gastrite chronique/métaplasie intestinale et l'atrophie est bien reconnue.

TABLE 3.3 – Matrice de confusion résultant de la classification avec le réseau VGG11.

		Prédiction				
		Normal	Gastrite active	Gastrite chronique	Métablasie intestinale	Atrophie
Label histologique	Normal	16	14	72	11	7
	Gastrite active	8	34	27	9	12
	Gastrite chronique	10	20	135	15	0
	Métablasie intestinale	7	8	42	52	1
	Atrophie	0	8	10	1	51

TABLE 3.4 – Détail des résultats de classification avec le réseau VGG11.

	Précision	Rappel	Score F1
Normal	0,39	0,13	0,20
Gastrite active	0,40	0,38	0,39
Gastrite chronique	0,47	0,75	0,58
Métablasie intestinale	0,59	0,47	0,53
Atrophie	0,72	0,73	0,72
Moyenne des cinq classes	0,52	0,49	0,48

3.3.2/ APPLICATION DU MODÈLE DICHROMATIQUE

L'application du modèle dichromatique, développé dans le chapitre 2 permet de séparer les composantes intrinsèques d'une image. Elle permet notamment de s'affranchir des variations d'illumination et de la géométrie de la surface de l'estomac. Il peut donc améliorer les résultats de classification en enlevant les confusions dues à ces deux phénomènes. Nous avons donc appliqué la décomposition par ce modèle aux patches sélectionnés précédemment pour vérifier si la décomposition en images intrinsèques aidait la classification. Plus précisément, nous extrayons les trois composantes g , S et k , respectivement le shading, la réflectance de l'estomac et la composante spéculaire.

À titre d'exemple, nous présentons sur la figure 3.4 les résultats de la décomposition sur un patch de métablasie. Nous remarquons que la trame se trouve effacée dans l'image S . La trame se retrouve davantage dans l'image de shading g , elle correspond en effet au relief présent dans l'estomac. La couleur de la trame et la couleur du fond sont sur le même plan dichromatique, c'est à dire sur un plan tel que les valeurs de deux pixels (un de la trame et un du fond) soient égales à un coefficient multiplicatif et un coefficient additif près.

La composante S doit contenir toutes les informations de couleur nécessaires à la classification des patches, la composante shading g contient les informations de forme et de texture. La composante k ne donne a priori aucune information utile à la classification comme il s'agit uniquement de la réflexion spéculaire, c'est-à-dire les tâches blanches dues au spéculaire. Nous avons donc testé trois configurations différentes, une en ne retenant que g , une en ne retenant que S et une en

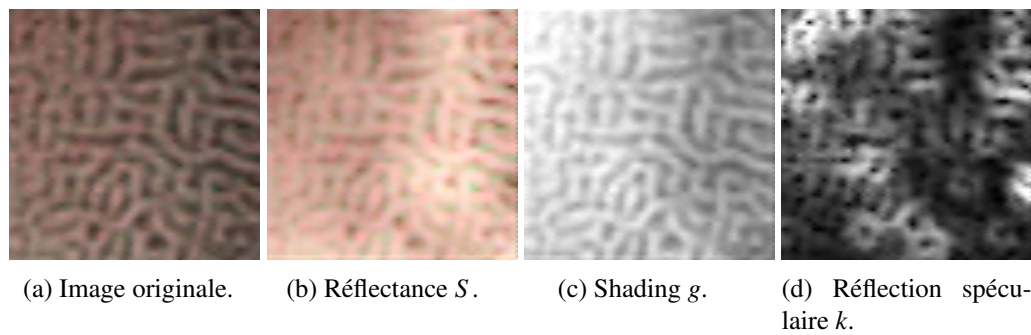


FIGURE 3.4 – Exemples de patch décomposé selon le modèle dichromatique.

retenant gS . Dans les trois cas, nous calculons les mêmes descripteurs que dans la partie précédente sur les composantes g , S et gS de chaque patch.

Le tableau 3.5 contient les résultats de classification, pour tous les descripteurs entraînés sur les composantes g , S et gS . Les résultats obtenus sont, de manière générale, légèrement plus bas que dans le tableau 3.2 à l'exception des descripteurs "Pixels vectorisés", "HOG" et "Daisy". Nous pouvons donc conclure que l'application du modèle dichromatique n'aide malheureusement pas la classification. Ceci peut être dû au fait que l'estomac a quasiment toujours une couleur rosée, l'espace des couleurs que l'on trouve dans les images NBI est relativement réduit et l'application du modèle dichromatique demanderait un certain pré-traitement pour augmenter les variations des couleurs visibles. On peut noter que l'on se trouve dans un cas limite dans lequel nous avons des images à trois canaux distincts mais dans lesquels les canaux vert et bleu sont très fortement corrélés. Or, dans un cas où les couleurs sont précisément toutes situées sur un plan dichromatique comme c'est le cas avec les images NBI, le modèle considère que tous les pixels appartiennent au même matériau.

TABLE 3.5 – Résultats de classification pour les différents descripteurs après décomposition par le modèle dichromatique.

	Avec g			Avec S			Avec gS		
	Précision	Rappel	Score F1	Précision	Rappel	Score F1	Précision	Rappel	Score F1
LBP	0,39	0,40	0,38	0,37	0,30	0,32	0,32	0,30	0,29
HOG	0,35	0,28	0,30	0,36	0,30	0,33	0,36	0,30	0,32
Daisy	0,35	0,25	0,28	0,36	0,29	0,31	0,36	0,30	0,32
GLCM	0,35	0,32	0,31	0,34	0,27	0,29	0,37	0,30	0,32
Pixels vectorisés	0,36	0,30	0,32	0,37	0,30	0,33	0,35	0,28	0,31
LBP + histogramme couleur	0,41	0,32	0,35	0,37	0,35	0,35	0,36	0,31	0,33
Alexnet	0,40	0,37	0,37	0,37	0,36	0,35	0,41	0,36	0,35
Squeezenet 1,1	0,38	0,33	0,35	0,40	0,40	0,38	0,36	0,33	0,33
Densenet	0,48	0,42	0,42	0,53	0,46	0,46	0,38	0,38	0,36
VGG 11	0,40	0,39	0,38	0,47	0,42	0,41	0,34	0,35	0,33

3.3.3/ ASSOCIATION DES RÉSULTATS À LA MOSAÏQUE

En général, les endoscopistes ne re-regardent pas les vidéos qu'ils ont acquises. Dans le cadre du projet EMMIE, l'équipe du laboratoire CRAN a développé un algorithme permettant de créer une mosaïque à partir d'une vidéo endoscopique. Il s'agit de créer une image qui est un condensé de la vidéo. La mosaïque sert alors à revoir des détails de la vidéo et de pouvoir échanger avec d'autres endoscopistes sur des cas litigieux.

Pour créer la mosaïque, les images de la vidéo doivent être recalées entre elles. La complexité de ce travail réside dans le peu de texture que l'on trouve dans les images endoscopiques ainsi que les forts changements d'illuminations. La méthode repose sur l'estimation du flux optique à partir de descripteurs invariants à l'illumination. À partir de ce flux optique, les images sont combinées pour former une mosaïque représentant la paroi de l'estomac [133, 139]. Ces travaux nous permettent en plus d'apposer les résultats de classification sur la mosaïque et de présenter la combinaison au gastroentérologue. Le spécialiste a alors une vue globale de la paroi de l'estomac.

Comme la classe la plus facilement prédictible est l'atrophie, nous avons choisi de montrer un exemple de la détection de l'atrophie sur une mosaïque. Pour cette démonstration, nous avons besoin d'une nouvelle classe que nous nommons "classe aléatoire". Il s'agit d'un ensemble de patches 112x112 pixels pris aléatoirement dans toutes les vidéos endoscopiques disponibles. Cette classe est essentielle dans le sens où les patches précédents sont caractéristiques des pathologies mais sont très minoritaires par rapport à tout ce qui est visible lors d'une endoscopie. Cette "classe aléatoire" contient plus de 21000 patches et sert à modéliser l'ensemble des patches qu'il est possible d'obtenir en endoscopie. Nous avons ainsi ré-entraîné un classifieur dont le but est de pouvoir discriminer la "classe aléatoire" de la classe "atrophie". Nous avons obtenu de très bons résultats avec ce classifieur avec une précision de 96%, un rappel de 99% et un score F1 de 97%.

Pour appliquer le classifieur à notre mosaïque, nous parcourons toute la mosaïque avec une fenêtre 112x112 (puisque le classifieur a été entraîné avec des patches de cette taille) en déplaçant la fenêtre par pas de 56 pixels (nous avons donc un recouvrement de 75%, 50% en largeur et 50% en longueur). Nous mettons en relief les zones classées positives à l'atrophie en les entourant en bleu comme illustré dans la figure 3.5.

Les trois zones bleues trouvées sont bien cohérentes avec ce que nous attendions. La zone en haut à gauche, bien que chevauchée par la fibre montre bien de l'atrophie et il en est de même pour la zone d'en bas au milieu. Seule la zone à droite n'est pas représentative de l'atrophie.

3.4/ DÉTERMINATION DE CRITÈRES VISUELS

Dans la mesure où les descripteurs usuels et les descripteurs basés sur l'apprentissage profond donnent une classification satisfaisante uniquement pour l'atrophie, nous tentons de comprendre d'où vient la confusion. L'annotation de patches n'est pas naturelle du point de vue des endoscopistes. D'une part parce que les patches ne sont qu'une petite partie des images endoscopiques et d'autre part parce que l'endoscopiste voit une vidéo et n'a donc pas le temps de repérer les indices des inflammations. Pour cela nous avons développé avec le gastroentérologue un ensemble de cinq nouveaux critères visuels sur lesquels se base l'endoscopiste pour établir un diagnostic. Ces cinq critères sont respectivement la couleur, la régularité de la trame, l'épaisseur de la trame, la texture à l'intérieur de la trame et la densité vasculaire. Pour chacun de ces critères, une note entre 0 et 3 est attribuée à chaque patch de manière aveugle, c'est-à-dire sans avoir connaissance du diagnostic patient. Pour des raisons de contraintes de temps, nous avons proposé à l'endoscopiste

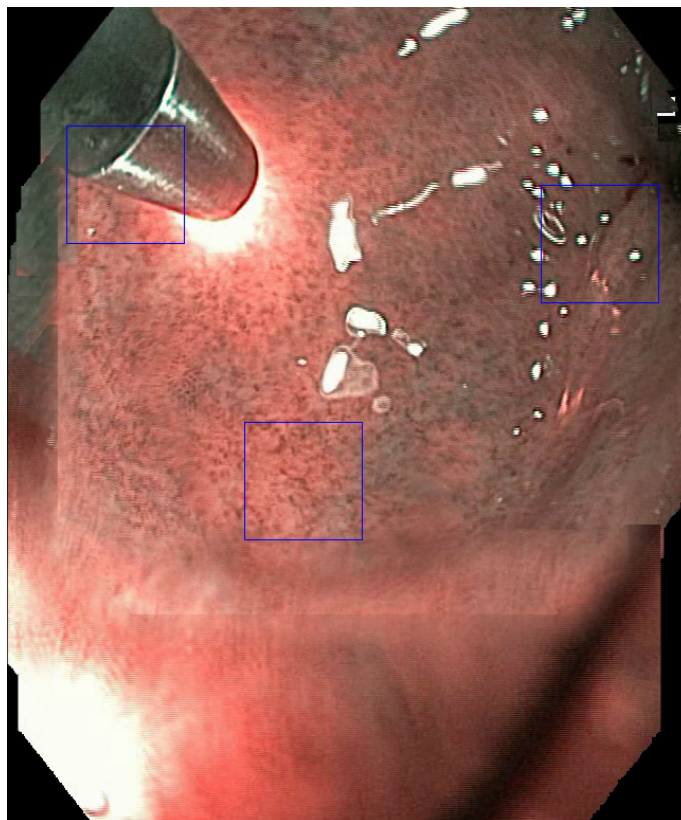
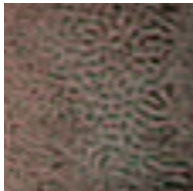
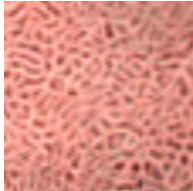
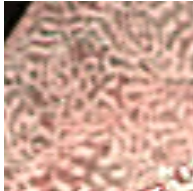
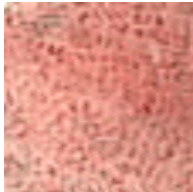
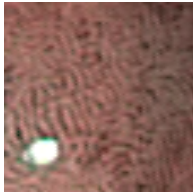
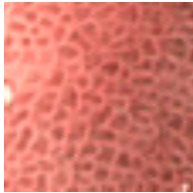
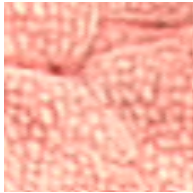
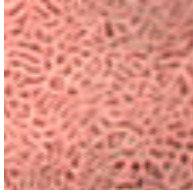


FIGURE 3.5 – Juxtaposition des résultats de classification sur la mosaïque fournie par le laboratoire CRAN.

d'annoter un sous-ensemble des patchs utilisés dans la section précédente constitué de 177 patchs significatifs des pathologies. La définition de ces critères visuels se trouvent dans le tableau 3.6.

TABLE 3.6 – Définition des critères visuels et exemples.

1. La couleur du fond • 0 - rose pale • 1 - rose foncé • 2 - violet • 3 - noir		
2. La régularité de la trame • 0 - parallélépipède - carré • 1 - parallélépipède - rectangle • 2 - disparition des angles • 3 - irrégulière • (-1) - absence de trame		
3. L'épaisseur de la trame • 0 - fine • 1 - peu épaisse • 2 - moyennement épaisse • 3 - très épaisse • (-1) - absence de trame		
4. Texture à l'intérieur de la trame • 0 - aucune • 1 - points • 2 - traits fins • 3 - traits épais		
5. Densité vasculaire • 0 - nulle • 1 - faible • 2 - moyenne • 3 - forte		

Il est alors intéressant de voir si ces critères visuels sont bien corrélés aux résultats histologiques. Pour cela, nous effectuons pour chaque critère et pour chaque pathologie le test de la somme des rangs [10]. Il s'agit d'un test statistique aussi appelé test de Wilcoxon ou test U de Mann-Whitney. Il est fait pour déterminer la probabilité que les médianes de deux ensembles d'échantillons sont les mêmes. Plus cette probabilité, appelée valeur-p est basse, plus le critère est efficace au regard d'une pathologie donnée.

TABLE 3.7 – Résultats des tests de la somme des rangs.

	Normal	Gastrite active	Gastrite chronique	Métaplasie intestinale	Atrophie
La couleur du fond	$4,64 \times 10^{-5}$	$9,96 \times 10^{-1}$	$3,77 \times 10^{-3}$	$1,72 \times 10^{-6}$	$2,37 \times 10^{-3}$
La régularité de la trame	$3,59 \times 10^{-3}$	$6,00 \times 10^{-2}$	$2,01 \times 10^{-3}$	$5,64 \times 10^{-18}$	$6,47 \times 10^{-22}$
L'épaisseur de la trame	$6,08 \times 10^{-2}$	$7,56 \times 10^{-1}$	$1,60 \times 10^{-6}$	$3,22 \times 10^{-13}$	$1,95 \times 10^{-21}$
Texture à l'intérieur de la trame	$7,00 \times 10^{-3}$	$2,11 \times 10^{-25}$	$1,24 \times 10^{-1}$	$4,67 \times 10^{-1}$	$7,77 \times 10^{-3}$
Densité vasculaire	$1,62 \times 10^{-5}$	$1,57 \times 10^{-2}$	$2,68 \times 10^{-4}$	$4,04 \times 10^{-3}$	$3,03 \times 10^{-38}$

Dans le tableau 3.7, nous avons consigné les valeurs-p obtenues avec le test de la somme des rangs pour tous les critères visuels au regard d'une pathologie donnée. Nous remarquons en premier que l'atrophie est facilement identifiable grâce à l'absence de trame. En effet nous obtenons une valeur-p de l'ordre de $\approx 10^{-22}$ pour la régularité de la trame et de $\approx 10^{-21}$ pour l'épaisseur de la trame. Dans ces deux cas, l'absence de trame est indiquée par la note -1. Nous remarquons également qu'une forte densité de vaisseau (valeur-p de $\approx 10^{-38}$) est significative de la présence de l'atrophie. Dans ce tableau, nous pouvons également lire que la gastrite active se caractérise par la présence de texture à l'intérieur de la trame, notamment des points blancs (valeur-p de $\approx 10^{-25}$). Enfin la métaplasie intestinale se caractérise par une trame irrégulière (valeur-p de $\approx 10^{-18}$) et épaisse (valeur-p de $\approx 10^{-13}$).

Certains critères sont donc plus importants que d'autres. Par exemple la couleur du fond est relativement peu significative. Ceci peut s'expliquer par le fait qu'une couleur est difficile à juger, elle dépend de l'illumination de l'écran. Les autres critères sont par contre intéressants pour au moins une pathologie.

À partir des notes données par l'endoscopiste, il est possible de modéliser la réflexion qu'il a eu. Pour cela, nous avons entraîné un classifieur de type arbre de décision avec les annotations du gastroendoscopiste comme données d'entraînement et essayons de prédire le label histologique. L'arbre de décision résultant est visible sur la figure 3.6.

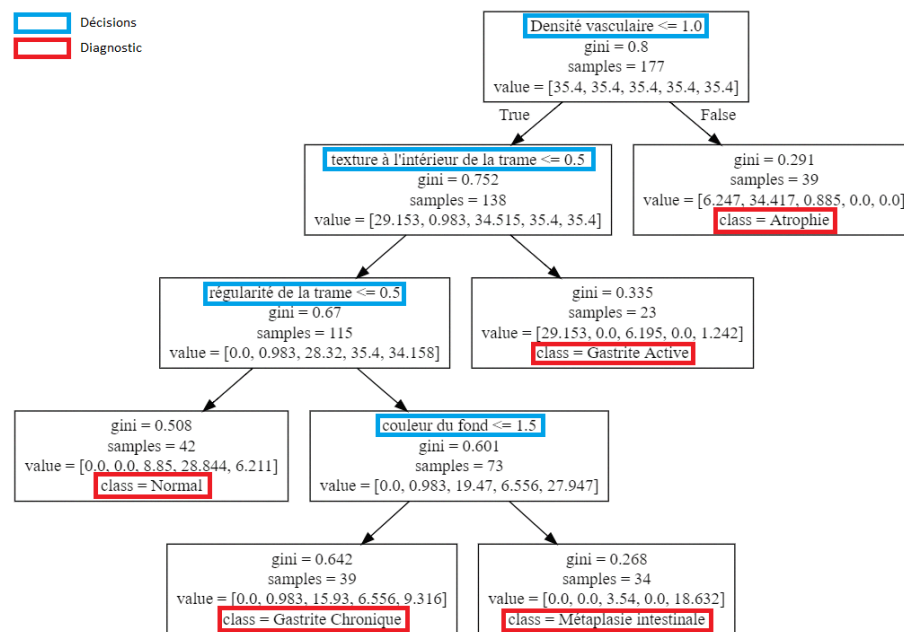


FIGURE 3.6 – Arbre de décision généré à partir des choix du gastroendoscopiste.

Pour lire ce schéma, nous partons du nœud le plus en haut. À chaque nœud de l'arbre, une condition est testée (en bleu sur le schéma), si cette condition est vraie, nous passons au nœud fils de gauche, sinon au nœud fils de droite. Ce principe est répété jusqu'à tomber sur un nœud terminal. On appelle ces nœuds terminaux des feuilles de l'arbre. Arrivé à une feuille, nous pouvons poser un diagnostic (en rouge sur le schéma). Si l'on devait réécrire le processus de pensée de l'endoscopiste. On aurait le raisonnement suivant :

- La première étape consiste à noter la densité vasculaire.
 - Si celle-ci est supérieure à 1, alors le diagnostic est l'atrophie.
- Dans le cas contraire on évalue l'intérieur de la trame.
 - Si celle-ci est supérieure à 0, alors le diagnostic est la gastrite active.
- Dans le cas contraire on évalue la régularité de la trame.
 - Si celle-ci est inférieure à 1, alors le diagnostic est "normal".
- Dans le cas contraire on évalue la couleur du fond.
 - Si celle-ci est inférieure à 2, alors le diagnostic est la gastrite chronique.
 - Sinon, le diagnostic est la métaplasie intestinale.

Avec cet arbre de décision, nous arriverions à prédire la classe histologique avec une précision de 67%, un rappel de 71% et un score F1 de 67%. Comme précédemment, l'atrophie est la classe la mieux prédite (90% de précision, 97% de rappel et un score F1 de 93%). Ces chiffres dépassent les résultats obtenus avec les descripteurs provenant de l'apprentissage par transfert. Les critères visuels sont donc cohérents avec l'histologie et sont donc très prometteurs si les outils d'apprentissage supervisé ou d'apprentissage profond arrivent à reproduire les annotations des experts endoscopistes.

Dans le tableau 3.9, nous détaillons également la matrice de confusion obtenue avec cet arbre de décision. Nous pouvons ainsi confirmer que l'atrophie est bien prédite (35 cas sur 36). La gastrite active est également bien prédite (13 cas sur 17). Nous observons aussi certaines confusions,

TABLE 3.8 – Résultats de classification avec les critères visuels.

	Précision	Rappel	Score F1
Normal	0,52	0,81	0,64
Gastrite active	0,59	0,76	0,67
Gastrite chronique	0,45	0,45	0,45
Métaplasie intestinale	0,88	0,53	0,66
Atrophie	0,90	0,97	0,93
Moyenne des cinq classes	0,67	0,71	0,67

en particulier pour la métaplasie intestinale, confondue avec la gastrite chronique (15 patches de métaplasie intestinale se retrouvent classés comme gastrite chronique) et le normal (10 patches de métaplasie classés comme normal). Comme nous l'avons déjà expliqué précédemment, la confusion entre gastrite chronique et métaplasie est acceptable d'un point de vue médical dans le sens où la métaplasie se trouve en pratique sous forme de tâches sur la gastrite chronique. Nous observons également une confusion entre la gastrite chronique et la classe "Normal", 10 patches de gastrite chronique sont prédits comme étant normaux par le classifieur.

TABLE 3.9 – Résultats de classification pour les différents descripteurs.

		Prédiction				
		Normal	Gastrite active	Gastrite chronique	Métaplasie intestinale	Atrophie
Label histologique	Normal	22	0	5	0	0
	Gastrite active	0	13	1	0	3
	Gastrite chronique	10	7	18	4	1
	Métaplasie intestinale	10	2	15	30	0
	Atrophie	0	0	1	0	35

Ce deuxième test nous a permis de confirmer les premiers résultats obtenus, à savoir, que l'on peut distinguer facilement l'atrophie mais que l'on a malheureusement plus de difficulté à distinguer les classes "normal", gastrite chronique et métaplasie autant d'un point de vue du traitement d'images qu'avec l'œil d'un spécialiste.

3.5/ CONCLUSION ET PERSPECTIVES

Dans ce chapitre, nous avons mis en place une méthode pour analyser les images endoscopiques NBI acquises à l'hôpital Ambroise Paré. Nous procédons en deux étapes : dans un premier temps, nous extrayons avec un gastroentérologue expérimenté des patches caractéristiques d'une muqueuse normale et de quatre pathologies différentes. Dans un second temps, un ensemble de descripteurs est utilisé pour entraîner et tester un classifieur. Nous faisons notamment la distinction entre les

descripteurs provenant de réseaux de neurones pré-entraînés et des descripteurs classiques en traitement d'images. La classification se trouve être difficile sur la base de données de patches créée. Les meilleurs scores sont obtenus avec le réseau VGG11 (51% de précision, 49% de rappel et un score F1 de 48%). Les résultats moyens peuvent peut-être s'expliquer par le fait que nous n'avons pas suffisamment de patches d'entraînement par rapport au nombre de descripteurs, provoquant ainsi un phénomène de sur-apprentissage. De plus, il y a une forte variabilité inter-patient qui rend la tâche plus difficile. Certaines pathologies sont pourtant plus faciles à reconnaître que d'autres, notamment l'atrophie grâce à l'absence de trame et la présence de petits vaisseaux sanguins.

Dans ce chapitre nous avons également mis en place avec un gastroentérologue des critères visuels sur lesquels il se base pour poser un premier diagnostic (le diagnostic final étant donné par l'histologie). Cinq critères ont ainsi été élaborés et visent à caractériser un patch d'image endoscopique. Le gastroentérologue a annoté à l'aveugle ces patches selon ces cinq critères, respectivement la couleur du fond, la régularité de la trame, l'épaisseur de la trame, la texture à l'intérieur de la trame et la densité vasculaire. En utilisant ces critères comme descripteurs, nous sommes capables d'élaborer un arbre de décision correspondant au cheminement de pensée que peut avoir un endoscopiste lorsqu'il analyse les images. Cet arbre de décision donne de meilleurs résultats comparé à ceux obtenus via l'apprentissage par transfert (67% de précision, 71% de rappel et un score F1 de 67%). Cela prouve que l'information est bien contenue dans l'image et que la classification est possible.

Des améliorations pourraient être obtenus en évaluant un nombre de patches plus significatif et d'analyser en profondeur les erreurs commises par les classifieurs avec un expert endoscopiste. Il serait également très intéressant d'implémenter, via des méthodes d'apprentissage automatique, l'évaluation des critères visuels donnée par l'endoscopiste pour pouvoir reproduire son jugement.

II

TRAITEMENT ET ANALYSE DES SPECTRES CHEZ LA SOURIS ET CHEZ L'HOMME

ÉTUDE DE LA RÉFLECTANCE DE L'ESTOMAC CHEZ LA SOURIS

P our aller plus loin dans la caractérisation des lésions de l'estomac, nous souhaitons améliorer la résolution spectrale des acquisitions par rapport au NBI. Avant de faire des acquisitions spectrales chez l'Homme, nous avons mené une campagne d'acquisitions de spectres sur des souris infectées par *Helicobacter pylori*, une bactérie impliquée dans le développement des lésions de l'estomac aussi bien chez l'Homme que chez la souris. L'analyse de ces spectres nous permet de définir un pipeline de classification et d'isoler des longueurs d'onde intéressantes pour la détection de l'inflammation. Les méthodes d'analyse et les résultats pourront être réutilisés pour la partie clinique avec les autres travaux concernant la caractérisation spectrale des lésions de l'estomac chez l'Homme.

Sommaire

4.1	<i>Helicobacter pylori</i>	54
4.2	Protocole d'acquisition	55
4.3	Pipeline de classification des spectres	58
4.4	Résultats de classification	61
4.5	Conclusion et perspectives	63

4.1/ *Helicobacter pylori*

Une des causes majeures du développement du cancer de l'estomac qui a été identifiée est l'infection à *Helicobacter pylori* (*H. pylori*). Cette bactérie est de forme hélicoïdale ce qui lui confère une motilité importante et lui permet de traverser la lumière (*lumen*) de l'estomac et d'atteindre la muqueuse gastrique. Cette bactérie est capable de survivre dans l'environnement gastrique très acide. Elle est actuellement la seule bactérie déclarée comme carcinogène de classe I par l'Organisation Mondiale de la Santé (OMS) [69], [50]. *H. pylori* colonise l'estomac humain et affecte les cellules de la muqueuse gastrique dont les cellules souches [103]. *H. pylori* est associée à de nombreux types de lésions inflammatoires de l'estomac. Elle induit une gastrite chronique qui peut évoluer en ulcère, métaplasie intestinale, dysplasie voire cancer [16], [17], [119]. Environ 50% de la population est infectée par cette bactérie, 10% d'entre eux développent un ulcère gastrique ou duodénal et entre 1 et 3% développent un cancer gastrique [127].

Il est actuellement reconnu que *H. pylori* est le facteur de risque majeur du cancer gastrique. On sait également que la bactérie est, entre autres, associée à l'anémie chez l'enfant [124], à des maladies cardio-vasculaires, au diabète et à des maladies neuro-dégénératives [94]. Plus récemment, Hernando-Harder *et al.* ont mis en évidence un lien entre l'infection à *H. pylori* et certaines pathologies cutanées telle l'urticaire chronique et le Purpura thrombotique thrombocytopénique [44]. Cette bactérie n'affecte donc pas que l'estomac et elle a un fort impact dans la recherche dans le domaine de la santé.

La bactérie provoque une inflammation chronique persistante et échappe au système immunitaire de l'hôte. C'est ce phénomène qui est à l'origine de l'initiation du processus de cancérogenèse. Cette inflammation chronique est à l'origine de la production d'espèces réactives de l'oxygène capable d'induire des dommages à l'ADN et être une source de mutagenèse. Chez la souris, il a été prouvé que la bactérie induit également des mutations génétiques associées à l'inflammation de l'estomac [28]. D'autre part, *H. pylori* déclenche plusieurs réactions immunitaires et altère le microbiote gastrique et intestinal [109]. L'étude des effets de la bactérie sur un modèle animal et en particulier la souris est donc intéressante pour pouvoir orienter les futures recherches sur l'Homme. En effet, les différentes études menées sur le modèle murin ont montré un parallèle avec les lésions gastriques induites par cette infection en pathologie humaine.

Des systèmes multispectraux ou hyperspectraux ont déjà été utilisés pour la détection de divers cancers chez la souris. Par exemple, Martin *et al.* ont ainsi développé un système basé sur des filtres LCTF (Liquid Crystal Tunable Filter) couplés à un endoscope pour caractériser le cancer de la peau chez la souris. Ils utilisent le principe d'autofluorescence : lorsqu'un tissu est excité avec une lumière d'une certaine longueur d'onde, certaines protéines, acides aminés, enzymes et coenzymes émettent en retour à une longueur d'onde spécifique. Cette longueur d'onde est différente et est toujours plus élevée que la longueur d'onde d'excitation. En l'occurrence, Martin *et al.* ont testé des longueurs d'onde d'excitation entre 420 nm and 480 nm et ont observé les réponses entre 530 nm et 580 nm. Ils ont alors calculé des ratios et des différences avec une réponse saine pour distinguer les tissus malades [32]. Une autre étude a également montré que l'utilisation de caméras hyperspectrales entre 450 et 950 nm permettait de discriminer les tissus cancéreux et non-cancéreux dans les ganglions lymphatiques et les poumons [64]. La discrimination était faite par un classifieur de type SVM dont la sensibilité et la spécificité ont été estimés à 97,7% et 92,6% pour les poumons et 98,3% et 96,2% pour les ganglions lymphatiques.

En collaboration avec l'équipe "Pathogenèse de *Helicobacter*" de l'institut Pasteur associée au projet EMMIE, nous avons étudié les propriétés des spectres de réflectance acquis sur des estomacs de souris. Les souris ont été divisées en deux groupes : les souris infectées au *H. pylori* et

les souris témoins. Le but étant de déterminer des longueurs d'onde utiles pour la détection de l'inflammation et de créer un pipeline de classification pour séparer les spectres dit "normaux" des spectres d'estomacs inflammés.

4.2/ PROTOCOLE D'ACQUISITION

Les travaux sur les souris ont été réalisés dans le respect et les recommandations du "Guide spécifique pour le bien-être et l'utilisation des animaux de laboratoire à l'institut Pasteur" en accord avec les directives européennes (2010/63/UE). Le projet a été approuvé par le comité central des conseils d'installation de l'institut Pasteur et par le comité d'éthique en expérimentation Animale (CETEA) de l'institut Pasteur (Ref 2013-0051).

Les souris de cette étude sont des mâles C57BL/6 âgés de 5 à 6 semaines (Charle River, France). 72 animaux ont été divisés en deux groupes, la moitié étant infectée par voie orale avec 100 μ l d'une suspension de la souche de *H. pylori* SS1 (10^8 colonies par unité cfu/ml). La souche SS1 est capable de coloniser l'estomac murin pendant plusieurs mois, provoquant une gastrite active [23]. Les réponses immunitaires locales et systémiques sont similaires à celles décrites chez l'Homme [24]. L'autre moitié des souris, servant de groupe contrôle, ont reçu 100 μ l d'eau peptonée comme décrit dans l'article de Touati *et al.* [28].

Les souris ont été sacrifiées à des moments prédéfinis, soit 1, 3, 6, 12 ou 18 mois après infection. Leur estomac est alors isolé et plusieurs acquisitions spectrales sont faites immédiatement après grâce à un spectromètre. En parallèle, un fragment d'estomac est prélevé afin de quantifier le niveau de colonisation de la bactérie comme illustré par les figures 4.1 et 4.2. Le sang est aussi collecté pour une analyse sérologique de l'infection.

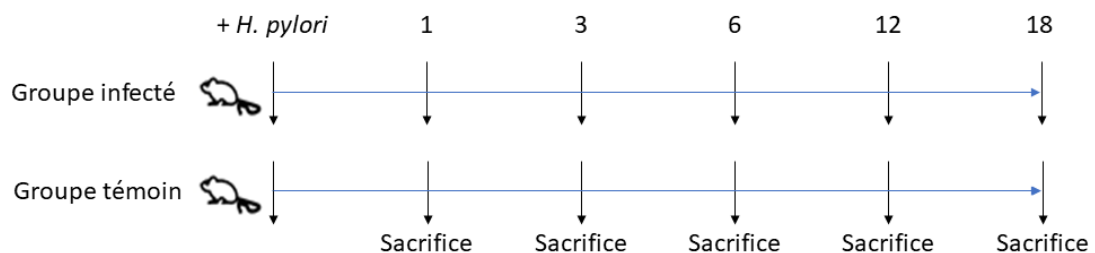


FIGURE 4.1 – Sacrifice des souris à 1, 3, 6, 12 et 18 mois après infection.

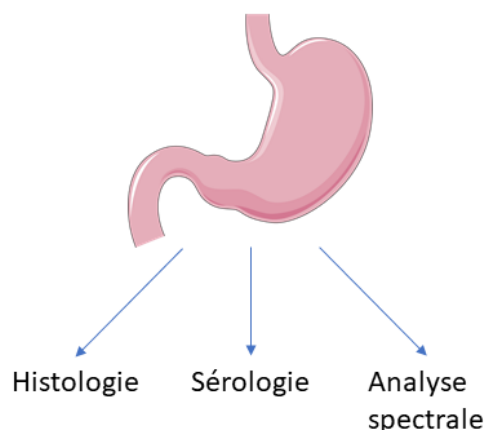


FIGURE 4.2 – L'estomac de souris est prélevé pour trois types de mesures : des mesures histologiques, sérologiques et spectrales.

Concernant l'analyse histologique, les fragments d'estomacs sont fixés en RCL2® (Alphelys, France) et inclus en paraffine (Poly Ethylene Glycol Distearate ; Sigma, USA). Des sections de 4 μm d'épaisseur sont colorées à l'hématoxyline et à l'éosine (H et E) et examinées à l'aveugle par un histopathologiste [28] afin d'identifier histologiquement les lésions inflammatoires induites et d'en évaluer leur intensité. Un exemple de cette coloration est donné par la figure 4.3. Les altérations histologiques de l'estomac (infiltration des cellules inflammatoires, métaplasie intestinale, dysplasie et perte de cellules pariétales) sont évaluées de manière semi-quantitative selon les méthodes décrites par Eaton *et al.* et Rogers *et al.* [20] [30]. Un score global est alors attribué à chaque souris :

1. - Absence de lésions
2. $\pm$ lésion minime
3. + lésion moyenne
4. ++ lésion modérée
5. +++ lésion marquée
6. ++++ lésion sévère

Un tableau récapitulatif des grades histologiques obtenus chez les souris se trouve figure 4.4.

Le sang est lui collecté pour déterminer la sérologie par la méthode immuno-enzymatique ELISA. La présence d'immunoglobuline G (IgG) spécifiques de *H. pylori* dans le sérum des souris infectées est analysée [24]. Cela nous permet de vérifier que les souris ont bien été infectées comme le montre la figure 4.5. La colonisation gastrique s'exprime en nombre de colonies de bactéries par gramme de tissu (colonies forming unit (cfu)/g) [24].

Pour les acquisitions spectrales, l'estomac est ouvert selon la grande courbure et placé entre lamelles en verre. L'ensemble est placé dans une boîte noire empêchant la lumière extérieure de perturber l'acquisition. Les spectres sont acquis sur la partie glandulaire de l'estomac grâce à un spectromètre Avantes. Les spectres sont obtenus dans l'intervalle 200 nm-1160 nm (à pas non constant). Une sphère intégrante, combinée à une source de lumière halogène (AvaSphere-50-LS-HAL-12V) nous permet d'avoir une lumière diffuse sur l'échantillon. 4 à 5 spectres sont acquis à différents endroits de l'estomac pour chaque souris, un spectre étant une moyenne sur un disque de 1 cm de diamètre. Le dispositif d'acquisition des spectres est schématisé par la figure 4.6.

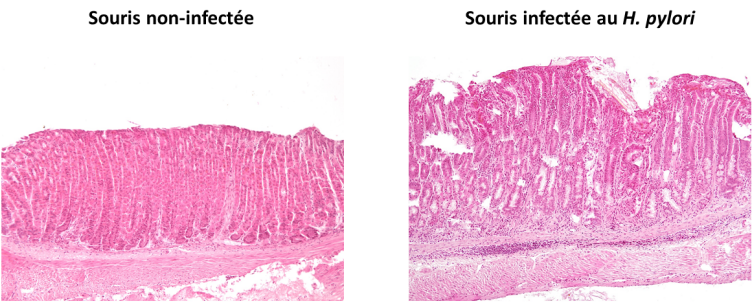


FIGURE 4.3 – Analyse des lésions inflammatoires induites par *H. pylori* dans la muqueuse des souris. Coloration H et E des lésions histologiques de deux souris, 12 mois après infection. La souris de gauche n’est pas infectée alors que la souris de droite l’est. On observe l’allongement des glandes et la présence d’agrégats de lymphocytes dans la sous-muqueuse (échelle : 50 μ m).

Degré d'inflammation		
Non-infectées		infectées
0	++++	1
0	+++	1
1	++	7
2	+	4
1	±	1
14	-	2

FIGURE 4.4 – Nombre de souris infectées et non-infectées à 12 et 18 mois selon leur grade histologique déterminé par les méthodes de Rogers *et al.* (2005) et Eaton *et al.* (1995).

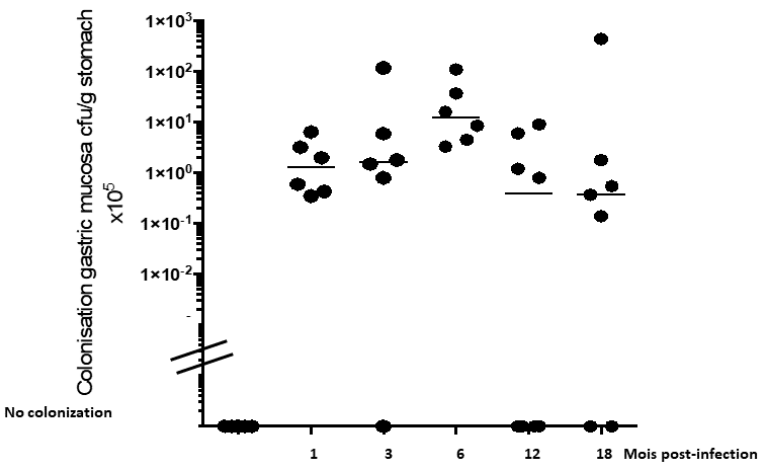


FIGURE 4.5 – Colonisation gastrique de *H. pylori* exprimée en cfu/g de tissu.

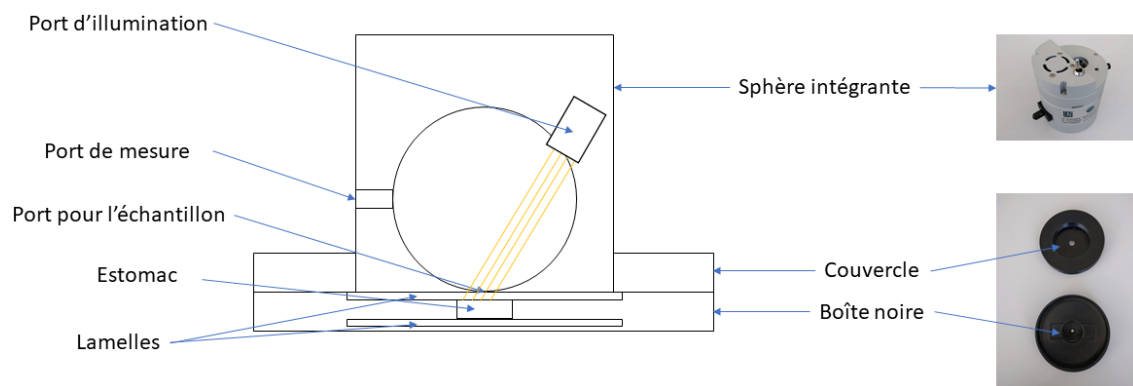


FIGURE 4.6 – Dispositif d'acquisition des spectres.

4.3/ PIPELINE DE CLASSIFICATION DES SPECTRES

Les spectres obtenus sont bruités pour les longueurs d'onde en dessous de 400 nm et au-dessus de 950 nm, nous ne gardons donc que l'information à l'intérieur de cet intervalle. La figure 4.7 représente la distribution des spectres des deux classes, non-inflammé en bleu et inflammé en rouge. Les deux zones colorées, respectivement en bleu et rouge sont délimitées par le premier et le troisième quartile de chaque classe. Nous observons que la distribution des spectres n'est pas la même pour les deux classes ce qui confirme la possibilité de classer les spectres selon les résultats histologiques des souris. On note en particulier que, pour certaines longueurs d'onde, il n'y a pas de différence entre les classes, par exemple pour les intervalles [400, 450 nm] et [800, 950 nm]. Il y a plus de différence entre 500 et 700 nm, les spectres de souris inflammées sont clairement au-dessus des spectres non-inflammés. Pour autant, la séparation n'est pas nette, les zones colorées se chevauchent et le premier quartile des spectres inflammés est toujours en-dessous du troisième quartile des spectres non-inflammés.

Un autre problème se pose : le nombre de points de mesure dans l'intervalle [400, 950 nm] est de 953. Ce nombre est trop grand par rapport au nombre d'acquisitions. Leur utilisation telle quelle comme descripteurs pour la classification conduirait à des problèmes de sur-apprentissage dans le pipeline de classification. Nous avons donc élaboré un algorithme de fusion de bandes pour réduire le nombre de descripteurs.

Nous sommes partis de l'observation que certaines bandes sont extrêmement corrélées entre elles comme le montre la figure 4.8a. On voit apparaître sur la diagonale, plusieurs carrés rouges de taille plus ou moins grande. Ces zones correspondent à des groupes de longueurs d'onde à forte corrélation qui peuvent être facilement regroupées sans qu'il y ait une grande perte d'information.

La stratégie employée est de fusionner de manière récursive les bandes spectrales qui sont les plus corrélées. À chaque itération, on calcule donc la corrélation entre bandes voisines et fusionnons celles qui ont le coefficient de corrélation le plus haut. L'algorithme s'arrête si le coefficient de corrélation le plus haut est en dessous d'un certain seuil. En l'occurrence, nous fixons de manière empirique un seuil à $\tau = 0,95$. Ce seuil est élevé mais il correspond tout de même à une réduction de 953 bandes à 6 bandes tout en préservant un maximum d'information.

La figure 4.8b présente les six groupes de longueurs d'onde obtenus ainsi qu'un spectre moyen de souris. On définit ensuite les λ_i comme étant les longueurs d'onde de coupure qui définissent nos six groupes : $\lambda_{[0...6]} = 400 \text{ nm}, 451 \text{ nm}, 516 \text{ nm}, 589 \text{ nm}, 620 \text{ nm}, 668 \text{ nm}$ et 950 nm. Notons que

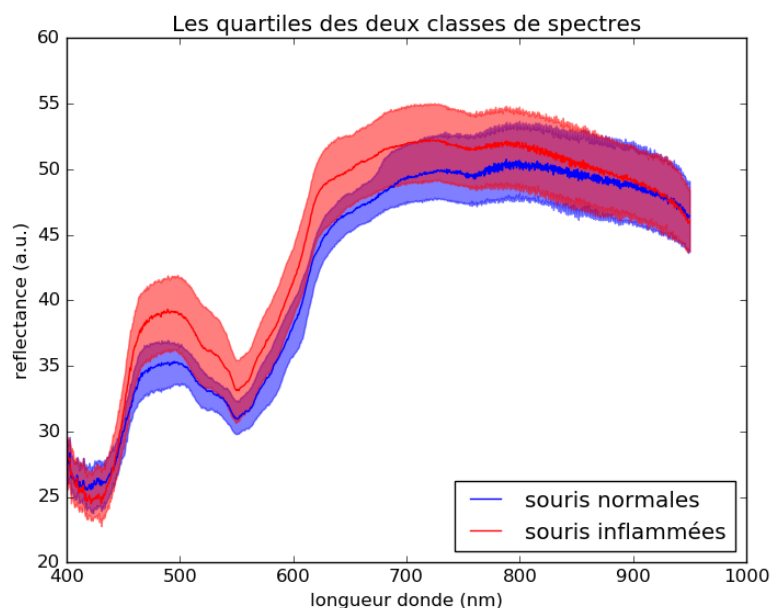
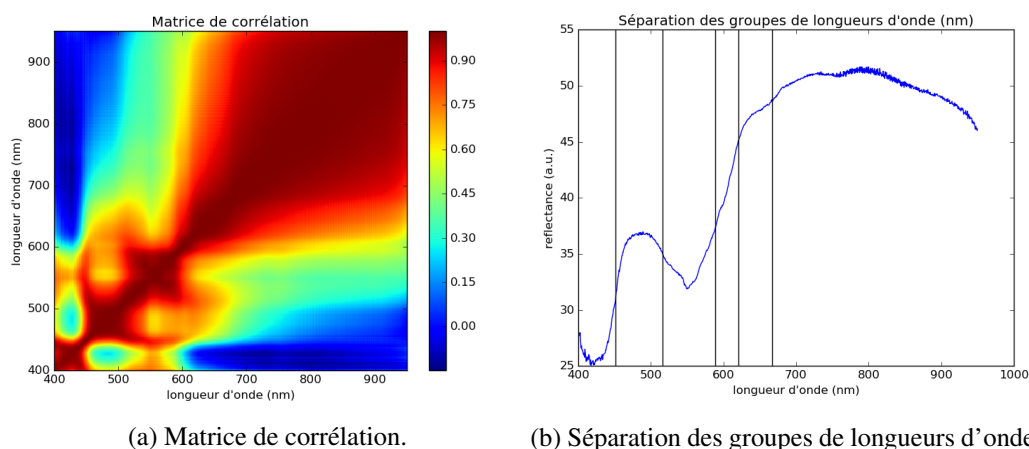


FIGURE 4.7 – Distribution des spectres des deux classes.



(a) Matrice de corrélation.

(b) Séparation des groupes de longueurs d'onde.

FIGURE 4.8 – Corrélation entre bandes spectrales.

le dernier groupe est relativement large, ce qui indique que les longueurs d'onde entre 668 nm et 950 nm sont fortement corrélées et donc l'information dans cet intervalle de longueurs d'onde est redondante.

L'algorithme de fusion de bandes est détaillé par l'algorithme en pseudo-code 1. Cet algorithme emprunte ses idées au regroupement hiérarchique descendant. Au début de l'algorithme, chaque élément est le seul représentant de son groupe (*i.e.* il y a 953 groupes de 1 longueur d'onde). À chaque itération, deux groupes sont fusionnés. Une version de cet algorithme existe avec des contraintes sur les groupes qui peuvent être fusionnés [61]. Dans notre cas, nous ajoutons une contrainte de connectivité : seules des bandes voisines peuvent être fusionnées. Cet algorithme est non-supervisé dans le sens où l'étiquette des souris (*i.e.* estomac normal ou inflammé) n'est pas utilisée pour trouver les longueurs d'onde de coupure. Ce choix est essentiel pour ne pas introduire de biais dans la classification.

Algorithme 1 : Pseudo code de l'algorithme de fusion de bandes

Entrée : M = une matrice de taille $(N_{spectres}, N_{bandes})$ qui contient tous les spectres
 τ = une constante égale à 0,95 par défaut

Sortie : M_{reduit} = une matrice de taille $(N_{spectres}, N_{comp})$ avec $N_{comp} \leq N_{bandes}$

$m \leftarrow 1,0$

$M_{reduit} \leftarrow M$

répéter

$c = \text{correlation}(M_{reduit})$, les corrélations entre bandes voisines

$m \leftarrow \max(c)$

$x, y \leftarrow \text{argmax}(c)$

 FUSIONNER les colonnes x et y dans M_{reduit}

tant que $m > \tau$

Les bandes qui ont été regroupées sont réduites à leur moyenne spectrale μ_i sur l'intervalle de longueurs d'onde considéré suivant l'équation 4.1 :

$$\mu_i = \frac{1}{N_i} \sum_{\lambda=\lambda_i}^{\lambda_{i+1}} S_{\lambda} \quad (4.1)$$

où N_i représente le nombre de longueurs d'onde dans l'intervalle $[\lambda_i, \lambda_{i+1}[$ et S_{λ} est la réflectance du spectre à la longueur d'onde λ .

Après avoir réduit le nombre de longueurs d'onde, nous extrayons des descripteurs. Nous avons testé trois types de descripteurs : soit nous utilisons les bandes "réduites" après l'algorithme de fusion, soit nous calculons des descripteurs de "division" et de "soustraction". Il s'agit simplement de calculer toutes les divisions ou toutes les soustractions possibles à partir des six bandes réduites. Ces deux types de descripteurs ont l'avantage d'être invariants, respectivement à la multiplication et à l'addition par un scalaire et sont donc des descripteurs robustes pour la classification. De manière plus formelle, nous notons $R_{i,j}$ et $D_{i,j}$ les ratios et différences correspondant aux descripteurs de division et de soustraction. Ils sont définis par les équations 4.2 et 4.3 :

$$R_{i,j} = \frac{\lambda_i}{\lambda_j}, \quad (4.2)$$

$$D_{i,j} = \lambda_i - \lambda_j. \quad (4.3)$$

Une étape de sélection univariée des descripteurs est ensuite appliquée pour réduire la complexité du classifieur. Réduire la complexité permet de réduire les risques de sur-apprentissage. Nous avons utilisé le test F de l'analyse de variance [62] (ANOVA : ANalysis Of VAriance en anglais) pour déterminer les k descripteurs les plus discriminants. Il s'agit d'une méthode de sélection univariée qui consiste à calculer une probabilité que les classes soient distribuées de la même manière. Plus cette probabilité, appelée valeur-p, est basse, plus les classes sont séparables.

La classification, quant à elle consiste à utiliser un classifieur de type SVM avec un noyau linéaire. Une validation croisée est effectuée : A chaque itération, les spectres d'une souris sont mis à l'écart et les autres sont utilisés pour entraîner le classifieur. La classe véritable de chaque souris est définie par son étiquette histologique. Les souris étiquetées (-) ou ($\pm$) sont classées comme normales et les autres comme inflammées en accord avec l'histologie présentée figure 4.4. En effet, il est plus judicieux de se baser sur les grades histologiques que sur le fait que la souris soit infectée ou non. Nous remarquons sur la figure 4.4 qu'il est possible qu'une souris infectée ne

développe pas d'inflammation et que, au contraire, des souris non-infectées développent tout de même une inflammation.

La totalité du pipeline a été créé en Python grâce à la librairie Scikit-Learn [63]. L'ensemble est résumé par l'algorithme en pseudo-code 2.

Algorithme 2 : Pipeline de classification

Entrée : M = une matrice de taille $(N_{spectres}, N_{bandes})$ qui contient tous les spectres
 y = un vecteur de taille $(N_{spectres})$ qui correspond aux labels de chaque spectre

Sortie : r = un vecteur de taille $N_{spectres}$, les résultats de classification

M_{reduit} = réduction de M par l'algorithme 1

pour chaque souris s faire

M_{test} = Spectres de la souris s dans M_{reduit}
 $M_{entraînement}$ = M_{reduit} privé de M_{test}
 $X_{entraînement}$ = CALCULER les descripteurs de $M_{entraînement}$
 X_{test} = CALCULER les descripteurs de M_{test}
 F = SELECTION des meilleurs descripteurs en fonction de $X_{entraînement}$
 $X_{F,entraînement}$ = descripteurs F de $X_{entraînement}$
 $X_{F,test}$ = descripteurs F de X_{test}
 clf = un classifieur SVM linéaire
 ENTRAINER clf avec $X_{F,entraînement}$
 $r[s]$ = TEST clf avec $X_{F,test}$

fin pour chaque

4.4/ RÉSULTATS DE CLASSIFICATION

Comme très peu de souris à 1, 3 et 6 mois avaient véritablement développé une inflammation, nous avons décidé que seules les données à 12 et 18 mois seraient utilisées pour entraîner et tester les classifieurs. Nous avons testé et comparé les trois types de descripteurs présentés dans la section précédente : les bandes réduites (*BR*), les descripteurs de division (*DIV*) et les descripteurs de soustractions (*SOU*).

Les figures 4.9b, 4.9d et 4.9f nous montrent les distributions statistiques de ces descripteurs en fonction de l'étiquette "normal" (en bleu) ou "inflammé" (en rouge). Nous observons en parallèle à ses figures, les valeurs-p associées qui donnent une indication sur le pouvoir discriminant des descripteurs. Dans les figures 4.9a, 4.9c et 4.9e, nous affichons le logarithme des valeurs-p pour une meilleure visualisation. Les meilleurs descripteurs sont les descripteurs avec une valeur-p très basse. La distribution des descripteurs de type "bandes réduites" nous montrent que μ_0 et μ_5 ne sont pas très informatifs car les distributions montrées par les boîtes à moustaches rouge et bleu sont quasiment identiques. Au contraire, μ_1 , μ_3 et μ_4 sont plus informatifs avec des valeurs-p en dessous de 10^{-8} . De la même manière, les divisions $R_{1,3}$ et $R_{2,4}$ ne sont pas informatives mais $R_{4,5}$ semble très intéressant. Enfin, les soustractions $D_{3,4}$, $D_{1,4}$ et $D_{1,3}$ seraient inefficaces pour la classification alors que $D_{0,1}$ et $D_{4,5}$ sont intéressantes pour séparer les deux classes.

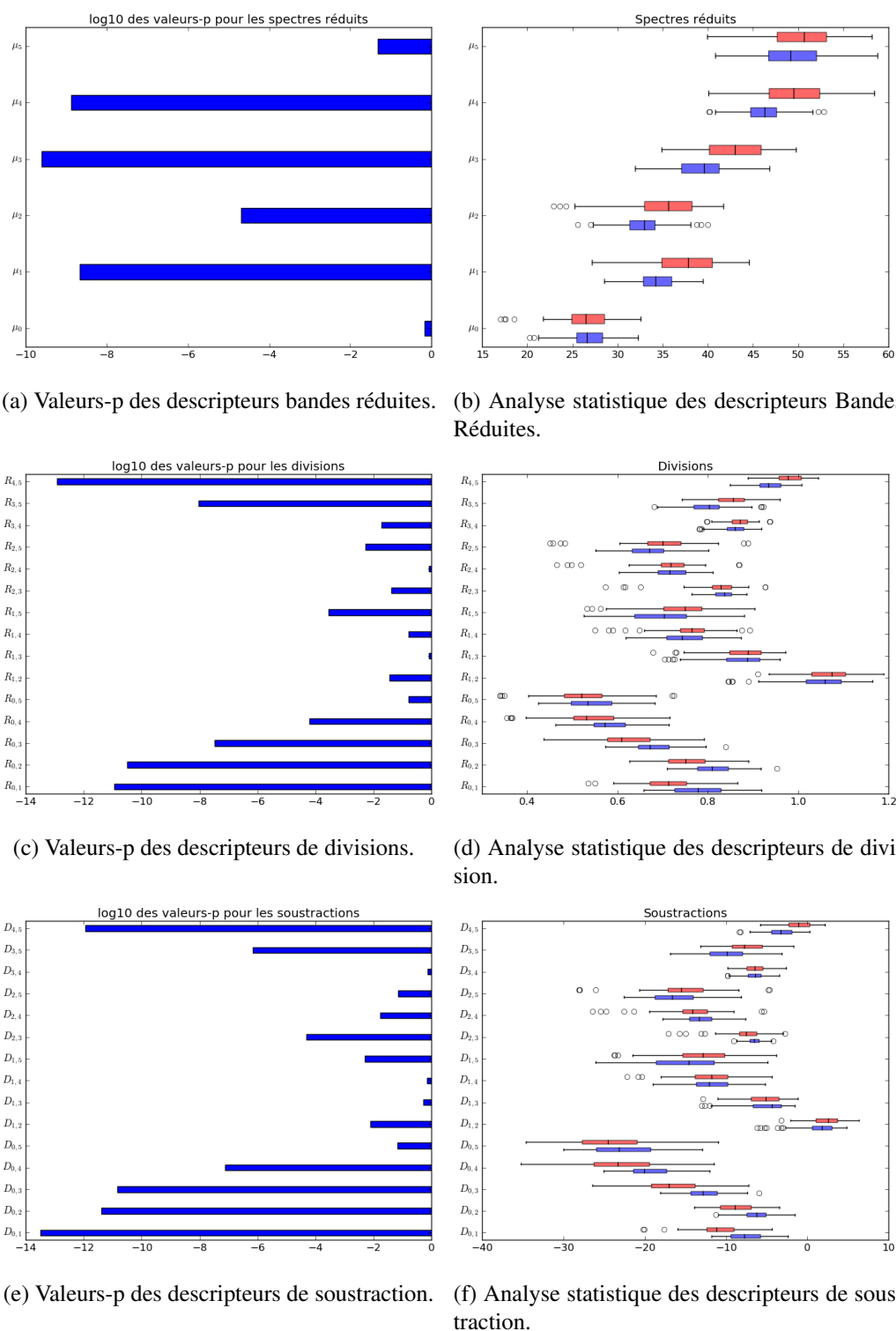


FIGURE 4.9 – Analyse statistique des descripteurs.

De manière plus concrète, le tableau 4.1 résume les résultats de classification obtenus avec un classifieur de type SVM avec noyau linéaire. Les meilleurs résultats obtenus avec les bandes réduites sont 79% de précision, 76% de rappel et un score F1 de 76%. Ces résultats sont proches des résultats obtenus avec les descripteurs de soustractions (81% de précision, 77% de rappel et un

score F1 de 77%). Cette similitude est attendue dans le sens où il s'agit simplement d'une transformation linéaire des descripteurs "bandes réduites". Les meilleurs résultats sont obtenus grâce aux descripteurs de divisions (98% de précision, 98% de rappel et un score F1 de 98%). Ces résultats montrent l'intérêt d'utiliser ce type de descripteurs. Les descripteurs de division ont l'avantage d'être invariant à l'échelle. Cela revient indirectement à normaliser les spectres. Le tableau indique également les valeurs des hyperparamètres utilisés pour la classification. D'une part k , le nombre de descripteurs retenus lors de la sélection univariée et c , le paramètre qui permet d'avoir une marge souple avec les SVM. Il s'agit d'une constante qui permet de contrôler le compromis entre nombre d'erreurs de classement et la largeur de la marge [19].

TABLE 4.1 – Résultats de classification des spectres.

	Hyperparamètres optimaux	Précision	Rappel	Score F1
<i>Bandes Réduites</i>	$c = 0,01$ $k = 6$	0,79	0,76	0,76
<i>Divisions</i>	$c = 1$ $k = 1$	0,98	0,98	0,98
<i>Soustractions</i>	$c = 0,01$ $k = 4$	0,81	0,77	0,77

De plus, nous avons fait varier le nombre de descripteurs sélectionnés via ANOVA. Les résultats présentés dans le tableau 4.1 sont les meilleurs résultats pris parmi l'ensemble des tests en faisant varier le nombre de descripteurs sélectionnés entre 1 et 15. Les résultats avec les descripteurs de division qui atteignent 98% en précision, rappel et score F1 ont été obtenus en sélectionnant un seul ratio (car $k=1$). Ce ratio est le quotient $R_{4,5}$ entre les groupes de longueurs d'onde [620 nm, 668 nm[et [668 nm, 950 nm[. Ce ratio a été choisi la plupart du temps comme étant le ratio le plus discriminant. En effet, avec la boucle de validation croisée, il est possible de ne pas avoir à chaque itération le même ratio. Ce ratio en particulier semble très intéressant pour la détection de l'inflammation d'autant plus qu'il est suffisant pour la classification. Une classification à un seul descripteur est avantageuse dans le sens où elle s'en trouve simplifiée et est plus robuste car la diminution du nombre de descripteurs réduit également le risque de sur-apprentissage.

4.5/ CONCLUSION ET PERSPECTIVES

En conclusion, les travaux sur les estomacs de souris nous ont permis de mettre en place une stratégie robuste pour analyser un ensemble de spectres. Nous avons atteint une précision, un rappel et un score F1 de 98% et nous avons isolé deux intervalles de longueurs d'onde pertinents pour la détection de l'inflammation due à *Helicobacter pylori* chez les souris. Ces intervalles sont [620 nm, 668 nm[et [668 nm, 950 nm[. Plus précisément c'est le ratio de ces deux intervalles qui permet de détecter l'inflammation.

Pour autant, ces résultats sont à confirmer et il serait intéressant de les corrélérer avec une explication biologique, c'est à dire relier la différence dans les spectres à un changement dans les chromophores de l'estomac. Nous pourrions également envisager d'acquérir des images multispectrales à la place d'un seul spectre sur les estomacs de souris ce qui nous permettrait d'étudier l'aspect spatial et de localiser les zones inflammatoires dans l'estomac.

PRÉSENTATION ET CALIBRAGE DU PROTOTYPE D'ACQUISITION

Pour pouvoir également mener nos recherches sur les pathologies de l'estomac chez l'Homme, nous avons élaboré et mis en place un système d'acquisition bimodal permettant d'acquérir simultanément des images endoscopiques classiques et des images multispectrales. Ce système a besoin d'être calibré spatialement et spectralement. Spatialement parce que l'on doit connaître la transformation géométrique qui permet de recalibrer les images endoscopiques et multispectrales, et spectralement pour pouvoir estimer le spectre réel de la surface de l'estomac à partir des images brutes produites par le capteur multispectral. Nous détaillons donc dans ce chapitre le système bimodal que nous avons mis en place ainsi que la manière dont nous calibrons ce système. Nous montrons également comment la méthode de calibrage peut être utilisée pour simuler des images en lumière blanche et en NBI.

Sommaire

5.1	Prototype d'acquisition	66
5.2	Calibrage spectral	68
5.2.1	Transformation linéaire pour le calibrage	69
5.2.2	Validation de l'estimation spectrale	70
5.3	Vers la simulation en lumière blanche et en NBI	73
5.3.1	Extension de la méthode pour la simulation d'image en vraie ou fausse couleur	73
5.3.2	Transformation en lumière blanche	73
5.3.3	Transformation en NBI	75
5.4	Conclusion et perspectives	76

5.1/ PROTOTYPE D'ACQUISITION

Au vu des systèmes existants et de leurs limites décrites dans le chapitre 1, à savoir que le NBI est limité en nombre de bandes et que le système FICE ne fait qu'une estimation des bandes spectrales, nous avons choisi de mettre en place un système qui permet d'acquérir en même temps des images endoscopiques et des images multispectrales. Le système utilise une colonne endoscopique pré-existante avec un endoscope Olympus Exera III de résolution 720x576 pixels. À cette colonne s'ajoute la partie permettant d'acquérir les images multispectrales ainsi qu'un ordinateur sur lequel nous avons développé le logiciel d'acquisition. Cette partie est visible sur l'image 5.2a.

Pour obtenir des images à l'intérieur de l'estomac, un fibroscope (microflex fiberscope de ITConcepts de 2,5 m de longueur et de 2,5 mm de diamètre [4]) est utilisé. Il est inséré dans le canal opérateur de l'endoscope, ce même canal opérateur qui est habituellement utilisé pour les biopsies comme illustré sur la figure 5.3b. Le fibroscope contient des fibres émettrices et des fibres réceptrices, les premières servent à illuminer l'estomac et les secondes à capturer l'image.

Nous avons souhaité avoir des informations spectrales dans le visible et dans le proche infrarouge. Malheureusement, il est très difficile d'avoir ces informations avec un seul et même capteur, nous avons donc trouvé une solution avec deux caméras. Les deux caméras multispectrales sont des caméras XIMEA basées sur la technologie CMOSIS CMV2000 [9]. Les deux caméras sont référencées CMV2K SSM4x4 et CMV2K SSM5x5. Elles peuvent respectivement acquérir dans les intervalles de longueur d'onde 450-620 nm et 600-1000 nm. Par abus de langage et pour les différencier, nous parlerons dans la suite de caméra visible (VIS) et de caméra proche infrarouge (NIR). Elles contiennent respectivement 16 et 25 bandes spectrales arrangées en grilles de 4x4 ou 5x5 filtres comme le montre la figure 5.1. Les nombres à l'intérieur des cases correspondent à la longueur d'onde (ou aux longueurs d'onde) du pic de sensibilité. Les capteurs ont respectivement une résolution de 2048x1024 pixels et 2045x1080. Comme des blocs de 4x4 ou 5x5 pixels sont utilisés pour avoir toutes les bandes spectrales, les véritables résolutions sont 512x256 et 409x216.

0	1	2	3
490	500	477	478
4	5	6	7
577	591	563	553
8	9	10	11
618	616 478	613	600
12	13	14	15
538	549	523	511

(a) Matrice de filtres pour la caméra CMV2K SSM4x4.

0	1	2	3	4
904 613	913 621	895 604	885 603	971
5	6	7	8	9
713	813	825	801	788
10	11	12	13	14
763	776	750	737	713
15	16	17	18	19
944 651	949 660	934 644	926 635	966 677
20	21	22	23	24
865	875	854	842	961 668

(b) Matrice de filtres pour la caméra CMV2K SSM5x5.

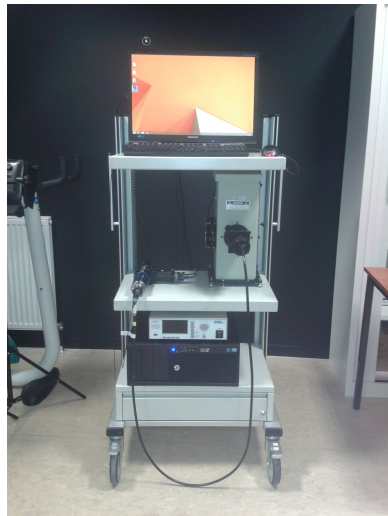
FIGURE 5.1 – Matrices de filtres des deux caméras.

Pour que les deux caméras reçoivent de la lumière, il nous faut un moyen de diviser les rayons en deux. Pour cela, nous utilisons un filtre dichroïque qui permet de séparer les basses longueurs d'onde des hautes longueurs d'onde (respectivement les intervalles 470-620 nm et 600-975 nm) avant d'être captées par les caméras. Le système séparant la lumière en deux est la "TwinCam" de

la société Cairn Research (figure 5.2b) [8].

Dans notre cas, il est également nécessaire d'utiliser une source externe de lumière de 1000W pour que les caméras multispectrales reçoivent suffisamment de lumière. Il y a en effet des pertes d'intensité, notamment à cause de la fibre, de la séparation des flux dans la Twincam et de la sensibilité des caméras. Il s'agit d'une source mercure-xénon (figure 5.2c) contrôlée par un boîtier (figure 5.2d) proposée par Newport® (référence : 66924-1000XF-R1 [3]).

Enfin, les caméras multispectrales sont reliées par port USB3 à l'ordinateur et l'endoscope est relié à l'ordinateur par un câble firewire. Nous avons développé un logiciel d'acquisition spécifique, en C# qui nous permet de capturer simultanément les flux vidéo des deux caméras multispectrales et de l'endoscope.



(a) Colonne complète avec l'ordinateur, les caméras, le miroir, la source de lumière et le fibroscope.



(b) Miroir semi-réfléchissant et les deux caméras.



(c) Source de lumière mercure-xénon.



(d) Contrôleur de la source de lumière.

FIGURE 5.2 – Ensemble des éléments du prototype.

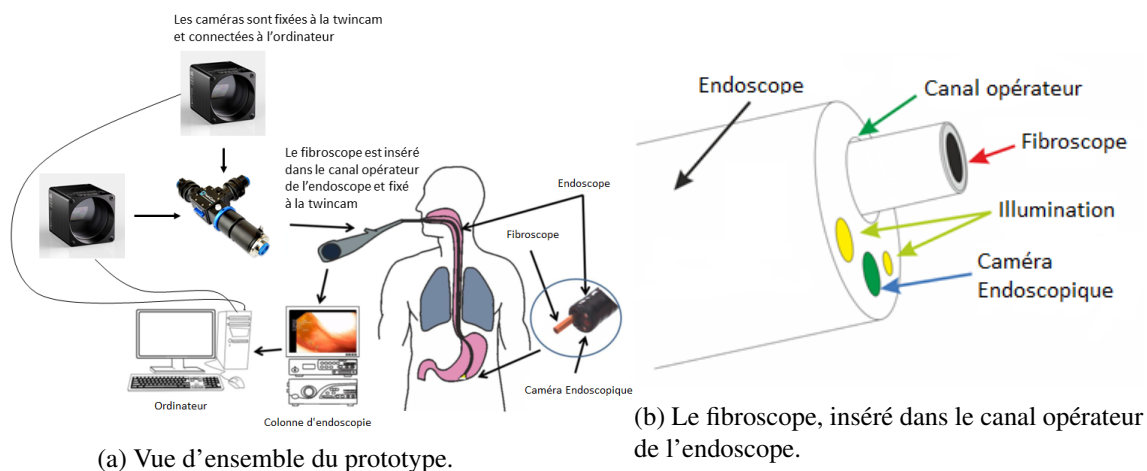


FIGURE 5.3 – Schéma fonctionnel du prototype.

5.2/ CALIBRAGE SPECTRAL

Le prototype que nous proposons doit être calibré spatialement et spectralement. Dans le cadre du projet, le calibrage spatial est réalisé par l'équipe PRISME. Il s'agit de trouver la transformation projective entre l'image endoscopique et l'image multispectrale. Pour cela, nous nous aidons du fait que le bout du fibroscope soit visible dans l'image endoscopique. La translation et le facteur d'échelle sont reliés quadratiquement à la longueur de fibre visible dans l'image endoscopique [E] et [G].

Le calibrage spectral, quant à lui, est nécessaire parce que les caméras multispectrales ne permettent pas d'avoir directement des spectres en fonction des longueurs d'onde. Les caméras que nous utilisons sont basées sur des structures de type Fabry-Pérot. Il s'agit de deux surfaces réfléchissantes espacées d'une certaine distance L . La longueur d'onde λ à laquelle le filtre est sensible dépend de cette distance. Ce principe est illustré par la figure 5.4a. En pratique, les filtres ont effectivement une réponse dite primaire à une certaine longueur d'onde λ , mais également des réponses secondaires aux harmoniques de cette longueur d'onde. De plus, il existe également un problème de diaphonie (crosstalk en anglais) : chaque filtre possède en réalité des réponses parasites en dehors de l'intervalle prévu qu'il faut corriger comme illustré figure 5.4b.

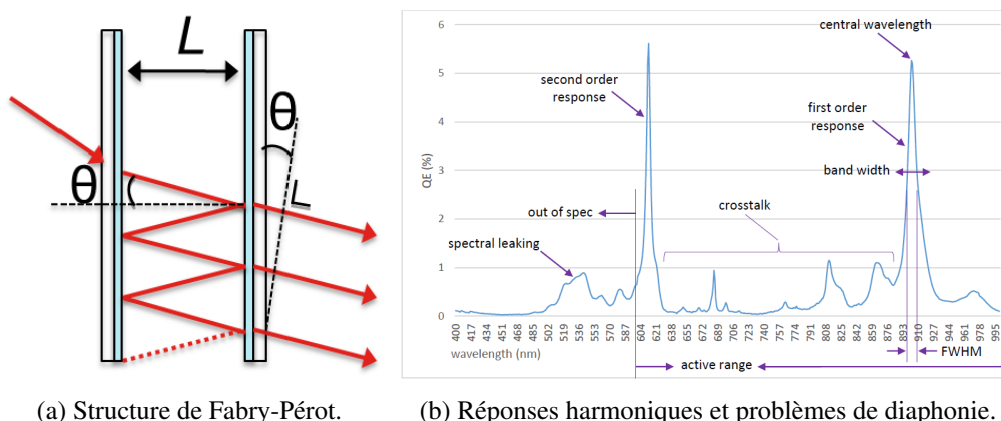


FIGURE 5.4 – Fonctionnement des caméras multispectrales.

5.2.1/ TRANSFORMATION LINÉAIRE POUR LE CALIBRAGE

Le calibrage des caméras se fait habituellement en acquérant un noir et un blanc de référence et en appliquant une matrice de transformation pour régler les problèmes de diaphonie. Malheureusement, il est impossible de réaliser ce calibrage lors d'une utilisation *in vivo*. En effet, nous faisons l'acquisition d'une vidéo, la scène n'est donc pas fixe. De plus, le fibroscope étant nettoyé entre chaque patient, il est continuellement débranché et rebranché entre chaque patient. Le calibrage spatial et le calibrage spectral devraient alors être systématiquement refaits, ce qui est une procédure trop lourde pour le personnel médical. Le calibrage doit donc être réalisé en post-traitement.

Pour réaliser le calibrage convenablement, nous avons besoin d'une part du temps d'exposition de chaque image, d'images noires de référence et de patches de couleurs dont la réflectance est connue. Le temps d'exposition n'est pas fixe, il varie dynamiquement dans le logiciel d'acquisition. Le but est de pouvoir répondre aux variations de l'intensité lumineuse reçue, causées par le rapprochement ou l'éloignement de la fibre par rapport à la paroi gastrique. Le calibrage dépend de ce temps d'exposition. Les images noires permettent de supprimer le bruit numérique de la caméra. Nous avons acquis plusieurs images noires à différents temps d'acquisitions pour caractériser ce bruit. Nous utilisons les patches d'un Digital ColorChecker® pour avoir des couleurs de référence. Le Digital ColorChecker® est un ensemble de patches de couleurs spécifiques qui couvrent l'ensemble des couleurs de l'espace colorimétrique. Plus précisément, nous utilisons les 96 patches centraux soit les lignes 2 à 9 et les colonnes B à M délimitées par le cadre rouge comme sur la figure 5.5. Des acquisitions sont faites avec les caméras multispectrales en se plaçant à environ 2 cm de distance au-dessus des patches pour être dans des conditions proches de celles dans l'estomac et avec un spectromètre pour connaître les spectres de réflectance de ces patches.



FIGURE 5.5 – X-Rite ColorChecker® Digital SG
(issue de <https://www.xrite.com/categories/calibration-profiling/colorchecker-digital-sg>).

De manière plus formelle, soit x la moyenne spatiale des pixels d'une image multispectrale brute

et y , le spectre que l'on désire obtenir. x est un vecteur de 16 ou 25 éléments (16 pour le calibrage de la caméra VIS et 25 pour le calibrage de la caméra NIR). Nous réalisons en effet le calibrage des deux caméras de manière indépendante. y est un vecteur de 61 éléments correspondant aux longueurs d'onde de 400 nm à 1000 nm espacées de 10 nm. Nous considérons qu'il existe une matrice de transformation W (de dimension 16x61 ou 25x61) tel que :

$$\frac{x - b}{\tau} W = x' W = y \quad (5.1)$$

où b est le spectre moyen de l'image noire qui corrige le décalage dû au bruit de fond de la caméra et τ le temps d'exposition de la caméra.

La matrice W peut être apprise grâce aux patches du Digital ColorChecker®. Pour cela, nous avons acquis le spectre de ces patches grâce à un spectromètre (le même que celui utilisé dans les travaux sur les estomacs de souris au chapitre 4), donnant ainsi 96 vecteurs vérité-terrain y . Nous avons également acquis les patches avec les caméras multispectrales, donnant les vecteurs x . La matrice W est alors déterminée par régression sur composantes principales.

La régression sur composantes principales diffère de la régression linéaire traditionnelle par le fait que l'on applique une Analyse en Composantes Principales (ACP) sur les variables explicatives avant d'effectuer la régression. Cela permet de décorrélérer les variables explicatives et de rendre la régression plus stable [11]. Du point de vue de l'algèbre linéaire, cela revient à imposer que le rang de la matrice W soit inférieur ou égal au nombre de composantes.

5.2.2/ VALIDATION DE L'ESTIMATION SPECTRALE

L'estimation de W dépend du nombre de composantes que nous retenons lors de la régression. La figure 5.6 montre le RMSE (Root Mean Square Error, qui est la racine de l'erreur quadratique moyenne) en fonction du nombre de composantes retenues. La figure se présente sous forme de boîtes à moustache, les traits rouges correspondent aux médianes et les rectangles bleus sont délimités par les deux autres quartiles (25% et 50%). Ces résultats sont obtenus par validation croisée, en divisant la base des 96 patches en 3 groupes. De manière alternée, 2 d'entre eux sont utilisés pour apprendre la matrice W et un est utilisé pour tester le classifieur.

Nous observons que l'erreur tend à diminuer avec le nombre de composantes. La médiane la plus basse est obtenue en gardant les 16 composantes. Pourtant ce choix n'est pas le meilleur, d'une part parce qu'il y a une forte variance comme peut en témoigner la hauteur des boîtes à moustache, d'autre part les résultats médians varient en réalité très peu au-delà de l'ajout de la quatrième composante. Enfin, en observant les valeurs propres des matrices que forment nos données, nous remarquons que 4 composantes suffisent à expliquer plus de 99,9% de la variance comme le montre la figure 5.7. Pour ces raisons, nous choisissons de ne garder que 4 composantes dans la suite des expériences.

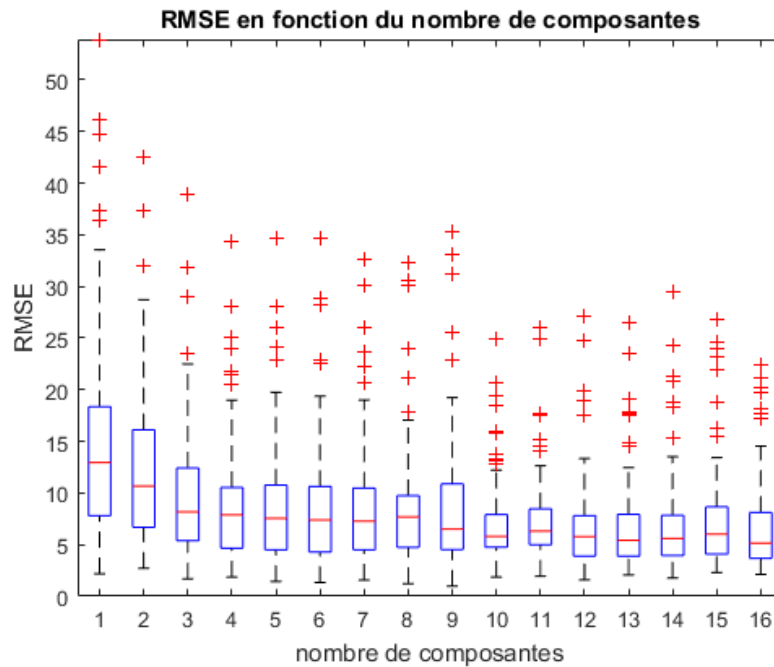
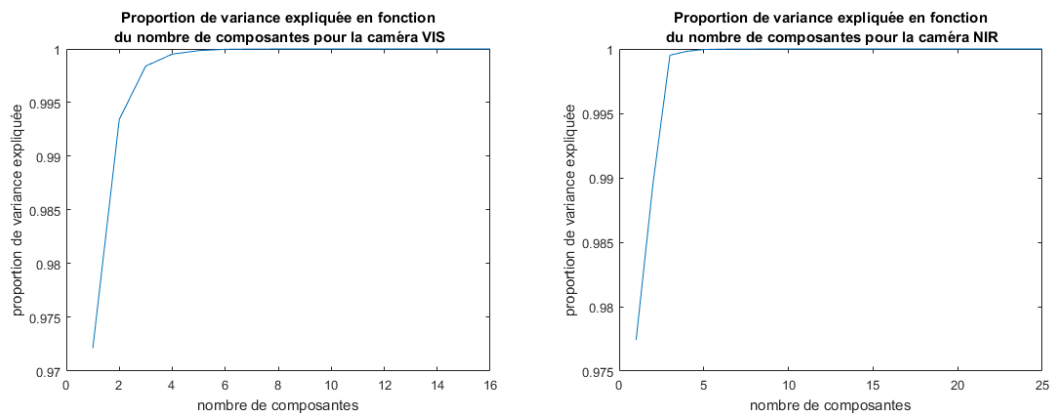


FIGURE 5.6 – RMSE en fonction du nombre de composantes.



(a) Proportion de la variance expliquée en fonction du nombre de composantes pour la caméra VIS. (b) Proportion de la variance expliquée en fonction du nombre de composantes pour la caméra NIR.

FIGURE 5.7 – Analyse des valeurs propres des données provenant des caméras VIS et NIR.

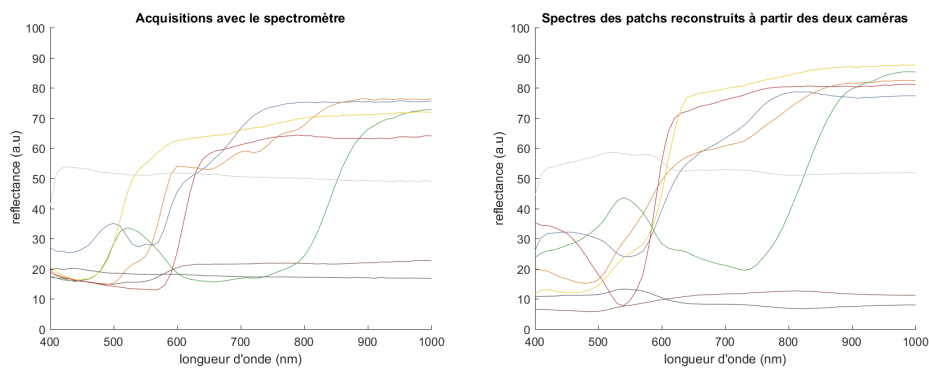
Les figures 5.8a et 5.8b montrent sur quelques exemples de patches (choisis pour être variés, gris, bleu, vert, jaune, orange, rouge et brun), respectivement les acquisitions faites avec le spectromètre et les mêmes spectres reconstruits à partir des caméras multispectrales. Nous retrouvons bien des aspects similaires entre les spectres. Nous observons tout de même des erreurs, par exemple pour le spectre du patch jaune qui ne monte pas suffisamment dans le proche infrarouge (650-1000 nm).

La figure 5.8c montre le RMSE pour tous les patches. Nous remarquons que les erreurs sont localisées sur quelques patches de la troisième ligne, dont le RMSE dépasse 25 alors que les autres patches ont des erreurs plus acceptables entre 0 et 20.

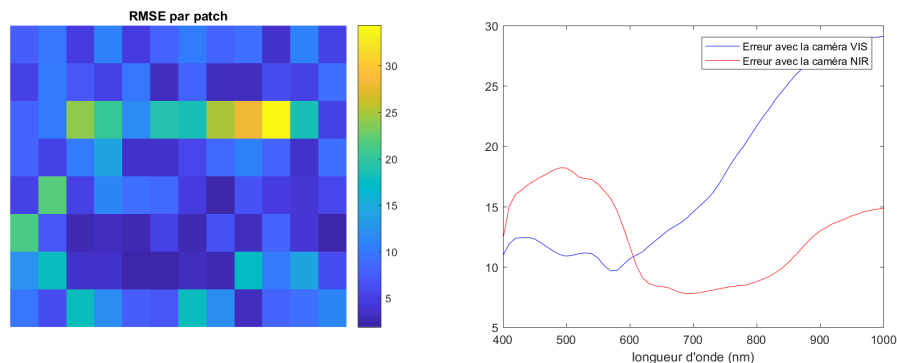
Il semble assez attendu que la caméra visible permette de bien prédire la réflectance aux longueurs

d'onde basses et la caméra proche infrarouge, celle au longueurs d'onde hautes. Il est alors intéressant de pouvoir avoir un intervalle commun qui nous permette de recalcr deux parties d'un même spectre. Nous déterminons donc simultanément les intervalles précis dans lesquels chacune des caméras peut être utilisée. En effet, les deux caméras n'ont pas la même précision pour estimer un spectre à une longueur d'onde donnée. Il s'agit alors de trouver quels sont les intervalles dans lesquels la caméra VIS est à privilégier et quels sont les intervalles dans lesquels la caméra NIR est à privilégier.

Les résultats sont visibles figure 5.8d. Les courbes représentent le RMSE sur tous les patches en fonction des longueurs d'onde. La courbe bleue est l'erreur de la caméra VIS seule et la courbe rouge, celle de la caméra NIR seule. Nous observons qu'entre 400 nm et 610 nm, il est préférable d'utiliser la caméra VIS, et au-delà, il est avantageux d'utiliser seulement la caméra NIR. Ces résultats sont cohérents avec l'intuition que l'on a, à savoir que la caméra VIS est faite pour les longueurs d'onde basses et la caméra NIR pour les longueurs d'onde hautes.



(a) Acquisitions avec le spectromètre (vérité-terrain). (b) Spectres des patches reconstruits à partir des deux caméras.



(c) RMSE par patch. (d) RMSE par longueur d'onde.

FIGURE 5.8 – Estimation de la réflectance à partir des caméras multispectrales.

Cette analyse nous permet de déterminer un intervalle que nous pouvons utiliser pour recalcr les deux parties de spectres. Nous choisissons l'intervalle [600, 630 nm] pour recalcr les deux parties de spectres. Dans l'exemple présenté figure 5.9, nous estimons donc la réflectance dans l'intervalle [400, 630 nm] grâce à la caméra VIS et la réflectance dans l'intervalle [600, 840 nm] grâce à la caméra NIR. L'intervalle [600, 630 nm] est utilisé pour recalcr les deux morceaux de spectres. Le recalage est effectué en ajustant un facteur multiplicatif. Le spectre est ensuite normalisé par sa norme l_1 , autrement dit, le spectre est divisé par la somme de ses éléments. La forme du spectre obtenu figure 5.9 ressemble bien à un spectre attendu, on retrouve le pic un peu avant 500 nm, le

creux à 560 nm et le plateau à partir de 650 nm que l'on a déjà observé chez la souris au chapitre 4.

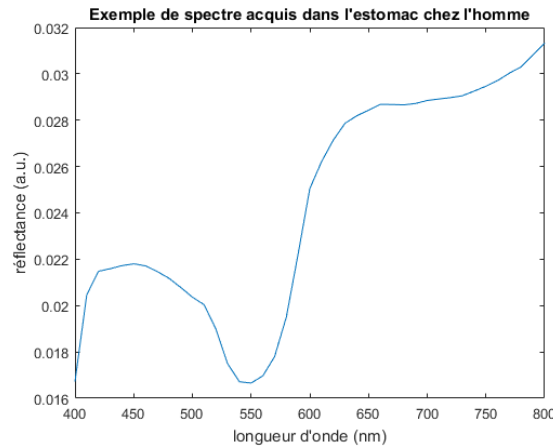


FIGURE 5.9 – Exemple de spectre acquis dans l'estomac chez l'Homme.

5.3/ VERS LA SIMULATION EN LUMIÈRE BLANCHE ET EN NBI

5.3.1/ EXTENSION DE LA MÉTHODE POUR LA SIMULATION D'IMAGE EN VRAIE OU FAUSSE COULEUR

Dans la méthode utilisée pour le calibrage, nous avons supposé l'existence d'une relation linéaire qui permet d'estimer un spectre à partir des données des caméras. À partir des spectres, nous pouvons également supposer qu'il existe une relation linéaire pour simuler respectivement les couleurs vues en NBI et en lumière blanche. Par transitivité, il existe donc une relation linéaire qui permet d'estimer les couleurs NBI ou en lumière blanche à partir des acquisitions multispectrales. Nous pouvons donc écrire qu'il existe deux matrices, W_{NBI} et W_{WL} telles que

$$x' W_{NBI} = y_{NBI} \quad (5.2)$$

et

$$x' W_{WL} = y_{WL} \quad (5.3)$$

où y_{NBI} est un vecteur à trois composantes correspondant à la couleur d'un patch sous NBI et y_{WL} un vecteur de trois composantes correspondant aux canaux R, G et B d'un patch acquis sous lumière blanche.

Même si l'on retrouve la même formulation mathématique, l'intérêt de ces transformations est différent. Pour la transformation des images multispectrales en spectre, le but est de s'affranchir de la lumière et de la sensibilité du capteur alors que les deux transformations 5.2 et 5.3 ont un but de simulation et de visualisation.

5.3.2/ TRANSFORMATION EN LUMIÈRE BLANCHE

Pour simuler ce que l'on verrait en lumière blanche à partir des images multispectrales, nous apprenons également une matrice de transformation grâce au ColorChecker®. Pour être précis,

nous cherchons une transformation entre les patches acquis avec la caméra VIS et les patches acquis avec l'endoscope en lumière blanche exprimés dans l'espace $L^*a^*b^*$. En effet, l'espace $L^*a^*b^*$ est plus adapté à la perception des couleurs par l'œil humain.

La figure 5.10b montre l'image du ColorChecker® simulée à partir des données de la caméra visible. Nous avons calculé le ΔE_{ab} pour quantifier l'erreur pour chacun des patches. Le ΔE_{ab} moyen trouvé est égal à 11,44. Nous pouvons voir sur l'image 5.10c qu'en pratique, quatre patches ont un fort ΔE_{ab} supérieur à 25 et les autres se situent entre 0 et 25. Trois patches apparaissent plus lumineux dans l'image 5.10b, il s'agit des patches beige, rose et bleu.



(a) Acquisitions sous lumière blanche. (b) Patches en lumière blanche reconstruits à partir de la caméra VIS. (c) RMSE par patch.

FIGURE 5.10 – Transformation en image endoscopique sous lumière blanche.

Nous pouvons également appliquer la transformation sur d'autres images acquises, par exemple sur les images de la figure 5.11 qui représentent respectivement la paume d'une main et une mire en noir et blanc. Les figures 5.11a et 5.11b sont des images endoscopiques sous lumière blanche et les images 5.11c et 5.11d sont les images multispectrales acquises au même moment auxquelles nous avons appliqué la transformation décrite par l'équation 5.3. La mire en noir et blanc est assez bien rendue et la main semble un peu plus rouge que sa couleur naturelle mais les résultats sont relativement réalistes.

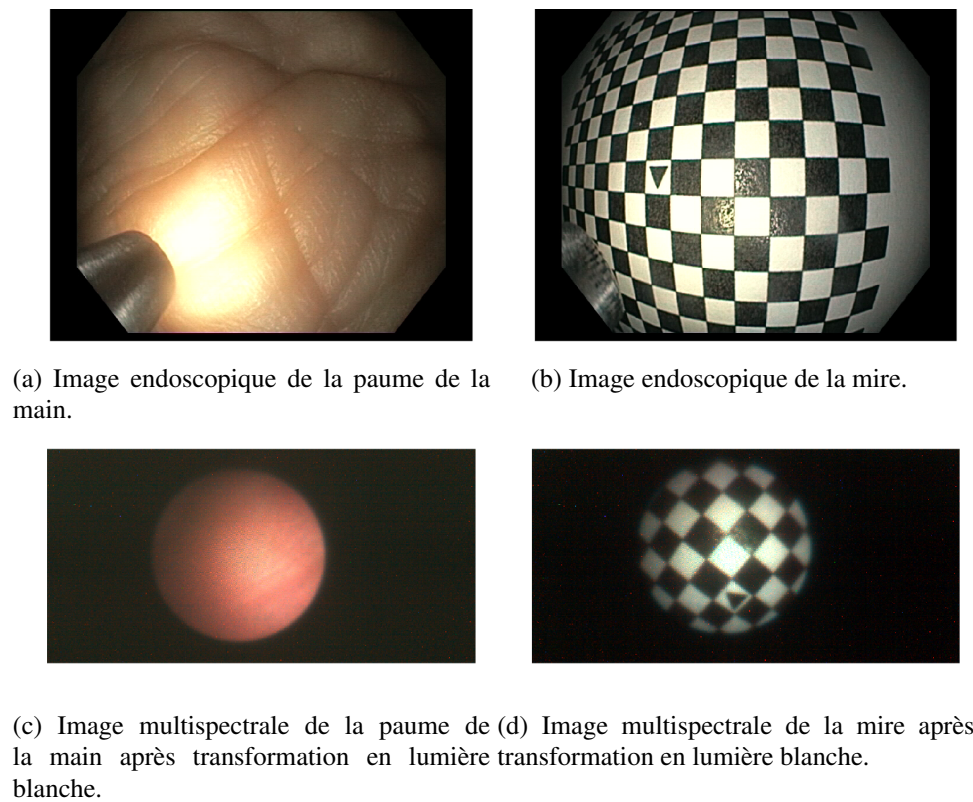


FIGURE 5.11 – Application de la transformation en image endoscopique sous lumière blanche sur des images de main et de mire.

5.3.3/ TRANSFORMATION EN NBI

Pour la simulation d'images en NBI, nous adoptons la même stratégie. Nous calculons la transformation entre les patches acquis avec la caméra VIS et les patches acquis sous NBI exprimés dans l'espace $L^*a^*b^*$.

La figure 5.12b montre l'image du ColorChecker® simulée à partir des données de la caméra visible comme s'il était vu sous NBI. Nous avons calculé pour tous les patches le ΔE_{ab} pour quantifier l'erreur pour chaque patch. Le ΔE_{ab} moyen trouvé est égal à 26,91. L'erreur est plus grande que pour la transformation en lumière blanche. Ceci peut s'expliquer en partie parce que les canaux des images NBI correspondent à des filtres relativement étroits, il est donc certainement difficile d'identifier les bandes de la caméra multispectrale visible qui ont une réelle influence parmi les 16 bandes. On peut supposer que seulement un petit nombre parmi elles réagissent dans les mêmes longueurs d'onde que le NBI et que les autres ajoutent un bruit conséquent. Nous observons également trois patches très lumineux dans l'image 5.12b. Il s'agit des mêmes patches que dans la section précédente. Ces patches ont probablement un bruit fort dû aux acquisitions.

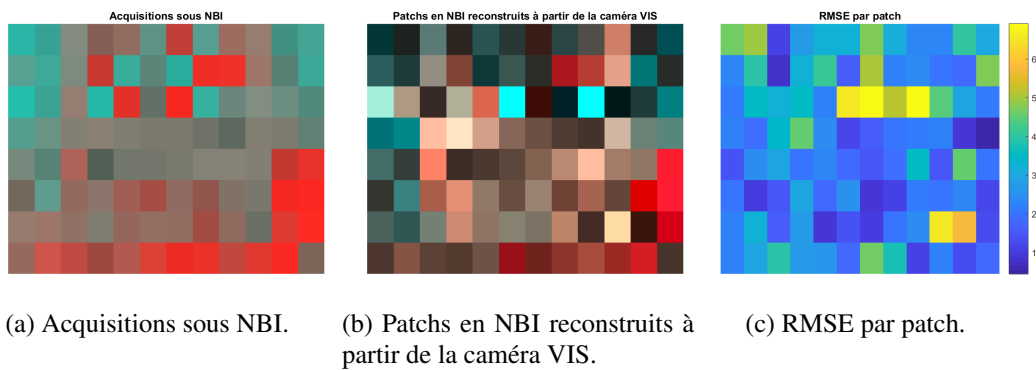


FIGURE 5.12 – Transformation en image endoscopique sous NBI.

De la même manière que précédemment, nous appliquons également la transformation sur l'image de la main et l'image de la mire (les images 5.13a et 5.13b sont prises en lumière blanche et ne sont données que pour l'illustration). Les résultats sont plus mitigés, la main apparaît bleue. S'il est difficile de prévoir la couleur obtenue sous lumière NBI, la couleur bleue n'est ici pas attendue. La mire montre elle un résultat tout à fait cohérent.

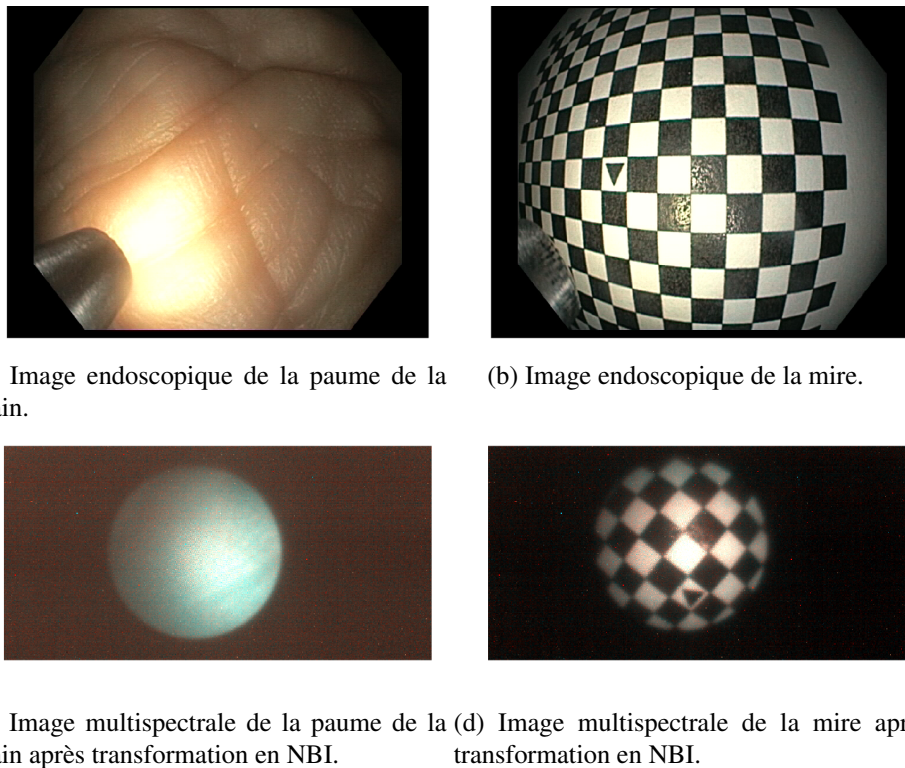


FIGURE 5.13 – Application de la transformation en image endoscopique sous NBI sur des images de main et de mire.

5.4/ CONCLUSION ET PERSPECTIVES

En conclusion, nous avons développé un système d'acquisition qui permet d'acquérir en même temps une vidéo endoscopique et une vidéo multispectrale. Le système consiste en un endoscope

déjà présent dans le bloc opératoire auquel on ajoute une colonne qui permet l'acquisition d'images multispectrales. Ce système doit être calibré spatialement et spectralement. Nous avons donc utilisé un modèle linéaire pour pouvoir estimer la transformation spectrale entre les données brutes acquises par les caméras multispectrales et le spectre correspondant. La transformation est estimée grâce aux patches d'un ColorChecker®. Le même procédé peut être utilisé pour simuler des images endoscopiques en lumière blanche ou en NBI, images qui sont plus usuelles du point de vue de l'endoscopiste.

Les calculs d'erreur sur l'estimation d'un spectre ou la simulation en lumière blanche ou en NBI montrent que le procédé de calibrage peut être amélioré. On peut notamment utiliser un éventail de patches plus diversifié, en particulier dans le proche infrarouge, les patches ayant tous une réponse relativement constante dans le proche infrarouge. Nous pourrions également utiliser un autre modèle qu'un simple modèle linéaire, qui serait plus juste par rapport à nos données ou bien encore utiliser d'autres termes de régularisation à la manière de la régression Lasso ou la régression Ridge qui ajoutent respectivement la norme l_1 et la norme l_2 comme terme de régularisation. Il serait également plus précis de considérer un système de calibrage plus évolué que la prise de photos de patches d'un ColorChecker®, par exemple avec des sources de lumière permettant de générer des couleurs sur des plages de longueurs d'onde fines.

Par rapport aux systèmes existant les plus utilisés, en l'occurrence le NBI et FICE, notre système possède tout de même des avantages conséquents : il contient plus de bandes que le NBI, il en contient 41 (25+16) et permet, tout comme FICE, d'estimer des spectres à différentes longueurs d'onde.

ÉTUDE PRÉLIMINAIRE DE LA RÉFLECTANCE DE LA PAROI GASTRIQUE CHEZ L'HOMME

Le prototype comprenant les caméras multispectrales, une fois calibré, nous permet d'acquérir des informations spectrales. Dans ce chapitre, il est question de différencier quatre types de tissus différents, à savoir, un estomac normal, un estomac avec gastrite active, un estomac avec gastrite chronique et un estomac avec métaplasie intestinale. Nous présentons les résultats de classification obtenus avec un pipeline de classification similaire à celui utilisé dans le chapitre 4 dans lequel nous avons à traiter des spectres issus des estomacs de souris. Ce pipeline nous permet d'identifier des longueurs d'onde discriminantes pour différencier les quatre types de muqueuse décrits précédemment.

Sommaire

6.1	Méthodes existantes d'analyse multispectrale	80
6.2	Acquisitions multispectrales chez l'Homme	81
6.2.1	Collecte des données	81
6.2.2	Première analyse des données	84
6.3	Classification des pathologies de l'estomac	85
6.4	Conclusion et perspectives	87

6.1/ MÉTHODES EXISTANTES D'ANALYSE MULTISPECTRALE

Il existe plusieurs travaux dans la littérature qui utilisent des systèmes multispectraux ou hyperspectraux pour détecter le cancer de l'estomac ou des lésions pré-cancéreuses. Certains d'entre eux ont déjà été présentés dans le chapitre 1 lors de la présentation des systèmes existants.

Kiyotoki *et al.* ont utilisé une caméra hyperspectrale entre 400 et 800 nm pour caractériser des spectres de tumeurs réséquées de 14 patients [80]. Ils ont sélectionné des spectres sur des zones saines et sur des zones de tumeurs. De bons résultats de classification ont été obtenus, notamment pour les longueurs d'onde entre 650 et 800 nm, les meilleurs résultats étant obtenus à 726 nm avec un score prédictif de 77,5%. Des travaux similaires ont été conduits par Ogihara *et al.*. Ils ont également utilisé une caméra hyperspectrale entre 400 et 800 nm [112] pour photographier des estomacs avec tumeurs réséquées. Ils ont observé une forte variabilité inter-patient qui les a conduit à normaliser l'ensemble des spectres en utilisant le spectre moyen des zones saines. Ils ont cherché la longueur d'onde la plus discriminante, qui, dans leur étude était 770 nm, en utilisant la distance de Mahalanobis.

Les études des systèmes hyperspectraux ne se limitent pas au visible et au proche infrarouge. Akbari *et al.* ont démontré que l'intervalle entre 1000 et 2500 nm permettait également de discriminer des spectres d'estomac sain et des spectres de tumeurs. Pour cela, ils ont utilisé des descripteurs comme l'intégrale ou la dérivée d'une portion des spectres, ou encore un écart type spatial-spectral [57]. Ces trois études donnent des arguments convaincants en faveur de l'utilisation de caméras multispectrales. Malheureusement, elles sont pour le moment limitées à un cadre *ex vivo* et sont focalisées sur le cancer.

Des travaux *in vivo* existent aussi dans la littérature. Martinez-Herrera *et al.* ont acquis des images multispectrales de la muqueuse de l'estomac *in vivo* grâce à un système endoscopique associé à une roue à filtres [110]. Ils ont ainsi collecté des images à six bandes de patients normaux et de patients avec lésions malignes. Les descripteurs utilisés pour discriminer les deux catégories de patients incluaient des statistiques sur les spectres, notamment l'écart type, l'écart à la médiane et l'entropie. Ces descripteurs ont permis d'entraîner un classifieur de type k-plus proche voisins ou SVM (avec noyau linéaire ou gaussien). Les meilleurs résultats ont été obtenus pour un classifieur SVM avec noyau gaussien, permettant de reconnaître un spectre de lésion maligne avec une précision de 77%.

En 2017, Hohmann *et al.* ont démontré l'intérêt des données multispectrales dans la détection d'adénocarcinomes dans l'estomac [122]. Leur système consiste en un endoscope standard couplé à une source de lumière modifiée qui permet d'extraire des données à six longueurs d'onde différentes entre 438 et 628 nm. Ils ont testé plusieurs types de classifieurs notamment des forêts d'arbre aléatoires, RobustBoost, Adaboost ou des SVM avec noyau linéaire ou gaussien leur permettant d'obtenir une précision de classification autour de 60%.

La spectroscopie Raman a déjà été utilisée plusieurs fois *in vivo* pour détecter des lésions de l'estomac. Lorsqu'un tissu est illuminé à une certaine longueur d'onde, la fréquence de la lumière est modifiée. L'analyse spectrale de la lumière réfléchie donne donc des informations sur la nature des tissus. Bergholt *et al.* ont ainsi réussi à distinguer des spectres normaux, d'ulcère bénin et d'ulcère malin avec une sensibilité entre 80% et 90% et une spécificité supérieure à 90% en utilisant une classification de type "analyse discriminante par moindres carrés partiels" (Partial Least Squares-Discriminant Analysis PLS - DA) [49]. Dans une autre étude, la même équipe de chercheurs a montré l'intérêt du démixage. Ils ont décomposé les spectres pour obtenir les contributions relatives de l'ADN, des protéines, des lipides, du sang et des glycoprotéines. Les contributions ont permis de reconnaître quatre types de tissus différents : les tissus normaux, la métaplasie intesti-

nale, la dysplasie et les adénocarcinomes [73].

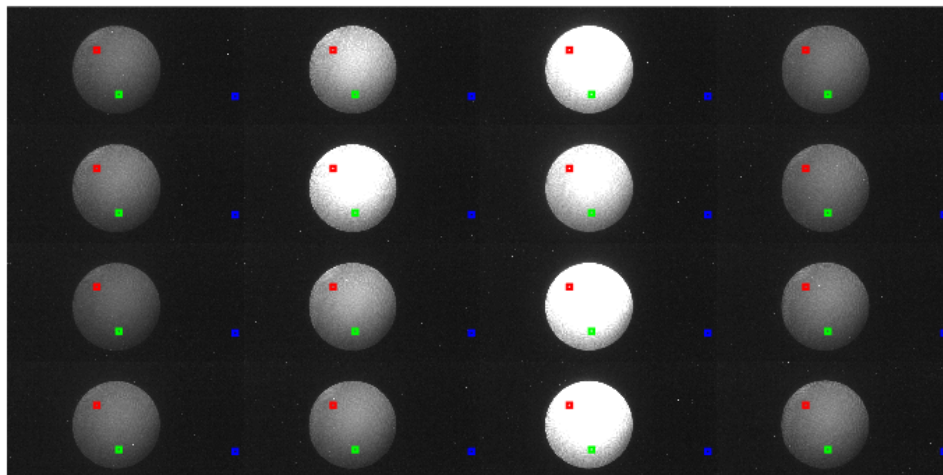
Dans un autre champ d'étude, Feng *et al.* ont réussi à détecter le cancer gastrique en analysant le plasma du sang de 32 malades et 33 personnes saines. Le plasma est analysé avec un système comprenant une fibre optique, un laser permettant une excitation à 785 nm et un polariseur. Les spectres obtenus sont classés via une analyse en composantes principales suivie d'une analyse discriminante linéaire. Ils ont réussi à montrer que la polarisation avait une influence sur les résultats, en l'occurrence, ils ont obtenu de meilleurs résultats avec une polarisation circulaire gauche [58]. L'utilisation de la polarisation donne des informations supplémentaires qui aident à la détection de pathologies et peut donc être une autre modalité intéressante à explorer.

Dans notre cas, nous ne nous intéressons pas au cancer comme la majorité des études mais à des stades pré-cancéreux, plus difficiles à traiter. Les spectres sont acquis *in vivo* grâce au système présenté au chapitre 5.

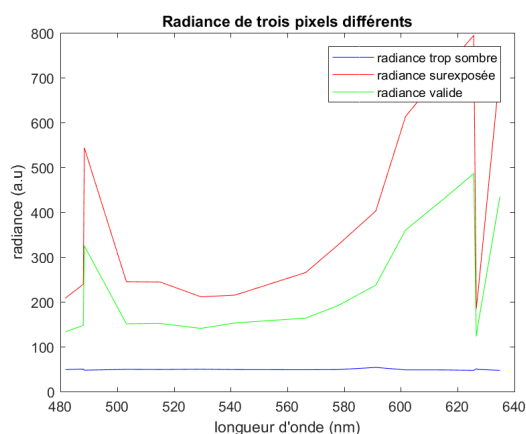
6.2/ ACQUISITIONS MULTISPECTRALES CHEZ L'HOMME

6.2.1/ COLLECTE DES DONNÉES

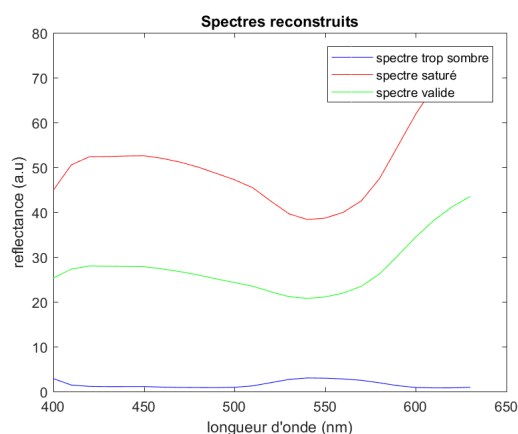
Le système bimodal que nous avons présenté fournit des images multispectrales à une cadence de 1 image par seconde environ. Les images multispectrales ont donc des temps d'acquisitions longs par rapport aux images endoscopiques (1 image par seconde contre 25-30 images par secondes). À titre d'exemple, une image multispectrale acquise à l'intérieur de l'estomac est présentée figure 6.1a. L'image est réarrangée en 16 imagerie pour montrer les 16 bandes de la caméra VIS. Nous pouvons observer le disque lumineux au milieu de l'image correspondant à la zone couverte par fibre. Toutes les images multispectrales ne contiennent pas forcément de l'information. De manière générale les images multispectrales sont souvent floues à cause des mouvements de l'endoscope, de la résolution spatiale, de la fibre optique et du fort temps d'exposition. Il est alors assez difficile d'exploiter la texture de l'image multispectrale. Certaines zones de l'image peuvent être soit trop sombres, soit sur-exposées, nous faisons donc attention à ne retenir que les radiances dont le maximum est en dessous de 512. Dans la documentation technique des caméras, il est en effet précisé que, si le 10^{ème} bit est à 1 alors, le pixel est considéré comme saturé. De la même manière, nous éliminons tous les pixels dont la moyenne dans l'axe spectral n'excède pas 128, ces pixels ont une amplitude trop faible et sont considérés comme trop sombres. Pour illustrer ce choix, nous mettons en valeur dans la figure 6.1b trois pixels : en bleu, un pixel trop sombre, en rouge un pixel saturé (dont le maximum excède 512) et en vert un pixel valide (moyenne au dessus de 128 et valeur maximale en dessous de 512). Nous appliquons ensuite à l'image la transformation décrite dans le chapitre 5 pour obtenir des spectres de réflectance (figure 6.1c).



(a) Image multispectrale acquise à l'intérieur de l'estomac.



(b) Valeurs brutes obtenues à partir des pixels d'une image multispectrale.



(c) Spectres estimés à partir d'une image multispectrale.

FIGURE 6.1 – Exemple d'extraction de spectres à partir d'une image multispectrale.

Comme les images multispectrales sont difficiles à interpréter au premier abord, nous avons donc demandé au gastroendoscopiste de sélectionner les images endoscopiques NBI dans lesquels la fibre est dirigée vers une zone significative de la pathologie. Nous pouvons ainsi sélectionner dans le même temps l'image multispectrale la plus proche temporellement pour les analyses. Cette image est filtrée par un filtre passe-bas, un filtre pyramidal de taille 7x7 pour supprimer le bruit :

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 \\ 4 & 8 & 12 & 16 & 12 & 8 & 4 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 \end{bmatrix}.$$

Après cela, nous sélectionnons des spectres disposés sur une grille régulière en prenant systématiquement

quement 1 pixel sur 7 en longueur et en largeur.

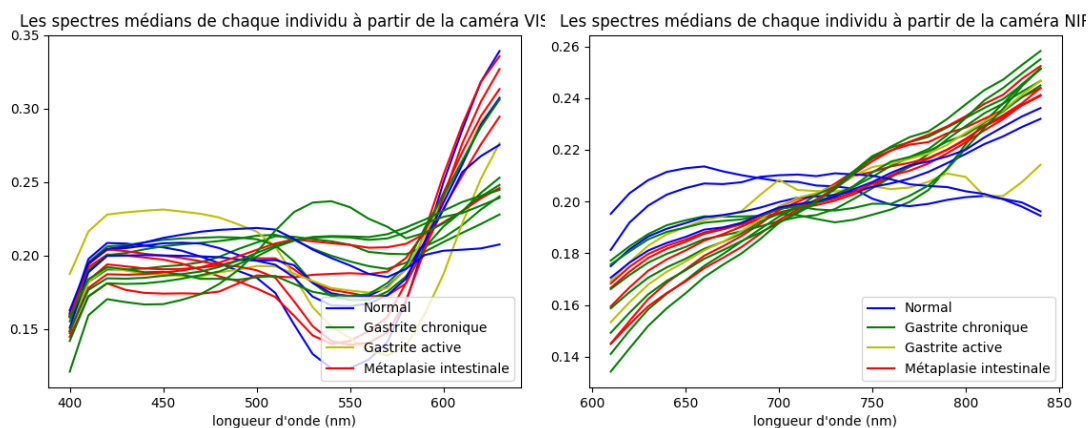
Nous n'avons pas de données multispectrales pour tous les patients pour lesquels nous avons réalisé des acquisitions NBI. Ceci est dû majoritairement aux contraintes de stérilisation de la fibre. À cause du temps de nettoyage, la fibre ne peut pas être utilisée à chaque endoscopie. Nous avons tout de même recueilli des vidéos multispectrales pour 18 d'entre eux. Ces 18 patients sont répartis dans le tableau 6.1 selon leurs résultats histologiques. Pour chacun de ces patients, nous prenons aléatoirement 20 spectres valides pour avoir le même nombre de spectres par patient.

TABLE 6.1 – Ensemble des données multispectrales obtenues selon les grades histologiques.

Normal	Gastrite Chronique	Gastrite Active	Métaplasie Dysplasie
4	7	2	5

Dans tout ce chapitre, nous séparons intentionnellement les figures et résultats provenant des deux caméras. La raison est que nous n'avons pas systématiquement deux images (une dans le visible et l'autre dans le proche infrarouge) simultanées, les deux caméras multispectrales acquièrent toutes les deux des images de manière indépendante. D'autre part, certaines images ne sont pas retenues si aucun pixel n'est valide, c'est à dire, des pixels qui sont soit saturés, soit trop sombres selon les critères énoncés plus haut.

Les deux figures 6.2 ci-après représentent les spectres médians des 18 patients après calibrage, respectivement pour la caméra VIS (figure 6.2a) et la caméra NIR (figure 6.2b). La couleur indique leur grade histologique : bleu pour la classe "Normal", vert pour la "Gastrite chronique", jaune pour la "Gastrite active" et rouge pour la "Métaplasie intestinale". S'il est difficile d'observer une tendance avec la caméra VIS, on observe que les spectres normaux sont regroupés avec la caméra NIR. Ils ont de manière générale une dérivée plus basse. Sur l'intervalle [610, 730 nm], les courbes correspondant à la gastrite chronique et à la métaplasie sont globalement en dessous des courbes de la classe "Normal" et inversement sur l'intervalle [730, 840 nm].



(a) Spectres médians provenant de la caméra VIS. (b) Spectres médians provenant de la caméra NIR.

FIGURE 6.2 – Visualisation des acquisitions spectrales avec les caméras VIS et NIR.

6.2.2/ PREMIÈRE ANALYSE DES DONNÉES

Dans un spectre les bandes voisines sont naturellement corrélées. Nous avons déjà fait cette observation dans le chapitre 4 sur le modèle animal de souris. Nous affichons ci-dessous les matrices de corrélation de manière indépendante pour les caméras VIS et NIR (figures 6.3a et 6.3b). Dans les deux images, nous observons des carrés de couleur rouge le long de la diagonale qui indiquent que certaines longueurs d'onde sont fortement corrélées, notamment les intervalles [410, 480 nm] et [520, 580 nm] dans les données de la caméra VIS et les intervalles [610, 690 nm] et [780, 850 nm] dans les données de la caméra NIR. Ces observations nous encouragent à utiliser à nouveau l'algorithme (1), déjà utilisé dans le chapitre 4 pour grouper les bandes spectrales qui sont corrélées. Nous utilisons le seuil $\tau = 0,9$, pour ne fusionner que les bandes voisines qui sont corrélées à plus de 90%.

Les figures 6.3c et 6.3d montrent les résultats de la fusion des bandes. Elles présentent à chaque fois un spectre médian et des lignes verticales noires qui permettent de visualiser les frontières entre les groupes de longueurs d'onde sélectionnées. Nous obtenons dans le cas VIS, 9 bandes de longueurs d'onde et dans le cas NIR, 7 bandes de longueurs d'onde.

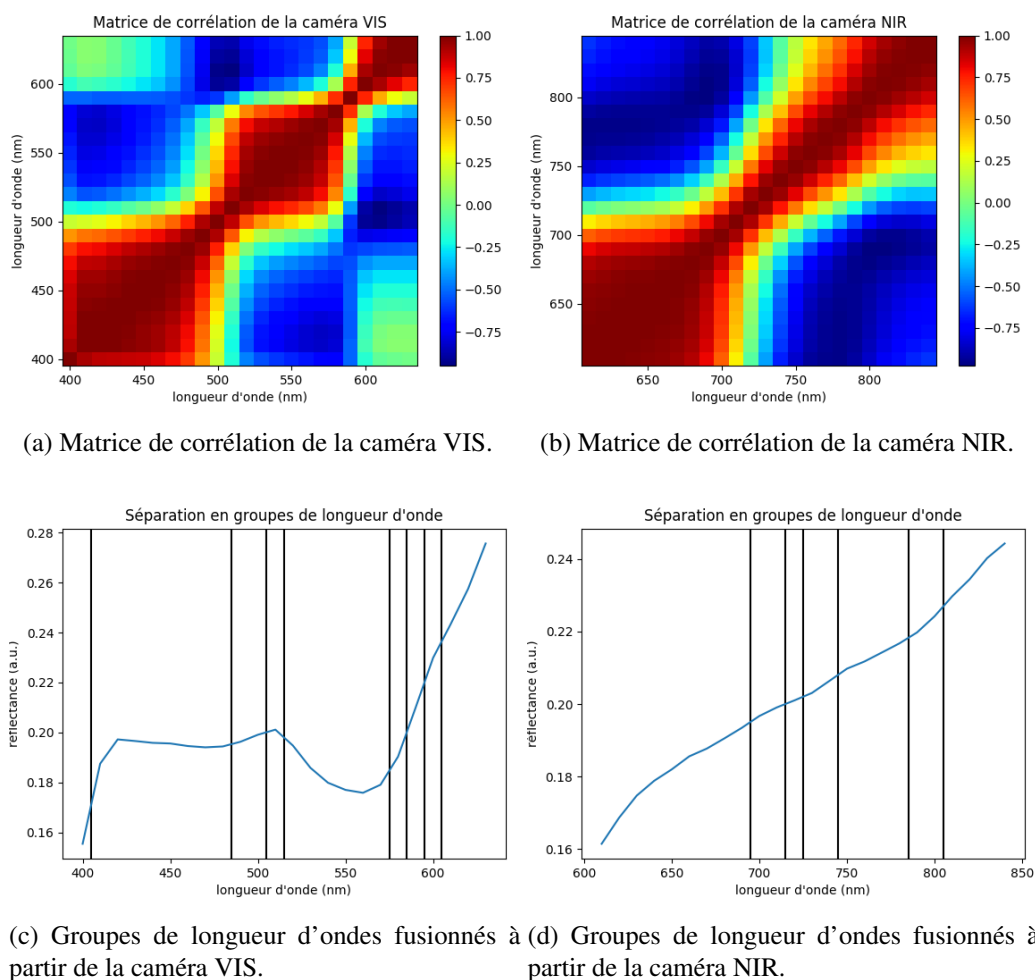


FIGURE 6.3 – Corrélation entre bandes de longueurs d'onde.

6.3/ CLASSIFICATION DES PATHOLOGIES DE L'ESTOMAC

Dans un souci d'utiliser une logique identique à celle employée chez le petit animal dans le chapitre 4, nous testons à nouveau trois types de descripteurs : soit nous utilisons les bandes réduites telles-elles, soit nous calculons des descripteurs de soustraction ou de division. Dans les deux derniers cas, il s'agit de calculer toutes les soustractions, respectivement, toutes les divisions possibles à partir des bandes réduites.

Pour résoudre les éventuels problèmes de sur-apprentissage, nous réduisons le nombre de descripteurs, en utilisant l'analyse de variance (ANOVA). Les descripteurs sélectionnés permettent alors d'entraîner un classifieur de type SVM avec noyau linéaire.

Pour ne pas introduire de biais dans la méthode, nous effectuons une validation croisée : le pipeline décrit précédemment (sélection univariée et entraînement d'un classifieur) est réalisé autant de fois que le nombre de patients. À chacune des itérations, les données d'un patient sont mises de côté et ne sont pas utilisées pour l'apprentissage mais gardées pour tester le classifieur.

Dans les tableaux 6.2 et 6.3, nous présentons les résultats de classification obtenus avec les caméras VIS et NIR. De manière générale, les résultats sont meilleurs avec la caméra NIR qu'avec la caméra VIS. Nous nous attardons donc plus sur ceux de la caméra NIR. Nous observons une grande différence entre les classes. Les classes "Gastrite Chronique" et "Métaplasie intestinale" sont bien reconnues avec une précision et un rappel de 99% ou 100% mais la gastrite active et la classe "Normal" sont souvent confondues. Ceci peut être dû au nombre trop peu élevé de patient chez lesquels nous avons observé de la gastrite active (2 patients). Nos données ne sont pas balancées et cela impacte les résultats de classification même si le classifieur donne des poids plus importants aux classes minoritaires (traditionnellement, inversement proportionnels aux fréquences des classes). Nous donnons également la valeur des hyperparamètres optimaux pour chaque cas, à savoir k , le nombre de descripteurs retenus lors de la sélection univariée et c , le paramètre qui permet d'avoir une marge souple avec les SVM.

TABLE 6.2 – Résultat de classification des spectres avec la caméra VIS.

	Hyperparamètres optimaux	Classe	Précision	Rappel	Score F1
Bandes Réduites	$c = 0,001$ $k = 7$	Normal	0,38	0,94	0,54
		Gastrite Active	0,00	0,00	0,00
		Gastrite Chronique	1,00	0,71	0,83
		Métaplasie Intestinale	1,00	0,49	0,66
Divisions	$c = 0,001$ $k = 1$	Normal	0,36	0,86	0,51
		Gastrite Active	0,95	0,47	0,63
		Gastrite Chronique	0,89	0,62	0,73
		Métaplasie Intestinale	1,00	0,49	0,66
Soustractions	$c = 0,001$ $k = 5$	Normal	0,33	0,89	0,48
		Gastrite Active	0,61	0,47	0,54
		Gastrite Chronique	0,93	0,50	0,65
		Métaplasie Intestinale	1,00	0,38	0,55

Nous pouvons nous intéresser plus en détail aux prédictions pour tous les patients, par exemple avec les descripteurs de divisions. Ces prédictions sont visibles sur la figure 6.4. Pour chaque patient, nous donnons la répartition des prédictions des 20 spectres. Le label de chaque patient est indiqué par la couleur du texte (bleu pour normal, vert pour gastrite chronique, jaune pour gastrite

TABLE 6.3 – Résultat de classification des spectres avec la caméra NIR.

	Hyperparamètres optimaux	Classe	Précision	Rappel	Score F1
Bandes Réduites	c = 0,001 k = 1	Normal	0,54	0,47	0,51
		Gastrite Active	0,17	0,23	0,20
		Gastrite Chronique	1,00	0,99	1,00
		Métaplasie Intestinale	1,0	0,99	0,99
Divisions	c = 0,001 k = 7	Normal	0,77	0,45	0,57
		Gastrite Active	0,65	0,28	0,39
		Gastrite Chronique	1,00	1,00	1,00
		Métaplasie Intestinale	0,63	0,98	0,77
Soustractions	c = 0,001 k = 3	Normal	0,84	0,46	0,60
		Gastrite Active	0,45	0,35	0,39
		Gastrite Chronique	0,85	1,00	0,92
		Métaplasie Intestinale	0,82	0,99	0,90

active et rouge pour métaplasie). Nous observons que les patients 9 et 17 sont mal classés, ces deux patients ont été diagnostiqués "Normaux" mais la majorité de leurs spectres ont été classés comme métaplasie. Nous notons également que le patient 10 a été mal classé, ceci dû probablement au fait que la classe "Gastrite active" est sous-représentée. Néanmoins, la plupart des patients montrent des résultats cohérents avec l'histologie. En considérant la classe la plus fréquemment prédite comme label pour chaque patient, 15 d'entre eux obtiennent le bon label contre 3 qui ne l'obtiennent pas (les patients 9, 10 et 17).

Au vue des premiers résultats du tableau 6.4, nous avons décidé de tester le pipeline de classification en enlevant les patients classés dans la classe "Gastrite active". Dans le tableau 6.4 se trouvent les résultats obtenus sans ces deux patients en utilisant les descripteurs de soustraction, qui donnent les meilleurs résultats. Le classifieur ne commet alors quasiment aucune erreur et, de manière intéressante, un seul descripteur est utile, il s'agit de la différence entre la bande [730, 740 nm] et la bande [810, 850 nm]. Dans le cadre des travaux sur le modèle animal de souris, nous avons déjà rencontré un cas similaire où seulement deux intervalles étaient utilisés, les longueurs d'onde pertinentes étaient également dans le proche infrarouge.

TABLE 6.4 – Résultat de classification des spectres après suppression des patients atteints de gastrite active.

	Hyperparamètres optimaux	Classe	Précision	Rappel	Score F1
Bandes Réduites	c = 0,001 k = 1	Normal	0,99	1,00	0,99
		Gastrite Chronique	1,00	1,00	1,00
		Métaplasie Intestinale	1,00	0,99	0,99
Divisions	c = 0,001 k = 1	Normal	0,99	0,82	0,9
		Gastrite Chronique	0,92	1,00	0,96
		Métaplasie Intestinale	0,99	0,99	0,99
Soustractions	c = 0,001 k = 1	Normal	0,99	1,00	0,99
		Gastrite Chronique	1,00	1,00	1,00
		Métaplasie Intestinale	1,00	0,99	0,99

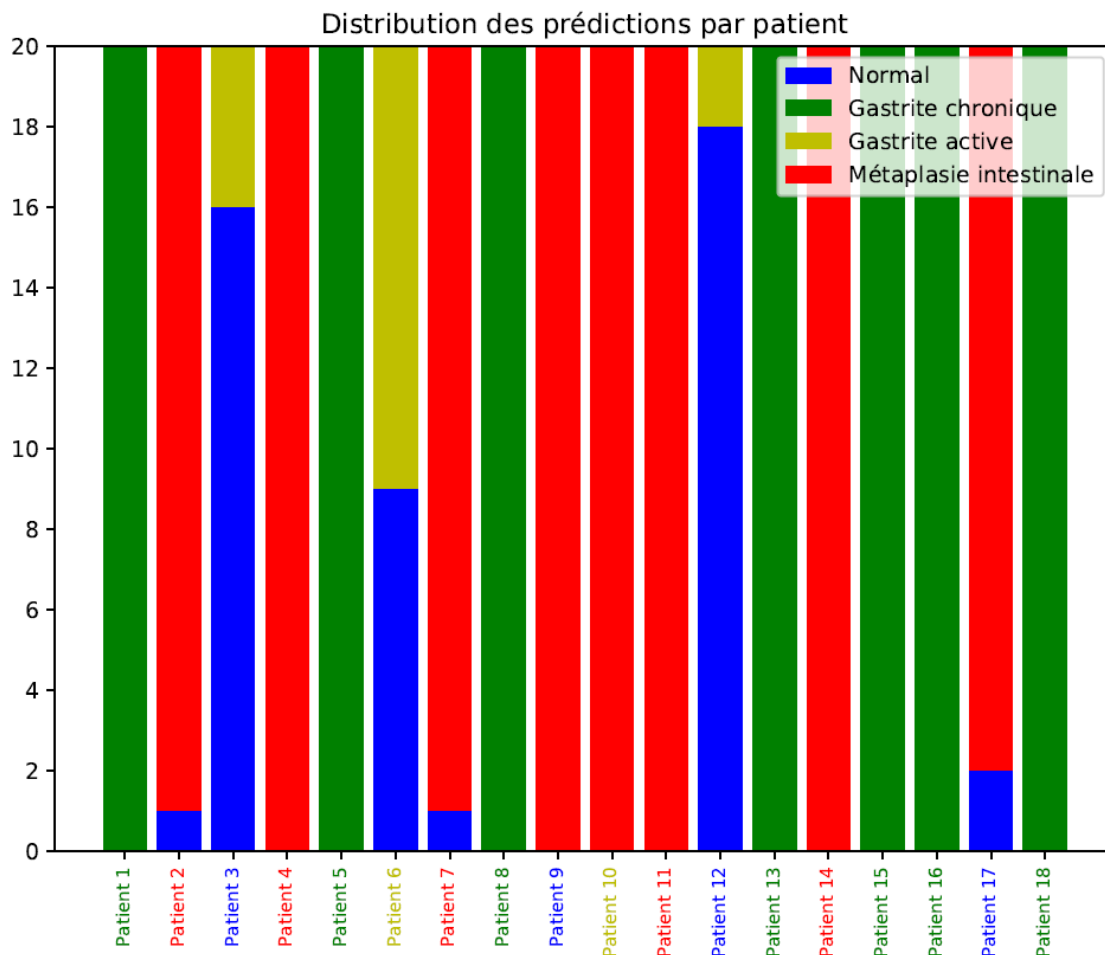


FIGURE 6.4 – Distribution des prédictions par patient.

6.4/ CONCLUSION ET PERSPECTIVES

Dans ce chapitre, nous utilisons les acquisitions spectrales de l'hôpital Ambroise Paré pour caractériser différents stades pré-cancéreux. Nous nous sommes basés sur la même méthodologie déjà utilisée dans l'analyse spectrale chez le petit animal au chapitre 4 pour classer des spectres de 3 pathologies différentes (gastrite active, gastrite chronique et métaplasie intestinale) par rapport à des spectres normaux. Nous manquons malheureusement de données en quantité suffisante concernant la gastrite active mais nous sommes arrivés à classifier les spectres des trois autres catégories. Les longueurs d'onde qui nous ont été utiles se trouvent dans le proche infrarouge, plus précisément les intervalles [730, 740 nm] et [810, 850 nm].

Pour assurer et confirmer ces résultats, il est important de multiplier le nombre d'acquisitions. Pour le moment nous sommes également limités par la pauvre résolution spatiale obtenue avec les caméras multispectrales. Les acquisitions nécessitent actuellement près d'une seconde de temps d'exposition ce qui est trop élevé. Durant ce laps de temps, la scène bouge grandement ce qui induit fatalement un flou de mouvement important.

CONCLUSION ET PERSPECTIVES GLOBALES

7.1/ APPORTS SCIENTIFIQUES

Dans ce manuscrit ont été présentés les travaux de thèse sur la détection de lésions pré-cancéreuses par traitement d'images endoscopiques NBI et imagerie multispectrale. La thèse s'inscrit dans le cadre du projet ANR EMMIE. Le but de ce projet est de proposer un prototype d'acquisition multimodal permettant l'exploration de l'estomac et plus précisément de l'antrum pylorique. Le système proposé doit servir à détecter des lésions pré-cancéreuses et ainsi alerter et guider les endoscopistes dans leur travail quotidien. Ce projet regroupe plusieurs partenaires ayant des expertises différentes et complémentaires dans les domaines de la biologie, de la médecine et du traitement d'images.

Dans ce projet, nous ne nous sommes pas intéressés directement au cancer mais aux lésions pré-cancéreuses. En particulier l'ensemble des pathologies considérées comprend la gastrite active, la gastrite chronique, la métaplasie intestinale et l'atrophie glandulaire. À notre connaissance, il n'y a à ce jour, aucun travail similaire qui considère cet ensemble de pathologies pour les distinguer.

Nous avons répondu aux objectifs du projet en proposant un prototype permettant d'acquérir des images NBI et des images multispectrales. Le prototype est constitué d'une colonne endoscopique déjà présente habituellement dans les blocs opératoires à laquelle sont ajoutés des caméras multispectrales, un système optique et un ordinateur pour collecter les données. Pour les acquisitions spectrales, un fibroscope est insérée dans le canal opérateur de l'endoscope. Ce prototype a été validé par le Comité de Protection des Personnes (CPP).

Ce prototype doit être calibré. Dans le cadre du manuscrit, nous sommes intervenus sur le calibrage spectral. Le calibrage est réalisé en acquérant les patches d'un ColorChecker® avec différentes modalités : avec un spectromètre, avec les caméras multispectrales et avec l'endoscope en lumière blanche et NBI. Ces acquisitions permettent de déterminer les coefficients d'une transformation linéaire modélisant le calibrage. Nous avons montré que le principe pouvait être étendu à l'estimation d'images endoscopiques sous lumière NBI ou en lumière blanche. Le système proposé peut être vu comme une alternative aux systèmes NBI et FICE dans le sens où il contient plus de bandes que le NBI et permet d'estimer des longueurs d'onde tout comme FICE à la différence que ce système les estime uniquement à partir de trois bandes.

Avec ce système, nous avons pu collecter une grande quantité de données. Les données d'une cinquantaine de patients sont exploitées dans l'analyse des images NBI et 18 le sont dans l'analyse multispectrale. Les données ainsi acquises ont été analysées de manière indépendante. Pour les images, nous avons sélectionné avec un expert gastroentérologue des patches significatifs des pathologies citées plus haut. Des réseaux de neurones pour l'apprentissage par transfert ont été utilisés pour classer ces patches. Les résultats sont prometteurs, en particulier, pour la détection

de l'atrophie avec une précision et un rappel d'environ 72%. Pour améliorer les résultats, nous avons mis en place des descripteurs visuels utilisés en pratique par les endoscopistes : il s'agit d'évaluer, en donnant une note entre 0 et 3, la couleur de la muqueuse, la régularité et l'épaisseur de la trame, la présence d'une texture particulière à l'intérieur de la trame et la densité vasculaire. Ces descripteurs n'ont, à notre connaissance, jamais été explicités dans la littérature. Ils ont pourtant montré de meilleurs résultats pour la reconnaissance des pathologies en comparaison aux méthodes d'apprentissage par transfert (67% contre 50% environ), démontrant ainsi qu'ils sont significatifs.

D'un point de vue théorique, nous avons proposé deux solutions innovantes à l'inversion d'un modèle mathématique, le modèle dichromatique. Ce modèle permet de séparer une image RGB ou multispectrale en ses composantes intrinsèques, notamment la composante diffuse et la composante spéculaire. Cette décomposition peut-être vue comme un pré-traitement des images endoscopiques. La méthodologie utilisée se veut originale dans le sens où elle est adaptable aux images couleur et aux images multispectrales, ne nécessite pas d'apprentissage et se distingue donc des méthodes utilisant l'apprentissage profond. La robustesse des méthodes a été démontrée en appliquant nos algorithmes sur trois bases de données différentes : une base d'images couleur provenant du MIT et deux bases d'images multispectrales, respectivement CAVE et SIID.

Dans le cadre de l'analyse des spectres chez la souris et chez l'Homme. Nous avons développé un pipeline de classification robuste qui a montré des résultats intéressants pour la détection de l'inflammation chez la souris et la détection de la gastrite chronique et de la métaplasie chez l'Homme. Le pipeline est composé d'un algorithme de fusion de bandes fortement corrélées entre elles, d'un calcul de descripteurs de soustraction ou de division, d'une étape de sélection univariée de descripteurs et enfin d'une classification avec validation croisée par patient. L'analyse spectrale des données a également mis en lumière l'intérêt d'utiliser des caméras dans le proche infrarouge. Chez la souris nous avons identifié les intervalles [620, 668 nm] et [668, 950 nm] comme étant discriminant pour détecter l'inflammation. Chez l'Homme, nous avons mis en avant les intervalles [730, 740 nm] et [810, 850 nm] pour discriminer la muqueuse saine d'une muqueuse avec gastrite chronique ou avec métaplasie. Tous ces intervalles se trouvent majoritairement dans le proche infrarouge. Ces résultats doivent être confirmés à plus grande échelle pour pouvoir être utilisés au quotidien par les gastroentérologues. Tout d'abord, en développant un système plus robuste puis en installant des systèmes identiques dans plusieurs hôpitaux qui seraient manipulés par des médecins indépendants.

7.2/ PERSPECTIVES DU PROJET

Nous avons développé un algorithme complet permettant une bonne inversion du modèle dichromatique. La méthodologie basée sur l'optimisation quadratique ou la déconvolution ouvre de nombreuses portes. En particulier l'approche par deux déconvolutions indépendantes ouvre la voie à des implémentations très rapides en utilisant par exemple la puissance des GPUs actuels (Graphic Processing Unit). Les solutions proposées peuvent être améliorées, notamment en ajoutant d'autres contraintes sous forme de termes de régularisation connus dans le domaine telles la minimisation de la norme l_1 ou l_2 . Dans un autre champ d'étude, nous pourrions utiliser plusieurs images d'un même objet. Les contraintes de lissage proposées pourraient alors être remplacées par des contraintes de similarité entre les pixels pointant vers le même point de l'espace. Ce dernier point serait particulièrement adapté à la décomposition en images intrinsèques appliquée à des flux vidéo comme dans le cadre des enregistrements en endoscopie. Il est à noter que dans les images NBI, un pré-traitement supplémentaire est nécessaire à cause de la forte corrélation entre deux des

trois canaux disponibles.

Nous avons vu que la reconnaissance de pathologies par imagerie NBI était difficile. Il y a plusieurs raisons à cela : nous pouvons encore largement augmenter la base de données pour obtenir un ensemble de patches plus conséquent, plus varié et donc plus représentatif des cas auxquels les endoscopistes peuvent être confrontés. Nous avons cependant une base déjà conséquente qui pourrait être valorisée en étant rendue publique avec l'accord des patients. L'association des deux modalités n'a malheureusement pas pu être exploitée dans ce manuscrit. Il serait d'autant plus intéressant de pouvoir lier les deux caméras VIS et NIR et de lier les acquisitions spectrales aux acquisitions NBI. D'autre part, l'histologie que nous utilisons comme vérité terrain pour les patients peut être faussée à cause du fait que les prélèvements soient non-orientés, ce qui explique en partie les résultats de classification des patches NBI. Il faudrait donc orienter la recherche de manière à avoir une corrélation directe entre les images endoscopiques et les sites des prélèvements pour les biopsies. Nous pourrions alors avoir un diagnostic par patch et non plus par patient, ce qui serait plus juste d'un point de vue médical. Nous pourrions alors annoter ces nouveaux patches selon les critères visuels que nous avons défini et, grâce à des outils d'apprentissage automatique, construire un programme capable d'annoter automatiquement les patches endoscopiques.

Le prototype d'acquisition multispectral peut lui aussi être amélioré. Il n'est malheureusement pas optimal pour l'exploration dans l'estomac dans la mesure où nous sommes obligés de travailler avec un système non calibré et nous sommes fortement influencés par les flous de mouvement qui empêchent l'exploitation d'une éventuelle texture dans les images multispectrales. Les futures solutions devraient donc contenir des caméras plus performantes en termes de vitesse d'acquisition et focus automatique. L'imagerie multispectrale a tout de même apporté des preuves convaincantes dans le domaine médical pour la détection précoce des lésions de l'estomac et les recherches doivent continuer dans ce sens pour que cette modalité puisse être utilisée en pratique par les endoscopistes. L'analyse spectrale pourrait être encore plus intéressante si l'on pouvait également remonter à des paramètres biologiques, c'est-à-dire aux quantités relatives d'hémoglobine, d'eau et potentiellement à d'autres composants non identifiés. D'un point de vue numérique, ceci peut être fait par des méthodes de séparation de sources.

ARTICLES DE CONFÉRENCES ET ARTICLES DE JOURNAUX ASSOCIÉS AUX TRAVAUX DE THÈSE

Ces différents travaux nous ont permis de publier un résumé, trois articles de conférence, un chapitre de livre et de soumettre deux articles journaux et deux résumés. Un article de conférence [F] et deux articles journaux [A] et [B] ont permis de valoriser les travaux sur le modèle dichromatique. L'article de conférence [F] a été élu meilleur papier de la conférence (CiSRA Best Paper Award). Ces publications sont propres au laboratoire ImViA. Les autres conférences impliquent d'autres membres du projet EMMIE et sont les résultats de collaboration avec l'institut Pasteur sur les travaux sur le petit animal [E] et le résumé [G] et avec l'équipe du laboratoire PRISME sur les travaux de calibrage du prototype d'acquisition [D]. Ce dernier article de conférence a été sélectionné pour qu'une version étendue apparaisse dans un chapitre du livre "Communications in Computer and Information Science, volume 997" [C]. Dans le cadre des acquisitions NBI et multispectrales chez l'Homme, deux résumés ont été soumis aux "Journées Francophones d'Hépatogastroentérologie et d'Oncologie Digestive" (JFHOD) [H] et [I].

ARTICLES DE JOURNAUX SOUMIS

- [A] **Krebs, Alexandre** and Benezeth, Yannick and Marzani, Franck, Intrinsic RGB and multispectral images recovery by independent quadratic programming, *PeerJ Computer Science*, 2019.
- [B] **Krebs, Alexandre** and Benezeth, Yannick and Marzani, Franck, Intrinsic image decomposition as two independent deconvolution problems, *Signal Processing : Image Communication*, 2019.

CHAPITRE DE LIVRE

- [C] Zenteno, Omar and **Krebs, Alexandre** and Treuillet, Sylvie and Lucas, Yves and Benezeth, Yannick and Marzani, Franck, Spatial and Spectral Calibration of a Multispectral-Augmented Endoscopic Prototype, *International Joint Conference on Computer Vision, Imaging and Computer Graphics*, 2018.

ARTICLES DE CONFÉRENCE

- [D] Zenteno, Omar and **Krebs, Alexandre** and Treuillet, Sylvie and Lucas, Yves and Benezeth, Yannick and Marzani, Franck, Dual-channel Geometric Registration of a Multispectral-augmented Endoscopic Prototype, *VISIGRAPP*, 2018.
- [E] **Krebs, Alexandre** and Camilo, Vania and Touati, Eliette and Benezeth, Yannick and Michel, Valérie and Jouvion, Grégory and Yang, Fan and Lamarque, Dominique and Marzani,

Franck, Detection of H. pylori induced gastric inflammation by diffuse reflectance analysis, *IEEE International Conference on BioInformatics and BioEngineering (BIBE)*, 2018.

- [F] **Krebs, Alexandre** and Benezeth, Yannick and Marzani, Franck, Quadratic Objective Functions for Dichromatic Model Parameters Estimation, *IEEE International Conference on Digital Image Computing : Techniques and Applications (DICTA)*, 2017.

RÉSUMÉS

- [G] Camilo, Vania and **Krebs, Alexandre** and Jobart-Malfait, Aude and Michel, Valérie and Benezeth, Yannick and Bazin, Thomas and Marzani, Franck and Touati, Eliette and Lamarque, Dominique, Détection des lésions inflammatoires de la muqueuse gastrique par analyse multispectrale de la réflectance dans un modèle murin d'infection à H. pylori, *Journées Francophones d'Hépatogastroentérologie et d'Oncologie Digestive*, 2019. (Accepté)
- [H] Bazin, Thomas and **Krebs, Alexandre** and Benezeth, Yannick and Moussata, Driffa and Julié, C. and Emile, J.-F. and Marzani, Franck and Lamarque, Dominique, Détection automatisée de l'atrophie glandulaire et de la métaplasie intestinale gastriques, *Journées Francophones d'Hépatogastroentérologie et d'Oncologie Digestive*, 2019. (Soumis, en cours de relecture)
- [I] Bazin, Thomas and **Krebs, Alexandre** and Benezeth, Yannick and Julié, C. and Emile, J.-F. and Marzani, Franck and Lamarque, Dominique, Prototype de détection de la gastrite chronique et de la métaplasie intestinale par imagerie multispectrale, *Journées Francophones d'Hépatogastroentérologie et d'Oncologie Digestive*, 2019. (Soumis, en cours de relecture)

BIBLIOGRAPHIE

- [1] **direct-intrinsics**. Accessed : 2019-09-23.
- [2] **Fice fuji intelligent chromo endoscopy**. Accessed : 2019-05-27.
- [3] **High power mercury-xenon light sources**. Accessed : 2019-05-27.
- [4] **Itconcepts microflex**. Accessed : 2019-05-27.
- [5] **Jon barron**. Accessed : 2019-09-23.
- [6] **Narrow band imaging (nbi)**. Accessed : 2019-06-18.
- [7] **Scyllarus**. Accessed : 2019-09-23.
- [8] **Twincam**. Accessed : 2019-05-25.
- [9] **Usb3 vision compliant cameras - xiq**. Accessed : 2019-05-27.
- [10] WILCOXON, F. **Individual comparisons by ranking methods**. *Biometrics bulletin* (1945).
- [11] MASSY, W. F. **Principal components regression in exploratory statistical research**. *Journal of the American Statistical Association* (1965).
- [12] HARALICK, R. M., SHANMUGAM, K., AND OTHERS. **Textural features for image classification**. *Transactions on systems, man, and cybernetics* (1973).
- [13] BUCHSBAUM, G. **A spatial processor model for object colour perception**. *Journal of the Franklin institute* (1980).
- [14] SHAFER, S. A. **Using color to separate reflection components**. *Color Research & Application* (1985).
- [15] FORSYTH, D. A. **A novel algorithm for color constancy**. *International Journal of Computer Vision* (1990).
- [16] PARSONNET, J., FRIEDMAN, G. D., VANDERSTEEN, D. P., CHANG, Y., VOGELMAN, J. H., ORENTREICH, N., AND SIBLEY, R. K. **Helicobacter pylori infection and the risk of gastric carcinoma**. *New England Journal of Medicine* (1991).
- [17] CORREA, P. **Human gastric carcinogenesis : a multistep and multifactorial process—first american cancer society award lecture on cancer epidemiology and prevention**. *Cancer research* (1992).
- [18] BARALDI, A., AND PARMIGGIANI, F. **An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters**. *Transactions on Geoscience and Remote Sensing* (1995).
- [19] CORTES, C., AND VAPNIK, V. **Support-vector networks**. *Machine learning* (1995).
- [20] EATON, K., RADIN, M., AND KRAKOWKA, S. **An animal model of gastric ulcer due to bacterial gastritis in mice**. *Veterinary pathology* (1995).
- [21] DIXON, M. F., GENTA, R. M., YARDLEY, J. H., CORREA, P., AND OTHERS. **Classification and grading of gastritis : the updated sydney system**. *The American journal of surgical pathology* (1996).

- [22] OJALA, T., PIETIKÄINEN, M., AND HARWOOD, D. **A comparative study of texture measures with classification based on featured distributions.** *Pattern Recognition* (1996).
- [23] LEE, A., O'ROURKE, J., DE UNGRIA, M. C., ROBERTSON, B., DASKALOPOULOS, G., AND DIXON, M. F. **A standardized mouse model of helicobacter pylori infection : introducing the sydney strain.** *Gastroenterology* (1997).
- [24] FERRERO, R. L., THIBERGE, J.-M., HUERRE, M., AND LABIGNE, A. **Immune responses of specific-pathogen-free mice to chronic helicobacter pylori (strain ss1) infection.** *Infection and immunity* (1998).
- [25] CARDEI, V. C., FUNT, B., AND BARNARD, K. **White point estimation for uncalibrated images.** In *Color and Imaging Conference* (1999), Society for Imaging Science and Technology.
- [26] KOBAYASHI, M., TAJIRI, H., SEIKE, E., SHITAYA, M., TOUNOU, S., MINE, M., AND OBA, K. **Detection of early gastric cancer by a real-time autofluorescence imaging system.** *Cancer letters* (2001).
- [27] BRUNO, M. **Magnification endoscopy, high resolution endoscopy, and chromoscopy; towards a better optical diagnosis.** *Gut* (2003).
- [28] TOUATI, E., MICHEL, V., THIBERGE, J.-M., WUSCHER, N., HUERRE, M., AND LABIGNE, A. **Chronic helicobacter pylori infections induce gastric mutations in mice.** *Gastroenterology* (2003).
- [29] DALAL, N., AND TRIGGS, B. **Histograms of oriented gradients for human detection.** In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2005), IEEE.
- [30] ROGERS, A. B., TAYLOR, N. S., WHARY, M. T., STEFANICH, E. D., WANG, T. C., AND FOX, J. G. **Helicobacter pylori but not high salt induces gastric intraepithelial neoplasia in b6129 mice.** *Cancer research* (2005).
- [31] IAKOVIDIS, D. K., MAROULIS, D. E., AND KARKANIS, S. A. **An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy.** *Computers in Biology and Medicine* (2006).
- [32] MARTIN, M. E., WABUYELE, M. B., CHEN, K., KASILI, P., PANJEHPOUR, M., PHAN, M., OVERHOLT, B., CUNNINGHAM, G., WILSON, D., DENOVO, R. C., AND OTHERS. **Development of an advanced hyperspectral imaging (hsi) system with applications for cancer detection.** *Annals of biomedical engineering* (2006).
- [33] NOCEDAL, J., AND WRIGHT, S. J. **Numerical Optimization, second edition.** World Scientific, 2006.
- [34] UEDO, N., ISHIHARA, R., IISHI, H., YAMAMOTO, S., YAMADA, T., IMANAKA, K., TAKEUCHI, Y., HIGASHINO, K., ISHIGURO, S., AND TATSUTA, M. **A new method of diagnosing gastric intestinal metaplasia : narrow-band imaging with magnifying endoscopy.** *Endoscopy* (2006).
- [35] BURGOS, H., PORRAS, M., BRENES, F., AND IZQUIERDO, E. **Fujinon fice electronic chromovideoendoscopy helps differentiate the type of metaplasia in patients with chronic atrophic gastritis.** *Gastrointestinal Endoscopy* (2007).
- [36] ABELOFF, M. D., ARMITAGE, J. O., NIEDERHUBER, J. E., KASTAN, M. B., AND MCKENNA, W. G. **Abeloff's Clinical Oncology E-Book.** Elsevier Health Sciences, 2008.
- [37] BANSAL, A., ULUSARAC, O., MATHUR, S., AND SHARMA, P. **Correlation between narrow band imaging and nonneoplastic gastric pathology : a pilot feasibility trial.** *Gastrointestinal Endoscopy* (2008).
- [38] BASHAR, M. K., MORI, K., SUENAGA, Y., KITASAKA, T., AND MEKADA, Y. **Detecting informative frames from wireless capsule endoscopic video using color and texture features.** In

International Conference on Medical Image Computing and Computer-Assisted Intervention (2008), Springer.

- [39] MAKHLOUF, H., GMITRO, A. F., TANBAKUCHI, A. A., UDOVICH, J. A., AND ROUSE, A. R. **Multispectral confocal microendoscope for in vivo and in situ imaging.** *Journal of biomedical optics* (2008).
- [40] REDDYMASU, S. C., AND SHARMA, P. **Advances in endoscopic imaging of the esophagus.** *Gastroenterology Clinics* (2008).
- [41] SONG, L. M. W. K., ADLER, D. G., CONWAY, J. D., DIEHL, D. L., FARRAYE, F. A., KANTSEVOY, S. V., KWON, R., MAMULA, P., RODRIGUEZ, B., SHAH, R. J., AND OTHERS. **Narrow band imaging and multiband imaging.** *Gastrointestinal Endoscopy* (2008).
- [42] DENG, J., DONG, W., SOCHER, R., LI, L.-J., LI, K., AND FEI-FEI, L. **Imagenet : A large-scale hierarchical image database.** In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2009), IEEE.
- [43] GROSSE, R., JOHNSON, M. K., ADELSON, E. H., AND FREEMAN, W. T. **Ground truth dataset and baseline evaluations for intrinsic image algorithms.** In *International Conference on Computer Vision* (2009), IEEE.
- [44] HERNANDO-HARDER, A. C., BOOKEN, N., GOERDT, S., SINGER, M. V., AND HARDER, H. **Helicobacter pylori infection and dermatologic diseases.** *European Journal of Dermatology* (2009).
- [45] LI, B., AND MENG, M. Q.-H. **Texture analysis for ulcer detection in capsule endoscopy images.** *Image and Vision computing* (2009).
- [46] PAN, S. J., AND YANG, Q. **A survey on transfer learning.** *Transactions on Knowledge and Data Engineering* (2009).
- [47] TREVOR, H., ROBERT, T., AND JH, F. **The elements of statistical learning : data mining, inference, and prediction,** 2009.
- [48] WU, Y.-L., TU, S.-X., CHENG, S.-D., AND ZHANG, S. **Computer aided analysis system of gastric cancer based on endoscopic image database.** *Computer Engineering and Design* (2009).
- [49] BERGHOLT, M. S., ZHENG, W., LIN, K., HO, K. Y., TEH, M., YEOH, K. G., SO, J. B. Y., AND HUANG, Z. **Raman endoscopy for in vivo differentiation between benign and malignant ulcers in the stomach.** *Analyst* (2010).
- [50] GROUP, I. W., AND OTHERS. **Ingested nitrate and nitrite, and cyanobacterial peptide toxins.** *IARC monographs on the evaluation of carcinogenic risks to humans* (2010).
- [51] HUYNH, C. P., AND ROBLES-KELLY, A. **A solution of the dichromatic model for multispectral photometric invariance.** *International Journal of Computer Vision* (2010).
- [52] JOLIVOT, R., MARZANI, F., AND VABRES, P. **An open system to generate hyperspectral cubes for skin optical reflectance analysis.** *Skin Research and Technology* (2010).
- [53] LI, B., AND MENG, M. Q.-H. **Capsule endoscopy images classification by color texture and support vector machine.** In *International Conference on Automation and Logistics* (2010), IEEE.
- [54] MITRA, J., JOLIVOT, R., MARZANI, F., AND VABRES, P. **Blind source separation of skin chromophores on a hyperspectral cube.** *Skin Research and Technology* (2010).
- [55] TOLA, E., LEPETIT, V., AND FUA, P. **Daisy : An efficient dense descriptor applied to wide-baseline stereo.** *Transactions on Pattern Analysis and Machine Intelligence* (2010).

- [56] YASUMA, F., MITSUNAGA, T., ISO, D., AND NAYAR, S. K. **Generalized assorted pixel camera : postcapture control of resolution, dynamic range, and spectrum.** *Transactions on Image Processing* (2010).
- [57] AKBARI, H., UTO, K., KOSUGI, Y., KOJIMA, K., AND TANAKA, N. **Cancer detection using infrared hyperspectral imaging.** *Cancer science* (2011).
- [58] FENG, S., CHEN, R., LIN, J., PAN, J., WU, Y., LI, Y., CHEN, J., AND ZENG, H. **Gastric cancer detection based on blood plasma surface-enhanced raman spectroscopy excited by polarized laser light.** *Biosensors and Bioelectronics* (2011).
- [59] GROSBERG, L. E., RADOSEVICH, A. J., ASFAHA, S., WANG, T. C., AND HILLMAN, E. M. **Spectral characterization and unmixing of intrinsic contrast in intact normal and diseased gastric tissues using hyperspectral two-photon microscopy.** *PLoS one* (2011).
- [60] JOLIVOT, R., VABRES, P., AND MARZANI, F. **Reconstruction of hyperspectral cutaneous data from an artificial neural network-based multispectral imaging system.** *Computerized medical imaging and graphics* (2011).
- [61] LI, J. **Agglomerative connectivity constrained clustering for image segmentation.** *Statistical Analysis and Data Mining : The ASA Data Science Journal* (2011).
- [62] NEUHAUSER, M. **Nonparametric statistical tests : A computational approach.** Chapman and Hall/CRC, 2011.
- [63] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., AND DUCHESNAY, E. **Scikit-learn : Machine learning in Python.** *Journal of Machine Learning Research* (2011).
- [64] AKBARI, H., HALIG, L. V., ZHANG, H., WANG, D., CHEN, Z. G., AND FEI, B. **Detection of cancer metastasis using a novel macroscopic hyperspectral method.** In *Medical Imaging : Biomedical Applications in Molecular, Structural, and Functional Imaging* (2012), International Society for Optics and Photonics.
- [65] BUTLER, D. J., WULFF, J., STANLEY, G. B., AND BLACK, M. J. **A naturalistic open source movie for optical flow evaluation.** In *European Conference on Computer Vision (ECCV)* (2012), A. Fitzgibbon et al. (Eds.), Ed., Part IV, LNCS 7577, Springer.
- [66] GALAL, A., HASAN, H., AND IMAM, I. **Learnable hyperspectral measures.** *Egyptian Informatics Journal* (2012).
- [67] GALEANO, J., JOLIVOT, R., AND MARZANI, F. **Quantification of melanin and hemoglobin in human skin from multispectral image acquisition : use of a neuronal network combined to a non-negative matrix factorization.** *Applied and Computational Mathematics* (2012).
- [68] HÄFNER, M., LIEDLGRUBER, M., UHL, A., VÉCSEI, A., AND WRBA, F. **Color treatment in endoscopic image classification using multi-scale local color vector patterns.** *Medical image analysis* (2012).
- [69] NOTO, J. M., AND PEEK, R. M. **Helicobacter pylori : an overview.** In *Helicobacter Species*. Springer, 2012.
- [70] ROBLES-KELLY, A., AND HUYNH, C. P. **Imaging spectroscopy for scene analysis.** Springer Science & Business Media, 2012.
- [71] SHEN, X., SUN, K., ZHANG, S., AND CHENG, S. **Lesion detection of electronic gastroscopie images based on multiscale texture feature.** In *International Conference on Signal Processing, Communication and Computing (ICSPCC)* (2012), IEEE.

- [72] WULFF, J., BUTLER, D. J., STANLEY, G. B., AND BLACK, M. J. **Lessons and insights from creating a synthetic optical flow benchmark**. In *ECCV Workshop on Unsolved Problems in Optical Flow and Stereo Estimation* (2012), A. Fusiello et al. (Eds.), Ed., Part II, LNCS 7584, Springer.
- [73] BERGHOLT, M. S., ZHENG, W., HO, K. Y., TEH, M., YEOH, K. G., SO, J. B. Y., SHABBIR, A., AND HUANG, Z. **Fiber-optic raman spectroscopy probes gastric carcinogenesis in vivo at endoscopy**. *Journal of biophotonics* (2013).
- [74] GALEANO, J., JOLIVOT, R., MARZANI, F., AND BENEZETH, Y. **Unmixing of human skin optical reflectance maps by non-negative matrix factorization algorithm**. *Biomedical Signal Processing and Control* (2013).
- [75] GU, L., ROBLES-KELLY, A., AND ZHOU, J. **Efficient estimation of reflectance parameters from imaging spectroscopy**. *Transactions on Image Processing* (2013).
- [76] HE, K., SUN, J., AND TANG, X. **Guided image filtering**. *Transactions on Pattern Analysis and Machine Intelligence* (2013).
- [77] HERRERA, S. M., AKKA, R., BENEZETH, Y., BOFFETY, M., EMILE, J.-F., MARZANI, F., LAMARQUE, D., AND GOUDAIL, F. **Intérêt de l'analyse multispectrale des images endoscopiques de la muqueuse gastrique pour reconnaître la gastrite inflammatoire et ses différents types : résultats d'une étude préliminaire (oral)**. In *Journées Francophones d'Hépatogastroenterologie et d'Oncologie Digestive* (2013).
- [78] JOLIVOT, R., BENEZETH, Y., AND MARZANI, F. **Skin parameter map retrieval from a dedicated multispectral imaging system applied to dermatology/cosmetology**. *Journal of Biomedical Imaging* (2013).
- [79] KASSAB, R., LAPAYRE, J.-C., AUPET, J.-B., MARZANI, F., AND PIERALLI, C. **Scars collaborative telediagnosis platform using adaptive image flow**. *Integrated Computer-Aided Engineering* (2013).
- [80] KIYOTOKI, S., NISHIKAWA, J., OKAMOTO, T., HAMABE, K., SAITO, M., GOTO, A., FUJITA, Y., HAMAMOTO, Y., TAKEUCHI, Y., SATORI, S., AND OTHERS. **New method for detection of gastric cancer by hyperspectral imaging : a pilot study**. *Journal of biomedical optics* (2013).
- [81] LIM, J. H., KIM, N., LEE, H. S., CHOE, G., JO, S. Y., CHON, I., CHOI, C., YOON, H., SHIN, C. M., PARK, Y. S., AND OTHERS. **Correlation between endoscopic and histological diagnoses of gastric intestinal metaplasia**. *Gut and liver* (2013).
- [82] NOURI, D., LUCAS, Y., AND TREUILLET, S. **Calibration and test of a hyperspectral imaging prototype for intra-operative surgical assistance**. In *Medical Imaging : Digital Pathology* (2013), International Society for Optics and Photonics.
- [83] NOURI, D., LUCAS, Y., TREUILLET, S., JOLIVOT, R., AND MARZANI, F. **Colour and multispectral imaging for wound healing evaluation in the context of a comparative preclinical study**. In *Medical Imaging : Image Processing* (2013), International Society for Optics and Photonics.
- [84] HERRERA, S. M., BENEZETH, Y., BOFFETY, M., MARZANI, F., GOUDAIL, F., AND LAMARQUE, D. **Development of a multispectral endoscope to improve the detection of precancerous lesions in digestive gastroendoscopy (oral)**. In *Doctoral Consortium of the International Conference on Computer Vision* (2014), Imaging and Computer Graphics Theory and Application.
- [85] KRIZHEVSKY, A. **One weird trick for parallelizing convolutional neural networks**. *ArXiv preprint* (2014).

- [86] MARTINEZ-HERRERA, S. E., BENEZETH, Y., BOFFETY, M., EMILE, J.-F., MARZANI, F., LAMARQUE, D., AND GOUDAIL, F. **Multispectral endoscopy to identify precancerous lesions in gastric mucosa**. In *International Conference on Image and Signal Processing* (2014), Springer.
- [87] NOURI, D., LUCAS, Y., AND TREUILLET, S. **Efficient tissue discrimination during surgical interventions using hyperspectral imaging**. In *International Conference on Information Processing in Computer-Assisted Interventions* (2014), Springer.
- [88] NOURI, D., LUCAS, Y., AND TREUILLET, S. **Infrared hyperspectral imaging for enhanced tissue visualization and discrimination during surgical operation**. In *International Conference on Image Processing (ICIP)* (2014), IEEE.
- [89] SIMONYAN, K., AND ZISSERMAN, A. **Very deep convolutional networks for large-scale image recognition**. *ArXiv preprint* (2014).
- [90] VAN DER WALT, S., SCHÖNBERGER, J. L., NUNEZ-IGLESIAS, J., BOULOGNE, F., WARNER, J. D., YAGER, N., GOUILLART, E., YU, T., AND THE SCIKIT-IMAGE CONTRIBUTORS. **scikit-image : image processing in Python**. *PeerJ* (2014).
- [91] BARRON, J. T. **Convolutional color constancy**. In *International Conference on Computer Vision* (2015), IEEE.
- [92] BARRON, J. T., AND MALIK, J. **Shape, illumination, and reflectance from shading**. *Transactions on Pattern Analysis and Machine Intelligence* (2015).
- [93] BIANCO, S., CUSANO, C., AND SCHETTINI, R. **Color constancy using cnns**. In *Conference on Computer Vision and Pattern Recognition Workshops* (2015), IEEE.
- [94] FRANCESCHI, F., GASBARRINI, A., POLYZOS, S. A., AND KOUNTOURAS, J. **Extragastric diseases and helicobacter pylori**. *Helicobacter* (2015).
- [95] HERRERA, S. E. M., TOUATI, E., BENEZETH, Y., BOFFETY, M., FIETTE, L., EMILE, J.-F., MARZANI, F., GOUDAIL, F., AND LAMARQUE, D. **L'analyse spectrale de la lumière réfléchie permet de différencier les différentes étapes des lésions inflammatoires de la muqueuse gastrique dans un modèle expérimental d'infection à h. pylori chez la souris (oral)**. In *Journées Francophones d'Hépatogastroenterologie et d'Oncologie Digestive* (2015).
- [96] MURAKAMI, T., KAGEYAMA, Y., AND NISHIDA, M. **Background replacement using chromatic adaptation transform for visual communication**. *Journal of Visual Communication and Image Representation* (2015).
- [97] NAMIKAWA, T., SATO, T., AND HANAZAKI, K. **Recent advances in near-infrared fluorescence-guided imaging surgery using indocyanine green**. *Surgery today* (2015).
- [98] NARIHIRA, T., MAIRE, M., AND YU, S. X. **Direct intrinsics : Learning albedo-shading decomposition by convolutional regression**. In *International Conference on Computer Vision* (2015), IEEE.
- [99] RASTGOO, M., GARCIA, R., MOREL, O., AND MARZANI, F. **Automatic differentiation of melanoma from dysplastic nevi**. *Computerized Medical Imaging and Graphics* (2015).
- [100] YANG, Q., TANG, J., AND AHUJA, N. **Efficient and robust specular highlight removal**. *Transactions on Pattern Analysis and Machine Intelligence* (2015).
- [101] ZHU, R., ZHANG, R., AND XUE, D. **Lesion detection of endoscopy images based on convolutional neural network features**. In *International Congress on Image and Signal Processing (CISP)* (2015), IEEE.
- [102] AKASHI, Y., AND OKATANI, T. **Separation of reflection components by sparse non-negative matrix factorization**. *Computer Vision and Image Understanding* (2016).

- [103] AMIEVA, M., AND PEEK JR, R. M. **Pathobiology of helicobacter pylori-induced gastric cancer**. *Gastroenterology* (2016).
- [104] FOURURE, D., ÉMONET, R., FROMONT, E., MUSELET, D., TRÉMEAU, A., AND WOLF, C. **Mixed pooling neural networks for color constancy**. In *International Conference on Image Processing (ICIP)* (2016), IEEE.
- [105] GU, L., AND ROBLES-KELLY, A. **A quadratic optimisation approach for shading and specular recovery from a single image**. In *International Conference on Image Processing (ICIP)* (2016), IEEE.
- [106] GU, X., YAO, L., ZHONG, Y., HAN, Z., SHI, Q., FU, Y., LIU, C., WANG, X., AND XIE, T. **Image enhancement based on in vivo hyperspectral gastroscopic images : a case study**. *Journal of biomedical optics* (2016).
- [107] HASSAN, T. M., AL-NAJJAR, S. I., AL-ZAHRANI, I. H., ALANAZI, F. I., AND ALOTIBI, M. G. **Helicobacter pylori chronic gastritis updated sydney grading in relation to endoscopic findings and h. pylori igg antibody : diagnostic methods**. *Journal of microscopy and ultrastructure* (2016).
- [108] IANDOLA, F. N., HAN, S., MOSKEWICZ, M. W., ASHRAF, K., DALLY, W. J., AND KEUTZER, K. **Squeezenet : Alexnet-level accuracy with 50x fewer parameters and < 0.5 mb model size**. *ArXiv preprint* (2016).
- [109] KIENESBERGER, S., COX, L. M., LIVANOS, A., ZHANG, X.-S., CHUNG, J., PEREZ-PEREZ, G. I., GORKIEWICZ, G., ZECHNER, E. L., AND BLASER, M. J. **Gastric helicobacter pylori infection affects local and distant microbial populations and host responses**. *Cell reports* (2016).
- [110] MARTINEZ-HERRERA, S. E., BENEZETH, Y., BOFFETY, M., EMILE, J.-F., MARZANI, F., LAMARQUE, D., AND GOUDAIL, F. **Identification of precancerous lesions by multispectral gastroendoscopy**. *Signal, Image and Video Processing* (2016).
- [111] MEKA, A., ZOLLHÖFER, M., RICHARDT, C., AND THEOBALT, C. **Live intrinsic video**. *ACM Transactions on Graphics (TOG)* (2016).
- [112] OGIHARA, H., HAMAMOTO, Y., FUJITA, Y., GOTO, A., NISHIKAWA, J., AND SAKAIDA, I. **Development of a gastric cancer diagnostic support system with a pattern recognition method using a hyperspectral camera**. *Journal of Sensors* (2016).
- [113] SHARMA, P., BERGMAN, J. J., GODA, K., KATO, M., MESSMANN, H., ALSOP, B. R., GUPTA, N., VENNALAGANTI, P., HALL, M., KONDA, V., AND OTHERS. **Development and validation of a classification system to identify high-grade dysplasia and esophageal adenocarcinoma in barrett's esophagus using narrow-band imaging**. *Gastroenterology* (2016).
- [114] SLAMA, S. B., GHACHEM, D. B., DHAOU, A., JOMNI, M. T., DOUGUI, M. H., AND BELLIL, K. **Gastrites chroniques à hélicobacter pylori : Évaluation des systèmes olga et olgim helicobacter pylori gastritis : Assessment of olga and olgim staging systems**. *Pan African Medical Journal* (2016).
- [115] SUO, J., AN, D., JI, X., WANG, H., AND DAI, Q. **Fast and high quality highlight removal from a single image**. *Transactions on Image Processing* (2016).
- [116] TAJBAKHSH, N., SHIN, J. Y., GURUDU, S. R., HURST, R. T., KENDALL, C. B., GOTWAY, M. B., AND LIANG, J. **Convolutional neural networks for medical image analysis : Full training or fine tuning?** *Transactions on Medical Imaging* (2016).
- [117] XIE, D., LIU, S., LIN, K., ZHU, S., AND ZENG, B. **Intrinsic decomposition for stereoscopic images**. In *International Conference on Image Processing (ICIP)* (2016), IEEE.
- [118] BASLAMISLI, A. S., LE, H.-A., AND GEVERS, T. **Cnn based learning using reflection and retinex models for intrinsic image decomposition**. *ArXiv preprint* (2017).

- [119] BURKITT, M. D., DUCKWORTH, C. A., WILLIAMS, J. M., AND PRITCHARD, D. M. **Helicobacter pylori-induced gastric pathology : insights from in vivo and ex vivo models**. *Disease models & mechanisms* (2017).
- [120] CHEN, X., ZHU, W., ZHAO, Y., YU, Y., ZHOU, Y., YUE, T., DU, S., AND CAO, X. **Intrinsic decomposition from a single spectral image**. *Applied optics* (2017).
- [121] GONO, K. **An introduction to high-resolution endoscopy and narrowband imaging**. *Comprehensive Atlas of High-Resolution Endoscopy and Narrowband Imaging* (2017).
- [122] HOHMANN, M., KANAWADE, R., KLÄMPFL, F., DOUPLIK, A., MUDTER, J., NEURATH, M., AND ALBRECHT, H. **In-vivo multispectral video endoscopy towards in-vivo hyperspectral video endoscopy**. *Journal of biophotonics* (2017).
- [123] HUANG, G., LIU, Z., VAN DER MAATEN, L., AND WEINBERGER, K. Q. **Densely connected convolutional networks**. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), IEEE.
- [124] HUDAK, L., JARAISY, A., HAJ, S., AND MUHSEN, K. **An updated systematic review and meta-analysis on the association between h elicobacter pylori infection and iron deficiency anemia**. *Helicobacter* (2017).
- [125] LI, C., LIN, S., ZHOU, K., AND IKEUCHI, K. **Specular highlight removal in facial images**. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), IEEE.
- [126] WANG, F., AINOUS, S., PETITJEAN, C., AND BENSRAIR, A. **Specularity removal : A global energy minimization approach based on polarization imaging**. *Computer Vision and Image Understanding* (2017).
- [127] GE, Z., SHEH, A., FENG, Y., MUTHUPALANI, S., GE, L., WANG, C., KURNICK, S., MANNION, A., WHARY, M. T., AND FOX, J. G. **Helicobacter pylori-infected c57bl/6 mice with different gastrointestinal microbiota have contrasting gastric pathology, microbial and host immune responses**. *Scientific reports* (2018).
- [128] HUANG, Q., ZHU, W., ZHAO, Y., CHEN, L., WANG, Y., YUE, T., AND CAO, X. **Multispectral image intrinsic decomposition via subspace constraint**. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), IEEE.
- [129] ITOH, T., KAWAHIRA, H., NAKASHIMA, H., AND YATA, N. **Deep learning analyzes helicobacter pylori infection by upper gastrointestinal endoscopy images**. *Endoscopy international open* (2018).
- [130] KANESAKA, T., LEE, T.-C., UEDO, N., LIN, K.-P., CHEN, H.-Z., LEE, J.-Y., WANG, H.-P., AND CHANG, H.-T. **Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging**. *Gastrointestinal Endoscopy* (2018).
- [131] LI, C., BROST, V., BENEZETH, Y., MARZANI, F., AND YANG, F. **Design and evaluation of a parallel and optimized light–tissue interaction-based method for fast skin lesion assessment**. *Journal of Real-Time Image Processing* (2018).
- [132] LIU, X., WANG, C., HU, Y., ZENG, Z., BAI, J., AND LIAO, G. **Transfer learning with convolutional neural network for early gastric cancer classification on magnifying narrow-band imaging images**. In *International Conference on Image Processing (ICIP)* (2018), IEEE.
- [133] TRINH, D. H., DAUL, C., BLONDEL, W., AND LAMARQUE, D. **Mosaicing of images with few textures and strong illumination changes : Application to gastroscopic scenes**. In *International Conference on Image Processing (ICIP)* (2018), IEEE.
- [134] WU, Y., LI, M., ZHANG, Q., AND LIU, Y. **A retinex modulated piecewise constant variational model for image segmentation and bias correction**. *Applied Mathematical Modelling* (2018).

- [135] **Les cancers en france en 2018 - l'essentiel des faits et chiffres (édition 2019)**, 2019.
- [136] LEENHARDT, R., VASSEUR, P., LI, C., SAURIN, J. C., RAHMI, G., CHOLET, F., BECQ, A., MARTEAU, P., HISTACE, A., DRAY, X., AND OTHERS. **A neural network algorithm for detection of gi angiectasia during small-bowel capsule endoscopy**. *Gastrointestinal Endoscopy* (2019).
- [137] LI, Y., XIE, X., YANG, X., GUO, L., LIU, Z., ZHAO, X., LUO, Y., JIA, W., HUANG, F., ZHU, S., AND OTHERS. **Diagnosis of early gastric cancer based on fluorescence hyperspectral imaging technology combined with partial-least-square discriminant analysis and support vector machine**. *Journal of biophotonics* (2019).
- [138] PIMENTEL-NUNES, P., LIBÂNIO, D., MARCOS-PINTO, R., AREIA, M., LEJA, M., ESPOSITO, G., GARRIDO, M., KIKUSTE, I., MEGRAUD, F., MATYSIAK-BUDNIK, T., AND OTHERS. **Management of epithelial precancerous conditions and lesions in the stomach (maps ii) : European society of gastrointestinal endoscopy (esge), european helicobacter and microbiota study group (ehmsg), european society of pathology (esp), and sociedade portuguesa de endoscopia digestiva (sped) guideline update 2019**. *Endoscopy* (2019).
- [139] TRINH, D.-H., AND DAUL, C. **On illumination-invariant variational optical flow for weakly textured scenes**. *Computer Vision and Image Understanding* (2019).
- [140] WU, L., ZHOU, W., WAN, X., ZHANG, J., SHEN, L., HU, S., DING, Q., MU, G., YIN, A., HUANG, X., AND OTHERS. **A deep neural network improves endoscopic detection of early gastric cancer without blind spots**. *Endoscopy* (2019).

TABLE DES FIGURES

1.1	Morphologie de l'estomac (issue de https://www.passeportsante.net/fr/parties-corps/Fiche.aspx?doc=fundus).	3
1.2	Paroi de l'estomac (issue de http://ressources.unisciel.fr/physiologie/co/grain6b2.html).	4
1.3	Score Sydney (issue de https://www.hindawi.com/journals/grp/2013/393015/fig3/).	4
1.4	Fonctionnement du NBI : deux filtres centrés sur les longueurs d'onde 415 et 540 nm sont placés devant la lumière d'un endoscope. La lumière filtrée pénètre le tissu et est absorbée par les vaisseaux sanguins et les capillaires (issue de https://www.barmherzige-brueder.at/site/wien/home).	6
1.5	Répartition des tâches entre les acteurs du projet comme initialement définie lors du dépôt de dossier scientifique.	9
3.1	Exemple de patch extrait d'une image endoscopique NBI	36
3.2	Exemples de patchs extraits des images NBI.	37
3.3	Exemples de patchs difficiles à évaluer.	37
3.4	Exemples de patch décomposé selon le modèle dichromatique.	41
3.5	Juxtaposition des résultats de classification sur la mosaïque fournie par le laboratoire CRAN.	44
3.6	Arbre de décision généré à partir des choix du gastroendoscopiste.	47
4.1	Sacrifice des souris à 1, 3, 6, 12 et 18 mois après infection.	55
4.2	L'estomac de souris est prélevé pour trois types de mesures : des mesures histologiques, sérologiques et spectrales.	56
4.3	Analyse des lésions inflammatoires induites par <i>H. pylori</i> dans la muqueuse des souris. Coloration H et E des lésions histologiques de deux souris, 12 mois après infection. La souris de gauche n'est pas infectée alors que la souris de droite l'est. On observe l'élongation des glandes et la présence d'agrégats de lymphocytes dans la sous-muqueuse (échelle : 50 μ m).	57
4.4	Nombre de souris infectées et non-infectées à 12 et 18 mois selon leur grade histologique déterminé par les méthodes de Rogers <i>et al.</i> (2005) et Eaton <i>et al.</i> (1995).	57
4.5	Colonisation gastrique de <i>H. pylori</i> exprimée en cfu/g de tissu.	57
4.6	Dispositif d'acquisition des spectres.	58
4.7	Distribution des spectres des deux classes.	59
4.8	Corrélation entre bandes spectrales.	59

4.9	Analyse statistique des descripteurs.	62
5.1	Matrices de filtres des deux caméras.	66
5.2	Ensemble des éléments du prototype.	67
5.3	Schéma fonctionnel du prototype.	68
5.4	Fonctionnement des caméras multispectrales.	68
5.5	X-Rite ColorChecker® Digital SG (issue de https://www.xrite.com/categories/calibration-profiling/colorchecker-digital-sg).	69
5.6	RMSE en fonction du nombre de composantes.	71
5.7	Analyse des valeurs propres des données provenant des caméras VIS et NIR.	71
5.8	Estimation de la réflectance à partir des caméras multispectrales.	72
5.9	Exemple de spectre acquis dans l'estomac chez l'Homme.	73
5.10	Transformation en image endoscopique sous lumière blanche.	74
5.11	Application de la transformation en image endoscopique sous lumière blanche sur des images de main et de mire.	75
5.12	Transformation en image endoscopique sous NBI.	76
5.13	Application de la transformation en image endoscopique sous NBI sur des images de main et de mire.	76
6.1	Exemple d'extraction de spectres à partir d'une image multispectrale.	82
6.2	Visualisation des acquisitions spectrales avec les caméras VIS et NIR.	83
6.3	Corrélation entre bandes de longueurs d'onde.	84
6.4	Distribution des prédictions par patient.	87

LISTE DES TABLES

2.1	Exemples d'images des trois bases de données : de gauche à droite et de haut en bas, une pomme, un raton-laveur et un soleil de la base du MIT ; un visage, de la nourriture réelle et factice et des balles de la base CAVE ; et Baymax, un cube et un cylindre de la base SIID.	24
2.2	Résultats quantitatifs sur la base de données du MIT.	27
2.3	Résultats sur trois images de la base du MIT : de gauche à droite, successivement la réflectance et le shading des images du dinosaure, du papier et de la poire. . . .	28
2.4	Résultats sur la base SIID.	29
2.5	Résultats sur trois images de la base SIID : de gauche à droite, successivement la réflectance et le shading des images de Baymax, du cube et d'un hibou.	30
2.6	Résultats sur trois images de la base CAVE : de gauche à droite, successivement la réflectance et le shading des images de ballons, de vrais et faux aliments et des plumes.	31
3.1	Répartition des données vidéos selon les grades histologiques.	35
3.2	Résultats de classification pour les différents descripteurs.	39
3.3	Matrice de confusion résultant de la classification avec le réseau VGG11.	40
3.4	Détail des résultats de classification avec le réseau VGG11.	40
3.5	Résultats de classification pour les différents descripteurs après décomposition par le modèle dichromatique.	42
3.6	Définition des critères visuels et exemples.	45
3.7	Résultats des tests de la somme des rangs.	46
3.8	Résultats de classification avec les critères visuels.	48
3.9	Résultats de classification pour les différents descripteurs.	48
4.1	Résultats de classification des spectres.	63
6.1	Ensemble des données multispectrales obtenues selon les grades histologiques. . .	83
6.2	Résultat de classification des spectres avec la caméra VIS.	85
6.3	Résultat de classification des spectres avec la caméra NIR.	86
6.4	Résultat de classification des spectres après suppression des patients atteints de gastrite active.	86

LISTE DES DÉFINITIONS

ACP (PCA) : Analyse en Composantes Principales (*Principal Component Analysis*)
Algorithme non supervisé de réduction des données.

ANOVA : (*ANalysis Of VAriance*)
ANOVA est un modèle statistique utilisé pour comparer les moyennes d'échantillons appartenant à des catégories différentes.

Apprentissage profond (*Deep Learning*) :
Famille de méthodes de l'intelligence artificielle. Sous-ensemble du Machine Learning qui comprend les réseaux de neurones à plusieurs couches. Ces couches peuvent être de nature variée. Ils ont en général un grand nombre de paramètres internes qui peuvent être optimisés et nécessitent donc une grande base de données d'entraînement.

Apprentissage supervisé (*Machine Learning*) :
Famille de méthodes d'intelligence artificielle qui permettent l'apprentissage automatique à partir d'exemples annotés.

CNN : (*Réseaux de neurones convolutionnels, Convolutional Neural Network*)
Réseau de neurones, particulièrement adapté au traitement du signal et de l'image comprenant des couches de neurones réalisant des opérations de convolution.

DAISY :
Descripteurs utilisés en traitement d'images, très proche des SIFT, développés pour être facilement calculables. Ils permettent entre autres d'estimer de manière dense des cartes de profondeur.

FICE : (*Fuji Intelligent Chromo Endoscopy*)
FICE est un système développé par Fuji qui permet d'estimer une image multispectrale à partir d'une image endoscopique classique.

GLCM : (*Gray Level Co-occurrence Matrix*)
Matrices générées à partir d'images sur lesquelles on peut calculer des descripteurs de texture utilisés en traitement d'images.

HOG : (*Histogram of Oriented Gradients*)
Descripteur utilisé en traitement d'images, notamment utilisé pour la détection d'objets.

LBP : (*Local Binary Patterns*)

Descripteur visuel utilisé en traitement d'images.

LCVP : (*Local Color Vector Patterns*)

Les LCVP sont des descripteurs inspirés des LBP mais étendu aux images couleurs. Ils ont notamment été utilisés pour le traitement des images endoscopiques.

LMSE : (*Local Mean Square Error*)

Métrique consistant à calculer le MSE normalisé sur des fenêtres de taille donnée.

MSE : (*Mean Square Error*)

Métrique consistant à calculer la moyenne des erreurs au carré sur un signal ou une image.

NBI : (*Narrow Band Imaging*)

Modalité d'imagerie basée sur des longueurs d'onde spécifiques correspondant aux pics d'absorption de l'hémoglobine.

NIR : (*Near InfraRed*)

Ensemble de longueur d'ondes situées entre le visible et l'infrarouge, soit entre 780 nm et 2500 nm.

PLS-DA : (*Partial Least Square - Discriminant Analysis*)

Le PLS-DA est un algorithme de classification dérivé du modèle de régression PLS, qui est un modèle linéaire ayant la particularité de prendre en compte le bruit sur les variables explicatives mais aussi sur les variables cibles (réponses).

Précision :

La précision (nommée aussi valeur prédictive positive) mesure la probabilité d'un système à donner le bon résultat lorsque celui-ci est positif.

RMSE : (*Root of mean square error*)

Racine carrée du MSE.

SAM : (*Spectral Angle Mapper*)

Métrique utilisée pour comparer des spectres. Comme son nom l'indique, elle calcule un angle et contrairement au MSE, elle a l'avantage d'être invariante à l'échelle.

Sensibilité :

La sensibilité (nommée aussi rappel) mesure la capacité d'un système à donner un résultat positif lorsqu'une hypothèse est vérifiée.

SIFT : (*Scale Invariant Feature Transform*)

SIFT est un ensemble de descripteurs en traitement d'images invariant à l'échelle qui sont utilisés pour trouver des points similaires entre deux images.

Spécificité :

La spécificité (nommée aussi sélectivité) mesure la capacité d'un système à donner un résultat négatif lorsqu'une hypothèse n'est pas vérifiée.

SVM : (*Séparateurs à Vastes Marges ou Machine à Vecteurs Supports, Support Vector Machine*)

Algorithme de classification.

VIS : (*VISible*)

Ensemble des longueurs d'onde observables par l'Homme, soit entre 380 nm et 780 nm.

WCS : (*Wavelet Correlation Signature*)

Les WCS sont des descripteurs de textures basés sur les ondelettes. Il s'agit d'une extension aux images couleur qui prend en compte la corrélation entre canaux.

