



HAL
open science

Modélisation et traitement statistique d'images de microscopie confocale : application en dermatologie

Abdelghafour Halimi

► **To cite this version:**

Abdelghafour Halimi. Modélisation et traitement statistique d'images de microscopie confocale : application en dermatologie. Autre [cs.OH]. Institut National Polytechnique de Toulouse - INPT, 2017. Français. NNT : 2017INPT0125 . tel-04228537

HAL Id: tel-04228537

<https://theses.hal.science/tel-04228537v1>

Submitted on 4 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Signal, Image, Acoustique et Optimisation

Présentée et soutenue par :

M. ABDELGHAFOUR HALIMI

le lundi 4 décembre 2017

Titre :

Modélisation et traitement statistique d'images de microscopie confocale:
application en dermatologie

Ecole doctorale :

Mathématiques, Informatique, Télécommunications de Toulouse (MITT)

Unité de recherche :

Institut de Recherche en Informatique de Toulouse (I.R.I.T.)

Directeur(s) de Thèse :

M. HADJ BATATIA

Rapporteurs :

M. ANDRE FERRARI, UNIVERSITE DE NICE SOPHIA ANTIPOLIS

Mme MIREILLE GARREAU, UNIVERSITE RENNES 1

Membre(s) du jury :

M. JEAN-YVES TOURNERET, INP TOULOUSE, Président

M. HADJ BATATIA, INP TOULOUSE, Membre

M. JIMMY LE DIGABEL, INSTITUT DE RECHERCHE PIERRE FABRE, Membre

M. MARCELO ALEXANDRO PEREYRA, HERIOT WATT UNIVERSITY, Membre

Remerciements

Ce travail de thèse a été réalisé au sein de l'équipe Traitement et Compréhension des Images (TCI) de l'Institut de Recherche en Informatique de Toulouse (IRIT).

Je remercie avant tout mon directeur de thèse, M. Hadj Batatia, Maître de conférence à l'université de Toulouse, pour son dynamisme, sa tolérance, sa disponibilité, ses compétences scientifiques et ses idées que j'ai pu apprécier tout au long de ma thèse.

Je remercie également M. Jean-Yves Tournet, Professeur à l'université de Toulouse, pour le soutien qu'il a pu m'apporter lorsque cela a été nécessaire. Il a toujours su me consacrer des moments de son temps, me guider et me conseiller. Je souhaite lui transmettre l'expression de ma plus profonde et sincère gratitude.

Durant ces trois années, j'ai eu la chance de collaborer avec des chargés de recherche qui m'ont apporté leurs conseils et expériences à travers de nombreuses réunions. Je pense notamment à M. Jimmy Le Digabel et M. Gwendal Josse des laboratoires Pierre Fabre. Je leur exprime mes sincères remerciements pour ces trois années de recherche.

Je suis profondément reconnaissant à mes deux rapporteurs, Mme. Mireille Garreau, Professeur à l'université de Rennes 1 et M. André Ferrari, Professeur à l'université de Nice, d'avoir accepté de lire et d'évaluer ma thèse.

Je remercie également les membres du jury : M. Marcelo Alexandro Pereyra, Professeur Assistant à l'université de Heriot-Watt en Ecosse et M. Lotfi Chaari, Maître de Conférence à l'université de Toulouse d'avoir accepté d'examiner cette thèse et de participer à ce jury.

Ma gratitude s'adresse aussi à tout le personnel de l'IRIT, enseignants, chercheurs, doctorants et plus spécialement à Sylvie Eichen et Sylvie Armengaud qui ont été disponibles à chaque fois que j'ai eu besoin d'une aide et pour leur sympathie et convivialité.

Ces remerciements ne seraient pas complets sans l'expression de toute ma sympathie envers mes camarades de bureau Zacharie et Facundo ainsi qu'à tous les jeunes que j'ai eu le plaisir de côtoyer durant ces années: Abdelaziz, Adel, Ahmed, Farouk, Jorge, Mohammed Amine, Mohammed, Mohanad, Pierre-Antoine, Vincent, ... Les moments magnifiques que nous avons passés ensemble au bureau, dans le couloir, à la cantine, au foyer et sur le terrain de foot, je les garde précieusement au fond de mon coeur. Je passe ensuite une dédicace spéciale pour mon cousin Abdelouahab et mes plus proches amis Baha Eddine Youcef et Mourad pour leurs encouragements et pour leur bonne humeur communicative, et aussi pour leur soutien qu'ils m'ont apporté durant toutes les périodes difficiles de ma thèse.

Je tiens finalement à remercier ma famille dont mes trois frères ainsi que mes parents sans qui rien de cela n'aurait été possible. Qu'ils trouvent dans ces lignes toute mon admiration, mon affection et ma gratitude.

Résumé

Dans cette thèse, nous développons des modèles et des méthodes statistiques pour le traitement d'images de microscopie confocale de la peau dans le but de détecter une maladie de la peau appelée lentigo. Une première contribution consiste à proposer un modèle statistique paramétrique pour représenter la texture dans le domaine des ondelettes. Plus précisément, il s'agit d'une distribution gaussienne généralisée dont on montre que le paramètre d'échelle est caractéristique des tissus sous-jacents. La modélisation des données dans le domaine de l'image est un autre sujet traité dans cette thèse. A cette fin, une distribution gamma généralisée est proposée. Notre deuxième contribution consiste alors à développer un estimateur efficace des paramètres de cette loi à l'aide d'une descente de gradient naturel. Finalement, un modèle d'observation de bruit multiplicatif est établi pour expliquer la distribution gamma généralisée des données. Des méthodes d'inférence bayésienne paramétrique sont ensuite développées avec ce modèle pour permettre la classification d'images saines et présentant un lentigo. Les algorithmes développés sont appliqués à des images réelles obtenues d'une étude clinique dermatologique.

Mots clés: Microscopie confocale par réflectance, modèle de texture, gaussienne généralisée, distribution gamma généralisée, estimation de paramètres, géométrie de l'information, gradient naturel, méthodes bayésiennes, MCMC

Abstract

In this work, we develop statistical models and processing methods for confocal microscopy images. The first contribution consists of a parametric statistical model to represent textures in the wavelet domain. Precisely, a generalized Gaussian distribution is proposed, whose scale parameter is shown to be discriminant of the underlying tissues. The thesis deals also with modeling data in the image domain using the generalized gamma distribution. The second contribution develops an efficient parameter estimator for this distribution based on a natural gradient approach. The third contribution establishes a multiplicative noise observation model to explain the distribution of the data. Parametric Bayesian inference methods are subsequently developed based on this model to classify healthy and lentigo images. All algorithms developed in this thesis have been applied to real images from a dermatologic clinical study.

Keywords: Reflectance confocal microscopy, texture modeling, generalized Gaussian, generalized gamma distribution, parameter estimation, information geometry, natural gradient, Bayesian methods, MCMC

Abréviations et notations

Abréviations

SVM	Support Vector Machine
SURF	Speeded Up Robust Features
CRB	Cramér-Rao bound
MCMC	Markov chain Monte-Carlo (Monte Carlo par chaînes de Markov)
EM	Expectation-maximisation
EQM	Erreur quadratique moyenne
MLE	Maximum likelihood estimator (estimateur du maximum de vraisemblance)
PDF	Fonction densité de probabilité
MAP	Maximum a posteriori
MMSE	Estimateur qui minimise l'erreur quadratique moyenne (Minimum mean square error)
DPMM	Dirichlet process mixture model
GGD	Distribution gamma généralisée
DEJ	Jonction dermo-épidermique

Fonctions mathématiques

$N(\cdot)$: Loi normale
$\Gamma(\cdot)$: Fonction gamma
$\Psi(\cdot)$: Fonction digamma
$\Psi'(\cdot)$: Fonction trigamma
$\mathcal{G}(\cdot)$: Loi gamma
$\mathcal{IG}(\cdot)$: Loi inverse gamma
$\mathcal{I}_{\mathbb{R}^+}$: Fonction indicatrice sur \mathbb{R}^+
$E(\cdot)$: Espérance

Table des matières

Remerciements	iii
Résumé	v
Abstract	vii
Abréviations et notations	ix
Introduction	1
1 Contexte médical et microscopie confocale à réflectance	7
1.1 Introduction	7
1.2 La peau	8
1.3 Composition de la peau	9
1.3.1 L'épiderme	10
1.3.2 La jonction dermo-épidermique	12
1.3.3 Le derme	13
1.3.4 L'hypoderme	14
1.4 Fonctions principales de la peau	15
1.4.1 Protection	15
1.4.2 Fonction sensorielle	15
1.4.3 Thermorégulation	15

1.4.4	Fonction métabolique: synthèse de la vitamine D	15
1.5	Pigmentation de la peau	15
1.5.1	Pigmentation mélanique	15
1.5.2	Hyperpigmentations	16
1.6	Techniques d'imagerie de la peau	18
1.6.1	Imagerie ultrasonore	19
1.6.2	Imagerie par résonance magnétique	20
1.6.3	Tomographie par cohérence optique	22
1.6.4	Microscopie confocale à réflectance MCR	23
1.7	Conclusions	28
2	Modélisation statistique et classification d'images de microscopie confocale à réflectance à l'aide d'une analyse multirésolution.	29
2.1	Introduction	29
2.2	Méthode proposée	32
2.2.1	Décomposition en ondelettes	32
2.2.2	Modélisation Statistique	33
2.2.3	Caractérisation utilisant le test paramétrique de Student et la pseudo-distance de Kullback-Leibler	36
2.2.4	Classification utilisant les machines à vecteurs support	41
2.3	Expériences	41
2.3.1	Données	41
2.3.2	Analyse statistique des coefficients d'ondelettes	42
2.3.3	Identification des profondeurs caractéristiques	45
2.3.4	Signification statistique avec le test T	54
2.3.5	Performance du classifieur SVM	58
2.4	Conclusions	61

3	Modélisation statistique et classification d’images de microscopie confocale à réflectance	63
3.1	Introduction	63
3.2	Méthode proposée	64
3.2.1	Estimation statistique	64
3.2.2	Caractérisation utilisant le test paramétrique de Student et la distance de Kullback-Leibler	67
3.3	Expériences	68
3.3.1	Performance de l’algorithme d’estimation proposé	68
3.3.2	Identification des profondeurs caractéristiques	76
3.3.3	Importance statistique avec T-Test	79
3.3.4	Classification d’images sains et avec lentigo	82
3.4	Conclusions	85
4	Approche bayésienne non supervisée pour la reconstruction et la classification conjointe d’images de microscopie confocale à réflectance cutanées	87
4.1	Introduction	87
4.2	Formulation du problème	88
4.2.1	Modèle d’observation	88
4.3	Modèle bayésien hiérarchique	89
4.3.1	Vraisemblance	89
4.3.2	Distribution a priori du signal d’intérêt	90
4.3.3	Distribution a priori des variances du bruit	90
4.3.4	Distribution a priori du vecteur d’étiquette \mathbf{z}	91
4.3.5	Lois a priori des hyperparamètres	91
4.3.6	Distribution conjointe <i>a posteriori</i>	91
4.4	Algorithme de Metropolis-within-Gibbs	92
4.4.1	Inférence bayésienne et estimation des paramètres	93
4.4.2	Convergence	94

4.5	Résultats de simulation	95
4.5.1	Données synthétiques	95
4.5.2	Données réelles	101
4.6	Conclusions	102
5	Conclusions et Perspectives	105
	Conclusions et perspectives	105
5.1	Perspectives	107
	Annexes	113
A	T-test et facteur de Bayes	113
B	Matrice de Fisher - Gaussienne Généralisée	115
C	Matrice de Fisher - Gamma Généralisée	119
D	Divergence de Kullback-Leibler, entropie, variance	123
D.1	Kullback Liebler	123
D.2	Variance	124
D.3	Entropie	124
E	Échantillonnage des paramètres	125
E.1	Echantillonnage du paramètre s_{nl}	125
E.2	Echantillonnage du paramètre θ_l	126
E.3	Echantillonnage du paramètre μ_k	126
E.4	Echantillonnage du paramètre σ^2_k	127
E.5	Echantillonnage du paramètre z_l	128
F	Échantillonnage des paramètres du modèle inverse gamma	129
F.1	Formulation du problème	129

F.1.1	Modèle d'observation	129
F.2	Modèle Bayésien Hiérarchique	129
F.2.1	Vraisemblance	130
F.2.2	Distribution a priori du signal d'intérêt	130
F.2.3	Distribution a priori des variances du bruit	130
F.2.4	Distribution a priori du vecteur d'étiquette \mathbf{z}	131
F.2.5	Lois a priori des hyperparamètres	131
F.2.6	Distribution conjointe <i>a posteriori</i>	131
F.3	Algorithme de Metropolis-within-Gibbs	133
F.4	Echantillonnage du paramètre θ_l	134
F.5	Echantillonnage du paramètre β_k	135
F.6	Echantillonnage du paramètre z_l	135
F.7	Résultats de simulation	135
F.7.1	Données synthétiques	135
F.7.2	Données réelles	138
G	Approche bayésienne non paramétrique	141
G.1	Formulation du problème	141
G.1.1	Modèle d'observation	141
G.2	Application du modèle de mélange de processus de Dirichlet (DPMM)	141
G.2.1	Classification en utilisant DPMM	142
G.2.2	Prior pour le vecteur d'étiquette label \mathbf{z}	142
G.2.3	Distribution a priori des paramètres moyennes et variances	143
G.3	Résultats	146
G.3.1	Test pour différents mélanges de gaussiennes	146
G.3.2	Données réelles	148
	Bibliographie	161

Table des figures

1.1	Représentation schématique de la peau humaine [Sage, 2012]. A) Illustration des différentes couches de la peau et des annexes cutanées (modifié de [MacNeil, 2007]). B) Représentation des différentes couches de cellules qui forment l'épiderme (modifié de [Fuchs and Raghavan, 2002])	9
1.2	Types de cellules et couches de l'épiderme d'une peau épaisse. Les caractéristiques d'une peau mince sont semblables, toutefois le stratum lucidum est absent et le stratum corneum est plus mince [Tortora and Grabowski, 1993]	10
1.3	Les couches de l'épiderme. L'épiderme est constitué de plusieurs couches cellulaires qui, de la plus profonde à la plus superficielle, sont les suivantes : la couche basale, la couche épineuse, la couche granuleuse, la couche de transition (seulement dans la peau épaisse) et la couche cornée (modifié de [Geras, 1990],[Gagnon, 2005])	13
1.4	Schéma de la peau. Les crêtes épidermiques et les papilles dermiques alternent dans la jonction dermo-épidermique [Laplante, 2002]	14
1.5	Melasma (http://www.asarchcenter.com/proced_kb/melasma/)	16
1.6	Taches café au lait [Oualid, 2011]	17
1.7	Lentigo simplex [Oualid, 2011]	18
1.8	Lentigo actinique [Oualid, 2011]	19
1.9	Image échographique de la peau avec une sonde de 20 MHz - Mode B [Naouri, 2008]	20
1.10	Anatomie de la peau par IRM (séquence FIESTA)	21

1.11	Résolution et profondeur de pénétration comparatives des ultrasons, de la microscopie confocale et de la tomographie par cohérence optique [Sacchet, 2010]	22
1.12	Principe de fonctionnement d'un microscope confocal (http://www3.jouy.inra.fr/miaj/public/imaste/Projets/AI2M_v2/node9.html)	24
1.13	Vivascope 1500 (http://www.vivascope.de/en/products/devices/in-vivo-devices/vivascope-1500/pictures.html)	24
1.14	Images de microscopie confocale à réflectance prises à différentes profondeurs de la peau	26
1.15	Images MCR (patient sain et patient atteint de lentigo) prises au niveau de la jonction dermo-épidermique	27
1.16	Couche cornée [Kanitakis et al., 2013]	27
2.1	Étapes de la méthode de caractérisation proposée pour une profondeur donnée.	33
2.2	Schéma de décomposition en sous-bandes pour une image 2D-DWT à deux niveaux. y_v est une image 2D acquise à la profondeur v	34
2.3	Exemple de décomposition en ondelettes pour une image MCR au premier niveau de décomposition.	34
2.4	Histogrammes et lois gaussiennes associées pour les quatre paramètres (échelle, forme, entropie, variance) avec leurs moyennes et écart-types en utilisant tous les patients (sains et lentigo)	37
2.5	Matrices des pseudo-distances Kulback-Leibler en fonction de la profondeur Z	40
2.6	Histogrammes des coefficients d'ondelettes à partir des bandes mélangées HVD aux quatre échelles et leurs distributions GGD estimées correspondantes. La Figure montre les données de deux patients arbitraires sains et lentigo (#6 and #38 respectivement) à trois profondeurs représentatives (une profondeur par colonne). On remarque le très bon ajustement pour les échelles de 1 à 3 pour les deux types de patients.	43
2.7	Évaluation de l'ajustement d'une loi GGD aux coefficients d'ondelettes. Statistique KS moyenne pour l'ensemble de la population à diverses profondeurs et pour toutes les échelles. Les scores sont excellents pour toutes les configurations.	44

2.8	Évolution des paramètres moyens $\hat{\alpha}$ et $\hat{\beta}$ en fonction de la profondeur, pour les différentes bandes et à toutes les échelles. Les valeurs de α sont très similaires pour les patients sains et lentigo et ne peuvent être utilisées pour la discrimination. Le paramètre β montre une différence significative pour les profondeurs entre 31 μm and 76 μm , avec une différence maximale aux environs de 50 μm . Notre conclusion est que ce paramètre β peut discriminer les tissus sains et lentigo de la peau.	46
2.9	Evolution des paramètres moyens de l'entropie et de la variance en fonction de la profondeur, pour les différentes bandes et à toutes les échelles. Les valeurs d'entropie et de variance sont très similaires pour les patients sains et lentigo et ne peuvent être utilisées pour la discrimination.	47
2.10	Évolution des distances Kulback-Leibler moyennes ainsi que leurs approximations en fonction de la profondeur, pour les différentes bandes à toutes les échelles. Ces distances montrent une différence significative pour les profondeurs entre 31 μm et 76 μm , avec une différence maximale aux environs de 50 μm	48
2.11	Évolution des paramètres moyens $\hat{\alpha}$ et $\hat{\beta}$ en fonction de la profondeur, pour les différentes bandes à toutes les échelles. Les valeurs de α sont très similaires pour les patients sains et lentigo et ne peuvent être utilisées pour la discrimination. Le paramètre β montre une différence significative pour les profondeurs entre 31 μm et 76 μm , avec une différence maximale aux environs de 50 μm . Notre conclusion est que ce paramètre β peut discriminer les tissus sains et lentigo de la peau.	49
2.12	Images des patients sains (patient #1, #2, #3, #4, #5, #6) et lentigo (patient #31, #33, #37, #38, #40, #44) à la profondeur DEJ. On peut observer la différence significative dans la texture, capturée par le paramètre β (comme l'illustre la Figure 2.8).	51
2.13	Images macro couleur des patients sains (patient #1, #2, #3, #4, #5, #6) et lentigo (patient #31, #33, #37, #38, #40, #44) affichées dans la Figure 2.12.	52
2.14	Images couleur des patients sains et lentigo.	53
2.15	P-value (in $-\log$ scale) et facteur de Bayes du test T pour α . Le faible score montre que α n'est manifestement pas un discriminant entre les images des patients sains et lentigo.	55

2.16	P-value (in $-\log$ scale) et facteur de Bayes (BF) du test T pour β . La troisième ligne représente des zooms sur les faibles scores du BF afin de clarifier le seuil de signification. Des scores forts peuvent être observés pour des profondeurs comprises entre $31 \mu m$ and $76 \mu m$. Les scores les plus élevés sont obtenus pour des profondeurs aux environs de $50 \mu m$. Cela confirme que β est une bonne fonction discriminante qui peut être utilisée pour séparer les images saines et lentigo à ces profondeurs.	55
2.17	P-value (in $-\log$ scale) et facteur de Bayes du test T pour l'entropie. Le faible score montre que l'entropie n'est manifestement pas un discriminant entre les images des patients sains et lentigo pour les échelles (1 et 2).	56
2.18	P-value (in $-\log$ scale) et facteur de Bayes du test T pour la variance. Le faible score montre que la variance n'est manifestement pas un discriminant entre les images des patients sains et lentigo pour les échelles (1 et 2).	56
2.19	Les profondeurs caractéristiques qui se situent entre $48 \mu m$ et $63 \mu m$ selon le test T et les profondeurs DEJ associées aux 45 patients.	57
2.20	Exemples d'images MCR de patients sains et lentigo classés par le classifieur SVM. . .	59
3.1	Méthode de classification proposée.	65
3.2	Evolution du RMSE et du temps total des trois méthodes pour N variant de 40 à 1000.	70
3.3	Evolution du biais des trois méthodes pour N variant de 40 à 1000.	70
3.4	Evolution du temps total des trois méthodes pour N=10000 et N=100 lorsque les paramètres (γ, β, ρ) varient	71
3.5	Evolution du paramètre RMSE des trois méthodes pour N=10000 lorsque les paramètres (γ, β, ρ) varient.	72
3.6	Evolution du paramètre biais des trois méthodes N=10000 lorsque les paramètres (γ, β, ρ) varient.	72
3.7	Evolution du paramètre RMSE des trois méthodes pour N=100 lorsque les paramètres (γ, β, ρ) varient.	73
3.8	Evolution du paramètre biais des trois méthodes pour for N=100 lorsque les paramètres (γ, β, ρ) varient.	73

3.9	Histogrammes des intensités d’images MCR. La Figure représente les données de deux patients arbitraires sain et lentigo (#6 and #38) à trois profondeurs représentatives (une profondeur par colonne).	74
3.10	Evaluation de l’adaptation du GGD aux pixels. Statistique KS moyenne pour l’ensemble de la population et pour toutes les profondeurs. Les scores sont excellents pour toutes les configurations.	75
3.11	Évolution des paramètres moyens $\hat{\gamma}$, $\hat{\beta}$ et $\hat{\rho}$ tout au long de la profondeur. Les valeurs de γ sont trop similaires pour les patients sains et lentigo et ne peuvent pas être utilisées pour la discrimination. Les paramètres β et ρ montrent une différence significative pour les profondeurs entre $30 \mu m$ et $76 \mu m$, avec une différence maximale aux environs de $50 \mu m$. Notre conclusion est que les paramètres β et ρ peuvent discriminer les tissus sains et lentigo de la peau.	76
3.12	Évolution des paramètres moyens entropie et variance tout au long de la profondeur. Les paramètres E et V montrent une différence significative pour les profondeurs entre $30 \mu m$ et $76 \mu m$, avec une différence maximale aux environs de $50 \mu m$	77
3.13	Évolution des distances Kulaback-Leibler moyennes tout au long de la profondeur. Les distances montrent une différence significative pour les profondeurs entre $30 \mu m$ et $76 \mu m$, avec une différence maximale aux environs de $50 \mu m$	77
3.14	Images de patients sains (patient #1, #2, #3, #4, #5, #6) et lentigo (patient #31, #33, #37, #38, #40, #44) à la profondeur DEJ ($54 \mu m$). On peut observer des images plus texturées en présence de lentigo.	78
3.15	P-value (en échelle logarithmique $-\log$) et facteur de Bayes (BF) du T test pour γ , β and ρ . Le score faible montre que γ n’est pas discriminant entre les images saines et lentigo. Des scores forts peuvent être observés pour des profondeurs comprises entre $30 \mu m$ et $76 \mu m$ et les scores les plus élevés sont obtenus pour des profondeurs autour de $50 \mu m$ pour les paramètres β et ρ . Cela confirme que β et ρ sont de bonnes fonctions discriminantes qui peuvent être utilisées pour séparer les images saines et lentigo à ces profondeurs.	79

3.16	P-value (en échelle logarithmique $-\log$) et facteur de Bayes (BF) du T test pour E et V. Des scores forts peuvent être observés pour des profondeurs comprises entre $30 \mu m$ et $76 \mu m$ et les scores les plus élevés sont obtenus pour des profondeurs autour de $50 \mu m$ pour les paramètres E et V.	80
3.17	Profondeurs caractéristiques estimées entre 40um et 60um selon le T-test (lignes vertes et rouges) et les vrais profondeurs DEJ annotées par les dermatologues (lignes bleues) associées aux 45 patients.	81
3.18	Exemples d'images MCR des patients sains et lentigo classées à l'aide du classifieur SVM.	84
4.1	DAG pour les lois a priori des paramètres et hyperparamètres. Les hyperparamètres fixés par l'utilisateur apparaissent dans des boîtes (ligne continue).	92
4.2	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des données synthétiques avec $SNR_Y = 0$ dB.	97
4.3	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des données synthétiques avec $SNR_Y = 10$ dB.	98
4.4	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des données synthétiques avec $SNR_Y = 20$ dB.	99
4.5	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des images MCR réelles.	100
4.6	Images (à la profondeur $54 \mu m$) du patient #8 qui est mal classé en comparaison au patient sain et patient atteint de lentigo (bien classés). On peut observer plus de similitude entre ce patient et le patient sain qu'avec celui atteint lentigo.	103
4.7	Exemples d'images bruitées (à la profondeur $54 \mu m$) et leurs vrais reflectivités estimées.	103
5.1	Exemples d'images blocs ainsi que les zones à segmenter	109
F.1	DAG pour les lois a priori des paramètres et hyperparamètres. Les hyperparamètres fixés par l'utilisateur apparaissent dans des boîtes (ligne continue).	132
F.2	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des données synthétiques avec $SNR_Y = 0$ dB.	136

F.3	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_Y = 10$ dB.	137
F.4	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_Y = 20$ dB.	137
F.5	Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des images MCR réelles.	138
G.1	(gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.	146
G.2	(gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.	147
G.3	(gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.	147
G.4	(gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.	149

Liste des tableaux

- 2.1 Profondeurs où β discrimine les patients ($H_0^{(\beta)}$ rejeté) obtenu pour $T_{\nu}^{(\beta)} > T_{0.05} = 2.02$; La p-value et le facteur de Bayes (BF) correspondants sont indiqués. La première ligne donne les intervalles de profondeur (profondeur minimale à la profondeur maximale) où les scores T sont significatifs. La deuxième ligne montre les profondeurs donnant les scores T les plus élevés (score maximal $\mp 10\%$). La troisième ligne montre les profondeurs correspondant au score T maximal. Les p-value et les facteurs de Bayes correspondants à chaque profondeur sont indiqués en dessous. 54

- 2.2 Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre d'échelle α . (Sensit, Specif) signifie (Sensibilité et Spécificité) 59

- 2.3 Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre de forme β . (Sensit, Specif) signifie (Sensibilité et Spécificité) 60

- 2.4 Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre entropie. (Sensit, Specif) signifie (Sensibilité et Spécificité) 60

- 2.5 Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre variance. (Sensit, Specif) signifie (Sensibilité et Spécificité) 61

- 3.1 Paramètres Γ GD estimés associés aux paramètres Γ GD illustrés dans la Figure 3.9. . . 75

3.2	Profondeurs où $T^{(\beta)}$ et $T^{(\rho)} > T_{0.05} = 2.02$; les valeurs correspondantes de p-value et du facteur de Bayes (BF) sont mentionnées. La première ligne donne les intervalles de profondeur (profondeur minimale à la profondeur maximale) où les scores T sont significatifs. La deuxième ligne montre les profondeurs donnant les scores T les plus élevés (score maximal $\mp 10\%$). La troisième ligne montre les profondeurs correspondant au score maximal T. Les P-values et les facteurs de Bayes correspondant à chaque profondeur sont présentés ci-dessous.	80
3.3	Matrices de confusion du classifieur SVM basé sur β and ρ (pour les profondeurs caractéristiques ($40 \mu m$ à $60 \mu m$)). On remarque la bonne précision pour les deux paramètres en particulier pour β	83
3.4	Performance de la classification sur des données réelles (45 patients) utilisant la méthode CART (pour les profondeurs caractéristiques ($40 \mu m$ à $60 \mu m$)).	83
4.1	Performance de l'algorithme proposé pour le débruitage et la classification des données synthétiques bruitées pour trois valeurs du signal/bruit $SNR_Y = [0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}]$	101
4.2	Performance de la classification sur des données réelles (45 patients) en utilisant la méthode proposée (pour la profondeur caractéristique $54 \mu m$).	102
4.3	Performance de la classification sur des données réelles (45 patients) en utilisant la méthode CART (pour la profondeur caractéristique $54 \mu m$).	102
F.1	Performance de l'algorithme proposé pour le débruitage et la classification des données synthétiques bruitées pour trois valeurs du signal/bruit $SNR_Y = [0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}]$	139
F.2	Performance de la classification sur des données réelles (45 patients) en utilisant la méthode proposée dans cette annexe.	139
F.3	Performance de la classification sur des données réelles (45 patients) en utilisant la méthode CART.	140
F.4	Performance de la classification sur des données réelles (45 patients) en utilisant la méthode proposée dans le chapitre 4.	140
G.1	Performance de l'algorithme proposé pour la classification en utilisant différents nombres de gaussiennes dans le mélange	148

G.2	Performance de classification sur des données réelles (45 patients) en utilisant la méthode proposée (pour la profondeur caractéristique 54 μm).	150
G.3	Performance de classification sur des données réelles (45 patients) en utilisant la méthode CART (pour la profondeur caractéristique 54 μm).	150

Introduction

Contexte et problématique de la thèse

Cette thèse s'intéresse à l'analyse et la modélisation statistique des images de microscopie confocale à réflectance (MCR) afin de les caractériser pour l'identification du Lentigo.

La microscopie confocale à réflectance est une modalité d'imagerie médicale de plus en plus utilisée. En effet, elle permet d'avoir une très bonne résolution de la peau (de l'ordre du μm en surface et en profondeur) sans altération de cette dernière. Sa facilité d'utilisation et sa rapidité d'acquisition tendent au fort développement de cet outil. Il faut cependant un long temps d'apprentissage pour que les dermatologues puissent utiliser pleinement les possibilités de cette technique à des fins de diagnostic. C'est pourquoi de plus en plus de méthodes sont utilisées pour minimiser ce temps d'apprentissage en automatisant certaines étapes nécessaires au diagnostic. Nous souhaitons dans cette recherche apprendre à caractériser et distinguer la peau saine de la peau atteinte de lentigo. Les lentigos sont des taches de vieillesse qui apparaissent principalement sur la main où sur les zones le plus souvent exposées au soleil. Sur la surface de la peau, ils apparaissent comme une tache plus foncée. A l'intérieur de la peau, c'est principalement au niveau de la jonction derme-épiderme que nous verrons des différences. En effet, celle-ci est plus grande et désordonnée sur une zone atteinte.

La microscopie confocale possède deux modes d'acquisition. On peut effectuer des prises *stack* : c'est à dire on positionne le capteur et on acquiert des images à différentes profondeurs de peau en faisant varier le plan focal. On peut aussi effectuer des prises bloc où le système est le même mais c'est la caméra qui bouge selon les trois axes x, y et z. Les deux acquisitions n'amènent pas les mêmes

questions. En effet, la taille d'une image étant de $500 \times 500 \mu m$, on va avoir des images *stack* qui vont représenter soit une peau saine soit une peau présentant un lentigo. Par contre, lors d'une analyse bloc, une image va représenter un champ large de taille $8 \times 8 mm$ où on peut faire la distinction entre la zone saine de la zone avec lentigo puisque celui-ci a généralement un diamètre inférieure à $5 mm$. La motivation de cette thèse est de proposer des modèles statistiques permettant de caractériser les images *stack* pour mieux identifier la peau atteinte de lentigo de la peau saine. Pour cela, comme première étape, nous avons utilisé les méthodes multi-résolution pour la modélisation des images MCR. Ces dernières ont montré de bonnes propriétés pour l'analyse d'images texturées ce qui a motivé leur utilisation pour la modélisation des images MCR. Dans une seconde étape, nous avons caractérisé les images MCR directement à partir des intensités des différents pixels en tenant compte du bruit multiplicatif affectant les observations. Dans une troisième étape, nous nous sommes intéressés à la classification des images confocales en adoptant des approches bayésiennes paramétriques permettant d'inclure des informations a-priori concernant les images considérées. Les modèles proposés ont été développés en collaboration avec le Centre de Recherche sur la peau, Pierre Fabre Dermo Cosmétique. Ils ont pu être testés sur des images MCR réelles collectées lors d'une étude nommée *cf010* menée sur 45 patients volontaires.

Organisation du manuscript

Le manuscrit est organisé de la manière suivante.

- **Chapitre 1:** Le chapitre 1 introduit le contexte médical qui motive ce travail en présentant des notions générales sur les tissus de la peau et différentes techniques non-invasives d'imagerie, comme l'Imagerie par Résonance Magnétique (IRM), l'Echographie, la Tomographie par Cohérence Optique, et la Microscopie Confocale à Réflectance (MCR) dont les deux paramètres importants sont la résolution et la profondeur de pénétration.
- **Chapitre 2:** Le chapitre 2 s'intéresse à l'analyse des images MCR pour caractériser la profondeur à laquelle on peut distinguer une peau saine de celle atteinte de lentigo. Chaque patient (sain ou lentigo) est représenté par une suite d'images MCR (appelée image *stack* dans

ce manuscrit) qui sont associées aux différentes profondeurs. La caractérisation des textures de ces images a été faite à travers l'utilisation de la décomposition en ondelettes de Daubechies 4 pour chaque profondeur et pour tous les patients (lentigo et sains). Un premier travail a consisté à vérifier que la loi des coefficients d'ondelettes d'une image MCR était une loi gaussienne généralisée comme cela a été proposé dans la littérature pour d'autres types d'images. Les paramètres de cette loi ont été ensuite estimés par une méthode du maximum de vraisemblance. Nous avons ensuite caractérisé la présence de lentigo à une profondeur donnée par une analyse des paramètres de cette loi gaussienne généralisée. Enfin, nous avons appliqué des tests paramétriques (T-tests, Machine à vecteurs de support (SVM)) sur les différents paramètres pour déterminer les profondeurs caractéristiques auxquelles on peut distinguer les images lentigo et non-lentigo de manière supervisée.

- **Chapitre 3:** Le chapitre 3 porte sur la modélisation statistique des intensités des images MCR. Nous montrons tout d'abord que la loi gamma généralisée est bien adaptée pour décrire les propriétés statistiques des intensités des images des patients sains et atteints de lentigo. Pour estimer les paramètres de cette loi, nous avons proposé un estimateur qui s'avère être plus rapide et plus performant (pour un faible nombre d'échantillons) que ceux disponibles dans la littérature. Cet estimateur utilise la méthode du maximum de vraisemblance combinée avec une descente de gradient naturel. Ce dernier permet d'avoir une convergence plus rapide en pondérant le gradient par l'inverse de la matrice de Fisher des paramètres. Nous avons ensuite étudié les performances de tests paramétriques appliqués aux paramètres de la loi gamma généralisée et nous avons comparé les résultats trouvés avec ceux du chapitre 2.
- **Chapitre 4:** Ce chapitre présente un modèle bayésien hiérarchique totalement non-supervisé permettant une restauration des images MCR et une classification entre patients sains et atteints de lentigo. L'approche bayésienne proposée tient compte de la loi du bruit multiplicatif (*speckle*) affectant les mesures et introduit des lois a-priori adaptées aux paramètres inconnus du modèle, afin de régulariser le problème d'estimation. Nous avons ensuite proposé des méthodes de Monte Carlo par Chaines de Markov (MCMC) afin d'estimer conjointement les différents paramètres

du modèle, incluant la réflectivité débruitée de l'image, un vecteur de label permettant de classer chaque pixel de l'image dans la classe "sain" ou "lentigo" et les paramètres de la loi du bruit. Plus précisément, un échantillonneur de Metropolis-within-Gibbs est utilisé pour générer des échantillons distribués asymptotiquement suivant la loi a-posteriori d'intérêt. Ces échantillons sont ensuite utilisés pour construire des estimateurs bayésiens, comme l'estimateur du maximum a posteriori (MAP) ou l'estimateur MMSE (minimum mean square error) des paramètres inconnus du modèle. L'algorithme résultant est validé sur des données synthétiques et réelles en l'appliquant aux profondeurs caractéristiques du lentigo, déterminées dans les précédents chapitres 2 et 3.

Principales contributions

Les principales contributions sont rappelées ci -dessous

- **Chapitre 2** Une première contribution consiste à proposer un modèle statistique paramétrique pour représenter la texture des images MCR dans le domaine des ondelettes. Précisément, il s'agit d'une distribution gaussienne généralisée dont on montre que le paramètre d'échelle est caractéristique des tissus sous-jacents. [[Halimi et al., 2017,Biomedical](#)]
- **Chapitre 3** Une deuxième contribution consiste à proposer un modèle statistique adapté au domaine de l'image pour la caractérisation des tissus des images MCR en utilisant un nouvel algorithme d'estimation pour les paramètres de la loi gamma généralisée, basé sur une approche de gradient naturel.[[Halimi et al., 2017,CAMSAP](#)], [[Halimi et al., SIVP](#)]
- **Chapitre 4** Dans ce chapitre, un modèle d'observation de bruit multiplicatif est établi pour expliquer la distribution gamma généralisée des données. Des méthodes d'inférence bayésienne paramétrique sont ensuite développées avec ce modèle pour permettre la reconstruction et la classification conjointe d'images de microscopie confocale à réflectance cutanées.[[Halimi et al., 2017 EUSIPCO](#)], [[Halimi et al., IEEE BE](#)]

Liste des publications

Journal papers

1. A. Halimi, H. Batatia, J.L. Digabel, G. Josse and J.Y. Tourneret, “ Statistical modeling of reflectance confocal microscopy images and characterization of skin lentigo ” Biomedical Optics Express, 8, 5450-5467 (2017)
2. A. Halimi, H. Batatia, J.L. Digabel, G. Josse and J.Y. Tourneret, “ Statistical Model-based Classification of Reflectance Confocal Microscopy Images ” Signal image and video processing, à soumettre
3. A. Halimi, H. Batatia, J.L. Digabel, G. Josse and J.Y. Tourneret, “ An unsupervised Bayesian approach for the joint reconstruction and classification of reflectance confocal microscopy images” IEEE Transactions on Biomedical Engineering, à soumettre

International conference papers

1. A. Halimi, H. Batatia, J.L. Digabel, G. Josse and J.Y. Tourneret, “ An unsupervised Bayesian approach for the joint reconstruction and classification of cutaneous reflectance confocal microscopy images ” 25th European Signal Processing Conference (EUSIPCO), Kos island, Greece, pp.251-255, 2017.
2. A. Halimi, H. Batatia, J.L. Digabel, G. Josse and J.Y. Tourneret, “ Statistical modeling and classification of reflectance confocal microscopy images ” Computational Advances in Multi-Sensor Adaptive Processing Conf. (CAMSAP), Curaçao, Dutch Antilles 2017, accepté.

Technical reports

1. A. Halimi, H. Batatia, J. L.Digabel, G. Josse, and J.-Y. Tourneret, Technical report associated

- with the paper “ An unsupervised Bayesian approach for the joint reconstruction and classification of cutaneous reflectance confocal microscopy images ”, University of Toulouse, France, Tech. Rep., Feb 2017. [Online]. Available: <https://arxiv.org/pdf/1703.01444.pdf>
2. A. Halimi, H. Batatia, J. L.Digabel, G. Josse, and J.-Y. Tourneret, Technical report associated with the paper “ Statistical modeling and classification of reflectance confocal microscopy images ”, University of Toulouse, France, Tech. Rep., April 2017. [Online]. Available: <https://arxiv.org/pdf/1707.00647.pdf>
 3. A. Halimi, H. Batatia, J L.Digabel, G. Josse, and J.-Y. Tourneret, 'Technical report associated with the paper “ Wavelet-based statistical classification of skin images acquired with reflectance confocal microscopy ”, Tech. Rep., University of Toulouse, France, Feb 2017.

CHAPITRE 1

Contexte médical et microscopie confocale à réflectance

Sommaire

1.1	Introduction	7
1.2	La peau	8
1.3	Composition de la peau	9
1.3.1	L'épiderme	10
1.3.2	La jonction dermo-épidermique	12
1.3.3	Le derme	13
1.3.4	L'hypoderme	14
1.4	Fonctions principales de la peau	15
1.4.1	Protection	15
1.4.2	Fonction sensorielle	15
1.4.3	Thermorégulation	15
1.4.4	Fonction métabolique: synthèse de la vitamine D	15
1.5	Pigmentation de la peau	15
1.5.1	Pigmentation mélanique	15
1.5.2	Hyperpigmentations	16
1.6	Techniques d'imagerie de la peau	18
1.6.1	Imagerie ultrasonore	19
1.6.2	Imagerie par résonance magnétique	20
1.6.3	Tomographie par cohérence optique	22
1.6.4	Microscopie confocale à réflectance MCR	23
1.7	Conclusions	28

1.1 Introduction

La peau humaine est un organe complexe, vaste et lourd du corps humain et qui peut être soumis à un certain nombre de maladies. C'est un organe complexe ayant 4 fonctions essentielles : une fonction de

protection, une fonction de sensibilité, une fonction métabolique et une fonction de thermorégulation. C'est le plus gros organe de l'être humain, représentant environ 8% de la masse corporelle et une surface de l'ordre de $2 m^2$ chez un adulte [Marieb., 1999]. On trouve également dans la peau des annexes cutanées telles que les poils, les glandes sébacées qui produisent le sébum (substance grasse qui protège la peau), les glandes sudoripares à l'origine de la sueur [C. Prost-Squarcioni and Fraitag, 2005, Prost-Squarcioni, 2006] et des récepteurs nerveux permettant les sensations tactiles et de pression. La microscopie confocale à réflectance (MCR) est une technique d'imagerie non-invasive qui permet une visualisation *in vivo* de l'épiderme jusqu'au derme papillaire en temps réel [Hofmann-Wellenhof et al., 2009, Nehal et al., 2008]. Le terme *in vivo* est une expression latine qualifiant des recherches ou des examens pratiqués sur un organisme vivant, par opposition à *in vitro* (en dehors de l'organisme vivant). De par ses caractéristiques, la MCR se situe entre l'examen dermatoscopique et l'examen histologique classique. En effet, son caractère totalement non-invasif, donc indolore, la possibilité de répéter l'examen sur une même zone cutanée de façon illimitée et l'orientation horizontale des images obtenues la rapprochent de la dermatoscopie; en revanche, le grossissement et la résolution des images la rapprochent de l'examen histologique classique [Kanitakis et al., 2013]. Son potentiel pour améliorer la détection du cancer et des tumeurs a été démontré dans diverses recherches et études cliniques dermatologiques [Alarcon et al., 2014b]. Dans [Menge et al., 2016], une corrélation entre la MCR et l'histologie a été signalée pour le diagnostic du mélanome. La MCR s'est également révélée utile pour le suivi et la surveillance du traitement du lentigo malin [Alarcon et al., 2014a, Champin et al., 2014, Guitera et al., 2014], et la chirurgie cutanée guidée [Hibler et al., 2015].

1.2 La peau

Généralités

La peau est l'enveloppe du corps. Elle est aussi une barrière physique à la plupart des microorganismes, à l'eau, et à une grande partie des rayons UV. L'acidité (pH 4,0 à 6,8) de la surface de la peau empêche le développement de la plupart des pathogènes. La peau protège le corps de la déshydratation dans les environnements secs et empêche l'absorption de l'eau lorsque le corps est immergé dans

l'eau. Les effets antagonistes du frisson et de la sudation, ainsi que la vasoconstriction et la vasodilatation des vaisseaux de la peau, permettent le maintien d'une température corporelle normale de 37°C . La peau absorbe de petites quantités de rayonnements UV nécessaires à la synthèse de la vitamine D.

1.3 Composition de la peau

La structure de la peau est complexe (Figure 1.1). Elle se subdivise en 3 régions superposées qui vont de la superficie vers la profondeur : l'épiderme (tissu d'origine épithéliale), le derme (un tissu conjonctif) et l'hypoderme. [Fuchs and Raghavan, 2002]. C'est la première ligne de défense de l'organisme contre l'environnement extérieur (micro-organismes, radiations).

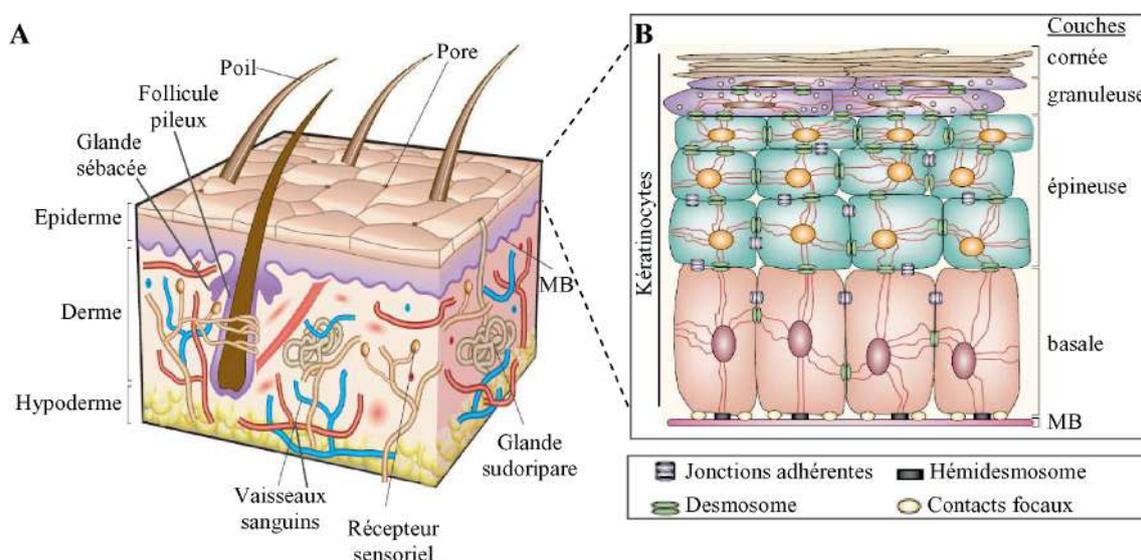


Figure 1.1: Représentation schématique de la peau humaine [Sage, 2012].

A) Illustration des différentes couches de la peau et des annexes cutanées (modifié de [MacNeil, 2007]). B) Représentation des différentes couches de cellules qui forment l'épiderme (modifié de [Fuchs and Raghavan, 2002])

Les cellules de l'épiderme forment des crêtes épidermiques qui se projettent dans le derme, et ce dernier présente également des projections dans l'épiderme, appelées papilles dermiques. Ces structures permettent d'augmenter la surface de contact et d'adhésion entre les deux compartiments. L'épiderme est séparé du derme par une membrane basale, nommée jonction dermo-épidermique.

L'hypoderme, dans le prolongement du derme, assure l'interface entre la peau et les tissus qu'elle recouvre (muscles, os).

1.3.1 L'épiderme

L'épiderme est un épithélium de revêtement. Il n'est pas vascularisé mais contient des terminaisons nerveuses sensibles. Il est constitué de quatre types cellulaires (Figure 1.2) : les kératinocytes, les mélanocytes, les cellules de Langerhans, et les cellules de Merkel [Prost-Squarcioni, 2006].

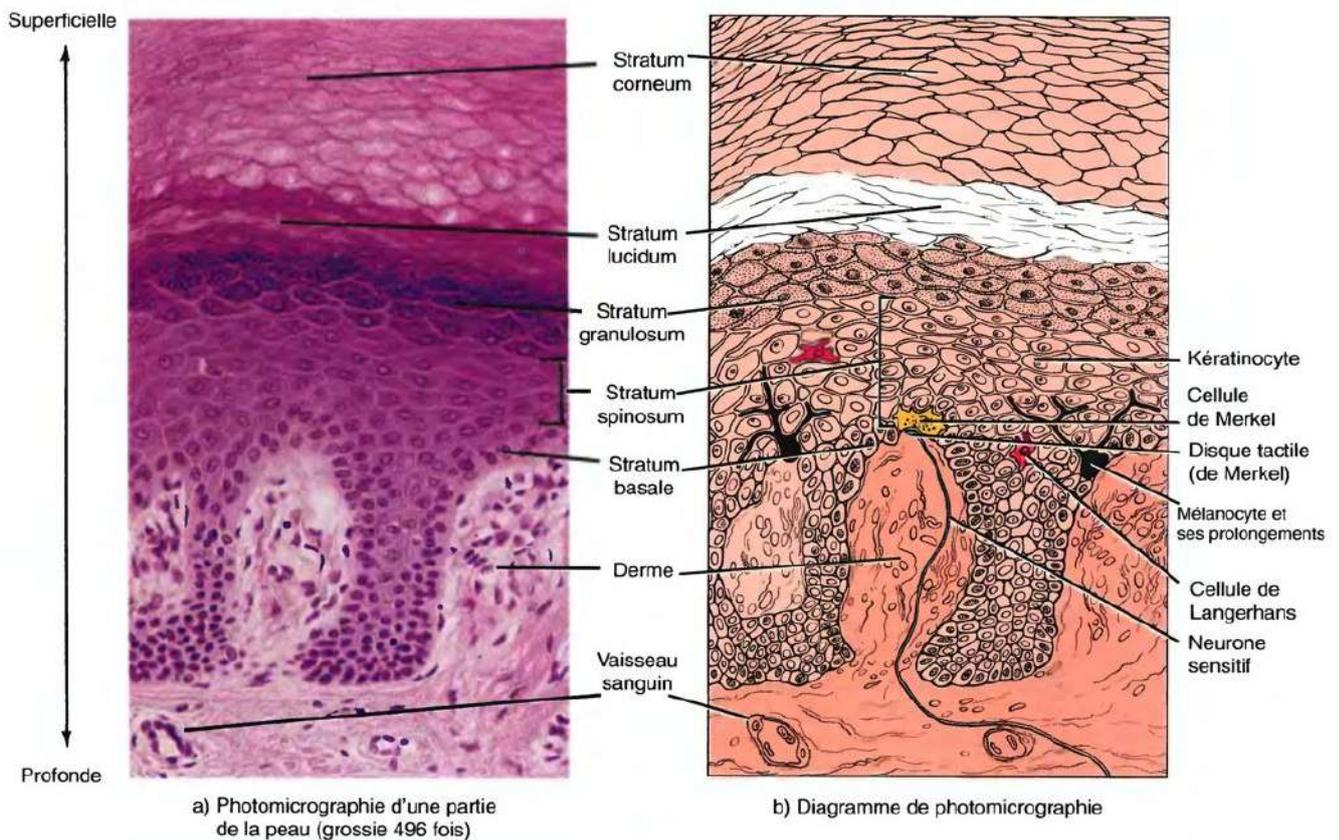


Figure 1.2: Types de cellules et couches de l'épiderme d'une peau épaisse. Les caractéristiques d'une peau mince sont semblables, toutefois le stratum lucidum est absent et le stratum corneum est plus mince [Tortora and Grabowski, 1993]

Les cellules de l'épiderme

- **Les kératinocytes** de l'épiderme qui représentent 80% des cellules de l'épiderme se répartissent dans 4 couches qui sont bien visibles en microscopie optique et appelées (de la profondeur à la surface) : couche basale (jonction dermo-épidermique), couche épineuse, couche granuleuse et couche cornée (Figure 1.1B). Les kératinocytes produisent de la kératine, une protéine fibreuse qui confère aux cellules de l'épiderme leurs propriétés protectrices. Les kératinocytes assurent trois grandes fonctions: la cohésion de l'épiderme et sa protection contre les agressions mécaniques, une fonction de barrière entre les milieux intérieur et extérieur et la protection contre les radiations lumineuses. Les kératinocytes sont étroitement liés les uns aux autres par des desmosomes. Ils proviennent de cellules qui se divisent de manière quasi continue et se situent au niveau de la couche la plus profonde de l'épiderme. A mesure que les kératinocytes sont poussés vers la surface de la peau par les nouvelles cellules, ils se mettent à produire de la kératine molle qui va devenir leur constituant majeur. Les kératinocytes meurent au cours de leur trajet jusqu'à la surface de la peau pour ne laisser plus qu'une membrane plasmique remplie de kératine qui tombent par desquamation. La peau se renouvelle en moyenne tous les 25 à 45 jours,
- **Les mélanocytes** sont moins nombreux que les kératinocytes. Ce sont des cellules épithéliales étoilées qui synthétisent un pigment, la mélanine. On les trouve dans les couches profondes de l'épiderme. La mélanine nouvellement synthétisée s'accumule dans des granules appelés mélanosomes qui sont acheminés par des protéines motrices le long des filaments d'actine pour rejoindre les extrémités des mélanocytes où ils sont absorbés par les kératinocytes avoisinants. Les mélanosomes s'accumulent près du noyau des kératinocytes, vers le milieu extérieur, et forment ainsi un bouclier pigmentaire qui protège le noyau contre les radiations ultraviolets (UV),
- **Les cellules de Langerhans** sont des cellules de l'immunité. Elles représentent 3 à 8% des cellules épidermiques. Elles contribuent à l'activation d'autres cellules de l'immunité et leur nombre augmente lors d'une infection chronique de la peau,

- *Les cellules de Merkel* sont, d'une part, liées à la terminaison nerveuse d'une neurofibre sensitive (disque de Merkel) et d'autre part, associées aux kératinocytes voisins par des microvillosités. Elles constituent la population cellulaire minoritaire de l'épiderme. Elles sont relativement abondantes au niveau des lèvres, des paumes et du dos des pieds. Avec les disques de Merkel, elles jouent le rôle de perceuteur sensoriel du toucher.

Les couches de l'épiderme

- *La couche basale, la plus profonde* est fixée au derme sous-jacent. elle est composée d'une couche unique de cellules, les kératinocytes, qui se multiplient rapidement et dont le rôle est de renouveler le contingent des cellules cutanées. On trouve aussi dans cette couche les cellules mélanocytaires, ou mélanocytes,
- *La couche épineuse* est constituée de la superposition de 4 à 5 épaisseurs de kératinocytes liés entre eux. Cet accrochage étroit favorise la fonction d'imperméabilité de la peau,
- *La couche granuleuse* est constituée de trois à cinq couches de cellules dans lesquelles les keratinocytes changent d'aspect. A ce stade, les cellules remplies de kératine se rapprochent de la surface de la peau et commencent à dégénérer et à se déstructurer (Figure 1.3),
- *La couche cornée* est la couche la plus superficielle de la peau. Les cellules devenues plates et translucides sont mortes et forment une couche résistante et imperméable.

1.3.2 La jonction dermo-épidermique

La jonction dermo-épidermique appelée aussi membrane basale épidermique est une structure complexe séparant l'épiderme du derme [Briggaman, 1982, Stanley et al., 1982], élaborée à la fois par les kératinocytes basaux et les fibroblastes dermiques. Les crêtes épidermiques (saillies de l'épiderme dans le derme) et les papilles dermiques (saillies du derme dans l'épiderme) alternent dans la jonction épidermique (Figure 1.4). Elle peut être divisée en 4 zones allant de l'épiderme vers le derme :

- *la membrane plasmique des kératinocytes basaux* avec leur structure d'attache : les hémidesmosomes,

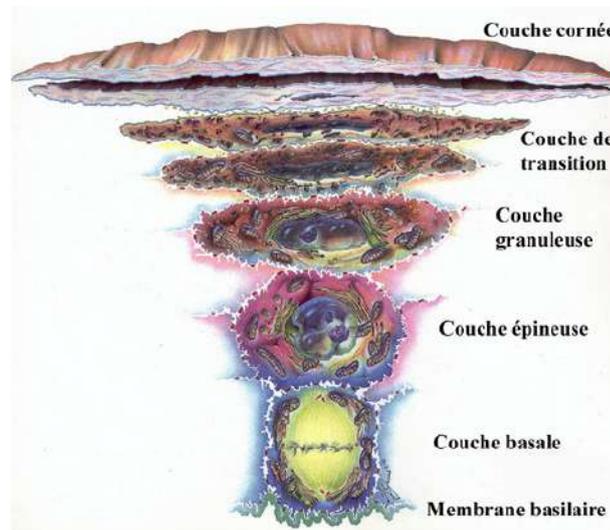


Figure 1.3: Les couches de l'épiderme. L'épiderme est constitué de plusieurs couches cellulaires qui, de la plus profonde à la plus superficielle, sont les suivantes : la couche basale, la couche épineuse, la couche granuleuse, la couche de transition (seulement dans la peau épaisse) et la couche cornée (modifié de [Geras, 1990],[Gagnon, 2005])

- *la lamina lucida* traversée par des filaments d'ancrage,
- *la lamina densa* formant la zone d'ancrage des filaments et des fibres issus de l'épiderme et de la zone fibrillaire,
- *la zone fibrillaire* qui contient des fibres d'ancrages reliant la lamina densa de la membrane basale à des plaques d'ancrage dans le derme papillaire.

La jonction dermo-épidermique assure dans la peau des fonctions fondamentales [Laurent, 2005]: elle possède un rôle de support mécanique pour l'adhésion de l'épiderme au derme et un rôle de barrière sélective permettant le contrôle des échanges moléculaires et cellulaires entre les deux compartiments.

1.3.3 Le derme

Le derme est un tissu de soutien de l'épiderme. Contrairement à l'épiderme, le derme est vascularisé, ce qui lui permet non seulement de fournir de l'énergie et des nutriments à l'épiderme, mais aussi de

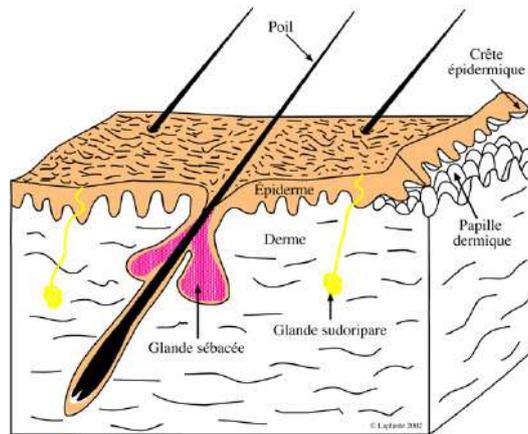


Figure 1.4: Schéma de la peau. Les crêtes épidermiques et les papilles dermiques alternent dans la jonction dermo-épidermique [Laplante, 2002]

jouer un rôle primordial dans la thermorégulation du corps et dans la guérison des blessures. Son épaisseur est d'environ 1mm à 4mm et varie selon les zones du corps. Il est formé de deux zones : la zone papillaire formée de tissu conjonctif lâche et la zone réticulaire formée d'un tissu conjonctif dense. Les cellules principales du derme sont les fibroblastes. Ils sont situés principalement dans les papilles cutanées, près de l'épiderme, et ne se trouvent qu'en faible quantité dans les couches profondes du derme. Ils sont spécialisés dans la production de deux types de fibres protéiniques, soit le collagène et l'élastine, qui forment le réseau fibreux du derme. Les fibres de collagène, qui constituent des protéines dans le derme, donnent à celui-ci sa résistance à la tension et à la traction, alors que l'élastine lui fournit ses propriétés élastiques.

1.3.4 L'hypoderme

L'hypoderme est un tissu adipeux séparant le derme et les plans aponévrotiques et musculaires sous-jacents. Les cellules adipeuses qui le constituent sont des cellules mésenchymateuses différenciées, rondes ou polygonales, possédant un noyau triangulaire périphérique et un cytoplasme riche en triglycérides et en acides gras. Outre son rôle énergétique dû à la réserve de graisse qui le constitue, l'hypoderme protège contre les chocs (mécanique et thermique).

1.4 Fonctions principales de la peau

1.4.1 Protection

La peau est soumise aux multiples agressions du milieu extérieur et se comporte comme un bouclier contre les agressions mécaniques (chocs, écrasements, tractions, etc.), qu'elle doit amortir et absorber afin de préserver son intégrité et celle des tissus et organes qu'elle enveloppe. Elle nous protège aussi contre les pollutions diverses, contre les micro-organismes et du soleil [Ratovoson, 2011, Tran, 2007].

1.4.2 Fonction sensorielle

La peau contient les récepteurs du sens du toucher qui captent toutes sortes de renseignements provenant du monde extérieur. Un rôle d'information qui ajoute une aide indirecte à la défense de l'organisme.

1.4.3 Thermorégulation

On appelle thermorégulation l'ensemble des mécanismes qui permettent à l'homme de maintenir constante sa température interne autour de 37°C . La peau nous protège de la chaleur grâce à la sueur qui a pour rôle d'éliminer la chaleur excédentaire de notre corps.

1.4.4 Fonction métabolique: synthèse de la vitamine D

C'est dans l'épiderme qu'est synthétisée la vitamine D, sous l'effet des UV du soleil. Un déficit en vitamine D a pour conséquence des troubles de croissance chez l'enfant (rachitisme) et une augmentation du risque de fractures chez l'adulte (ostéomalacie).

1.5 Pigmentation de la peau

1.5.1 Pigmentation mélanique

La couleur de la peau est normalement définie par le phototype d'une personne et par son exposition solaire, les zones fréquemment exposées au soleil étant plus sombres que les zones cachées. Le pigment naturel qui détermine la couleur de la peau est la mélanine, que l'organisme sécrète pour protéger la peau du rayonnement UV, ce qui donne le bronzage. Les troubles de la pigmentation sont liés à un

mauvais fonctionnement de la mélanine qui provoque soit une hyperpigmentation (taches sombres) soit une dépigmentation de la peau (taches plus claires). Dans ce dernier cas, il s'agit généralement d'une maladie, comme le vitiligo ou l'albinisme, qui relève d'une prise en charge médicale.

1.5.2 Hyperpigmentations

Les hyperpigmentations résultent d'une accumulation ou d'une mauvaise répartition du pigment au niveau de la peau. Elles prennent différentes formes [Chaouat, 2011, Oualid, 2011]:

Le melasma

Le mélasma est une affection fréquente de cause inconnue. C'est une affection bénigne de la peau se présentant sous la forme de taches hyperpigmentées apparaissant sur les zones exposées au soleil, surtout au niveau du visage, du décolleté et du cou. Le mélasma atteint principalement les femmes (Figure 1.5).



Figure 1.5: Melasma (http://www.asarchcenter.com/proced_kb/melasma/)

Les éphélides ou taches de rousseur

Fréquentes chez les sujets roux ou blonds, les taches de rousseur sont des manifestations cutanées physiologiques, isolées, dénuées de signification pathologique. Elles apparaissent comme des hyperpigmentations maculeuses, localisées, de petite taille (1 à 3 mm), de teinte brun clair ou ocre, siégeant au niveau du visage (nez, joues) et sur les zones photo-exposées (dos des mains, décolleté, haut du dos).

Les taches café au lait

Les taches café-au-lait se présentent comme des macules brunes, de couleur homogène, dont la taille peut varier de quelques centimètres à plusieurs dizaines de centimètres. Histologiquement, elles sont semblables aux éphélides, en rapport avec une hyperméláninose (Figure 1.6). Ces lésions, le plus souvent présentes à la naissance, sont en général isolées.



Figure 1.6: Taches café au lait [Oualid, 2011]

Naevus naevo-cellulaires

Egalement appelés naevus mélanocytaires, pigmentaires ou communément « grains de beauté », ils sont banals chez les sujets caucasiens. Ces naevus sont des lésions rondes ou ovales, planes ou légèrement en relief, de coloration homogène, brun clair à brun noir. Ils peuvent être présents à la naissance (naevus congénitaux) ou apparaître dans l'enfance avec des poussées à la puberté.

Le lentigo

Le lentigo est un type de lésion qui provient de la jonction entre le derme et l'épiderme en raison d'une forte concentration de mélanocytes au niveau des parois des papilles dermiques. Cette lésion conduit à la destruction du réseau cellulaire régulier à certaines couches de l'épiderme, principalement à la jonction dermo-épidermique [Menge et al., 2016]. Les lentigos sont des macules de petites tailles, bien limitées, de teinte brun foncé à noir. Très fréquents, ils affectent l'enfant et l'adulte et touchent plus fréquemment les personnes à peau claire. Ces taches pigmentaires sont sans danger mais témoignent

d'une altération de la peau liée au soleil. Histologiquement, les lentigos correspondent à une accumulation de mélanine, pigment qui colore la peau et la protège des ultraviolets, dans l'épiderme. Les deux principaux types de lentigos sont le lentigo simplex et le lentigo actinique.

- ***lentigo simplex*** Le terme lentigo simplex désigne tout lentigo survenant indépendamment de l'exposition solaire. Le lentigo simplex apparaît dans les 10 premières années de vie. Il se présente sous forme de petites macules brunes ou noires, isolées et réparties sur toute la peau, les muqueuses ou encore les ongles (Figure 1.7).



Figure 1.7: Lentigo simplex [Oualid, 2011]

- ***lentigo actinique*** ou lentigo solaire se retrouve sur les zones exposées au soleil (dos des mains, visages, bas des jambes ..). Il se traduit par des taches plus grandes que le lentigo simplex, de couleur brun clair. Il est très fréquent chez la population caucasienne de plus de 60 ans et est corrélé au vieillissement cutané induit par le soleil (Figure 1.8).

1.6 Techniques d'imagerie de la peau

En dermatologie, la décision médicale est basée sur l'examen clinique où l'inspection et la palpation permettent souvent de porter un diagnostic. En cas de doute, on pratique une biopsie qui reste toutefois un examen très local, invasif et ne peut être répété. On peut donc légitimement souhaiter disposer d'outils complémentaires (techniques non-invasives d'imagerie) qui peuvent offrir une vue en profondeur de la peau. Ces outils peuvent provenir des techniques non-invasives d'imagerie, comme



Figure 1.8: Lentigo actinique [Oualid, 2011]

l'Imagerie par Résonance Magnétique (IRM), l'Echographie, la Tomographie par Cohérence Optique, et la Microscopie Confocale, dont les deux paramètres importants sont la résolution et la profondeur de pénétration. Ces techniques d'imagerie vont être présentées ci-dessous afin de nous aider à mieux comprendre leur principe et leurs applications sur l'imagerie de la peau humaine *in vivo* (expression latine qualifiant des recherches ou des examens pratiqués sur un organisme vivant).

1.6.1 Imagerie ultrasonore

Les premières études sur l'échographie de la peau datent de 1979 [Alexander and Miller, 1979]. Les auteurs ont réalisé des coupes de la peau en mode A (A-scan), permettant de détecter des échos provenant de la surface de la peau et du derme. L'échographie mesure l'intensité et le temps de propagation d'une onde acoustique ultrasonore réfléchi par les structures de l'échantillon. Les appareils d'échographie font appel à des sources d'énergie (des ondes sonores de haute fréquence, ou ultrasons). Les impulsions ultrasonores qui traversent le corps sont réfléchies et déviées par les divers types de tissus [Juin, 2014]. A partir des échos ainsi produits, un ordinateur reconstruit des images du contour des organes examinés [Mcheik, 2010, Tran, 2007]. En échographie clinique humaine et vétérinaire, les fréquences utilisées oscillent entre 2 MHz et 15 MHz. L'échographie haute résolution utilise par définition des fréquences supérieures à 15 MHz (20-100 MHz). L'utilisation d'ondes sonores de fréquence plus élevée, une sensibilité accrue au niveau de la réception, une meilleure analyse des signaux, etc.,

permettent d'obtenir maintenant des images de qualité supérieure, en trois dimensions, et de visualiser des mouvements en temps réel. Plusieurs modes d'acquisition sont actuellement utilisés en échographie cutanée haute résolution [Naouri, 2008] :

- le mode A (amplitude) est un mode archaïque de représentation, où à chaque écho est attribué un pic d'amplitude.
- Le mode B (brillance) fournit une image bidimensionnelle. Les échos sont représentés sous forme de points dont la brillance est proportionnelle à l'énergie reçue. Il permet ainsi d'obtenir *in vivo* et en temps réel, l'image d'une coupe transversale de peau dans l'axe de la sonde (Figure 1.9).
- Le mode C permet de reconstituer une image tridimensionnelle de la lésion étudiée.

L'échographie cutanée haute résolution permet de différencier des structures dont l'écart est inférieur à $100\mu\text{m}$ sur l'axe du faisceau (résolution axiale) et $200\mu\text{m}$ sur l'axe de balayage (résolution latérale).

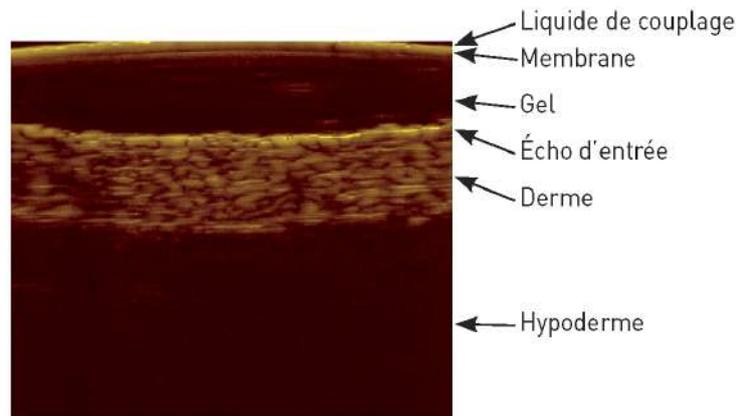


Figure 1.9: Image échographique de la peau avec une sonde de 20 MHz - Mode B [Naouri, 2008]

1.6.2 Imagerie par résonance magnétique

L'imagerie par résonance magnétique nucléaire (IRM), est extrêmement intéressante parce qu'elle produit des images très contrastées des tissus mous, pour lesquels les techniques utilisant les rayons X ne sont pas d'une grande utilité [Osseni, 2010]. Sous sa forme originale, la résonance magnétique

nucléaire donne avant tout une image de l'hydrogène, le plus petit des atomes ; dans notre organisme, la plus grande partie de l'hydrogène fait partie des molécules d'eau et des molécules de lipides. On applique à l'organisme des champs magnétiques ayant de 10000 à 60000 fois l'intensité du magnétisme terrestre qui varie de $30\mu T$ à $60\mu T$ (microTesla) selon la localisation sur la terre. Le patient est étendu à l'intérieur d'un espace entouré d'un aimant. Les molécules d'hydrogène se comportent comme de petits aimants et tournent comme des toupies dans le champ magnétique. Ils s'orientent alors dans une même direction parallèle au champ magnétique appliqué; des émissions très brèves d'ondes radio sont alors appliquées, créant un second champ magnétique perpendiculaire au premier et ayant pour effet de modifier cette orientation (c'est la « résonance magnétique »). Lorsque l'émission des ondes radio cesse, les molécules contenant de l'hydrogène retournent à leur orientation initiale et l'énergie libérée est transformée en image [Decoupigny, 2011, Selb, 2002]. L'IRM est une technique qui utilise un rayonnement non ionisant, permet d'avoir une forte résolution spatiale ($25 - 100\mu m$) et un bon contraste des tissus mous [Kozłowska et al., 2009]. La technique peut aussi fournir des informations sur le fonctionnement de certains organes [Glunde et al., 2007], et est également utilisée pour le diagnostic des maladies [Rudin and Weissleder, 2003]. La Figure 1.10 représente l'anatomie de la peau par IRM.

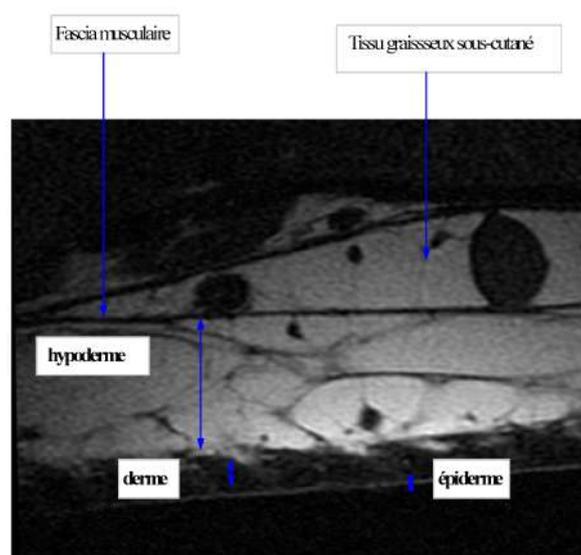


Figure 1.10: Anatomie de la peau par IRM (séquence FIESTA)

<http://pe.sfrnet.org/Data/ModuleConsultationPoster/pdf/2004/1/d958caf4-945c-4174-b542-f1cf3157c06c.pdf>

1.6.3 Tomographie par cohérence optique

La Tomographie par Cohérence Optique (OCT) permet de produire de manière non invasive des images tomographiques d'échantillons biologiques. C'est au début des années 1990 que les premières applications de la tomographie à cohérence optique en mode 2B (B-scan) ont lieu avec les travaux de [Huang et al., 1991] sur un oeil humain, *in vitro*, et de [Fercher et al., 1993] cette fois-ci *in vivo*. Par analogie avec l'échographie, le signal obtenu en OCT dépend de l'amplitude et du chemin optique parcourue par une onde électromagnétique lumineuse réfléchi et/ou rétrodiffusée par le milieu biologique. Le contraste des images en OCT résulte des inhomogénéités d'indice de réfraction du milieu. la mesure en OCT ne peut se faire que par corrélation de l'onde revenue avec une onde de référence car les fréquences optiques (de l'ordre de 10^{14} Hz) sont très élevées par rapport aux fréquences acoustiques. La corrélation est réalisée par interférométrie [Juin, 2014, Nahas, 2014, Ouadour-abbbar, 2009]. L'OCT possède une profondeur de pénétration et des résolutions spatiales situées entre les techniques non optiques (échographie, rayons X et IRM) possédant de faibles résolutions spatiales et la microscopie confocale, possédant une résolution élevée mais une profondeur de pénétration très limitée (Figure 1.11).

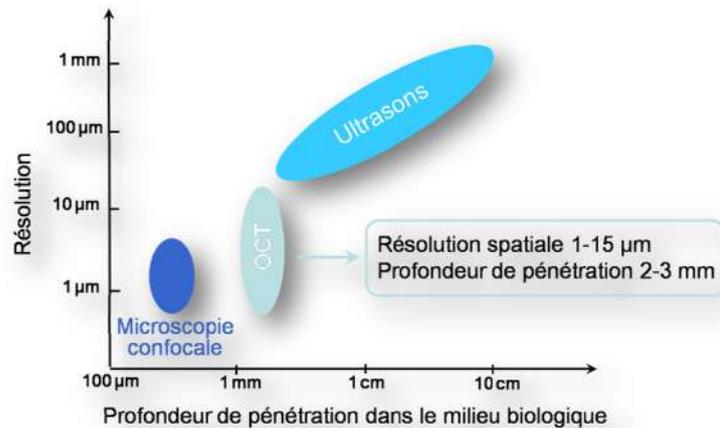


Figure 1.11: Résolution et profondeur de pénétration comparatives des ultrasons, de la microscopie confocale et de la tomographie par cohérence optique [Sacchet, 2010]

1.6.4 Microscopie confocale à réflectance MCR

La technique d'imagerie confocale qui a été décrite pour la première fois par M. Minsky en 1957 [Minsky, 1988] permet d'obtenir l'image d'un spécimen en trois dimensions. Le terme de réflectance traduit le fait que c'est la réflexion du rayon laser sur les différentes molécules contenues dans la peau, notamment la mélanine, la kératine et le collagène, qui produit les images [Kanitakis et al., 2013]. Dans son principe, le microscope confocal ressemble beaucoup à un microscope conventionnel, à l'exception d'un diaphragme placé devant le photodétecteur. Ce diaphragme élimine les faisceaux de lumière provenant des plans situés au-dessus et au-dessous du plan focal. En effet, l'une des limitations de la microscopie conventionnelle est la trop grande profondeur de champ de l'image. La zone de mise au point apparaît nette, mais les zones immédiatement au-dessus et en-dessous sont floues et perturbent l'image observée. De plus, un microscope conventionnel acquiert parallèlement tous les points du spécimen, alors qu'un microscope confocal acquiert l'image du spécimen par balayage, c'est-à-dire point par point [Fouard, 2005]. Cette technique de visualisation a l'avantage sur les technologies optiques classiques, de pouvoir effectuer une observation tridimensionnelle directe du spécimen analysé *in vivo* sans l'endommager. De plus, elle permet à l'expert de réaliser une acquisition 4D (3D spatiale + 1D temporelle) des données, facilitant l'observation complète de structures biologiques en déformation au cours du temps [Tran, 2007]. La Figure 1.12 représente le schéma optique du microscope confocal : un rayon lumineux émis par le laser vient éclairer un point du spécimen dans le plan focal de l'objectif, après avoir été dirigé par le miroir dichroïque (miroir qui réfléchit les rayons lumineux dont la longueur d'ondes se situe dans une certaine plage et se laisse traverser par les rayons lumineux dont la longueur d'onde n'appartient pas à cette portion de spectre). Le rayon réémis vient ensuite frapper un photodétecteur. Le microscope confocal acquiert ainsi un point du spécimen. Pour acquérir tout le spécimen, il effectue un balayage ligne à ligne, puis plan par plan.

Appareils de microscopie confocale

La microscopie confocale *in vivo* (MCiv) a été développée commercialement par la société LUCID avec un microscope confocal en réflectance à balayage laser, le Vivascope 1000 puis la nouvelle version, le Vivascope 1500 (Figure 1.13).

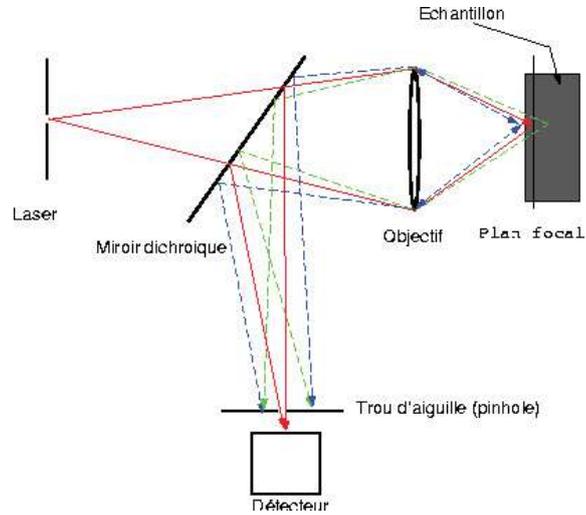


Figure 1.12: Principe de fonctionnement d'un microscope confocal (http://www3.jouy.inra.fr/miaj/public/imaste/Projets/AI2M_v2/node9.html)



Figure 1.13: *Vivascope 1500* (<http://www.vivascope.de/en/products/devices/in-vivo-devices/vivascope-1500/pictures.html>)

VivaScope 1500 - Données techniques

Résolution optique : horizontale: $<1.25 \mu\text{m}$ au centre du champ d'image ; verticale : $<5.0 \mu\text{m}$ au centre du champ d'image

Profondeur d'image : couches supérieures de la couche réticulaire

Section visualisable, image individuelle : $500 \mu\text{m} \times 500 \mu\text{m}$

Mouvement d'axe X & Y : $\pm 4.0 \text{ mm}$

Résolution d'image : 1000×1000 pixels (optimisé Nyquist)

Conversion numérique-analogique d'image : 8 bits

Cadence d'images : 9 images/seconde

Ecran de contrôle : 19", 1280×1024 pixels, écran plat couleur

Logiciel : VivaScan version 7

Puissance d'exploitation optique : CDRH classe 1, EU classe 1M

Longueur d'onde imagerie : 830 nm

Installation tête imagerie : une suspension à la Cardan fournit $\pm 90^\circ$ de rotation en direction orthogonale

Objectif inclus : Lucid StableView TM F.L. = 4.3 mm , 0.9 NA immersion

Plage de températures de service : 55°F à 85°F (13°C à 30°C)

Source d'alimentation : 110-230 VAC 50-60 Hz

Humidité d'exploitation : sans condensation

Certifications : FCC classe A, marqué CE

Le VivaScope 1500 offre aux chercheurs et aux utilisateurs dans les secteurs de la médecine et des cosmétiques, la possibilité de faire des examens optiques de la peau, en temps réel. Cela fournit une vue *in vivo*, non-invasive de l'épiderme et du derme, jusqu'à la couche réticulaire supérieure. Un laser infrarouge proche (830 nm) est dirigé sur les sections individuelles de la peau, où il est réfléchi. La mélanine et la kératine agissent comme des agents de contraste naturels en raison de leur indice de réfraction relativement élevé. L'appareil produit des images de la peau, en noir et blanc, dans une qualité optimale.

Applications médicales

Les applications médicales de la microscopie confocale *in vivo* sont multiples. En dermatologie, la MCR est utilisée pour réaliser par exemple des tests de dépistage de cancer (les mélanomes malins sont différents des naevus épidermiques "normaux" par de nombreux aspects, lorsqu'ils sont examinés par MCR) ou le contrôle des bords après l'excision des mélanomes malins. Cela permet de vérifier que toute la zone pathologique a bien été enlevée et permet donc de diminuer le risque de récurrence dans la zone du mélanome. La MCR est également utilisée dans le secteur de la recherche cosmétique, pour examiner la peau. Les recherches dans ce domaine concernent, par exemple, la pénétration et la diffusion des substances contenues dans les lotions, crèmes, etc... A titre d'exemple, la Figure 1.14 représente des images de microscopie confocale à réflectance prises à différentes profondeurs de la peau, la Figure 1.15 représente deux images MCR (patient sain et patient atteint de lentigo) prises au niveau de la jonction dermo-épidermique et enfin la Figure 1.16 représente la couche cornée d'une image MCR. La couche la plus superficielle de la peau apparaît brillante (réflétante). Les cornéocytes sont agencés en îlots séparés par des dépressions sombres (dermatoglyphes).

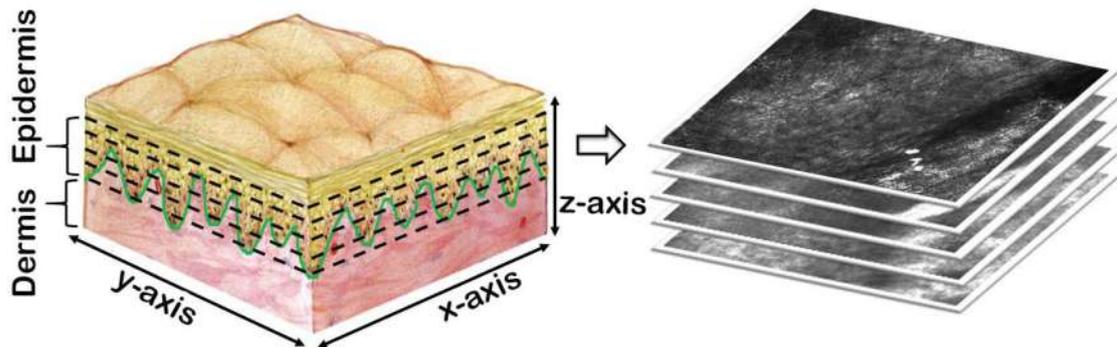


Figure 1.14: Images de microscopie confocale à réflectance prises à différentes profondeurs de la peau

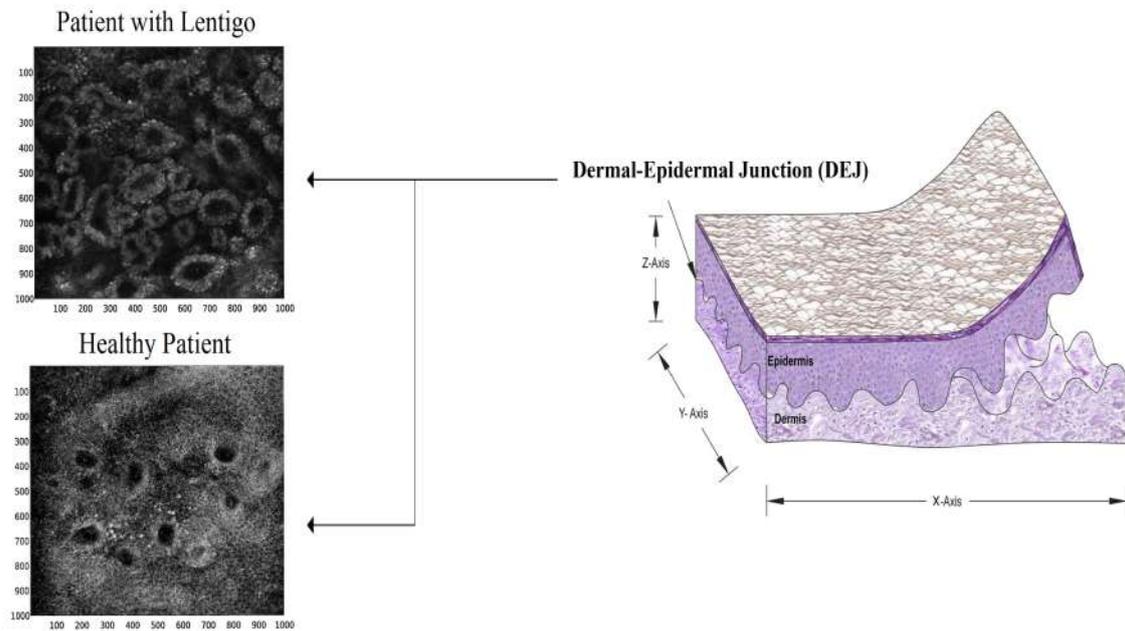


Figure 1.15: Images MCR (patient sain et patient atteint de lentigo) prises au niveau de la jonction dermo-épidermique

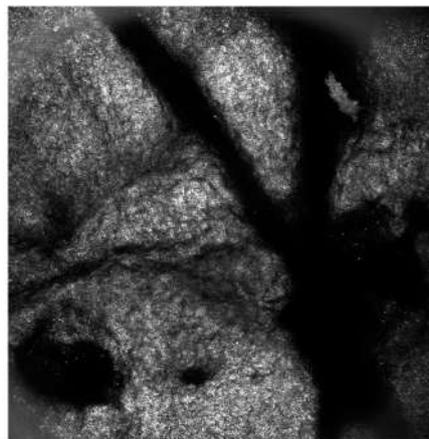


Figure 1.16: Couche cornée [Kanitakis et al., 2013]

1.7 Conclusions

Nous avons commencé par décrire dans ce chapitre quelques propriétés de la peau humaine. Nous avons ensuite exposé sa composition, ses différentes fonctions principales, et les types de pigmentation: mélanique ou hyperpigmentation. Les hyperpigmentations cutanées sont nombreuses et d'origine variée, entraînant une gêne esthétique plus ou moins importante et étant parfois mal supportée par les patients. Nous avons ensuite présenté plusieurs techniques d'imagerie médicale ainsi que leur principe de fonctionnement. Nous nous sommes principalement intéressés dans cette thèse à la microscopie confocale à réflectance qui est une technique d'imagerie non-invasive permettant une visualisation *in vivo* de l'épiderme jusqu'au derme papillaire en temps réel. Le but final est de caractériser ces images de microscopie confocale à réflectance afin de distinguer les patients sains de ceux atteints de lentigo qui est une hyperpigmentation cutanée.

CHAPITRE 2

Modélisation statistique et classification d'images de microscopie confocale à réflectance à l'aide d'une analyse multirésolution.

Sommaire

2.1	Introduction	29
2.2	Méthode proposée	32
2.2.1	Décomposition en ondelettes	32
2.2.2	Modélisation Statistique	33
2.2.3	Caractérisation utilisant le test paramétrique de Student et la pseudo-distance de Kullback-Leibler	36
2.2.4	Classification utilisant les machines à vecteurs support	41
2.3	Expériences	41
2.3.1	Données	41
2.3.2	Analyse statistique des coefficients d'ondelettes	42
2.3.3	Identification des profondeurs caractéristiques	45
2.3.4	Signification statistique avec le test T	54
2.3.5	Performance du classifieur SVM	58
2.4	Conclusions	61

2.1 Introduction

Dans ce chapitre, on s'intéresse à caractériser et modéliser les textures des images MCR à l'aide d'une analyse multirésolution afin d'identifier la profondeur de la peau à laquelle on peut distinguer une peau saine de celle atteinte de lentigo.

La nature complexe des images MCR nécessite des méthodes automatiques de traitement d'images pour élaborer des stratégies de diagnostic précises. Peu de techniques automatiques sont disponibles

dans la littérature. La recherche subséquente sur les images MCR s'est principalement concentrée sur trois aspects : i) études cliniques (non abordées dans notre thèse) pour évaluer leur utilité, ii) segmentation des noyaux, et iii) classification des tissus de la peau.

Luck et al. [Luck et al., 2005] ont d'abord développé une méthode automatique de traitement d'images MCR pour segmenter les noyaux. Leur méthode est basée sur un modèle gaussien de l'image qui tient compte de la réflectivité des noyaux et d'une distribution gaussienne tronquée pour représenter l'intensité des fibres de cytoplasme. Un champ aléatoire gaussien de Markov a également été utilisé pour la corrélation spatiale, et un algorithme de classification bayésienne a été étudié pour étiqueter les tissus. Harris et al. [Harris et al., 2015] ont développé un algorithme basé sur les réseaux neuronaux pour segmenter les noyaux dans les images MCR. Les noyaux et arrière-plans, à partir d'images segmentées manuellement, ont été représentés par des distributions gaussiennes. Le réseau neuronal a été formé à l'aide de caractéristiques telles que la taille, le contraste et la densité du noyau. Diverses caractéristiques extraites des images MCR ont été utilisées pour les applications citées ci-dessus. Kurugol et al. [Kurugol et al., 2011, 2012] ont développé et validé une méthode semi-automatique pour localiser la jonction dermoépidermique (DEJ) à l'aide d'un classifieur statistique basé sur des caractéristiques de texture. Ces caractéristiques sont principalement liées aux mesures de luminosité associées aux cellules basales. Les auteurs de [Somoza et al., 2014] ont développé une méthode automatique pour localiser les couches de la peau dans les images MCR en fonction de l'analyse de la texture.

Hames et al. [Hames et al., 2015, 2016] ont développé un classifieur à base de régression logistique pour segmenter automatiquement les différentes couches de la peau dans les images MCR. Dans [Kose et al., 2016], une méthode de classification SVM basée sur les caractéristiques de texture SURF (Speeded Up Robust Features) a été utilisée pour identifier les motifs morphologiques de la peau dans les images MCR. Enfin, pour les applications de diagnostic, Koller et al. [Koller et al., 2011] ont étudié une méthode de classification par arbre de décision basée sur les ondelettes pour distinguer les tumeurs bénignes et malines des mélanocytes cutanées dans les images MCR. Cette méthode, sera utilisée comme référence dans notre étude.

Très peu de travaux de recherche disponibles dans la littérature ont porté sur la détermination des

marqueurs quantitatifs pour la caractérisation des tissus dans les images MCR. Parmi ces travaux, Raphael et al. [Raphael et al., 2013] ont étudié une méthode de caractérisation des images MCR pour évaluer le photovieillissement. Dans leur étude, les intensités des images, leurs coefficients d'ondelettes 2D, leurs coefficients de Fourier 2D et les formes, segmentés avec un algorithme adhoc, ont été corrélés avec les données cliniques. Les résultats obtenus dans [Raphael et al., 2013] ont montré que l'intensité de l'image et les coefficients d'ondelettes n'avaient aucune corrélation significative, contrairement aux coefficients de Fourier et aux résultats de segmentation.

La plupart des recherches concernant le traitement automatique des images MCR sont basés sur une sorte d'analyse de la texture de ces images. De nombreux algorithmes basés sur les distributions temps-fréquence et temps-échelle des images MCR ont montré des propriétés intéressantes pour l'analyse de la peau. En particulier, les coefficients d'ondelettes résultant d'une décomposition en ondelettes discrètes multirésolution ont été considérés pour la caractérisation et la classification des tissus dans de nombreuses applications pratiques [Dima et al., 2002, Hu et al., 2008, Patwardhan et al., 2003, Yuan et al., 2005]. Cependant, la caractérisation des images MCR en utilisant des modèles statistiques associés aux coefficients d'ondelettes n'a à notre connaissance jamais été étudiée. Étant donné que le lentigo apparaît visuellement comme une modification de la structure régulière de la texture de l'image MCR, le chapitre actuel étudie une caractérisation statistique des textures basée sur des modèles statistiques adaptés au domaine des ondelettes. Plus précisément, une décomposition en ondelettes bidimensionnelle est effectuée à quatre échelles pour chaque profondeur d'un *stack* d'images MCR. Une loi gaussienne généralisée est ensuite ajustée aux coefficients en ondelettes. Les paramètres de forme et d'échelle de cette loi sont considérés comme des marqueurs pour caractériser le lentigo. L'objectif est ensuite d'identifier la profondeur de la peau à laquelle cette caractérisation est la plus importante. Un test paramétrique de Student ainsi que la distance de Kulback-Leibler ont été mis en œuvre afin d'évaluer la signification statistique de la caractérisation. Une analyse quantitative soutenue par un classifieur basé sur une machine à vecteur de support (SVM) montre en particulier que le paramètre de forme de ces coefficients multi-résolution permet la distinction entre les tissus sains et les lésions aux profondeurs caractéristiques. Les résultats des expériences effectuées au cours d'une étude clinique montrent que le paramètre de forme est un bon indicateur pour distinguer le

lentigo des tissus sains à des profondeurs entre $50 \mu m$ et $60 \mu m$. Le reste du chapitre est organisé comme suit. Section 2.2 présente la méthode proposée pour la caractérisation du lentigo. Elle décrit le processus de décomposition pyramidal d'une analyse multi-résolution à l'aide d'ondelettes, le modèle statistique utilisé pour les coefficients d'ondelettes et la méthode statistique adoptée pour estimer les paramètres de ce modèle. Cette Section présente enfin la méthode de caractérisation de texture proposée basée sur les tests T et le facteur de Bayes. Les résultats expérimentaux et les discussions sont présentés dans la Section 2.3. Les conclusions et les perspectives pour les travaux futurs sont finalement rapportées dans la Section 2.4.

2.2 Méthode proposée

Cette Section présente l'approche proposée pour caractériser et classifier les images MCR saines et lentigo. Soit $\mathbf{y} = (y_1, \dots, y_V)$ le *stack* d'images acquis pour un patient donné, où l'image 2D y_v avec $v = 1, \dots, V$ correspond à une profondeur fixe de la peau (mesurée en μm). La méthode d'analyse proposée se compose de quatre étapes qui sont résumées dans la Figure 2.1 et décrites dans les Sections suivantes.

2.2.1 Décomposition en ondelettes

Les ondelettes ont été utilisées avec succès pour la caractérisation de la texture en raison de leurs propriétés de multirésolution qui nous permettent de capturer différentes structures à plusieurs niveaux de détails et d'échelles. La décomposition en ondelettes d'une image bidimensionnelle peut être calculée selon les directions horizontale et verticale à l'aide d'un banc de filtres séparables (avec des réponses impulsionnelles de filtre passe-bas et passe-haut désignées par h et g) comme indiqué sur la Figure 2.2. Les images d'entrée sont traitées à l'aide de deux étapes de filtres d'analyse qui sont associés à des ondelettes orthogonales ou biorthogonales (ondelettes de Daubechies d'ordre 4). À chaque échelle, l'image est traitée par des filtres horizontaux, sous-échantillonnée par deux et suivie de filtres verticaux. Les coefficients obtenus représentent l'approximation de l'image MCR et les détails horizontaux (H), verticaux (V) et diagonaux (D). Dans cette étude, nous considérons les échelles associées à $K = 4$ niveaux de décomposition.

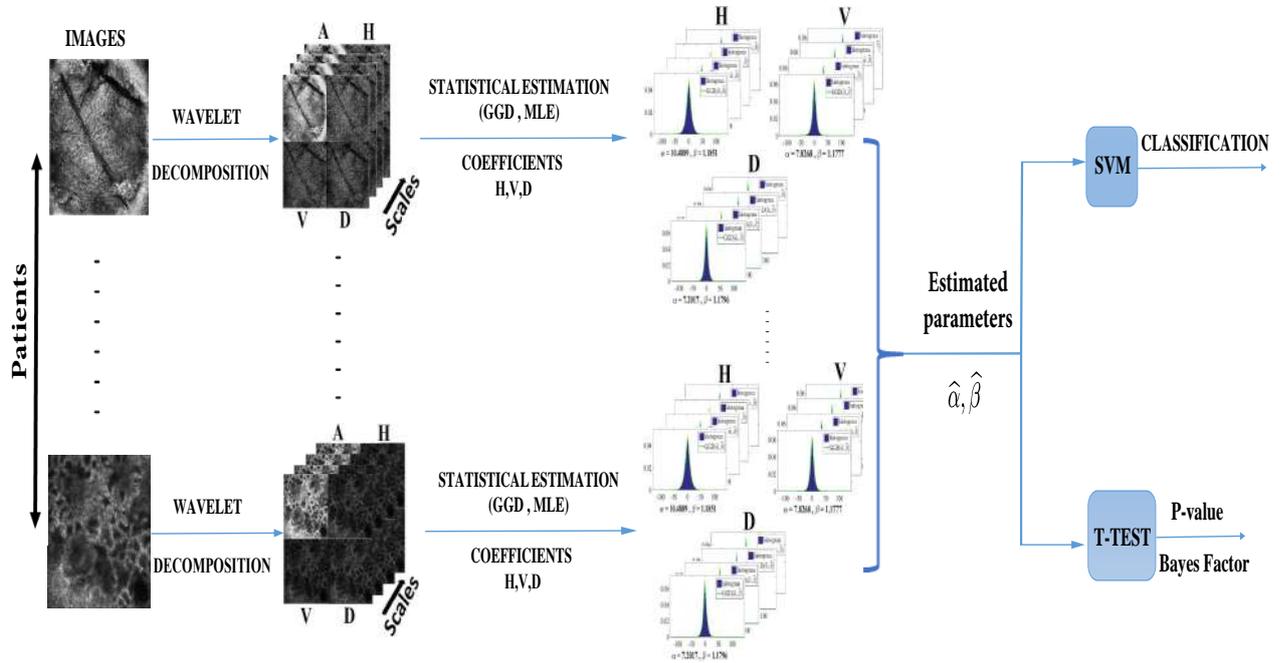


Figure 2.1: Étapes de la méthode de caractérisation proposée pour une profondeur donnée.

Soit $\mathbf{x}_i^{(d)}$ désignant le vecteur $L \times 1$ des coefficients ordonnés arbitrairement obtenus à l'échelle i pour la bande $d \in \{H, V, D, HVD\}$. On note que la bande HVD est la combinaison de tous les coefficients à partir des trois bandes de détails. Un exemple de décomposition au premier niveau d'une image MCR est représenté à la Figure 2.3.

2.2.2 Modélisation Statistique

Distribution gaussienne Généralisée

Dans cette sous-Section, nous proposons d'utiliser les propriétés statistiques des coefficients d'ondelettes $\mathbf{x}_i^{(d)}, \forall d \in \{H, V, D, HVD\}, \forall i \in \{1, \dots, K\}$ pour détecter la présence de lentigo dans les images RCM. Plus précisément, nous supposons comme dans [Do and Vetterli, 2002, P.Moulin and Liu, 1999, Wouwer et al., 1999, Yu et al., 2012] que la loi des coefficients d'ondelettes à chaque échelle est une loi

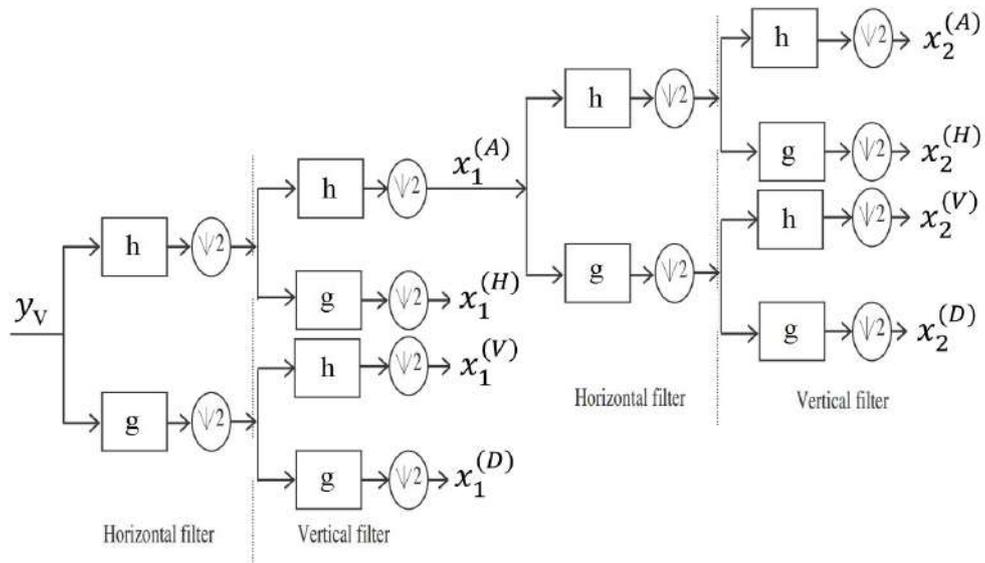


Figure 2.2: Schéma de décomposition en sous-bandes pour une image 2D-DWT à deux niveaux. y_v est une image 2D acquise à la profondeur v .

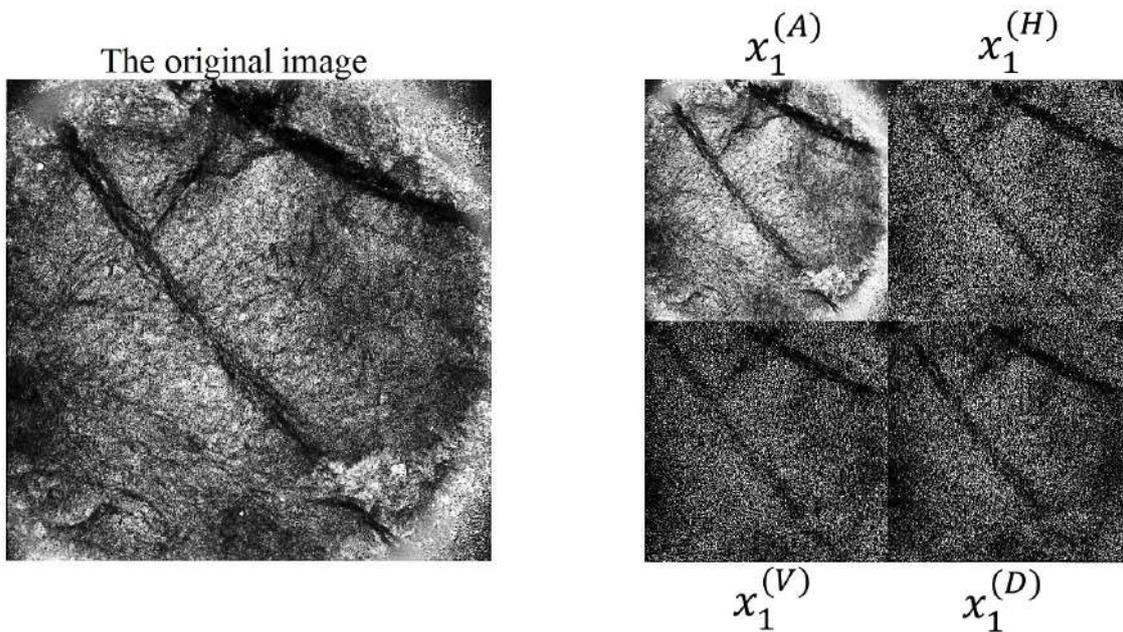


Figure 2.3: Exemple de décomposition en ondelettes pour une image MCR au premier niveau de décomposition.

gaussienne généralisée (GGD). Cette loi est connue pour capturer les propriétés statistiques des ondelettes pour une grande classe d'images [Do and Vetterli, 2002, Sharifi and Leon-Garcia, 1995], et son utilisation nous permet de réduire la taille de plusieurs milliers de coefficients à certains paramètres seulement (dans notre cas deux, les paramètres d'échelle et de forme) et de mieux saisir la variabilité de ces coefficients. La loi GGD a la densité de probabilité (ddp) suivante $p_{\alpha_i^{(d)}, \beta_i^{(d)}}(x_i^{(d)})$ qui dépend de deux paramètres $\alpha_i^{(d)} > 0$ et $\beta_i^{(d)} > 0$

$$p_{\alpha_i^{(d)}, \beta_i^{(d)}}(x_i^{(d)}) = \frac{\beta_i^{(d)}}{2 \alpha_i^{(d)} \Gamma\left(\frac{1}{\beta_i^{(d)}}\right)} \exp\left[-\left(\frac{|x_i^{(d)}|}{\alpha_i^{(d)}}\right)^{\beta_i^{(d)}}\right], \quad x_i^{(d)} \in \mathbb{R} \quad (2.1)$$

où $\Gamma(\cdot)$ est la fonction gamma. Le paramètre $\beta_i^{(d)}$ gère la forme de la loi GGD (qui est plus ou moins pointue) et est classiquement appelé **paramètre de forme**. Le paramètre $\alpha_i^{(d)}$ caractérise la largeur (ou la dispersion) de la densité et est désigné par **paramètre d'échelle**.

Nous montrerons dans la Section 2.3.2 que la densité (2.1) est parfaitement adaptée aux coefficients d'ondelettes associés à la texture des images MCR. La Section suivante présente une méthode d'estimation statistique qui a d'abord été étudiée dans [Do and Vetterli, 2002] permettant d'estimer les paramètres $\alpha_i^{(d)}$ et $\beta_i^{(d)}$ à partir des coefficients d'ondelettes $\mathbf{x}_i^{(d)}$.

Estimateur du maximum de vraisemblance

La méthode d'estimation du maximum de vraisemblance (MV) consiste à maximiser la vraisemblance des échantillons observés par rapport aux paramètres du modèle inconnu [Kay, 1993]. En supposant que les coefficients d'ondelettes associés aux différentes bandes et échelles sont indépendants, la fonction de vraisemblance de l'échantillon $\mathbf{x}_i^{(d)} = (x_{i1}^{(d)}, \dots, x_{iL}^{(d)})$ est définie par

$$L(\mathbf{x}_i^{(d)}; \alpha_i^{(d)}, \beta_i^{(d)}) = \prod_{l=1}^L p_{\alpha_i^{(d)}, \beta_i^{(d)}}[x_{i,l}^{(d)}] \quad (2.2)$$

où $\alpha_i^{(d)}$ et $\beta_i^{(d)}$ sont les paramètres d'échelle et de forme associés à la bande d et à l'échelle i . Ces paramètres seront notés par la suite α et β par souci de concision. Les dérivées partielles de la fonction

log-vrasemblance par rapport à α et β peuvent être calculées aisément, ce qui nous donne

$$\frac{\partial \log L(\mathbf{x}_i^{(d)}; \alpha, \beta)}{\partial \alpha} = -\frac{L}{\alpha} + \frac{\beta}{\alpha^{\beta+1}} \sum_{l=1}^L |x_{i,l}^{(d)}|^\beta \quad (2.3)$$

$$\frac{\partial \log L(\mathbf{x}_i^{(d)}; \alpha, \beta)}{\partial \beta} = \frac{L}{\beta} + L \frac{\psi\left(\frac{1}{\beta}\right)}{\beta^2} - \sum_{l=1}^L \left(\frac{|x_{i,l}^{(d)}|}{\alpha}\right)^\beta \log\left(\frac{|x_{i,l}^{(d)}|}{\alpha}\right) \quad (2.4)$$

où $\psi(z) = \Gamma'(z)/\Gamma(z)$ est la fonction digamma [Abramowitz and Stegun, 1970]. L'annulation de la première dérivée (2.3) nous donne

$$\alpha = \left(\frac{\beta}{L} \sum_{l=1}^L |x_{i,l}^{(d)}|^\beta\right)^{1/\beta}. \quad (2.5)$$

En substituant cette valeur de α dans (2.4) on obtient l'équation non-linéaire suivante pour β

$$1 + \frac{\psi(1/\beta)}{\beta} - \frac{\sum_{l=1}^L |x_{i,l}^{(d)}|^\beta \log |x_{i,l}^{(d)}|}{\sum_{l=1}^L |x_{i,l}^{(d)}|^\beta} + \frac{\log\left(\frac{\beta}{L} \sum_{l=1}^L |x_{i,l}^{(d)}|^\beta\right)}{\beta} = 0 \quad (2.6)$$

qui peut être résolue numériquement. Nous avons utilisé la méthode itérative de Newton Raphson [Kay, 1993] avec comme initialisation la méthode des moments décrite dans [Sharifi and Leon-Garcia, 1995] pour résoudre (2.6). La valeur estimée résultante de β est alors introduite dans (2.5) afin d'estimer α . La Section suivante étudie un test statistique en utilisant l'entropie et la variance des coefficients d'ondelettes et des paramètres estimés (α, β) pour discriminer les patients sains de lentigo.

2.2.3 Caractérisation utilisant le test paramétrique de Student et la pseudo-distance de Kullback-Leibler

En supposant que les paramètres α , β , entropie et variance (estimés pour tous les patients) suivent tous des distributions normales comme représentées dans la Figure 2.4, nous proposons alors d'appliquer le test paramétrique du T de Student (appelé T-test) ainsi que le Facteur de Bayes (BF) (Annexe A) à ces paramètres pour les différentes bandes et échelles afin d'évaluer la signification statistique de leurs capacités respectives à séparer les patients sains et lentigo.

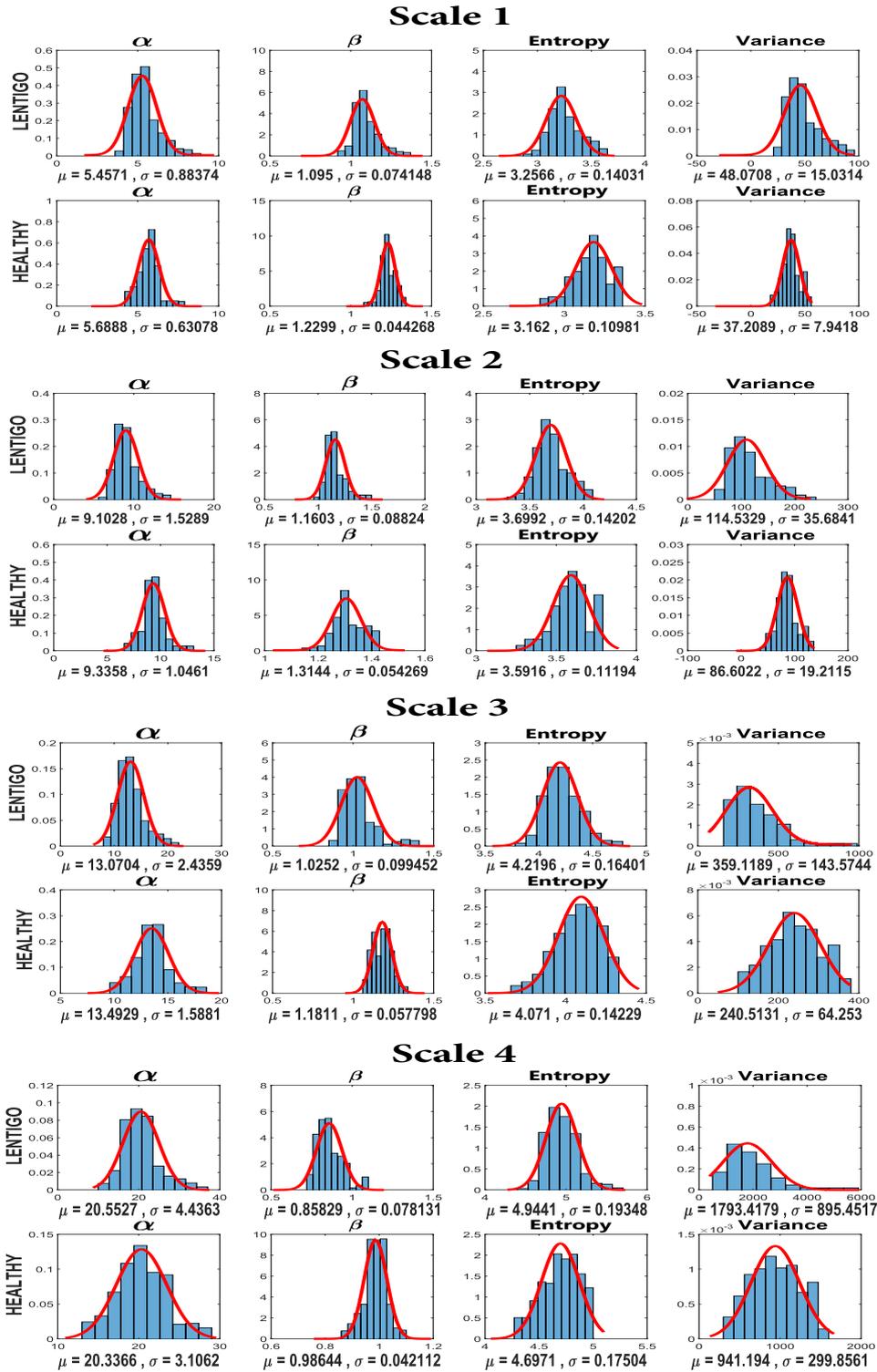


Figure 2.4: Histogrammes et lois gaussiennes associées pour les quatre paramètres (échelle, forme, entropie, variance) avec leurs moyennes et écart-types en utilisant tous les patients (sains et lentigo)

Il convient de noter que l'entropie (dénotée par E) et la variance (dénotée par V) de la loi GGD sont données par

$$E = \frac{1}{\beta} - \log \left[\frac{\beta}{2 \alpha \Gamma\left(\frac{1}{\beta}\right)} \right], \quad V = \frac{\alpha^2 \Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)}.$$

Après avoir caractérisé et séparé les populations à l'aide des différents paramètres (α , β , variance, entropie) et avoir analysé leur pertinence pour la détection de lentigo avec le T-test, nous avons utilisé une pseudo-distance (divergence de Kullback-Leibler) qui est un cas particulier de la famille de divergences appelée α -divergences définies comme suit [Cichocki and Amari, 2010]:

Soient $p(x)$ et $q(x)$ deux densités de probabilité définies de \mathbb{R}^m dans l'intervalle $[0, 1]$ ($m = 1$ dans le cas des histogrammes monodimensionnels).

On appelle α -divergence entre $p(x)$ et $q(x)$ la fonction D_α mesurant la (dis)similarité entre ces deux densités de probabilité et définie par:

$$D_\alpha(p \parallel q) = \frac{\int \alpha p(x) + (1 - \alpha) q(x) - [p(x)]^\alpha [q(x)]^{1-\alpha} dx}{\alpha(1 - \alpha)}, \quad \alpha \in [-\infty, +\infty] \quad (2.7)$$

Les α -divergences vérifient les propriétés suivantes de

- convexité
- positivité: $D_\alpha \geq 0$ et $D_\alpha = 0$ quand les deux densités de probabilité sont égales;
- continuité: la famille est continue par rapport au paramètre α même dans le cas des points singuliers $\alpha = 1$ et $\alpha = 0$ où les limites $\alpha \rightarrow 1$ et $\alpha \rightarrow 0$ peuvent être calculées grâce à la règle de l'Hôpital;

Respectant la propriété de positivité, les divergences peuvent être considérées comme des pseudo-distances. La divergence la plus populaire en traitement du signal et de l'image est la divergence de Kullback-Leibler (ainsi que sa version symétrique).

- Divergence de Kullback-Leibler:

$$D_{KL}(p \parallel q) = \lim_{\alpha \rightarrow 1} D_\alpha(p \parallel q) = \int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx \quad (2.8)$$

$$D_{KL}(q \parallel p) = \lim_{\alpha \rightarrow 0} D_{\alpha}(p \parallel q) = \int q(x) \log \left(\frac{q(x)}{p(x)} \right) dx \quad (2.9)$$

- Divergence de Kullback-Leibler symétrisée:

$$D_{KL}(p, q) = \frac{1}{2} \left(\lim_{\alpha \rightarrow 0} D_{\alpha}(p \parallel q) + \lim_{\alpha \rightarrow 1} D_{\alpha}(p \parallel q) \right) \quad (2.10)$$

$$D_{KL}(p, q) = \int \left[p(x) \log \left(\frac{p(x)}{q(x)} \right) + q(x) \log \left(\frac{q(x)}{p(x)} \right) \right] dx$$

Dans le cas où les deux densités de probabilité $p(x)$ et $q(x)$ ayant respectivement comme paramètres $\theta_1 = (\alpha_1 \ \beta_1)$ et $\theta_2 = (\alpha_2 \ \beta_2)$ sont proches, la divergence de Kullback-Leibler symétrisée peut se mettre sous une forme quadratique avec les coefficients définis par la matrice d'information de Fisher [Jeffreys, 1946].

$$\lim_{q \rightarrow p} D_{KL}(p, q) = (\theta_1 - \theta_2) \times F_{\theta_1} \times (\theta_1 - \theta_2)^T \quad (2.11)$$

où F représente la matrice de Fisher (pour plus de détails voir Annexe B) donnée par :

$$F = - \begin{pmatrix} E \left(\frac{\partial^2 \ln(p_{\alpha, \beta})}{\partial \alpha^2} \right) & E \left(\frac{\partial^2 \ln(p_{\alpha, \beta})}{\partial \alpha \partial \beta} \right) \\ E \left(\frac{\partial^2 \ln(p_{\alpha, \beta})}{\partial \beta \partial \alpha} \right) & E \left(\frac{\partial^2 \ln(p_{\alpha, \beta})}{\partial \beta^2} \right) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\beta}{\alpha^2} & -\frac{1}{\alpha \beta} \left[1 + \psi \left(\frac{1}{\beta} \right) + \beta \right] \\ -\frac{1}{\alpha \beta} \left[1 + \psi \left(\frac{1}{\beta} \right) + \beta \right] & \frac{1}{\beta^2} + \frac{2}{\beta^2} \left[1 + \frac{1}{\beta} \right] \psi \left(\frac{1}{\beta} \right) + \frac{1}{\beta^3} \left[1 + \frac{1}{\beta} \right] \psi' \left(\frac{1}{\beta} \right) + \frac{1}{\beta^3} \left[\psi \left(\frac{1}{\beta} \right) \right]^2 \end{pmatrix}$$

Après avoir utilisé l'approximation au second ordre (forme quadratique) de la divergence de Kullback-Leibler symétrisée, on a calculé la divergence de Kullback-Leibler entre 2 distributions GGD et qui peut se définir comme suit :

$$D(p(x; \alpha_1, \beta_1) \parallel p(x; \alpha_2, \beta_2)) = \int p(x; \alpha_1, \beta_1) \log \left(\frac{p(x; \alpha_1, \beta_1)}{p(x; \alpha_2, \beta_2)} \right) dx \quad (2.12)$$

$$D(p(\cdot; \alpha_1, \beta_1) | p(\cdot; \alpha_2, \beta_2)) = \log \left(\frac{\beta_1 \alpha_2 \Gamma\left(\frac{1}{\beta_2}\right)}{\beta_2 \alpha_1 \Gamma\left(\frac{1}{\beta_1}\right)} \right) + \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma\left(\frac{\beta_2+1}{\beta_1}\right)}{\Gamma\left(\frac{1}{\beta_1}\right)} - \frac{1}{\beta_1} \quad (2.13)$$

Ces pseudo-distances ont été utilisées afin de séparer les deux populations sains (représentée par $p(x)$) et lentigo (représentée par $q(x)$). L'objectif est d'avoir des valeurs faibles de distance dans les cas patients sains-sains et patients lentigo-lentigo et des valeurs élevées de distance entre patients sains-lentigo. On a commencé d'abord par calculer toutes les pseudo-distances Kullback Leibler entre tous les patients (lentigo avec lentigo, lentigo avec sains, et sains avec sains) pour chaque bande, chaque échelle et profondeur. Ensuite on a calculé la moyenne des pseudo-distances (lentigo-lentigo, lentigo-sains, sains-sains) pour chaque profondeur. Un exemple de la matrice regroupant toutes les pseudo-distances Kullback Leibler entre tous les patients (sains et lentigo) pour la bande H scale 1 est donné ci-dessous:

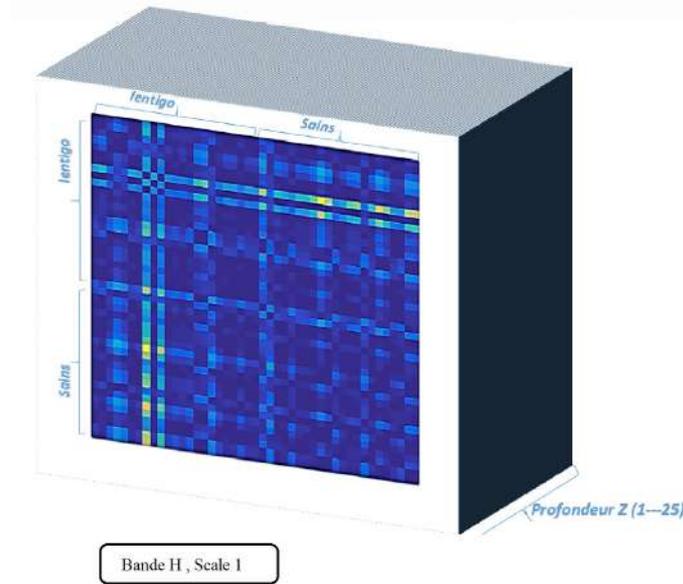


Figure 2.5: Matrices des pseudo-distances Kulback-Leibler en fonction de la profondeur Z.

Ces deux tests nous ont fourni une estimation des profondeurs qui caractérisent les patients sains de ceux atteints de lentigo. Sur la base de ces profondeurs, nous avons considéré un algorithme de

classification pour confirmer leur validité. La Section suivante utilise les paramètres estimés pour les images de test afin d'identifier les patients sains de lentigo en utilisant une méthode de classification.

2.2.4 Classification utilisant les machines à vecteurs support

La méthode des machines à vecteurs support (SVM) est une technique de classification importante qui a fourni de bons résultats dans de nombreuses applications pratiques [Vapnik, 1998]. Le classifieur SVM utilise une méthode d'optimisation permettant d'identifier les vecteurs support \mathbf{s}_p , les poids a_p et le biais b en utilisant un ensemble de données d'apprentissage expertisée, c'est-à-dire telle qu'on connaisse la classe de chaque vecteur d'apprentissage comme expliqué dans la Section 2.3.5. Soit $\theta = [\alpha, \beta, E, V]$ un vecteur de taille (1×4) qui contient $j = 4$ paramètres. Ces paramètres sont utilisés pour classer les vecteurs $\theta^{depths}(j)$ à l'aide de la fonction discriminante suivante

$$c[\theta^{depths}(j)] = \sum_p a_p \mathbf{k}[\mathbf{s}_p, \theta^{depths}(j)] + b \quad (2.14)$$

où \mathbf{k} est la fonction noyau et $\theta^{depths}(j)$ contient le paramètre j aux profondeurs caractéristiques (noté *depths*) de l'image du patient que nous voulons classer. Dans notre cas, nous avons utilisé un noyau linéaire défini par $\mathbf{k}[\mathbf{s}_p, \theta^{depths}(j)] = \mathbf{s}_p^T \theta^{depths}(j)$ [Tourneret, 2003]. Le classifieur SVM affecte le vecteur $\theta^{depths}(j)$ à la classe 1 si $c[\theta^{depths}(j)] \geq 0$ et la classe 2 si $c[\theta^{depths}(j)] \leq 0$. Cette procédure est appliquée à chaque bande et à chaque échelle de façon indépendante, afin d'analyser leur potentiel de classification des tissus sains et de lésions.

2.3 Expériences

2.3.1 Données

Les images MCR réelles ont été obtenues avec l'appareil Vivascope 1500. Les images in-vivo ont été acquises à partir de la couche cornée, de la couche épidermique, de la jonction derme-épiderme (DEJ) et enfin du derme papillaire supérieur. Chaque image MCR offre un champ de vision $500 \times 500 \mu m$ avec 1000×1000 pixels. Un ensemble de $L = 45$ femmes âgées de 60 ans et plus ont été recrutées. L'étude a été menée selon les principes de la déclaration d'Helsinki et a été approuvée par le *Comité de Protection des Personnes Sud-Ouest et Outre-mer III* à Bordeaux, France, No. CPP 2011/36.

Tous les volontaires ont donné leur consentement en parfaite connaissance de cause pour l'examen de la peau par MCR. Selon l'évaluation clinique réalisée par un médecin, les volontaires ont été divisés en deux groupes. Le premier groupe est constitué de 27 femmes avec au moins 3 lentigos au dos de la main tandis que 18 femmes sans lentigo constituaient le groupe témoin. Par conséquent, notre base de données contient $L = 45$ patients. Deux acquisitions ont été effectuées pour chaque volontaire et pour les 25 profondeurs, chaque acquisition fournissait une pile (noté comme *stack*) contenant les 25 images. Des images ont été prises sur des lésions lentigo pour les volontaires du premier groupe et au dos de la main sur une peau saine pour le groupe témoin. Un examen de chaque acquisition a été effectué afin de localiser la couche cornée et le DEJ précisément pour chaque image. Pour chaque patient, nous avons retenu deux piles de 25 images MCR, ce qui nous donne un total de 2250 images.

2.3.2 Analyse statistique des coefficients d'ondelettes

Cette Section illustre la qualité de l'ajustement de la loi gaussienne généralisée pour les coefficients d'ondelettes des images MCR. La Figure 2.6 compare les histogrammes des coefficients d'ondelettes des bandes mélangées *HVD* à une échelle donnée avec les lois GGD estimées pour les quatre échelles à 3 profondeurs représentatives. Cette Figure représente deux patients arbitraires sains et lentigo, à savoir les patients #6 et #38 respectivement. Elle montre le bon ajustement entre les histogrammes observés et les lois estimées pour toutes les échelles et pour les images des patients sains et lentigo. Pour une échelle égale à 4, la distribution manque le mode mais s'ajuste bien à la forme. Cela peut s'expliquer par le manque de données car les coefficients à cette échelle sont très parcimonieux et tendent vers un histogramme plat. Des différences légères dans la forme des lois peuvent être observées entre les patients sains et lentigo, comme en témoignent les différences dans les paramètres correspondants α et β . Ces différences constituent la base de la méthode de caractérisation proposée, comme nous allons le montrer dans les expériences suivantes. De bons résultats ont également été obtenus pour les bandes H, V, D séparément mais ils ne sont pas présentés ici car ils sont très similaires à ceux obtenus pour une analyse conjointe avec les trois bandes (H,V,D).

La Figure 2.7 montre l'évaluation quantitative de l'ajustement à l'aide du test de Kolmogorov-Smirnov (KS). Le score statistique moyen de *KS* de l'ensemble de la population (45 patients) a été

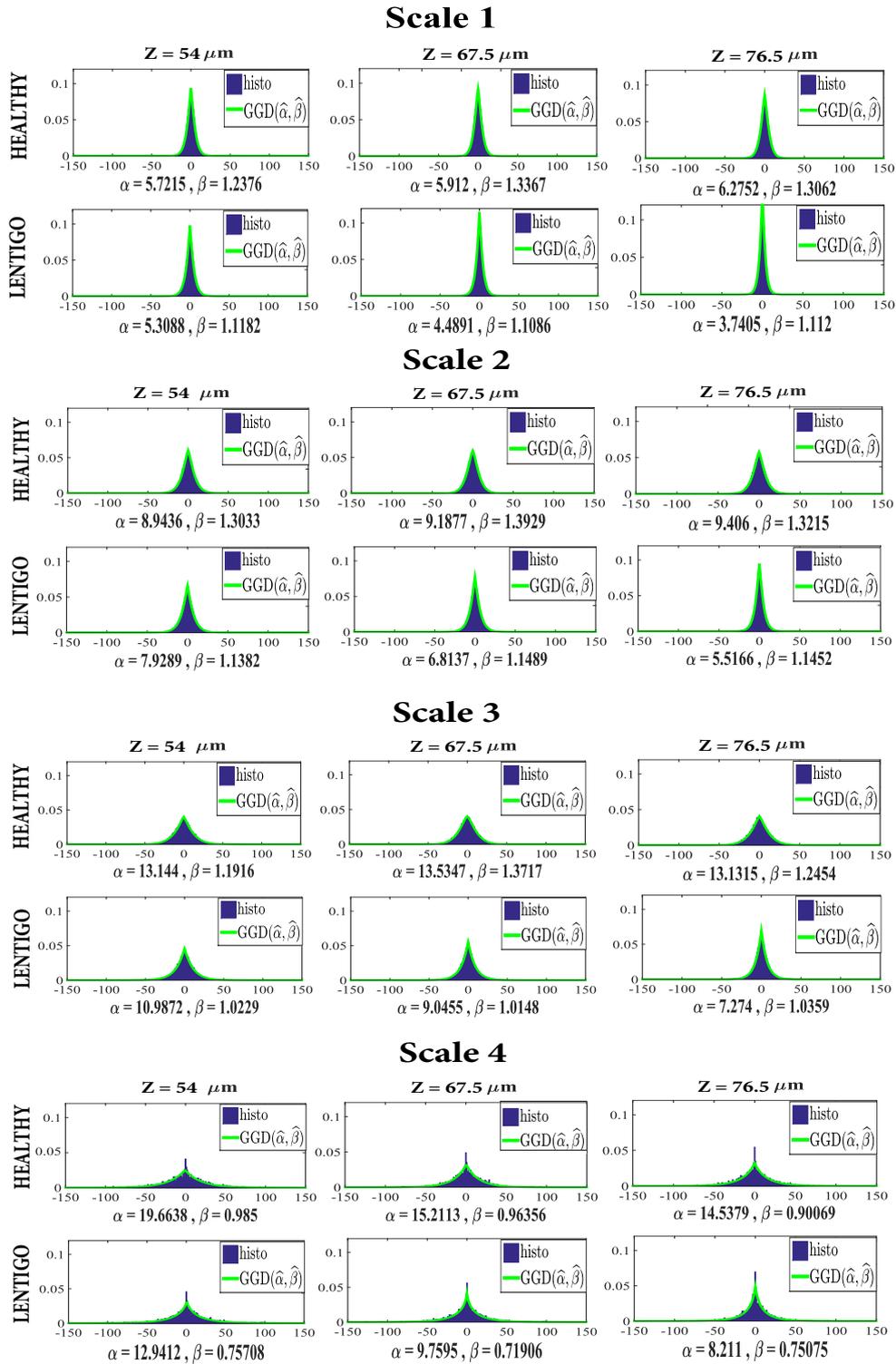


Figure 2.6: Histogrammes des coefficients d'ondelettes à partir des bandes mélangées HVD aux quatre échelles et leurs distributions GGD estimées correspondantes. La Figure montre les données de deux patients arbitraires sains et lentigo (#6 and #38 respectivement) à trois profondeurs représentatives (une profondeur par colonne). On remarque le très bon ajustement pour les échelles de 1 à 3 pour les deux types de patients.

calculé à chaque profondeur et pour toutes les échelles. On remarque les excellents scores avec des valeurs de KS très proches de zéro, bien que légèrement plus élevés pour la quatrième échelle.

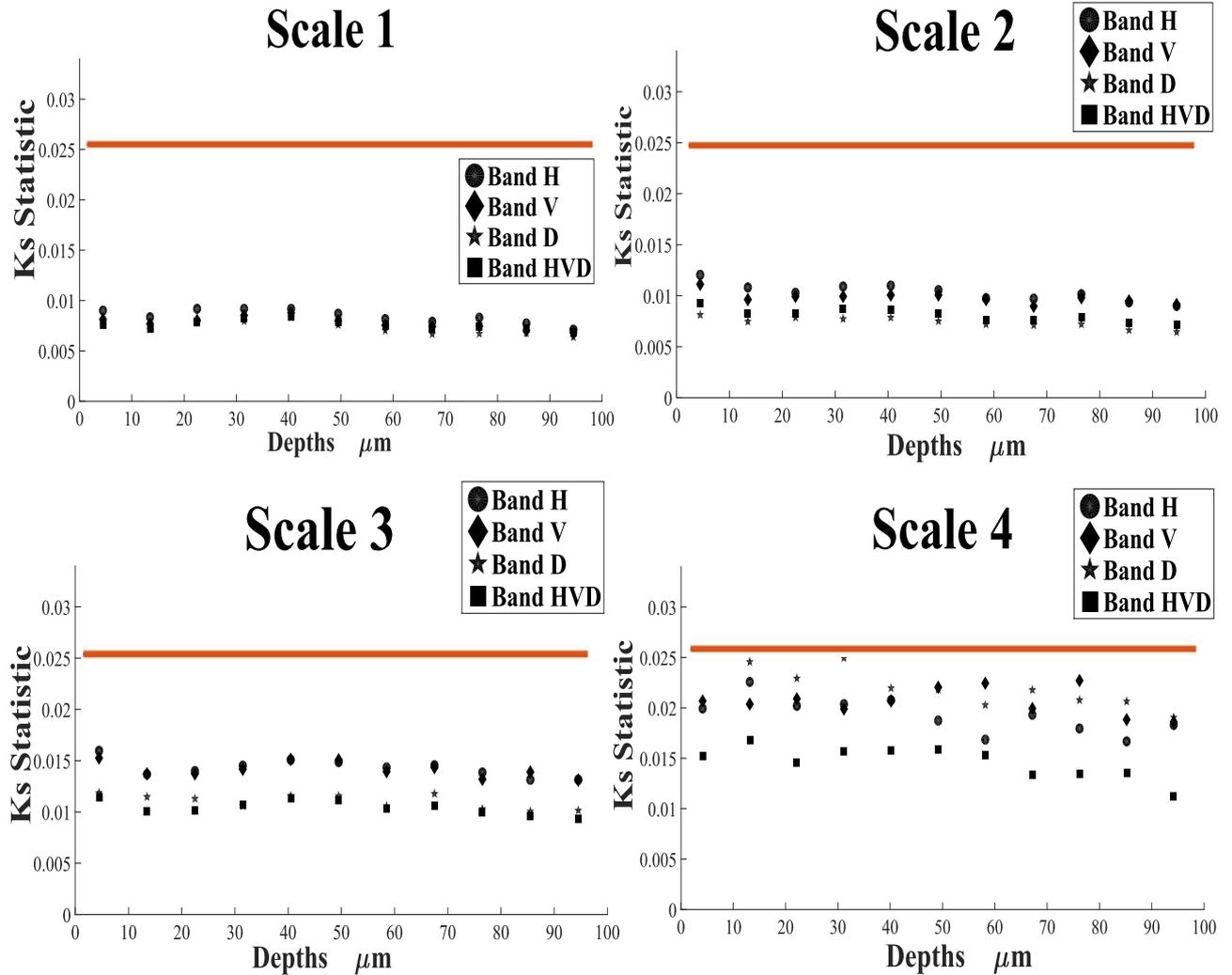


Figure 2.7: Evaluation de l'ajustement d'une loi GGD aux coefficients d'ondelettes. Statistique KS moyenne pour l'ensemble de la population à diverses profondeurs et pour toutes les échelles. Les scores sont excellents pour toutes les configurations.

2.3.3 Identification des profondeurs caractéristiques

Chacune des 2250 images a été décomposée selon l'analyse multi-résolution présentée dans la Section 2.2.1 et les lois GGD ont été ajustées à chaque échelle et à chaque profondeur. Ayant acquis deux *stack* de 25 images pour chaque patient, l'un des deux *stack* a été sélectionnée au hasard pour l'analyse. Les paramètres moyens $\hat{\theta}^{(H)}(j)$ et $\hat{\theta}^{(L)}(j)$ ont ensuite été calculés à chaque profondeur, pour chaque échelle et bande et pour les patients sains et lentigo, respectivement. La courbe montrant l'évolution des paramètres $\hat{\theta}^{(\cdot)}(j)$ en fonction de la profondeur de l'image a finalement été calculée. L'approximation au second ordre de la divergence de Kullback-Leibler symétrisée ainsi que la divergence de Kullback-Leibler ont été aussi calculées à partir de ces paramètres pour chaque profondeur afin d'identifier les plus caractéristiques. Pour tenir compte de la variabilité, le processus de sélection d'une pile pour chaque patient a été répété 300 fois et les courbes moyennes et leurs écart-types ont été calculés. Ces courbes moyennes vont nous permettre de trouver les profondeurs auxquelles la séparation entre les deux populations (lentigo et sains) est significative. Les résultats sont présentés dans les Figures 2.8, 2.9 et 2.10. Elles montrent clairement que l'entropie, la variance et α ne permettent pas la discrimination d'images des patients sains et lentigo pour toutes les profondeurs, les bandes et les échelles 1 à 2.

À l'inverse, la distance de Kullback-Leibler ainsi que le paramètre β permettent la discrimination d'images des patients sains et lentigo pour toutes les bandes à des échelles de 1 à 3 et pour des profondeurs comprises entre $31 \mu m$ and $76 \mu m$. Les profondeurs aux environs de $50 \mu m$ donnent la plus grande discrimination. Les bandes H et V pour la quatrième échelle montrent une grande variabilité des paramètres en raison de la faible quantité de données. Le choix de l'ondelette et de son échelle ne changent pas les résultats obtenus de manière significative. Pour confirmer cela, nous avons mené une étude similaire avec les ondelettes Haar et Symlets, ce qui a mené à des conclusions similaires. Les résultats obtenus sont présentés sur les Figures 2.11(a) and 2.11(b). Ces résultats sont semblables à ceux présentés dans la Figure 2.8.

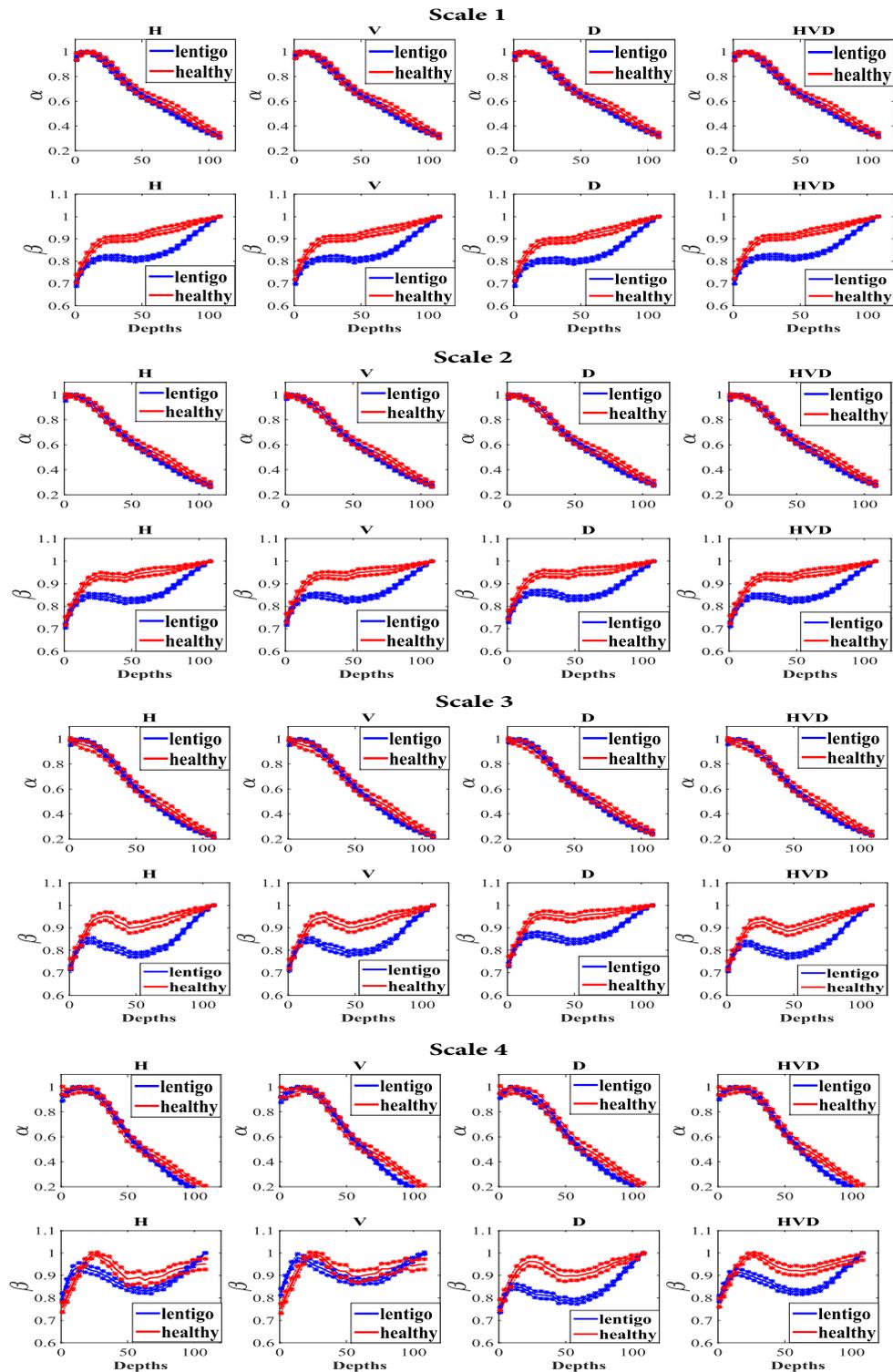


Figure 2.8: Évolution des paramètres moyens $\hat{\alpha}$ et $\hat{\beta}$ en fonction de la profondeur, pour les différentes bandes et à toutes les échelles. Les valeurs de α sont très similaires pour les patients sains et lentigo et ne peuvent être utilisées pour la discrimination. Le paramètre β montre une différence significative pour les profondeurs entre 31 μm and 76 μm , avec une différence maximale aux environs de 50 μm . Notre conclusion est que ce paramètre β peut discriminer les tissus sains et lentigo de la peau.

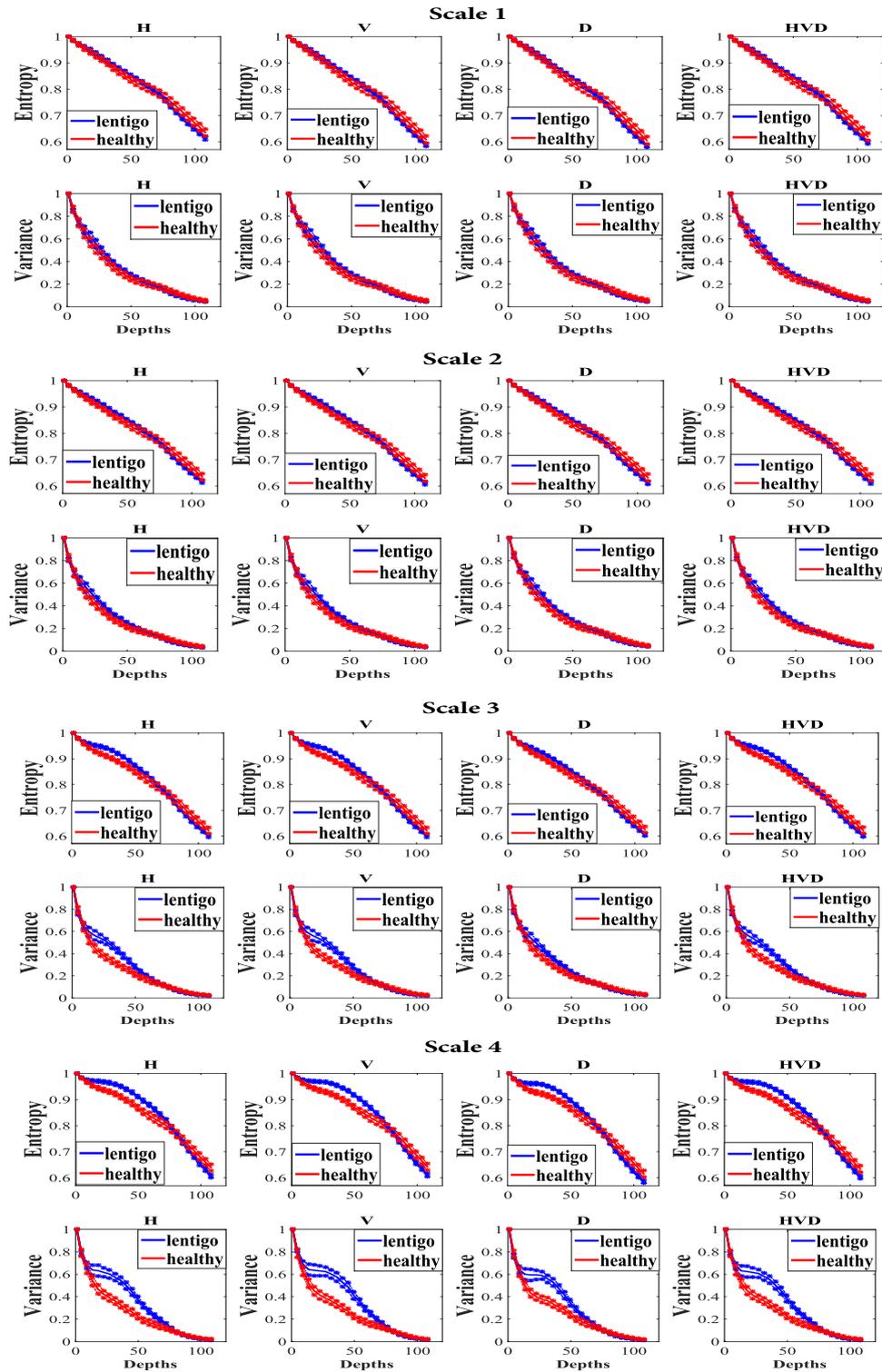
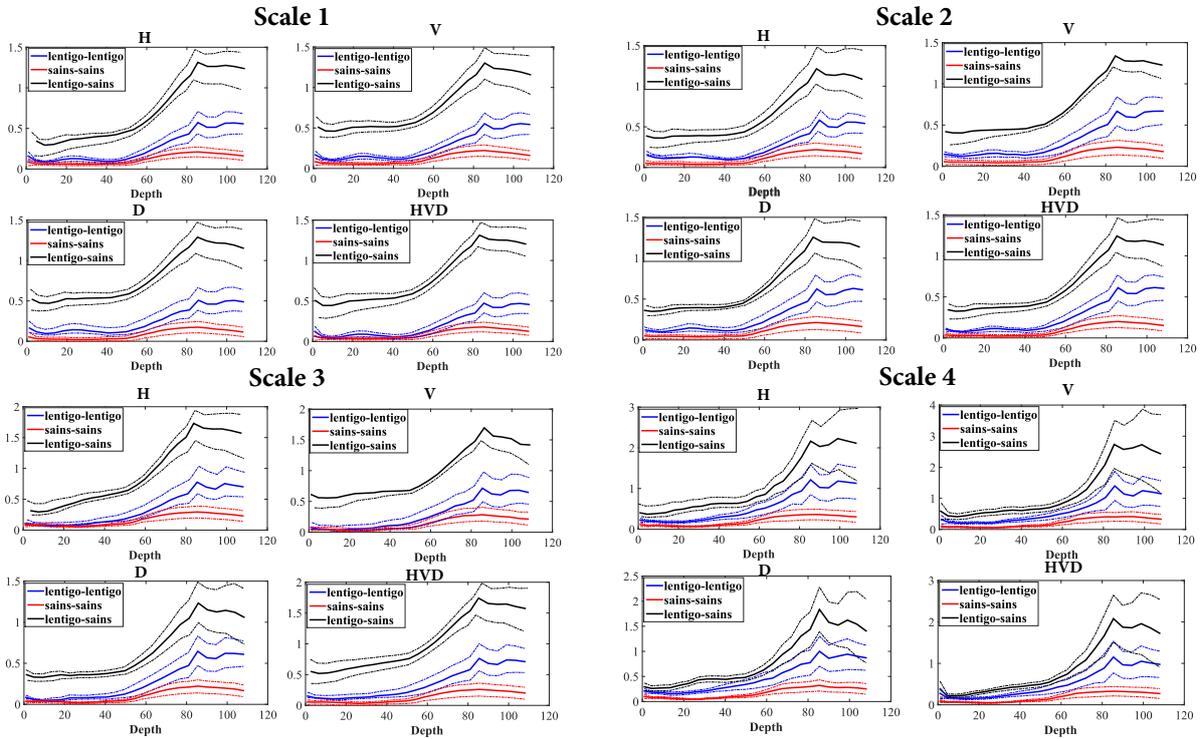
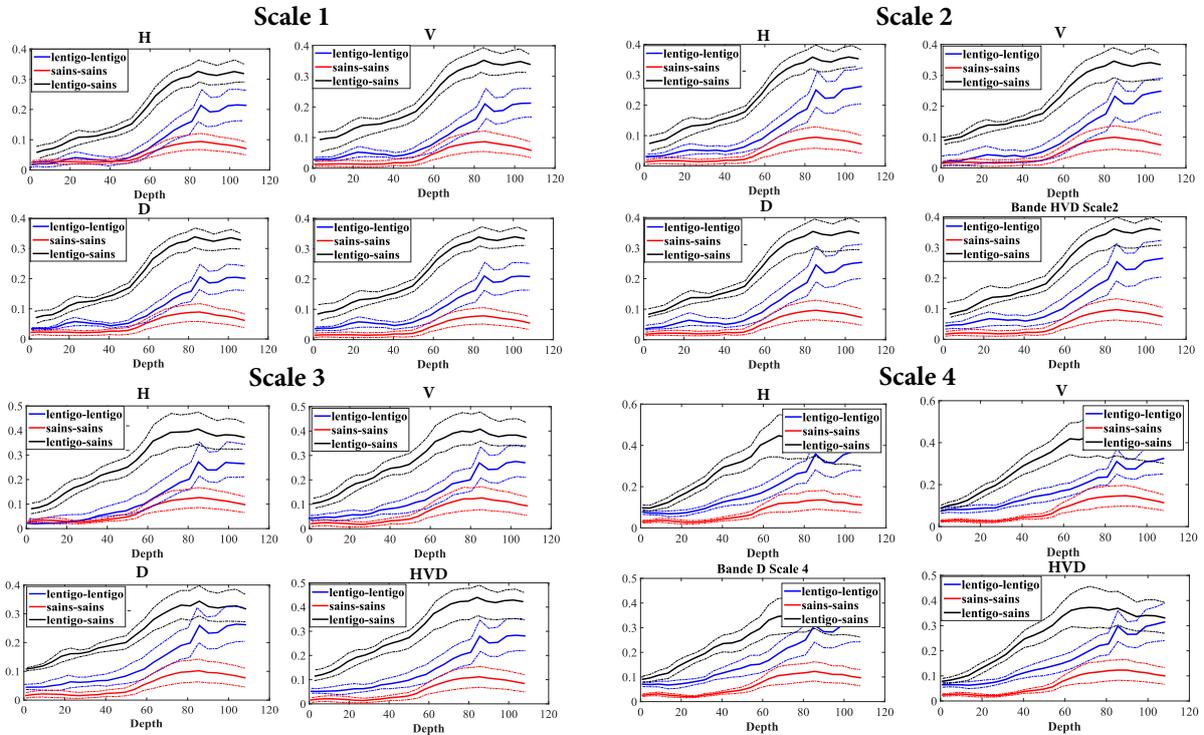


Figure 2.9: Evolution des paramètres moyens de l'entropie et de la variance en fonction de la profondeur, pour les différentes bandes et à toutes les échelles. Les valeurs d'entropie et de variance sont très similaires pour les patients sains et lentigo et ne peuvent être utilisées pour la discrimination.

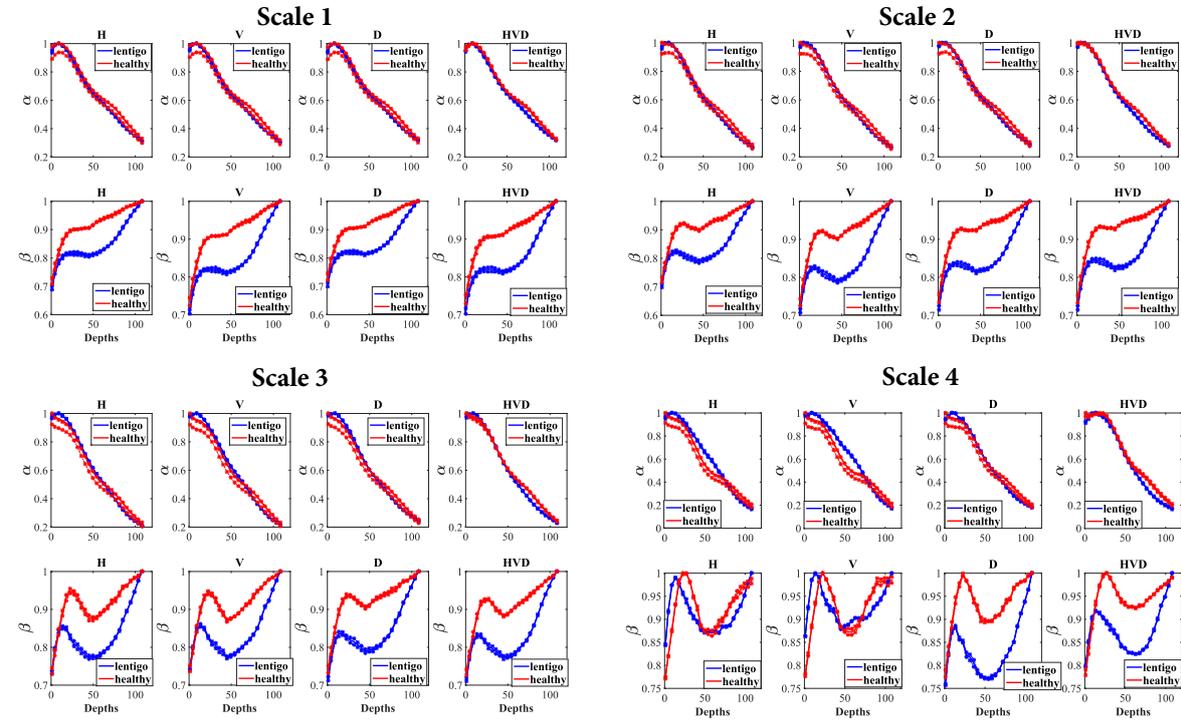


(a) Approximation de Kulback-Leibler

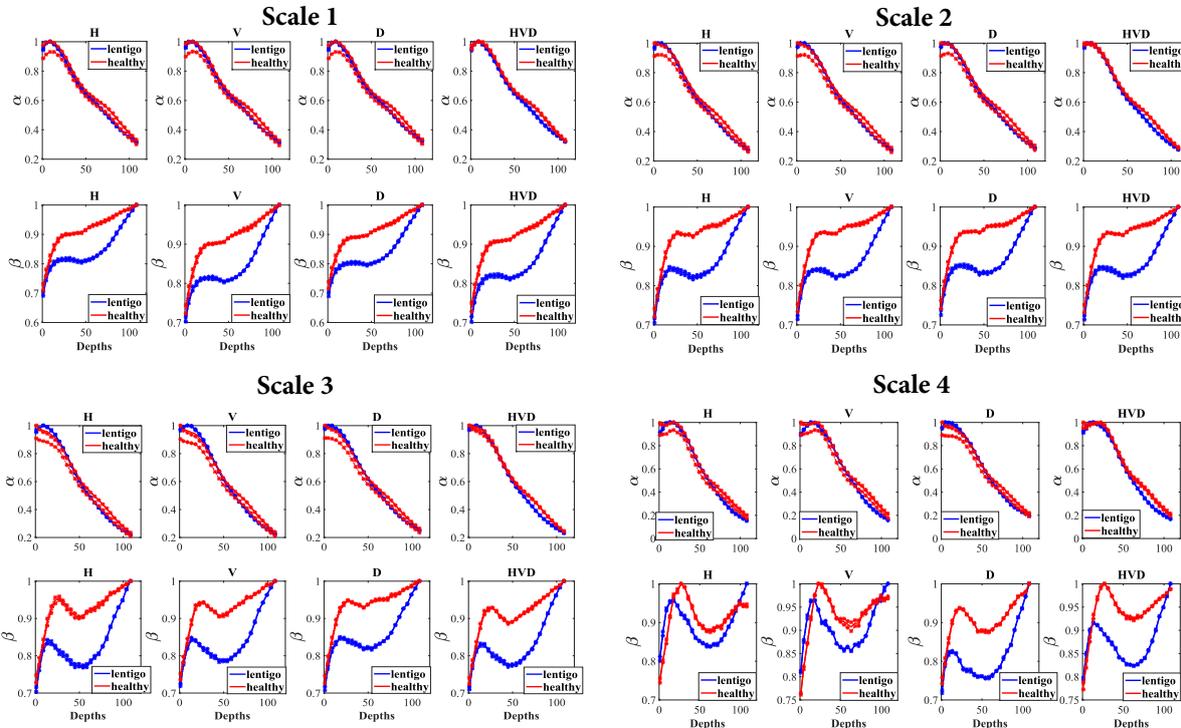


(b) La distance de Kulback-Leibler

Figure 2.10: Évolution des distances Kulback-Leibler moyennes ainsi que leurs approximations en fonction de la profondeur, pour les différentes bandes à toutes les échelles. Ces distances montrent une différence significative pour les profondeurs entre 31 μm et 76 μm , avec une différence maximale aux environs de 50 μm .



(a) Haar decomposition



(b) Symlets decomposition

Figure 2.11: Évolution des paramètres moyens $\hat{\alpha}$ et $\hat{\beta}$ en fonction de la profondeur, pour les différentes bandes à toutes les échelles. Les valeurs de α sont très similaires pour les patients sains et lentigo et ne peuvent être utilisées pour la discrimination. Le paramètre β montre une différence significative pour les profondeurs entre 31 μm et 76 μm , avec une différence maximale aux environs de 50 μm . Notre conclusion est que ce paramètre β peut discriminer les tissus sains et lentigo de la peau.

La Figure 2.12 montre deux séries d'images de six patients sains (#1, #2, #3, #4, #5, #6) et six patients lentigo (#31, #33, #37, #38, #40, #44) qui illustrent la différence de texture à la profondeur DEJ. La Figure 2.13 représente les images macro couleur des différents patients (lentigo et sains) affichées dans la Figure 2.12. La Figure 2.14 montre aussi d'autres exemples d'images couleur de patients sains et lentigo, qui souligne le fait que dans notre cas la pigmentation de la peau ne peut pas être utilisée pour différencier les patients atteints de lentigo et les patients en bonne santé. Cependant, même s'il y a une variation dans la pigmentation de la peau nos résultats ne devraient pas être affectés car notre méthode repose sur les lois des coefficients d'ondelettes et non pas sur l'intensité de la lumière réfléchiée sur les images macro couleurs.

En résumé, dans cette Section précédente, on a montré qu'il y avait un bon ajustement entre la loi gaussienne généralisée et la distribution des coefficients d'ondelettes des images MCR, et que l'entropie, la variance et α ne permettent pas la discrimination d'images des patients sains et lentigo pour toutes les profondeurs, les bandes et les échelles 1 à 2. Par contre, la pseudo-distance de Kulback-Leibler ainsi que le paramètre β permettent la discrimination d'images des patients sains et lentigo pour toutes les bandes à des échelles de 1 à 3 et pour des profondeurs comprises entre 31 μm and 76 μm . Les profondeurs aux environs de 50 μm donnent la plus grande discrimination.

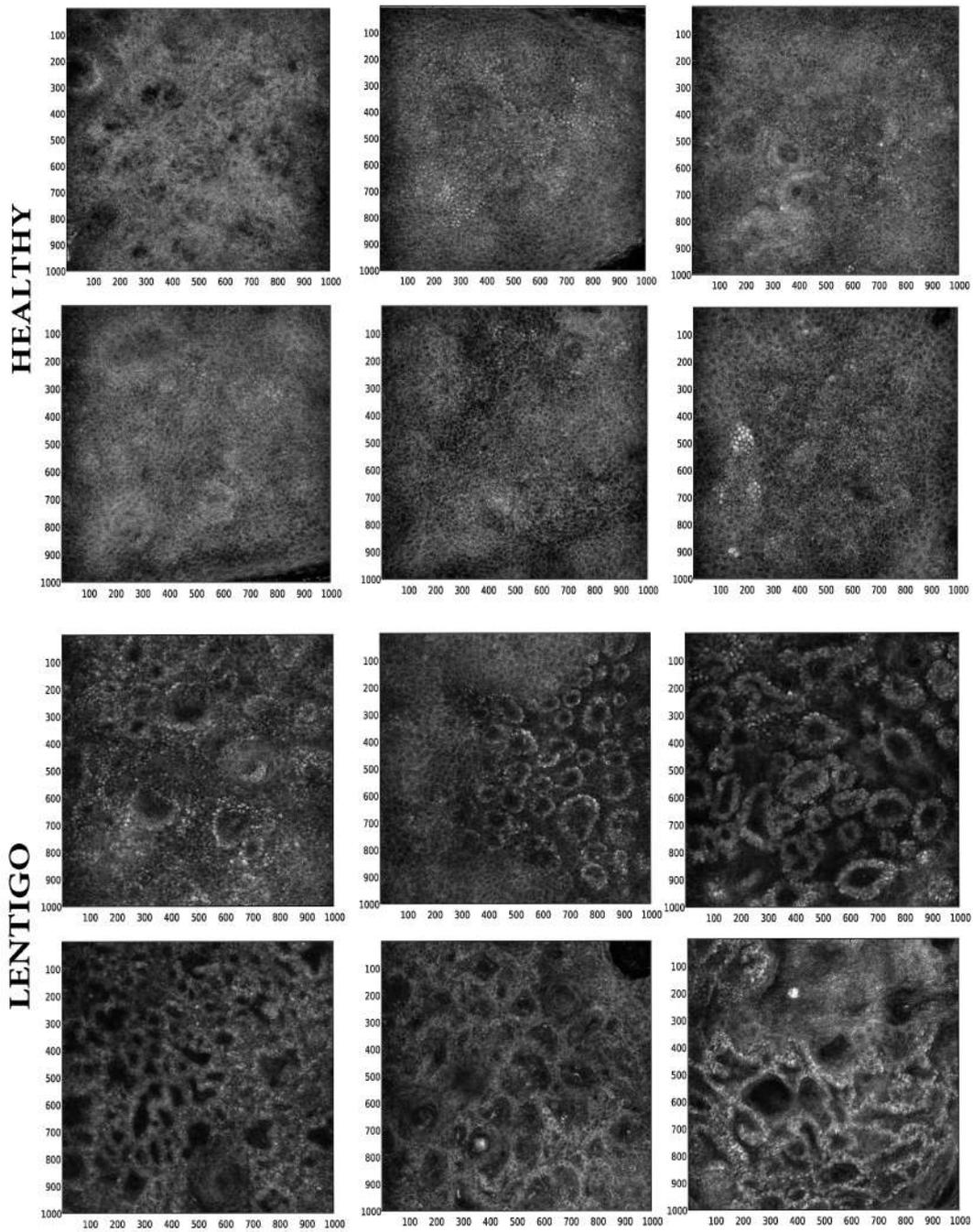


Figure 2.12: Images des patients sains (patient #1, #2, #3, #4, #5, #6) et lentigo (patient #31, #33, #37, #38, #40, #44) à la profondeur DEJ. On peut observer la différence significative dans la texture, capturée par le paramètre β (comme l'illustre la Figure 2.8).

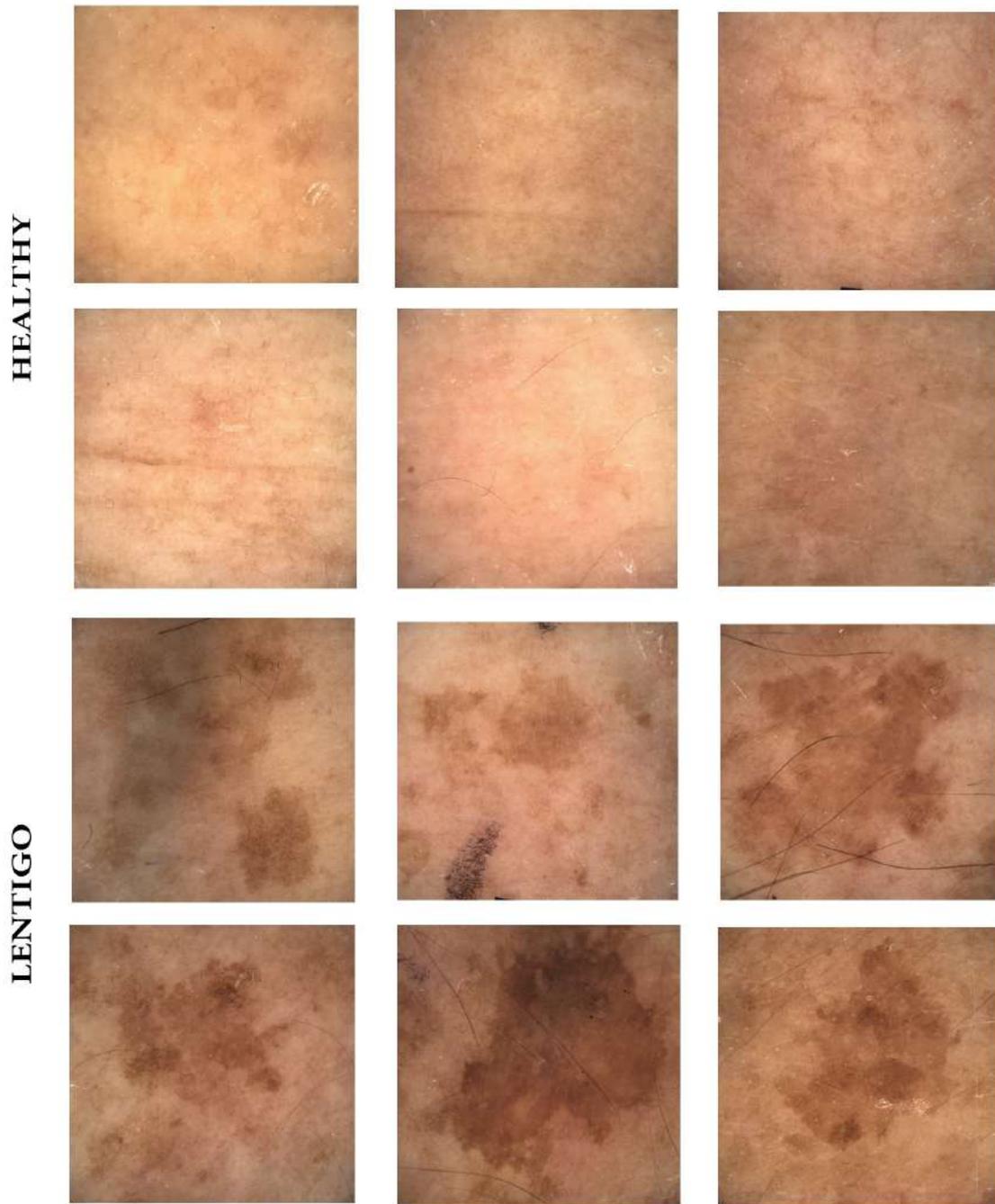


Figure 2.13: Images macro couleur des patients sains (patient #1, #2, #3, #4, #5, #6) et lentigo (patient #31, #33, #37, #38, #40, #44) affichées dans la Figure 2.12.

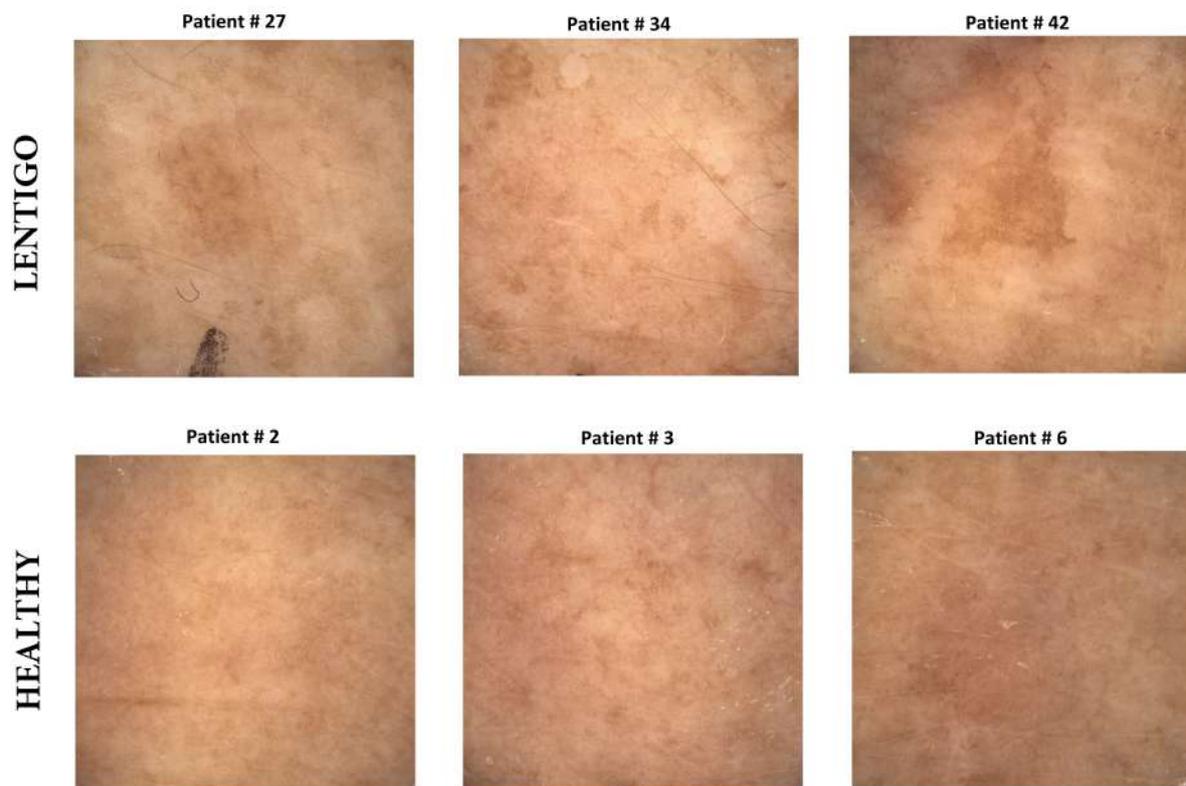


Figure 2.14: Images couleur des patients sains et lentigo.

2.3.4 Signification statistique avec le test T

Le test paramétrique décrit dans l'Annexe A a été appliqué aux paramètres de chaque bande, échelle et profondeur pour évaluer la signification des résultats obtenues dans la Section précédente. Les Figures 2.15 à 2.18 représentent les p-values et les facteurs de Bayes associés aux tests de Student à deux échantillons (plus de détails sur le test dans Annexe A) effectués respectivement avec α , β , l'entropie et la variance, à différentes profondeurs. La p-value a été représentée en échelle logarithmique $-\log$ dans toutes les Figures pour plus de lisibilité. Les Figures 2.15, 2.17 et 2.18 présentent des valeurs faibles à la fois pour la p-value et le facteur de Bayes confirmant que α , l'entropie et la variance ne peuvent pas discriminer les patients sains et lentigo. La Figure 2.16 montre des scores élevés pour les deux indicateurs (p-value et facteur de Bayes) et pour une gamme de profondeurs. Le tableau 2.1 présente des profondeurs qui donnent des valeurs de $T_{\nu}^{(\beta)}$ plus élevées que le seuil, confirmant ainsi l'hypothèse selon laquelle β discrimine les patients sains et lentigo ($H_0^{(\beta)}$ rejetée) à des profondeurs caractéristiques. Ce tableau montre aussi les profondeurs qui fournissent des p-values inférieures à la probabilité de fausse alarme $t = 0,05$ et leurs facteurs de Bayes correspondants. On peut voir que pour toutes les profondeurs où $H_0^{(\beta)}$ est rejetée, les p-values sont inférieures à 0.01.

		scale 1		scale 2		scale 3		scale 4	
		min	max	min	max	min	max	min	max
$T_{\nu}^{(\beta)} > 2.02$	profondeurs (μm)	14	90	18	90	18	90	18	94
	T-score	2.10	2.22	2.13	2.27	2.10	2.20	2.14	2.11
	p-value	0.040	0.037	0.040	0.030	0.040	0.037	0.038	0.040
	BF	11.00	13.40	7.18	15.23	23.10	29.70	13.24	10.20
$T_{\nu}^{(\beta)}$ les plus significatif	profondeurs (μm)	48	60	48	63	48	68	48	68
	T-score	4.80	4.84	4.54	4.73	4.02	4.07	3.99	4.26
	p-value	0.0002	0.0002	0.0004	0.0003	0.002	0.002	0.003	0.001
	BF	877	970	891	949	454	460	359	435
Maximal $T_{\nu}^{(\beta)}$	profondeurs (μm)	54		54		59		59	
	T-score	5.11		4.94		4.69		4.41	
	p-value	0.00008		0.00016		0.00031		0.00074	
	BF	1246		1459		436		507	

Tableau 2.1: Profondeurs où β discrimine les patients ($H_0^{(\beta)}$ rejeté) obtenu pour $T_{\nu}^{(\beta)} > T_{0.05} = 2.02$; La p-value et le facteur de Bayes (BF) correspondants sont indiqués. La première ligne donne les intervalles de profondeur (profondeur minimale à la profondeur maximale) où les scores T sont significatifs. La deuxième ligne montre les profondeurs donnant les scores T les plus élevés (score maximal $\mp 10\%$). La troisième ligne montre les profondeurs correspondant au score T maximal. Les p-value et les facteurs de Bayes correspondants à chaque profondeur sont indiqués en dessous.

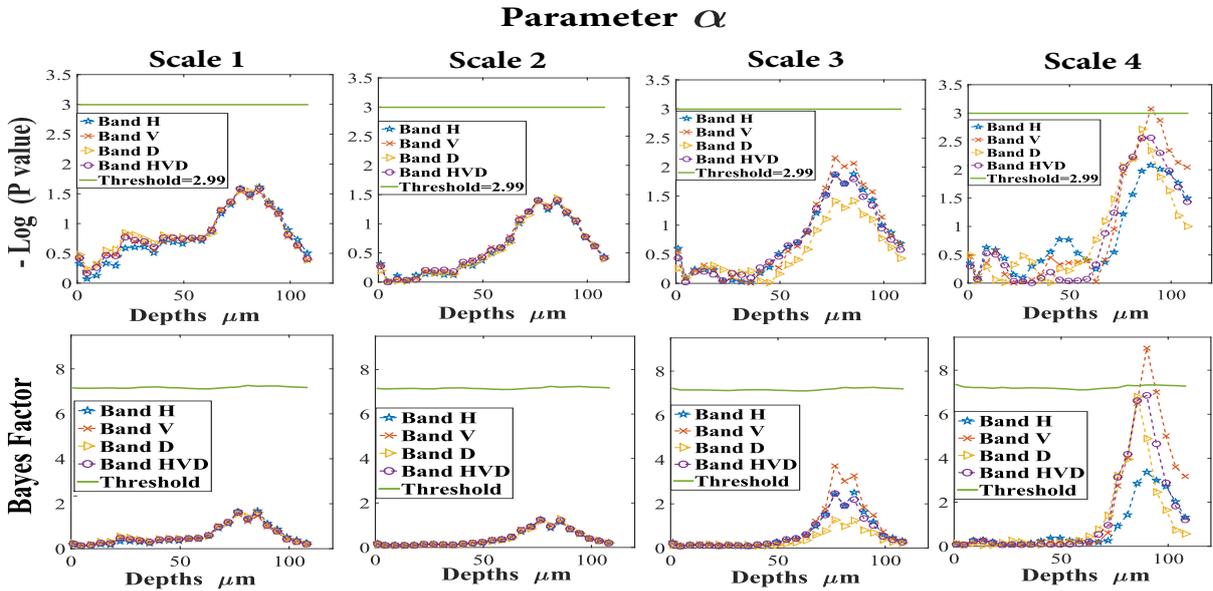


Figure 2.15: P-value (in $-\log$ scale) et facteur de Bayes du test T pour α . Le faible score montre que α n'est manifestement pas un discriminant entre les images des patients sains et lentigo.

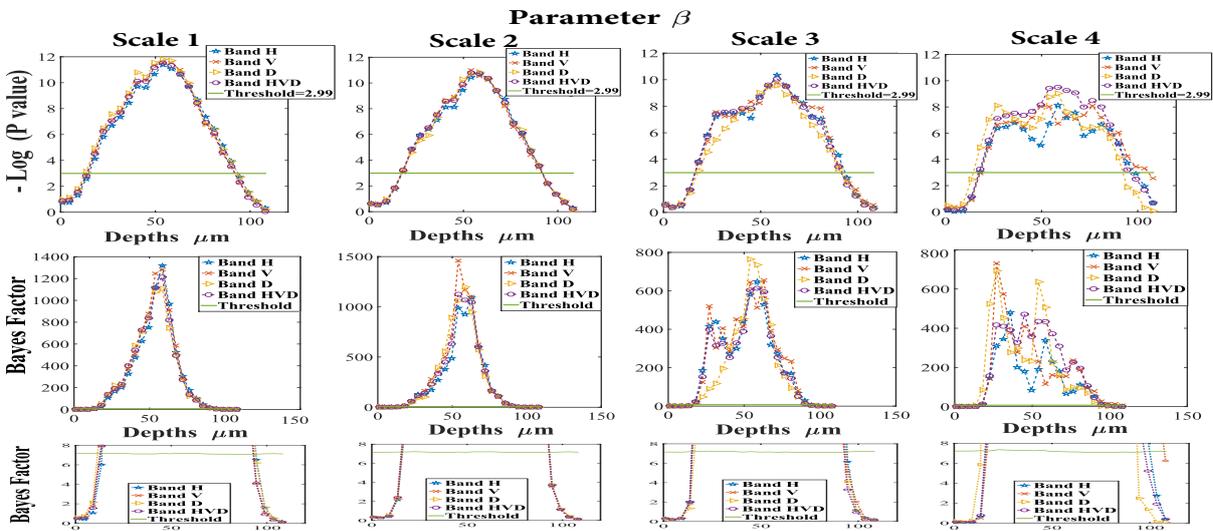


Figure 2.16: P-value (in $-\log$ scale) et facteur de Bayes (BF) du test T pour β . La troisième ligne représente des zooms sur les faibles scores du BF afin de clarifier le seuil de signification. Des scores forts peuvent être observés pour des profondeurs comprises entre $31 \mu m$ and $76 \mu m$. Les scores les plus élevés sont obtenus pour des profondeurs aux environs de $50 \mu m$. Cela confirme que β est une bonne fonction discriminante qui peut être utilisée pour séparer les images saines et lentigo à ces profondeurs.

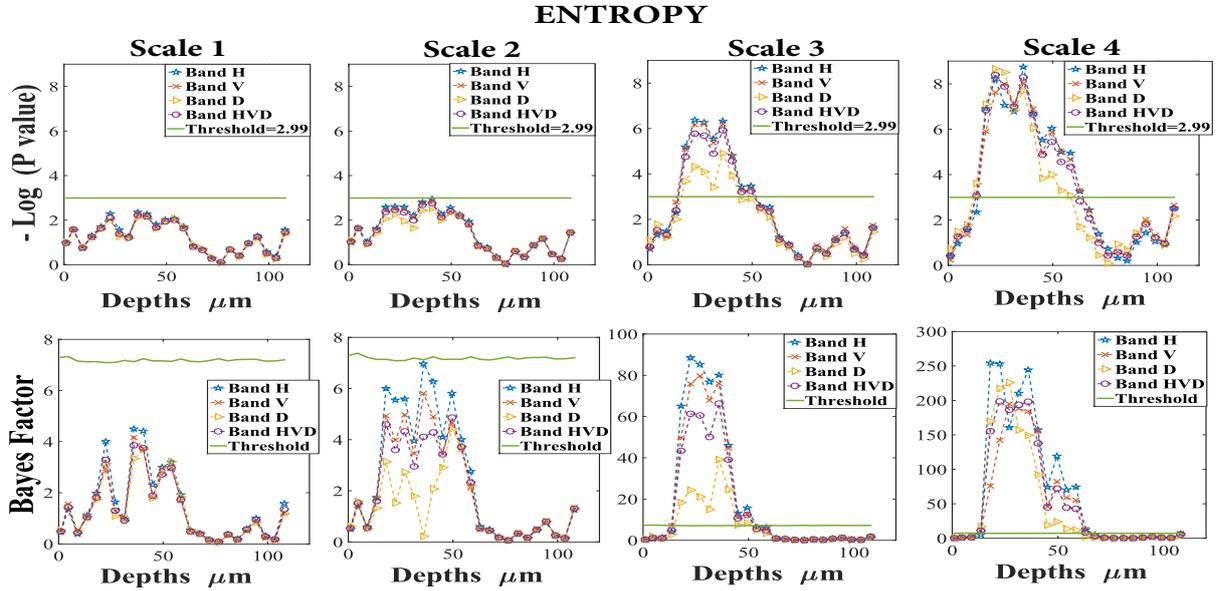


Figure 2.17: P-value (in $-\log$ scale) et facteur de Bayes du test T pour l'entropie. Le faible score montre que l'entropie n'est manifestement pas un discriminant entre les images des patients sains et lentigo pour les échelles (1 et 2).

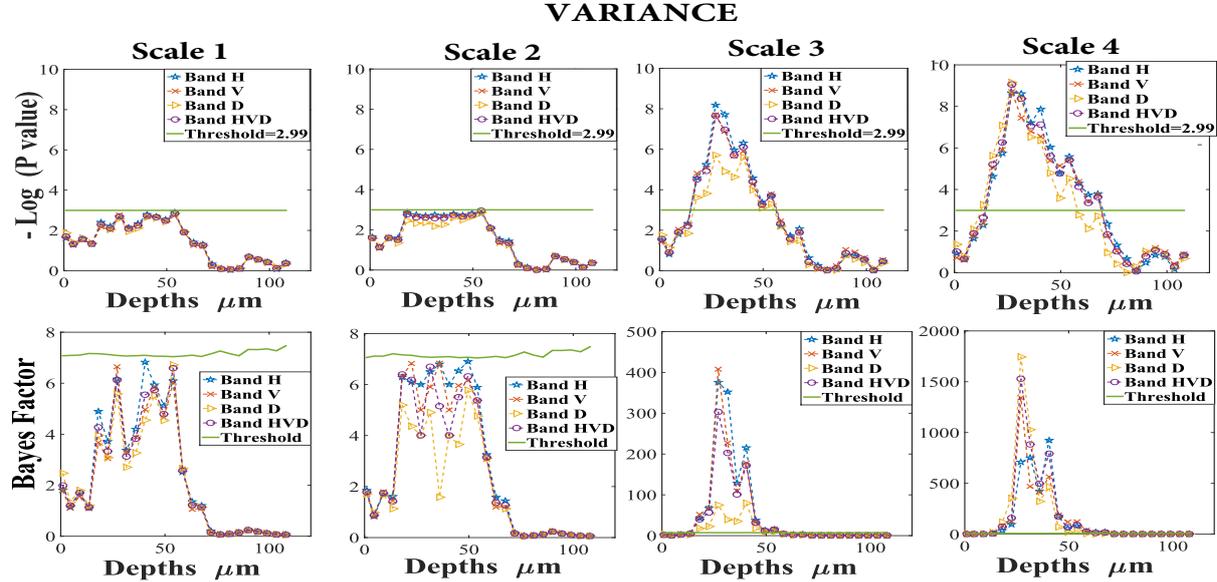


Figure 2.18: P-value (in $-\log$ scale) et facteur de Bayes du test T pour la variance. Le faible score montre que la variance n'est manifestement pas un discriminant entre les images des patients sains et lentigo pour les échelles (1 et 2).

Selon nos règles de décision décrites dans l'Annexe A, les résultats sont *hautement significatifs* pour les profondeurs de $49 \mu\text{m}$ et $60 \mu\text{m}$ et pour la première échelle (scale 1), avec le score le plus élevé à $54 \mu\text{m}$. Ceux-ci sont en parfait accord avec les différences quantitatives indiquées dans la Figure 2.8. Ces résultats confirment que β est une bonne statistique de test pour discriminer une peau saine de lentigo, en particulier à des profondeurs autour de $50 \mu\text{m}$. Comme mentionné dans l'introduction, les lentigos se caractérisent principalement en MCR par la désorganisation de la jonction dermo-épidermique (DEJ). Il n'est donc pas surprenant de ne trouver aucune différence significative entre la peau saine et lentigo dans le stratum corneum et les couches supérieures d'épiderme (inférieures à $40 \mu\text{m}$). D'une manière cohérente, le paramètre β est très discriminant à des profondeurs proches de $50 \mu\text{m}$, ce qui correspond à l'emplacement moyen du DEJ comme c'est indiqué sur la Figure 2.19.

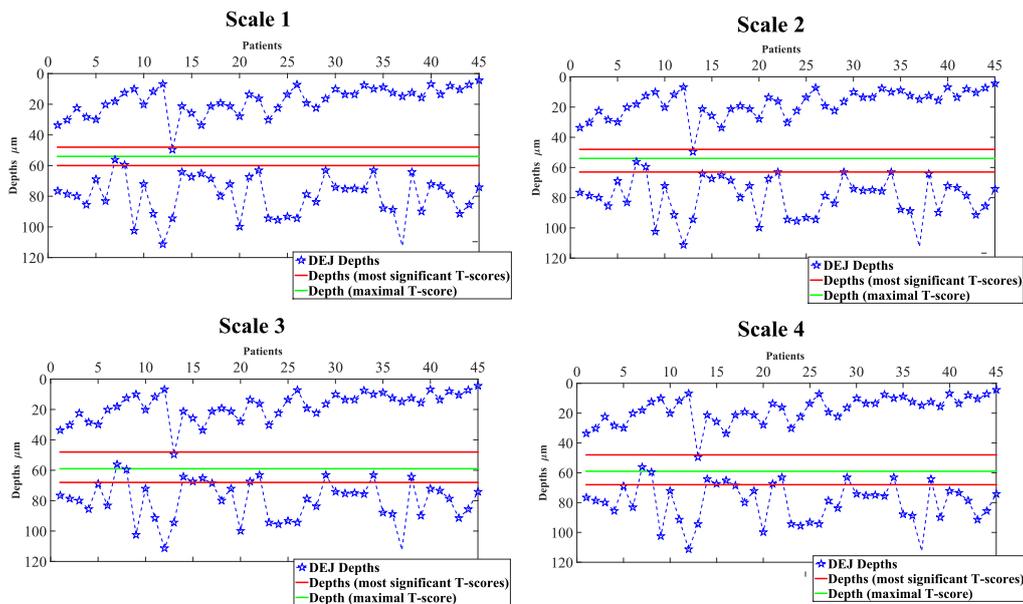


Figure 2.19: Les profondeurs caractéristiques qui se situent entre $48 \mu\text{m}$ et $63 \mu\text{m}$ selon le test T et les profondeurs DEJ associées aux 45 patients.

2.3.5 Performance du classifieur SVM

Les paramètres des lois GGD associées à chaque échelle, bande (H,V,D,HVD) et aux profondeurs caractéristiques ($40 \mu m$ à $60 \mu m$) pour un patient donné ont ensuite été utilisés pour classer les patients dans 2 classes appelées “lentigo” et “sain”. La méthode leave-one-out a été utilisée pour calculer les différentes probabilités d’erreurs. Cette méthode utilise $L - 1$ images pour l’apprentissage (où L est le nombre de patients dans la base de données) et l’image restante pour le test. Cette opération a été exécutée $M = 1000$ fois. Pour chaque expérience, nous n’avons considéré que les images d’une acquisition sur les deux disponibles (pour chaque patient). Les M résultats obtenus ont été utilisés pour calculer la matrice de confusion moyenne indiquée dans le tableau 2.3 et pour évaluer les indicateurs moyens (Sensibilité, Spécificité, Précision et Accuracy). Ces indicateurs sont définis comme Sensibilité = $TP/(TP+FN)$, Spécificité = $TN/(FP+TN)$, Précision = $TP/(TP+FP)$, Accuracy = $(TP+TN)/(TP+FN+FP+TN)$, où TP, TN, FP et FN sont les nombres des vrais positifs, vrais négatifs, faux positifs et faux négatifs.

Ce tableau nous permet d’évaluer la performance de classification en utilisant le paramètre β pour chaque échelle et pour les trois bandes séparément et conjointement. Les résultats montrent que la classification des tissus sains et des lésions est similaire pour un seul détail (H, V ou D) et nous recommandons donc d’utiliser le vecteur conjoint HVD (contenant tous les détails) pour cette classification. En ce qui concerne les différentes échelles utilisées pour l’analyse, le tableau 2.3 montre que les trois premières échelles fournissent des résultats similaires et qu’elles sont légèrement meilleures que celles obtenues avec la 4^{ème} échelle. Nous recommandons donc d’utiliser les trois premières échelles de la décomposition des ondelettes pour la classification des images saines et lentigo. La Figure 2.20 montre des exemples d’images MCR classées en utilisant la méthodologie proposée. Finalement, les tableaux 2.2, 2.4 et 2.5 confirment que les paramètres (alpha, entropie, variance) ne permettent pas de discriminer les images sains-lentigo, ce qui est en accord avec les résultats précédents obtenus avec le T-test.

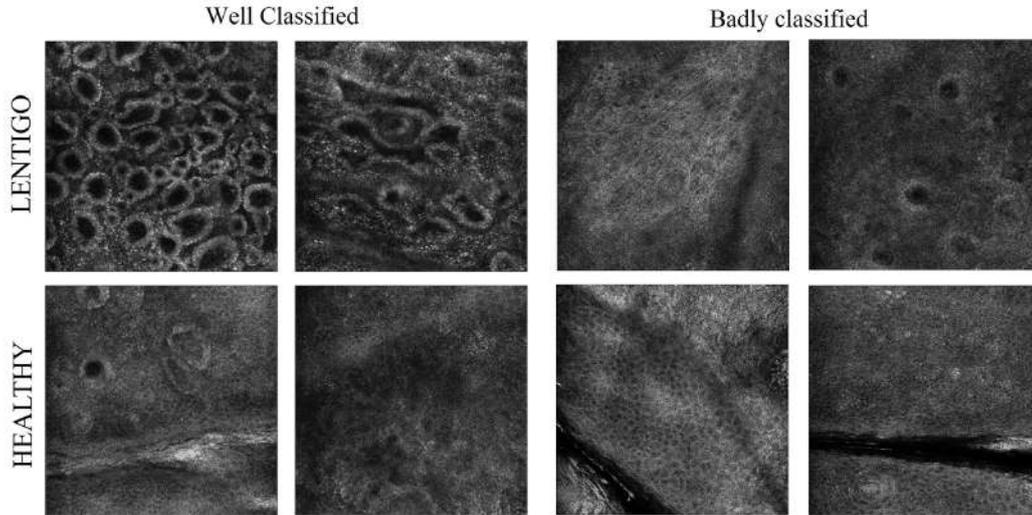


Figure 2.20: Exemples d'images MCR de patients sains et lentigo classés par le classifieur SVM.

α		Scale 1			Scale 2			Scale 3			Scale 4		
Matrice de confusion		\hat{L}	\hat{h}	Sensit Specif									
H	Lentigo	16	11	59.2 %	14	13	51.8 %	14	13	51.8%	11	16	40.7 %
	Sain	11	7	38.8 %	11	7	38.8 %	11	7	38.8%	8	10	55.5 %
	Précision	59.2%	38.8%		56%	35%		56 %	35 %		57.8%	38.4%	
	Accuracy	51.1%			46.6%			46.6 %			46.6 %		
V	Lentigo	16	11	59.2%	15	12	55.5%	13	14	48.1%	11	16	40.7%
	Sain	11	7	38.8%	11	7	38.8%	10	8	44.4%	9	9	50 %
	Précision	59.2%	38.8%		57.6 %	36.8 %		56.5%	36.3%		55%	36%	
	Accuracy	51.1%			48.8 %			46.6 %			44.4 %		
D	Lentigo	16	11	59.2 %	15	12	55.5 %	12	15	44.4%	13	14	48.1 %
	Sain	11	7	38.8 %	11	7	38.8 %	9	9	50%	10	8	44.4%
	Précision	59.2 %	38.8 %		57.6 %	36.8 %		57.1%	37.5%		56.5 %	36.3 %	
	Accuracy	51.1 %			48.8 %			46.6%			46.6%		
HVD	Lentigo	16	11	59.2%	15	12	55.5%	14	13	51.8%	12	15	44.4%
	Sain	11	7	38.8%	11	7	38.8%	10	8	44.4%	9	9	50%
	Précision	59.2%	38.8%		57.6%	36.8%		58.3%	38%		57.1%	37.5%	
	Accuracy	51.1%			48.8%			48.8 %			46.6 %		

Tableau 2.2: Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre d'échelle α . (Sensit, Specif) signifie (Sensibilité et Spécificité)

β		Scale 1			Scale 2			Scale 3			Scale 4		
Confusion matrix		\hat{L}	\hat{h}	Sensit Specif									
H	Lentigo	22	5	81.4 %	21	6	77.7 %	20	7	74%	19	8	70.3 %
	Sain	3	15	83.3 %	4	14	77.7 %	4	14	77.7%	6	12	66.6 %
	Précision	88%	75%		84%	70%		83.3 %	66.6 %		76%	60%	
	Accuracy	82.2%			77.7%			75.5 %			68.8 %		
V	Lentigo	22	5	81.4%	20	7	74%	20	7	74%	20	7	74%
	Sain	3	15	83.3%	4	14	77.7%	4	14	77.7%	5	13	72.2 %
	Précision	88%	75%		83.3 %	66.6 %		83.3%	66%		80%	65%	
	Accuracy	82.2%			75.5 %			75.5 %			73.3 %		
D	Lentigo	22	5	81.4 %	21	6	77.7 %	20	7	74%	19	8	70.3 %
	Sain	3	15	83.3 %	5	13	72.2 %	5	13	72.2%	5	13	72.2%
	Précision	88 %	75 %		80.7 %	68.4 %		80%	65%		79.1 %	61.9 %	
	Accuracy	82.2 %			75.5 %			73.3 %			71.1 %		
HVD	Lentigo	22	5	81.4%	21	6	77.7%	20	7	74%	19	8	70.3%
	Sain	3	15	83.3%	4	14	77.7%	4	14	77.7%	5	13	72.2%
	Précision	88%	75%		84%	70%		83.3%	66%		79.1%	61.9%	
	Accuracy	82.2%			77.7%			75.5 %			71.1 %		

Tableau 2.3: Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre de forme β . (Sensit, Specif) signifie (Sensibilité et Spécificité)

E		Scale 1			Scale 2			Scale 3			Scale 4		
Matrice de confusion		\hat{L}	\hat{h}	Sensit Specif									
H	Lentigo	15	12	55.5 %	15	12	55.5 %	15	12	55.5%	18	9	66.6 %
	Sain	9	9	50 %	7	11	61.1 %	7	11	61.1%	6	12	66.6 %
	Précision	62.5%	42.8%		68.1%	47.8%		68.1 %	47.8 %		75%	57.1%	
	Accuracy	53.3%			57.7%			57.7 %			66.6 %		
V	Lentigo	12	15	44.4%	12	15	44.4%	15	12	55.5%	17	10	62.9%
	Sain	7	11	61.1%	7	11	61.1%	7	11	61.1%	6	12	66.6 %
	Précision	63.1%	42.3%		63.1 %	42.3 %		68.1%	47.8%		73.9%	54.4%	
	Accuracy	51.1%			51.1 %			57.7 %			64.4 %		
D	Lentigo	12	15	44.4 %	12	15	44.4 %	14	13	51.8%	15	12	55.5 %
	Sain	7	11	61.1 %	7	11	61.1 %	6	12	66.6%	6	12	66.6%
	Précision	63.1 %	42.3 %		63.1 %	42.3 %		70%	48%		71.4 %	50 %	
	Accuracy	51.1 %			51.1 %			57.7%			60%		
HVD	Lentigo	15	12	44.4%	14	13	51.8%	15	12	55.5%	17	10	66.6%
	Sain	8	10	55.5%	7	11	61.1%	7	11	61.1%	6	12	66.6%
	Précision	65.2%	45.4%		66.6%	45.8%		68.1%	47.8%		73.9%	54.5%	
	Accuracy	55.5%			55.5%			57.7 %			64.4 %		

Tableau 2.4: Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre entropie. (Sensit, Specif) signifie (Sensibilité et Spécificité)

V		Scale 1			Scale 2			Scale 3			Scale 4		
Matrice de confusion		\hat{L}	\hat{h}	Sensit Specif									
H	Lentigo	8	19	29.6 %	8	19	29.6 %	12	15	44.4%	16	11	59.2 %
	Sain	5	13	66.6 %	5	13	66.6 %	6	12	66.6%	5	13	72.2 %
	Précision	61.5%	40.6%		61.5%	40.6%		66.6 %	44.4 %		76.1%	54.1%	
	Accuracy	46.6%			46.6%			53.3 %			64.4 %		
V	Lentigo	9	18	33.3%	10	17	37%	13	14	48.1%	15	12	55.5%
	Sain	5	13	72.2%	5	13	72.2%	5	13	72.2%	4	14	77.7 %
	Précision	64.2%	41.9%		66.6 %	43.3 %		72.2%	48.1%		78.9%	53.8%	
	Accuracy	48.8%			51.1 %			57.7 %			64.4 %		
D	Lentigo	8	19	29.6 %	8	19	29.6 %	10	17	37%	12	15	44.4 %
	Healthy	4	14	77.7 %	4	14	77.7 %	4	14	77.7%	3	15	83.3%
	Precision	66.6 %	42.4 %		66.6 %	42.4 %		71.4%	45.1%		80 %	50 %	
	Accuracy	48.8 %			48.8 %			53.3%			60%		
HVD	Lentigo	8	19	29.6%	8	19	29.6%	11	16	40.7%	13	14	48.1%
	Sain	4	14	77.7%	4	14	77.7%	4	14	77.7%	3	15	83.3%
	Précision	66.6%	42.4%		66.6%	42.4%		73.3%	46.6%		81.2%	51.7%	
	Accuracy	48.8%			42.4%			55.5 %			62.2 %		

Tableau 2.5: Résultats de classification SVM pour les coefficients H, V, D, HVD à 4 échelles différentes pour le paramètre variance. (Sensit, Specif) signifie (Sensibilité et Spécificité)

2.4 Conclusions

Ce chapitre a étudié l'utilisation potentielle des propriétés statistiques des coefficients d'ondelettes des images de microscopie confocale à réflectance pour caractériser la peau lentigo. La méthode proposée détermine les paramètres d'échelle et de forme d'une distribution gaussienne généralisée associée à chaque bande et chaque échelle d'images acquises à différentes profondeurs. Ces paramètres sont ensuite utilisés pour distinguer les tissus sains de ceux atteints de lentigo. Un test T paramétrique a été conçu pour évaluer la signification statistique des différences observées entre les images des patients sains et lentigo. Le test d'hypothèse proposé a été effectué sur une base de données de 2250 images réelles associées à 45 patients. Les résultats obtenus ont montré que le paramètre de forme β est bien adapté pour discriminer les tissus sains et lentigo. Ils ont également montré qu'il n'y avait pas de direction préférée, et donc qu'il vaut mieux utiliser le vecteur conjoint *HVD* avec les premières échelles pertinentes pour avoir une bonne performance de caractérisation. Il est intéressant de noter que cette étude a révélé que le paramètre β caractérise le lentigo à des profondeurs aux environs de $50 \mu m$ ce qui correspond à l'emplacement de la jonction dermo-épidermique. Ce résultat est en accord avec le fait clinique que le lentigo désorganise significativement cette couche de peau. Dans le

chapitre suivant on va s'intéresser à l'analyse des propriétés statistiques des images MCR directement dans le domaine de l'image. Les paramètres associés à ces lois pourraient alors compléter et confirmer les résultats obtenus avec le modèle statistique défini par les paramètres d'échelle et de la forme de la loi gaussienne généralisée associée aux coefficients d'ondelettes considérés dans ce chapitre.

CHAPITRE 3

Modélisation statistique et classification d'images de microscopie confocale à réflectance

Sommaire

3.1	Introduction	63
3.2	Méthode proposée	64
3.2.1	Estimation statistique	64
3.2.2	Caractérisation utilisant le test paramétrique de Student et la distance de Kullback-Leibler	67
3.3	Expériences	68
3.3.1	Performance de l'algorithme d'estimation proposé	68
3.3.2	Identification des profondeurs caractéristiques	76
3.3.3	Importance statistique avec T-Test	79
3.3.4	Classification d'images sains et avec lentigo	82
3.4	Conclusions	85

3.1 Introduction

Dans ce chapitre, nous avons proposé une méthode permettant la modélisation statistique des intensités des images MCR à l'aide d'un estimateur qui s'avère être plus rapide et plus performant que ceux disponibles dans la littérature.

La première contribution de notre travail est un modèle statistique qui permet la caractérisation des tissus sous-jacents. La variabilité des intensités de pixels d'une image MCR est représentée par une distribution gamma généralisée (FGD), dont les paramètres sont utilisés comme caractéristiques pour la classification des images confocales saines et lentigo. La représentation des images confocales dans un espace 3D des paramètres agit comme une technique de réduction de dimension intéressante

permettant d'implémenter des algorithmes de classification en temps quasi-réel. Le modèle statistique FGD est ajusté aux intensités des images MCR à différentes profondeurs, afin d'identifier les profondeurs de la peau à laquelle la détection et la caractérisation de lentigo sont les plus importantes. Une analyse quantitative soutenue par un classificateur SVM est menée pour évaluer la performance de la caractérisation proposée. Une deuxième contribution de ce travail est un nouvel algorithme d'estimation pour les paramètres FGD, basé sur une approche de gradient naturel [Amari and Douglas, 1998]. La principale propriété de cet algorithme est sa convergence rapide par rapport à d'autres techniques existantes, permettant de traiter de grandes bases de données avec des coûts de calcul réduits. Cette approche est également connue sous le nom de Fisher scoring [Halimi et al., 2016]. Elle met à jour les paramètres dans un espace Riemannien, ce qui entraîne une convergence rapide vers un minimum local de la fonction coût d'intérêt [Pereyra et al., 2013]. Le modèle proposé et l'algorithme d'estimation sont validés en utilisant des images MCR synthétiques et réelles, résultant d'une étude clinique contenant des patients sains et lentigo. Les résultats obtenus sont très prometteurs.

Le chapitre est organisé de la manière suivante. La Section 2 présente la méthode proposée pour l'identification de lentigo. Les résultats de la simulation sont présentés et analysés dans la Section 3. Les conclusions et perspectives pour les travaux futurs sont finalement rapportées dans la Section 4.

3.2 Méthode proposée

Cette Section présente l'approche proposée basée sur l'utilisation de la distribution gamma généralisée pour classer et caractériser les images MCR saines et celles atteintes de lentigo. Elle est constituée de trois étapes qui sont résumées dans la Figure 3.1 et décrites dans les Sections suivantes.

3.2.1 Estimation statistique

Distribution gamma généralisée

Dans cette partie, nous proposons d'utiliser les propriétés statistiques des intensités de pixels pour détecter la présence de lentigo dans les images MCR. Plus précisément, nous considérons une image MCR vectorielle décrite comme $\boldsymbol{x} = (x_1, \dots, x_N)^T$ associé à N pixels et nous supposons que la distribution de ces intensités de pixels est une distribution gamma généralisée [Johnson et al., 1994,

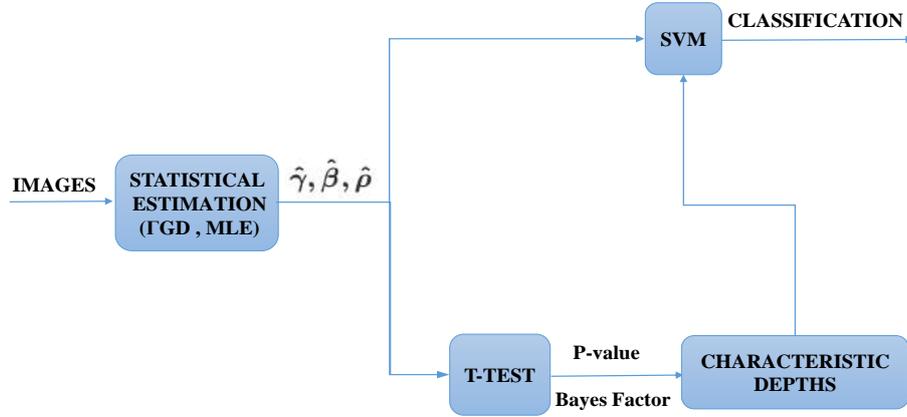


Figure 3.1: Méthode de classification proposée.

[Jonkman et al., 2009]. La Γ GD dépend du vecteur de paramètre $\boldsymbol{\theta} = (\gamma, \beta, \rho)$ et est définie comme suit :

$$f(x_n; \boldsymbol{\theta}) = \frac{(x_n - \gamma)^{\rho-1}}{\beta^\rho \Gamma(\rho)} \exp\left(-\frac{x_n - \gamma}{\beta}\right) \mathcal{I}_{] \gamma, +\infty[}(x_n) \quad (3.1)$$

où \mathcal{I}_A est la fonction indicatrice définie sur l'ensemble A , $\beta > 0, \rho > 0$ et $\Gamma(\cdot)$ est la fonction gamma [Abramowitz and Stegun, 1970].

Nous montrerons dans la Section 3.3.1 que la densité (3.1) est parfaitement adaptée à la distribution des valeurs d'intensité d'une image MCR. La Section suivante introduit une nouvelle méthode d'estimation statistique basée sur le principe du maximum de vraisemblance (ML) et la descente du gradient naturel. Cette méthode permet d'estimer le paramètre $\boldsymbol{\theta}$ à partir du vecteur \boldsymbol{x} .

Méthode du maximum de vraisemblance

La méthode d'estimation du maximum de vraisemblance (ML) consiste à maximiser la probabilité des échantillons observés par rapport aux paramètres du modèle inconnu [Kay, 1993]. En supposant que les observations x_n sont indépendantes, la fonction de vraisemblance de l'échantillon $\boldsymbol{x} = (x_1, \dots, x_N)^T$ est définie comme suit :

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \prod_{n=1}^N f(x_n; \boldsymbol{\theta}). \quad (3.2)$$

La fonction log-vraisemblance est alors définie par :

$$\mathcal{L}(\boldsymbol{\theta}) = \log [f(\mathbf{x}; \boldsymbol{\theta})] = -N\rho \log(\beta) - N \log [\Gamma(\rho)] + (\rho - 1) \sum_{n=1}^N \log(x_n - \gamma) - \frac{1}{\beta} \sum_{n=1}^N (x_n - \gamma). \quad (3.3)$$

Les dérivées partielles de la fonction log-vraisemblance par rapport à γ , β et ρ peuvent être facilement calculées, conduisant à :

$$\begin{cases} \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \gamma} = -(\rho - 1) \sum_{n=1}^N (x_n - \gamma)^{-1} + \frac{N}{\beta} \\ \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \beta} = \frac{-N\rho}{\beta} + \frac{1}{\beta^2} \sum_{n=1}^N (x_n - \gamma) \\ \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \rho} = -N\Psi(\rho) - N \log(\beta) + \sum_{n=1}^N \log(x_n - \gamma) \end{cases} \quad (3.4)$$

où $\Psi(z) = \Gamma'(z) / \Gamma(z)$ est la fonction digamma [Abramowitz and Stegun, 1970].

La maximisation du log-vraisemblance dans (3.3) par rapport à $\boldsymbol{\theta}$ peut être déterminée en utilisant différents algorithmes numériques. Jonkman et al. [Jonkman et al., 2009] ont proposé d'utiliser un algorithme de descente du gradient défini par la récursion suivante:

$$\boldsymbol{\theta}^{t+1} = \boldsymbol{\theta}^t + \lambda A(\boldsymbol{\theta}^t) \nabla \mathcal{L}(\boldsymbol{\theta}^t) \quad (3.5)$$

où ∇ est l'opérateur gradient, λ est le pas et A une matrice de préconditionnement qui dépend de $\boldsymbol{\theta}^t$ (e.g., définie à partir du Hessien du log-vraisemblance).

Dans notre travail, nous avons considéré une autre voie définie par la méthode de descente du gradient naturel afin d'estimer le vecteur de paramètres $\boldsymbol{\theta}$. Contrairement à la méthode de Newton, la récursion du gradient naturel ne suppose pas une fonction de coût localement quadratique et permet alors de maximiser la vraisemblance qui est asymptotiquement efficace [Amari and Douglas, 1998]. L'idée principale est de mettre à jour les paramètres dans un espace Riemannien, ce qui entraîne une convergence rapide vers un minimum local de la fonction objective [Pereyra et al., 2013]. La récursion de gradient naturel est définie par :

$$\boldsymbol{\theta}^{t+1} = \boldsymbol{\theta}^t + \frac{\lambda}{\|F^{-1}(\boldsymbol{\theta}^t) \nabla \mathcal{L}(\boldsymbol{\theta}^t)\|} F^{-1}(\boldsymbol{\theta}^t) \nabla \mathcal{L}(\boldsymbol{\theta}^t) \quad (3.6)$$

où $\nabla \mathcal{L}(\boldsymbol{\theta}^t)$ est le gradient défini par (3.4) et $F(\boldsymbol{\theta}^t)$ est la matrice d'information de Fisher (FIM) définie par :

$$F(\boldsymbol{\theta}^t) = - \begin{pmatrix} E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \gamma^2} \right) & E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \gamma \partial \beta} \right) & E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \gamma \partial \rho} \right) \\ E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \beta \partial \gamma} \right) & E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \beta^2} \right) & E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \beta \partial \rho} \right) \\ E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \rho \partial \gamma} \right) & E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \rho \partial \beta} \right) & E \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}^t)}{\partial \rho^2} \right) \end{pmatrix}. \quad (3.7)$$

Les calculs directs (plus de détail dans Annexe C) conduisent à :

$$F(\boldsymbol{\theta}) = N \begin{pmatrix} \frac{1}{\beta^2(\rho-2)} & \frac{1}{\beta^2} & \frac{1}{\beta(\rho-1)} \\ \frac{1}{\beta^2} & \frac{\rho}{\beta^2} & \frac{1}{\beta} \\ \frac{1}{\beta(\rho-1)} & \frac{1}{\beta} & \Psi'(\rho) \end{pmatrix} \quad (3.8)$$

où Ψ' désigne la fonction trigamma. L'intérêt de l'utilisation de cette méthode de gradient naturel pour les distributions gamma généralisées sera clarifié dans la Section 3.3.1. À notre connaissance, c'est la première fois qu'une méthode de gradient naturel est appliquée aux distributions gamma généralisées.

3.2.2 Caractérisation utilisant le test paramétrique de Student et la distance de Kullback-Leibler

Afin de caractériser la présence de lentigo dans les images MCR, nous pouvons adapter les distributions FGD aux images MCR à différentes profondeurs. Un test paramétrique T-test (Annexe A) est alors appliqué aux paramètres estimés γ, β, ρ , ainsi qu'à l'entropie et la variance qui en résultent afin d'identifier les profondeurs de la peau à laquelle la détection et la caractérisation de lentigo sont les plus significatives. Il convient de noter que l'entropie (dénotée par E) et la variance (dénotée par V) sont données par

$$E = \log [\beta \Gamma(\rho)] + (1 - \rho) \psi(\rho) + \rho, \quad V = \rho \beta^2.$$

. - Plus de détails sur les calculs de l'entropie et variance peuvent être trouvés dans l'annexe D.

La pseudo-distance de Kullback-Leibler a aussi été appliquée aux différentes distributions afin de trouver les profondeurs qui permettent la séparation des deux populations.

Les paramètres Γ GD estimés associés à ces profondeurs caractéristiques peuvent ensuite être utilisés pour classer les patients sains de lentigo en utilisant un algorithme de classification SVM [Gunn et al., 1998]. Les résultats de simulation confirmant ces propriétés sont présentés dans la Section 3.3.2 .

3.3 Expériences

3.3.1 Performance de l'algorithme d'estimation proposé

Données synthétiques

Cette Section évalue la performance de l'algorithme d'estimation proposé sur des données synthétiques. Le matériel utilisé est constitué d'un PC de qualité grand public avec un processeur Intel(R) Core(TM) i7-4860HQ CPU 2.4 GHz, 32 Go de RAM et une carte graphique Nvidia GeForce GTX 980m. L'algorithme a été exécuté à l'aide du logiciel MATLAB R2014b. Deux expériences ont été menées : la première est basée sur la génération de vecteurs distribués suivant une Γ GD en utilisant 3.1. La taille de ces vecteurs varie de $N = 40$ à 1000 et les paramètres ont été fixés à $\boldsymbol{\theta} = (\gamma, \beta, \rho)^T = (2, 15, 4)^T$. La deuxième expérience a été faite en considérant un paramètre variable et les deux autres constants et cela pour pour deux cas $N=100$ et $N=10000$. L'algorithme a été exécuté pour $M=1000$ réalisations pour la première expérience et $M=100$ pour la seconde. Les paramètres estimés $\hat{\boldsymbol{\theta}}$ ont ensuite été utilisés pour calculer le biais, la variance et l'erreur quadratique moyenne (RMSE) qui sont définis comme suit :

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \\ \hat{\rho} \end{pmatrix} \quad (3.9)$$

$$\text{Bias} = \boldsymbol{\theta} - \frac{\sum_{i=1}^M \hat{\boldsymbol{\theta}}(i)}{M} \quad (3.10)$$

$$\text{Variance} = \frac{1}{M} \sum_{i=1}^M \left(\hat{\boldsymbol{\theta}}(i) - \bar{\boldsymbol{\theta}} \right)^2, \quad \text{avec } \bar{\boldsymbol{\theta}} = \frac{1}{M} \sum_{i=1}^M \hat{\boldsymbol{\theta}}(i) \quad (3.11)$$

$$\text{RMSE} = \text{Bias}^2 + \text{Variance} = \frac{1}{M} \sum_{i=1}^M \left(\hat{\boldsymbol{\theta}}(i) - \boldsymbol{\theta} \right)^2 \quad (3.12)$$

où $\hat{\boldsymbol{\theta}}(i)$ désigne les paramètres estimés pour la i th réalisation.

Notre méthode a été comparée à deux autres méthodes, la première a été proposée dans [Jonkman et al., 2009] et elle consiste à appliquer la méthode de descente de gradient de Newton en utilisant le Hessien comme matrice pré-conditionnée. La deuxième (désignée par méthode analytique) a utilisé la solution donnée lorsqu'on annule les dérivées (voir (3.4)) pour avoir une nouvelle formulation des paramètres β and ρ uniquement en fonction du paramètre γ . Le problème résultant dépend uniquement du paramètre γ comme suit:

$$-N\Psi(\rho) - N \log(\beta) + \sum_{n=1}^N \log(x_n - \gamma) = 0 \quad (3.13)$$

avec

$$\beta = \frac{\left[\sum_{n=1}^N (x_n - \gamma) \sum_{n=1}^N \left(\frac{1}{x_n - \gamma} \right) \right] - N^2}{N \sum_{n=1}^N \left(\frac{1}{x_n - \gamma} \right)}$$

$$\rho = \frac{\sum_{n=1}^N (x_n - \gamma) \sum_{n=1}^N \left(\frac{1}{x_n - \gamma} \right)}{\left[\sum_{n=1}^N (x_n - \gamma) \sum_{n=1}^N \left(\frac{1}{x_n - \gamma} \right) \right] - N^2}.$$

La résolution de (3.13) a ensuite été effectuée à l'aide d'un algorithme de descente de gradient de Newton. Le biais et le RMSE des estimateurs sont représentés (en échelle logarithmique pour une meilleure lisibilité) dans les Figures 3.2 et 3.3 avec les temps d'exécution associés. Les trois méthodes ont utilisé la même initialisation pour les paramètres $\hat{\boldsymbol{\theta}}$ donnés par la "pseudo méthode des moments" (voir [Jonkman et al., 2009] pour plus de détails).

La méthode proposée basée sur une récursion de gradient naturel fournit des RMSE et des biais inférieurs pour un petit nombre d'échantillons, c'est-à-dire pour $N \in \{40, \dots, 300\}$. Cela prouve que notre méthode est plus robuste que les autres méthodes pour un petit nombre d'échantillons. La descente de gradient naturel fournit également une convergence plus rapide par rapport aux autres méthodes avec une réduction significative du coût de calcul pour toute taille d'échantillon. Ce résultat

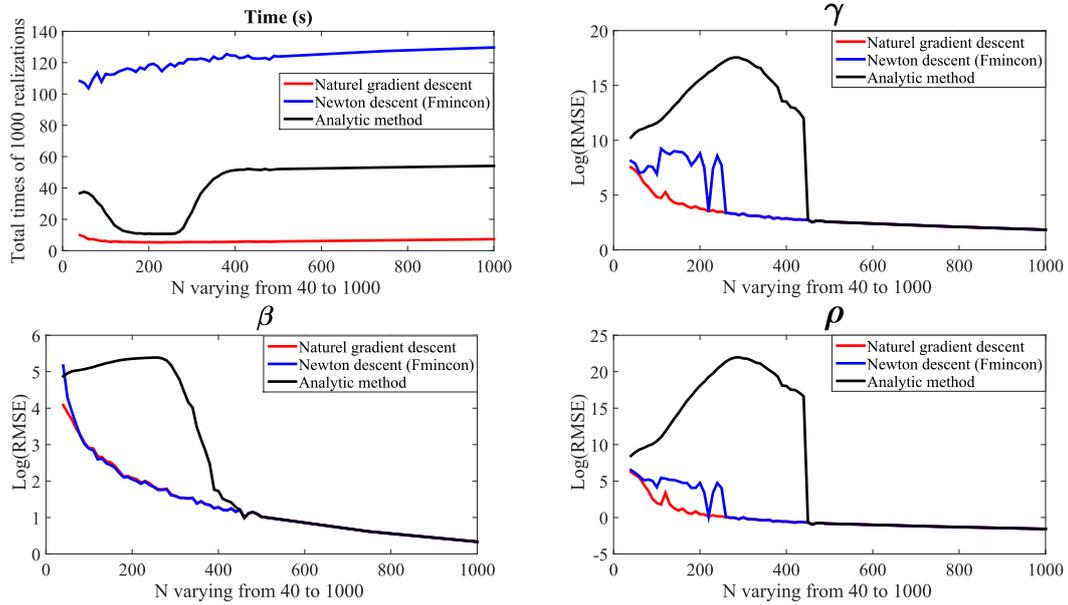


Figure 3.2: Evolution du RMSE et du temps total des trois méthodes pour N variant de 40 à 1000.

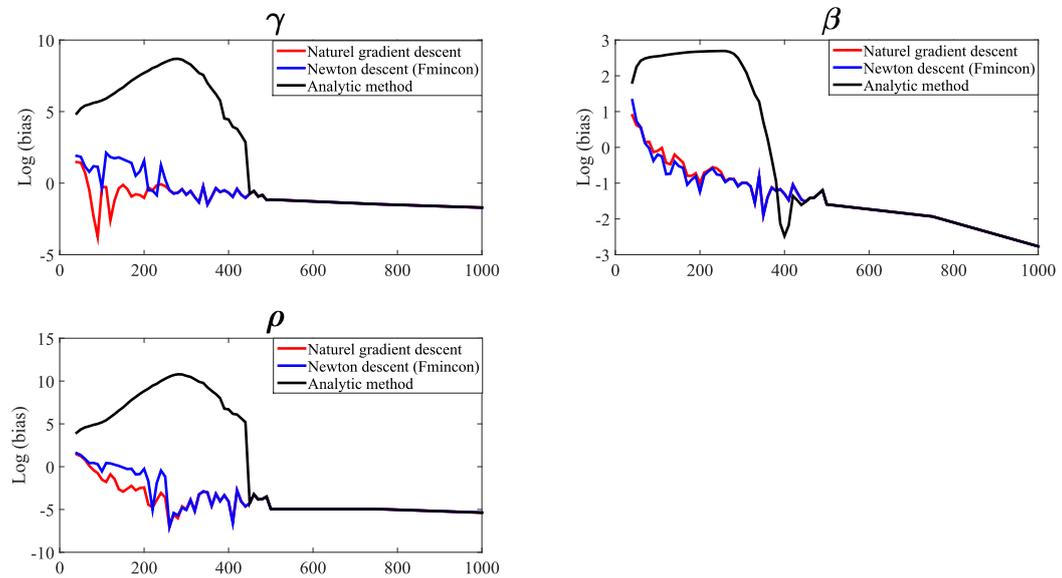


Figure 3.3: Evolution du biais des trois méthodes pour N variant de 40 à 1000.

est intéressant car il permet de traiter plus facilement de grandes bases de données d'images MCR. Ces résultats mettent en évidence les bonnes performances de la stratégie proposée pour l'estimation des paramètres de la Γ GD.

La deuxième expérience a été faite en considérant $N=100$ et $N=10000$ échantillons et en faisant varier un des paramètres (γ , β , ρ) tout en fixant les deux autres.

Nous avons considéré 10 valeurs pour le paramètre variable et avons réalisé $M = 100$ réalisations pour estimer $\hat{\theta}$.

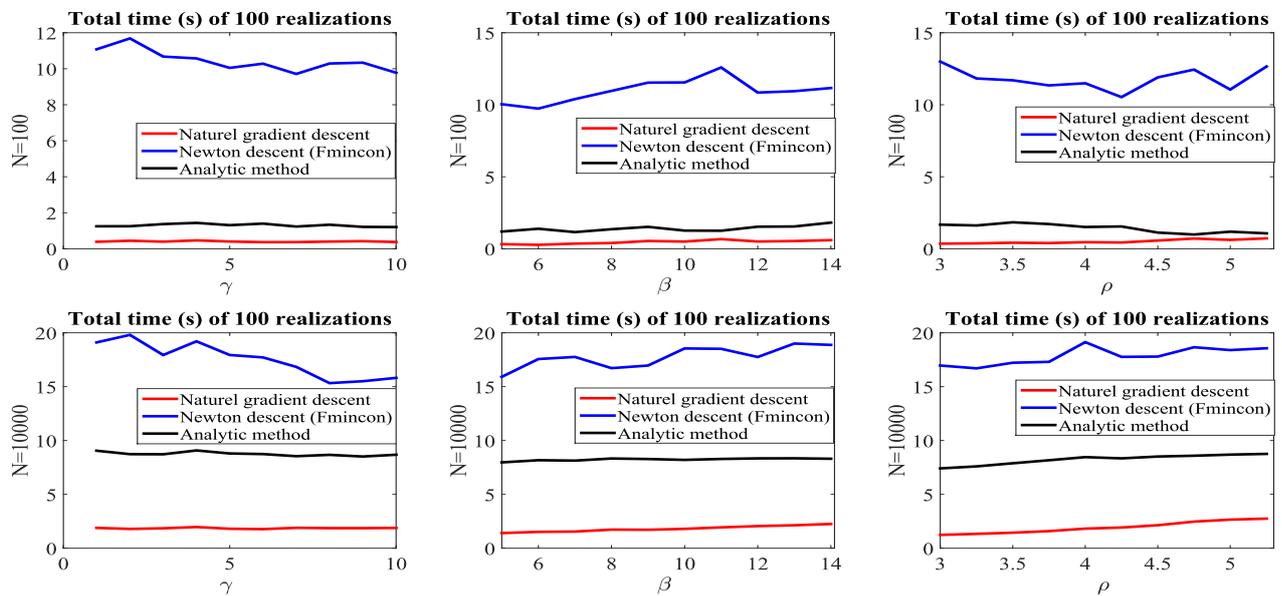


Figure 3.4: Evolution du temps total des trois méthodes pour $N=10000$ et $N=100$ lorsque les paramètres (γ , β , ρ) varient .

Les résultats obtenus pour la deuxième expérience (Figures 3.4 à 3.8) confirment les conclusions de la première expérience. Notre méthode a une convergence plus rapide par rapport aux autres méthodes et ceci indépendamment du nombre d'échantillons et est plus robuste que les autres méthodes dans le cas d'un petit nombre d'échantillons ($N=100$).

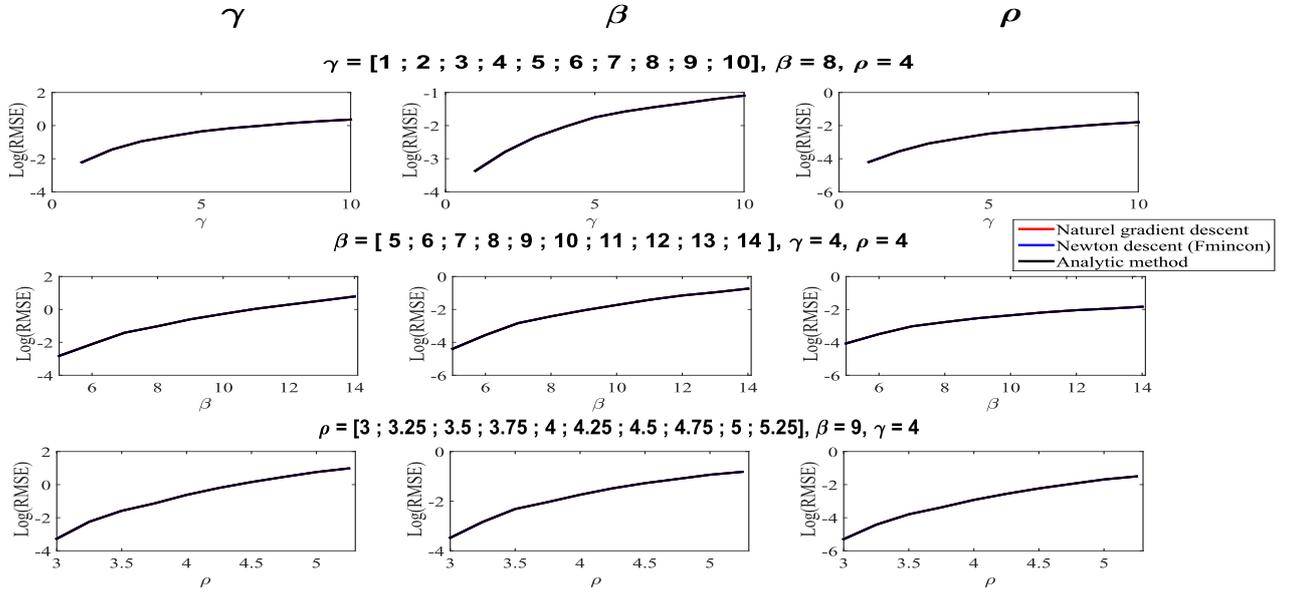


Figure 3.5: Evolution du paramètre RMSE des trois méthodes pour $N=10000$ lorsque les paramètres (γ, β, ρ) varient.

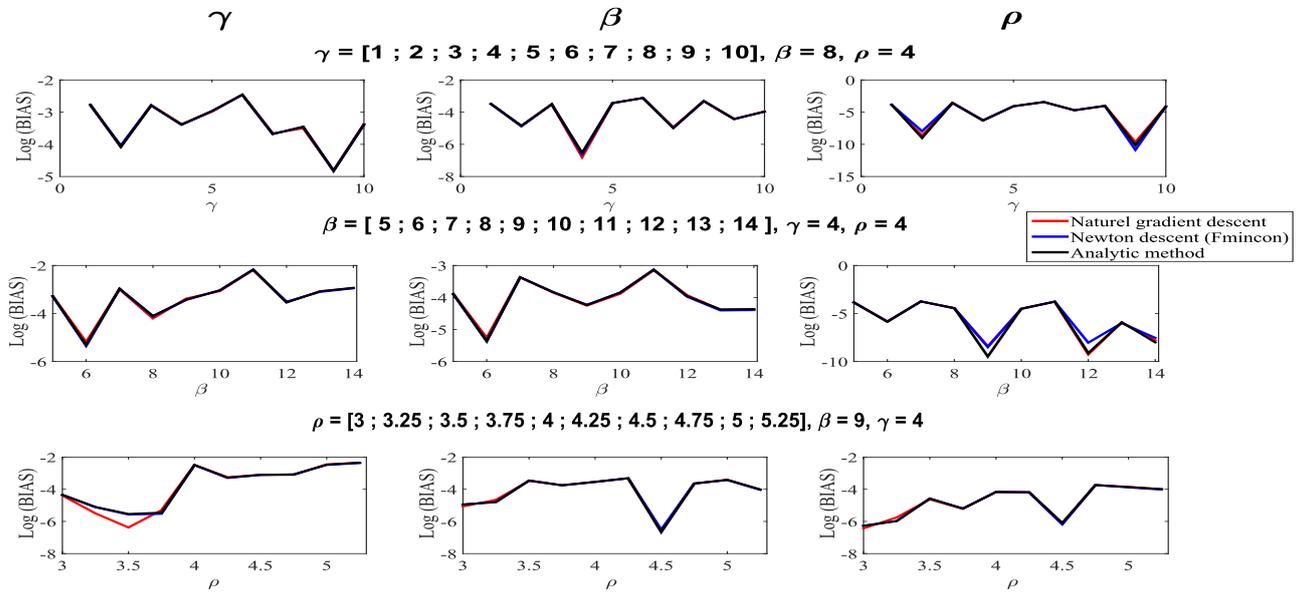


Figure 3.6: Evolution du paramètre biais des trois méthodes $N=10000$ lorsque les paramètres (γ, β, ρ) varient.

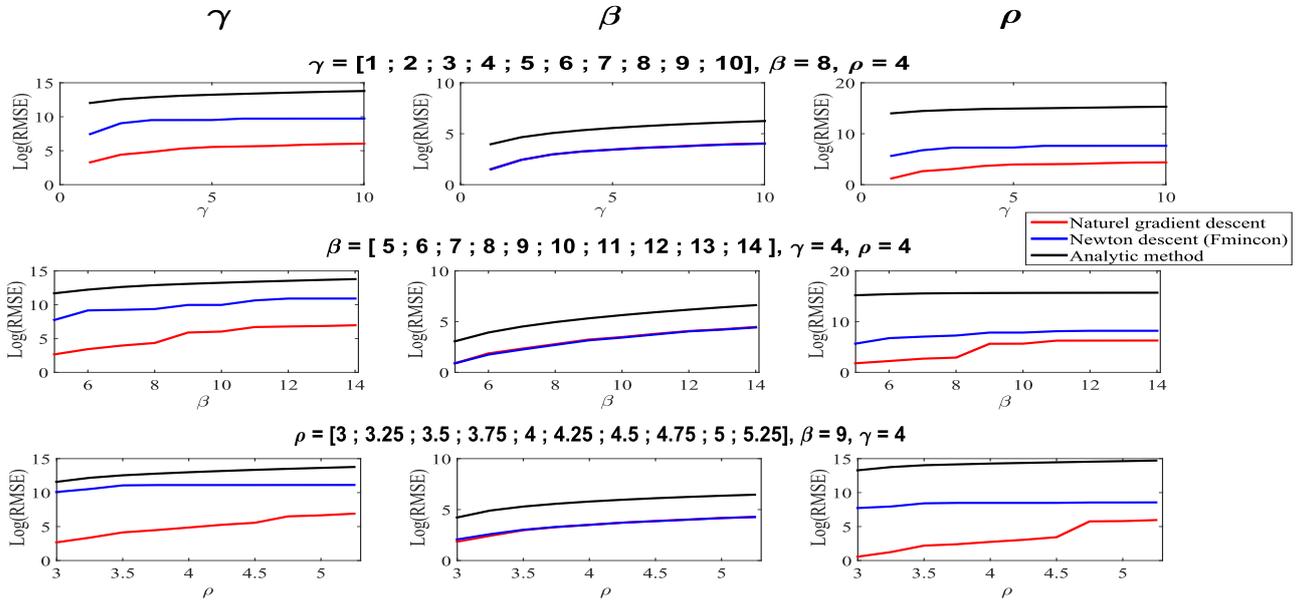


Figure 3.7: Evolution du paramètre RMSE des trois méthodes pour $N=100$ lorsque les paramètres (γ, β, ρ) varient.

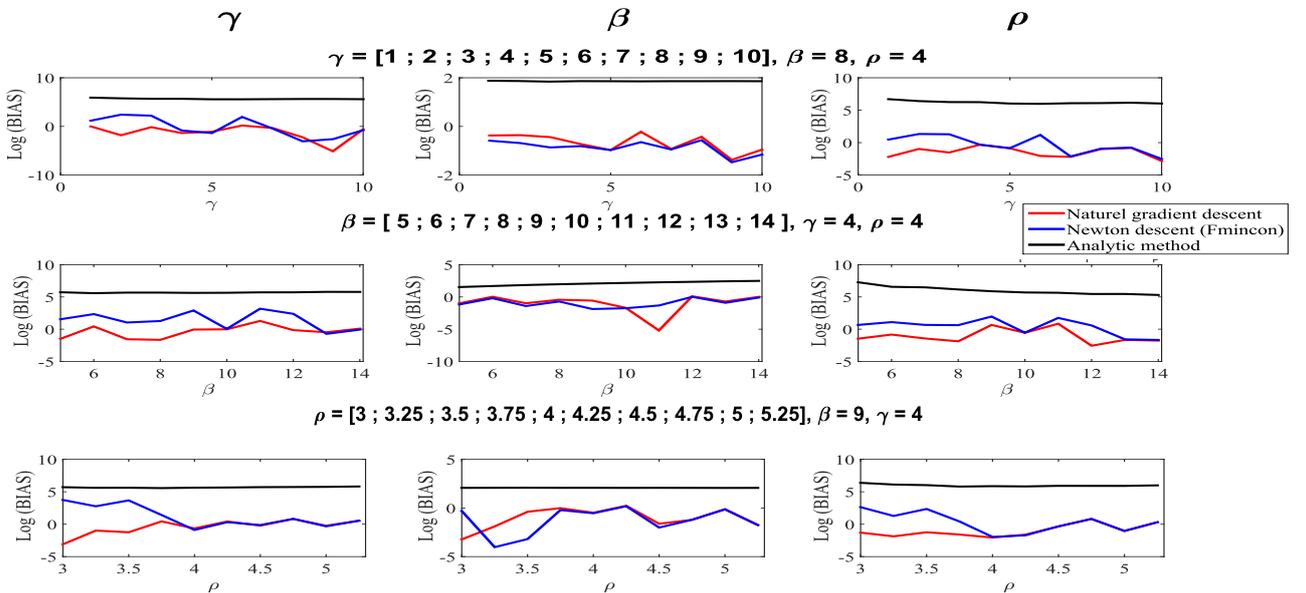


Figure 3.8: Evolution du paramètre biais des trois méthodes pour $N=100$ lorsque les paramètres (γ, β, ρ) varient.

Données réelles

Cette Section est consacrée à la validation de l'algorithme proposé lorsqu'il est appliqué à des images MCR réelles déjà défini dans la Section 2.3.1. La Figure 3.9 compare les histogrammes des intensités des images MCR avec les distributions Γ GD estimées à 3 profondeurs représentatives. Cette Figure concerne deux patients arbitraires sain et lentigo. La loi Γ GD est clairement adaptée aux intensités des images MCR, tant pour les images saines que lentigo. Ceci est confirmé quantitativement dans la Figure (3.10) qui montre les résultats de l'ajustement des données à la distribution Γ GD en utilisant le test Kolmogorov-Smirnov (KS). Le score statistique KS moyen de l'ensemble de la population (45 patients) a été calculé à chaque profondeur. On remarque les excellents scores avec des valeurs KS très proches de zéro ce qui montre que la modélisation par la loi Γ GD capte bien les propriétés statistiques des images MCR. Il est important de noter que les paramètres d'échelle et de forme estimés des distributions associées aux images des patients sains et lentigo sont significativement différents comme le montre le tableau 3.1, ce qui va nous permettre par la suite de classifier ces images.

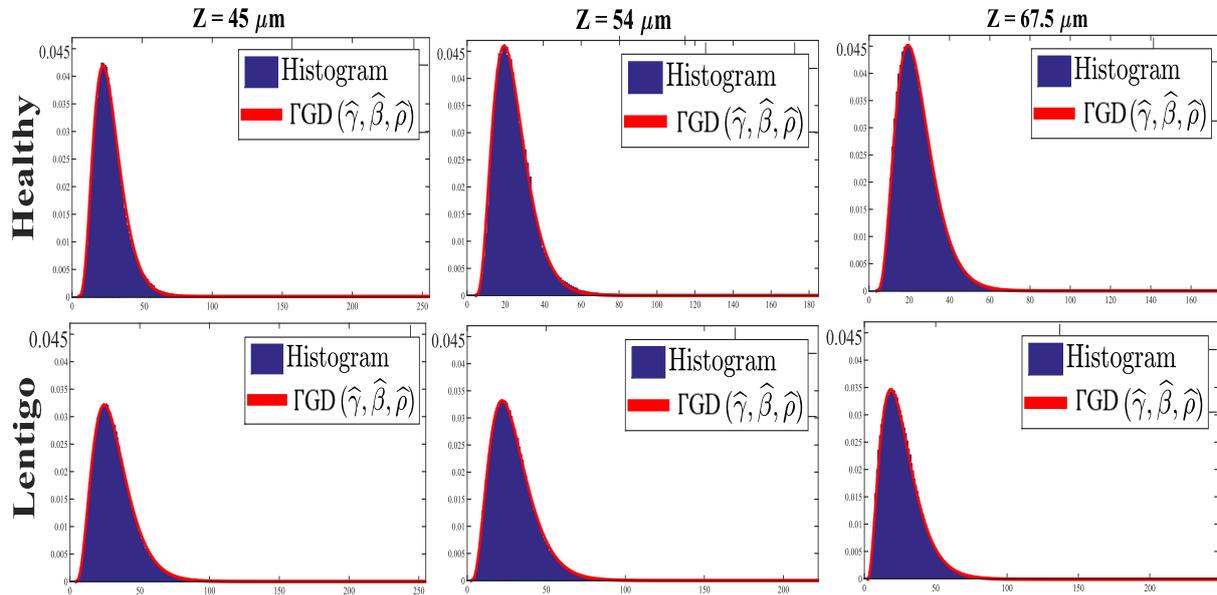


Figure 3.9: Histogrammes des intensités d'images MCR. La Figure représente les données de deux patients arbitraires sain et lentigo (#6 and #38) à trois profondeurs représentatives (une profondeur par colonne).

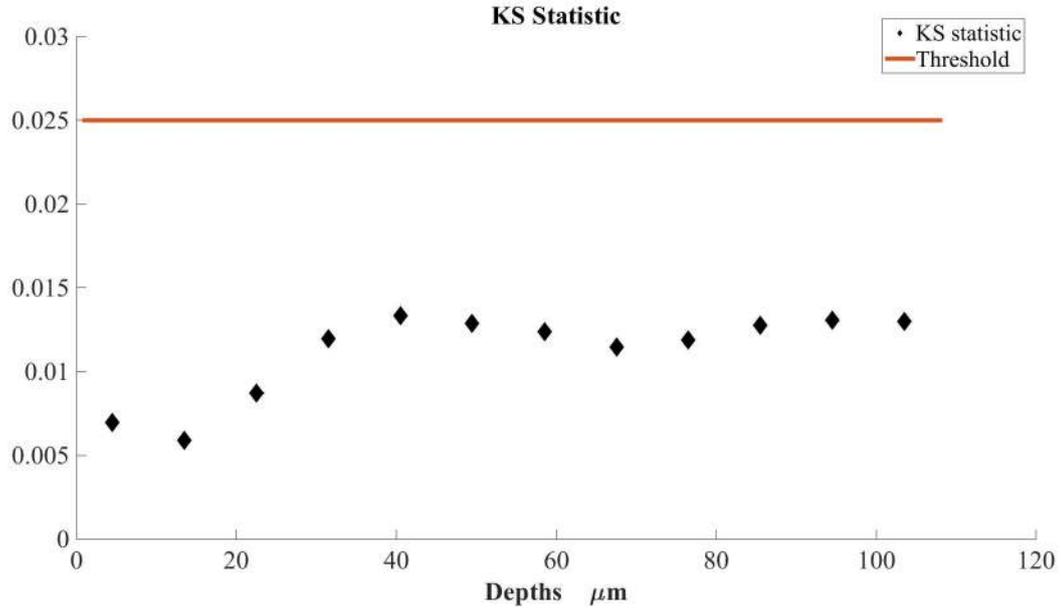


Figure 3.10: Evaluation de l'adaptation du FGD aux pixels. Statistique KS moyenne pour l'ensemble de la population et pour toutes les profondeurs. Les scores sont excellents pour toutes les configurations.

Tableau 3.1: Paramètres FGD estimés associés aux paramètres FGD illustrés dans la Figure 3.9.

Profondeur	Sain			Lentigo		
	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\rho}$
$Z = 45 \mu\text{m}$	4.59	4.96	4.51	4.27	7.19	3.81
$Z = 54 \mu\text{m}$	4.63	4.79	4.09	2.70	7.03	3.74
$Z = 67.5 \mu\text{m}$	3.68	4.73	4.30	2.51	7.59	3.11

3.3.2 Identification des profondeurs caractéristiques

Les distributions Γ GD ont été adaptées à chacune des 2250 images et à chaque profondeur. Ayant acquis deux groupes de 25 images pour chaque patient, l'un des deux groupes a été sélectionné au hasard pour l'analyse. Les paramètres moyens $\bar{\theta}_{Healthy}$ et $\bar{\theta}_{Lentigo}$ ainsi que la distance de Kulback-Leibler ont ensuite été calculés à chaque profondeur pour les patients sains et lentigo, respectivement.

Pour tenir compte de la variabilité, le processus de sélection d'un groupe pour chaque patient a été répété 300 fois. Les résultats moyens et les écarts-types sont présentés dans les Figures 3.11, 3.12, 3.13 où il est clairement montré que les distances Kulback-Leibler ainsi que les paramètres β et ρ permettent de discriminer des images saines de lentigo pour des profondeurs comprises entre $40 \mu m$ et $60 \mu m$, avec une différence maximale aux environs de $50 \mu m$. A l'inverse, le paramètre γ ne permet pas de distinguer les patients sains et lentigo. La Figure 3.14 montre deux ensembles d'images associés à six patients sains (#1, #2, #3, #4, #5, #6) et six patients avec lentigo (#31, #33, #37, #38, #40, #44). On peut observer des images plus texturées en présence de lentigo à la profondeur DEJ (correspondant à $54 \mu m$), comme prévu.

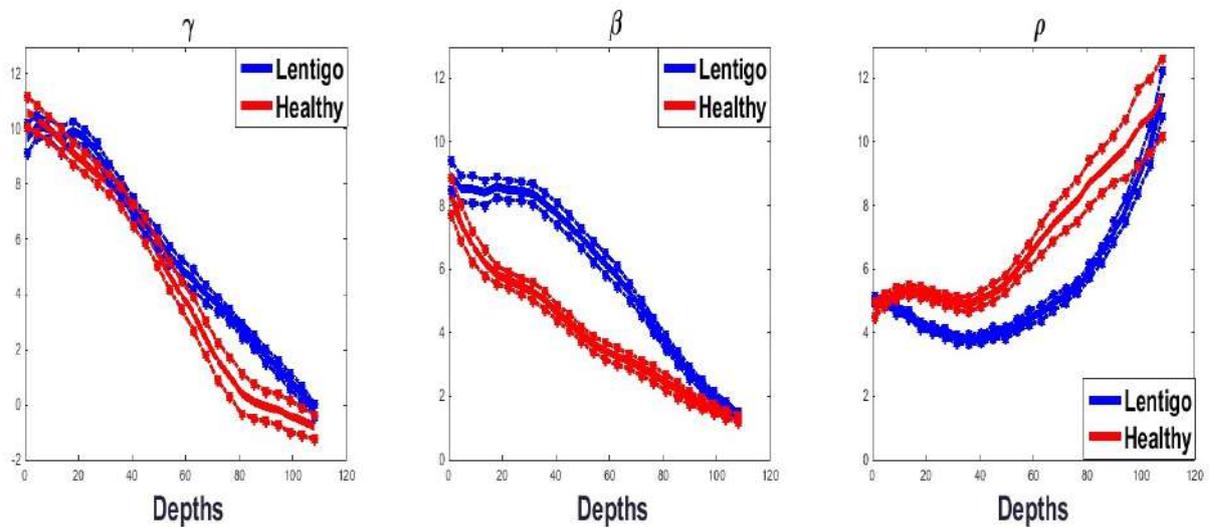


Figure 3.11: Évolution des paramètres moyens $\hat{\gamma}$, $\hat{\beta}$ et $\hat{\rho}$ tout au long de la profondeur. Les valeurs de γ sont trop similaires pour les patients sains et lentigo et ne peuvent pas être utilisées pour la discrimination. Les paramètres β et ρ montrent une différence significative pour les profondeurs entre $30 \mu m$ et $76 \mu m$, avec une différence maximale aux environs de $50 \mu m$. Notre conclusion est que les paramètres β et ρ peuvent discriminer les tissus sains et lentigo de la peau.

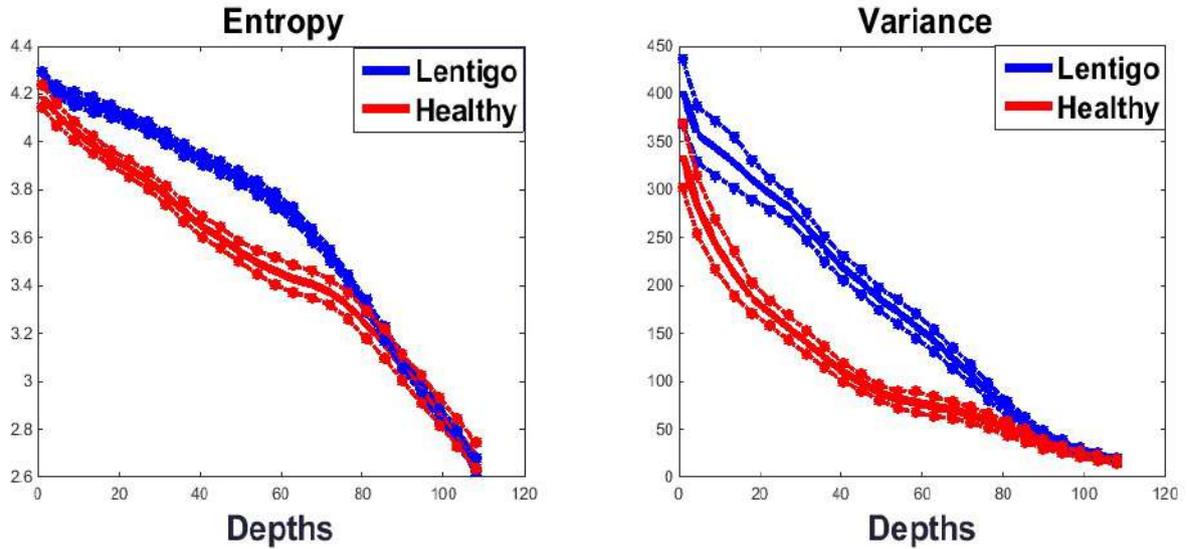


Figure 3.12: Évolution des paramètres moyens entropie et variance tout au long de la profondeur. Les paramètres E et V montrent une différence significative pour les profondeurs entre $30 \mu m$ et $76 \mu m$, avec une différence maximale aux environs de $50 \mu m$.

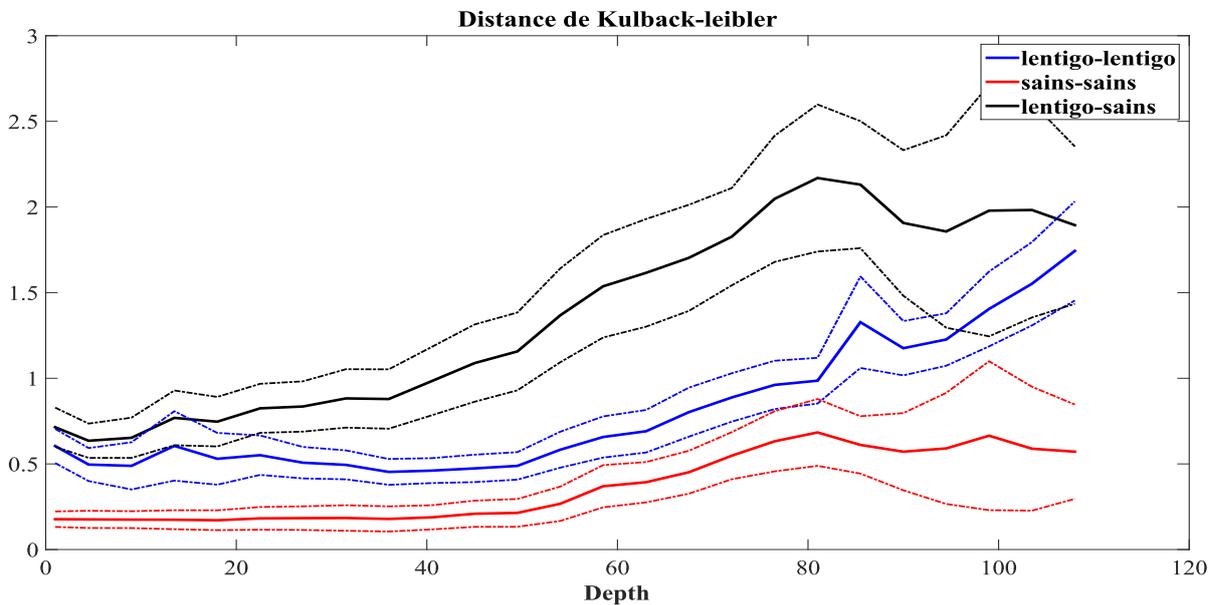


Figure 3.13: Évolution des distances Kulaback-Leibler moyennes tout au long de la profondeur. Les distances montrent une différence significative pour les profondeurs entre $30 \mu m$ et $76 \mu m$, avec une différence maximale aux environs de $50 \mu m$.

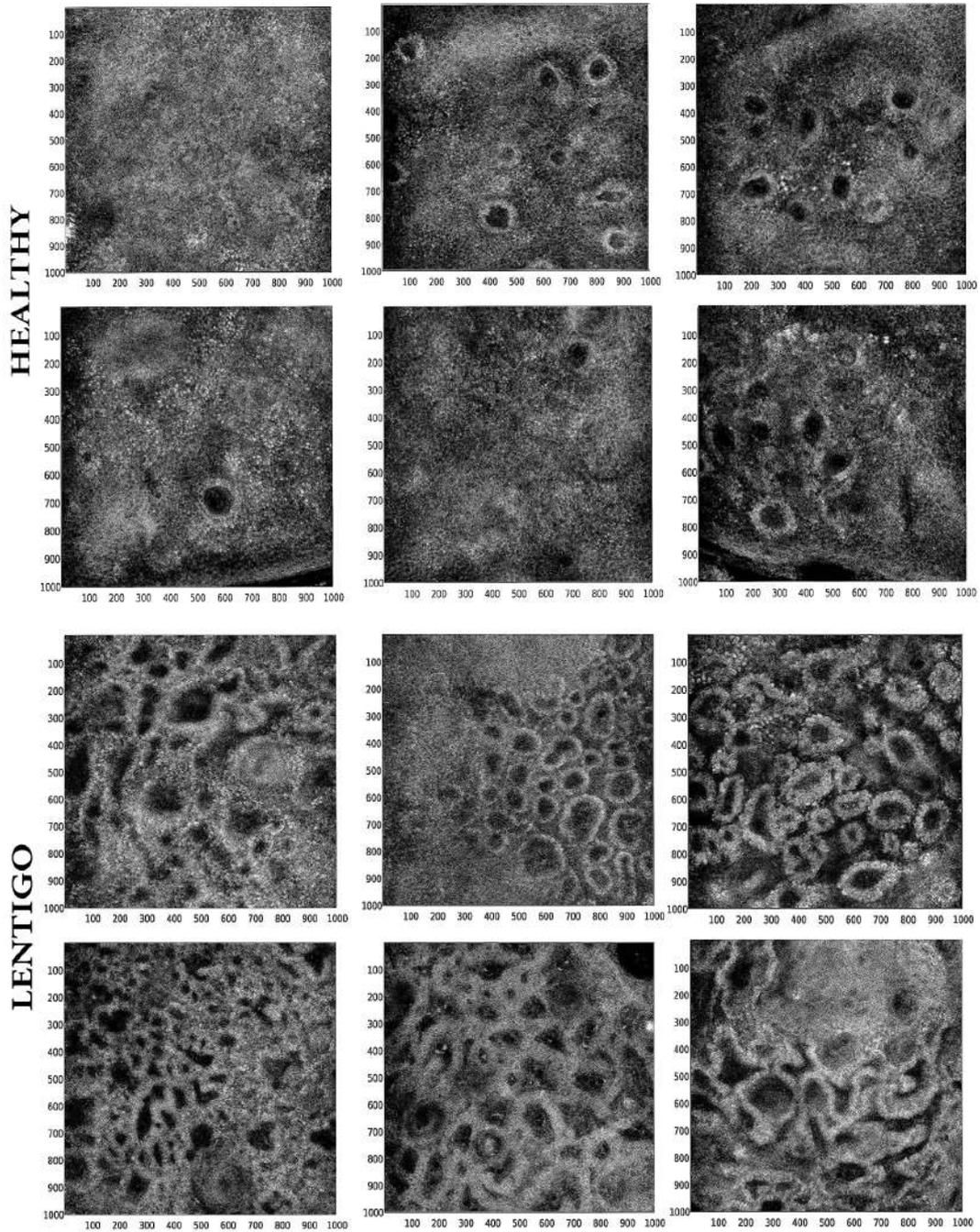


Figure 3.14: Images de patients sains (patient #1, #2, #3, #4, #5, #6) et lentigo (patient #31, #33, #37, #38, #40, #44) à la profondeur DEJ ($54 \mu\text{m}$). On peut observer des images plus texturées en présence de lentigo.

3.3.3 Importance statistique avec T-Test

Le test paramétrique décrit dans l'Annexe A a été appliqué aux paramètres γ , β , ρ , E et V à chaque profondeur pour évaluer la significativité des résultats. Les Figures 3.15, 3.16 montrent les p-values et les facteurs de Bayes associés au test de Student à deux échantillons effectué respectivement avec γ , β , ρ , E et V à différentes profondeurs.

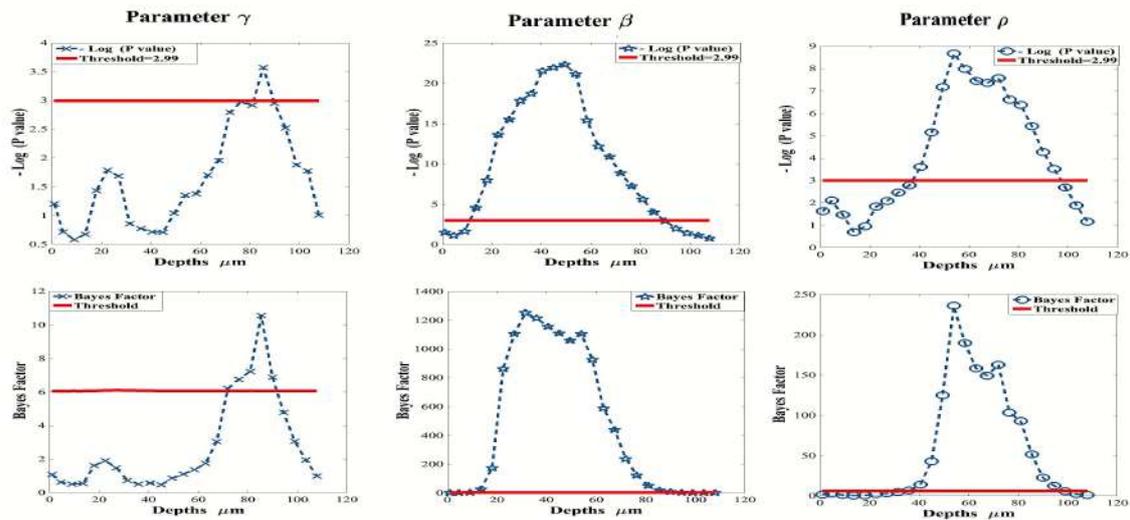


Figure 3.15: P-value (en échelle logarithmique $-\log$) et facteur de Bayes (BF) du T test pour γ , β and ρ . Le score faible montre que γ n'est pas discriminant entre les images saines et lentigo. Des scores forts peuvent être observés pour des profondeurs comprises entre $30 \mu\text{m}$ et $76 \mu\text{m}$ et les scores les plus élevés sont obtenus pour des profondeurs autour de $50 \mu\text{m}$ pour les paramètres β et ρ . Cela confirme que β et ρ sont de bonnes fonctions discriminantes qui peuvent être utilisées pour séparer les images saines et lentigo à ces profondeurs.

La p-value a été représentée en échelle logarithmique ($-\log$) pour une meilleure lisibilité. La Figure 3.15 montre des valeurs faibles pour γ à la fois pour la p-value et le facteur de Bayes confirmant que γ ne peut pas discriminer les images saines et lentigo. D'autre part, cette Figure montre des scores élevés pour les deux indicateurs et pour une large gamme de profondeurs pour les paramètres β et ρ .

Le Tableau 3.2 présente des profondeurs qui donnent des valeurs de T_β et T_ρ supérieures au seuil $T_{PFA} = 2.02$, confirmant ainsi l'hypothèse selon laquelle les paramètres β and ρ peuvent être utilisés pour discriminer les patients sains et lentigo. Ce tableau présente également les profondeurs qui fournissent des p-value inférieures à la probabilité de fausse alarme $PFA = 0.05$ et leur facteur

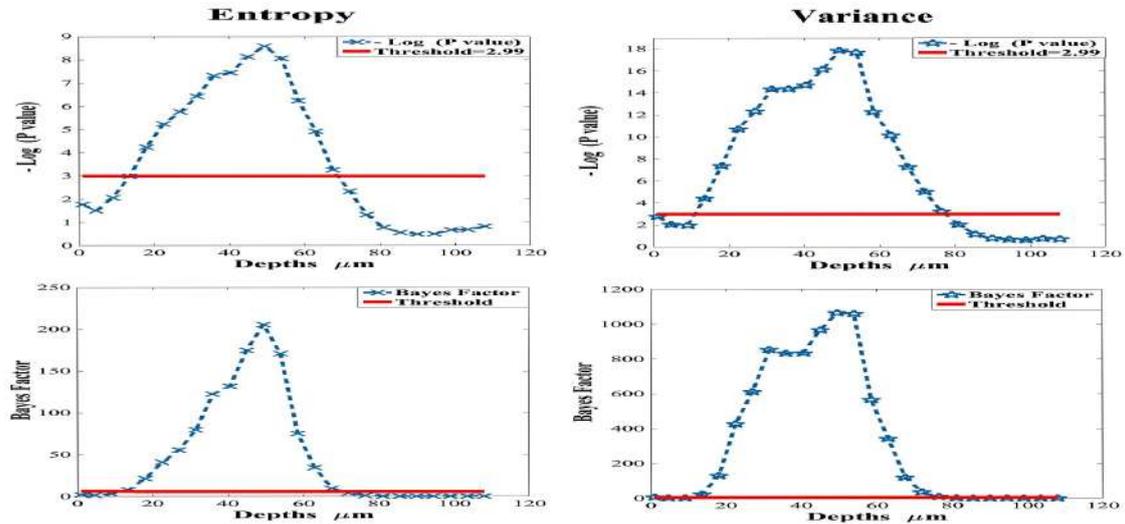


Figure 3.16: P-value (en échelle logarithmique $-\log$) et facteur de Bayes (BF) du T test pour E et V. Des scores forts peuvent être observés pour des profondeurs comprises entre $30 \mu\text{m}$ et $76 \mu\text{m}$ et les scores les plus élevés sont obtenus pour des profondeurs autour de $50 \mu\text{m}$ pour les paramètres E et V.

		β		ρ	
		min	max	min	max
$T_{\text{scores}} > 2.02$	profondeurs	13.5	86	40	94
	T_{scores}	2.6	2.46	2.33	2.27
	p-value	0.0106	0.018	0.0273	0.0302
	BF	26.5	18	14.03	12.64
T_{scores} les plus significatif	profondeurs	38.25	60.75	42.75	65.25
	T_{scores}	7.90	6.11	2.8	3.8
	p_value	$5.6 \cdot 10^{-5}$	$5.31 \cdot 10^{-4}$	0.0101	0.0067
	BF	1177	862	40	151
Maximal T_{scores}	profondeurs	49.5		54	
	T_{scores}	8.54		4.21	
	p-value	$2.78 \cdot 10^{-5}$		0.0018	
	BF	1060		236	

Tableau 3.2: Profondeurs où $T^{(\beta)}$ et $T^{(\rho)} > T_{0.05} = 2.02$; les valeurs correspondantes de p-value et du facteur de Bayes (BF) sont mentionnées. La première ligne donne les intervalles de profondeur (profondeur minimale à la profondeur maximale) où les scores T sont significatifs. La deuxième ligne montre les profondeurs donnant les scores T les plus élevés (score maximal $\mp 10\%$). La troisième ligne montre les profondeurs correspondant au score maximal T. Les P-values et les facteurs de Bayes correspondant à chaque profondeur sont présentés ci-dessous.

de Bayes correspondant. Selon nos règles de décision, les résultats sont *très significatifs* pour des profondeurs comprises entre $40 \mu m$ et $60 \mu m$, avec le score le plus élevé à $49.5 \mu m$ pour β et $54 \mu m$ pour ρ . Ces résultats sont en accord avec les différences quantitatives indiquées dans la Figure 3.11. Ces résultats confirment que β et ρ donnent une bonne statistique de test pour discriminer une peau lentigo d'une peau saine, en particulier à des profondeurs aux environs de $50 \mu m$. Comme mentionné dans l'introduction, les lentigos se caractérisent principalement en imagerie MCR par la désorganisation de la jonction dermo-épidermique (DEJ).

Cela explique pourquoi le paramètre β est très discriminant à des profondeurs proches de $50 \mu m$, ce qui correspond à la profondeur moyenne qui représente le DEJ (l'intervalle vrai de DEJ a été annoté par les dermatologues et est délimité par les deux lignes bleues) comme indiqué sur la Figure 3.17. Ce résultat est compatible avec les pratiques cliniques pour analyser le DEJ, confirmant la validité de notre approche.

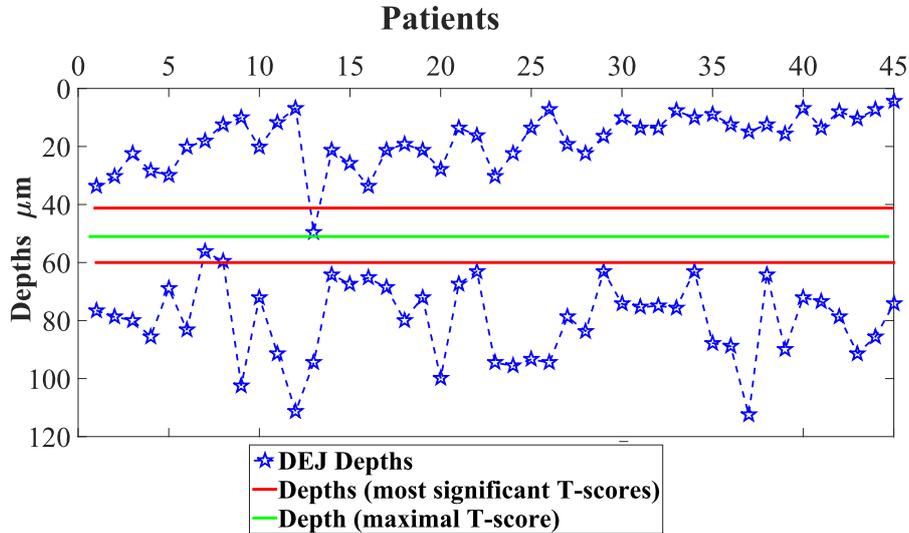


Figure 3.17: Profondeurs caractéristiques estimées entre $40 \mu m$ et $60 \mu m$ selon le T-test (lignes vertes et rouges) et les vrais profondeurs DEJ annotées par les dermatologues (lignes bleues) associées aux 45 patients.

3.3.4 Classification d'images sains et avec lentigo

La Section 3.3.3 a fourni une estimation pour les profondeurs caractérisant les patients sains de ceux atteints de lentigo. Sur la base de ces profondeurs, nous avons considéré un algorithme de classification basé sur la méthode des machines à vecteurs supports (méthode SVM pour “support vector machine”) pour confirmer leur validité. Les paramètres FGD associés à ces profondeurs caractéristiques ($40 \mu m$ à $60 \mu m$) ont ensuite été utilisés pour classer les patients dans 2 classes appelées “lentigo” et “sain”. La méthode *leave-one-out* déjà présentée dans la Section 2.3.5 a été utilisée afin de calculer les différentes probabilités d'erreurs. La matrice de confusion moyenne obtenues et indiquée dans le tableau 3.3 nous permet d'évaluer la performance de classification pour ces profondeurs caractéristiques. Les résultats montrent que la classification des tissus sains et ceux avec lésion a une accuracy égale à 84.4 % pour β et égale à 82.2 % pour ρ , nous recommandons donc d'utiliser le paramètre β pour cette classification. La Figure 3.18 montre des exemples d'images MCR classées en utilisant la méthodologie proposée. Pour évaluer l'importance de nos résultats, notre algorithme a ensuite été comparé à la méthode présentée dans [Koller et al., 2011]. Cette méthode consiste à extraire à partir de chaque image MCR un ensemble de 39 paramètres (d'autres détails techniques sont disponibles dans [Wiltgen et al., 2008]) et à leur appliquer une procédure de classification basée sur un arbre de classification et de régression (CART). Il faut noter que l'algorithme CART a été testé sur les images MCR réelles en utilisant une procédure de *leave one out*. Comme le montre le tableau 3.4, l'accuracy obtenue avec l'algorithme CART est de 80 %, c'est-à-dire légèrement inférieure à celle obtenue avec la méthode proposée et conduit à un nombre de patients mal classés plus élevé (deux patients additionnel sur les 45 considérés). En outre, la méthode proposée fournit les paramètres FGD qui peuvent être utilisés pour la caractérisation des images MCR, ce qui n'est pas possible avec l'algorithme CART.

Matrice de confusion	β			ρ		
	\hat{L}	\hat{H}	Sensibilité Spécificité	\hat{L}	\hat{H}	Sensibilité Spécificité
Lentigo	22	5	81.4 %	21	6	77.7 %
Sain	2	16	88.8 %	2	16	88.8 %
Précision	91.6 %	76.1 %		87.5 %	72.7 %	
Accuracy	84.4 %			82.2 %		

Tableau 3.3: Matrices de confusion du classifieur SVM basé sur β and ρ (pour les profondeurs caractéristiques (40 μm à 60 μm)). On remarque la bonne précision pour les deux paramètres en particulier pour β .

Tableau 3.4: Performance de la classification sur des données réelles (45 patients) utilisant la méthode CART (pour les profondeurs caractéristiques (40 μm à 60 μm)).

Matrice de confusion	\hat{L}	\hat{H}	Sensibilité Spécificité
Lentigo	21	6	77.7 %
Sain	3	15	83.3 %
Précision	87.5 %	71.4 %	
Accuracy	80 %		

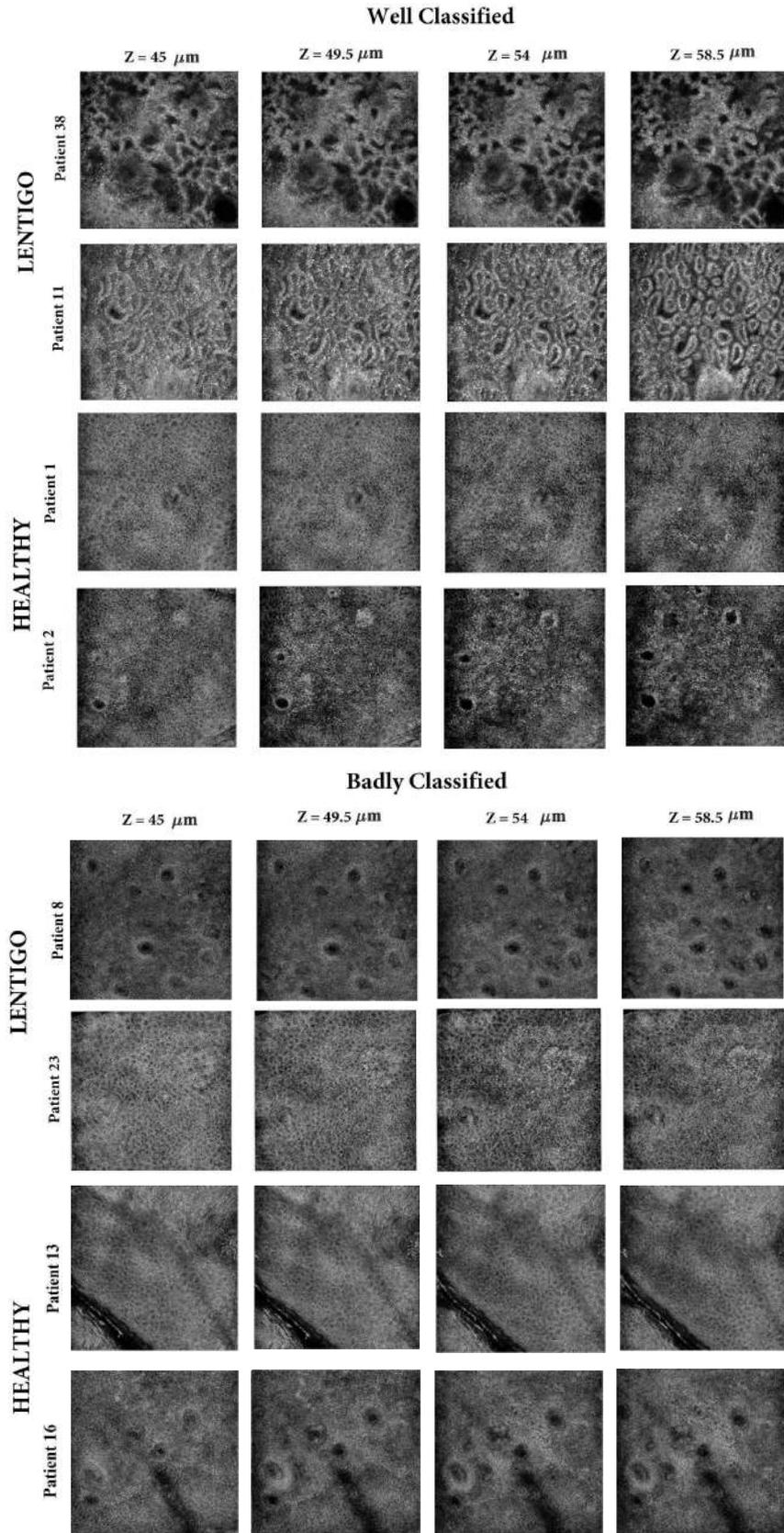


Figure 3.18: Exemples d'images MCR des patients sains et lentigo classées à l'aide du classifieur SVM.

3.4 Conclusions

Dans ce chapitre, nous avons étudié le potentiel d'utilisation des propriétés statistiques des intensités de pixels associées aux images de microscopie confocale à réflectance pour classifier les patients sains et les patients atteints de lentigo. La méthode proposée consiste à calculer les paramètres de localisation, d'échelle et de forme d'une distribution gamma généralisée pour les images MCR. Il a été montré que le lentigo se caractérise principalement à la profondeur DEJ. Les paramètres obtenus à cette profondeur ont ensuite été utilisés pour l'apprentissage d'un classifieur SVM. Le classifieur proposé a été testé sur une base de données de 2250 images réelles associées à 45 patients. Les résultats obtenus ont montré que le paramètre d'échelle β est bien adapté à la caractérisation et à la classification des tissus sains et des lésions.

CHAPITRE 4

Approche bayésienne non supervisée pour la reconstruction et la classification conjointe d'images de microscopie confocale à réflectance cutanées

Sommaire

4.1	Introduction	87
4.2	Formulation du problème	88
4.2.1	Modèle d'observation	88
4.3	Modèle bayésien hiérarchique	89
4.3.1	Vraisemblance	89
4.3.2	Distribution a priori du signal d'intérêt	90
4.3.3	Distribution a priori des variances du bruit	90
4.3.4	Distribution a priori du vecteur d'étiquette z	91
4.3.5	Lois a priori des hyperparamètres	91
4.3.6	Distribution conjointe <i>a posteriori</i>	91
4.4	Algorithme de Metropolis-within-Gibbs	92
4.4.1	Inférence bayésienne et estimation des paramètres	93
4.4.2	Convergence	94
4.5	Résultats de simulation	95
4.5.1	Données synthétiques	95
4.5.2	Données réelles	101
4.6	Conclusions	102

4.1 Introduction

Nous proposons dans ce chapitre une méthode bayésienne permettant de reconstruire les réflectivités réelles des images MCR tout en les classant dans deux catégories associées aux patients sains et atteint de lentigo.

La première contribution de ce chapitre est un modèle bayésien hiérarchique qui permet de classer un ensemble d’images MCR en classes saines et atteintes de lentigo. Chaque image est supposée être corrompue par un bruit *speckle* multiplicatif avec une distribution gamma. Une distribution gaussienne tronquée est ensuite associée à chaque image pour la classification, contraignant ces images à être positives. Des distributions a priori sont finalement affectées aux moyennes et variances de ces distributions gaussiennes tronquées, aux variances de bruit et aux étiquettes de l’image. La loi conjointe *a posteriori* du modèle proposé est finalement déterminée et sera utilisée pour la classification des images confocales et l’estimation de leurs paramètres. La deuxième contribution de ce chapitre est un algorithme d’estimation associé au modèle bayésien hiérarchique proposé. Comme les estimateurs bayésiens de la moyenne a posteriori (estimateur MMSE) et du maximum a posteriori (MAP) de ce modèle ne peuvent être facilement calculés à partir de la loi conjointe *a posteriori*, nous étudions un échantillonneur hybride de Gibbs permettant d’échantillonner la loi *a posteriori* d’intérêt (voir [Chib and Greenberg, 1995, Gilks et al., 1995] pour plus de détails). Le modèle bayésien proposé et l’algorithme d’estimation sont validés en utilisant des images MCR synthétiques et réelles résultant d’une étude clinique contenant des patients sains et atteints de lentigo. Les résultats obtenus sont très prometteurs et montrent le potentiel de la stratégie de classification et de débruitage proposée. Le chapitre est structuré comme suit. Le problème de classification étudié dans ce travail est introduit dans la Section G.1. Le modèle hiérarchique bayésien proposé et son algorithme d’estimation sont présentés dans les Sections 4.3 et 4.4. La Section 4.5.1 valide la technique proposée à l’aide de données simulées avec des niveaux de bruit différents. La Section 4.5.2 indique les résultats obtenus à l’aide de données réelles tirées d’une étude clinique. Les conclusions et les travaux futurs sont finalement présentés dans la Section 4.6.

4.2 Formulation du problème

4.2.1 Modèle d’observation

Considérons L images non bruitées, contenant N pixels, rassemblées dans une matrice $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_L] \in \mathbb{R}^{N \times L}$, où \mathbf{s}_l , $l \in \{1, \dots, L\}$ désigne l’image qui correspond au $l^{\text{ième}}$ patient. Désignons par $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_L] \in \mathbb{R}^{N \times L}$ les images bruitées correspondantes. En utilisant ces notations, le modèle

d'observation est donné par :

$$\mathbf{y}_l = \mathbf{s}_l \odot \mathbf{b}_l, \text{ avec } \mathbf{b}_l \sim \mathcal{G}(\rho_l, \theta_l) \quad (4.1)$$

où \mathbf{y}_l et \mathbf{s}_l sont des vecteurs ($N \times 1$) représentant les $l^{\text{ième}}$ images observées et non bruitées, \mathbf{b}_l est un vecteur ($N \times 1$) représentant le bruit multiplicatif de la loi gamma avec un paramètre de forme ρ_l et un paramètre d'échelle θ_l , et où \odot désigne le produit terme à terme. Afin de garantir que le modèle proposé (G.1) soit identifiable, la moyenne du bruit gamma est supposée égale à 1, ce qui conduit à :

$$\mathbb{E}(\mathbf{b}_l) = 1 \quad \Rightarrow \quad \rho_l = \frac{1}{\theta_l}. \quad (4.2)$$

Le problème traité dans ce chapitre est de classer ces images \mathbf{y}_l , $l \in \{1, \dots, L\}$ en deux classes représentant les patients sains et ceux atteints de lentigo. Ce problème est mal posé et nécessite l'introduction d'information a priori pour le régulariser. La Section suivante présente un modèle bayésien hiérarchique qui est utilisé pour régulariser le problème et effectuer cette classification.

4.3 Modèle bayésien hiérarchique

Cette Section présente un modèle bayésien hiérarchique qui peut être utilisé pour estimer les $N \times L$ images de \mathcal{S} , les $L \times 1$ vecteurs $(\mathbf{z}, \boldsymbol{\theta})$ contenant les étiquettes associées aux classes des éléments de \mathcal{S} ainsi que les variances de bruit associées aux L images observées à partir de la matrice \mathbf{Y} . Ce modèle est défini par la vraisemblance, et par les paramètres et hyperparamètres a priori définis ci-dessous.

4.3.1 Vraisemblance

Le bruit *speckle* multiplicatif \mathbf{b}_l est connu pour avoir une distribution gamma. Ainsi, le modèle d'observation (G.1) conduit à

$$y_{nl} | s_{nl}, \theta_l \sim \mathcal{G} \left(\frac{1}{\theta_l}, s_{nl} \theta_l \right) \quad (4.3)$$

où \sim signifie " est distribué selon", \mathcal{G} est la distribution gamma dont la densité de probabilité (pdf) est définie par

$$f(y_{nl} | s_{nl}, \theta_l) \propto \frac{(y_{nl})^{\frac{1}{\theta_l} - 1} \exp \left(-\frac{y_{nl}}{s_{nl} \theta_l} \right)}{\Gamma \left(\frac{1}{\theta_l} \right) (s_{nl} \theta_l)^{\frac{1}{\theta_l}}} I_{\mathbb{R}^+}(y_{nl}) \quad (4.4)$$

avec $I_{\mathbb{R}^+}(y_{nl})$ la fonction indicatrice sur \mathbb{R}^+ , \propto signifie “proportionnel à” et Γ représente la fonction gamma. En supposant l’indépendance entre les signaux observés, la vraisemblance des L images observées peut s’écrire

$$f(\mathbf{Y}|\mathbf{S}, \boldsymbol{\theta}) \propto \prod_{n=1}^N \prod_{l=1}^L f(y_{nl}|s_{nl}, \theta_l).$$

où L et N représentent le nombre de patient et de pixels, respectivement.

4.3.2 Distribution a priori du signal d’intérêt

Afin d’assurer la positivité des images débruitées, une distribution gaussienne tronquée est affectée à \mathbf{s}_l pour $l \in \{1, \dots, L\}$

$$\mathbf{s}_l | z_l = k, \mu_k, \sigma_k^2 \sim \mathcal{N}_{\mathbb{R}^+}(\mu_k, \sigma_k^2) \quad (4.5)$$

où $\mathcal{N}_{\mathcal{S}}$ représente la distribution normale tronquée sur \mathcal{S} , k prenant les deux valeurs 1 et 2 selon la classe du patient, et (μ_k, σ_k^2) sont les moyennes et variances des deux distributions gaussiennes tronquées.

On n’oubliera pas de mentionner que d’autres distributions a-priori peuvent être supposées sur \mathbf{S} . Par exemple, une loi a priori conjuguée (inverse gamma) a été étudiée pour le paramètre \mathbf{s}_l (plus de détails sur ce modèle et ses résultats peuvent être trouvés dans l’annexe F). Néanmoins, les résultats obtenus dans ce cas ne sont pas aussi performant que le modèle décrit par 4.5.

4.3.3 Distribution a priori des variances du bruit

Une loi conjuguée inverse gamma non informative (dénommée \mathcal{IG}) est généralement choisie comme densité a priori pour le paramètre d’échelle θ_l [Fink, 1997]

$$\theta_l | a, b \sim \mathcal{IG}(a, b) \quad (4.6)$$

où a and b sont les hyperparamètres fixés, qui sont ajustés pour tenir compte de l’absence d’information sur θ_l , i.e., les moyenne et variance de θ_l sont fixées à 1 et 100 afin d’avoir une loi non informative. En supposant l’indépendance entre les éléments de $\boldsymbol{\theta}$, on obtient la loi $f(\boldsymbol{\theta} | a, b)$ comme le produit des densités conditionnelles $f(\theta_i | a, b)$.

4.3.4 Distribution a priori du vecteur d'étiquette \mathbf{z}

Le vecteur de paramètre $\mathbf{z} = (z_1, \dots, z_L)^T$ est un vecteur d'étiquettes qui associe chaque image à une peau saine ou lentigo. A cause de l'absence d'information a priori concernant ce paramètre, une loi a priori uniforme lui a été affectée comme suit :

$$P(z_l = k) = \frac{1}{2}, \forall l = 1, \dots, L. \quad (4.7)$$

Les étiquettes associées à différents patients étant supposées indépendantes a priori, i.e., la loi conjointe a priori de \mathbf{z} notée par $f(\mathbf{z})$ est le produit des probabilités définies en (4.7).

4.3.5 Lois a priori des hyperparamètres

Afin de compléter la description du modèle hiérarchique bayésien proposé et d'estimer directement les hyperparamètres à partir des données, on propose de définir des lois a priori pour les différents hyperparamètres. On a ainsi sélectionné une loi a priori gaussienne pour la moyenne μ_k et une loi a priori conjuguée inverse gamma pour la variance σ_k^2 (voir [Dobigeon et al., 2008, Fink, 1997] pour motivations)

$$\mu_k \mid \mu_0, \sigma_0 \sim \mathcal{N}(\mu_0, \sigma_0^2) \quad (4.8)$$

$$\sigma_k^2 \mid \alpha_0, \beta_0 \sim \mathcal{IG}(\alpha_0, \beta_0) \quad (4.9)$$

où $\mu_0, \sigma_0^2, \alpha_0, \beta_0$ sont fixés afin d'obtenir des lois non informatives, i.e., $\mu_0 = 100, \sigma_0^2 = 10^5$ alors que les moyennes et variances de σ_k^2 sont fixées à 1 et 1000. Les densités de probabilité conjointes $f(\boldsymbol{\mu} \mid \mu_0, \sigma_0)$ et $f(\boldsymbol{\sigma}^2 \mid \alpha_0, \beta_0)$ sont finalement obtenues comme le produit de leurs densités marginales en supposant l'indépendance des a priori entre les composantes de ces deux vecteurs.

4.3.6 Distribution conjointe *a posteriori*

Le modèle bayésien proposé est illustré par le graphe orienté acyclique (DAG) représenté en Figure. 4.1. Ce dernier met en évidence la relation entre les observations \mathbf{Y} , les paramètres $\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}$, et les hyperparamètres μ_k, σ_k^2 . En supposant l'indépendance des a priori entre les différentes composantes

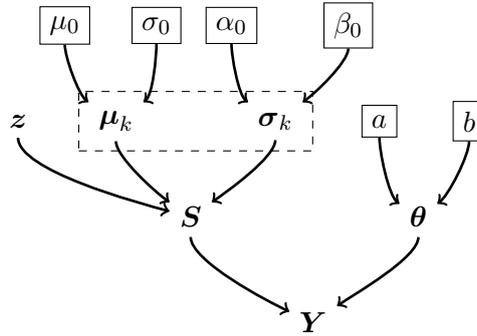


Figure 4.1: DAG pour les lois a priori des paramètres et hyperparamètres. Les hyperparamètres fixés par l'utilisateur apparaissent dans des boîtes (ligne continue).

du vecteur de paramètre $\mathbf{X} = (\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2)$, la distribution conjointe *a posteriori* de ce modèle bayésien peut être calculée en utilisant la structure hiérarchique suivante

$$f(\mathbf{X} | \mathbf{Y}) \propto f(\mathbf{Y} | \mathbf{S}, \boldsymbol{\theta}) f(\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) \quad (4.10)$$

avec

$$f(\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) = f(\mathbf{S} | \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) f(\boldsymbol{\theta} | a, b) f(\boldsymbol{\mu} | \mu_0, \sigma_0) f(\boldsymbol{\sigma}^2 | \alpha_0, \beta_0) f(\mathbf{z}). \quad (4.11)$$

La complexité du modèle bayésien proposé est illustré dans le DAG de la Figure. 4.1 et sa distribution *a posteriori* résultante (4.10) rendent difficile le calcul explicite des estimateurs bayésiens (MAP ou MMSE) des paramètres du modèle. La Section suivante étudie une méthode MCMC qui sert à échantillonner la loi *a posteriori* (4.10) et à construire des estimateurs des paramètres intervenant dans le modèle bayésien proposé en utilisant les échantillons générés.

4.4 Algorithme de Metropolis-within-Gibbs

Cette Section étudie un échantillonneur de Gibbs hybride qui génère des échantillons asymptotiquement distribués suivant la loi cible (4.10). L'échantillonneur de Gibbs décrit dans Algo. 1, génère itérativement des échantillons distribués selon les lois conditionnelles de la loi conjointe définie dans (4.10). Ces lois conditionnelles sont détaillées dans le reste de cette Section. A cause de la complexité de ces lois conditionnelles, nous considérons un algorithme de Metropolis-Hastings à marche aléatoire (RWMH) [Chib and Greenberg, 1995, Gilks et al., 1995] dans l'échantillonneur de Gibbs et

qui va consister à générer des échantillons distribués selon la loi conditionnelle complexe de chaque paramètre d'intérêt. Ceci est réalisé en utilisant les lois conditionnelles $f_j(\cdot)$, pour $j \in \{1, \dots, J\}$, et leurs lois de proposition associées $g_j(\cdot)$.

La boucle principale de l'algorithme RWMH se compose de trois parties, après l'initialisation des paramètres $\mathbf{X}_j^{(0)}$, pour $j \in \{1, \dots, J\}$, :

1. Générer un candidat $\mathbf{X}_j^{\text{cand}}$ à partir de la loi de proposition $g_j(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})$. La distribution $g_j(\cdot)$ est la distribution gaussienne tronquée $\mathcal{N}_{\mathbb{R}^+}(\mathbf{X}_j^{(i-1)}, \epsilon_j^2)$ pour les paramètres $\mathbf{S}, \boldsymbol{\theta}, \boldsymbol{\sigma}^2$, (voir [Robert, 1995]), et la distribution gaussienne $\mathcal{N}(\mathbf{X}_\mu^{(i-1)}, \epsilon_\mu^2)$ pour $\boldsymbol{\mu}$.
2. Calculer la probabilité d'acceptation en utilisant la fonction d'acceptation $\alpha(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})$ basée sur la loi de proposition et la densité marginale pour chaque paramètre

$$\alpha(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)}) = \min \left\{ \frac{f_j(\mathbf{X}_j^{\text{cand}})}{f_j(\mathbf{X}_j^{(i-1)})} \frac{g_j(\mathbf{X}_j^{(i-1)} | \mathbf{X}_j^{\text{cand}})}{g_j(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})}, 1 \right\}$$

3. Accepter le candidat avec la probabilité $\alpha(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})$.

Afin de maximiser l'efficacité de l'algorithme, les variances ϵ_j^2 des lois de proposition ont été ajustées de telle sorte que le taux d'acceptation soit compris entre 0,3 et 0,6 comme suggéré dans [Chib and Greenberg, 1995] et [Roberts et al., 2001].

Les densités conditionnelles des paramètres d'intérêt sont obtenues en multipliant la vraisemblance avec les différentes lois a priori et en enlevant le terme multiplicatif qui ne dépend pas de la variable d'intérêt. L'algorithme met à jour itérativement chaque paramètre en utilisant sa loi conditionnelle qui est détaillée dans l'Annexe E.

4.4.1 Inférence bayésienne et estimation des paramètres

Les principales étapes de l'échantillonneur de Metropolis-dans-Gibbs proposé sont résumées dans l'algorithme Algo. 1. Cet algorithme fournit un ensemble de vecteurs $\mathbf{X} = (\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k^2)$ notée

Algorithm 1 algorithme de Metropolis-within-Gibbs

-
- 1: Entrées: $N_{\text{bi}}, N_{\text{MC}}, \mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2$
 - 2: Initialisation
 - 3: Initialiser $\mathbf{S}^{(0)}, \boldsymbol{\theta}^{(0)}, \mathbf{z}^{(0)}, \boldsymbol{\mu}^{(0)}, \boldsymbol{\sigma}^{2(0)}$
 - 4: **for** $i=1$ to N_{MC} **do**
 - 5: mise à jour des paramètres
 - 6: Échantillonner $\mathbf{S}^{(i)} \mid \mathbf{Y}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2$ selon (E.1) en utilisant l'algorithme RWMH avec la loi de proposition gaussienne tronquée
 - 7: Échantillonner $\boldsymbol{\theta}^{(i)} \mid \mathbf{Y}, \mathbf{S}, a, b$ selon (E.2) en utilisant l'algorithme RWMH avec la loi de proposition gaussienne tronquée
 - 8: Échantillonner $\boldsymbol{\mu}^{(i)} \mid \mathbf{S}, \boldsymbol{\sigma}^2, \mu_0, \sigma_0^2$ selon (E.3) en utilisant l'algorithme RWMH avec la loi de proposition gaussienne
 - 9: Échantillonner $\boldsymbol{\sigma}^{2(i)} \mid \mathbf{S}, \boldsymbol{\mu}, \alpha_0, \beta_0$ selon (E.4) en utilisant l'algorithme RWMH avec la loi de proposition gaussienne tronquée
 - 10: Échantillonner $\mathbf{z}^{(i)} \mid \mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2$ à partir de la pdf (E.5)
 - 11: **end for**
 - 12: Sorties : $\mathbf{S}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathbf{z}^{(i)}, \boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{2(i)}$ for $i = 1, \dots, N_{\text{MC}}$.
-

$\mathbf{X}_j^{(i)}$ qui seront utilisés pour calculer les estimateurs MMSE des divers paramètres $\mathbf{S}, \boldsymbol{\theta}, \mu_k, \sigma_k^2$ comme suit

$$\widehat{\mathbf{X}}_j^{\text{MMSE}} \simeq \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{i=N_{\text{bi}}+1}^{N_{\text{MC}}} \widehat{\mathbf{X}}_j^{(i)} \quad (4.12)$$

où N_{bi} et N_{MC} représentent respectivement le nombre d'itérations de la période de chauffe et le nombre total d'itérations de Monte Carlo. Finalement, l'estimateur maximum a-posteriori (MAP) suivant est considéré pour l'étiquette \mathbf{z}

$$\hat{\mathbf{z}}_l^{\text{MAP}} \simeq \begin{cases} 1 & \text{if } \left[\hat{\mathbf{z}}_l^{(i)} = 1 \right]_{i=N_{\text{bi}}+1}^{N_{\text{MC}}} \geq \left[\hat{\mathbf{z}}_l^{(i)} = 2 \right]_{i=N_{\text{bi}}+1}^{N_{\text{MC}}} \\ 2 & \text{otherwise} \end{cases} \quad (4.13)$$

où $[x = 1]_i^j$ and $[x = 2]_i^j$ représente le nombre d'échantillons satisfaisant les conditions $x = 1$ et $x = 2$ dans l'intervalle $[i, j]$.

4.4.2 Convergence

L'exécution de plusieurs chaînes en parallèle avec différentes initialisations permet d'avoir diverses mesures de convergence pour les méthodes MCMC [Robert and Cellier, 1998]. Le critère populaire

permettant de déterminer des variances inter et intra chaines a montré des propriétés intéressantes pour l'analyse de la convergence des méthodes MCMC. Ce critère a été initialement étudié par Gelman et Rubin dans [Gelman and Rubin, 1992] et a ensuite été utilisé dans de nombreuses études [Robert and Cellier, 1998, p. 33], [Godsill and Rayner, 1998] et [Djurić and Chun, 2002]. L'idée principale est de construire M chaines de longueur $N_r + N_{bi}$ en parallèle, initialisées différemment, pour chaque ensemble de données et d'évaluer la dispersion des estimations obtenues pour les différentes chaines. La variance inter-chaîne B et la variance intra-chaîne W pour les M chaines de Markov sont définies par :

$$B = \frac{N_r}{M-1} \sum_{m=1}^M (\bar{k}_m - \bar{k})^2 \quad (4.14)$$

$$W = \frac{1}{M} \sum_{m=1}^M \frac{1}{N_r} \sum_{t=1}^{N_r} (k_m^{(t)} - \bar{k}_m)^2 \quad (4.15)$$

avec

$$\bar{k}_m = \frac{1}{N_r} \sum_{t=1}^{N_r} k_m^{(t)}, \quad \bar{k} = \frac{1}{M} \sum_{m=1}^M \bar{k}_m, \quad N_r = N_{MC} - N_{bi}. \quad (4.16)$$

où k est le paramètre d'intérêt et $k_m^{(t)}$ son estimé à la $t^{ième}$ exécution pour la $m^{ième}$ chaîne. La convergence de l'algorithme a été évaluée grâce au critère appelé *potential scale reduction factor* $\hat{\rho}$ défini par [Gelman et al., 1995, p. 332]

$$\sqrt{\hat{\rho}} = \sqrt{\frac{1}{W} \left(\frac{N_r - 1}{N_r} W + \frac{1}{N_r} B \right)}. \quad (4.17)$$

4.5 Résultats de simulation

4.5.1 Données synthétiques

L'algorithme proposé a été testé sur des données synthétiques afin de valider son efficacité. Différentes expériences ont été réalisées et ce pour trois valeurs du rapport signal/bruit $SNR \in \{0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}\}$, ce qui permet d'apprécier la performance de l'algorithme pour différents niveaux de bruit. Cette Section considère $L = 100$ images synthétiques. Chaque image contient $N = 2000$ pixels et a été générée d'après (4.3). Ces images ont été séparées en classes saines et lentigo contenant 50 images. Les

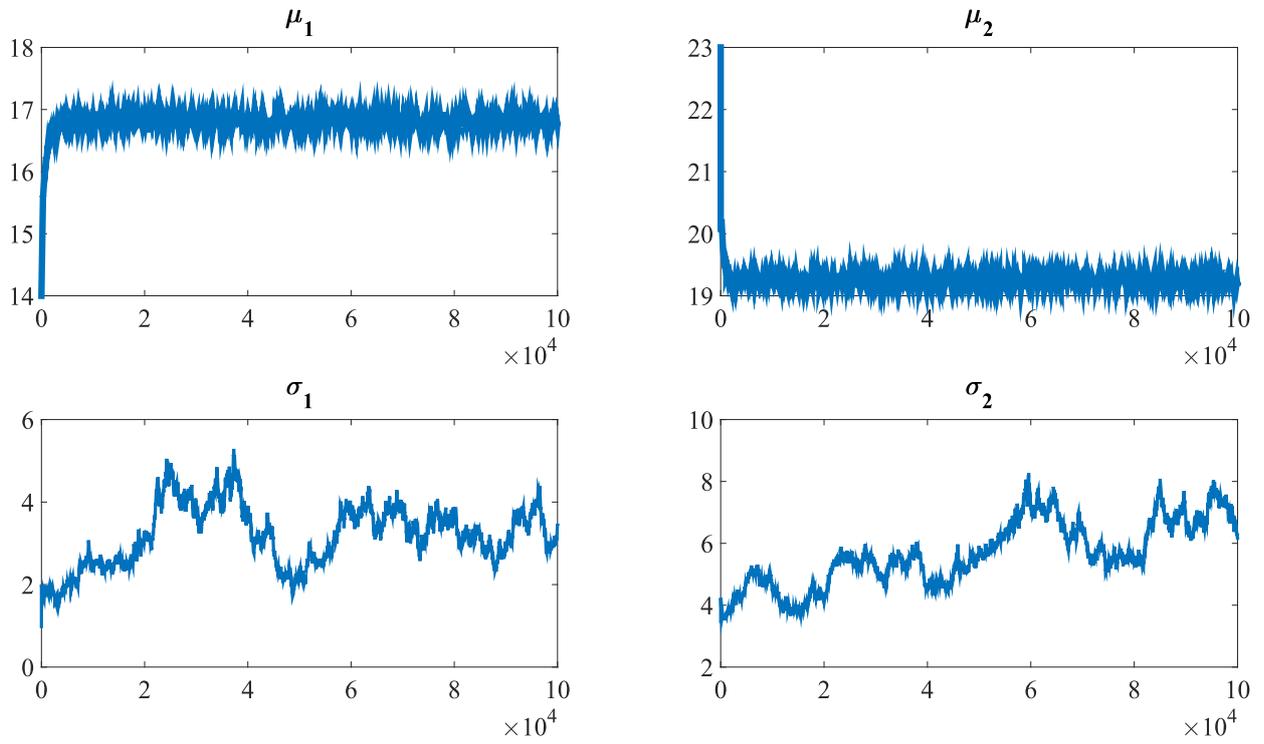
images non bruitées des deux classes ont été respectivement générées à l'aide de lois gaussiennes tronquées $\mathcal{N}_{\mathbb{R}^+}(\mu_1, \sigma_1^2)$ et $\mathcal{N}_{\mathbb{R}^+}(\mu_2, \sigma_2^2)$, avec $\mu_1 = 17, \mu_2 = 20, \sigma_1^2 = 2, \sigma_2^2 = 4$. La convergence de l'échantillonneur de l'algorithme est surveillée en calculant le *potential scale reduction factor* introduit en (4.17) pour un paramètre d'intérêt. Plusieurs choix du paramètre k peuvent être considérés pour la méthode proposée. La convergence de l'échantillonneur Metropolis-within-Gibbs à été surveillée ici en calculant le *psrf* à l'aide du paramètre $\sqrt{\hat{\rho}}$ de la variance du bruit θ (voir [Dobigeon et al., 2008, Godsill and Rayner, 1998] pour un choix similaire). Le *potential scale reduction factor* du paramètre θ est calculé pour $M = 10$ chaînes de Markov est égal à 1.01. Cette valeur de $\sqrt{\hat{\rho}}$ confirme la bonne convergence de l'échantillonneur (une recommandation pour l'évaluation de la convergence est une valeur de $\sqrt{\hat{\rho}} \leq 1.2$ [Gelman et al., 1995, p. 332]). Les Figures 4.2, 4.3, 4.4 and 4.5 montrent l'évolution des chaînes de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés respectivement pour des données synthétiques avec $\text{SNR}_Y = [0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}]$ et des données réelles à partir des images MCR.

Algo. 1 a été exécuté pour $N_{\text{MC}} = 100000$ itérations et les différents paramètres du modèle ont été estimés d'après (4.12) et (4.13) en utilisant une période de chauffe de longueur $N_{\text{bi}} = 99900$. La performance de l'algorithme a été évaluée en calculant les erreurs quadratiques moyennes (MSEs) des différents paramètres et les rapports signal à bruit (SNRs) définis par:

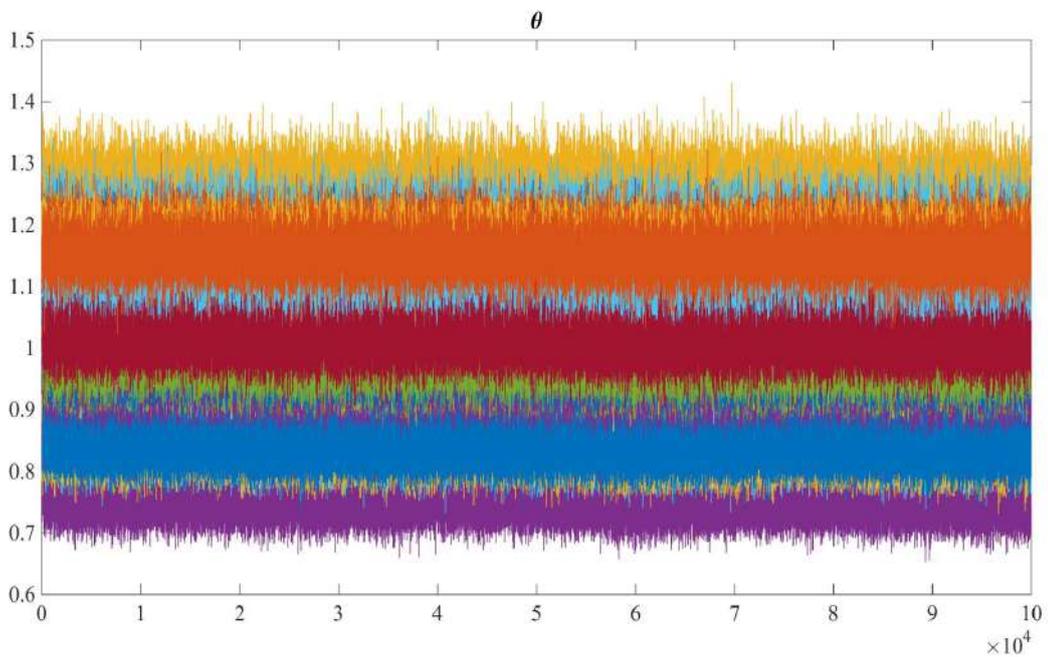
$$\text{MSE}_j = \|\hat{\mathbf{X}}_j - \mathbf{X}_j\|^2 \quad (4.18)$$

$$\text{SNR}_j = 20 \log_{10} \left(\frac{\|\mathbf{X}_j\|}{\|\hat{\mathbf{X}}_j - \mathbf{X}_j\|} \right). \quad (4.19)$$

Des résultats quantitatifs sont présentés dans le tableau 4.1 pour les trois expériences. Ce tableau montre les bon résultats de classification et d'estimation des paramètres pour différents niveaux de bruit. Ces résultats mettent en évidence l'efficacité de la stratégie proposée pour le débruitage et la classification des images obtenues à partir du modèle (4.3).

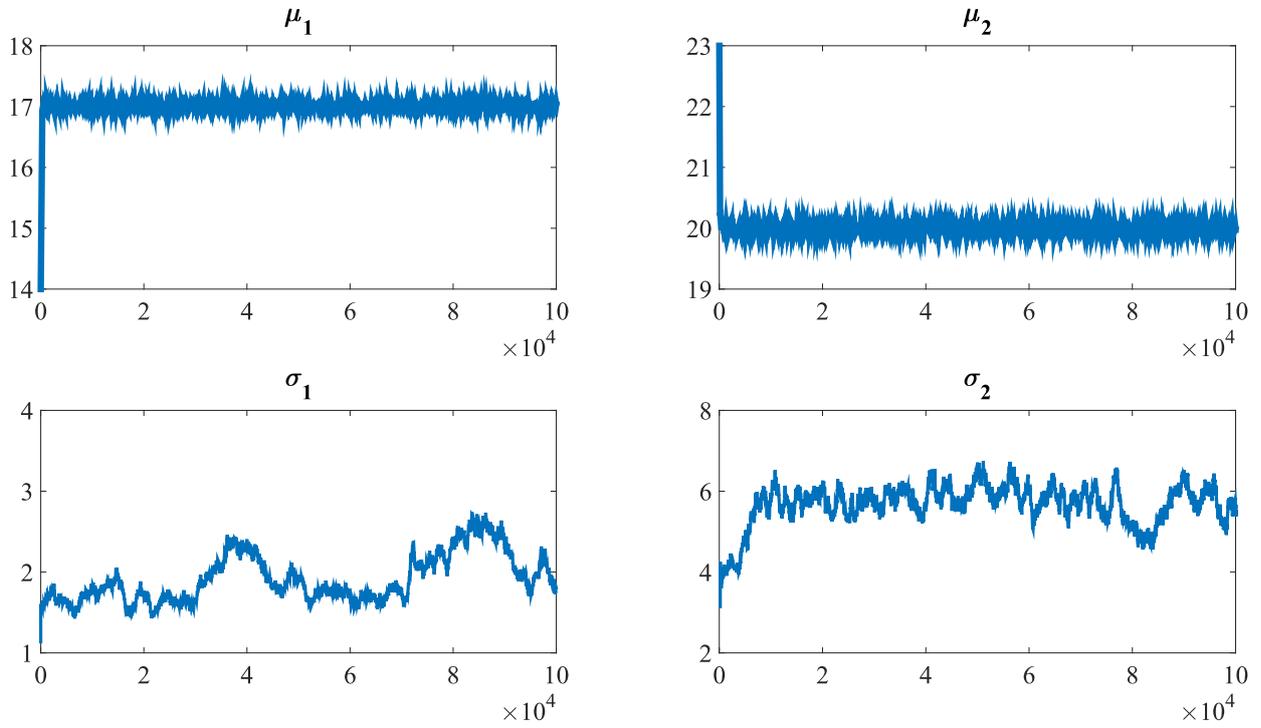


(a)

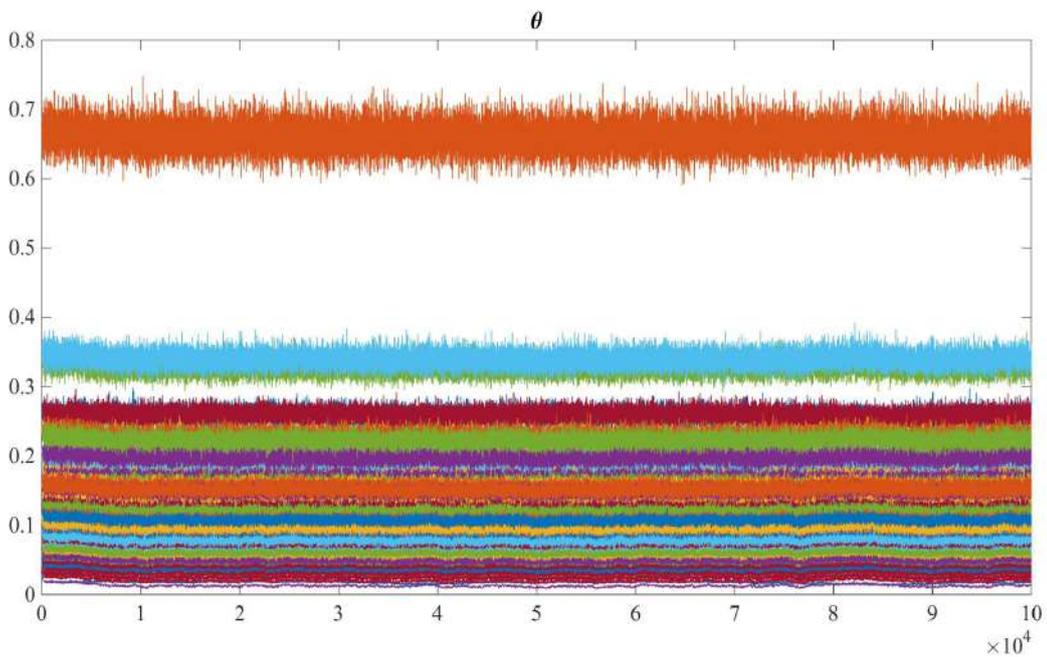


(b)

Figure 4.2: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_Y = 0$ dB.

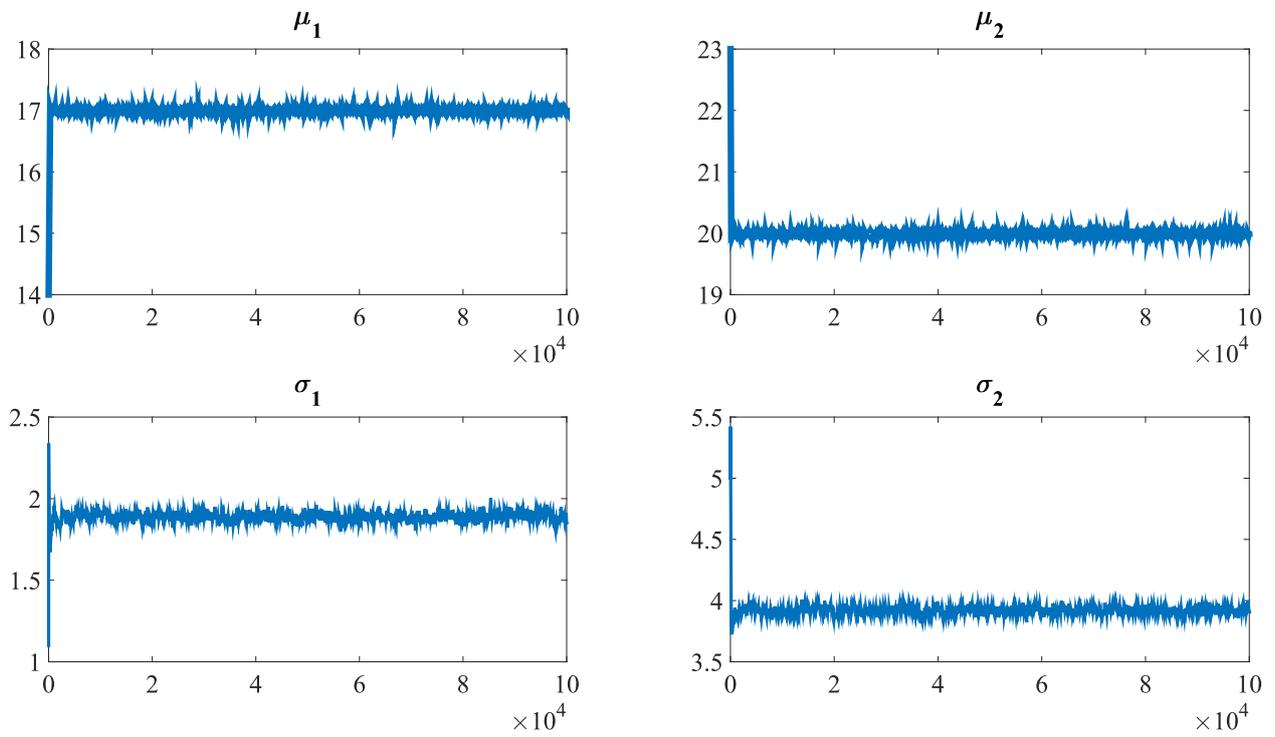


(a)

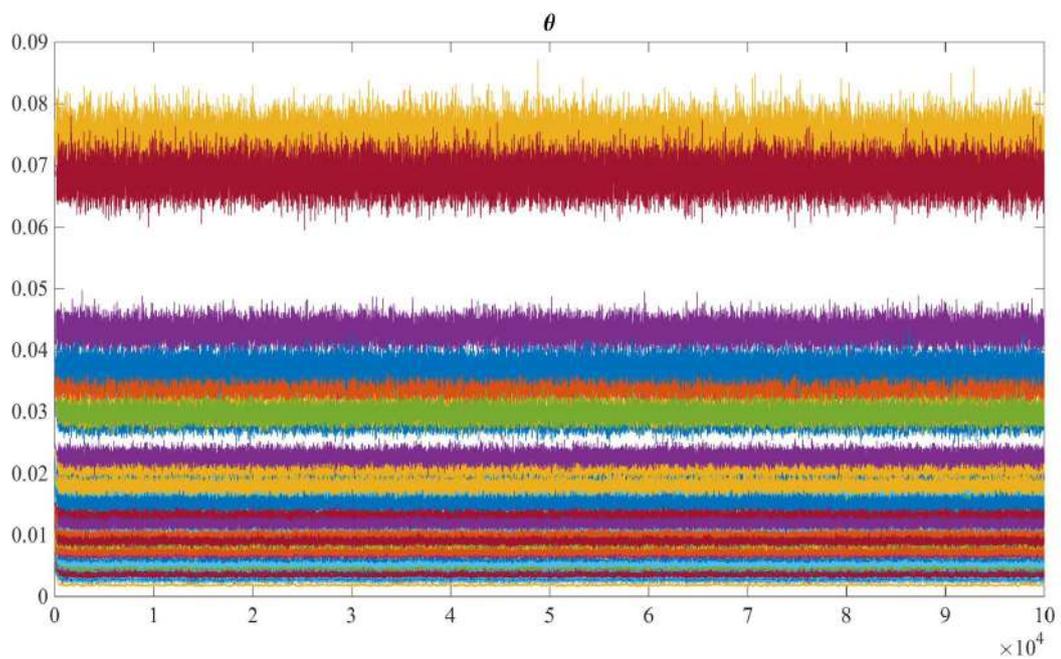


(b)

Figure 4.3: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_Y = 10$ dB.

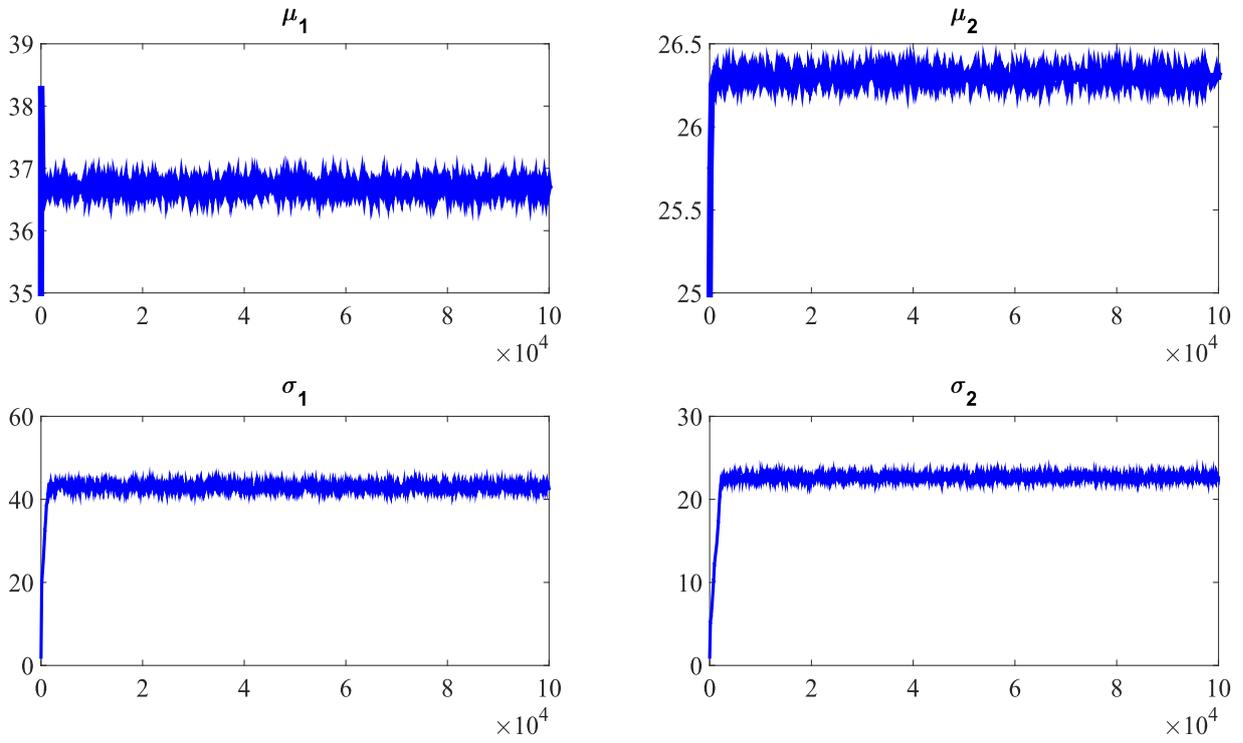


(a)

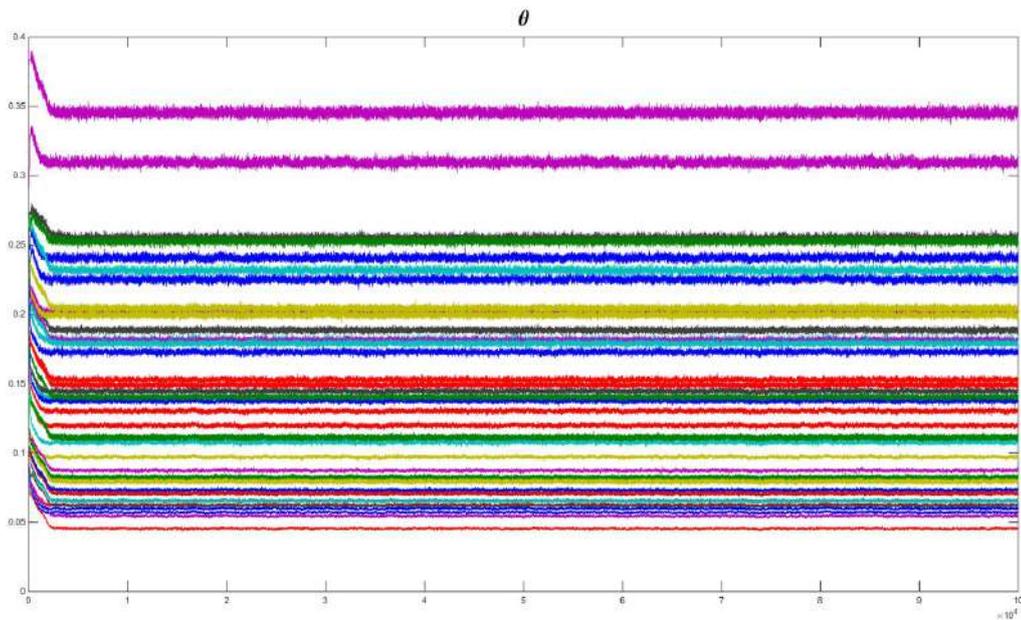


(b)

Figure 4.4: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_Y = 20$ dB.



(a)



(b)

Figure 4.5: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\theta}$ estimés pour des images MCR réelles.

Tableau 4.1: Performance de l'algorithme proposé pour le débruitage et la classification des données synthétiques bruitées pour trois valeurs du signal/bruit $\text{SNR}_Y = [0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}]$.

	$\text{SNR}_Y = 0 \text{ dB}$		$\text{SNR}_Y = 10 \text{ dB}$		$\text{SNR}_Y = 20 \text{ dB}$	
	MSE^2	SNR (dB)	MSE^2	SNR (dB)	MSE^2	SNR (dB)
μ_1	0.56	30.12	$1.54 \cdot 10^{-4}$	62.72	$2.63 \cdot 10^{-5}$	70.4
μ_2	0.95	21.42	$1.89 \cdot 10^{-5}$	73.24	$6.64 \cdot 10^{-5}$	67.79
σ_1^2	2.91	1.01	0.015	18.07	0.011	25.7
σ_2^2	7.14	2.57	4.58	5.42	0.006	22.07
θ	$1.14 \cdot 10^{-3}$	20.44	$4.74 \cdot 10^{-5}$	26.56	$5.68 \cdot 10^{-7}$	30.44
\mathcal{S}	5.48	16.53	2.88	20.81	0.7093	26.87
Accuracy	91%		100%		100%	
Accuracy (CART)	83%		100%		100%	

4.5.2 Données réelles

Cette Section est consacrée à la validation de l'algorithme de débruitage et de classification lorsqu'il est appliqué à des images MCR réelles déjà définies dans la Section 2.3.1, mais cette fois-ci en utilisant seulement comme profondeur caractéristique celle qui permet de distinguer avec une différence maximale les images des patients sains de ceux atteints de lentigo. D'après les résultats des chapitres 2 et 3, cette dernière a été fixée à $54 \mu\text{m}$. Compte tenu de la grande taille de ces images 1000×1000 pixels, nous avons préféré sélectionner et appliquer notre algorithme à des parcelles de 250×250 pixels pour chaque image afin de réduire le coût de calcul. Les résultats obtenus ont ensuite été utilisés pour calculer la matrice de confusion et les quatre indicateurs (sensibilité, spécificité, précision, accuracy) présentés dans les tableaux 4.2 et 4.3. Ces tableaux nous permettent d'évaluer les performances de classification de la stratégie proposée. Le paramètre "accuracy" de la méthode proposée est égale à 97.7%, ce qui correspond à une seule erreur pour le patient #8 atteint de lentigo. Figure. 4.6 montre que la texture de cette image mal classée n'est pas très déstructurée comme pour les autres patients atteints de lentigo et est visuellement similaire à la texture des patients en bonne santé. Figure. 4.7 montre des exemples d'images bruitées MCR et leurs réflectivités réelles estimées, illustrant la partie débruitage de l'algorithme proposé. On peut observer que les images estimées ont des intensités peu élevées par rapport aux images bruitées, ce qui est dû à la présence du bruit multiplicatif.

Tableau 4.2: Performance de la classification sur des données réelles (45 patients) en utilisant la méthode proposée (pour la profondeur caractéristique $54 \mu m$).

Matrice de confusion	\hat{L}	\hat{H}	Sensibilité Specificité
Lentigo	26	1	96.2 %
Sain	0	18	100 %
Précision	100 %	94.7 %	
Accuracy	97.7 %		

Tableau 4.3: Performance de la classification sur des données réelles (45 patients) en utilisant la méthode CART (pour la profondeur caractéristique $54 \mu m$).

Matrice de Confusion	\hat{L}	\hat{H}	Sensibilité Specificité
Lentigo	24	3	88.8 %
Sain	5	13	72.2 %
Précision	82.7 %	81.2 %	
Accuracy	82.2 %		

Afin d'évaluer l'importance de nos résultats, notre algorithme a été comparé à la méthode CART déjà présentée dans la Section 3.3.4. Comme montré dans le Tableau 4.3, le paramètre "accuracy" obtenu avec l'algorithme CART est égal à 82.2% , i.e., il est légèrement inférieur à celui obtenu avec la méthode proposée. De plus, le modèle bayésien proposé peut être utilisé pour la caractérisation des images MCR grâce à ses paramètres estimés.

4.6 Conclusions

Dans ce chapitre nous avons présenté une nouvelle stratégie bayésienne ainsi qu'un algorithme MCMC pour la classification d'images RCM pour des patients sains ou atteints de lentigo. Un modèle bayésien, basé sur l'hypothèse que chaque image est perturbée par un bruit multiplicatif de loi gamma, et utilisant diverses lois a priori a été introduit. Un échantillonneur de Gibbs hybride a ensuite été utilisé pour échantillonner la loi *a posteriori* de ce modèle bayésien et pour construire des estimateurs

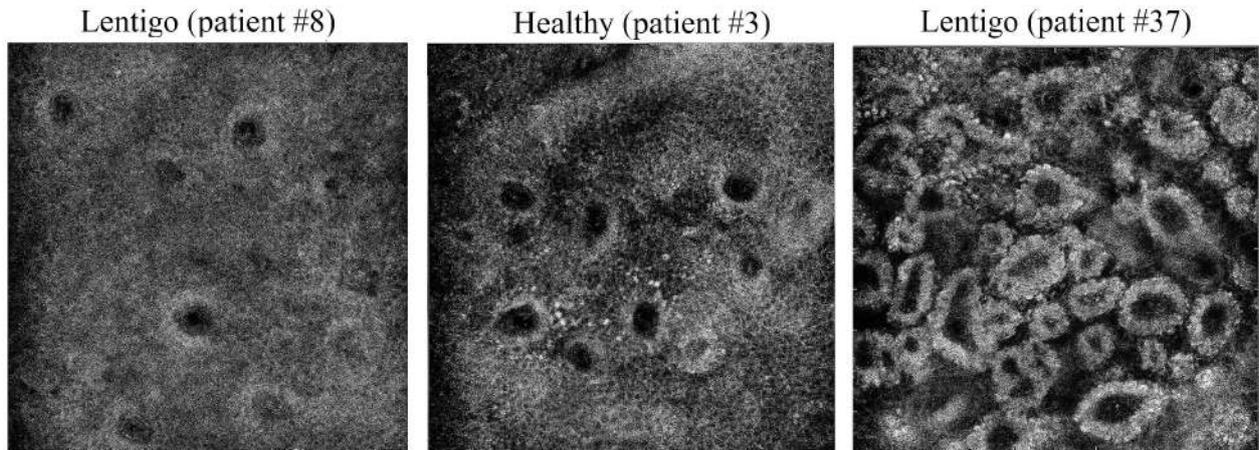


Figure 4.6: Images (à la profondeur $54 \mu m$) du patient #8 qui est mal classé en comparaison au patient sain et patient atteint de lentigo (bien classés). On peut observer plus de similitude entre ce patient et le patient sain qu'avec celui atteint lentigo.

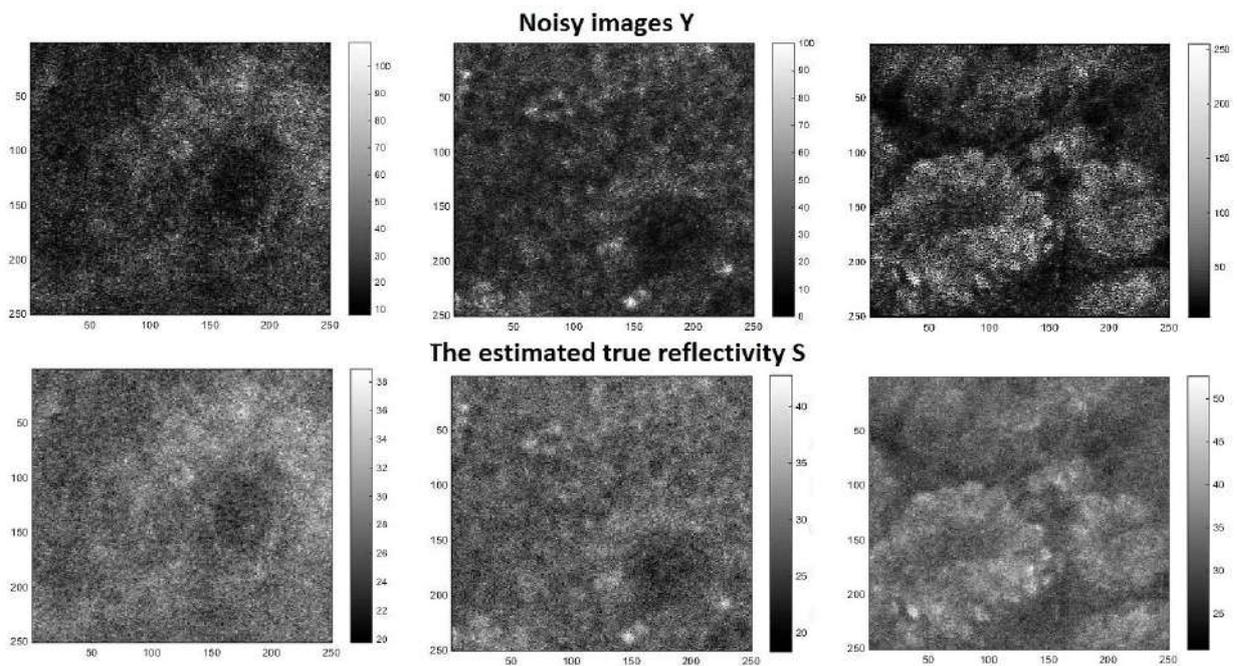


Figure 4.7: Exemples d'images bruitées (à la profondeur $54 \mu m$) et leurs vrais reflectivités estimées.

de ses paramètres. Les résultats de la simulation obtenues à l'aide de données synthétiques et réelles ont permis d'apprécier la bonne performance du classifieur proposé.

CHAPITRE 5

Conclusions et Perspectives

Cette thèse s'est intéressée à l'analyse et la modélisation statistique des images de microscopie confocale à réflectance (MCR) en dermatologie. Nous avons commencé par décrire dans le premier chapitre quelques propriétés de la peau humaine. Nous avons ensuite exposé sa composition, ses différentes fonctions principales, et les types de pigmentation: mélanique ou hyperpigmentation. Les hyperpigmentations cutanées sont nombreuses et d'origine variée, entraînant une gêne esthétique plus ou moins importante et étant parfois mal supportée par les patients. Nous avons ensuite présenté plusieurs techniques d'imagerie médicale ainsi que leur principe de fonctionnement. Nous nous sommes principalement intéressés dans cette thèse à la microscopie confocale à réflectance (MCR) qui est une technique d'imagerie non invasive permettant une visualisation *in vivo* de l'épiderme jusqu'au derme papillaire en temps réel. Le but final est de distinguer les patients sains de ceux atteints de lentigo qui est une hyperpigmentation cutanée.

Nous nous sommes intéressés dans le deuxième chapitre à l'analyse des images MCR pour caractériser la profondeur à laquelle on peut distinguer une peau saine de celle atteinte de lentigo. Chaque patient (sain ou lentigo) est représenté par une suite d'images MCR (appelée image *stack* dans ce manuscrit) qui sont associées aux différentes profondeurs. La caractérisation des textures de ces images a été faite à travers l'utilisation de la décomposition en ondelettes de Daubechies 4 pour chaque profondeur et pour tous les patients (lentigo et sains). Un premier travail a consisté à vérifier que la loi des coefficients d'ondelettes d'une image MCR était une loi gaussienne généralisée comme cela a été proposé dans la littérature pour d'autres types d'images. Les paramètres de cette loi ont été ensuite estimés par une méthode du maximum de vraisemblance. Un test paramétrique de Student a ensuite été conçu pour évaluer la signification statistique des différences observées entre les images des patients sains et lentigo. Le test d'hypothèses proposé a été effectué sur une base de données de

2250 images réelles associées à 45 patients. On a ainsi montré qu'il y avait un bon ajustement de la loi gaussienne généralisée pour les coefficients d'ondelettes des images MCR, et que l'entropie, la variance et le paramètre α ne permettent pas la discrimination d'images des patients sains et lentigo pour toutes les profondeurs, les bandes et les échelles. Par contre, la pseudo-distance de Kulback-Leibler ainsi que le paramètre β permettent la discrimination d'images des patients sains et lentigo pour toutes les bandes à des échelles de 1 à 3 et pour des profondeurs comprises entre 31 μm and 76 μm . Les profondeurs aux environs de 50 μm donnent la plus grande discrimination. Les résultats ont également montré qu'il n'y avait pas de direction préférée, et donc qu'il vaut mieux utiliser le vecteur conjoint *HVD* avec les premières échelles pertinentes pour avoir une bonne performance de caractérisation. Il est intéressant de noter que cette étude a révélé que le paramètre β caractérise le lentigo à des profondeurs aux environs de 50 μm ce qui correspond à l'emplacement de la jonction dermo-épidermique. Ce résultat est en accord avec le fait clinique que le lentigo désorganise significativement cette couche de la peau.

Le troisième chapitre a porté sur la modélisation statistique des intensités des images MCR. Nous avons montré tout d'abord que la loi gamma généralisée est bien adaptée pour décrire les propriétés statistiques des intensités des images des patients sains et atteints de lentigo. Pour estimer les paramètres de cette loi, nous avons proposé un estimateur qui s'avère être plus rapide et plus performant (pour un faible nombre d'échantillons) que ceux disponibles dans la littérature. Cet estimateur utilise la méthode du maximum de vraisemblance combinée avec une descente de gradient naturel. Ce dernier permet d'avoir une convergence plus rapide en pondérant le gradient par l'inverse de la matrice de Fisher des paramètres. Il a été montré que le lentigo se caractérise principalement à la profondeur DEJ. Les paramètres obtenus à cette profondeur ont ensuite été utilisés pour l'apprentissage d'un classifieur SVM. Le classifieur proposé a été testé sur une base de données de 2250 images réelles associées à 45 patients. Les résultats obtenus ont montré que le paramètre d'échelle β est bien adapté à la caractérisation et à la classification des tissus sains et des lésions. Notre algorithme a ensuite été comparé à la méthode présentée dans [Koller et al., 2011]. Cette méthode consiste à extraire à partir de chaque image MCR un ensemble de 39 paramètres et à leur appliquer une procédure de classification basée sur un arbre de classification et de régression (CART). Il faut noter que l'algorithme

CART a été testé sur les images MCR réelles en utilisant une procédure de *leave one out*. L'accuracy obtenue avec l'algorithme CART (en utilisant les profondeurs caractéristiques ($40 \mu m$ à $60 \mu m$)) est de 80 %, c'est-à-dire légèrement inférieure à celle obtenue avec la méthode proposée et conduit à un nombre de patients mal classés plus élevé (deux patients additionnel sur les 45 considérés). En outre, les paramètres FGD estimés peuvent être utilisés pour la caractérisation des images MCR, ce qui n'est pas possible avec l'algorithme CART.

Le quatrième chapitre présente un modèle bayésien hiérarchique totalement non-supervisé permettant une restauration des images MCR et une classification entre patients sains et atteints de lentigo. L'approche bayésienne proposée tient compte de la loi du bruit multiplicatif (*speckle*) affectant les mesures et introduit des lois a-priori adaptées aux paramètres inconnus du modèle, afin de régulariser le problème d'estimation. Nous avons ensuite proposé des méthodes de Monte Carlo par Chaines de Markov afin d'estimer conjointement les différents paramètres du modèle, incluant la vraie réflectivité de l'image, un vecteur de label permettant de classifier chaque pixel de l'image dans la classe "sain" ou "lentigo" et les paramètres de la loi du bruit. Plus précisément, un échantillonneur de Metropolis-within-Gibbs est utilisé pour générer des échantillons distribués asymptotiquement suivant la loi a-posteriori d'intérêt. Ces échantillons sont ensuite utilisés pour construire des estimateurs bayésiens, comme l'estimateur du maximum a posteriori (MAP) ou l'estimateur du minimum mean square error (MMSE) des paramètres inconnus du modèle. L'algorithme résultant est validé sur des données synthétiques et réelles en l'appliquant cette fois-ci seulement pour la profondeur caractéristique ($54 \mu m$) qui permet de distinguer avec une différence maximale les images des patients sains de ceux atteints de lentigo (cf chapitres précédents).. L'algorithme a été comparé à la méthode CART. Le paramètre "accuracy" obtenu avec l'algorithme CART (en utilisant seulement la profondeur $54 \mu m$) est de 82.2%. Il est légèrement inférieur à celui obtenu avec la méthode proposée. De plus, le modèle bayésien proposé peut être utilisé pour la caractérisation des images MCR grâce à ses paramètres estimés.

5.1 Perspectives

Le travail présenté dans cette thèse ouvre plusieurs perspectives.

Concernant la première partie de la thèse (chapitre 2 et 3), une première piste serait de prendre en compte différents types de peau ainsi que d'autres lésions cutanées pigmentées qui entraînent la destruction de la jonction dermo-épidermique comme le mélanome par exemple. Il serait également intéressant de combiner les paramètres discriminants associés à la gaussienne généralisée (associée aux coefficients d'ondelettes des images de la peau) avec les paramètres discriminants associés à la gamma généralisée (domaine de l'image) pour segmenter les zones avec lentigo sur les images blocs. La taille d'une image bloc étant de 8×8 mm, elle représente un champ large (16000×16000 pixel) où on peut faire la distinction entre la zone saine de la zone avec lentigo puisque celui-ci a généralement un diamètre inférieur à 5 mm. Un travail futur consisterait à décomposer cette image bloc en plusieurs imagerie et d'utiliser les paramètres des deux lois gaussienne généralisée et gamma généralisée estimées pour chaque imagerie afin de segmenter les zones avec lentigo de ceux saines. Un exemple d'image bloc et de la zone à segmenter est présenté dans la Figure 5.1

Concernant la deuxième partie de la thèse (chapitre 4), une première piste consisterait à améliorer la partie débruitage du modèle bayésien proposé. En effet on pourrait introduire la corrélation spatiale comme information a priori sur les images estimées sans bruit pour améliorer leur qualité, par exemple à l'aide d'un champs de Markov de loi gamma (gamma Markov random field). Une deuxième piste serait d'améliorer le temps de calculs de la partie classification de notre modèle bayésien. L'idée est de développer un nouveau modèle bayésien non-paramétrique pour classifier les images MCR. Ce dernier exploitera uniquement les paramètres de la loi gamma ajustée à ces images et non pas les intensités des pixels utilisées dans notre modèle bayésien proposé. Ce modèle va permettre aussi l'estimation du nombre de classes et de détecter éventuellement s'il y a des données aberrantes. L'approche bayésienne proposée dans cette piste sera basée sur la supposition que les paramètres de la loi gamma sont distribuées suivant un modèle de mélange de plusieurs lois gaussiennes bivariées. Un mélange de processus de Dirichlet (Dirichlet process mixture model, DPMM) peut alors être utilisé pour estimer les paramètres des différentes gaussiennes ainsi que leur nombre. Plus précisément, un échantillonneur de Gibbs pourrait être utilisé pour générer des échantillons distribués asymptotiquement suivant la loi a-posteriori des différents paramètres. Ces échantillons peuvent ensuite être utilisés pour estimer les différents paramètres inconnus du modèle de mélange, via les estimateurs bayésiens MAP et MMSE.

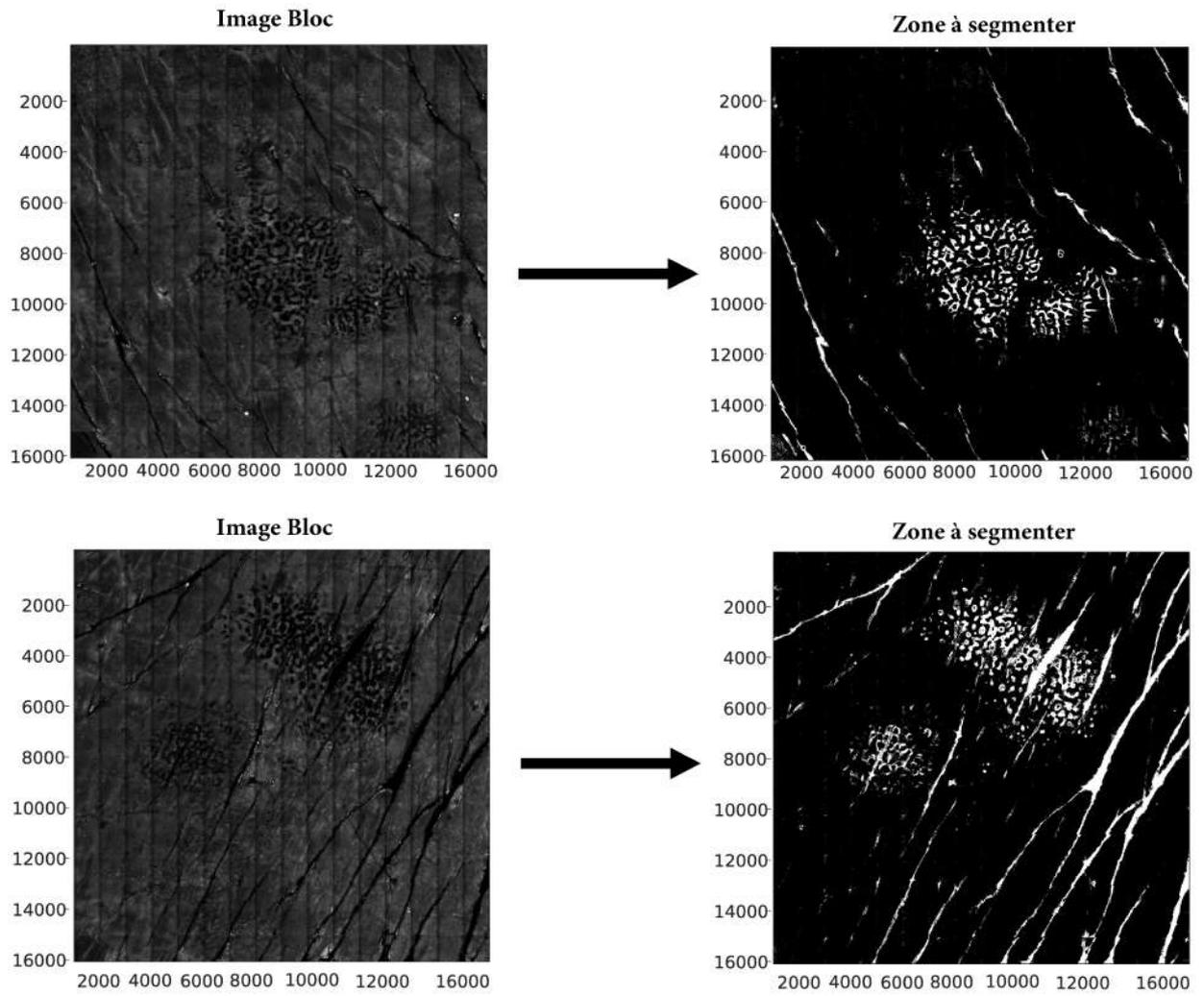


Figure 5.1: Exemples d'images blocs ainsi que les zones à segmenter

Nous avons déjà commencer à travailler sur ce modèle. Un résumé du modèle ainsi que quelques premiers résultats obtenus sont présentés dans l'annexe [G](#).

Annexes

ANNEXE A

T-test et facteur de Bayes

Nous considérons un corpus d'images constitué de n_1 patients sains et de n_2 patients avec lentigo, tel que annoté par un expert. Soit $\boldsymbol{\theta}_H = \{\boldsymbol{\theta}_H^1, \dots, \boldsymbol{\theta}_H^{n_1}\}$, désignant l'ensemble des paramètres estimés pour les patients sains et $\boldsymbol{\theta}_S = \{\boldsymbol{\theta}_S^1, \dots, \boldsymbol{\theta}_S^{n_2}\}$, pour les patients lentigo à une profondeur donnée. L'estimateur du maximum de vraisemblance (EMV) est connu pour être asymptotiquement gaussien et asymptotiquement efficace [Hurlin, 2013, Scholz, 1985], et on suppose que ces paramètres sont indépendants et suivent des distributions normales.

Un test T à deux échantillons [Cressie and Whitford, 1986, Rakotomalala, 2013, Wendorf, 2004] peut alors être appliqué pour comparer les moyennes μ_H and μ_S .

$$H_0 : \mu_H = \mu_S, \tag{A.1}$$

$$H_1 : \mu_H \neq \mu_S. \tag{A.2}$$

Les variances σ_H^2 et σ_S^2 (des lois GGD sous les 2 hypothèses) étant inconnues et inégales, nous proposons d'appliquer le test comme suit: Désignons par $\bar{\boldsymbol{\theta}}_H$ et $\bar{\boldsymbol{\theta}}_S$ les moyennes empiriques de $\boldsymbol{\theta}_H$ et $\boldsymbol{\theta}_S$ et par s_H^2 et s_S^2 les estimateurs non biaisés des variances pour les patients sains et lentigo.

La statistique du T-test associées à (A.1) et (A.2) est donc :

$$T = \frac{\bar{\boldsymbol{\theta}}_H - \bar{\boldsymbol{\theta}}_S}{\sqrt{\frac{s_H^2}{n_1} + \frac{s_S^2}{n_2}}}. \tag{A.3}$$

qui est distribuée selon une distribution de Student T avec ν degrés de liberté sous l'hypothèse H_0 , où ν est défini comme

$$\nu = \frac{\left[\frac{s_H^2}{n_1} + \frac{s_S^2}{n_2} \right]^2}{\frac{s_H^4}{n_1^2(n_1-1)} + \frac{s_S^4}{n_2^2(n_2-1)}}. \quad (\text{A.4})$$

L'hypothèse H_0 est rejetée si $|T_\nu| > T_{PFA}$. Dans ce travail, nous avons choisi une probabilité de fausse alarme $PFA = 0.05$ ce qui correspond à un seuil de $T_{PFA} = 2.02$ pour $n_1 = 18$ et $n_2 = 27$. Pour évaluer la signification statistique, la p-value de chaque test a également été calculée et les règles de décision suivantes ont été appliquées

- Quand $p \text{ value} > 0.10$ → la différence observée “ n'est pas significative ”
- Quand $p \text{ value} \in [0.05, 0.10]$ → la différence observée est “ légèrement significative ”
- Quand $p \text{ value} \in [0.01, 0.05[$ → la différence observée est “ significative ”
- Quand $p \text{ value} < 0.01$ → la différence observée est “ hautement significative ”

Remarque : compte tenu du débat récent sur la p-value et la reproductibilité des résultats scientifiques, une méthode a été développée dans [Johnson, 2013] pour établir une correspondance entre les tests de signification classiques, tels que celui conçu ici, et les tests Bayésiens. Cette méthode permet de rapporter la taille des tests d'hypothèses classiques avec des seuils dans les tests Bayésiens. Suite à ce travail et en supposant des variances égales, nous avons calculé le facteur de Bayes (BF) donné par:

$$BF = \left(\frac{\nu + T}{\nu + \left(T - \sqrt{\nu \alpha^*} \right)^2} \right)^{(n_1+n_2)/2} \quad (\text{A.5})$$

où l'hypothèse H_0 est rejetée quand $BF > \sqrt{\nu \alpha^*}$ with $\alpha^* = \alpha^{2/(n_1+n_2-1)} - 1$ et $\alpha = [(T_{PFA}^2/\nu) + 1]^{(n_1+n_2)/2}$. Ce facteur BF a fourni la même preuve que les p-values considérées comme cela apparaîtra dans la partie expérimentale.

ANNEXE B

Matrice de Fisher - Gaussienne Généralisée

Distribution Gaussienne Généralisée Une gaussienne généralisée est une fonction $G_{\alpha,\beta}(x)$ dépendant de deux paramètres $\alpha > 0$ et $\beta > 0$:

$$G_{\alpha,\beta}(x) = \frac{\beta}{2 \alpha \Gamma\left(\frac{1}{\beta}\right)} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} \quad \alpha > 0 \quad \beta > 0$$

Quand $\beta = 2$, on retrouve l'équation d'une gaussienne, quand $\beta = 1$, celle d'une distribution de Laplace. β est le paramètre qui régit la forme, plus ou moins pointue, de la gaussienne généralisée, c'est le **paramètre de forme**. α , lui, régit l'étalement de la courbe : c'est le **paramètre d'échelle**.

Rappel [Abramowitz and Stegun, 1970]

$$y = \alpha^x \quad \Rightarrow \quad y' = \frac{dy}{dx} = \log(\alpha) \times \alpha^x$$

$$\Gamma(n+1) = n! \quad \text{pour } n \text{ entier} \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$\Gamma(z+1) = z \Gamma(z)$$

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \quad \Rightarrow \quad \Gamma'(z) = \frac{d\Gamma(z)}{dz} = \int_0^\infty e^{-t} \frac{d}{dz}(t^{z-1}) dt = \int_0^\infty e^{-t} \ln(t) t^{z-1} dt$$
$$\Gamma''(z) = \int_0^\infty e^{-t} \ln(t) \frac{d}{dz}(t^{z-1}) dt = \int_0^\infty e^{-t} \ln(t)^2 t^{z-1} dt$$

$$\Gamma'(z) = \int_0^\infty e^{-t} \ln(t) t^{z-1} dt, \quad \Gamma''(z) = \int_0^\infty e^{-t} \ln(t)^2 t^{z-1} dt$$

la première dérivée $\frac{d \ln(\Gamma(z))}{dz} = \frac{\Gamma'(z)}{\Gamma(z)} = \psi(z)$ est appelée fonction *digamma*(z)

La dérivée seconde $\frac{d^2 \ln(\Gamma(z))}{dz^2} = \psi'(z)$ est appelée fonction *trigamma*(z)

Formule de récurrence

$$\psi(1+z) = \psi(z) + \frac{1}{z}$$

$$\psi^{(n)}(1+z) = \psi^{(n)}(z) + (-1)^n n! z^{-n-1}$$

ex: $\psi'(1+z) = \psi'(z) - z^{-2}$

Matrice de fisher d'une distribution $GGD(\alpha, \beta)$:

$$G(\alpha, \beta) = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} \quad \alpha > 0 \quad \beta > 0$$

$$\log(G(\alpha, \beta)) = \log(\beta) - \log(2\alpha) - \log\left(\Gamma\left(\frac{1}{\beta}\right)\right) - \left(\frac{|x|}{\alpha}\right)^\beta \quad (\text{B.1})$$

Calcul des dérivées

dérivée par rapport à α

$$\frac{\partial \log(G)}{\partial \alpha} = -\frac{2}{2\alpha} - \beta \left(\frac{|x|}{\alpha}\right)^{\beta-1} \left(-\frac{|x|}{\alpha^2}\right) = -\frac{1}{\alpha} + \beta \frac{|x|^\beta}{\alpha^{\beta+1}}$$

dérivée par rapport à α^2

$$\frac{\partial^2 \log(G)}{\partial \alpha^2} = \frac{1}{\alpha^2} + \beta |x|^\beta \left(\frac{-(\beta+1)\alpha^\beta}{\alpha^{2(\beta+1)}}\right) = \frac{1}{\alpha^2} - \beta(\beta+1) |x|^\beta \frac{1}{\alpha^{\beta+2}}$$

$$\frac{\partial^2 \log(G)}{\partial \alpha^2} = \frac{1}{\alpha^2} - \beta(\beta+1) \frac{|x|^\beta}{\alpha^{\beta+2}}$$

dérivée par rapport à β

$$\frac{\partial \log(G)}{\partial \beta} = \frac{1}{\beta} - \left(-\frac{1}{\beta^2}\right) \frac{\Gamma'\left(\frac{1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} - \log\left(\frac{|x|}{\alpha}\right) \left(\frac{|x|}{\alpha}\right)^\beta$$

$$\frac{\partial \log(G)}{\partial \beta} = \frac{1}{\beta} + \frac{1}{\beta^2} \text{digamma}\left(\frac{1}{\beta}\right) - \left(\frac{|x|}{\alpha}\right)^\beta \log\left(\frac{|x|}{\alpha}\right) = \frac{1}{\beta} + \frac{1}{\beta^2} \psi\left(\frac{1}{\beta}\right) - \left(\frac{|x|}{\alpha}\right)^\beta \log\left(\frac{|x|}{\alpha}\right)$$

dérivée par rapport à β^2

$$\frac{\partial^2 \log(G)}{\partial \beta^2} = -\frac{1}{\beta^2} + \left(-\frac{2\beta}{\beta^4}\right) \psi\left(\frac{1}{\beta}\right) + \frac{1}{\beta^2} \left(-\frac{1}{\beta^2}\right) \text{trigamma}\left(\frac{1}{\beta}\right) - \left(\frac{|x|}{\alpha}\right)^\beta \log\left(\frac{|x|}{\alpha}\right) \log\left(\frac{|x|}{\alpha}\right)$$

$$\frac{\partial^2 \log(G)}{\partial \beta^2} = -\frac{1}{\beta^2} - \frac{2}{\beta^3} \psi\left(\frac{1}{\beta}\right) - \frac{1}{\beta^4} \psi'\left(\frac{1}{\beta}\right) - \left(\frac{|x|}{\alpha}\right)^\beta \left[\log\left(\frac{|x|}{\alpha}\right)\right]^2$$

dérivée par rapport à α, β

$$\frac{\partial}{\partial \alpha} \left(\frac{\partial \log(G)}{\partial \beta} \right) = - \left(\frac{-\frac{|x|}{\alpha^2}}{\frac{|x|}{\alpha}} \right) \left(\frac{|x|}{\alpha} \right)^\beta - \log\left(\frac{|x|}{\alpha}\right) \beta \left(\frac{|x|}{\alpha} \right)^{\beta-1} \left(-\frac{|x|}{\alpha^2} \right)$$

$$\frac{\partial}{\partial \alpha} \left(\frac{\partial \log(G)}{\partial \beta} \right) = \frac{|x|^\beta}{\alpha^{\beta+1}} + \beta \log \left(\frac{|x|}{\alpha} \right) \frac{|x|^\beta}{\alpha^{\beta+1}} = \frac{|x|^\beta}{\alpha^{\beta+1}} \left(1 + \beta \log \left(\frac{|x|}{\alpha} \right) \right)$$

$$\frac{\partial}{\partial \alpha} \left(\frac{\partial \log(G)}{\partial \beta} \right) = \frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} = \frac{\partial^2 \log(G)}{\partial \beta \partial \alpha} = \frac{|x|^\beta}{\alpha^{\beta+1}} \left(1 + \beta \log \left(\frac{|x|}{\alpha} \right) \right)$$

Matrice de Fisher

$$\text{Fisher matrix} = - \begin{pmatrix} E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) & E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) \\ E \left(\frac{\partial^2 \log(G)}{\partial \beta \partial \alpha} \right) & E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) \end{pmatrix}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) = -E \left(\frac{1}{\alpha^2} - \beta(\beta+1) \frac{|x^\beta|}{\alpha^{\beta+2}} \right) = E \left(-\frac{1}{\alpha^2} + \beta(\beta+1) \frac{|x^\beta|}{\alpha^{\beta+2}} \right)$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) = \int_{-\infty}^{+\infty} \left(-\frac{1}{\alpha^2} + \beta(\beta+1) \frac{|x^\beta|}{\alpha^{\beta+2}} \right) G(\alpha, \beta) dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) = \int_{-\infty}^{+\infty} -\frac{1}{\alpha^2} G(\alpha, \beta) dx + \int_{-\infty}^{+\infty} \beta(\beta+1) \frac{|x^\beta|}{\alpha^{\beta+2}} G(\alpha, \beta) dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) = -\frac{1}{\alpha^2} + \int_{-\infty}^{+\infty} \beta(\beta+1) \frac{|x^\beta|}{\alpha^{\beta+2}} \frac{\beta}{2\alpha \Gamma\left(\frac{1}{\beta}\right)} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) = -\frac{1}{\alpha^2} + \frac{(\beta+1)}{\alpha^2} = \frac{\beta}{\alpha^2}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) = \frac{\beta}{\alpha^2}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) = -E \left(\frac{\partial^2 \log(G)}{\partial \beta \partial \alpha} \right) = E \left[-\frac{|x|^\beta}{\alpha^{\beta+1}} \left(1 + \beta \log \left(\frac{|x|}{\alpha} \right) \right) \right]$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) = \int_{-\infty}^{+\infty} \left[-\frac{|x|^\beta}{\alpha^{\beta+1}} \left(1 + \beta \log \left(\frac{|x|}{\alpha} \right) \right) \right] G(\alpha, \beta) dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) = \int_{-\infty}^{+\infty} \left[-\frac{|x|^\beta}{\alpha^{\beta+1}} \left(1 + \beta \log \left(\frac{|x|}{\alpha} \right) \right) \right] \frac{\beta}{2\alpha \Gamma\left(\frac{1}{\beta}\right)} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) = -\frac{1}{\alpha} \frac{\Gamma\left(1 + \frac{1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} - \frac{1}{\alpha} \frac{\Gamma'\left(1 + \frac{1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} = -\frac{1}{\alpha} \frac{\Gamma\left(1 + \frac{1}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \left[1 + \frac{\Gamma'\left(1 + \frac{1}{\beta}\right)}{\Gamma\left(1 + \frac{1}{\beta}\right)} \right]$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) = -\frac{1}{\alpha \beta} \left[1 + \text{digamma} \left(1 + \frac{1}{\beta} \right) \right] = -\frac{1}{\alpha \beta} \left[1 + \psi \left(1 + \frac{1}{\beta} \right) \right]$$

$$\Psi(1+z) = \psi(z) + \frac{1}{z}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) = -\frac{1}{\alpha \beta} \left[1 + \psi \left(\frac{1}{\beta} \right) + \beta \right]$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) = -E \left(-\frac{1}{\beta^2} - \frac{2}{\beta^3} \psi \left(\frac{1}{\beta} \right) - \frac{1}{\beta^4} \psi' \left(\frac{1}{\beta} \right) - \left(\frac{|x|}{\alpha} \right)^\beta \left[\log \left(\frac{|x|}{\alpha} \right) \right]^2 \right)$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) = E \left(\frac{1}{\beta^2} + \frac{2}{\beta^3} \psi \left(\frac{1}{\beta} \right) + \frac{1}{\beta^4} \psi' \left(\frac{1}{\beta} \right) + \left(\frac{|x|}{\alpha} \right)^\beta \left[\log \left(\frac{|x|}{\alpha} \right) \right]^2 \right)$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) = \int_{-\infty}^{+\infty} \left(\frac{1}{\beta^2} + \frac{2}{\beta^3} \psi \left(\frac{1}{\beta} \right) + \frac{1}{\beta^4} \psi' \left(\frac{1}{\beta} \right) + \left(\frac{|x|}{\alpha} \right)^\beta \left[\log \left(\frac{|x|}{\alpha} \right) \right]^2 \right) G(\alpha, \beta) dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) = \frac{1}{\beta^2} + \frac{2}{\beta^3} \psi \left(\frac{1}{\beta} \right) + \frac{1}{\beta^4} \psi' \left(\frac{1}{\beta} \right) + \int_{-\infty}^{+\infty} \left(\frac{|x|}{\alpha} \right)^\beta \left[\log \left(\frac{|x|}{\alpha} \right) \right]^2 G(\alpha, \beta) dx$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) = \frac{1}{\beta^2} + \frac{2}{\beta^2} \left[1 + \frac{1}{\beta} \right] \psi \left(\frac{1}{\beta} \right) + \frac{1}{\beta^3} \left[1 + \frac{1}{\beta} \right] \psi' \left(\frac{1}{\beta} \right) + \frac{1}{\beta^3} \left[\psi \left(\frac{1}{\beta} \right) \right]^2$$

$$\text{Fisher matrix} = - \begin{pmatrix} E \left(\frac{\partial^2 \log(G)}{\partial \alpha^2} \right) & E \left(\frac{\partial^2 \log(G)}{\partial \alpha \partial \beta} \right) \\ E \left(\frac{\partial^2 \log(G)}{\partial \beta \partial \alpha} \right) & E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\beta}{\alpha^2} & -\frac{1}{\alpha \beta} \left[1 + \psi \left(\frac{1}{\beta} \right) + \beta \right] \\ -\frac{1}{\alpha \beta} \left[1 + \psi \left(\frac{1}{\beta} \right) + \beta \right] & \frac{1}{\beta^2} + \frac{2}{\beta^2} \left[1 + \frac{1}{\beta} \right] \psi \left(\frac{1}{\beta} \right) + \frac{1}{\beta^3} \left[1 + \frac{1}{\beta} \right] \psi' \left(\frac{1}{\beta} \right) + \frac{1}{\beta^3} \left[\psi \left(\frac{1}{\beta} \right) \right]^2 \end{pmatrix}$$

ANNEXE C

Matrice de Fisher - Gamma Généralisée

Distribution gamma généralisée

$$g_{\gamma, \beta, \rho}(x) = \begin{cases} \frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} & \text{si } x > \gamma \quad \rho > 0 \\ 0 & \text{sinon} \end{cases}$$

Calcul du logarithme de la densité de probabilité *Log - likelihood* ($G = \prod g(\gamma, \beta, \rho)$ représente le produit des gamma généralisée)

$$\text{Log - likelihood} = \log(G) = \log\left(\prod_{i=1}^N \left[\frac{\beta^{-\rho}}{\Gamma(\rho)} (x_i - \gamma)^{\rho-1} e^{-\frac{(x_i-\gamma)}{\beta}}\right]\right)$$

$$\log(G) = -N\rho \log(\beta) - N \log(\Gamma(\rho)) + (\rho - 1) \sum_{i=1}^N \log(x_i - \gamma) - \sum_{i=1}^N \frac{(x_i - \gamma)}{\beta}$$

Calcul des dérivées

dérivée par rapport à γ

$$\frac{\partial \log(G)}{\partial \gamma} = - \sum_{i=1}^N \frac{(\rho - 1)}{(x_i - \gamma)} + \frac{N}{\beta}$$

dérivée par rapport à γ^2

$$\frac{\partial^2 \log(G)}{\partial \gamma^2} = - \sum_{i=1}^N \frac{(\rho - 1)}{(x_i - \gamma)^2}$$

dérivée par rapport à β

$$\frac{\partial \log(G)}{\partial \beta} = -\frac{N\rho}{\beta} + \sum_{i=1}^N \frac{(x_i - \gamma)}{\beta^2}$$

dérivée par rapport à β^2

$$\frac{\partial^2 \log(G)}{\partial \beta^2} = \frac{N\rho}{\beta^2} - 2 \sum_{i=1}^N \frac{(x_i - \gamma)}{\beta^3}$$

dérivée par rapport à ρ

$$\frac{\partial \log(G)}{\partial \rho} = -N \log(\beta) - N \frac{\Gamma'(\rho)}{\Gamma(\rho)} + \sum_{i=1}^N \log(x_i - \gamma)$$

$$\frac{\partial \log(G)}{\partial \rho} = -N \log(\beta) - N \text{digamma}(\rho) + \sum_{i=1}^N \log(x_i - \gamma) = -N \log(\beta) - N \psi(\rho) + \sum_{i=1}^N \log(x_i - \gamma)$$

dérivée par rapport à ρ^2

$$\frac{\partial^2 \log(G)}{\partial \rho^2} = -N \psi'(\rho) = -N \text{trigamma}(\rho)$$

dérivée par rapport à γ, β

$$\frac{\partial}{\partial \gamma} \left(\frac{\partial \log(G)}{\partial \beta} \right) = \frac{\partial}{\partial \beta} \left(\frac{\partial \log(G)}{\partial \gamma} \right) = \frac{\partial^2 \log(G)}{\partial \gamma \partial \beta} = \frac{\partial^2 \log(G)}{\partial \beta \partial \gamma} = -\frac{N}{\beta^2}$$

dérivée par rapport à γ, ρ

$$\frac{\partial}{\partial \gamma} \left(\frac{\partial \log(G)}{\partial \rho} \right) = \frac{\partial}{\partial \rho} \left(\frac{\partial \log(G)}{\partial \gamma} \right) = \frac{\partial^2 \log(G)}{\partial \gamma \partial \rho} = \frac{\partial^2 \log(G)}{\partial \rho \partial \gamma} = -\sum_{i=1}^N \frac{1}{(x_i - \gamma)}$$

dérivée par rapport à β, ρ

$$\frac{\partial}{\partial \beta} \left(\frac{\partial \log(G)}{\partial \rho} \right) = \frac{\partial}{\partial \rho} \left(\frac{\partial \log(G)}{\partial \beta} \right) = \frac{\partial^2 \log(G)}{\partial \beta \partial \rho} = \frac{\partial^2 \log(G)}{\partial \rho \partial \beta} = -\frac{N}{\beta}$$

Matrice de Fisher

$$\text{Matrice de Fisher} = - \begin{pmatrix} E\left(\frac{\partial^2 \log(G)}{\partial \gamma^2}\right) & E\left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \beta}\right) & E\left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \rho}\right) \\ E\left(\frac{\partial^2 \log(G)}{\partial \beta \partial \gamma}\right) & E\left(\frac{\partial^2 \log(G)}{\partial \beta^2}\right) & E\left(\frac{\partial^2 \log(G)}{\partial \beta \partial \rho}\right) \\ E\left(\frac{\partial^2 \log(G)}{\partial \rho \partial \gamma}\right) & E\left(\frac{\partial^2 \log(G)}{\partial \rho \partial \beta}\right) & E\left(\frac{\partial^2 \log(G)}{\partial \rho^2}\right) \end{pmatrix}$$

Calcul des différents termes de la Matrice de Fisher

$$\begin{aligned}
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma^2} \right) &= -E \left(- \sum_{i=1}^N \frac{(\rho-1)}{(x_i - \gamma)^2} \right) = E \left(\sum_{i=1}^N \frac{(\rho-1)}{(x_i - \gamma)^2} \right) \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma^2} \right) &= \int_{\gamma}^{+\infty} \left(\sum_{i=1}^N \frac{(\rho-1)}{(x_i - \gamma)^2} \right) g_{\gamma, \beta, \rho}(x_i) dx_i \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma^2} \right) &= (\rho-1) \int_{\gamma}^{+\infty} \left(\sum_{i=1}^N \frac{1}{(x_i - \gamma)^2} \right) \frac{\beta^{-\rho}}{\Gamma(\rho)} (x_i - \gamma)^{\rho-1} e^{-\frac{(x_i - \gamma)}{\beta}} dx_i \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma^2} \right) &= \frac{N}{\beta^2 (\rho-2)} \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \beta} \right) &= -E \left(\frac{\partial^2 \log(G)}{\partial \beta \partial \gamma} \right) = -E \left[-\frac{N}{\beta^2} \right] = E \left[\frac{N}{\beta^2} \right] = \frac{N}{\beta^2} \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \rho} \right) &= -E \left(\frac{\partial^2 \log(G)}{\partial \rho \partial \gamma} \right) = -E \left[- \sum_{i=1}^N \frac{1}{(x_i - \gamma)} \right] = E \left[\sum_{i=1}^N \frac{1}{(x_i - \gamma)} \right] \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \rho} \right) &= \int_{\gamma}^{+\infty} \left(\sum_{i=1}^N \frac{1}{(x_i - \gamma)} \right) g_{\gamma, \beta, \rho}(x_i) dx_i \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \rho} \right) &= \int_{\gamma}^{+\infty} \left(\sum_{i=1}^N \frac{1}{(x_i - \gamma)} \right) \frac{\beta^{-\rho}}{\Gamma(\rho)} (x_i - \gamma)^{\rho-1} e^{-\frac{(x_i - \gamma)}{\beta}} dx_i \\
-E \left(\frac{\partial^2 \log(G)}{\partial \gamma \partial \rho} \right) &= \frac{N}{\beta} \frac{\Gamma(\rho-1)}{\Gamma(\rho)} = \frac{N}{\beta} \frac{\Gamma(\rho-1)}{(\rho-1)\Gamma(\rho-1)} = \frac{N}{\beta(\rho-1)} \\
-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) &= -E \left(\frac{N\rho}{\beta^2} - 2 \sum_{i=1}^N \frac{(x_i - \gamma)}{\beta^3} \right) = E \left(-\frac{N\rho}{\beta^2} + 2 \sum_{i=1}^N \frac{(x_i - \gamma)}{\beta^3} \right) \\
-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) &= \int_{\gamma}^{+\infty} \left(-\frac{N\rho}{\beta^2} + 2 \sum_{i=1}^N \frac{(x_i - \gamma)}{\beta^3} \right) g_{\gamma, \beta, \rho}(x_i) dx_i \\
-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) &= -\frac{N\rho}{\beta^2} + \frac{2}{\beta^3} \int_{\gamma}^{+\infty} \sum_{i=1}^N (x_i - \gamma) g_{\gamma, \beta, \rho}(x_i) dx_i
\end{aligned}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta^2} \right) = -\frac{N\rho}{\beta^2} + \frac{2}{\beta^2} N \frac{\Gamma(\rho+1)}{\Gamma(\rho)} = -\frac{N\rho}{\beta^2} + \frac{2N}{\beta^2} \frac{\rho \Gamma(\rho)}{\Gamma(\rho)} = -\frac{N\rho}{\beta^2} + \frac{2N\rho}{\beta^2} = \frac{N\rho}{\beta^2}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \beta \partial \rho} \right) = -E \left(\frac{\partial^2 \log(G)}{\partial \rho \partial \beta} \right) = -E \left[-\frac{N}{\beta} \right] = E \left[\frac{N}{\beta} \right] = \frac{N}{\beta}$$

$$-E \left(\frac{\partial^2 \log(G)}{\partial \rho^2} \right) = -E \left(-N \psi'(\rho) \right) = E \left(N \psi'(\rho) \right) = N \psi'(\rho) = N \text{ trigamma}(\rho)$$

On obtient finalement (en écrivant *digamma* (z) comme $\psi(z)$ et *trigamma* (z) comme $\psi'(z)$)

$$\text{Matrice de Fisher} = N \begin{pmatrix} \frac{1}{\beta^2(\rho-2)} & \frac{1}{\beta^2} & \frac{1}{\beta(\rho-1)} \\ \frac{1}{\beta^2} & \frac{\rho}{\beta^2} & \frac{1}{\beta} \\ \frac{1}{\beta(\rho-1)} & \frac{1}{\beta} & \psi'(\rho) \end{pmatrix}$$

ANNEXE D

Divergence de Kullback-Leibler, entropie, variance

Distribution Gamma Généralisée

$$g_{\gamma, \beta, \rho}(x) = \begin{cases} \frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} & \text{si } x > \gamma \quad \rho > 0 \\ 0 & \text{sinon} \end{cases}$$

D.1 Kullback Liebler

$$KL = \int_{-\infty}^{+\infty} p(x) \ln \left(\frac{p(x)}{q(x)} \right) dx$$

$$KL = \int_{-\infty}^{+\infty} \frac{\beta_1^{-\rho_1}}{\Gamma(\rho_1)} (x - \gamma_1)^{\rho_1-1} e^{-\frac{(x-\gamma_1)}{\beta_1}} \ln \left(\frac{\frac{\beta_1^{-\rho_1}}{\Gamma(\rho_1)} (x - \gamma_1)^{\rho_1-1} e^{-\frac{(x-\gamma_1)}{\beta_1}}}{\frac{\beta_2^{-\rho_2}}{\Gamma(\rho_2)} (x - \gamma_2)^{\rho_2-1} e^{-\frac{(x-\gamma_2)}{\beta_2}}} \right) dx$$

$$KL = \int_{-\infty}^{+\infty} \frac{\beta_1^{-\rho_1}}{\Gamma(\rho_1)} (x - \gamma_1)^{\rho_1-1} e^{-\frac{(x-\gamma_1)}{\beta_1}} \ln \left(\frac{\beta_2^{\rho_2} \Gamma(\rho_2)}{\beta_1^{\rho_1} \Gamma(\rho_1)} \frac{(x - \gamma_1)^{\rho_1-1}}{(x - \gamma_2)^{\rho_2-1}} e^{-\frac{(x-\gamma_1)}{\beta_1}} e^{\frac{(x-\gamma_2)}{\beta_2}} \right) dx$$

$$\begin{aligned} KL &= \ln \left(\frac{\beta_2^{\rho_2} \Gamma(\rho_2)}{\beta_1^{\rho_1} \Gamma(\rho_1)} \right) + (\rho_1 - 1) [\ln(\beta_1) + \psi(\rho_1)] + \rho_1 \left(\frac{\beta_1}{\beta_2} - 1 \right) + \left(\frac{\gamma_1 - \gamma_2}{\beta_2} \right) \\ &\quad - (\rho_2 - 1) \ln(\beta_1) - \frac{(\rho_2 - 1)}{\Gamma(\rho_1)} \int_0^{+\infty} t^{\rho_1-1} e^{-t} \ln \left(t + \frac{\gamma_1 - \gamma_2}{\beta_1} \right) dt \end{aligned}$$

Remarque: si $\gamma_1 = \gamma_2$

$$KL = \ln \left(\frac{\beta_2^{\rho_2} \Gamma(\rho_2)}{\beta_1^{\rho_1} \Gamma(\rho_1)} \right) + (\rho_1 - 1) [\ln(\beta_1) + \psi(\rho_1)] + \rho_1 \left(\frac{\beta_1}{\beta_2} - 1 \right) - (\rho_2 - 1) \ln(\beta_1) - (\rho_2 - 1) \frac{\Gamma'(\rho_1)}{\Gamma(\rho_1)}$$

$$\begin{aligned}
KL &= \ln \left(\frac{\beta_2^{\rho_2} \Gamma(\rho_2)}{\beta_1^{\rho_1} \Gamma(\rho_1)} \right) + (\rho_1 - 1) [\ln(\beta_1) + \psi(\rho_1)] + \rho_1 \left(\frac{\beta_1}{\beta_2} - 1 \right) - (\rho_2 - 1) [\ln(\beta_1) + \psi(\rho_1)] \\
KL &= \ln \left(\frac{\Gamma(\rho_2) \beta_2^{\rho_2}}{\Gamma(\rho_1) \beta_1^{\rho_1}} \right) + (\rho_1 - \rho_2) [\ln(\beta_1) + \psi(\rho_1)] + \rho_1 \left(\frac{\beta_1}{\beta_2} - 1 \right) \quad (D.1)
\end{aligned}$$

D.2 Variance

$$\begin{aligned}
\text{Variance} &= E(X^2) - E(X)^2 \\
E(X^2) &= \int_{\gamma}^{+\infty} x^2 \frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} dx \\
E(X^2) &= \beta^2 \rho(\rho + 1) + \gamma^2 + 2\beta\gamma\rho \\
E(X) &= \int_{\gamma}^{+\infty} x \frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} dx \\
E(X) &= \beta\rho + \gamma \\
\text{Variance} &= E(X^2) - E(X)^2 = \beta^2 \rho(\rho + 1) + \gamma^2 + 2\beta\gamma\rho - [\beta\rho + \gamma]^2 = \beta^2 \rho \\
\text{Variance} &= \beta^2 \rho \quad (D.2)
\end{aligned}$$

D.3 Entropie

$$\begin{aligned}
\text{Entropie} &= - \int_{-\infty}^{+\infty} p(x) \ln[p(x)] dx \\
\text{Entropie} &= - \int_{\gamma}^{+\infty} \frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} \ln \left[\frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} \right] dx \\
\text{Entropie} &= - \int_{\gamma}^{+\infty} \frac{\beta^{-\rho}}{\Gamma(\rho)} (x - \gamma)^{\rho-1} e^{-\frac{(x-\gamma)}{\beta}} \left[\ln \left(\frac{\beta^{-\rho}}{\Gamma(\rho)} \right) + (\rho - 1) \ln(x - \gamma) - \frac{(x - \gamma)}{\beta} \right] dx \\
\text{Entropie} &= \ln(\beta \Gamma(\rho)) + (1 - \rho) \psi(\rho) + \rho \quad (D.3)
\end{aligned}$$

ANNEXE E

Échantillonnage des paramètres

Les densités conditionnelles des paramètres d'intérêt sont obtenues en multipliant la vraisemblance avec les différents a priori et en enlevant le terme multiplicatif qui ne dépend pas de la variable d'intérêt. On obtient les résultats détaillés dans les paragraphes suivants:

E.1 Echantillonnage du paramètre s_{nl}

On obtient la loi conditionnelle suivante

$$f(s_{nl} | z_l = k, y_{nl}, \theta_l, \sigma_k^2, \mu_k) \propto \frac{1}{(s_{nl})^{1/\theta}} \exp\left(-\frac{y_{nl}}{s_{nl} \theta_l}\right) \times \exp\left[-\frac{1}{2 \sigma_k^2} (s_{nl} - \mu_k)^2\right] I_{\mathbb{R}^+}(s_{nl}) \quad (\text{E.1})$$

En utilisant la loi de proposition (gaussienne tronquée positive [Chari et al., 2015]) suivante

$$g(x | s_{nl}^t) \propto \frac{\frac{1}{\sqrt{2\pi} \epsilon_2} \exp\left[-\frac{(x-s_{nl}^t)^2}{2 \epsilon_2^2}\right]}{1 - \Phi\left(\frac{-s_{nl}^t}{\epsilon_2}\right)}, \text{ où } \Phi \text{ est la fonction de répartition d'une loi normale}$$

On obtient la règle d'acceptation/rejet suivante

$$s_{nl}^{t+1} = \begin{cases} s_{nl}^* & \text{with prob } \min\left\{\frac{f(s_{nl}^*)}{f(s_{nl}^t)} \frac{g(s_{nl}^t | s_{nl}^*)}{g(s_{nl}^* | s_{nl}^t)}, 1\right\} \\ s_{nl}^t & \text{else} \end{cases}$$

avec

$$\frac{f(s_{nl}^*)}{f(s_{nl}^t)} \frac{g(s_{nl}^t | s_{nl}^*)}{g(s_{nl}^* | s_{nl}^t)} = \left(\frac{s_{nl}^*}{s_{nl}^t}\right)^{-\rho_l} \left[\frac{1 - \Phi\left(\frac{-s_{nl}^t}{\epsilon_2}\right)}{1 - \Phi\left(\frac{-s_{nl}^*}{\epsilon_2}\right)} \right] \times \exp\left[\frac{-2 y_{nl} \sigma_k^2 - s_{nl}^* \theta_l (s_{nl}^* - \mu_k)^2}{2 \sigma_k^2 s_{nl}^* \theta_l} + \frac{2 y_{nl} \sigma_k^2 + s_{nl}^t \theta_l (s_{nl}^t - \mu_k)^2}{2 \sigma_k^2 s_{nl}^t \theta_l} \right].$$

E.2 Echantillonnage du paramètre θ_l

On obtient la loi conditionnelle suivante

$$f(\theta_l | y_{nl}, s_{nl}, a, b) \propto \frac{1}{\theta_l^{N/\theta_l + a + 1}} \prod_{n=1}^N \left(\frac{y_{nl}}{s_{nl}} \right)^{\frac{1}{\theta_l}} \times \exp \left[-\frac{1}{\theta_l} \left(\sum_{n=1}^N \frac{y_{nl}}{s_{nl}} - b \right) \right] [\Gamma(1/\theta_l)]^{-N} \quad (\text{E.2})$$

En utilisant la loi de proposition (gaussienne tronquée positive [Chaari et al., 2015]) suivante

$$g(x | \theta_l^t) \propto \frac{\frac{1}{\sqrt{2\pi}\epsilon_1} \exp \left[-\frac{(x - \theta_l^t)^2}{2\epsilon_1^2} \right]}{1 - \Phi \left(\frac{-\theta_l^t}{\epsilon_1} \right)}, \text{ où } \Phi \text{ est la fonction de répartition d'une loi normale}$$

On obtient la règle d'acceptation/rejet suivante

$$\theta_l^{t+1} = \begin{cases} \theta_l^* & \text{with prob } \min \left\{ \frac{f(\theta_l^*)}{f(\theta_l^t)} \frac{g(\theta_l^t | \theta_l^*)}{g(\theta_l^* | \theta_l^t)}, 1 \right\} \\ \theta_l^t & \text{else} \end{cases}$$

avec

$$\begin{aligned} \frac{f(\theta_l^*)}{f(\theta_l^t)} \frac{g(\theta_l^t | \theta_l^*)}{g(\theta_l^* | \theta_l^t)} &= \frac{\theta_l^*}{\theta_l^t} \frac{\theta_l^{* - \frac{N}{\theta_l^*} - a - 1}}{\theta_l^{t - \frac{N}{\theta_l^t} - a - 1}} \left[\frac{\Gamma \left(\frac{1}{\theta_l^*} \right)}{\Gamma \left(\frac{1}{\theta_l^t} \right)} \right]^N \exp \left[-\frac{\left(\sum_{n=1}^N \frac{y_{nl}}{s_{nl}} \right) - b}{\theta_l^*} + \frac{\left(\sum_{n=1}^N \frac{y_{nl}}{s_{nl}} \right) + b}{\theta_l^t} \right] \\ &\times \frac{\left(\prod_{n=1}^N \frac{1}{s_{nl}} \right)^{\frac{1}{\theta_l^*}} \left(\prod_{n=1}^N y_{nl} \right)^{\frac{1}{\theta_l^*} - 1}}{\left(\prod_{n=1}^N \frac{1}{s_{nl}} \right)^{\frac{1}{\theta_l^t}} \left(\prod_{n=1}^N y_{nl} \right)^{\frac{1}{\theta_l^t} - 1}} \left[\frac{1 - \Phi \left(\frac{-\theta_l^t}{\epsilon_1} \right)}{1 - \Phi \left(\frac{-\theta_l^*}{\epsilon_1} \right)} \right]. \end{aligned}$$

E.3 Echantillonnage du paramètre μ_k

On obtient la loi conditionnelle suivante

$$f(\mu_k | s_{nl}, \sigma_k^2, \mu_0, \sigma_0) \propto \frac{\exp \left[\frac{-\sum_{n=1}^N \sum_{l=1}^{L_k} (s_{nl} - \mu_k)^2}{2\sigma_k^2} - \frac{(\mu_k - \mu_0)^2}{2\sigma_0^2} \right]}{\left(1 - \Phi \left(-\frac{\mu_k}{\sigma_k} \right) \right)^{NL_k}} \quad (\text{E.3})$$

En utilisant la loi de proposition (gaussienne [Chari et al., 2016]) suivante

$$g(x | \mu_k^t) \propto \frac{1}{\sqrt{2\pi} \epsilon_3} \exp\left[-\frac{(x - \mu_k^t)^2}{2 \epsilon_3^2}\right].$$

On obtient la règle d'acceptation/rejet suivante

$$\mu_k^{t+1} = \begin{cases} \mu_k^* & \text{with prob } \min\left\{\frac{f(\mu_k^*)}{f(\mu_k^t)}, \frac{g(\mu_k^t | \mu_k^*)}{g(\mu_k^* | \mu_k^t)}, 1\right\} \\ \mu_k^t & \text{else} \end{cases}$$

avec

$$\begin{aligned} \frac{f(\mu_k^*)}{f(\mu_k^t)} \frac{g(\mu_k^t | \mu_k^*)}{g(\mu_k^* | \mu_k^t)} &= \left(\frac{1 - \Phi\left(-\frac{\mu_k^t}{\sigma_k}\right)}{1 - \Phi\left(-\frac{\mu_k^*}{\sigma_k}\right)} \right)^{NL_k} \exp\left[\frac{-\sum_{n=1}^N \sum_{l=1}^{L_k} (s_{nl} - \mu_k^*)^2}{2 \sigma_k^2} - \frac{(\mu_k^* - \mu_0)^2}{2 \sigma_0^2} \right] \\ &\times \exp\left[\frac{\sum_{n=1}^N \sum_{l=1}^{L_k} (s_{nl} - \mu_k^t)^2}{2 \sigma_k^2} + \frac{(\mu_k^t - \mu_0)^2}{2 \sigma_0^2} \right]. \end{aligned}$$

E.4 Echantillonnage du paramètre σ_k^2

On obtient la loi conditionnelle suivante

$$f(\sigma_k^2 | s_{nl}, \mu_k, \alpha_0, \beta_0) \propto \frac{\left(\frac{1}{\sigma_k^2}\right)^{\frac{NL_k}{2} + \alpha_0 + 1} \exp\left[\frac{-\sum_{n=1}^N \sum_{l=1}^{L_k} (s_{nl} - \mu_k)^2}{2 \sigma_k^2} - \frac{\beta_0}{\sigma_k^2} \right]}{\left(1 - \Phi\left(-\frac{\mu_k}{\sigma_k}\right)\right)^{NL_k}} \quad (\text{E.4})$$

En utilisant la loi de proposition (gaussienne tronquée positive [Chari et al., 2015]) suivante

$$g(x | (\sigma_k^2)^t) \propto \frac{\frac{1}{\sqrt{2\pi} \epsilon_4} \exp\left[-\frac{(x - (\sigma_k^2)^t)^2}{2 \epsilon_4^2}\right]}{1 - \Phi\left(-\frac{(\sigma_k^2)^t}{\epsilon_4}\right)}.$$

On obtient la règle d'acceptation/rejet suivante

$$(\sigma_k^2)^{t+1} = \begin{cases} \sigma_k^{2*} & \text{with prob } \min \left\{ \frac{f(\sigma_k^{2*})}{f((\sigma_k^2)^t)} \frac{g((\sigma_k^2)^t | \sigma_k^{2*})}{g(\sigma_k^{2*} | (\sigma_k^2)^t)}, 1 \right\} \\ (\sigma_k^2)^t & \text{else} \end{cases}$$

avec

$$\begin{aligned} \frac{f(\sigma_k^{2*})}{f((\sigma_k^2)^t)} \frac{g((\sigma_k^2)^t | \sigma_k^{2*})}{g(\sigma_k^{2*} | (\sigma_k^2)^t)} &= \left(\frac{1 - \Phi\left(-\frac{\mu_k}{\sigma_k^t}\right)}{1 - \Phi\left(-\frac{\mu_k}{\sqrt{\sigma_k^{2*}}}\right)} \right)^{NL_k} \left[\frac{1 - \Phi\left(\frac{-(\sigma_k^2)^t}{\epsilon_3}\right)}{1 - \Phi\left(\frac{-\sigma_k^{2*}}{\epsilon_3}\right)} \right] \times \exp \left[\frac{\sum_{n=1}^N \sum_{l=1}^{L_k} (s_{nl} - \mu_k)^2}{2 (\sigma_k^2)^t} + \frac{\beta_0}{(\sigma_k^2)^t} \right] \\ &\times \left(\frac{(\sigma_k^2)^t}{\sigma_k^{2*}} \right)^{\frac{NL_k}{2} + \alpha_0 + 1} \exp \left[\frac{-\sum_{n=1}^N \sum_{l=1}^{L_k} (s_{nl} - \mu_k)^2}{2 x^t} - \frac{\beta_0}{\sigma_k^{2*}} \right]. \end{aligned}$$

E.5 Echantillonnage du paramètre z_l

On a

$$P(z_l = k | s_{nl}, \sigma_k^2, \mu_k) \propto \frac{\frac{1}{(\sqrt{2\pi}\sigma_k)^N} \exp\left[-\sum_{n=1}^N \frac{(s_{nl} - \mu_k)^2}{2\sigma_k^2}\right]}{\left[1 - \Phi\left(\frac{-\mu_k}{\sigma_k}\right)\right]^N} \quad (\text{E.5})$$

On doit d'abord calculer les proportions associées à chaque valeur de k en utilisant E.5, et leurs associer des intervalles dont la longueur est proportionnelle à ces valeurs dans l'intervalle $[0,1]$.

On génère après une variable uniforme comme suit

$$u \sim U[0, 1]$$

Et enfin on met à jour z_l par la valeur de k associée à l'intervalle sélectionné par u .

ANNEXE F

Échantillonnage des paramètres du modèle inverse gamma

F.1 Formulation du problème

F.1.1 Modèle d'observation

Considérons L images non bruitées, contenant N pixels, rassemblées dans une matrice $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_L] \in \mathbb{R}^{N \times L}$, où \mathbf{s}_l , $l \in \{1, \dots, L\}$ désigne l'image qui correspond au $l^{\text{ième}}$ patient. Désignons par $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_L] \in \mathbb{R}^{N \times L}$ les images bruitées correspondantes.

En utilisant ces notations, le modèle d'observation est donné par :

$$\mathbf{y}_l = \mathbf{s}_l \odot \mathbf{b}_l, \text{ avec } \mathbf{b}_l \sim \mathcal{G}(\rho_l, \theta_l) \quad (\text{F.1})$$

où \mathbf{y}_l et \mathbf{s}_l sont des vecteurs ($N \times 1$) représentant les $l^{\text{ième}}$ images observées et non bruitées, \mathbf{b}_l est un vecteur ($N \times 1$) représentant le bruit multiplicatif de la loi gamma avec un paramètre de forme ρ_l et un paramètre d'échelle θ_l , et où \odot désigne le produit terme à terme. Afin de garantir que le modèle proposé (F.1) soit identifiable, la moyenne du bruit gamma est supposée égale à 1, ce qui conduit à :

$$\mathbb{E}(\mathbf{b}_l) = 1 \quad \Rightarrow \quad \rho_l = \frac{1}{\theta_l}. \quad (\text{F.2})$$

Le problème traité dans ce chapitre est de classer ces images \mathbf{y}_l , $l \in \{1, \dots, L\}$ en deux classes représentant les patients sains et ceux atteints de lentigo. La Section suivante présente un modèle bayésien hiérarchique qui est utilisé pour cette classification.

F.2 Modèle Bayésien Hiérarchique

Cette Section présente un modèle bayésien hiérarchique qui peut être utilisé pour estimer les $N \times L$ images de \mathbf{S} , les $L \times 1$ vecteurs $(\mathbf{z}, \boldsymbol{\theta})$ contenant les étiquettes associées aux classes des éléments de \mathbf{S}

ainsi que les variances de bruit associées aux L images observées à partir de la matrice \mathbf{Y} . Ce modèle est défini par la vraisemblance, et par les paramètres et hyperparamètres a priori définis ci-dessous.

F.2.1 Vraisemblance

Le bruit *speckle* multiplicatif \mathbf{b}_l est connu pour avoir une distribution gamma. Ainsi, le modèle d'observation (F.1) conduit à

$$y_{nl}|s_{nl}, \theta_l \sim \mathcal{G}\left(\frac{1}{\theta_l}, s_{nl} \theta_l\right) \quad (\text{F.3})$$

où \sim signifie “ est distribué selon”, \mathcal{G} est la distribution gamma dont la densité de probabilité (pdf) est définie par

$$f(y_{nl} | s_{nl}, \theta_l) \propto \frac{(y_{nl})^{\frac{1}{\theta_l}-1} \exp\left(-\frac{y_{nl}}{s_{nl} \theta_l}\right)}{\Gamma\left(\frac{1}{\theta_l}\right) (s_{nl} \theta_l)^{\frac{1}{\theta_l}}} I_{\mathbb{R}^+}(y_{nl}) \quad (\text{F.4})$$

avec $I_{\mathbb{R}^+}(y_{nl})$ la fonction indicatrice sur \mathbb{R}^+ , \propto signifie “proportionnel à” et Γ représente la fonction gamma. En supposant l'indépendance entre les signaux observés, la vraisemblance des L images observées peut s'écrire

$$f(\mathbf{Y}|\mathbf{S}, \boldsymbol{\theta}) \propto \prod_{n=1}^N \prod_{l=1}^L f(y_{nl}|s_{nl}, \theta_l).$$

L = nombre de patient, N = nombre de pixels.

F.2.2 Distribution a priori du signal d'intérêt

Afin d'assurer la positivité des images débruitées, une loi conjuguée inverse gamma est affectée à \mathbf{s}_l pour $l \in \{1, \dots, L\}$

$$\mathbf{s}_l | z_l = k, \alpha_k, \beta_k \sim \mathcal{IG}(\alpha_k, \beta_k) \quad (\text{F.5})$$

où (α_k, β_k) sont les paramètres de forme et d'échelle des deux distributions inverse gamma, k prenant les deux valeurs 1 et 2 selon la classe du patient.

F.2.3 Distribution a priori des variances du bruit

Une loi conjuguée inverse gamma non informative (dénommé \mathcal{IG}) est généralement choisie comme densité a priori pour le paramètre d'échelle θ_l [Fink, 1997]

$$\theta_l | a, b \sim \mathcal{IG}(a, b) \quad (\text{F.6})$$

où a and b sont les hyperparamètres fixés, qui sont ajustés pour tenir compte de l'absence d'information sur θ_l , i.e., les moyenne et variance de θ_l sont fixées à 1 et 100 afin d'avoir une loi non informative. La loi conjointe a priori pour le vecteur des variances de bruit $f(\boldsymbol{\theta} \mid a, b)$ est finalement obtenue à partir du produit des densités marginales $f(\theta_i \mid a, b)$.

F.2.4 Distribution a priori du vecteur d'étiquette \mathbf{z}

Le vecteur de paramètre $\mathbf{z} = (z_1, \dots, z_L)^T$ est un vecteur d'étiquettes qui associe chaque image à une peau saine ou lentigo. A cause de l'absence d'information a priori concernant ce paramètre, une loi a priori uniforme lui est affectée et est définie par :

$$P(z_l = k) = \frac{1}{2}, \forall l = 1, \dots, L. \quad (\text{F.7})$$

Les étiquettes associées à différents patients étant supposées indépendantes a priori, i.e., la loi conjointe a priori de \mathbf{z} notée par $f(\mathbf{z})$ est le produit des probabilités définies en (F.7).

F.2.5 Lois a priori des hyperparamètres

Afin de compléter la description du modèle hiérarchique bayésien proposé et d'estimer directement les hyperparamètres à partir des données, on propose de définir une loi a priori pour l'hyperparamètre β_k . On a ainsi sélectionné une loi a priori conjuguée gamma pour le paramètre d'échelle β_k (voir [Fink, 1997] pour motivations)

$$\beta_k \mid e, v \sim \mathcal{G}(e, v) \quad (\text{F.8})$$

où e et v sont fixés afin d'obtenir des lois non informatives, i.e., les moyennes et variances de β_k sont fixées à 1 et 1000. La densité de probabilité conjointe $f(\boldsymbol{\beta} \mid e, v)$ est finalement obtenue comme le produit de leur densité marginale en supposant l'indépendance des a priori entre les composante de ces vecteur.

F.2.6 Distribution conjointe *a posteriori*

Le modèle Bayésien proposé est illustré par le graphe orienté acyclique (DAG) représenté en Figure. F.1, et qui met en évidence la relation entre les observations \mathbf{Y} , les paramètres $\boldsymbol{\theta}, \mathbf{z}$, et les hyperparamètres α, β_k .

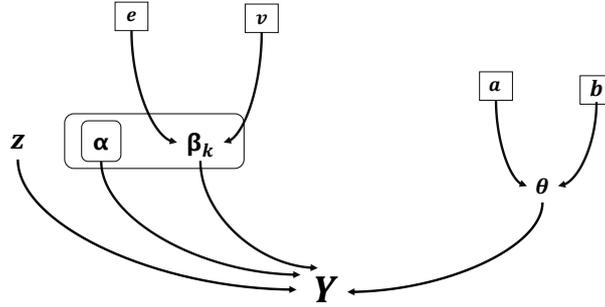


Figure F.1: DAG pour les lois a priori des paramètres et hyperparamètres. Les hyperparamètres fixés par l'utilisateur apparaissent dans des boîtes (ligne continue).

En supposant l'indépendance des a priori entre les différentes composantes du vecteur de paramètre $\mathbf{X} = (\boldsymbol{\theta}, \mathbf{z}, \alpha, \beta_k)$, la distribution conjointe *a posteriori* de ce modèle Bayésien peut être calculée en utilisant la structure hiérarchique suivante

$$f(\mathbf{X} | \mathbf{Y}) \propto f(\mathbf{Y} | \mathbf{S}, \boldsymbol{\theta}) f(\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \alpha, \beta) \quad (\text{F.9})$$

avec

$$f(\mathbf{S}, \boldsymbol{\theta}, \mathbf{z}, \alpha, \beta) = f(\mathbf{S} | \mathbf{z}, \alpha, \beta) f(\boldsymbol{\theta} | a, b) f(\beta | e, v) f(\mathbf{z}). \quad (\text{F.10})$$

Distribution conjointe *a posteriori* marginalisée

La distribution conjointe *a posteriori* peut être marginalisée par rapport à \mathbf{S} comme suit:

$$\begin{aligned} f(\boldsymbol{\theta}, \mathbf{z}, \alpha, \beta | \mathbf{y}) &\propto \int f(\mathbf{y} | \boldsymbol{\theta}, \mathbf{S}) f(\mathbf{S} | \mathbf{z}, \alpha, \beta) d\mathbf{S} f(\boldsymbol{\theta} | a, b) f(\alpha | c, d, t) f(\beta | e, v) f(\mathbf{z}) \\ &\propto f(\mathbf{y} | \boldsymbol{\theta}, \mathbf{z}, \alpha, \beta) f(\boldsymbol{\theta} | a, b) f(\alpha | c, d, t) f(\beta | e, v) f(\mathbf{z}) \end{aligned} \quad (\text{F.11})$$

où

$$y_{nl} | \theta_l, z_l = k, \alpha, \beta_k \sim \prod_{l=1}^L \prod_{n=1}^N \left[\frac{(y_{nl})^{\frac{1}{\theta_l} - 1}}{\left(\beta_k + \frac{y_{nl}}{\theta_l}\right)^{\alpha_k + \frac{1}{\theta_l}}} \frac{(\beta_k)^{\alpha_k}}{(\theta_l)^{\frac{1}{\theta_l}}} \frac{\Gamma\left(\alpha_k + \frac{1}{\theta_l}\right)}{\Gamma\left(\frac{1}{\theta_l}\right) \Gamma(\alpha_k)} \right]$$

La complexité du modèle bayésien proposé est illustré dans le DAG de la Figure. F.1 et sa distribution *a posteriori* résultante (F.11) rendent difficile le calcul explicite des estimateurs Bayésiens

(MAP ou MMSE) des paramètres du modèle. La Section suivante étudie une méthode MCMC qui sert à échantillonner la loi *a posteriori* (F.11) et à construire des estimateurs des paramètres intervenant dans le modèle bayésien proposé en utilisant les échantillons générés.

F.3 Algorithme de Metropolis-within-Gibbs

Cette Section étudie un échantillonneur de Gibbs hybride qui génère des échantillons asymptotiquement distribués suivant la loi cible (F.11). L'échantillonneur de Gibbs décrit dans Algo. 2, génère itérativement des échantillons distribués selon les lois conditionnelles de la loi conjointe définie dans (F.11). Ces lois conditionnelles sont détaillées dans le reste de cette Section. A cause de la complexité de ces lois conditionnelles, nous considérons un algorithme de Metropolis-Hastings à marche aléatoire (RWMH) [Chib and Greenberg, 1995, Gilks et al., 1995] dans l'échantillonneur de Gibbs et qui va consister à générer des échantillons distribués selon la loi conditionnelle complexe de chaque paramètre d'intérêt. Ceci est atteint en utilisant les lois conditionnelles $f_j(\cdot)$, pour $j \in \{1, \dots, J\}$, et leurs lois de proposition associées $g_j(\cdot)$. La première étape est d'initialiser la valeur de l'échantillon pour chaque paramètre $\mathbf{X}_j^{(0)}$, pour $j \in \{1, \dots, J\}$.

La boucle principale de l'algorithme RWMH se compose de trois parties:

1. Générer un candidat $\mathbf{X}_j^{\text{cand}}$ à partir de la loi de proposition $g_j(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})$. La distribution $g_j(\cdot)$ est la distribution gaussienne tronquée [Robert, 1995] $\mathcal{N}_{\mathbb{R}^+}(\mathbf{X}_j^{(i-1)}, \epsilon_j^2)$ pour les paramètres $\boldsymbol{\theta}, \boldsymbol{\beta}$,
2. Calculer la probabilité d'acceptation en utilisant la fonction d'acceptation $\alpha(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})$ basée sur la loi de proposition et la densité marginale pour chaque paramètre

$$\alpha(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)}) = \min \left\{ \frac{f_j(\mathbf{X}_j^{\text{cand}})}{f_j(\mathbf{X}_j^{(i-1)})} \frac{g_j(\mathbf{X}_j^{(i-1)} | \mathbf{X}_j^{\text{cand}})}{g_j(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})}, 1 \right\}$$

3. Accepter le candidat avec la probabilité $\alpha(\mathbf{X}_j^{\text{cand}} | \mathbf{X}_j^{(i-1)})$.

Afin de maximiser l'efficacité de l'algorithme, les variances ϵ_j^2 des lois de proposition ont été ajustées de telle sorte que le taux d'acceptation soit compris entre 0,3 et 0,6 comme suggéré dans [Chib and Greenberg, 1995] et [Roberts et al., 2001].

Algorithm 2 algorithme de Metropolis-within-Gibbs

```

1: Entrées:  $N_{bi}, N_{MC}, \boldsymbol{\theta}, \mathbf{z}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ 
2: Initialisation
3: Initialiser  $\boldsymbol{\theta}^{(0)}, \mathbf{z}^{(0)}, \boldsymbol{\alpha}^{(0)}, \boldsymbol{\beta}^{(0)}$ 
4: for  $i=1$  to  $N_{MC}$  do
5:   mise à jour des paramètres
6:   Échantillonner  $\boldsymbol{\theta}^{(i)} \mid \mathbf{Y}, a, b, \boldsymbol{\alpha}, \boldsymbol{\beta}$  selon (F.12) en utilisant l'algorithme RWMH avec la loi de proposition gaussienne tronquée
7:   Échantillonner  $\boldsymbol{\beta}^{(i)} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\alpha}, e, v$  selon (F.13) en utilisant l'algorithme RWMH avec la loi de proposition gaussienne tronquée
8:   Échantillonner  $\mathbf{z}^{(i)} \mid \mathbf{Y}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}$  à partir de la pdf (F.14)
9: end for
10: Résultat:  $\boldsymbol{\theta}^{(i)}, \mathbf{z}^{(i)}, \boldsymbol{\beta}^{(i)}$  for  $i = 1, \dots, N_{MC}$ .

```

Les densités conditionnelles des paramètres d'intérêt sont obtenues en multipliant la vraisemblance avec les différentes lois a priori et en enlevant le terme multiplicatif qui ne dépend pas de la variable d'intérêt. Les résultats obtenus sont détaillés dans les paragraphes suivants:

F.4 Échantillonnage du paramètre θ_l

On obtient la loi conditionnelle suivante

$$f(\theta_l \mid y_{nl}, a, b, \alpha, \beta_k) \propto \left(\prod_{n=1}^N y_{nl} \right)^{\frac{1}{\theta_l} - 1} \left[\frac{\Gamma\left(\alpha_k + \frac{1}{\theta_l}\right)}{\Gamma\left(\frac{1}{\theta_l}\right)} \right]^N \frac{\theta_l^{-\frac{N}{\theta_l} - a - 1} \exp\left(\frac{b}{\theta_l}\right)}{\left[\prod_{n=1}^N \left(\beta_k + \frac{y_{nl}}{\theta_l}\right) \right]^{\alpha_k + \frac{1}{\theta_l}}} \quad (\text{F.12})$$

F.5 Echantillonnage du paramètre β_k

On obtient la loi conditionnelle suivante

$$f(\beta_k | y_{nl}, \theta_l, \alpha, e, v) \sim \frac{\beta_k^{\alpha(N-L_k)-e-1} \exp\left(-\frac{\beta_k}{v}\right)}{\prod_{l=1}^{L_k} \prod_{n=1}^N \left[\left(\beta_k + \frac{y_{nl}}{\theta_l}\right)\right]^{\alpha + \frac{1}{\theta_l}}} \quad (\text{F.13})$$

F.6 Echantillonnage du paramètre z_l

On a

$$P(z_l = k | y_{nl}, \theta_l, \alpha, \beta_k) \propto \frac{1}{K} \frac{\prod_{n=1}^N (y_{nl})^{\frac{1}{\theta_l}-1}}{\left[\prod_{n=1}^N \left(\beta_k + \frac{y_{nl}}{\theta_l}\right)\right]^{\alpha + \frac{1}{\theta_l}}} \frac{(\beta_k)^{N\alpha}}{(\theta_l)^{\frac{N}{\theta_l}}} \left[\frac{\Gamma\left(\alpha + \frac{1}{\theta_l}\right)}{\Gamma\left(\frac{1}{\theta_l}\right) \Gamma(\alpha)} \right]^N \quad (\text{F.14})$$

On doit d'abord calculer les proportions associées à chaque valeur de k en utilisant F.14, et leurs associer des intervalles dont la longueur est proportionnelle à ces valeurs dans l'intervalle $[0,1]$.

On génère après une variable uniforme comme suit

$$u \sim U[0, 1]$$

Et enfin on met à jour z_l par la valeur de k associée à l'intervalle sélectionné par u .

F.7 Résultats de simulation

F.7.1 Données synthétiques

L'algorithme proposé a été testé sur des données synthétiques afin de valider son efficacité. Différentes expériences ont été réalisées et ce pour trois valeurs du rapport signal/bruit $\text{SNR} \in \{0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}\}$, ce qui permet d'apprécier la performance de l'algorithme pour différents niveaux de bruit. Cette Section considère $L = 100$ images synthétiques. Chaque image contient $N = 2000$ pixels et a été générée d'après (F.3). Ces images ont été séparées en classes saines et lentigo contenant 50 images. Les images non bruitées des deux classes ont été respectivement générées à l'aide de lois inverse gamma

$\mathcal{IG}(\alpha_1, \beta_1)$ et $\mathcal{IG}(\alpha_2, \beta_2)$, avec comme moyenne et variance égale à $\mu_1 = 17, \mu_2 = 20, \sigma_1^2 = 2, \sigma_2^2 = 4$. La convergence de l'échantillonneur de l'algorithme est surveillée en calculant le *potential scale reduction factor* introduit en 4.17 pour un paramètre d'intérêt. Plusieurs choix du paramètre k peuvent être considérés pour la méthode proposée. La convergence de l'échantillonneur Metropolis-within-Gibbs à été surveillée ici en calculant le *psrf* à l'aide du paramètre $\sqrt{\hat{\rho}}$ de la variance du bruit θ (voir [Dobigeon et al., 2008, Godsill and Rayner, 1998] pour un choix similaire). Le *potential scale reduction factor* du paramètre θ est calculé pour $M = 10$ chaines de Markov est égal à 1.07. Cette valeur de $\sqrt{\hat{\rho}}$ confirme la bonne convergence de l'échantillonneur (une recommandation pour l'évaluation de la convergence est une valeur de $\sqrt{\hat{\rho}} \leq 1.2$ [Gelman et al., 1995, p. 332]). Les Figures F.2, F.3, F.4 and F.5 montrent l'évolution des chaines de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés respectivement pour des données synthétiques avec $\text{SNR}_Y = [0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}]$ et des données réelles à partir des images MCR.

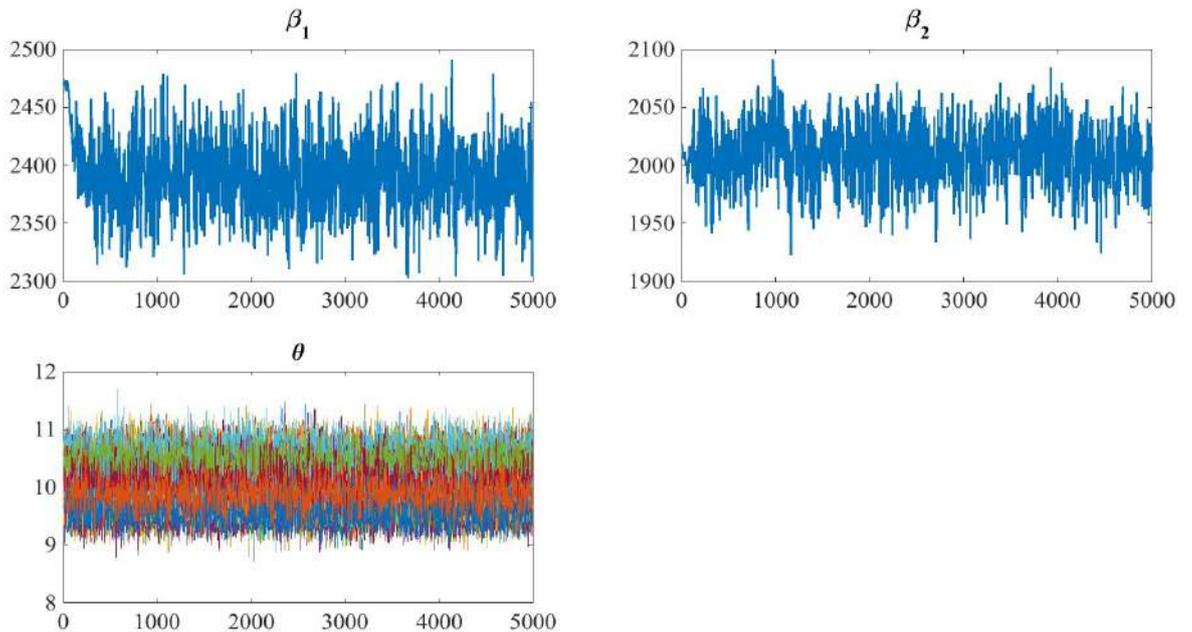


Figure F.2: Evolution de la convergence des chaines de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_Y = 0 \text{ dB}$.

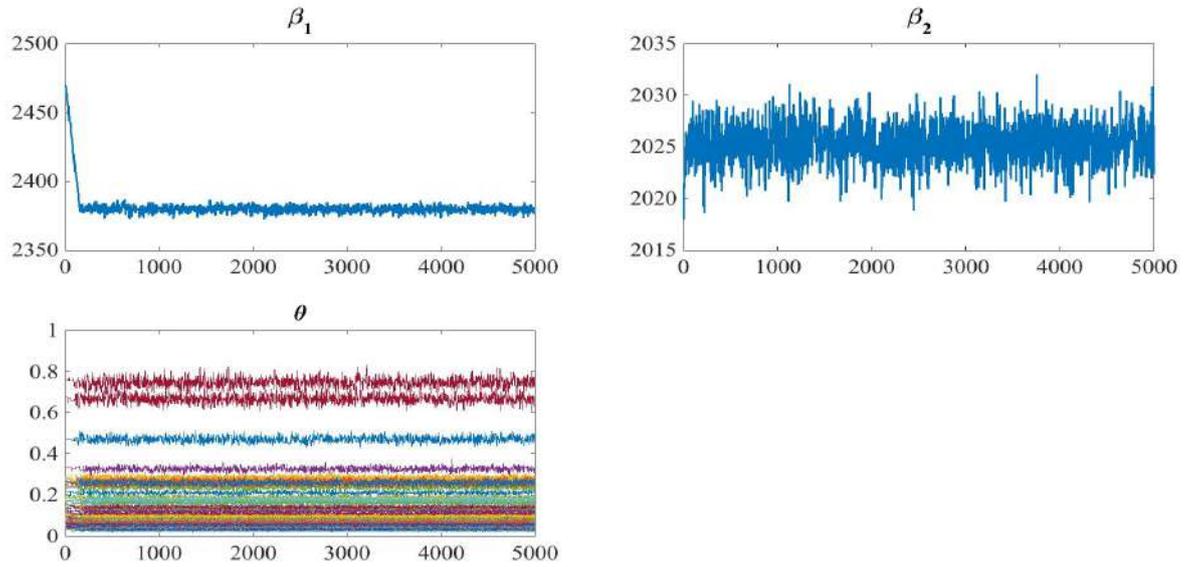


Figure F.3: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_\gamma = 10$ dB.

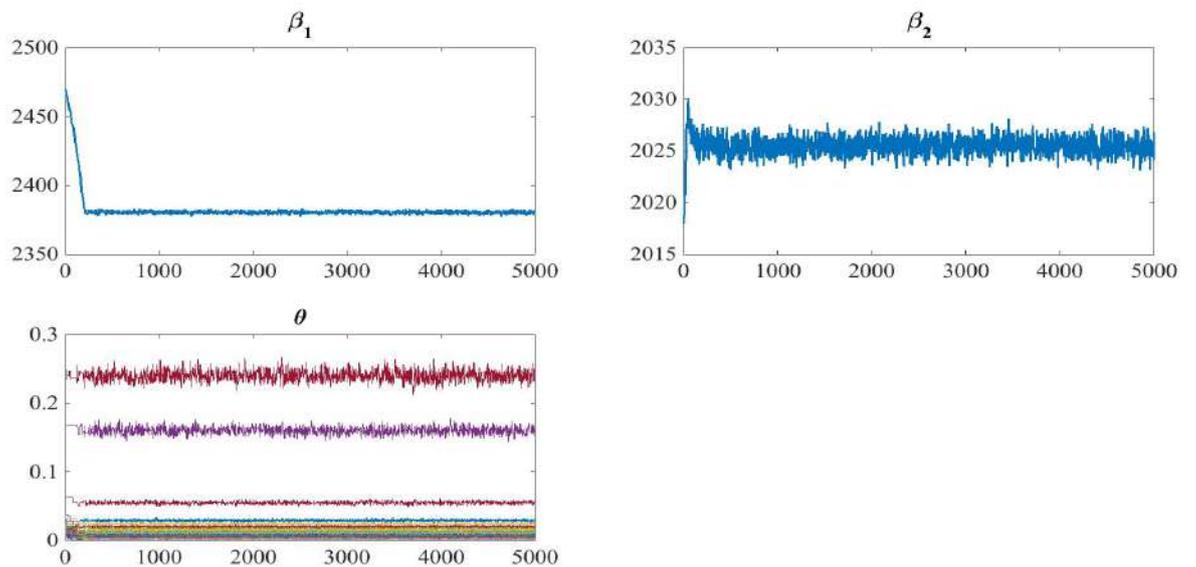


Figure F.4: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des données synthétiques avec $\text{SNR}_\gamma = 20$ dB.

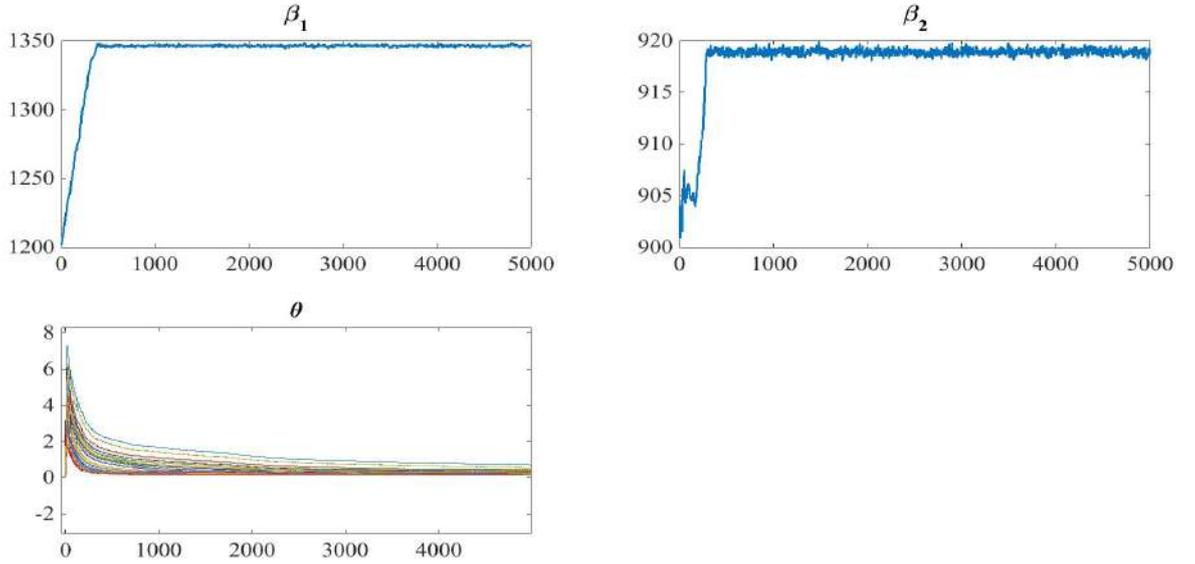


Figure F.5: Evolution de la convergence des chaînes de Markov pour les différents paramètres $\hat{\beta}_1, \hat{\beta}_2, \hat{\theta}$ estimés pour des images MCR réelles.

Algo. 1 a été exécuté pour $N_{MC} = 100000$ itérations et les différents paramètres du modèle ont été estimés d'après (4.12) et (4.13) en utilisant une période de chauffe de longueur $N_{bi} = 99900$. La performance de l'algorithme a été évaluée en calculant les rapports signal à bruit (SNRs) des différents paramètres définis par:

$$\text{SNR}_j = 20 \log_{10} \left(\frac{\|\mathbf{X}_j\|}{\|\mathbf{X}_j - \widehat{\mathbf{X}}_j\|} \right). \quad (\text{F.15})$$

Des résultats quantitatifs sont présentés dans le tableau F.1 pour les trois expériences. Ce tableau montre les bons résultats de classification et d'estimation des paramètres pour différents niveaux de bruit.

F.7.2 Données réelles

Cette Section est consacrée à la validation de l'algorithme de débruitage et de classification lorsqu'il est appliqué à des images MCR réelles déjà définies dans 2.3.1, mais cette fois-ci en utilisant comme profondeur caractéristique (qui se trouve être égale à $54 \mu m$ à partir des résultats obtenus dans les

Tableau F.1: Performance de l'algorithme proposé pour le débruitage et la classification des données synthétiques bruitées pour trois valeurs du signal/bruit $\text{SNR}_Y = [0 \text{ dB}, 10 \text{ dB}, 20 \text{ dB}]$.

	$\text{SNR}_Y = 0 \text{ dB}$	$\text{SNR}_Y = 10 \text{ dB}$	$\text{SNR}_Y = 20 \text{ dB}$
	SNR (dB)	SNR (dB)	SNR (dB)
β_1	18.0219	28.3990	28.4679
β_2	28.5427	41.8416	51.4867
θ	23.3134	25.1243	30.1780
Accuracy	88%	100%	100 %
Accuracy (CART)	83%	100%	100%

Tableau F.2: Performance de la classification sur des données réelles (45 patients) en utilisant la méthode proposée dans cette annexe.

Matrice de confusion	$\hat{\mathbf{L}}$	$\hat{\mathbf{H}}$	Sensibilité Spécificité
Lentigo	24	3	88.8 %
Sain	3	15	83.3 %
Précision	88.8 %	83.3 %	
Accuracy	86.6 %		

chapitres 2 et 3) celle qui permet de distinguer avec une différence maximale les images des patients sains de ceux atteints de lentigo. Compte tenu de la grande taille de ces images 1000×1000 pixels, nous avons préféré sélectionner et appliquer notre algorithme à des parcelles de 250×250 pixels pour chaque image afin de réduire le coût de calcul. Les résultats obtenus ont ensuite été utilisés pour calculer la matrice de confusion et les quatre indicateurs (sensibilité, spécificité, précision, accuracy) présentés dans les tableaux F.2 et F.3. Ces tableaux nous permettent d'évaluer les performances de classification de la stratégie proposée. Le paramètre "accuracy" de la méthode proposée ici est égale à 86.6%. Cette algorithme a été comparé à la méthode CART déjà présentée dans 3.3.4 et à notre méthode proposé dans le chapitre 4. Comme cela est mentionné dans les Tableaux F.3,F.4, le paramètre "accuracy" obtenu avec l'algorithme CART est égal à 82.2% , i.e., il est légèrement inférieur à celui obtenu avec la méthode qui est proposée dans cette annexe, par contre l'accuracy obtenu avec notre méthode proposée dans 4 est largement supérieur à celui obtenu avec la méthode utilisée ici.

Tableau F.3: Performance de la classification sur des données réelles (45 patients) en utilisant la méthode CART.

Matrice de Confusion	\hat{L}	\hat{H}	Sensibilité Specificité
Lentigo	24	3	88.8 %
Sain	5	13	72.2 %
Précision	82.7 %	81.2 %	
Accuracy	82.2 %		

Tableau F.4: Performance de la classification sur des données réelles (45 patients) en utilisant la méthode proposée dans le chapitre 4.

Matrice de confusion	\hat{L}	\hat{H}	Sensibilité Specificité
Lentigo	26	1	96.2 %
Sain	0	18	100 %
Précision	100 %	94.7 %	
Accuracy	97.7 %		

ANNEXE G

Approche bayésienne non paramétrique

G.1 Formulation du problème

G.1.1 Modèle d'observation

Considérons L images bruitées, contenant N pixels, regroupées dans la matrice $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_L] \in \mathbb{R}^{N \times L}$, où $\mathbf{y}_l, l \in \{1, \dots, L\}$ représente l'image associée au l ème patient. En utilisant ces notations, le modèle d'observation est donné par:

$$\mathbf{y}_l \sim \mathcal{G}(\beta_l, \rho_l) \tag{G.1}$$

où \mathbf{y}_l est un vecteur ($N \times 1$) représentant la l ème image observée distribuée suivant une loi gamma de paramètre de forme β_l et un facteur d'échelle ρ_l . Le problème posé dans cette partie est de classer ces images $\mathbf{y}_l, l \in \{1, \dots, L\}$ en deux classes représentant les patients sains et ceux atteints de lentigo. La prochaine section introduit le modèle bayésien hiérarchique qui sera utilisé pour cette classification.

G.2 Application du modèle de mélange de processus de Dirichlet (DPMM)

Dans cette partie nous allons appliquer le modèle de mélange de processus de Dirichlet (DPMM) pour le problème de classification. Nous utilisons le processus de Dirichlet afin de modéliser le problème et faire la classification par la suite. Le processus de Dirichlet nous permet de ne pas fixer le nombre de composantes du mélange mais plutôt de donner un a-priori sur le nombre allant de 1 à l'infini.

G.2.1 Classification en utilisant DPMM

Nous supposons que le vecteur de paramètres $\mathbf{x}_l = (\beta_l, \rho_l)^T$, $l \in \{1, \dots, L\}$ est distribué selon un modèle de mélange de gaussiennes bivariées donné par l'équation (G.2)

$$p(\mathbf{x}_l | \pi_i, \mu_i, \Sigma_i) = \sum_{i=1}^M \pi_i F(\mathbf{x}_l | \mu_i, \Sigma_i) \quad (\text{G.2})$$

où M est le nombre de gaussiennes, $\pi_i, i = 1, \dots, M$ sont les poids du mélange, et $F(\mathbf{x}_l | \mu_i, \Sigma_i), i = 1, \dots, M$ représentent les densités gaussiennes de moyennes μ_i et de covariance Σ_i , comme suit:

$$F(\mathbf{x}_l | \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{|\Sigma_i|} |\Sigma_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_l - \mu_i)' \Sigma_i^{-1} (\mathbf{x}_l - \mu_i) \right\} \quad (\text{G.3})$$

où $|\cdot|$ représente l'opérateur déterminant. Nous supposons que la matrice de covariance $\Sigma_i = 1/\sigma \times I$ est diagonale et isotropique, par conséquent ce que nous ignorons est le facteur scalaire σ . Les poids du mélange satisfont la contrainte $\sum_1^M \pi_i = 1$. Le modèle complet du mélange gaussien est paramétrisé par les vecteurs moyens, les matrices de covariance et les poids du mélange comme ceci:

$$\theta = \{\pi_i, \mu_i, \Sigma_i\}, \quad i = 1, \dots, M$$

G.2.2 Prior pour le vecteur d'étiquette label \mathbf{z}

Le vecteur d'étiquette $\mathbf{z} = (z_1, \dots, z_L)$ associe chaque image à une peau saine ou atteinte de lentigo. Nous proposons d'utiliser le processus du restaurant chinois (CRP) comme a priori de ce paramètre:

$$P(z_{n+1} = k | z_1, z_2, \dots, z_n) = \frac{\alpha}{\alpha + n} \delta(K+1) + \sum_{k=1}^K \frac{n_k}{\alpha + n} \delta(k) \quad (\text{G.4})$$

où

$$P(z_{n+1} = k | z_1, z_2, \dots, z_n) = \left\{ \begin{array}{l} \frac{n_k}{\alpha + n} \quad \text{if } k \leq K \\ \text{(i.e, k est la classe occupée précédemment)} \\ \frac{\alpha}{\alpha + n} \quad \text{otherwise} \\ \text{(i.e, k est la classe suivante inoccupée)} \end{array} \right\}$$

où n_k représente le nombre de patients attribué à la classe k .

G.2.3 Distribution a priori des paramètres moyennes et variances

La vraisemblance $F(x|\mu_i, \Sigma_i)$ étant gaussienne, nous choisissons une distribution Normal-Gamma comme a priori conjuguée $G_0(\theta)$ pour les moyennes et variances.

Étant donné les hyperparamètres $\lambda, \beta, \tau > 0$ et ν , la loi a-priori conjuguée $G_0(.,.)$ est définie par

$$G_0(\mu, \sigma | \lambda, \beta, \tau, \nu) = \begin{cases} \frac{\sigma^{\lambda-1} \exp(-\frac{\sigma}{\beta})}{\Gamma(\lambda) \beta^\alpha} \left(\frac{\sigma\tau}{2\pi}\right)^{1/2} \exp(-\frac{\sigma\tau}{2}(\mu - \nu)^2) & \text{où } \sigma > 0 \\ 0 & \text{sinon} \end{cases}. \quad (\text{G.5})$$

Plus précisément, la loi marginale de a-priori du paramètre de précision σ suit une loi gamma

$$f(\sigma | \lambda, \beta) = \begin{cases} \frac{\sigma^{\lambda-1} \exp(-\frac{\sigma}{\beta})}{\Gamma(\lambda) \beta^\alpha} & \text{où } \sigma > 0 \\ 0 & \text{sinon} \end{cases}$$

et la loi marginale de l'a-priori du paramètre de la moyenne μ possède une distribution de Student-t multivariée avec 2λ degrés de liberté, location ν et précision $\lambda \beta \tau$:

$$f(\mu | \lambda, \beta, \tau, \nu) = \sqrt{\frac{\beta\tau}{2\pi}} \frac{\Gamma(\frac{2\lambda+1}{2})}{\Gamma(\lambda)} \left(1 + \frac{\tau\beta}{2}(\mu - \nu)^2\right)^{-\frac{2\lambda+1}{2}}$$

Nous pouvons donc définir notre modèle comme suit:

$$x_i | \mu_i, \sigma_i \sim N(\mu_i, \sigma_i I_2)$$

$$\mu_i, \sigma_i | G \sim G(\mu, \sigma)$$

$$G \sim DP(\alpha G_0(\mu, \sigma))$$

$$G_0 \sim NG(\mu, \sigma | \lambda, \beta, \tau, \nu)$$

où $DP(\alpha G_0(\mu, \sigma))$ est le processus de Dirichlet avec la base G_0 et la dispersion α , et G est une distribution aléatoire tirée du processus de Dirichlet.

Echantillonneur de Gibbs

Il y a deux techniques essentielles dans la littérature qui permettent de générer des échantillons distribués selon une loi a posteriori: l'algorithme de Metropolis-Hastings et l'échantillonneur de Gibbs. Nous suggérons d'utiliser l'échantillonneur de Gibbs afin de générer une suite $\theta^{(i)}$ distribuée selon la loi a posteriori $f(\theta|x)$. Pour cela, nous devons échantillonner les paramètres θ_k selon leurs lois conditionnelles comme présenté dans l'algorithme [3](#)

Algorithm 3 Echantillonnage de Gibbs pour le modèle de mélange de processus de Dirichlet

Etant donné $\alpha^{(t-1)}$, $\{\theta_k^{(t-1)}\}_{k=1}^K$ et $\{z_i^{(t-1)}\}_{i=1}^n$ de l'itération précédente, échantillonner un nouvel ensemble de $\{\theta_k^{(t)}\}_{k=1}^K$ et $\{z_i^{(t)}\}_{i=1}^n$ comme suit:

1. Poser $z = z^{(t-1)}$, $\alpha = \alpha^{(t-1)}$
2. Pour $i = 1, \dots, n$
 - a) Enlever la donnée x_i du cluster z_i vu que nous allons échantillonner un nouveau z_i pour x_i .
 - b) Si x_i est la seule donnée dans son cluster courant, ce cluster deviendra vide après l'étape (2.a). Ce cluster est donc enlevé, avec son paramètre, et K est diminué de 1.
 - c) Réarranger les indices des clusters de telle manière que $1, \dots, K$ soient actives (i.e ,pas vides)
 - d) Tirer un nouvel échantillon pour z_i des probabilités suivantes:

$$p(z_i = k, k \leq K) \propto \frac{n_{k,-i}}{n + \alpha - 1} F(x_i | \theta_k^{(t-1)}) \quad n_{k,-i} = \sum_{j \neq i} \delta(z_j - k)$$

$$p(z_i = K + 1) \propto \frac{\alpha}{n + \alpha - 1} \int F(x_i | \theta) G_0(\theta) d\theta$$

- e) Si $z_i = K + 1$, nous obtenons un nouveau cluster. Indexer ce cluster comme $K + 1$, échantillonner un nouveau paramètre du cluster ϕ_i de $H(\phi_i | x_i)$ défini comme:

$$H(\phi | x_i) = \frac{F(x_i | \phi) \times G_0(\phi)}{\int F(x_i | \phi) \times G_0(\phi) d\phi}$$

l'assigner à θ_{K+1} et augmenter K de 1.

3. Pour $k = 1, \dots, K$
Echantillonner le paramètre de cluster θ_k de la loi a posteriori suivante :

$$\theta_k^{(t)} \propto G_0(\theta_k | x) \text{ Likelihood } (x_k^{(t)} | \theta_k^{(t-1)})$$

4. Poser $z^{(t)} = z$
-

G.3 Résultats

G.3.1 Test pour différents mélanges de gaussiennes

Cette section évalue les performances de l'algorithme proposé sur des données synthétiques. L'algorithme 3 a été testé sur les données qui ont été générés selon le modèle suivant :

$$p(x_i|\pi, \theta) = \sum_{i=1}^M \pi_i F(x|\mu_i, \Sigma_i)$$

Afin d'évaluer l'algorithme proposé, différentes expérimentations ont été réalisées en utilisant $L = 200$ points synthétiques et trois mélanges de gaussiennes bidimensionnelles.

L'algorithme 3 a été appliqué en utilisant des valeurs fixes pour les hyperparamètres $\lambda = 5$, $\beta = 0.9$, $\tau = 0.05$, $\nu = 0$ et le paramètre de concentration $\alpha = 10^{-5}$. Les Figures G.1-G.3 montrent les résultats obtenus en variant le nombre du mélange de 3 à 5 gaussiennes avec différentes probabilités $p=[0.3 \ 0.2 \ 0.5]$ pour 3 gaussiennes, $p=[0.4 \ 0.1 \ 0.3 \ 0.2]$ pour 4 gaussiennes et $p=[0.4 \ 0.1 \ 0.3 \ 0.05 \ 0.15]$ pour 5 gaussiennes.

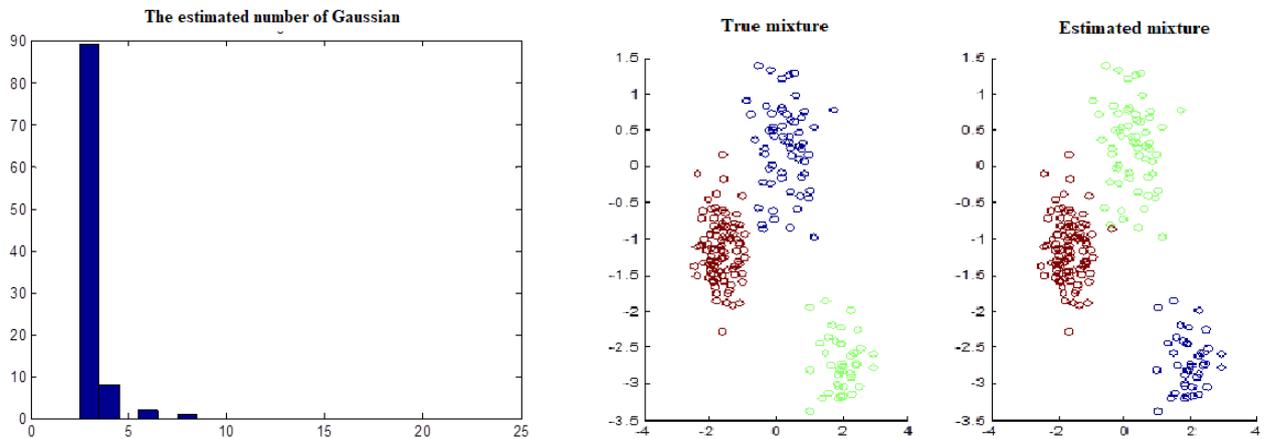


Figure G.1: (gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.

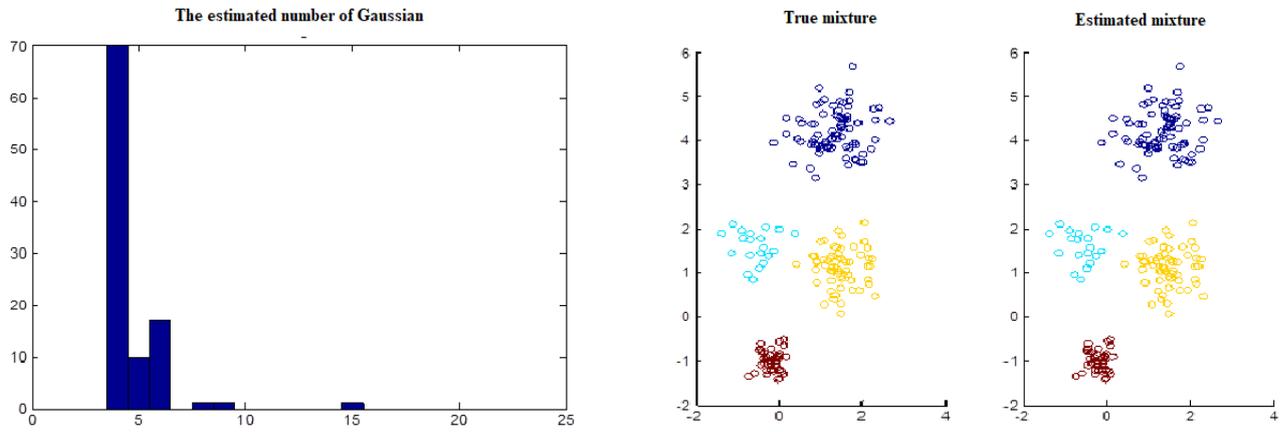


Figure G.2: (gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.

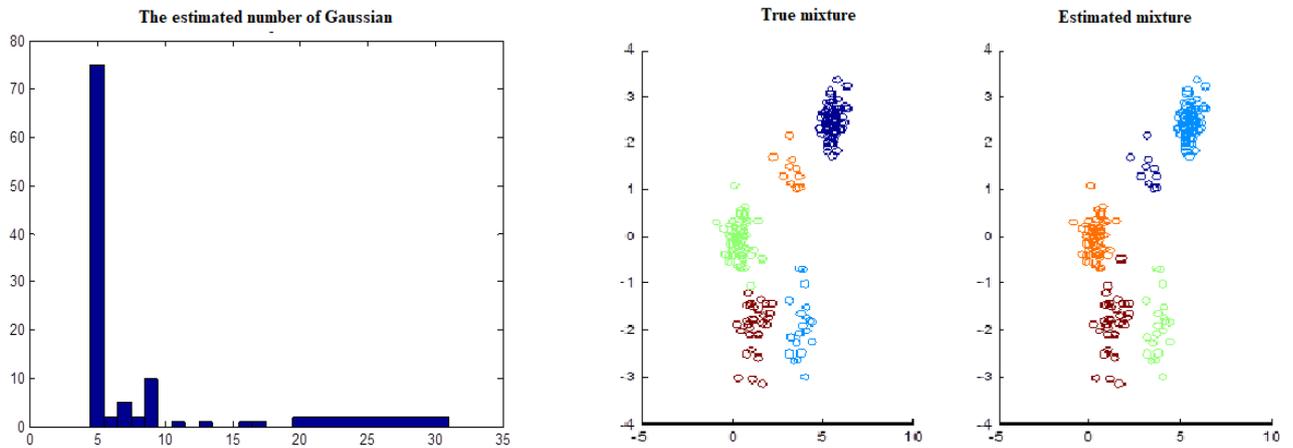


Figure G.3: (gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.

		Moyennes(μ_i)		précision(σ)		Poids du mélange(π_i)	
		$MSE^2(\mu)$	$SAD(\mu)$	$MSE^2(\sigma)$	$SAD(\sigma)$	$GMSE^2$	RMSE
Mélange de 3 gaussiennes bidimensionnelles	1	0.0018	0.1095	0.0024	0	1.0e-04*0.25	0.0041
	2	0.0155	0.01	0.0025	0	0	
	3	0.0025	0.0212	0.0013	0	1.0e-04*0.25	
Mélange de 4 gaussiennes bidimensionnelles	1	0.0001	0.0008	0.0006	0	0	0
	2	0.0689	0.0605	0.0001	0	0	
	3	0.0021	0.0086	0.0015	0	0	
	4	0.0007	0.0245	0.0003	0	0	
Mélange de 5 gaussiennes bidimensionnelles	1	0.00001	0.0001	0.0006	0	0.0576	0.1364
	2	0.0464	0.0462	0.0001	0	0.0036	
	3	0.0027	0.0635	0.0001	0	0.0196	
	4	0.0182	0.0103	0.0008	0	0.0121	
	5	0.0074	0.0070	0.0016	0	0.0001	

Tableau G.1: Performance de l'algorithme proposé pour la classification en utilisant différents nombres de gaussiennes dans le mélange

La Table 1 représente l'erreur quadratique moyenne (MSE^2), l'erreur quadratique moyenne globale (GMSE), la racine de l'erreur quadratique moyenne (RMSE) et la distance angulaire spectrale (SAD) calculé entre les vrais paramètres et les paramètres correspondants estimés.

$$MSE^2(\mu) = \|\hat{\mu}_i - \mu_i\|^2 \quad i = 1, \dots, M ; \quad MSE^2(\sigma) = \|\hat{\sigma}_i - \sigma_i\|^2 \quad i = 1, \dots, M$$

$$SAD(\mu) = \arccos\left(\frac{\langle \hat{\mu}_i, \mu_i \rangle}{\|\hat{\mu}_i\| \|\mu_i\|}\right) ; \quad SAD(\sigma) = \arccos\left(\frac{\langle \hat{\sigma}_i, \sigma_i \rangle}{\|\hat{\sigma}_i\| \|\sigma_i\|}\right)$$

$$GMSE^2(\pi) = (\hat{\pi}_i - \pi_i)^2 ; \quad RMSE(\pi) = \sqrt{(1/M) * \|\hat{\pi} - \pi\|^2},$$

où π est le vecteur de probabilités du mélange.

G.3.2 Données réelles

Cette section concerne la validation de l'algorithme proposé sur des images MCR réelles. Figure G.4 présente les résultats obtenus avec notre algorithme pour la classification des images MCR réelles. Ces résultats sont par la suite utilisés pour calculer la matrice de confusion et quatre indicateurs (sensitivité, spécificité, précision, accuracy) représentés dans la Table G.2. Cette table nous permet d'évaluer la performance de classification de la stratégie proposée. L'accuracy de la méthode proposée

est égale à 97.7%, ce qui correspond à une seule erreur pour le patient lentigo #8. Pour évaluer l'importance de nos résultats, notre algorithme a été comparé à la méthode CART présentée dans [Koller et al., 2011]. Cette méthode consiste à extraire pour chaque image MCR un ensemble de 39 paramètres d'analyse (de plus amples détails techniques sont disponibles dans [Wiltgen et al., 2008]) et d'appliquer à ces caractéristiques une procédure de classification basée sur la classification et l'arbre de régression (CART). On note que l'algorithme CART a été testé sur des images MCR réelles en utilisant la procédure leave one out. Comme indiqué dans la Table G.3, l'accuracy obtenue avec l'algorithme CART est égale à 82.2% ce qui montre la supériorité des performances de l'algorithme proposé.

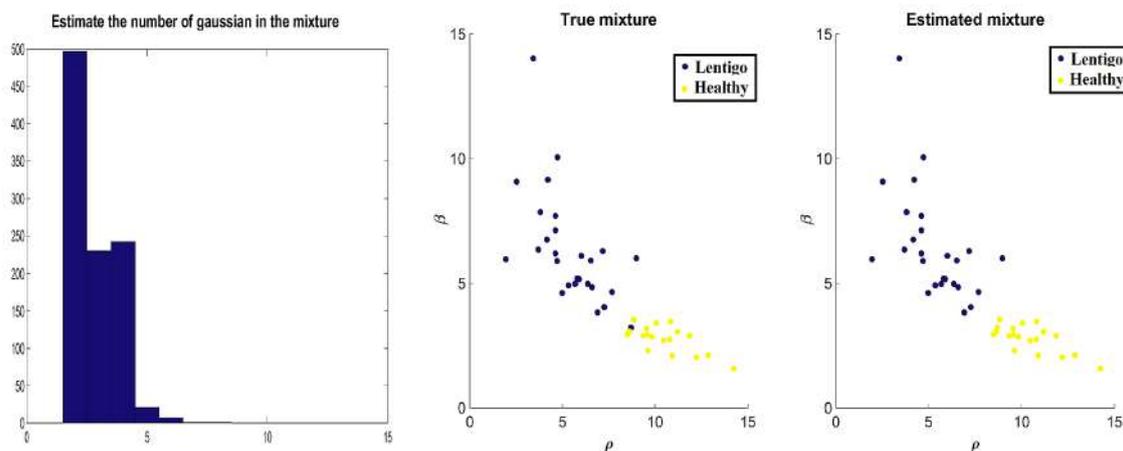


Figure G.4: (gauche) distribution a posteriori du nombre estimé des gaussiennes , (droite) le vrai mélange et son estimé.

Tableau G.2: Performance de classification sur des données réelles (45 patients) en utilisant la méthode proposée (pour la profondeur caractéristique $54 \mu m$).

Matrice de confusion	\hat{L}	\hat{H}	Sensitivité Spécificité
Lentigo	26	1	96.2 %
Sain	0	18	100 %
Précision	100 %	94.7 %	
Accuracy	97.7 %		

Tableau G.3: Performance de classification sur des données réelles (45 patients) en utilisant la méthode CART (pour la profondeur caractéristique $54 \mu m$).

Matrice de confusion	\hat{L}	\hat{H}	Sensitivité Spécificité
Lentigo	24	3	88.8 %
Sain	5	13	72.2 %
Précision	82.7 %	81.2 %	
Accuracy	82.2 %		

Bibliographie

- M. Abramowitz and I. A. Stegun. *Handbook of Mathematical functions with Formulas, Graphs and Mathematical Tables*. Dover Publications, 1970.
- I. Alarcon, C. Carrera, L. Alos, J. Palou, J. Malveyh, and S. Puig. In vivo reflectance confocal microscopy to monitor the response of lentigo maligna to imiquimod. *J Am Acad Dermatol.*, 71: 49–55, 2014a.
- I. Alarcon, C. Carrera, J. Palou, L. Alos, J. Malveyh, and S. Puig. Impact of in vivo reflectance confocal microscopy on the number needed to treat melanoma in doubtful lesions. *British journal of Dermatology.*, 170:802–808, 2014b.
- H. Alexander and D. L. Miller. Determining skin thickness with pulsed ultra sound. *The Journal of Investigate Dermatology*, 72(1):17–19, 1979.
- S. Amari and S. C. Douglas. Why natural gradient? In *IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, volume 2, pages 1213–1216, Seattle, WA, May 1998.
- R. A. Briggaman. Biochemical composition of the epidermal-dermal junction and other basement membrane. *The Journal of Investigate Dermatology*, 78(1):1–6, 1982.
- M. H. C. Prost-Squarcioni and S. Fraitag. Histologie et histophysiologie de la peau et de ses annexes. *Ann Dermatol. Venereol.*, 132:8S5–48, 2005.
- L. Chaari, J.-Y. Tומרet, and C. Chaux. Sparse signal recovery using a bernoulli generalized gaussian prior. In *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pages 1711–1715. IEEE, 2015.

- L. Chaari, J.-Y. Tourneret, C. Chaux, and H. Batatia. A hamiltonian monte carlo method for non-smooth energy sampling. *IEEE Transactions on Signal Processing*, 64(21):5585–5594, 2016.
- J. Champin, J.-L. Perrot, E. Cinotti, B. Labeille, C. Douchet, G. Parrau, F. Cambazard, P. Seguin, and T. Alix. In vivo reflectance confocal microscopy to optimize the spaghetti technique for defining surgical margins of lentigo maligna. *Dermatologic Surgery*, 40(3):247–256, 2014.
- C. Chaouat. *Conception de nouveaux systèmes de formulations d’actifs dépigmentants, en vue de leur utilisation par voie cutanée*. PhD thesis, Université Toulouse III, Paul Sabatier, France, 2011.
- S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):32–335, 1995.
- A. Cichocki and S. Amari. Families of alpha - beta - and gamma- divergences: Flexible and robust measures of similarities. *Entropy*, 12(6):1532–1568, 2010.
- N. A. C. Cressie and H. J. Whitford. How to use the two sample t-test. *Biometrical Journal*, 28(2):131–148, 1986.
- B. H. A. Decoupigny. *Le développement des technologies numériques d’imagerie médicale en france. Etude bibliographique et résultats d’une enquête auprès d’un échantillon de cliniques vétérinaires*. PhD thesis, Faculté de Médecine de Créteil, France, 2011.
- A. Dima, M. Scholz, and K. Obermayer. Automatic segmentation and skeletonization of neurons from confocal microscopy images based on the 3-d wavelet transform. *IEEE Trans. Image Process*, 11:790–801, 2002.
- P. M. Djurić and J.-H. Chun. An MCMC sampling approach to estimation of nonstationary hidden markov models. *IEEE Trans. Signal Processing*, 50(5):1113–1123, 2002.
- M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. Image Processing*, 11(2):146–158, 2002.

- N. Dobigeon, J.-Y. Tourneret, and C.-I Chang. Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery. *IEEE Trans. Signal Processing*, 56(7):2684–2695, July 2008.
- A. F. Fercher, C. K. Hitzenberger, W. Drexler, G. Kamp, and H. Sattmann. In vivo optical coherence tomography. *American Journal of Ophthalmology*, 116(1):113–114, July 1993.
- D. Fink. A compendium of conjugate priors. See [http://www. people. cornell. edu/pages/df36/CONJINTRnew% 20TEX. pdf](http://www.people.cornell.edu/pages/df36/CONJINTRnew%20TEX.pdf), page 46, 1997.
- C. Fouard. *Extraction de paramètres morphométriques pour l'étude du réseau micro-vasculaire cérébral*. PhD thesis, Université de Nice Sophia Antipolis, France, 2005.
- E. Fuchs and S. Raghavan. Getting under the skin of epidermal morphogenesis. *Nature Reviews. Genetics*, 3:199–209, March 2002.
- V. Gagnon. Etude des interactions entre les nerfs sensoriels et les follicules pileux dans un modèle in vitro de peau reconstruite par génie tissulaire. Master's thesis, Université Laval, Québec, 2005.
- A. Gelman and D. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–511, 1992.
- A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall, London, 1995.
- A. Geras. Dermatology: A medical artist's interpretation. *Sandoz Medical Publications*, 445, 1990.
- W. R. Gilks, S. Richardson, and D. Spiegelhalter. *Markov chain Monte Carlo in practice*. CRC press, 1995.
- K. Glunde, A. P. Pathak, and Z. M. Bhujwalla. Molecular-functional imaging of cancer: to image and imagine. *TRENDS in Molecular Medicine*, 13(7):287–297, 2007.

- S. J. Godsill and P. J. W. Rayner. Statistical reconstruction and analysis of autoregressive signals in impulsive noise using the Gibbs sampler. *IEEE Trans. Speech, Audio Processing*, 6(4):352–372, 1998.
- P. Guitera, L. Haydu, S. Menzies, R. Scolyer, A. Hong, G. Fogarty, F. Gallardo, and S. Segura. Surveillance for treatment failure of lentigo maligna with dermoscopy and in vivo confocal microscopy: new descriptors. *British journal of Dermatology.*, 170:1305–1312, 2014.
- S. R. Gunn et al. Support vector machines for classification and regression. *ISIS Technical Report*, 14:85–86, 1998.
- A. Halimi, C. Mailhes, J.-Y. Tourneret, and H. Snoussi. Bayesian estimation of smooth altimetric parameters: Application to conventional and delay/doppler altimetry. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4):2207–2219, 2016.
- A. Halimi, H. Batatia, J. Le Digabel, J. Gwendal, and J.-Y. Tourneret. An unsupervised Bayesian approach for the joint reconstruction and classification of cutaneous reflectance confocal microscopy images. In *25th European Signal Processing Conference (EUSIPCO)*, pages 251–255, Kos Island, Greece, 28 August 2017 EUSIPCO. URL <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2017/papers/1570347087.pdf>.
- A. Halimi, H. Batatia, J. Le Digabel, J. Gwendal, and J.-Y. Tourneret. Wavelet-based statistical classification of skin images acquired with reflectance confocal microscopy. *Biomedical optics express*, 2017,Biomedical.
- A. Halimi, H. Batatia, J. Le Digabel, J. Gwendal, and J.-Y. Tourneret. Statistical modeling and classification of reflectance confocal microscopy images. In *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pages , accepté, Curaçao, Dutch Antilles, December 2017,CAMSAP.
- A. Halimi, H. Batatia, J. Le Digabel, J. Gwendal, and J.-Y. Tourneret. An unsupervised Bayesian

- approach for the joint reconstruction and classification of reflectance confocal microscopy images. *IEEE Transactions on Biomedical Engineering*, *À soumettre*, IEEE BE.
- A. Halimi, H. Batatia, J. Le Digabel, J. Gwendal, and J.-Y. Tournieret. Statistical model-based classification of reflectance confocal microscopy images. *Signal image and video processing*, *À soumettre*, SIVP.
- S. C. Hames, M. Ardigo, H. P. Soyer, A. P. Bradley, and T. W. Prow. Anatomical skin segmentation in reflectance confocal microscopy with weak labels. In *Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on.*, pages 1–8. IEEE, 2015.
- S. C. Hames, M. Ardigo, H. P. Soyer, A. P. Bradley, and T. W. Prow. Automated segmentation of skin strata in reflectance confocal microscopy depth stacks. *PloS one*, 11(4):e0153208, 2016.
- M. A. Harris, A. N. Van, B. H. Malik, J. M. Jabbour, and K. C. Maitland. A pulse coupled neural network segmentation algorithm for reflectance confocal images of epithelial tissue. *PloS one*, 10(3):e0122368, 2015.
- B. P. Hibler, M. Cordova, R. J. Wong, and A. M. Rossi. Intraoperative real-time reflectance confocal microscopy for guiding surgical margins of lentigo maligna melanoma. *Dermatol Surg.*, 41:980–983, 2015.
- R. Hofmann-Wellenhof, E. M. Wurm, V. Ahlgrim-Siess, S. K. E. Richtig, J. Smolle, and A. Gerger. Reflectance confocal microscopy-state-of-art and research overview. *Seminars in Cutaneous Medicine and Surgery, Elsevier*, 28:172–179, 2009. doi: 10.1016/j.sder.2009.06.004.
- Y. Hu, A. Shen, T. Jiang, Y. Ai, and J. Hu. Classification of normal and malignant human gastric mucosa tissue with confocal raman microspectroscopy and wavelet analysis. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 69(2):378–382, 2008.
- D. Huang, E. A. Swanson, C. P. Lin, J. S. S. and W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. Puliafito, and J. G. Fujimoto. Optical coherence tomography. *Sciences, New Series*, 254(5035):1178–1181, November 1991.

- C. Hurlin. Maximum likelihood estimation. 2013. URL http://www.univ-orleans.fr/deg/masters/ESA/CH/Chapter2_MLE.pdf.
- H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. of the Royal Society of London. Series A*, 186(1007):453–461, 1946.
- N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, vol. 1*. Wiley Series in Probability and Statistics, 1994.
- V. E. Johnson. Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences*, 110(48):19313–19317, 2013.
- J. N. Jonkman, P. D. Gerard, and W. H. Swallow. Estimating probabilities under the three-parameter gamma distribution using composite sampling. *Computational Statistics & Data Analysis*, 53(4):1099–1109, February 2009.
- J. A. L. Juin. *Imagerie de la peau: actualités en médecine humaine et applications envisageables en médecine vétérinaire*. PhD thesis, Faculté de médecine de Créteil, France, 2014.
- J. Kanitakis, P. Bahadoran, R. Braun, S. Debarbieux, B. Labeille, J.-L. Perrot, and P. Vabres. La microscopie confocale par réflectance in vivo en dermatologie : proposition de terminologie en langue française. *Annales de dermatologie et de vénéréologie*, 140:678–686, 2013.
- S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- S. Koller, M. Wiltgen, V. Ahlgrimm-Siess, W. Weger, R. Hofmann-Wellenhof, E. Richtig, J. Smolle, and A. Gerger. In vivo reflectance confocal microscopy: automated diagnostic image analysis of melanocytic skin tumours. *Journal of the European Academy of Dermatology and Venereology*, 25(5):554–558, 2011.
- K. Kose, C. Alessi-Fox, M. Gill, J. G. Dy, D. H. Brooks, and M. Rajadhyaksha. A machine learning method for identifying morphological patterns in reflectance confocal microscopy mosaics of

- melanocytic skin lesions in-vivo. In *SPIE BiOS*, pages 968908–968908. International Society for Optics and Photonics, 2016.
- D. Kozłowska, P. Foran, P. MacMahon, M. J. Shelly, S. Eustace, and R. O. Kennedy. Molecular and magnetic resonance imaging: The value of immunoliposomes. *Advanced Drug Delivery Reviews*, 61: 1402–1411, 2009.
- S. Kurugol, J. G. Dy, M. Rajadhyaksha, K. W. Gossage, J. Weissmann, and D. H. Brooks. Semi-automated algorithm for localization of dermal/epidermal junction in reflectance confocal microscopy images of human skin. In *SPIE BiOS*, pages 79041A–79041A. International Society for Optics and Photonics, 2011.
- S. Kurugol, M. Rajadhyaksha, J. G. Dy, and D. H. Brooks. Validation study of automated dermal/epidermal junction localization algorithm in reflectance confocal microscopy images of skin. In *SPIE BiOS*, pages 820702–820702. International Society for Optics and Photonics, 2012.
- A. Laplante. *Mécanismes de réépithélialisation des plaies cutanées : expression des protéines de stress chez la souris et analyse à l'aide d'un nouveau modèle tridimensionnel humain développé par génie tissulaire*. PhD thesis, Université Laval, Québec, 2002.
- C. Laurent. *Rôle du stress dans le développement des effets cellulaires radio-induits au niveau cutané: application aux irradiations localisées accidentelles*. PhD thesis, l'Université de Versailles Saint-Quentin-en-Yvelines, France, 2005.
- B. L. Luck, K. D. Carlson, A. C. Bovik, and R. R. Richards-Kortum. An image model and segmentation algorithm for reflectance confocal images of in vivo cervical tissue. *IEEE Trans. Image Processing*, 14(9):1265–1276, Sept. 2005.
- S. MacNeil. Progress and opportunities for tissue-engineered skin. *Nature*, 445:874–880, February 2007.
- E. N. Marieb. *Anatomie et physiologie humaines*. De Boeck Université, 1999.

- A. Mcheik. *Modélisation statistique du Speckle en OCT : application à la segmentation d'images de la peau*. PhD thesis, Institut National Polytechnique de Toulouse, Toulouse, France, 2010.
- T. D. Menge, B. P. Hibler, M. A. Cordova, K. S. Nehal, and A. M. Rossi. Concordance of handheld reflectance confocal microscopy (rcm) with histopathology in the diagnosis of lentigo maligna (lm): A prospective study. *Journal of the American Academy of Dermatology*, 74(6):1114–1120, 2016.
- M. Minsky. Memoir on inventing the confocal scanning microscope. *Sanning*, 10:128–138, 1988. URL http://www.barnesreport.net/WFBZ/arcanum/confocal%20microscope_/Inventing%20the%20Confocal%20Scanning%20Microscope.pdf.
- A. Nahas. *Imagerie multimodale optique, élastique et photo-thermique des tissus biologiques par OCT plein champ*. PhD thesis, Université Pierre et Marie Curie, France, 2014.
- M. Naouri. Echographie cutanée haute résolution. *Images en Dermatologie*, 1:110–114, 2008. URL <http://www.edimark.fr/Front/frontpost/getfiles/14718.pdf>.
- K. S. Nehal, D. Gareau, and M. Rajadhyaksha. Skin imaging with reflectance confocal microscopy. *Seminars in Cutaneous Medicine and Surgery, Elsevier*, 27:37–43, 2008. doi: 10.1016/j.sder.2008.01.006.
- S. A. Osseni. *Nanoplateformes hybrides multimodales pour l'imagerie médicale*. PhD thesis, Université de Toulouse III - Paul Sabatier, France, 2010.
- M. Ouadour-abbbar. *Principe de tomographie et spectrotomographie optique de cohérence par intercorrélation sans balayage basée sur un réseau de diffraction*. PhD thesis, Université de Franche-Comté, France, 2009.
- E. Oualid. *Les produits dépigmentants : le point en 2011*. PhD thesis, Université Joseph Fourier, Faculté de Médecine de Grenoble, France, 2011.
- S. V. Patwardhan, A. P. Dhawan, and P. A. Relue. Classification of melanoma using tree structured wavelet transforms. *Computer Methods and Programs in Biomedicine*, 72(3):223–239, 2003.

- M. Pereyra, H. Batatia, and S. McLaughlin. Exploiting information geometry to improve the convergence properties of variational active contours. *IEEE Journal of Selected Topics in Signal Processing*, 7(4):700–707, 2013.
- P. Moulin and J. Liu. Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors. *IEEE Trans. Inform. Theory.*, 45:909–919, 1999.
- C. Prost-Squarcioni. Histologie de la peau et des follicules pileux. *MEDECINE/SCIENCES*, 22:131–137, 2006.
- R. Rakotomalala. Comparaison de populations: Tests paramétriques. *Bartlett test, Version*, 1:27–29, 2013.
- A. P. Raphael, T. A. Kelf, E. M. T. Wurm, A. V. Zvyagin, H. P. Soyer, and T. W. Prow. Computational characterization of reflectance confocal microscopy features reveals potential for automated photoageing assessment. *Experimental dermatology*, 22(7):458–463, 2013.
- D. Ratovoson. *Propriétés thermomécaniques de la peau et de son environnement direct*. PhD thesis, Université Montpellier II, France, 2011.
- C. P. Robert. Simulation of truncated normal variables. *Statistics and computing*, 5(2):121–125, 1995.
- C. P. Robert and D. Cellier. Convergence control of MCMC algorithms. In C. P. Robert, editor, *Discretization and MCMC Convergence Assessment*, pages 27–46. Springer Verlag, New York, 1998.
- G. O. Roberts, J. S. Rosenthal, et al. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical science*, 16(4):351–367, 2001.
- M. Rudin and R. Weissleder. Molecular imaging in drug discovery and development. *Nature Reviews, Drug discovery*, 2:123–131, February 2003.
- D. Sacchet. *Tomographie par cohérence optique plein champ linéaire et non linéaire*. PhD thesis, Université Paris Sud XI, France, 2010.

- J. Sage. *Rôle de la cathepsine S dans le remodelage de la membrane basale et régulation de son activité par des glycosaminoglycanes*. PhD thesis, Université François Rabelais de Tours, France, 2012.
- F. W. Scholz. Maximum likelihood estimation. *Encyclopedia of statistical sciences*, 1985.
- J. Selb. *Source virtuelle acousto-optique pour l'imagerie des milieux diffusants*. PhD thesis, Université Paris XI, France, 2002.
- K. Sharifi and A. Leon-Garcia. Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. *IEEE Trans. Circ. Syst., Video Techn.*, 5:52–56, 1995.
- E. Somoza, G. O. Cula, C. Correa, and H. J. B. *Automatic Localization of Skin Layers in Reflectance Confocal Microscopy*, pages 141–150. Springer International Publishing, Cham, 2014. ISBN 978-3-319-11755-3. doi: 10.1007/978-3-319-11755-3_16. URL http://dx.doi.org/10.1007/978-3-319-11755-3_16.
- J. R. Stanley, D. T. Woodley, S. I. Katz, and G. R. Martin. Structure and function of basement membrane. *The Journal of Investigate Dermatology*, 79:69s–72s, 1982.
- G. J. Tortora and S. R. Grabowski. *Principes d'anatomie et de physiologie*. De Boeck Université, 1993.
- J.-Y. Tourneret. Classification. Technical report, University of Toulouse, 2003. URL <http://tourneret.perso.enseeiht.fr/Classification/poly2003.pdf>.
- H.-V. Tran. *Caractérisation des propriétés mécaniques de la peau humaine in vivo via l'IRM*. PhD thesis, Université de Technologie de Compiègne, France, 2007.
- V. N. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, New-York, 1998.
- C. Wendorf. Manuals for univariate and multivariate statistics. *Stevens Point, WI: University of Wisconsin*, 2004.

- M. Wiltgen, A. Gerger, C. Wagner, J. Smolle, et al. Automatic identification of diagnostic significant regions in confocal laser scanning microscopy of melanocytic skin tumors. *Methods of Information in Medicine*, 47(1):14–25, 2008.
- G. V. Wouwer, P. Scheunders, and D. Dyck. Statistical texture characterization from discrete wavelet representation. *IEEE Trans. Image Proc.*, 8:592–598, Avril 1999.
- S. Yu, A. Zhang, and H. Li. A review of estimating the shape parameter of generalized Gaussian distribution. *J. Comput. Inf. Syst.*, 8(21):9055–9064, 2012.
- C. Yuan, Z. Peng, and X. Yan. Surface characterisation using wavelet theory and confocal laser scanning microscopy. *J. tribol. ,ASME*, 127:394–404, 2005.