



UNIVERSITÉ DU QUÉBEC À RIMOUSKI

**Développement des méthodes
d'intelligence artificielle pour le
diagnostic automatisé des maladies et
cancers de la peau**

*Mémoire présenté dans le cadre du programme de Maîtrise en
informatique – volet recherche (3017) en vue de l'obtention du
grade de Maître ès sciences (M.Sc.)*

Département de mathématique, d'informatique et génie

PAR

© ZAKARIAOU MOUNMIE

mai 2026

Composition du jury :

- **Pr. Mohamed Tarik Moutacalli**, président du jury, Université du Québec à Rimouski
- **Pr. Yacine Yaddaden**, directeur de recherche, Université du Québec à Rimouski
- **Pr. Abdenour Bouzouane**, codirecteur de recherche, Université du Québec à Chicoutimi
- **Pr. Bruno Bouchard**, examinateur externe, Université du Québec à Chicoutimi

Moi, Zakariaou MOUNMIE, déclare que ce mémoire intitulé, « Développement des méthodes d'intelligence artificielle pour le diagnostic automatisé des maladies et cancers de la peau » et le travail qui y est présenté sont de ma propre rédaction. Je confirme que :

- Ce travail a été réalisé entièrement et principalement pendant ma candidature à un diplôme de maîtrise en informatique - volet recherche (3017) au sein de cette université.
- Lorsque toute partie de ce mémoire a été précédemment soumis pour l'obtention d'un diplôme dans cette université, cela a été clairement indiqué.
- Lorsque j'ai consulté les travaux publiés par d'autres, cela est toujours clairement attribué.
- Lorsque j'ai cité la source est toujours indiquée. À l'exception de telles citations, le mémoire est entièrement le fruit de mon propre travail.
- J'ai reconnu toutes les principales sources d'aide.
- Lorsque le mémoire est basé sur un travail que j'ai réalisé en collaboration avec d'autres, j'ai clairement indiqué ce qui a été fait par les autres et ce à quoi j'ai contribué moi-même.

Signé :

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

Service de la bibliothèque

Avertissement

La diffusion de ce mémoire ou de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire « Autorisation de reproduire et de diffuser un rapport, un mémoire ou une thèse ». En signant ce formulaire, l'auteur concède à l'Université du Québec à Rimouski une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de son travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, l'auteur autorise l'Université du Québec à Rimouski à reproduire, diffuser, prêter, distribuer ou vendre des copies de son travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de la part de l'auteur à ses droits moraux ni à ses droits de propriété intellectuelle. Sauf entente contraire, l'auteur conserve la liberté de diffuser et de commercialiser ou non ce travail dont il possède un exemplaire.

« Si vous ne pouvez pas voler, alors courez ; si vous ne pouvez pas courir, alors marchez ; si vous ne pouvez pas marcher, alors rampez, mais quoi que vous fassiez, continuez d'avancer. »

Martin Luther King Jr.

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

Résumé

Département de mathématique, d'informatique et génie

Maîtrise en informatique – volet recherche (3017)

**Développement des méthodes d'intelligence artificielle pour le diagnostic
automatisé des maladies et cancers de la peau**

par Zakariaou MOUNMIE

Ces dernières années, l'incidence du cancer de la peau connaît une augmentation préoccupante, représentant un enjeu majeur de santé publique. Les principaux facteurs de risque identifiés incluent l'exposition excessive aux rayons ultraviolets et le recours croissant aux séances de bronzage artificiel. Des prédispositions génétiques ainsi que certaines pathologies immunosuppressives, telles que le VIH/SIDA, constituent également des facteurs de risque significatifs. Le pronostic de cette pathologie reste étroitement lié à la précocité du diagnostic, un dépistage précoce permettant dans de nombreux cas, un traitement efficace.

Le développement récent de systèmes de détection assistés par l'intelligence artificielle a ouvert des perspectives prometteuses dans le domaine du diagnostic dermatologique. Si la majorité des approches existantes se concentre sur la classification binaire (lésion maligne vs bénigne), peu d'entre elles abordent la reconnaissance spécifique des différents types de lésions cutanées. Cette limitation souligne la nécessité de développer des modèles capables d'assurer à la fois une haute performance diagnostique et une identification fine des variétés lésionnelles, adaptée à tous les phototypes cutanés.

Dans ce mémoire, nous proposons une méthodologie articulée en deux phases complémentaires. La première phase consiste en un système de segmentation des images dermoscopiques fondé sur des réseaux de neurones profonds, permettant la détection et l'extraction précise des régions d'intérêt. Cette étape préliminaire vise à optimiser la préparation des données pour la tâche de classification. La seconde phase introduit un système hiérarchique innovant qui optimise les performances par une combinaison des caractéristiques extraites, mais aussi par la fusion stratégique de plusieurs classificateurs, permettant une reconnaissance multiclassées des lésions cutanées.

L'évaluation rigoureuse de ces systèmes est conduite sur plusieurs bases de données publiques, facilitant la comparaison avec l'état de l'art et garantissant la reproductibilité des résultats. Cette approche intégrée vise à surmonter les limitations des méthodes conventionnelles en combinant segmentation précise et classification avancée.

En conclusion, ce travail de recherche ambitionne de contribuer à l'amélioration du diagnostic dermatologique par la proposition d'une solution optimisée et fiable, associant une segmentation robuste des lésions à une reconnaissance fine de leur typologie, avec pour objectif ultime d'augmenter la précision diagnostique et d'améliorer la prise en charge des patients.

UNIVERSITY OF QUEBEC AT RIMOUSKI

Abstract

Department of Mathematics, Computer Science and Engineering

Research Master's degree in Computer Science (3017)

Improving Skin Disease Diagnosis Using Artificial Intelligence

by Zakariaou MOUNMIE

In recent years, the incidence of skin cancer has shown a concerning increase, representing a major public health challenge. The main identified risk factors include excessive exposure to ultraviolet rays and the growing use of artificial tanning sessions. Genetic predispositions as well as certain immunosuppressive pathologies, such as HIV/AIDS, also constitute significant risk factors. The prognosis of this pathology remains closely linked to early diagnosis, with early detection enabling effective treatment in many cases.

The recent development of artificial intelligence-assisted detection systems has opened promising perspectives in the field of dermatological diagnosis. While the majority of existing approaches focus on binary classification (malignant vs. benign lesions), few address the specific recognition of different types of skin lesions. This limitation highlights the need to develop models capable of ensuring both high diagnostic performance and fine identification of lesion varieties, adapted to all skin phototypes.

In this thesis, we propose a methodology structured in two complementary phases. The first phase consists of a dermoscopic image segmentation system based on deep neural networks, enabling the detection and precise extraction of regions of interest. This preliminary step aims to optimize data preparation for the classification task. The second phase introduces an innovative hierarchical system that optimizes performance through a combination of extracted features, as well as the strategic fusion of multiple classifiers, enabling multiclass recognition of skin lesions.

The rigorous evaluation of these systems is conducted on several public databases, facilitating comparison with the state of the art and ensuring the reproducibility of results. This integrated approach aims to overcome the limitations of conventional methods by combining precise segmentation and advanced classification.

In conclusion, this research work aims to contribute to the improvement of dermatological diagnosis by proposing an optimized and reliable solution, combining robust lesion segmentation with fine recognition of their typology, with the ultimate goal of increasing diagnostic accuracy and improving patient care.

Remerciements

Je tiens à exprimer ma sincère gratitude à toutes les personnes qui ont accompagné et soutenu la réalisation de ce travail.

Mes premiers remerciements vont à mon directeur de recherche, **Pr. Yacine Yaddaden**, dont le soutien constant, les conseils avisés et le professionnalisme ont été essentiels tout au long de ces deux années. Sa guidance éclairée a été précieuse pour l'aboutissement de ce mémoire.

Je remercie également mon co-directeur de recherche, **Pr. Abdenour Bouzouane** pour son engagement, ses retours constructifs et son expertise qui ont grandement contribué à la qualité de ce travail.

Que les membres du jury, **Pr. Tarik Moutacalli** et **Pr. Bruno Bouchard** retrouvent ici l'expression de mes sincères remerciements pour l'honneur qu'ils me font en prenant le temps de lire et d'évaluer ce travail.

Une gratitude particulière à **Mademoiselle Amandine Basirico** pour ses précieux conseils médicaux et son soutien qui ont enrichi la dimension clinique de cette recherche.

Ma reconnaissance s'adresse aussi à tous les professeurs dont l'enseignement a nourri ma réflexion et rendu ce projet possible.

À ma famille, mes plus profonds remerciements : à mes parents pour leur amour et leur soutien indéfectible, à mes frères et sœurs pour leur encouragement constant.

Je garde une pensée émue pour mon ami **Njih Mboumbouo Issah Landry**, dont le soutien fut précieux durant ces années.

Je remercie aussi mes amis de Lévis, Québec, Montréal, Rimouski et Gatineau pour leur chaleur et leur présence, qui ont atténué la distance.

Certaines reformulations de ce manuscrit ont été assistées par un outil d'IA, utilisé uniquement pour améliorer la clarté et la cohérence du texte, sans impact sur le contenu scientifique.

Ce mémoire est le fruit de ces rencontres et de ces soutiens qui ont marqué mon parcours.

Table des matières

Résumé	xii
Résumé	xiv
Remerciements	xv
1 Introduction générale	1
1.1 Mise en contexte	1
1.1.1 Exposition aux rayonnement ultraviolets	2
1.1.2 Facteurs génétiques et antécédents familiaux	2
1.1.3 Type de peau et phototype	3
1.1.4 Age et vieillissement cutanée	3
1.1.5 Affaiblissement du système immunitaire	3
1.2 Problématique	4
1.2.1 Variabilité et complexité des lésions cutanées	5
1.2.2 Les artefacts et bruits dans les images	5
1.2.3 Les conditions d'acquisition variables des images	6
1.2.4 Performance et robustesse du système de détection	6
1.2.5 Défis éthiques et acceptabilité	7
1.3 Objectifs	7
1.3.1 Segmentation des lésions cutanées	7
1.3.2 Reconnaissance des lésions cutanées	8
1.4 Méthodologie suivie	10
1.4.1 Revue de littérature	10
1.4.2 Définition des axes de travail	11
1.4.3 Identification et récupération des données	11
1.4.4 Système de la détection du cancer de la peau	12
1.5 Contributions	15
1.5.1 Conférences Internationales	15
1.5.2 Revues Internationales	15
1.6 Structure du mémoire	16

2	Revue de littérature	17
2.1	Les fondamentaux	17
2.1.1	Structure et fonction de la peau	17
2.1.2	Les lésions cancéreuses et non cancéreuses de la peau	20
2.1.3	Diagnostic du mélanome cutané	24
2.2	Bases de données	30
2.2.1	HAM10000 : <i>Human Against Machine with 10 000 training images</i> [61]	31
2.2.2	PH2 : Hôpital Pedro Hispano [43]	32
2.2.3	ISIC : International Skin Imaging Collaboration [31, 19, 18, 51]	33
2.2.4	Dermatology Information System (DermIS) [22]	34
2.2.5	Base de données DermQuest [23]	35
2.2.6	Bases de données utilisées	35
2.3	Méthodes et approches existantes	36
2.3.1	Système de segmentation des lésions cutanées	36
2.3.2	Système de classification des lésions cutanées	39
2.4	Bilan et discussion	41
3	Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing	46
	Submission Summary	47
	Declaration of interrests	48
	Cover Letter	48
	Highlights	50
	Abstract	51
3.1	Introduction	52
3.2	Background	54
3.2.1	Semantic segmentation	54
3.2.2	Related Work	56
3.3	Proposed System	57
3.3.1	Pre-processing	57
3.3.2	Model Architecture and Training Strategy	59
3.4	Experimentation and Evaluation	61
3.4.1	Dataset Description	61
3.4.2	Evaluation Metrics	62
3.4.3	Training and Validation Strategy	63
3.5	Results and Discussion	64

3.5.1	Segmentation Results	64
3.5.2	Training Dynamics Analysis	65
3.5.3	Comparison with Existing Methods	67
3.6	Concluding Remarks	71
	Acknowledgements	74
	References	74
4	Hierarchical Two-Phase Skin Lesion Classification Using Reused CNN Feature Spaces and Cross-Stage Classifier Fusion	80
	Submission Summary	81
	Declaration of inerrests	82
	Highlights	84
	Abstract	85
4.1	Introduction	86
4.2	Background	88
4.2.1	Fundamentals	88
4.2.2	Related Work	89
4.2.3	Validity with Regard to Previous Work	91
4.3	Features Extraction Techniques	92
4.3.1	Models Based on Pre-trained CNN Architectures	92
4.3.2	Conventional ML-based Feature Extraction	93
4.4	Classifiers Techniques	93
4.4.1	LDA (Linear Discriminant Analysis)	93
4.4.2	NuSVM (Support Vector Machine Linear Kernel)	94
4.4.3	LR (Logistic Regression)	94
4.4.4	BNB (Bernoulli Naive Bayes)	95
4.5	Proposed System	95
4.5.1	Pre-processing	96
4.5.2	Features Extraction	97
4.5.3	Reduction of Dimensionality	97
4.5.4	Classification	97
4.6	Evaluation Protocol	98
4.6.1	Data Description	98
4.6.2	Evaluation Metrics	98
4.6.3	Validation Strategy	101
4.7	Results & Discussion	102
4.7.1	Experimental Results	102
4.7.2	Comparison with Existing Methods	106

4.7.3	Limitations	112
4.8	Concluding Remarks	113
	Acknowledgements	116
	References	118
5	A Hybrid Deep Learning Pipeline for Melanoma Detection Using Modified U-Net and VGG-19 Features	127
	Abstract	128
5.1	Introduction	128
5.2	Fundamentals	129
5.2.1	Semantic segmentation	129
5.2.2	Lesion classification	129
5.3	Related Work	129
5.3.1	Segmentation approaches	129
5.3.2	Classification approaches	129
5.4	Proposed System	129
5.4.1	Pre-processing	130
5.4.2	Segmentation using modified U-Net	130
5.4.3	Feature Extraction with Pre-trained VGG-19	130
5.4.4	Classification	130
5.5	Evaluation Protocol	130
5.5.1	Data description	130
5.5.2	Evaluation strategy	131
5.5.3	Evaluation metrics	131
5.6	Results & Discussion	131
5.6.1	Segmentation Results	131
5.6.2	Classification Results	131
5.7	Concluding Remarks	132
	Acknowledgment	133
	References	133
6	Conclusion Générale	135
6.1	Objectifs accomplis	136
6.2	Comparaison des approches développées	137
6.2.1	Segmentation des lésions cutanées	137
6.2.2	Classification des lésions cutanées	138
6.3	Limitations connues	138
6.4	Perspectives et travaux futurs	139
6.5	Retour d'expérience	140

xx

Annexe 143

Bibliographie 147

Table des figures

1.1	Système de segmentation des lésions cutanées	13
1.2	système de classification des lésions cutanées	14
2.1	Structure de la peau [57]	18
2.2	Coupe transversale de l'épiderme [57]	19
2.3	les vaisseaux du derme [42]	20
2.4	Exemples des lésions kératoses bénignes [61]	21
2.5	Exemples des lésions Dermato fibromes [61]	22
2.6	Exemples des lésions Nævi mélanocytaires [61]	22
2.7	Exemples des lésions vasculaires [61]	22
2.8	Exemples des lésions carcinomes basocellulaires [61]	23
2.9	Exemples des lésions mélanome [61]	24
2.10	Exemples des lésions Kératoses actiniques [61]	24
2.11	Méthode ABCDE pour la détection du mélanome [41]	25
2.12	Le dermoscope [21]	27
2.13	Microscopie confocale à réflectance et dispositif associé [30, 60]	29
2.14	Exemples de lésions dermoscopiques du jeu de données HAM10000	32
2.15	Exemples de lésions dermoscopiques du jeu de données PH2	32
2.16	Exemples d'images dermoscopiques issues de la base de données ISIC	34
2.17	Exemples d'images de la base de données DermIS	34
2.18	Exemples d'images de la base de données DermQuest	35
3.1	Overview of the proposed segmentation framwork	58
3.2	Overview of the U-Net architecture	60
3.3	Overview of benign lesions samples	62
3.4	Overview of malignant lesions samples	62
3.5	Training and validation loss and accuracy on the HAM10000 dataset	66
3.6	Training and validation loss and accuracy on the PH2 dataset	67
3.7	Training and validation loss and accuracy on the DermIS dataset	67
3.8	Training and validation loss and accuracy on the DermQuest dataset	68

3.9	Training and validation loss and accuracy on the Global dataset	68
3.10	Qualitative segmentation results on the HAM10000 dataset . . .	69
4.1	Overview of a common skin lesion classification system	89
4.2	Overview of the Proposed system	95
4.3	Representative dermoscopic images and corresponding seg- mentation masks for benign and malignant lesions	100
5.1	Overview of the proposed system	130
5.2	Benign lesions : HAM10000 (a-h) and PH2 (i-l)	131
5.3	Malignant lesions : HAM10000 (a-f) and PH2 (g-h)	132
6.1	Comparaison des performances de segmentation sur différents jeux de données.	137
6.2	Comparaison des performances en classification binaire sur différents jeux de données	138
6.3	Comparaison des performances en classification multi-classes sur différents jeux de données	139

Liste des tableaux

2.1	Description des jeux de données de lésions cutanées et de leurs distributions par classes	37
2.2	Comparaison des performances des méthodes de segmentation sur différentes bases de données dermatologiques	42
2.3	Comparaison des performances des méthodes de classification sur différentes bases de données dermatologiques.	43
3.1	Details about HAM10000 dataset before and after data augmentation	59
3.2	Details about PH2 dataset before and after data augmentation	59
3.3	Details about DermIS dataset before and after data augmentation	59
3.4	Details about DermQuest dataset before and after data augmentation	59
3.5	Training configuration and hyperparameters	64
3.6	Detailed segmentation performance of the proposed model on validation and test sets	65
3.7	Comparison of the proposed method with state-of-the-art segmentation methods on the HAM10000 dataset	69
4.1	Details about HAM10000 dataset before and after data augmentation	99
4.2	Details about PH2 dataset before and after data augmentation	99
4.3	Details about DermIS dataset before and after data augmentation	99
4.4	Details about DermQuest dataset before and after data augmentation	99
4.5	Comparison of binary classification methods on HAM10000 dataset	103
4.6	Comparison of binary classification methods on PH2 dataset .	104
4.7	Comparison of binary classification methods on DermIS dataset	105
4.8	Comparison of binary classification methods on DermQuest dataset	106
4.9	benign lesions classification results (ANV, CNV) on PH2 dataset	107

4.10	benign lesions classification results (BKL, DF, NV, VASC) on HAM10000 dataset	108
4.11	Malignant lesions classification results (AKIEC, BCC, MEL) on HAM10000 dataset	109
4.12	Multi-class classification results (seven classes) on HAM10000 dataset	110
4.13	Multi-class classification results (three classes) on PH2 dataset	111
4.14	two-phase multi-class classification results on HAM10000 dataset	111
4.15	two-phase multi-class classification results on PH2 dataset. . .	112
4.16	HAM10000 (Binary)	112
4.17	PH2 (Binary)	112
4.18	DermIS (Binary)	112
4.19	DermQuest (Binary)	112
4.20	HAM10000 (Malignant)	112
4.21	HAM10000 (Benign)	112
4.22	PH2 (Benign)	112
4.23	PH2 (Multi-class)	113
4.24	HAM10000 (Multi-class)	113
4.25	PH2 (Multi-class, Two-phase)	113
4.26	HAM10000 (Multi-class, Two-phase)	114
4.27	Comparative evaluation of the performance of the proposed system against state-of-the-art methods on the HAM10000 dataset for binary classification	116
4.28	Comparative evaluation of the performance of the proposed system against state-of-the-art methods on the HAM10000 dataset for multi-class classification	116
4.29	Comparative evaluation of the performance of the proposed system against state-of-the-art methods on the PH2 dataset for binary classification	116
4.30	Comparative evaluation of the performance of the proposed system against state-of-the-art methods on the PH2 dataset for multi-class classification	117
4.31	Comparative evaluation of the performance of the proposed system against state-of-the-art methods on the DermIS dataset for binary classification	117

4.32	Comparative evaluation of the performance of the proposed system against state-of-the-art methods on the DermQuest dataset for binary classification	117
5.1	Details about HAM10000 and PH2	131
5.2	Results and comparative analysis with state-of-the-art methods using the HAM10000 dataset	132
5.3	Results and comparative analysis with state-of-the-art methods using the PH2 dataset	132
5.4	Performance of the VGG-19-based representation on the HAM10000 dataset	132
5.5	Performance of the VGG-19-based representation on the PH2 dataset	132
5.6	Confusion matrix for HAM10000	132
5.7	Confusion matrix for PH2	132
5.8	Comparative analysis of our method with existing approaches over recent years	132

Liste des abréviations

OMS	Organisation Mondiale de la Santé
ML	Machine Learning
DL	Deep Learning
ABCDE	Asymétrie, Bords , Couleur, Diamètre, Evolution
ADN	Acide Désoxyribonucléique
AJCC	American Joint Committee on Cancer
ALM	Acral Lentiginous Melanoma
BRAF	B-Raf proto-oncogene, serine/threonine kinase
BNB	Bernoulli Naïve Bayes
CDKN2A	Cyclin-Dependent Kinase Inhibitor 2A
CNN	Convolutional Neural Network
HAM10000	Human Against Machine with 10000 images
HOG	Histogram of Oriented Gradient
IA	Intelligence Artificielle
IDAT	Inquiétude, Durée, Apparition, Transformation
KIT	Proto-oncogène KIT (ou CD117)
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis
LMM	Lentigo Maligna Melanoma
LR	Logistic Regression
NM	Nodular Melanoma
NRAS	Neuroblastoma RAS viral oncogene homolog
OIT	Organisation Internationale du Travail
PCA	Principal Component Analysis
PH2	Pedro Hispano Hospital dataset
PRAME	Preferentially Expressed Antigen in Melanoma
RCM	Reflectance Confocal Microscopy
RF	Random Forest
ROI	Region Of Interest
SIDA	Syndrome d'Immunodéficience Acquise
SSM	Superficial Spreading Melanoma
SVC	Support Vector Classifier
TCO	Tomographie en Cohérence Optique
UVA	Ultraviolet A
UVB	Ultraviolet B
UV	Ultraviolet
VIH	Virus de l'Immunodéficience Humaine
FCN	Fully Convolutional Network
ViT	Vision Transformer

GAN	Generative Adversarial Network
ESRGAN	Enhanced Super-Resolution GAN
SAM	Segment Anything Model
MAE	Masked AutoEncoder
YOLO	You Only Look Once
FPN	Feature Pyramid Network
TL	Transfer Learning
ODC	Open Data Collection

*À ma famille et à tous ceux qui me sont chers —
votre amour, votre foi et vos encouragements ont été
la lumière qui a guidé mes pas jusqu'au bout de ce
travail.*

Chapitre 1

Introduction générale

1.1 Mise en contexte

Le cancer de la peau est l'un des cancers les plus fréquents et représente un enjeu majeur de santé publique. Ses causes sont multifactorielles, pouvant être liées à l'environnement, à la génétique ou encore au mode de vie. Parmi les principaux facteurs de risque, l'exposition aux rayonnements ultraviolets (UV) occupe une place centrale, et son impact s'est accentué avec l'évolution des habitudes de vie et le vieillissement de la population.

En 2016, selon le gouvernement du Canada, près de 6 800 Canadiens ont été diagnostiqués avec un mélanome – la forme la plus agressive du cancer de la peau – et environ 1 200 en sont décédés [28, 58]. Ces chiffres soulignaient déjà l'importance d'un diagnostic précoce pour améliorer le pronostic.

Trois ans plus tard, en 2019, l'ampleur de ce problème à l'échelle mondiale a été mise en lumière : près de 19 000 personnes dans 183 pays sont décédées d'un cancer de la peau non-mélanome lié à une exposition professionnelle prolongée au soleil. La majorité (65 %) de ces décès concernait des hommes [46].

Plus récemment, en 2023, l'Organisation mondiale de la santé (OMS) et l'Organisation internationale du Travail (OIT) ont publié un communiqué de presse confirmant que les activités professionnelles exercées en extérieur sont responsables d'un décès sur trois dus au cancer de la peau non-mélanome [46].

Ces données, de plus en plus alarmantes, mettent en évidence la nécessité urgente de développer des méthodes innovantes et fiables capables d'assurer un diagnostic précoce et précis, afin de réduire la mortalité associée à cette maladie. De ce fait, les causes du cancer de la peau sont multifactorielles et peuvent être liées à l'environnement, à la génétique, voire au mode de vie.

1.1.1 Exposition aux rayonnement ultraviolets

L'exposition prolongée et répétée aux UV, qu'ils soient d'origine solaire ou artificielle (Cabine de bronzage), est la principale cause du cancer de la peau [58]. Les UV endommagent l'ADN des cellules cutanée, provoquant des mutations génétiques pouvant conduire à une prolifération incontrôlée des cellules cancéreuses.

- **Rayons UVB** : Responsables des coups de soleils et de la plupart des mutations de l'ADN
- **Rayons UVA** : pénètre plus profondément dans la peau et favorisent le vieillissement cutané, augmentant ainsi le risque du cancer.
- **Brulures solaire répétées** : notamment pendant l'enfance, elles augmentent le risque de mélanome à l'âge adulte.

1.1.2 Facteurs génétiques et antécédents familiaux

Certaines personnes sont génétiquement prédisposées à développer le cancer de la peau et cela peut se produire de plusieurs façons.

- **Hérédité** : Un antécédent familial de mélanome augmente le risque de développer cette maladie.
- **Syndrome des naevus atypiques** : certaines personnes présentent un grand nombre de grains de beauté atypiques multipliant ainsi leur risque de mélanome
- **Mutations génétiques** : des mutations génétiques comme celle du gène CDKN2A et BRAF, sont fréquemment associées aux mélanomes [58]

1.1.3 Type de peau et phototype

Les personnes à peau claire et aux yeux clairs sont plus sensibles aux UV et ont un risque plus élevé de développer un cancer de la peau [58].

- Les **Phototypes 1 et 2** : généralement, les peaux très claires qui brûlent facilement et bronzent peu sont les plus vulnérables.
- **Peau foncée** : bien que plus résistante aux UV, elle n'est pas totalement protégée et peut aussi développer le cancer de la peau qui, dans la plupart des cas, est détecté tardivement.

1.1.4 Age et vieillissement cutané

Le cancer de la peau est plus fréquent chez les personnes âgées en raison de l'accumulation progressive de dommages à l'ADN des cellules cutanées causés par les UV au fil du temps. Cependant, le mélanome touche aussi les jeunes adultes, en particulier ceux ayant eu une forte exposition solaire dans leur enfance [58, 24].

1.1.5 Affaiblissement du système immunitaire

Un système immunitaire affaibli réduit la capacité du corps à réparer les cellules endommagées et à éliminer les cellules cancéreuses naissantes [58, 24]. Nous pouvons citer entre autres.

- Patients transplantés sous immunosuppresseurs
- Personnes atteintes du VIH/SIDA
- Traitements prolongés par corticostéroïdes

Le diagnostic des lésions cutanées repose principalement sur l'examen clinique et la **dermoscopie**, des outils essentiels permettant d'évaluer visuellement les lésions suspectes. Cependant, cette évaluation dépend fortement de l'expertise du clinicien, ce qui introduit une variabilité inter-observateur et augmente le risque d'erreurs diagnostiques. Dans la majorité des cas, une biopsie est nécessaire pour confirmer le diagnostic [28].

La reconnaissance des lésions cutanées reste une tâche complexe en raison de leur grande variabilité morphologique : elles peuvent différer en *couleur*, *forme*, *texture* et *structure interne*, rendant leur identification et leur classification difficiles, même pour les professionnels expérimentés [61].

Pour améliorer la détection précoce du cancer de la peau, plusieurs approches ont été développées. Parmi les plus courantes figure la règle ABCDE, largement utilisée en dermatologie pour évaluer les lésions pigmentées suspectes, en particulier les mélanomes [24, 41]. D'autres techniques viennent compléter cette approche, telles que la photographie numérique évolutive, la microscopie confocale et, plus récemment, l'application de méthodes d'IA à l'analyse des images dermoscopiques.

Dans ce contexte, l'intelligence artificielle (IA) émerge comme un outil prometteur pour améliorer la précision du diagnostic et automatiser le processus de détection du cancer de la peau. En exploitant les méthodes d'apprentissage automatique (ML) et profond (DL), ces systèmes peuvent analyser des images dermoscopiques et identifier des signes précoces de cancer avec une précision comparable, voire supérieure à celle des dermatologues expérimentés.

1.2 Problématique

Le diagnostic du cancer de la peau repose encore essentiellement sur l'examen clinique, la dermoscopie et, dans certains cas, la biopsie [41]. Bien que ces méthodes soient efficaces, elles restent fortement tributaires de l'expertise du dermatologue, ce qui peut entraîner une variabilité dans les diagnostics. Pour pallier ces limitations et améliorer le diagnostic précoce, des systèmes d'aide au diagnostic basés sur l'IA ont été récemment développés, notamment pour l'identification des mélanomes. Toutefois, malgré leur intégration croissante en dermatologie, ces systèmes présentent encore certaines limites [50].

1.2.1 Variabilité et complexité des lésions cutanées

L'un des principaux défis dans la détection du cancer de la peau réside dans la grande variabilité des lésions [61]. Contrairement à d'autres pathologies où les caractéristiques sont relativement homogènes, les lésions cutanées présentent des variations considérables en termes de :

- **Taille** : les lésions peuvent être de quelques millimètres à plusieurs centimètres.
- **Formes** : certaines sont bien délimitées tandis que d'autres présentent des contours irréguliers
- **Couleur** : une même lésion peut avoir plusieurs variations de couleurs (noir, brun, rouge, bleu, etc.)
- **Texture** et structure interne : certaines lésions apparaissent lisses, d'autres rugueuses, ulcérées ou nodulaires.

Ces facteurs rendent la reconnaissance des lésions difficile, d'autant plus que certaines lésions bénignes (les kératoses séborrhéiques ou les angiomes) peuvent visuellement ressembler à des mélanomes [24], et que ces caractéristiques peuvent varier d'un type de peau à l'autre. Certaines recherches déjà menées dans ce domaine montrent que cette complexité peut augmenter le taux de faux positifs et de faux négatifs dans le diagnostic [49, 15, 37].

1.2.2 Les artefacts et bruits dans les images

Un autre défi majeur réside dans la présence d'artefacts et de bruit au sein des images dermoscopiques. Les poils, les bulles de gel, ou encore les marqueurs de mesure peuvent perturber le diagnostic et être confondus avec des caractéristiques propres à la lésion [16]. De plus, certaines structures anatomiques normales, comme les vaisseaux sanguins, peuvent interférer avec le processus de segmentation, ajoutant une difficulté supplémentaire à la tâche. Le bruit intrinsèque généré par la technique d'imagerie elle-même peut également dégrader la qualité de la segmentation, tout comme les ombres et reflets issus des conditions d'éclairage, qui modifient l'apparence des lésions

et compliquent leur délimitation précise [16], ce qui pourrait affecter les performances d'un système de détection du cancer de la peau.

1.2.3 Les conditions d'acquisition variables des images

Les conditions d'acquisition des images varient considérablement d'un environnement clinique à un autre, ce qui influence directement la fiabilité des systèmes de détection du cancer de la peau. Les différences d'éclairage, tant en intensité qu'en température de couleur, peuvent altérer la perception visuelle des caractéristiques d'une lésion. De même, les variations de distance et d'angle lors de la prise de vue modifient l'apparence et les dimensions apparentes des lésions, compliquant la standardisation des traitements d'image. La qualité des images dépend également du dispositif utilisé, chaque dermoscope présentant des performances optiques et techniques spécifiques [3]. Par ailleurs, l'absence de protocoles d'acquisition uniformisés entre les différents centres médicaux engendre des incohérences dans les données collectées, limitant ainsi la capacité de généralisation des modèles de détection automatisée.

1.2.4 Performance et robustesse du système de détection

Un système de détection du cancer de la peau basé sur l'IA doit répondre à plusieurs critères essentiels pour une intégration clinique efficace. Il doit assurer une précision diagnostique élevée, appuyée par des techniques de détection des contours ou de segmentation, permettant l'extraction automatique des régions d'intérêt (ROI) et l'amélioration des caractéristiques utilisées [5]. Sa capacité de généralisation est primordiale afin de s'adapter à la diversité des types de peau, des conditions d'acquisition et des artefacts visuels [20, 3]. L'interprétabilité est également cruciale pour garantir des résultats compréhensibles et exploitables par les professionnels de la santé. De plus, le système doit être robuste face aux biais, grâce à un entraînement sur des bases diversifiées, et offrir un temps de réponse rapide afin de permettre une prise en charge en temps réel des patients.

1.2.5 Défis éthiques et acceptabilité

Le déploiement de l'IA dans le processus de détection du cancer de la peau soulève également des défis éthiques majeurs. L'une des principales préoccupations concerne la fiabilité des systèmes et le risque de mauvaise classification d'une lésion (faux négatifs ou faux positifs) [49]. Un faux négatif peut entraîner un retard de diagnostic et une aggravation de la maladie, tandis qu'un faux positif peut provoquer une anxiété inutile chez le patient, conduisant à des biopsies non justifiées. Par ailleurs, la collecte et l'utilisation des données médicales pour l'entraînement des modèles nécessitent des mesures strictes en matière de protection de la vie privée et de respect de la confidentialité des patients. Il est donc essentiel d'assurer une transparence dans le processus décisionnel des modèles et d'établir des protocoles de validation clinique rigoureux avant leur intégration en pratique médicale.

1.3 Objectifs

Le principal objectif de cette étude est de développer un système fiable de détection des lésions cutanées en tirant parti des techniques d'IA, dans le but d'améliorer les taux de diagnostic du cancer de la peau. Pour mieux cerner la finalité de cette recherche, des objectifs spécifiques ont été définis en fonction des différents axes de travail.

1.3.1 Segmentation des lésions cutanées

La segmentation d'images constitue une étape essentielle du processus de diagnostic assisté par intelligence artificielle du cancer de la peau. Son objectif principal est de délimiter automatiquement la région d'intérêt — généralement la lésion cutanée — avec la plus grande précision possible, tout en éliminant les éléments indésirables tels que le bruit et le fond de l'image. Les objectifs de la segmentation sont :

- **Délimitation précise des zones d'intérêt** : cette délimitation a pour but d'identifier et de séparer les ROI, ce qui permettra une meilleure analyse des caractéristiques spécifiques de la zone d'intérêt détectée.
- **Réduction de la complexité des images et amélioration des modèles d'IA** : en transformant l'image brute en une représentation simplifiée, cela facilitera l'interprétation par les modèles d'IA, voire par un humain. En précisant les contours et les ROI, la segmentation permettra une amélioration de la précision des modèles de classification en facilitant l'extraction des caractéristiques pertinentes, telles que la taille, la forme et la texture des ROI. De plus, une segmentation précise permet d'améliorer la robustesse et la généralisation du modèle en réduisant l'influence des variations liées aux conditions d'acquisition des images (éclairage, angle de prise de vue, qualité de l'appareil). En normalisant la zone d'analyse, elle garantit une meilleure cohérence des données et optimise ainsi la performance des algorithmes de DL.
- **Optimisation des données d'entrée pour la phase de reconnaissance des lésions** : La segmentation est une tâche cruciale dans notre recherche, car elle constitue une étape essentielle de prétraitement des données avant même la construction de notre modèle de reconnaissance des lésions cutanées. Elle permet d'isoler avec précision les ROI, en séparant les lésions cutanées de la peau saine, ce qui améliore considérablement la qualité des images analysées par le modèle.

1.3.2 Reconnaissance des lésions cutanées

La reconnaissance des lésions cutanées constitue une étape essentielle du processus de diagnostic assisté par IA. Elle a pour objectif d'identifier et de classer les anomalies cutanées à partir des caractéristiques extraites des images dermatologiques. Cette tâche peut être binaire — distinguant les lésions bénignes des lésions malignes — ou multicatégorie, lorsqu'il s'agit de différencier les sous-types de cancers ou d'autres affections cutanées. L'enjeu majeur de cette étape réside dans la conception d'un système automatisé,

précis et robuste, capable de soutenir efficacement le diagnostic clinique et de contribuer à une détection précoce des cancers de la peau. Les principaux objectifs de la classification ou reconnaissance des lésions cutanées sont :

- **Diagnostic précoce et précis** : L'objectif principal de la classification des images dans le domaine du cancer de la peau est d'identifier avec précision et rapidité les lésions suspectes en distinguant les lésions malignes des lésions bénignes. En automatisant cette tâche, les systèmes basés sur l'IA réduisent les erreurs de diagnostic.
- **Réduction de la subjectivité et de la variation inter-observateur** : le diagnostic traditionnel des lésions cutanées repose sur l'expertise du clinicien, mais demeure sujet à une variabilité inter-observateur. L'automatisation par l'IA permet de standardiser ce processus en s'appuyant sur des données objectives et des caractéristiques quantifiables (forme, couleur, texture). Cette approche favorise des résultats plus cohérents et reproductibles, tout en réduisant les erreurs liées à la subjectivité humaine et en renforçant la fiabilité du diagnostic.
- **Classification fine des types de lésions cutanées** : La classification d'images vise à distinguer les différents types de cancer de la peau, tels que le mélanome, le carcinome basocellulaire, le carcinome spinocellulaire, etc. Chacun nécessite une prise en charge spécifique. Elle permet aussi d'identifier les lésions bénignes courantes, comme les grains de beauté atypiques, les kératoses séborrhéiques ou les angiomes, qui peuvent visuellement ressembler à des lésions malignes. Cela permettra d'orienter les dermatologues vers des traitements appropriés et d'éviter les interventions inutiles.
- **Amélioration de la généralisation et de la robustesse du modèle** : Pour être efficaces, les modèles de classification doivent être capables de fonctionner dans divers contextes et sur des populations variées. Cela nécessite un entraînement sur des jeux de données variés incluant différents phototypes de peau, des conditions d'éclairage variées et des artefacts courants. L'objectif est de garantir que le système

fonctionne de manière fiable dans des environnements réels, indépendamment des variations techniques ou démographiques, tout en maintenant une haute précision et une robustesse face aux défis pratiques.

1.4 Méthodologie suivie

Dans ce travail, nous proposons une approche méthodologique progressive pour approfondir la problématique et obtenir des résultats fiables, en nous appuyant sur les recherches antérieures. Le système automatisé se compose de deux étapes principales : la segmentation et la classification.

La segmentation a pour objectif d'identifier et de délimiter avec précision la ROI dans les images dermoscopiques, fournissant ainsi des données exploitables pour les étapes suivantes. La classification s'effectue en deux niveaux : le premier, binaire, permet de distinguer les lésions bénignes des lésions malignes et constitue une étape clé, car il guide la seconde phase, qui vise à déterminer le type spécifique de chaque lésion. Les étapes suivies pour atteindre nos objectifs de recherche sont présentées ci-dessous.

1.4.1 Revue de littérature

Cette étape consiste à analyser de manière approfondie des articles scientifiques et des travaux de recherche publiés dans le domaine de la détection des maladies de la peau. Ce travail nous permettra de dresser un panorama des avancées actuelles et des méthodes déjà développées dans ce champ d'étude. En parallèle, il nous aidera à mieux comprendre les étapes clés du développement d'un système de détection automatisée du cancer de la peau, à évaluer les performances des solutions existantes, à repérer leurs limites et à identifier les améliorations nécessaires pour concevoir un système performant et fiable. Cette phase est cruciale pour orienter notre recherche et les axes de travail à explorer, comme définis précédemment. Plus précisément, notre revue de littérature se focalisera sur les approches fondées sur le

ML, le DL et, récemment, les Vision Transformers (ViT), qui aujourd'hui, sont au cœur de l'innovation dans la détection et la classification des lésions cutanées. Nous étudierons également les techniques de segmentation d'images médicales essentielles pour identifier et délimiter avec précision les ROI sur la peau, afin de faciliter le diagnostic et le traitement.

1.4.2 Définition des axes de travail

Dans le cadre de ce projet sur la détection du cancer de la peau, nous avons identifié deux grands axes de travail principaux, chacun visant à répondre à des besoins spécifiques et complémentaires. Ces axes sont conçus pour explorer et améliorer les méthodes actuelles de détection et de classification des lésions cutanées en tirant parti des avancées récentes en IA.

- *Segmentation des lésions cutanées* : le but est de développer un système de segmentation capable d'identifier et de délimiter avec précision les zones suspectes de la peau. C'est une étape clé pour isoler la lésion et fournir des informations détaillées sur son étendue et ses caractéristiques.
- *Classification des lésions cutanées* : le but ici est de développer un système de classification capable de distinguer les lésions cutanées malignes des lésions bénignes et d'affiner ce système pour identifier les différents types de lésions cutanées en tenant compte de leur variabilité morphologique.

1.4.3 Identification et récupération des données

En raison d'un manque d'expertise dans le domaine de la dermatologie et de contraintes de temps, il ne sera pas possible de procéder à une collecte de données spécifique pour ce projet. Toutefois, lors de la revue de la littérature, nous avons identifié quatre bases de données de référence [61, 23, 43, 22] qui seront utilisées pour notre étude. Ces ensembles de données publiques, largement reconnus dans le domaine de la détection du cancer de la peau,

nous permettent de travailler sur des informations fiables et déjà validées par la communauté scientifique.

L'utilisation de ces bases de données publiques présente plusieurs avantages. Elle permet une comparaison directe avec les travaux antérieurs, renforçant ainsi la validité et la reproductibilité de nos résultats. Cette approche nous assure une base solide pour développer et évaluer nos modèles de manière rigoureuse, tout en respectant les normes académiques et scientifiques.

1.4.4 Système de la détection du cancer de la peau

La conception d'un tel système passe par deux phases, comme mentionné plus haut : notamment une segmentation des zones suspectes ou des lésions cutanées, suivie d'une phase de classification de ces lésions.

Conception d'un système de segmentation des lésions cutanées

Avant d'entamer la tâche de segmentation, plusieurs techniques de pré-traitement ont été appliquées aux données afin d'optimiser les performances du système. Une approche de segmentation sémantique a été adoptée, consistant à attribuer une étiquette à chaque pixel de l'image pour identifier et isoler les ROI, en particulier les lésions cutanées, de la peau saine. Cette démarche permet une analyse fine des caractéristiques morphologiques des lésions, facilitant ainsi leur classification ultérieure. La méthode de segmentation retenue s'appuie sur une architecture adaptée aux images médicales, capable de combiner des informations globales et locales pour une détection précise des contours, même dans le cas de lésions présentant des bords flous ou irréguliers. Grâce à ce cadre méthodologique rigoureux et à un pré-traitement soigneusement conçu, la segmentation obtenue pourrait offrir une qualité élevée, essentielle pour la détection et l'analyse des lésions cutanées dans le cadre de cette étude.

Contrairement à plusieurs travaux antérieurs reposant sur des architectures classiques de détection de contours, telles que les méthodes de Sobel, de watershed ou les contours actifs (Snakes), la présente approche introduit

des ajustements spécifiques au domaine dermatologique afin d'améliorer la détection des bords complexes et de réduire les erreurs liées aux artefacts ou aux variations d'éclairage. En outre, en s'appuyant sur des bases de données variées, elle vise à renforcer la capacité de généralisation et la fiabilité du modèle dans un contexte clinique réel.

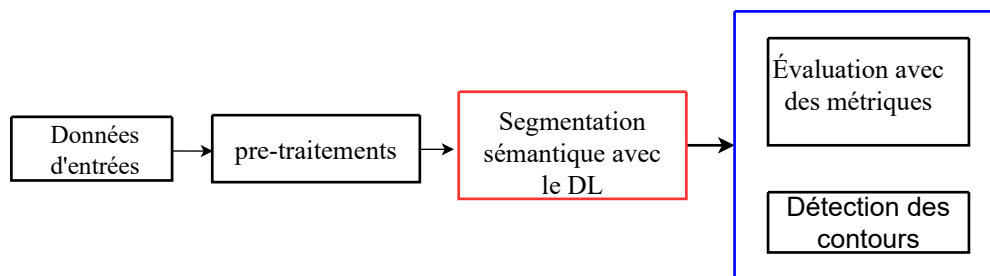


FIGURE 1.1 – Système de segmentation des lésions cutanées

La figure 1.1 présente le déroulement des étapes de segmentation des lésions cutanées qui seront expliquées en détail ci-dessous.

- **Prétraitement des données** : cette étape consiste à affiner les données pour une meilleure performance du modèle. Pour ce faire, nous avons commencé par une augmentation des données, accompagnée d'un équilibre des classes d'images, suivie de la normalisation et du redimensionnement de ces images.
- **Conception du modèle de segmentation** : l'utilisation d'une architecture de segmentation connue pour construire la nôtre a permis de gagner du temps, mais aussi d'améliorer les performances en l'adaptant à la tâche spécifique de segmentation des lésions cutanées.
- **Évaluation et interprétation des résultats** : mesurer l'efficacité du modèle, comprendre ses prédictions et garantir sa robustesse ainsi que sa généralisation à de nouvelles images.

Conception d'un système de classification des lésions cutanées

Dans le but d'optimiser notre système, nous avons appliqué diverses techniques de prétraitement, incluant une étape d'extraction de la ROI à l'aide du modèle de segmentation proposé, afin de préparer efficacement les données

pour la tâche de classification. Les images ainsi préparées servent à l'élaboration d'un modèle capable d'identifier les lésions cutanées bénignes et malignes, tout en distinguant les différents types de cancers associés à chaque lésion. Pour ce faire, nous exploitons des approches d'IA, notamment celles basées sur les réseaux de neurones convolutifs (CNN) et les techniques traditionnelles. Le modèle développé est ensuite évalué et validé à l'aide de métriques couramment utilisées dans les tâches de classification.

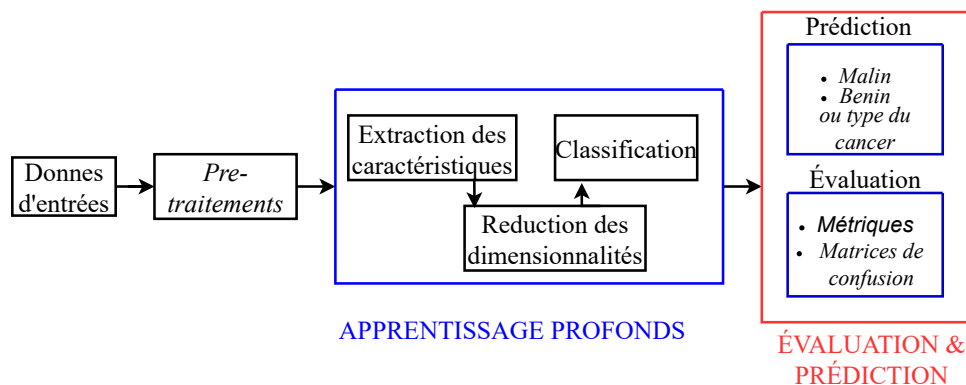


FIGURE 1.2 – système de classification des lésions cutanées

La figure 1.2 présente le déroulement des étapes de classification des lésions cutanées, qui seront expliquées en détail ci-dessous :

- **Prétraitements des données** : la première étape consiste en une augmentation de données, suivie d'un équilibrage des classes d'images, d'une extraction des **ROI**, d'une égalisation d'histogramme et d'un redimensionnement des images des lésions cutanées.
- **Extraction des caractéristiques** : les données sont extraites à l'aide des techniques basées sur les CNN et des méthodes traditionnelles.
- **Réduction de la dimensionnalité** : il s'agit ici d'une étape non obligatoire mais essentielle, qui vise à supprimer les caractéristiques redondantes et non pertinentes pour la tâche de classification.
- **Évaluation et validation du modèle** : des tests statistiques adaptés, combinés à un processus d'apprentissage et de validation d'un modèle d'apprentissage automatique, ont permis d'analyser la relation entre

les caractéristiques extraites, la nature des lésions (cancéreuses ou non cancéreuses) et les différents types de lésions cancéreuses identifiées.

1.5 Contributions

Le présent travail de recherche a pour objectif principal de proposer un système de détection basé sur l'IA pour le diagnostic automatisé des maladies et cancers de la peau

1.5.1 Conférences Internationales

Dans le cadre de ce mémoire, nous avons eu l'occasion de présenter nos avancées ainsi que nos méthodes proposées lors de conférences internationales. Voici une liste de nos contributions acceptées et présentées lors d'événements de vulgarisation scientifique :

Z. Mounmie, Y. Yaddaden, A. Bouzouane, "A Hybrid Deep Learning Pipeline for Melanoma Detection Using Modified U-Net and VGG-19 Features," in *Les 2Nd Conférence internationale sur les technologies avancées en génie électronique et électrique (ICATEEE2025)*, IEEE, 2025, pp. 1-7, (Statut : Publié)

1.5.2 Revues Internationales

Nous avons également soumis nos travaux à des revues scientifiques internationales de renom. Voici les papiers actuellement en cours de traitement :

Z. Mounmie, Y. Yaddaden, A. Bouzouane, "Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing" in *Biomedical Signal Processing and Control (BSPC-D-26-03612)*. (Statut : Soumis)

Z. Mounmie, Y. Yaddaden, A. Bouzouane, "Hierarchical Two-Phase Skin Lesion Classification Using Reused CNN Feature Spaces and Cross-Stage Classifier Fusion" in *Engineering Applications of Artificial Intelligence (EAAI-S-26-07398)*. (Statut : Soumis)

1.6 Structure du mémoire

Ce mémoire est structuré comme suit :

- La Revue de littérature présente le sujet dans son ensemble avec des avancées dans le domaine en élaborant les limites identifiées.
- Le Chapitre 3 illustre la segmentation des lésions cutanées à l'aide d'une architecture de CNN améliorée.
- Le Chapitre 4 illustre la détection de la présence ou non du cancer et la classification des types du cancer de la peau à l'aide des techniques de ML et de DL.
- Le Chapitre 5 présente l'intégration du modèle de segmentation avec le modèle de classification, dans le but d'assurer une détection du cancer de la peau plus robuste et plus précise.
- La Conclusion Générale présente une synthèse des méthodes étudiées en mettant en évidence leurs avantages et leurs limites, tout en proposant des pistes d'amélioration et des perspectives pour notre recherche.
- L'Annexe met en valeur les deux posters de mes participations au Forum Innovation, Ingénierie, Informatique et Entrepreneuriat (FI3E).

Chapitre 2

Revue de littérature

Ce chapitre examine les travaux de recherche et les efforts menés par diverses équipes sur la problématique de la détection du cancer de la peau. La section 2.1 introduit la structure de la peau, les types de cancers existants et les types de diagnostics du mélanome cutané, tandis que la section 2.2 détaille les principales bases de données utilisées. La section 2.3 analyse les contributions d'autres études déjà menées, et enfin, la section 2.4 propose un bilan des approches présentées, en mettant en évidence leurs limites et en ouvrant la discussion sur les axes d'amélioration possibles.

2.1 Les fondamentaux

Dans cette section, les principes de base relatifs à la structure de la peau et aux différents types de cancer cutané sont abordés ainsi que les diagnostics du cancer de la peau largement utilisés chez les dermatologues.

2.1.1 Structure et fonction de la peau

La peau, organe le plus étendu du corps humain, constitue une barrière dynamique entre l'organisme et son environnement. Elle assure des fonctions vitales de protection, de régulation thermique, de perception sensorielle et de synthèse de molécules essentielles, comme la vitamine D [1]. Sa structure complexe se compose de trois couches principales (voir figure 2.1) : l'épiderme (superficiel), le derme (intermédiaire) et l'hypoderme (profond).

ANATOMIE ET PHYSIOLOGIE DE LA PEAU NORMALE

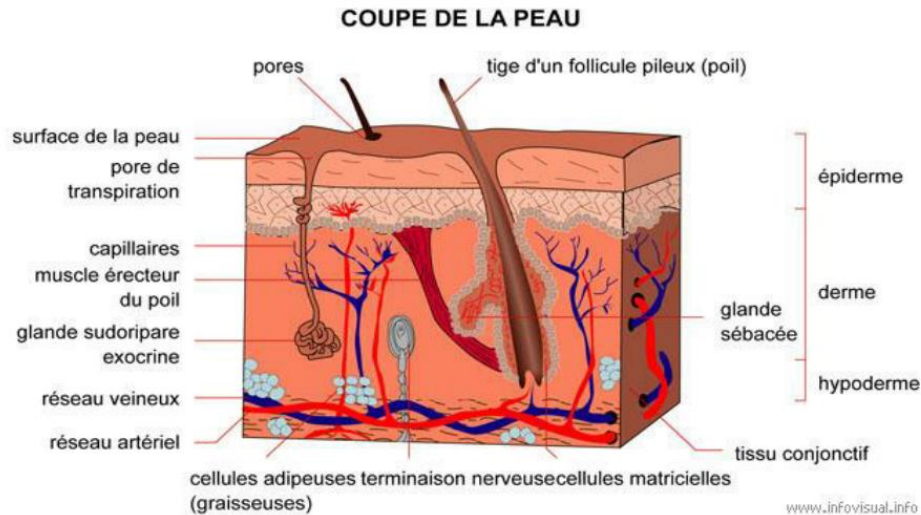


FIGURE 2.1 – Structure de la peau [57]

Épiderme : Couche la plus superficielle de la peau, il joue un rôle fondamental dans la protection contre les agressions extérieures, les rayonnements UV et la déshydratation [42, 25]. Il est organisé en plusieurs strates aux fonctions spécifiques :

- **Couche basale (stratum basale) :** contient les cellules souches et les mélanocytes, producteurs de mélanine.
- **Couche épineuse (stratum spinosum) :** kératinocytes liés par des desmosomes assurant la cohésion.
- **Couche granuleuse (stratum granulosum) :** production de kératine et de lipides épidermiques.
- **Couche claire (stratum lucidum) :** présente uniquement dans les zones de peau épaisse.
- **Couche cornée (stratum corneum) :** formée de cornéocytes morts, barrière imperméable contre toxines, microbes et perte de fluides.

L'épiderme se renouvelle en continu avec un cycle d'environ 28 jours [25]. Son épaisseur varie de 0,05 mm (paupières) à 1,5 mm (paumes,

plantes) [25, 11].

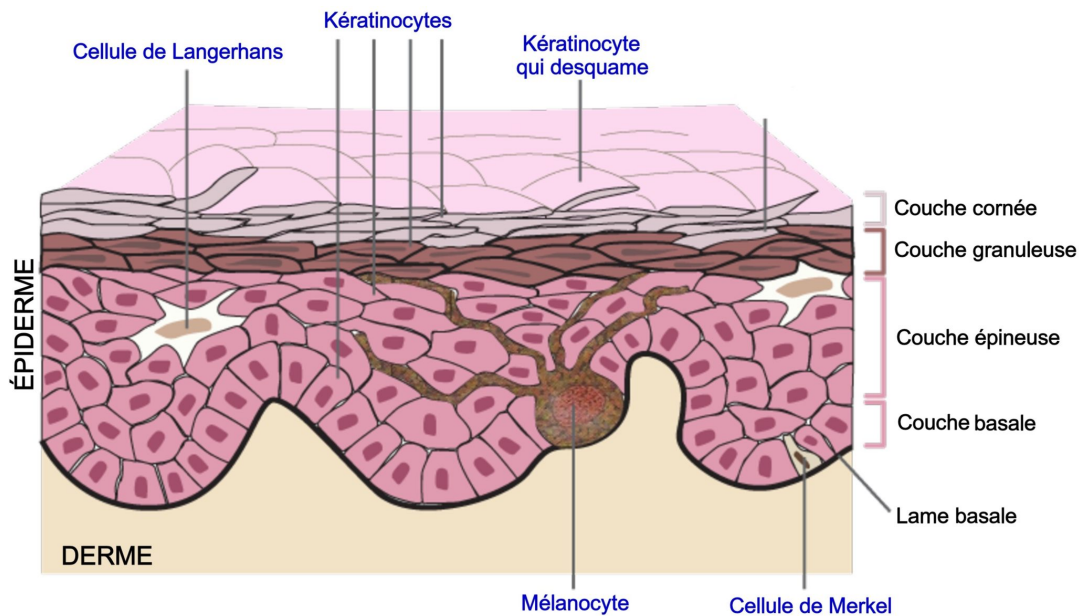


FIGURE 2.2 – Coupe transversale de l'épiderme [57]

Outre les kératinocytes, il comprend aussi :

- des **mélanocytes** (5–10%) produisant la mélanine,
- des **cellules de Langerhans** pour l'immunité cutanée,
- des **cellules de Merkel** pour la sensibilité tactile [11, 9].

Derme : Situé sous l'épiderme, le derme constitue la couche intermédiaire, plus épaisse, mesurant entre 0,6 et 3 mm selon les régions du corps [11]. Il est composé de tissu conjonctif riche en fibres de collagène et d'élastine, assurant élasticité et résistance mécanique à la peau [42]. la figure 2.3 est une représentation des vaisseaux du derme. On y distingue deux zones principales :

- **Derme papillaire** : couche superficielle, en contact avec l'épiderme, contenant des capillaires sanguins et des terminaisons nerveuses responsables de la sensibilité cutanée.
- **Derme réticulaire** : couche plus profonde, contenant des fibres de collagène épaisses, des follicules pileux, des glandes sudoripares et sébacées, ainsi que des vaisseaux sanguins et lymphatiques.

Le derme joue un rôle majeur dans la thermorégulation, la nutrition de l'épiderme et la cicatrisation.

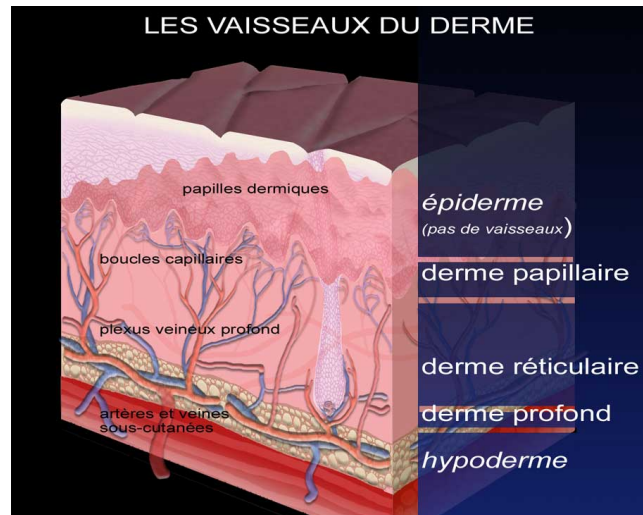


FIGURE 2.3 – les vaisseaux du derme [42]

Hypoderme : Également appelé tissu sous-cutané, il constitue la couche la plus profonde de la peau [42]. Il est formé principalement de tissu adipeux et de fibres conjonctives lâches. Ses fonctions principales sont le stockage énergétique, l'isolation thermique et l'amortissement mécanique face aux traumatismes [11, 42]. L'hypoderme contient également des vaisseaux sanguins de calibre plus important et des nerfs, participant à la vascularisation et à l'innervation des couches supérieures de la peau.

2.1.2 Les lésions cancéreuses et non cancéreuses de la peau

Le cancer de la peau est l'un des types de cancer les plus courants dans la population caucasienne. C'est l'un des types les plus dangereux et dont la croissance est la plus rapide, et représente donc un problème de santé important [46]. Nous avons mené des recherches qui nous ont permis d'identifier plusieurs de ces types de cancer de la peau regroupés en deux catégories, notamment les lésions non cancéreuses aussi appelées lésions bénignes et les lésions cancéreuses aussi appelées lésions malignes.

Les lésions non cancéreuses ou bénignes

Une lésion bénigne ou non cancéreuse est une anomalie ou une masse présente dans un tissu ou un organe qui ne relève pas d'un cancer [1]. Elle se caractérise généralement par son développement lent et sa localisation limitée et circonscrite, elle ne se propage pas à d'autres parties du corps et n'envahit pas les tissus voisins, ce qui signifie qu'elle ne produit pas de métastases [14]. Bien qu'elle ne soit habituellement pas mortelle et généralement asymptomatique, elle peut néanmoins entraîner des complications si sa croissance exerce une pression sur les structures environnantes [14]. Dans la majorité des cas, ce type de lésions n'augmente pas le risque de cancer, mais certaines situations particulières, comme des lésions atypiques, peuvent nécessiter un suivi médical plus attentif. Parmi les lésions non cancéreuses existantes, nous pouvons citer entre autres :

- **Kératoses bénignes** : ce sont des entités cutanées fréquentes qui se manifestent généralement par une lésion asymptomatique, solitaire, située le plus souvent sur le tronc (voir figure 2.4).

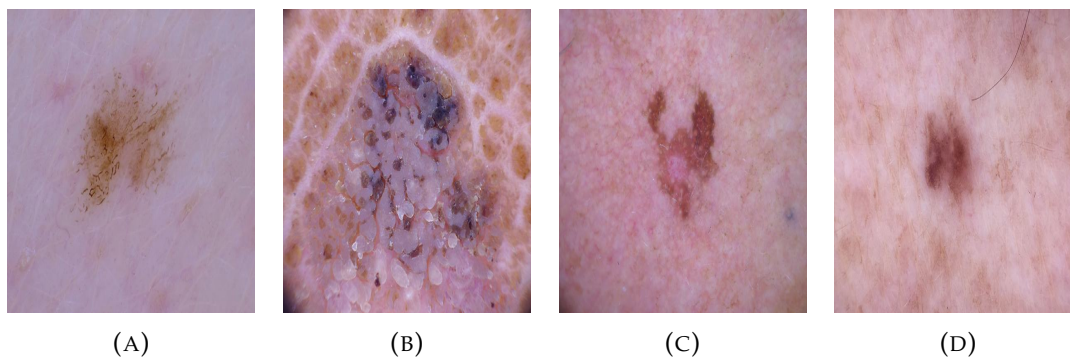


FIGURE 2.4 – Exemples des lésions kératoses bénignes [61]

- **Dermatofibrome** : c'est une lésion papuleuse ou nodulaire ferme, le plus souvent de petite taille et de couleur brun rougeâtre (voir figure 2.5).
- **Nævi mélanocytaires** : c'est une lésion dont les contours sont généralement réguliers, mal définis, qui sont de couleurs variées, brun clair à brun, avec des zones maculaires et papuleuses (voir figure 2.6).

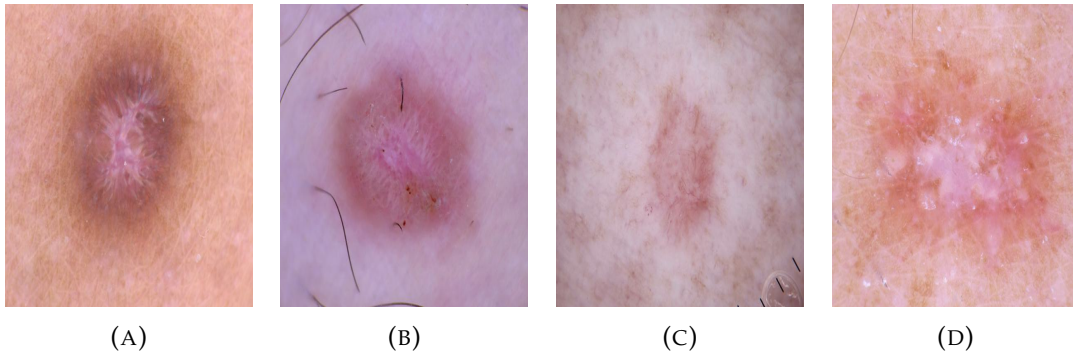


FIGURE 2.5 – Exemples des lésions Dermato fibromes [61]

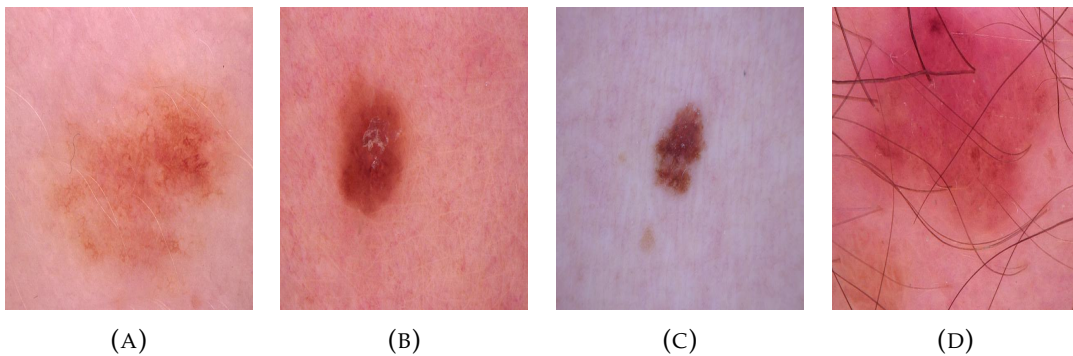


FIGURE 2.6 – Exemples des lésions Nævi mélanocytaires [61]

— **Lésions vasculaires** : les lésions vasculaires cutanées peuvent être bénignes ou malignes. Les bénins incluent les hémangiomes et les angiomes, tandis que les malins, plus rares, comprennent les angiosarcomes, qui apparaissent souvent sur le visage et le cuir chevelu chez l'adulte. La figure 2.7 présente les lésions vasculaires bénignes.

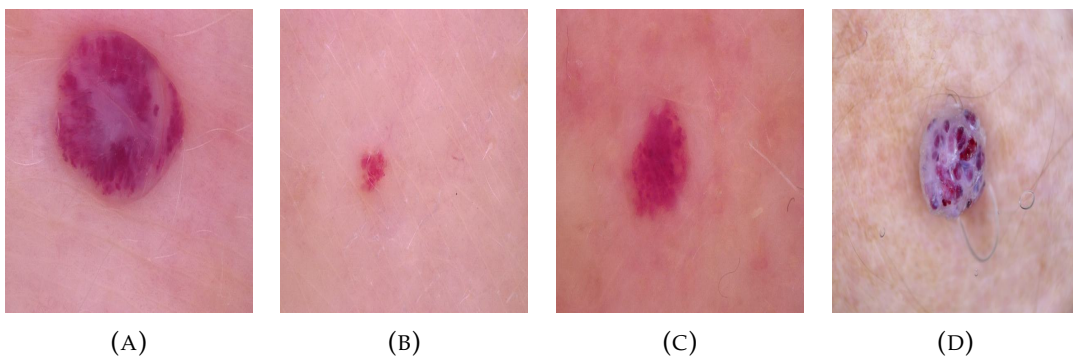


FIGURE 2.7 – Exemples des lésions vasculaires [61]

Les lésions cancéreuses ou malignes

Les lésions cancéreuses ou malignes sont des anomalies cellulaires caractérisées par une croissance incontrôlée des cellules, leur capacité à envahir les tissus environnants et à se propager à d'autres parties du corps par le biais du sang ou du système lymphatique (métastase) [27, 38]. Contrairement aux lésions bénignes, elles représentent un véritable danger pour la santé et peuvent être mortelles si elles ne sont pas diagnostiquées et traitées rapidement.

Elles se distinguent par une croissance rapide et désorganisée, une capacité à infiltrer les structures voisines, une tendance à récidiver même après un traitement efficace, parfois à distance du site initial, et une possible résistance aux traitements, qui peut se développer au fil du temps, souvent après une exposition aux thérapies en raison de l'apparition des mutations génétiques rendant certaines cellules cancéreuses plus agressives et moins sensibles aux traitements [14, 10].

Parmi les plus courants, on retrouve :

- **Carcinome basocellulaire** : il est formé à partir des cellules squameuses de l'épiderme (voir figure 2.8).

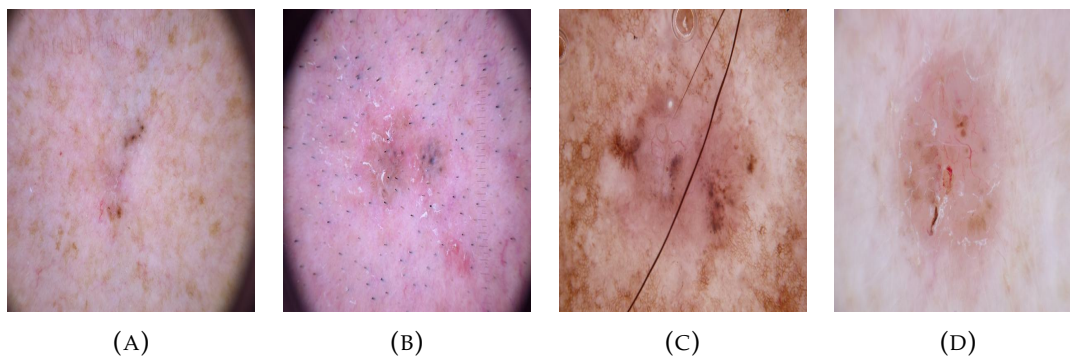


FIGURE 2.8 – Exemples des lésions carcinomes basocellulaires [61]

- **Mélanome** : il se développe à partir des mélanocytes qui sont des cellules responsables de la mélanine (voir la figure 2.9).
- **Kératoses actiniques** : c'est une lésion formée de petites plaques squameuses qui se forme sur la peau abîmée par le soleil.

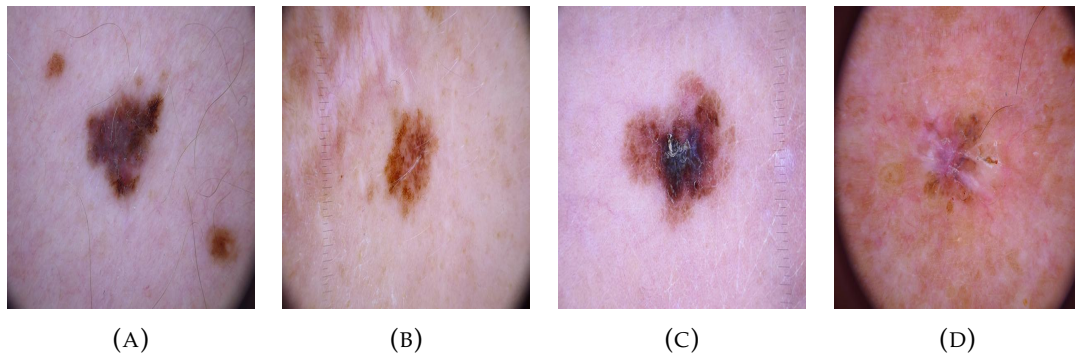


FIGURE 2.9 – Exemples des lésions mélanome [61]

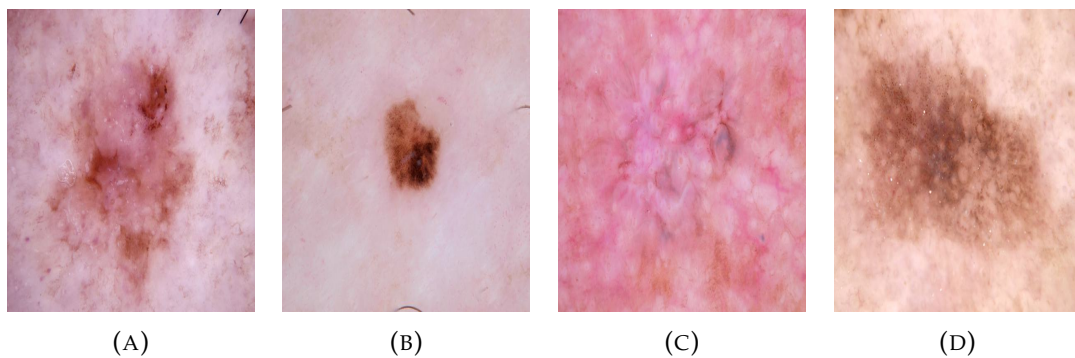


FIGURE 2.10 – Exemples des lésions Kératoses actiniques [61]

Face à la diversité des lésions cutanées, les experts ont mis au point plusieurs approches diagnostiques, combinant techniques cliniques, dermatoscopiques et plus récemment celles assistées par ordinateur, afin de favoriser une détection précoce et fiable du cancer de la peau.

2.1.3 Diagnostic du mélanome cutané

Le diagnostic du mélanome cutané repose sur une approche multidisciplinaire combinant l'examen clinique, la dermoscopie, l'analyse histopathologique, ainsi que les apports récents de l'imagerie médicale et de la biologie moléculaire [59, 1]. Cette complémentarité permet d'optimiser la détection précoce, d'affiner le pronostic et de guider la stratégie thérapeutique.

Diagnostic clinique

L'évaluation clinique débute par une inspection visuelle minutieuse des lésions pigmentées. Le système mnémotechnique ABCDE [41, 24] est couramment utilisé comme outil de repérage :

- **A** (Asymétrie) : un grain de beauté bénin est généralement symétrique, tandis qu'un mélanome a une forme irrégulière.
- **B** (Bords irréguliers) : les mélanomes présentent souvent des contours flous, dentelés ou mal définis, contrairement aux grains de beauté normaux aux bords lisses.
- **C** (Couleur non homogène) : une lésion suspecte peut contenir plusieurs nuances (brun, noir, rouge, blanc, voire bleu), ce qui est un signe d'alerte.
- **D** (Diamètre supérieur à 6 mm) : les mélanomes sont souvent plus grands que 6 mm, bien que certaines formes agressives puissent être plus petites.
- **E** (Évolution rapide) : un grain de beauté qui change d'apparence (taille, couleur, forme) en quelques semaines ou mois doit être examiné.

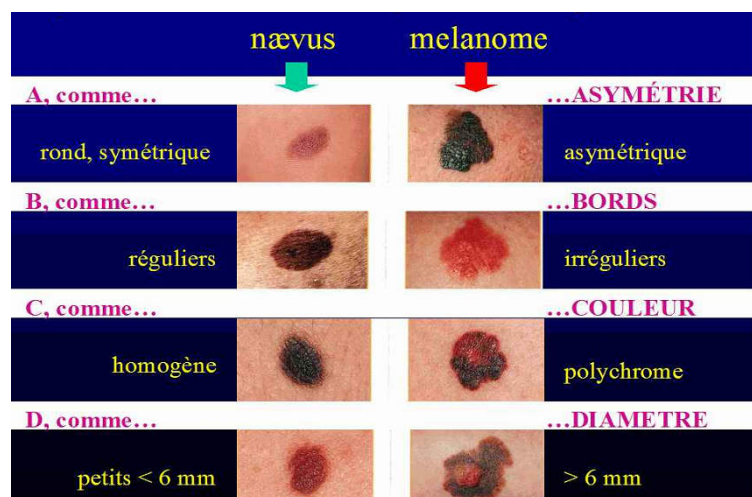


FIGURE 2.11 – Méthode **ABCDE** pour la détection du mélanome [41]

Ce système présente toutefois des limites, en particulier pour certains sous-types de mélanomes tels que le mélanome nodulaire, qui peut ne pas

suivre ces critères classiques [14]. Ainsi, la vigilance face à une lésion nouvelle ou évolutive reste un élément déterminant, quel que soit l'aspect initial.

Des critères additionnels, comme le « *vilain petit canard* » (lésion différente des autres sur le corps du patient), ou l'approche « IDAT » (Inquiétude du patient, Durée, Apparition, Transformation) peuvent également orienter la décision clinique [59, 27]. L'examen physique complet, incluant le cuir chevelu, les muqueuses, les zones acrales et sous-unguéales, est essentiel pour ne pas méconnaître une localisation atypique [27].

Diagnostic histopathologique

L'analyse histopathologique reste la pierre angulaire du diagnostic. Elle permet de confirmer la nature maligne de la lésion et d'en préciser le type, le stade et le potentiel évolutif [27, 32].

- **Mélanome superficiel extensif (SSM)** : le plus fréquent, caractérisé par une croissance radiale initiale suivie d'une invasion verticale.
- **Mélanome nodulaire (NM)** : forme agressive à croissance verticale rapide, souvent sans phase radiale.
- **Mélanome lentigineux acral (ALM)** : survenant sur les paumes, plantes et sous les ongles, fréquent chez les populations à peau foncée.
- **Mélanome de Dubreuilh (LMM)** : encore appelé Mélanome Lentigineux Malin, typiquement sur peau photo-exposée du visage chez le sujet âgé.

Les paramètres histopronostiques essentiels incluent l'épaisseur de *Breslow*, l'ulcération, l'index mitotique, la présence d'une invasion lymphocytaire et le niveau de Clark [27, 38]. Ces éléments conditionnent le stade selon la classification **AJCC** « *American Joint Commission on Cancer* » et influencent le traitement [27].

Diagnostic par imagerie et biologie moléculaire

L'imagerie non invasive et la biologie moléculaire constituent des outils complémentaires de plus en plus utilisés dans les cas de mélanomes avancés

ou atypiques que dans certains cas superficiels, avec pour but d'affiner les cas douteux, guider la prise en charge et orienter vers des traitements ciblés.

Imagerie non invasive : Certaines modalités d'imagerie avancée permettent une exploration in vivo de la lésion, avec une résolution quasi histologique.

- **Diagnostic dermoscopique :** la dermoscopie est une technique d'imagerie non invasive qui permet d'examiner les structures sous-épidermiques invisibles à l'œil nu (voir figure 2.12). Elle améliore significativement la sensibilité et la spécificité du diagnostic par rapport à l'examen clinique seul [27].



FIGURE 2.12 – Le dermoscope [21]

Les principaux critères en faveur d'un mélanome incluent un réseau pigmentaire atypique, des globules ou points irréguliers, la présence d'un voile bleu-blanc, de structures blanches brillantes visibles en lumière polarisée, ainsi que des zones de régression. La polychromie (plus de trois couleurs distinctes) et l'asymétrie de structure ou de couleur sur plusieurs axes renforcent également la suspicion [27, 38].

Ces manifestations varient selon le type de peau et la localisation. Sur peau claire, les motifs multicomposants ou réticulaires atypiques sont fortement évocateurs. Sur peau foncée, un motif parallèle aux crêtes est suspect de mélanome, contrairement au motif parallèle aux vallées observé dans les nævus bénins. Enfin, pour les lésions unguéales, l'élargissement irrégulier d'une bande pigmentée associé au signe de Hutchinson constitue un indicateur majeur [38].

- **La microscopie confocale à réflectance (RCM) :** la microscopie confocale à réflectance (RCM) est une technique d'imagerie non invasive offrant une visualisation cellulaire en temps réel de l'épiderme et du derme superficiel. Elle est particulièrement utile pour l'évaluation des lésions pigmentées du visage, des nævus atypiques ou des sites difficiles à biopsier [27]. Elle permet notamment de détecter des critères évocateurs de mélanome tels que l'atypie kératinocytaire, la présence de cellules dendritiques épidermiques atypiques ou encore des structures pagétoïdes [27, 1].

Le principe repose sur la réflexion de la lumière par les molécules endogènes de la peau. À l'aide d'un laser monochromatique et cohérent, combiné à un système optique complexe (lentilles, filtres, miroirs de balayage rapide) et à des photomultiplicateurs, la RCM capte exclusivement la lumière issue du plan focal, produisant ainsi des images horizontales de la peau avec une résolution cellulaire, depuis la couche cornée jusqu'au derme papillaire [32, 60].

Le système le plus utilisé, tel que le *Vivascope 1500* (figure 2.13), fonctionne avec un laser diode à 830 nm de faible puissance (5–10 mW), sans risque pour les tissus [1]. Certains modèles multi-lasers (488 nm, 658 nm, 785 nm) permettent également l'utilisation de colorants fluorescents afin d'améliorer le contraste des images [1].

- **Tomographie par cohérence optique (TCO) :** La TCO, ou tomographie optique cohérente, est une technique d'imagerie non invasive reposant sur l'interférométrie à faible cohérence [27]. Elle génère des

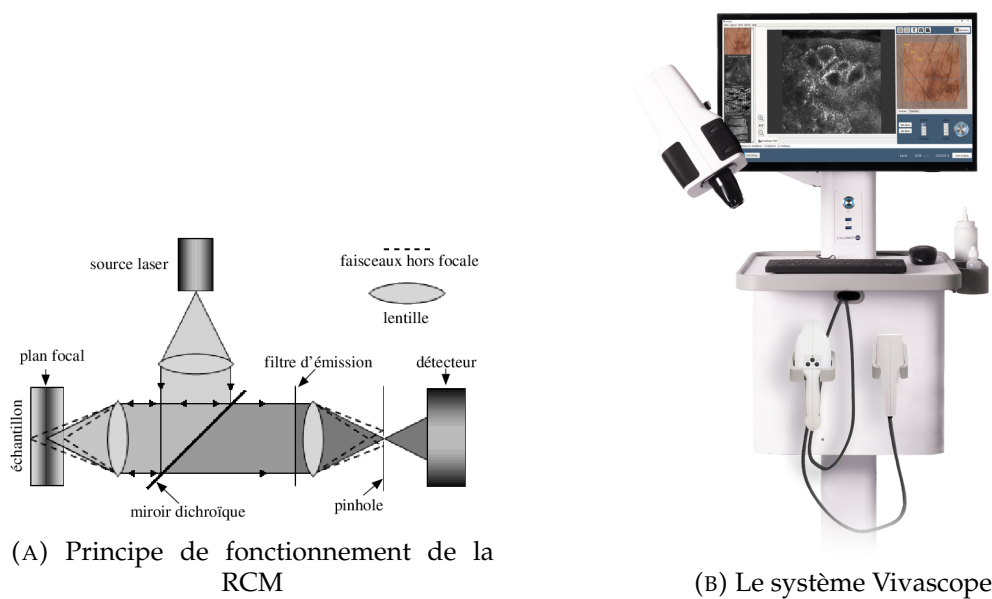


FIGURE 2.13 – Microscopie confocale à réflectance et dispositif associé [30, 60]

coupes verticales des tissus cutanés avec une résolution axiale d'environ $5\ \mu\text{m}$ et latérale proche de $10\ \mu\text{m}$, pour une profondeur de pénétration d'environ $1\ \text{mm}$ [27, 32]. Cette performance offre des images structurales quasi histologiques, comparables aux coupes histopathologiques traditionnelles.

Le principe consiste à diviser un faisceau lumineux à faible cohérence en deux voies : l'une dirigée vers la zone cutanée à examiner, l'autre vers un miroir de référence ajustable [32]. Les deux signaux réfléchis sont recombinaés et produisent une figure d'interférence uniquement si leurs trajets optiques sont équivalents dans la limite de la longueur de cohérence [32, 1]. Ce signal, capté par un photodétecteur, est converti en image en coupe. Le balayage en profondeur et le déplacement latéral du faisceau permettent ensuite de reconstruire des images bidimensionnelles de la peau.

Ces techniques non invasives sont particulièrement précieuses pour limiter les biopsies inutiles et optimiser les décisions cliniques, notamment dans les cas à localisation anatomique complexe ou à haut risque de cicatrice.

Biologie moléculaire et génomique : L'analyse moléculaire des lésions cutanées suspectes fournit des éléments supplémentaires pour le diagnostic différentiel et la stratification pronostique :

- Le gène **PRAME** (Preferentially Expressed Antigen in Melanoma), exprimé préférentiellement dans les cellules malignes, est un marqueur immunohistochimique prometteur[13]. Son expression est généralement absente dans les nævus bénins mais présente dans les mélanomes, ce qui en fait un outil diagnostique utile dans les cas histologiquement ambigus [13].
- Des panels multigéniques ou des signatures transcriptomiques sont en cours d'évaluation pour prédire le risque de récurrence ou la réponse au traitement [48]. Leur utilisation reste toutefois limitée à des contextes spécialisés ou à la recherche clinique.

Ces techniques, bien qu'efficaces pour le diagnostic du cancer de la peau, se heurtent encore à plusieurs difficultés, notamment le manque de personnel suffisamment formé et expérimenté dans l'utilisation d'outils allant de la dermoscopie aux méthodes d'imagerie et de biologie moléculaire. Afin de pallier ces limites et d'améliorer la rapidité ainsi que la fiabilité du diagnostic, de nombreuses recherches ont été menées pour mettre à disposition des jeux de données structurés. Ces bases de données ne servent pas uniquement à la formation des professionnels de santé, mais constituent également un support essentiel pour le développement de systèmes de diagnostic assistés par ordinateur, en particulier ceux basés sur la vision par ordinateur. Dans ce qui suit, nous présentons les principaux jeux de données disponibles et utilisés dans ce domaine.

2.2 Bases de données

Pour aborder efficacement le problème de la détection du cancer de la peau, l'utilisation de bases de données bien structurées est essentielle. Ces

bases de données, souvent issues de recherches antérieures et d'études empiriques, fournissent une base solide pour le développement et la validation des modèles de détection du cancer de la peau.

2.2.1 HAM10000 : *Human Against Machine with 10 000 training images* [61]

La base de données HAM10000¹ constitue une référence incontournable dans le domaine de la détection des lésions cutanées à l'aide de l'IA. Il regroupe environ 10,015 images dermoscopiques en couleur, standardisées au format 600*450 pixels. Les images proviennent de différentes sources et couvrent sept catégories de lésions cutanées couramment observées en dermatologie : le *carcinome kératinocytaire intraépithélial (akiec)*, le *carcinome basocellulaire (bcc)*, les *lésions kératosiques bénignes (bkl)*, le *dermatofibrome (df)*, les *nævi mélanocytaires (nv)*, les *lésions vasculaires (vasc)* ainsi que le *mélanome (mel)*, considérée comme la forme la plus critique de cancer de la peau. L'annotation de ces données a été validée par des experts dermatologues à l'aide des méthodes cliniques et histopathologiques, garantissant leur fiabilité scientifique. En plus des images brutes, la base de données inclut des masques de segmentation binaires. Ces masques, généralement au format PNG, représentent la région d'intérêt correspondant à la zone exacte de la lésion dans l'image. Chaque pixel du masque indique si la zone appartient ou non à la lésion. Toutefois, le jeu de données présente un déséquilibre marqué entre les classes : les nævi mélanocytaires représentent la majorité des échantillons, tandis que les mélanomes et certaines autres lésions malignes sont beaucoup plus rares. Ce déséquilibre constitue un défi pour les modèles d'apprentissage automatique.

Des exemples d'images des lésions de la base de données HAM10000 [61] sont présentés dans la figure 2.14

1. <https://doi.org/10.7910/DVN/DBW86T>

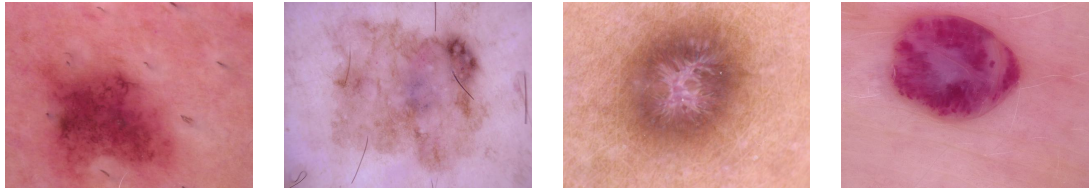


FIGURE 2.14 – Exemples de lésions dermoscopiques du jeu de données HAM10000

2.2.2 PH2 : Hôpital Pedro Hispano [43]

La base de données PH2² a été développée grâce à une collaboration entre l'*Universidade do Porto*, l'*Instituto Superior Técnico de Lisboa* et le service de dermatologie de l'*Hôpital Pedro Hispano* à Matosinhos, au Portugal. Elle contient 200 images dermoscopiques en couleur (résolution 768×560 pixels, RVB 8 bits), acquises dans des conditions contrôlées avec un grossissement de $20\times$.

Les images sont réparties en trois catégories principales, nævi communs, nævi atypiques et mélanomes, toutes issues de patients ayant un type de peau II ou III selon l'échelle de **Fitzpatrick** [1]. Chaque image est accompagnée de :

- des masques de segmentation manuels de la lésion cutanée, fournis par des dermatologues experts ;
- un diagnostic clinique (et histologique pour les cas suspects) ;
- une annotation des critères dermoscopiques les plus pertinents (asymétrie, couleurs, réseau pigmentaire, points/globules, stries, zones de régression, voile bleu-blanchâtre).

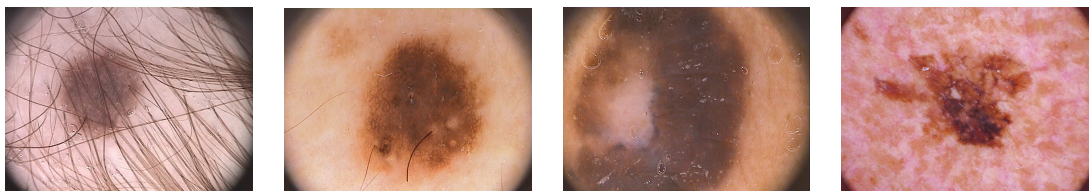


FIGURE 2.15 – Exemples de lésions dermoscopiques du jeu de données PH2

Grâce à la haute qualité des images et à la précision des annotations, PH2

2. <https://doi.org/10.7910/DVN/DBW86T>

constitue une base de référence pour l'évaluation des méthodes de segmentation et de classification des lésions cutanées. Elle est toutefois confrontée à un déséquilibre entre les différentes classes d'images. La figure 2.15 donne un aperçu des images présentes dans cette base de données.

2.2.3 ISIC : International Skin Imaging Collaboration [31, 19, 18, 51]

L'International Skin Imaging Collaboration (ISIC)³ constitue aujourd'hui une initiative de référence pour la mise à disposition de bases de données publiques en dermatologie numérique. Depuis 2016, cette collaboration internationale organise chaque année les ISIC Challenges, compétitions visant à standardiser l'évaluation des méthodes d'IA appliquées à la détection du cancer de la peau. Les jeux de données proposés rassemblent des images dermoscopiques issues de multiples institutions cliniques et annotées par des dermatologues experts [31, 19, 18, 51].

Au fil des éditions, la taille et la complexité des jeux de données ISIC ont considérablement augmenté. L'édition ISIC 2016 proposait environ 900 images et portait sur la segmentation et la classification binaire des lésions cutanées [31]. L'édition 2017 élargit ce corpus à environ 2 000 images avec l'introduction d'une classification en trois classes diagnostiques [19]. L'ISIC 2018 Challenge, avec plus de 10 000 images réparties en 7 catégories, constituera une étape majeure vers des benchmarks de grande échelle et donnera naissance au jeu de données HAM10000 [61].

Plus récemment, l'édition ISIC-2019 et ISIC-2020, avec respectivement plus de 25 000 et 33 000 images chacune, ont continué à intégrer HAM10000 en complément d'autres sources d'images, confirmant son rôle central de référence dans le domaine [18]. L'ISIC 2021, regroupant plus de 70 000 images, confirme cette dynamique d'expansion et d'intégration multisources [51]. La figure 2.16 donne un aperçu des images et des masques de la base de données ISIC.

3. <https://challenge.isic-archive.com/data/>

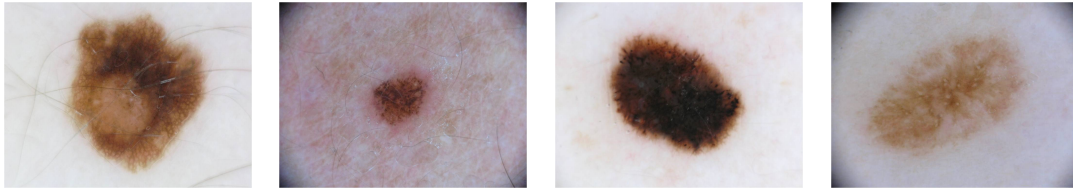


FIGURE 2.16 – Exemples d’images dermoscopiques issues de la base de données ISIC

2.2.4 Dermatology Information System (DermIS) [22]

La base de données DermIS⁴ (Dermatology Information System) provient d’une initiative clinique de l’Université de Heidelberg en Allemagne. Elle propose un vaste répertoire d’images dermatologiques couvrant différentes pathologies cutanées, notamment les lésions pigmentées [22]. Environ 9 000 images sont accessibles en ligne, dont plusieurs centaines de lésions dermoscopiques. Ces images sont de haute qualité clinique et accompagnées de diagnostics confirmés par des dermatologues. Cependant, la base ne fournit que quelques images dermoscopiques annotées avec leurs masques de segmentation (voir figure 2.17), regroupées en deux classes principales : mélanome et non-mélanome. Ainsi, son exploitation est particulièrement adaptée aux tâches de classification binaire (bénin vs. malin), mais reste limitée pour les approches modernes de segmentation profonde en raison du faible volume d’annotations disponibles.



FIGURE 2.17 – Exemples d’images de la base de données DermIS

4. <http://www.dermIS.com/>

2.2.5 Base de données DermQuest [23]

La base de données DermQuest⁵ a été développée par une plateforme collaborative dédiée à la formation médicale et au partage de cas dermatologiques. Elle regroupe plus de 20 000 images de diverses pathologies cutanées et met un accent particulier sur les cancers de la peau, notamment le mélanome. Les images dermoscopiques proposées sont de bonne qualité et accompagnées d’annotations, réalisées par des experts, couvrant une large diversité de cas cliniques [23]. Tout comme DermIS, DermQuest ne propose que quelques masques de segmentation (voir figure 2.18), ce qui limite son utilisation dans les tâches de segmentation avancée. En revanche, elle demeure une ressource précieuse pour les travaux de classification et de recherche clinique.



FIGURE 2.18 – Exemples d’images de la base de données DermQuest

2.2.6 Bases de données utilisées

Dans le cadre de notre étude, plusieurs bases de données de référence en dermatologie ont été exploitées afin d’entraîner et d’évaluer le système proposé. Ces bases de données présentent des caractéristiques complémentaires en termes de diversité des classes diagnostiques, de qualité des images et de disponibilité des annotations. Elles comprennent notamment :

- **HAM10000** : ce jeu de données constitue aujourd’hui l’une des ressources les plus utilisées dans le domaine. En plus de sa diversité et de la qualité des images, il fournit également des masques de segmentation, ce qui le rend particulièrement adapté à l’entraînement de modèles de segmentation et de classification robustes.

5. <http://vip.uwaterloo.ca/demos/skin-cancer-detection>

- **PH2** : cette base de données, bien que de taille plus réduite, se distingue par la qualité élevée de ses images dermoscopiques. Elle couvre trois classes principales de lésions cutanées et propose des masques de segmentation rigoureusement annotés par des experts, ce qui en fait une ressource précieuse pour l'évaluation de modèles sur des données de grande fiabilité.
- **DermIS et DermQuest** : ces deux bases, souvent mobilisées dans les travaux antérieurs, rassemblent un large corpus d'images dermoscopiques. Bien qu'elles ne distinguent généralement que deux classes (lésions bénignes et malignes), elles demeurent des ressources importantes pour les tâches de segmentation et de classification, en raison de la qualité des images proposées et de la disponibilité d'annotations exploitables.

Le tableau 2.1 décrit l'ensemble des jeux de données utilisés pour le développement du système que nous proposons dans cette recherche en fonction des classes d'images représentées et du nombre d'échantillons.

2.3 Méthodes et approches existantes

2.3.1 Système de segmentation des lésions cutanées

La segmentation des lésions cutanées est une étape clé pour la reconnaissance automatique des lésions dermoscopiques. Le but est de délimiter efficacement la lésion, malgré la variation des couleurs, des textures ou de forme, et les artefacts fréquents. Sa fiabilité repose sur des vérités de terrain annotées par des experts en dermatologie et sur les choix méthodologiques : des approches classiques (seuillage, détection des contours, méthodes régionales) aux modèles récents de CNNs et de ViT, nettement plus performants. Elle constitue ainsi le socle des tâches ultérieures de classification et de diagnostic assisté par ordinateur.

Wen Yin *et al.* [64] ont proposé une méthode de segmentation des lésions cutanées reposant sur une architecture U-Net interactive à double branche.

TABLE 2.1 – Description des jeux de données de lésions cutanées et de leurs distributions par classes

JEU DE DONNÉES	CLASSE	ÉCHANTILLONS	TOTAL
HAM10000	Naevus mélanocytaires (NV)	6705	10015
	Mélanome (MEL)	1113	
	Lésions kératosiques B (BKL)	1099	
	Carcinome basocellulaire (BCC)	514	
	Kératoses actiniques (AKIEC)	327	
	Lésions vasculaires (VASC)	142	
	Dermatofibrome (DF)	115	
PH²	Naevus commun (CNV)	80	200
	Naevus atypique (ANV)	80	
	Mélanome (MEL)	40	
DermIS	Mélanome (MEL)	43	69
	Non-mélanome (NMEL)	26	
DermQuest	Mélanome (MEL)	76	137
	Non-mélanome (NMEL)	61	

Cette approche intègre simultanément deux stratégies d'échantillonnage distinctes au sein du réseau, formant ainsi une structure verticale bifurquée. Par ailleurs, un nouveau bloc de convolution symétrique à double canal (DCS_Conv) est introduit en parallèle, conférant au réseau une structure horizontale également à double branche. Cette combinaison vise à accroître à la fois la profondeur et la largeur du modèle, favorisant une diversité et une richesse accrues des représentations multi-échelles. De plus, un module de fusion et de sélection résiduel (RFS) est intégré en sortie afin de permettre au modèle de se concentrer sur les caractéristiques les plus pertinentes, améliorant ainsi sa robustesse. Cependant, malgré ses apports notables, cette approche présente certaines limites. En particulier, sa capacité de généralisation demeure restreinte puisqu'elle a été évaluée sur des bases de données de taille relativement réduite. De plus, l'absence de recours à des techniques avancées d'augmentation de données limite la robustesse et la capacité du modèle à s'adapter à des scénarios cliniques variés.

Ahmed et al. [2] ont proposé une méthode hybride de détection et de segmentation des lésions de mélanome combinant RetinaNet et Mask R-CNN. Cette approche se concentre sur la région d'intérêt (ROI) correspondant à la lésion, tout en traitant les zones environnantes comme arrière-plan, ce qui réduit l'influence des artefacts cliniques. Le réseau intègre un réseau de caractéristiques en pyramide (FPN) afin de préserver les informations spatiales et d'améliorer la précision de la délimitation. L'architecture associe un module de classification, qui estime la probabilité de chaque région candidate, et un module de régression, qui affine la localisation des lésions.

Araujo et al. [6] ont proposé une méthode de segmentation des lésions cutanées combinant le détecteur YOLO et l'algorithme GrabCut. Le pipeline comporte quatre étapes principales : (i) élimination des artefacts (notamment les poils), (ii) détection de la localisation de la lésion, (iii) segmentation du contour à l'aide de GrabCut et (iv) post-traitement morphologique. Une telle fusion permet de garantir une détection rapide et une précision élevée. Toutefois, cette approche présente des limites, notamment une dépendance au bon fonctionnement de l'étape de détection initiale avec YOLO, mais aussi une sensibilité aux variations des contrastes et aux lésions mal définies.

Mingzhe Hu et al [36] ont proposé une approche qui exploite le Segment Anything Model (SAM) s'appuyant sur un ViT pré-entraîné avec un Masked Autoencoder (MAE), capable de traiter des images haute résolution (1024×1024) et de capturer à la fois des détails fins et des dépendances globales. Cette segmentation est guidée par des prompts (points, boîtes englobantes, croquis) encodés puis fusionnés avec les représentations visuelles, avant que le mask decoder ne génère le masque final. Cette approche offre une segmentation flexible et précise, adaptée aux images médicales. Toutefois, ses performances dépendent fortement de la qualité des prompts et de la diversité des données d'entraînement, ce qui peut limiter sa généralisation à certains contextes cliniques.

Bi et al [12] ont proposé une architecture basée sur les FCN optimisée pour la segmentation automatique des lésions cutanées. Leur approche intègre des étapes séquentielles ainsi qu'un mécanisme d'intégration parallèle,

exploitant la capacité des FCN à apprendre à différents niveaux d'abstraction. Cette conception permet de mieux traiter les défis liés aux lésions présentant des caractéristiques ambiguës. La combinaison entre segmentation multi-étapes et intégration parallèle favorise ainsi une localisation précise et une délimitation plus nette des contours, même dans des cas complexes.

2.3.2 Système de classification des lésions cutanées

La classification des lésions cutanées représente l'étape ultime du processus de détection du cancer de la peau. Elle a pour objectif d'identifier si une lésion est bénigne ou maligne, voire d'en préciser la nature. Ce domaine a connu une évolution majeure, passant des méthodes classiques d'extraction manuelle de caractéristiques et de classification traditionnelle, aux approches fondées sur les CNN pré-entraînés, puis plus récemment aux modèles de type ViT. Ces progrès ont considérablement renforcé la précision, la robustesse et la capacité de généralisation des systèmes de diagnostic assisté par ordinateur.

Ashraf et al.[7] ont proposé une approche basée sur la ROI afin de discriminer le mélanome du nævus en s'appuyant sur l'apprentissage par transfert via le CNN AlexNet. Un algorithme k-means amélioré est utilisé pour extraire automatiquement les ROI, ce qui permet de mettre en évidence les caractéristiques discriminantes. Par la suite, un modèle CNN pré-entraîné AlexNet, associé à une stratégie d'augmentation de données, est appliqué aux images ROI. Toutefois, cette méthode peut souffrir de limitations liées au surapprentissage induit par l'augmentation des données et à la généralisation restreinte sur des ensembles plus variés.

Alwakid et al [7] ont proposé une approche basée sur le DL pour améliorer la précision de la détection des zones lésionnelles. Le processus débute par une étape de rehaussement d'image à l'aide de l'ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), qui permet d'améliorer la qualité visuelle et de conserver davantage de détails pertinents. Par la suite, une segmentation est appliquée afin d'extraire automatiquement

les ROI. Une stratégie d’augmentation des données est ensuite utilisée pour compenser le déséquilibre des classes et enrichir la diversité des échantillons. Enfin, l’analyse et la classification des lésions cutanées reposent sur un CNN associé à une version modifiée de ResNet-50. Cette approche, bien que performante, présente des limites, notamment la complexité computationnelle liée à l’utilisation du GAN et les techniques d’augmentation de données.

Kemal Polat et al [47] ont proposé deux stratégies pour la classification automatique des maladies cutanées à partir d’images dermatologiques. La première repose sur un CNN appliqué directement aux images normalisées, permettant une classification multiclassées. La seconde améliore cette approche en combinant le CNN avec une stratégie One-Versus-All (OVA), où chaque classe est distinguée par rapport à toutes les autres via des modèles binaires, puis fusionnée pour obtenir une prédiction finale. Cette combinaison permet de mieux gérer la complexité interclasses et d’améliorer la robustesse du système. Toutefois, l’absence de prétraitements avancés peut limiter la capacité du modèle à exploiter pleinement les particularités des images médicales.

Aminur Rab et al. [50] ont proposé un modèle de classification des lésions cutanées reposant sur le TL en exploitant un CNN pré-entraîné de type VGG-19. Afin de renforcer la robustesse du système, plusieurs techniques de prétraitement ont été mises en œuvre, notamment le redimensionnement des images pour les adapter au format d’entrée du réseau, ainsi que l’augmentation des données destinée à pallier le manque d’images et le déséquilibre interclasses. Bien que ce modèle ait montré des performances encourageantes, il présente encore un potentiel d’amélioration, notamment par l’intégration de techniques avancées de prétraitement telles que la segmentation des ROI ou la réduction du bruit, susceptibles d’optimiser davantage la précision de la classification.

Kirill et al. [56] introduisent une approche d’apprentissage auto-supervisé séquentiel appliquée à la classification des lésions cutanées. Trois tâches prétextes distinctes — Relative Location, ODC et MoCo v2 [56] — sont mises en œuvre à l’aide de l’architecture ResNet-50, puis évaluées individuellement

avant d'être combinées selon différents ordonnancements (curriculum, anti-curriculum, mixed). Le modèle final est ensuite initialisé avec les poids issus du dernier prétexte de la séquence. Les résultats expérimentaux montrent que les curricula ordonnés par performances croissantes (ODC $\rightarrow$ RL $\rightarrow$ MoCo v2) produisent les meilleures performances, validant l'hypothèse selon laquelle l'ordre d'apprentissage des tâches prétextes influence de manière significative la qualité des représentations et la précision de classification. Cependant, bien que prometteuse, cette approche présente certaines limites méthodologiques. En particulier, l'absence de techniques d'augmentation des données peut accroître le risque de surapprentissage, et le non-recours à des méthodes de prétraitement classiques restreint le potentiel d'amélioration de la robustesse et de la précision du modèle.

2.4 Bilan et discussion

Afin de synthétiser les différentes études examinées dans cette revue de littérature, nous présentons ci-dessous un tableau récapitulatif pour chaque section, regroupant les travaux récents sélectionnés et mettant en évidence les auteurs, les méthodes utilisées, les bases de données exploitées et les taux d'exactitude obtenus. Ce résumé nous permet de faire une analyse générale des approches de segmentation et de classification existantes, mais aussi d'identifier les tendances actuelles et les lacunes des méthodes proposées.

Dans le tableau 2.2, nous présentons un bilan récapitulatif des travaux récents consacrés à la tâche de segmentation des lésions cutanées dans les images dermatologiques. Ce tableau met en évidence les méthodes classiques et profondes de segmentation appliquées sur différentes bases de données de référence, ainsi que leurs performances mesurées selon divers indicateurs tels que l'exactitude (ACC) et le Coefficient de similarité de Dice (DSC).

Dans le tableau 2.3, nous présentons un bilan récapitulatif des travaux récents consacrés à la tâche de classification des lésions cutanées. Ce tableau

TABLE 2.2 – Comparaison des performances des méthodes de segmentation sur différentes bases de données dermatologiques

DONNÉES	RÉFÉRENCE	ANNÉE	MÉTHODE	ACC. (%)	Dsc. (%)
HAM10000	Mingzhe et al. [36]	2023	SkinSAM	94.5	88.79
	Badawi et al. [8]	2024	U-Net	96.14	92.55
	Himel et al. [34]	2024	SkinSAM	97.00	94.40
	Monica et al. [44]	2024	Otsu Thresholding	–	92.00
	Anup et al. [39]	2022	U-RP Net	81.50	–
PH ²	Wen Yin et al. [64]	2023	DI-U-Net	94.34	93.48
	Noor Ahmed et al. [2]	2023	Mask-RCNN	94.30	–
	Halil et al. [62]	2019	YOLO+GrabCut	92.99	88.13
	Manu et al. [29]	2020	Mask-RCNN+DeeplabV3+	94.10	87.10
	Araujo et al. [6]	2021	U-NET	–	93.30
DermIS	Araujo et al. [6]	2021	U-Net	–	87.20
	Elineide et al. [54]	2020	SLICO	97.10	83.40
	Khan et al. [37]	2019	K-Means	94.00	–
	Santhosh et al. [53]	2021	ACS	–	86.26
	Chakkara et al. [17]	2020	Watershed+Sobel	89.60	–
DermQuest	Santhosh et al. [53]	2021	ACS	–	86.83
	Narayan et al. [45]	2024	CNN	91.40	–
	Chakkara et al. [17]	2020	Watershed+Sobel	95.30	–
	Ranjita et al. [52]	2020	Seuillage Global	98.20	80.20

met en évidence les méthodes classiques et profondes de classification appliquées à différentes bases de données de référence, ainsi que leur taux d'exactitude.

Les performances obtenues dans la littérature montrent une amélioration constante de la précision des modèles de classification et de segmentation au fil des travaux. Les approches récentes fondées sur l'apprentissage automatique et profond ont permis de renforcer la capacité des systèmes à extraire les ROI, mais aussi à distinguer efficacement les lésions bénignes des lésions malignes. Ces résultats traduisent une avancée significative dans la fiabilité des outils d'aide au diagnostic dermatologique, confirmant le potentiel des techniques d'IA pour automatiser et affiner l'analyse des images cutanées.

Cependant, malgré ces progrès, d'importantes limites subsistent. Les écarts

TABLE 2.3 – Comparaison des performances des méthodes de classification sur différentes bases de données dermatologiques.

DONNÉES	RÉFÉRENCE	ANNÉE	MÉTHODE	ACC (%)
HAM10000	Alwakid et al. [4]	2022	ESRGAN+ResNet50	84.00
	Aminur et al. [50]	2020	MobileNet	89.81
	Kemal et al. [47]	2020	CNN	77.00
	Carcagni et al. [15]	2019	DenseNet	88.00
	Kirill et al. [56]	2021	RL+ODC+MoCov2	82.30
PH ²	Zghal et al. [66]	2020	ABCD	90.00
	Wang et al. [63]	2023	SSD-KD	86.65
	Lattoofi et al. [40]	2019	ABCD	93.20
	Sena et al. [55]	2021	LBP+KNN	96.05
DermIS	Ashraf et al. [7]	2020	K-Means+AlexNet	97.90
	Khan et al. [37]	2019	Espace RVB + SVM	96.00
	Gamil et al. [26]	2024	EfficientNetB0+AdaBoost	93.00
	Narayan et al. [45]	2024	K-Means+CNN	70.40
DermQuest	Ashraf et al. [7]	2020	K-Means+AlexNet	97.40
	Narayan et al. [45]	2024	K-Means+CNN	91.40
	Hemalatha et al. [33]	2023	InceptionV3	84.00
	Hosny et al. [35]	2019	AlexNet	97.70

observés entre les taux de précision indiquent une forte dépendance à la qualité et à l'équilibre des données utilisées, ainsi qu'à la stratégie de prétraitement adoptée. Certains modèles restent sensibles au surapprentissage, tandis que leur capacité de généralisation demeure restreinte face à la diversité des types de lésions et aux variations d'éclairage, de texture ou de teinte de peau. Ces insuffisances réduisent encore la robustesse clinique et la reproductibilité des résultats obtenus.

Dans cette optique, la méthode proposée dans ce travail s'inscrit dans une logique d'optimisation des performances globales à travers une approche en deux étapes complémentaires. Elle repose sur une segmentation préalable destinée à isoler avec précision la région d'intérêt, suivie d'une classification

à deux phases étroitement liées, où la première oriente la seconde pour renforcer la cohérence décisionnelle. Cette stratégie intégrée vise à améliorer la qualité de l'apprentissage, à limiter la perte d'information contextuelle et à offrir un système plus stable et plus fiable pour la détection automatique des lésions cutanées.

Chapitre 3

Titre :

Segmentation robuste multi-bases de données des lésions cutanées à l'aide d'un U-Net régularisé entraîné à partir de zéro avec une stratégie de prétraitement améliorée.

Résumé - La segmentation précise des lésions cutanées constitue une étape de prétraitement essentielle pour les systèmes de diagnostic assisté par ordinateur (CAD) en dermatologie. Cependant, les images dermoscopiques présentent des défis importants en raison de la variabilité des lésions en termes de taille, de forme, de texture et de conditions d'illumination, ainsi que de la présence d'artefacts tels que les poils et le bruit. De plus, la disponibilité limitée de jeux de données médicaux annotés restreint souvent la capacité de généralisation des modèles dans différents contextes cliniques. Cette étude propose un cadre entièrement automatique de segmentation des lésions cutanées basé sur une architecture U-Net régularisée entraînée à partir de zéro, combinée à une chaîne de prétraitement sensible au bruit. La stratégie d'amélioration du signal proposée intègre un filtrage gaussien pour la réduction des artefacts, une normalisation de l'intensité, un redimensionnement standardisé, ainsi qu'une augmentation de données géométrique et photométrique symétrique afin d'atténuer le déséquilibre des classes et d'améliorer la robustesse du modèle. Le modèle a été évalué sur quatre ensembles de données dermoscopiques publics (HAM10000, PH2, DermIS et DermQuest), ainsi que sur un ensemble de données global composite construit à partir de leur agrégation afin d'évaluer la capacité de généralisation inter-bases de données. Les résultats expérimentaux démontrent des performances de segmentation cohérentes et compétitives sur l'ensemble des jeux de données, avec des scores Dice atteignant 92,54 %, un IoU pouvant atteindre 87,40 % et une précision allant jusqu'à 99,14 %. Le modèle maintient également une sensibilité et une spécificité élevées, confirmant une délimitation fiable des contours des lésions tout en minimisant les fausses détections. Les résultats obtenus mettent en évidence la robustesse et la capacité de généralisation du cadre proposé, démontrant ainsi sa pertinence en tant que module de prétraitement efficace pour les systèmes automatisés d'analyse du cancer de la peau et pour de futures applications d'aide à la décision clinique.

Mots clés :

Segmentation des lésions cutanées, Analyse dermoscopique, Traitement d'images biomédicales, Architecture U-Net, Apprentissage profond, Généralisation inter-bases de données, Diagnostic assisté par ordinateur.

Contributions associées :

Z. Mounmie, Y. Yaddaden, A. Bouzouane, "Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing" in *Biomedical Signal Processing and Control (BSPC-D-26-03612)*. (Statut : Soumis)

Biomedical Signal Processing and Control

Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing

--Manuscript Draft--

Manuscript Number:	BSPC-D-26-03612
Article Type:	Research Paper
Keywords:	Skin lesion segmentation; Dermoscopic image analysis; Biomedical image processing; U-Net architecture; Deep learning; Cross-dataset generalization; Computer-aided diagnosis
Corresponding Author:	Yacine Yaddaden University of Quebec at Rimouski
First Author:	Moumnie Zakariaou
Order of Authors:	Moumnie Zakariaou
	Yacine Yaddaden
	Abdenour Bouzouane
Abstract:	Accurate skin lesion segmentation is a critical preprocessing step for computer-aided diagnosis (CAD) systems in dermatology. However, dermoscopic images present significant challenges due to lesion variability in size, shape, texture, illumination conditions, and the presence of artifacts such as hair and noise. In addition, limited annotated medical datasets often restrict model generalization across diverse clinical settings. This study proposes a fully automatic skin lesion segmentation framework based on a regularized U-Net architecture trained from scratch, combined with a noise-aware preprocessing pipeline. The proposed signal enhancement strategy integrates Gaussian filtering for artifact reduction, intensity normalization, standardized resizing, and symmetric geometric and photometric data augmentation to mitigate class imbalance and improve robustness. The model was evaluated on four publicly available dermoscopic datasets (HAM10000, PH2, DermIS, and DermQuest), as well as on a composite global dataset constructed from their aggregation to assess cross-dataset generalization. Experimental results demonstrate consistent and competitive segmentation performance across all datasets, achieving Dice scores up to 92.54%, IoU up to 87.40%, and accuracy up to 99.14%. The model maintains high sensitivity and specificity, confirming reliable lesion boundary delineation while minimizing false detections. The obtained results highlight the robustness and generalization capability of the proposed framework, demonstrating its suitability as an effective preprocessing module for automated skin cancer analysis systems and future clinical decision-support applications.

Dear Editor,

We are pleased to submit our manuscript entitled “*Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing*” for consideration for publication in *Biomedical Signal Processing and Control*.

This work addresses an important challenge in dermatological computer-aided diagnosis systems: achieving reliable skin lesion segmentation under heterogeneous dermoscopic imaging conditions. Dermoscopic images often exhibit strong variability in illumination, lesion morphology, and acquisition artifacts, which can significantly affect segmentation accuracy and downstream diagnostic tasks. In addition, many recent approaches rely heavily on pretrained models or increasingly complex architectures, which may introduce domain bias and require substantial computational resources.

To address these challenges, we propose a robust segmentation framework that integrates a **signal-aware preprocessing pipeline** with a **dropout-regularized U-Net architecture trained from scratch**. The proposed preprocessing strategy combines Gaussian noise reduction, intensity normalization, spatial standardization, and controlled geometric and photometric data augmentation to improve signal consistency and segmentation robustness. By deliberately training the network from scratch, the framework promotes domain-specific feature learning and avoids potential biases introduced by natural-image pretraining.

The proposed method was evaluated on **four publicly available dermoscopic datasets (HAM10000, PH2, DermIS, and DermQuest)** as well as on an aggregated **global dataset** designed to assess cross-dataset generalization. Experimental results demonstrate stable and competitive performance across heterogeneous datasets, achieving **Dice scores up to 92.54% and IoU up to 87.40%**, while maintaining high sensitivity and specificity. Importantly, the proposed architecture remains computationally lightweight and can be trained using standard workstation hardware, making it suitable for practical biomedical imaging environments.

We believe this manuscript aligns well with the scope of *Biomedical Signal Processing and Control*, as it highlights the importance of **signal enhancement strategies combined with efficient deep learning architectures** for robust biomedical image segmentation. The study also emphasizes **cross-dataset robustness and computational efficiency**, which are key considerations for the development of reliable clinical decision-support systems.

This manuscript is original, has not been published previously, and is not under consideration for publication elsewhere. All authors have approved the manuscript and agree with its submission to *Biomedical Signal Processing and Control*.

Thank you for considering our manuscript. We look forward to your feedback.

Sincerely,

Prof. Yacine Yaddaden, Ph. D
Université du Québec à Rimouski
Campus de Lévis

Graphical Abstract

Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing

Zakariaou Moumnie, Yacine Yaddaden, Abdenour Bouzouane

Highlights

Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing

Zakariaou Moumnie, Yacine Yaddaden, Abdenour Bouzouane

- Noise-aware preprocessing pipeline for dermoscopic images
- Regularized U-Net trained from scratch for lesion segmentation
- Multi-dataset validation across four public benchmarks
- Dice up to 92.54% and IoU up to 87.40%
- Strong cross-dataset generalization with balanced sensitivity/specificity

Robust Multi-Dataset Skin Lesion Segmentation Using a Regularized U-Net Trained from Scratch with Enhanced Preprocessing

Zakariaou Moumnie^a, Yacine Yaddaden^{a,*}, Abdenour Bouzouane^b

^a *Université du Québec à Rimouski, 1595 Bd Alphonse-Desjardins, Lévis, G6V
0A6, QC, Canada*

^b *Université du Québec à Chicoutimi, 555 Bd de l'Université, Chicoutimi, G7H
2B1, QC, Canada*

Abstract

Accurate skin lesion segmentation is a critical preprocessing step for computer-aided diagnosis (CAD) systems in dermatology. However, dermoscopic images present significant challenges due to lesion variability in size, shape, texture, illumination conditions, and the presence of artifacts such as hair and noise. In addition, limited annotated medical datasets often restrict model generalization across diverse clinical settings.

This study proposes a fully automatic skin lesion segmentation framework based on a regularized U-Net architecture trained from scratch, combined with a noise-aware preprocessing pipeline. The proposed signal enhancement strategy integrates Gaussian filtering for artifact reduction, intensity normalization, standardized resizing, and symmetric geometric and photometric data augmentation to mitigate class imbalance and improve robustness.

The model was evaluated on four publicly available dermoscopic datasets (HAM10000, PH2, DermIS, and DermQuest), as well as on a composite global dataset constructed from their aggregation to assess cross-dataset generalization. Experimental results demonstrate consistent and competitive segmentation performance across all datasets, achieving Dice scores up to 92.54%, IoU up to 87.40%, and accuracy up to 99.14%. The model main-

*Corresponding author

Email addresses: zakariaou.moumnie@uqar.ca (Zakariaou Moumnie),
yacine_yaddaden@uqar.ca (Yacine Yaddaden), abouzoua@uqac.ca (Abdenour
Bouzouane)

tains high sensitivity and specificity, confirming reliable lesion boundary delineation while minimizing false detections.

The obtained results highlight the robustness and generalization capability of the proposed framework, demonstrating its suitability as an effective preprocessing module for automated skin cancer analysis systems and future clinical decision-support applications.

Keywords:

Skin lesion segmentation, Dermoscopic image analysis, Biomedical image processing, U-Net architecture, Deep learning, Cross-dataset generalization, Computer-aided diagnosis

1. Introduction

Skin cancer is among the most prevalent malignancies worldwide, with melanoma representing its most aggressive form. According to global cancer statistics, incidence rates continue to increase, reinforcing the need for reliable and automated diagnostic support systems [1]. Early and accurate lesion delineation plays a critical role in computer-aided diagnosis (CAD), as segmentation quality directly impacts downstream classification performance and malignancy assessment [2].

From a biomedical signal processing perspective, dermoscopic image segmentation remains a challenging task. Skin lesions exhibit significant intra- and inter-class variability in terms of color distribution, texture patterns, shape irregularity, and size. Moreover, dermoscopic images frequently contain acquisition-related artifacts such as illumination inhomogeneity, hair occlusion, low contrast, and sensor noise, which degrade boundary sharpness and reduce region separability [3, 4]. These signal-level degradations complicate precise region-of-interest (ROI) extraction and limit the robustness of traditional segmentation techniques.

Conventional image processing methods, including thresholding, watershed algorithms, region growing, and GrabCut-based approaches, have been explored for lesion segmentation [5]. Although computationally efficient, these techniques are often sensitive to contrast variations and require hand-crafted parameter tuning. Their performance typically degrades in heterogeneous clinical environments.

Recent advances in deep learning have significantly improved medical image segmentation performance. Convolutional Neural Networks (CNNs),

introduced by LeCun et al. [6], have demonstrated strong capabilities in hierarchical feature extraction for biomedical imaging [7]. Among them, the U-Net architecture proposed by Ronneberger et al. [8] has become a reference model for biomedical image segmentation due to its encoder–decoder structure and skip connections that preserve spatial information.

Several adaptations of U-Net and related architectures have been proposed for skin lesion segmentation. Fully Convolutional Networks (FCNs) were introduced for multi-stage dermoscopic segmentation [9]. Transfer learning strategies using pretrained encoders such as ResNet have shown improved performance compared to random initialization [10]. Dual-branch and attention-enhanced U-Net variants, such as DI-U-Net [11], aim to enhance multi-scale feature representation. More recently, transformer-based approaches and Segment Anything Model (SAM)-based frameworks have been explored for dermatological segmentation tasks [12]. Optimization-driven and hybrid architectures have also been proposed to improve lesion boundary delineation [13].

Despite these advances, several limitations remain. First, many state-of-the-art methods rely heavily on pretrained models trained on natural image datasets, which may introduce domain bias and reduce adaptability to dermoscopic imaging characteristics. Second, generalization across heterogeneous datasets remains insufficiently addressed [14], as numerous approaches are validated on limited or single-source datasets. Third, preprocessing strategies are often treated as secondary components, although signal enhancement (noise reduction, normalization, augmentation) plays a critical role in segmentation stability [15].

To address these limitations, this work proposes a fully automatic skin lesion segmentation framework based on a regularized U-Net architecture trained from scratch. Unlike many recent approaches relying on heavy pretrained architectures or transformer-based models, the proposed framework focuses on a lightweight yet robust segmentation strategy that prioritizes signal enhancement and domain-specific feature learning. This design choice enables stable cross-dataset generalization while maintaining low computational complexity, which is particularly desirable for practical biomedical imaging applications.

To rigorously assess cross-dataset generalization, the model is evaluated on four publicly available dermoscopic benchmarks—HAM10000 [16], PH2 [17], DermIS [18], and DermQuest [19]—as well as on a composite global dataset constructed by aggregating these sources. This multi-dataset valida-

tion framework enables comprehensive robustness analysis under heterogeneous imaging conditions.

The main contributions of this study can be summarized as follows:

- A signal enhancement-oriented preprocessing pipeline designed to improve dermoscopic segmentation robustness under heterogeneous acquisition conditions.
- A dropout-regularized U-Net architecture trained from scratch to reduce domain transfer bias.
- A comprehensive multi-dataset evaluation strategy to assess cross-domain robustness.
- A competitive segmentation performance across heterogeneous benchmarks using standard biomedical metrics.

The remainder of this paper is organized as follows: Section 2 reviews related work, Section 3 describes the proposed methodology, Section 4 details the experimental protocol, Section 5 presents the results and comparative analysis, and Section 6 concludes the paper with perspectives for clinical integration.

2. Background

2.1. Semantic segmentation

Semantic segmentation plays a central role in dermoscopic image analysis, as it enables precise delineation of lesion boundaries and reliable discrimination between pathological and healthy tissues. In CAD systems, segmentation quality directly influences downstream feature extraction and classification performance.

Among deep learning architectures, U-Net [8] has become a reference framework for biomedical image segmentation due to its ability to learn effective representations from limited annotated data. Its symmetric encoder-decoder structure consists of:

- an encoder (contracting path) that extracts hierarchical semantic features while reducing spatial resolution;

- a decoder (expansive path) that progressively reconstructs the segmentation map; and
- skip connections that preserve high-resolution spatial information and improve boundary localization.

Beyond architectural design, segmentation performance is strongly influenced by the data preprocessing pipeline, which directly affects signal quality and model generalization. Common preprocessing strategies include:

- Normalization, which standardizes pixel distributions to stabilize gradient-based optimization;
- Data augmentation, including geometric and photometric transformations to reduce overfitting and improve robustness;
- Spatial standardization (cropping/resizing), ensuring compatibility with network input constraints.

In this context, training from scratch represents a deliberate design choice to learn domain-specific representations without transfer bias from natural image datasets. This process involves:

- careful hyperparameter tuning (learning rate, batch size, regularization strength);
- regularization mechanisms such as Dropout, Weight Decay, and Early Stopping;
- appropriate weight initialization strategies to ensure stable convergence;
- the use of loss functions adapted to class imbalance, which is common in dermoscopic segmentation tasks.

These components jointly establish a robust framework for learning directly from dermatological data and form the methodological foundation of the proposed system.

2.2. Related Works

Skin lesion segmentation has been extensively studied using both classical image processing and deep learning approaches.

Early methods relied on handcrafted pipelines combining thresholding, morphological operations, and region-based segmentation [20]. Although computationally efficient, these techniques remain sensitive to illumination variability and poorly defined lesion boundaries.

With the advent of deep learning, Fully Convolutional Networks (FCNs) were introduced to enable end-to-end dermoscopic segmentation [9]. Subsequent research focused on improving boundary precision and multi-scale feature representation.

Transfer learning strategies were explored to leverage pretrained encoders. Tschandl et al. [10] demonstrated that initializing U-Net encoders with ImageNet-pretrained ResNet weights improves segmentation performance compared to random initialization. Similarly, Karri et al. [14] proposed a two-phase cross-domain transfer learning framework designed to enhance generalization across heterogeneous datasets.

More advanced architectures have been proposed to address feature diversity and contextual modeling. Yin et al. [11] introduced DI-UNet, a dual-branch architecture enhancing multi-scale representation. Kumar et al. [13] proposed U-RP-Net, combining dual segmentation networks with optimization-based fusion strategies. While effective, such approaches often increase computational complexity.

Recently, transformer-based models have gained attention in medical segmentation. SkinSAM [12] adapts the Segment Anything Model for dermoscopic images, while Himel et al. [21] explored Vision Transformers (ViT) for lesion segmentation and classification. Although promising, these approaches typically rely on pretrained models and substantial computational resources.

Hybrid detection–segmentation pipelines have also been explored. Araújo et al. [22] combined YOLO-based localization with GrabCut refinement, achieving competitive results but remaining sensitive to detection errors and contrast variability.

Despite these advances, several limitations persist:

1. **Dependence on pretrained models**, which may introduce domain bias;
2. **Limited cross-dataset validation**, restricting conclusions about generalization;

3. **Increased architectural complexity**, sometimes at the expense of interpretability and computational efficiency;
4. **Underemphasis on preprocessing robustness**, although signal enhancement strongly impacts segmentation stability.

These observations motivate the development of a streamlined yet robust segmentation framework that prioritizes domain-specific learning and systematic multi-dataset evaluation.

Unlike many recent works that focus on increasingly complex architectures, such as transformer-based or multi-branch segmentation networks, the objective of this study is to investigate whether a carefully designed signal-oriented preprocessing strategy combined with a regularized convolutional architecture can achieve competitive segmentation performance. This design choice prioritizes robustness, interpretability, and computational efficiency, which are important considerations for practical biomedical imaging applications where computational resources and training data may be limited.

3. Proposed system

The objective of this study is to design a robust and computationally efficient skin lesion segmentation framework capable of accurately delineating the ROI under heterogeneous dermoscopic acquisition conditions. The overall architecture of the proposed system is illustrated in Figure 1.

3.1. Preprocessing Strategy

Due to variability in illumination, acquisition devices, contrast levels, and the presence of artifacts (e.g., hair and minor noise), a carefully designed preprocessing stage is essential to improve signal consistency and enhance segmentation robustness.

The proposed preprocessing pipeline consists of three principal stages:

1. **Noise Reduction:** A Gaussian smoothing filter is applied to attenuate high-frequency noise while preserving lesion contours. This operation improves gradient stability during backpropagation and reduces sensitivity to minor intensity fluctuations. An ablation analysis was conducted to evaluate the impact of this step. Empirical observations during preliminary experiments suggested that Gaussian smoothing contributes to improved boundary regularity and training stability, particularly in low-contrast cases.

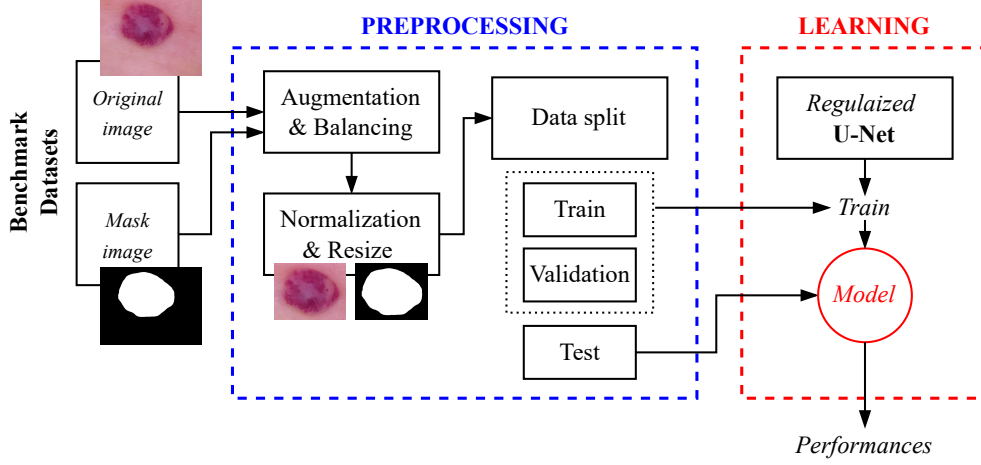


Figure 1: Overview of the proposed segmentation framework.

2. **Spatial Standardization:** All images are resized to 128×128 pixels to ensure homogeneous input dimensions and computational efficiency. The choice of this resolution was experimentally validated by comparing 128×128 and 256×256 configurations. Although higher resolution provided marginal improvements in boundary smoothness, it significantly increased memory consumption and training time without substantial gains in Dice or IoU performance. Therefore, 128×128 was retained as an optimal trade-off between accuracy and computational cost.
3. **Data Augmentation:** To mitigate overfitting and improve cross-dataset generalization, both geometric and photometric transformations are applied. Geometric transformations include controlled rotations, translations, horizontal and vertical flipping, and shearing operations to simulate positional variability. Photometric transformations include brightness modulation (80–120%), contrast adjustment, and mild Gaussian blur to emulate diverse acquisition conditions and enhance invariance to illumination changes.

Each dermoscopic image is paired with its corresponding binary segmentation mask to form image-mask training pairs. Tables 1, 2, 3 and 4 summarize dataset distributions before and after augmentation.

Table 1: Details about **HAM10000** dataset before and after data augmentation.

CLASS	SUBCLASS	BEFORE AUG.	AFTER AUG.
Benign lesions	Melanocytic nevi (NV)	6,705	2,000
	Benign keratosis-like (BKL)	1,099	2,000
	Vascular lesions (VASC)	142	1,778
	Dermatofibroma (DF)	115	1,656
Malignant lesions	Melanoma (MEL)	1,113	2,000
	Basal cell carcinoma (BCC)	514	2,000
	Actinic keratoses (AKIEC)	327	2,000
Total		10,015	13,434

Table 2: Details about **PH2** dataset before and after data augmentation.

CLASS	SUBCLASS	BEFORE AUG.	AFTER AUG.
Benign lesions	Common nevus (CNV)	80	400
	Atypical nevus (ANV)	80	400
Malignant lesions	Melanoma (MEL)	40	387
Total		200	1,187

Table 3: Details about **DermIS** dataset before and after data augmentation.

CLASS	BEFORE AUG.	AFTER AUG.
Melanoma (MEL)	43	400
Not melanoma (NMEL)	26	359
Total	69	759

Table 4: Details about **DermQuest** dataset before and after data augmentation.

CLASS	BEFORE AUG.	AFTER AUG.
Melanoma (MEL)	76	600
Not melanoma (NMEL)	61	600
Total	137	1,200

3.2. Model Architecture and Training Strategy

The proposed segmentation model is based on the U-Net architecture [8] (see Figure 2), selected for its effectiveness in biomedical image segmentation and its ability to preserve fine-grained spatial details through skip connections.

Input images are normalized to the range $[0, 1]$, while segmentation masks

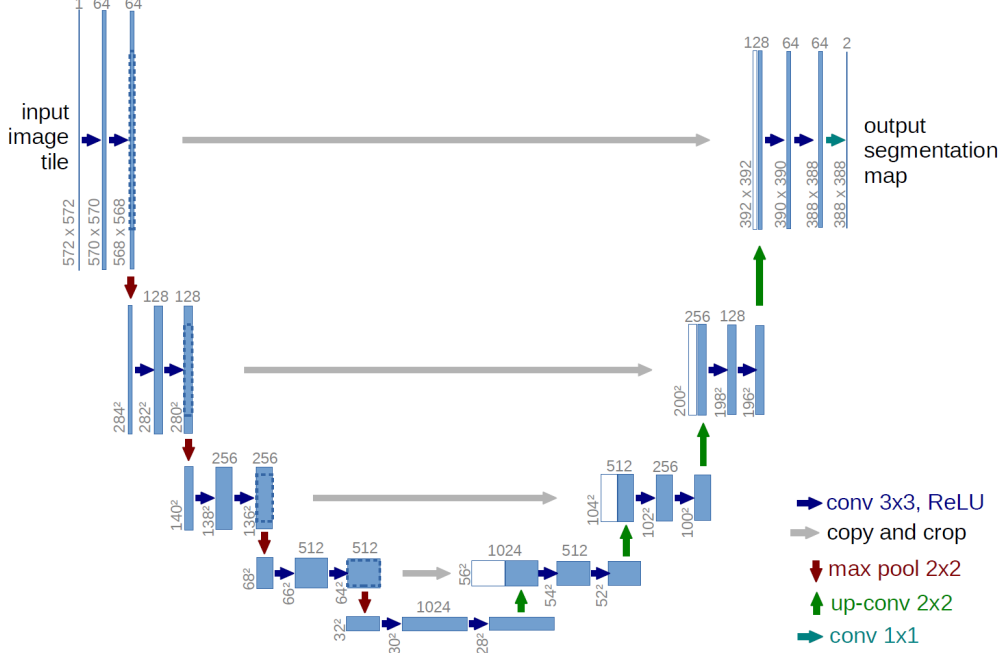


Figure 2: Overview of the **U-Net** architecture [8].

are binarized (background vs. lesion) to ensure stable optimization behavior and consistent learning dynamics.

Encoder: The encoder consists of successive convolutional blocks using 3×3 kernels with ReLU activation [23], followed by max-pooling layers for progressive hierarchical feature abstraction. Dropout layers (0.1–0.2) are integrated to reduce overfitting, particularly relevant in limited-size dermoscopic datasets.

Bottleneck: The bottleneck layer employs 256 filters with an increased dropout rate (0.3) to capture high-level contextual representations of lesion morphology and global structural information.

Decoder: The decoder reconstructs the segmentation map through transposed convolution operations combined with skip connections. Additional convolutional layers refine spatial details and improve boundary delineation. The final layer applies a 1×1 convolution followed by a sigmoid activation, producing a pixel-wise probability map.

Training Strategy: The model is trained from scratch to enable domain-specific feature learning and reduce potential bias introduced by natural-

image pretraining. Hyperparameters, including learning rate, batch size, and regularization strength, were systematically tuned to ensure stable convergence.

From a computational standpoint, the proposed architecture remains relatively lightweight compared to recent transformer-based segmentation models. Training at a resolution of 128×128 significantly reduces memory consumption and computational cost while maintaining competitive segmentation accuracy, making the framework suitable for practical biomedical image analysis scenarios.

4. Experimentation and Evaluation

This section describes the datasets used in this study, the evaluation metrics employed to assess segmentation performance, and the training and validation strategies adopted to ensure robust experimental evaluation.

4.1. Dataset Description

To evaluate the proposed skin lesion segmentation framework, four publicly available dermoscopic datasets were used: **HAM10000** [16], **PH2** [17], **DermIS** [18], and **DermQuest** [19]. These datasets represent diverse acquisition conditions and lesion distributions, making them suitable for assessing the robustness and generalization capability of segmentation models.

1. **HAM10000** dataset (Human Against Machine with 10000 training images) proposed by Tschandl *et al.* [16] contains **10,015** dermoscopic images of pigmented skin lesions belonging to seven diagnostic categories (see Tables 1, 2, 3 and 4). Due to its large scale and clinical diversity, HAM10000 is widely used for benchmarking automated dermoscopic analysis systems.
2. **PH2** dataset [17] contains **200** dermoscopic images acquired at the Pedro Hispano Hospital in Portugal. Each image is accompanied by expert annotations and clinical metadata, making PH2 a valuable benchmark for segmentation validation.
3. **DermIS** [18] is an online dermatology image repository containing annotated dermoscopic images. In this study, **69** images were used, including **43 melanomas** and **26 non-melanoma lesions**.
4. **DermQuest** [19] is a dermatological image database developed by Galderma. The subset used in this work contains **137 dermoscopic images**, including **76 melanomas** and **61 benign lesions**.

To further evaluate the generalization capability of the proposed model, a **global dataset** was constructed by aggregating all four datasets. This strategy introduces increased variability in acquisition conditions, lesion types, and imaging characteristics, enabling a more realistic assessment of segmentation robustness.

Figure 3 and Figure 4 illustrate examples of benign and malignant dermoscopic images along with their corresponding segmentation masks.

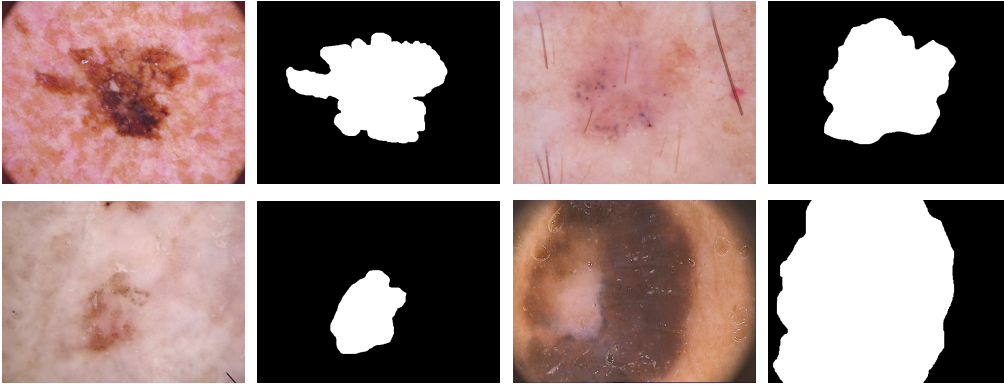


Figure 3: Overview of *malignant* lesion samples.

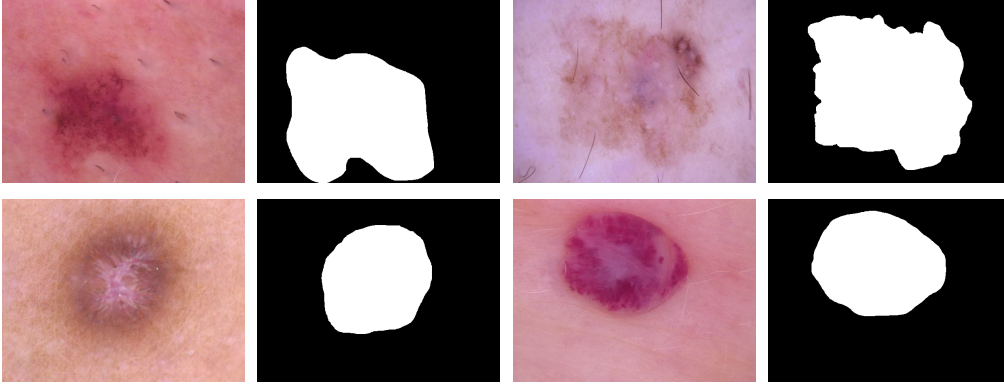


Figure 4: Overview of *benign* lesion samples.

4.2. Evaluation Metrics

Segmentation performance was evaluated using several widely adopted metrics in biomedical image segmentation.

1. **Dice Similarity Coefficient (DSC)**: Measures the overlap between predicted and ground truth segmentation masks (see Equation 1).
2. **Jaccard Index (Intersection over Union - IoU)**: Evaluates the similarity between predicted and reference masks (see Equation 2).
3. **Sensitivity (Recall / True Positive Rate)**: Quantifies the model's ability to correctly identify lesion pixels (see Equation 3).
4. **Specificity (True Negative Rate)**: Measures the model's ability to correctly identify background pixels (see Equation 4).
5. **Accuracy**: Represents the overall proportion of correctly classified pixels (see Equation 5).
6. **Structural Similarity Index (SSIM)**: Assesses structural similarity between predicted and reference masks (see Equation 6).
7. **Peak Signal-to-Noise Ratio (PSNR)**: Measures reconstruction fidelity between predicted and ground truth masks (see Equation 7).

The corresponding mathematical formulations are defined as follows:

$$\text{Dice} = \frac{2 |y_{\text{true}} \cap y_{\text{pred}}|}{|y_{\text{true}}| + |y_{\text{pred}}|}, \quad (1)$$

$$\text{IoU} = \frac{|y_{\text{true}} \cap y_{\text{pred}}|}{|y_{\text{true}} \cup y_{\text{pred}}|}, \quad (2)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (3)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad (4)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (5)$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (6)$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right). \quad (7)$$

4.3. Training and Validation Strategy

To ensure reliable experimental evaluation, the dataset was partitioned into three subsets: **70% for training**, **15% for validation**, and **15% for testing**. This configuration allows the model to learn robust representations while preventing overfitting through validation monitoring.

In addition to individual dataset experiments, the aggregated global dataset was also used to evaluate the robustness of the proposed model under heterogeneous data conditions.

The model was trained using the hyperparameters summarized in Table 5.

Experiments were conducted on a workstation equipped with an **Intel Core i7-9750H processor** (6 cores, 2.6 GHz), **32 GB RAM**, and a **64-bit operating system**. In addition to segmentation accuracy, computational efficiency is an important factor for biomedical image analysis systems. The proposed architecture remains lightweight compared with transformer-based or multi-branch segmentation networks. Training and inference can be performed on standard workstation hardware without requiring high-end GPU infrastructures, which improves the practicality of the method for clinical research environments.

Table 5: Training configuration and *hyperparameters*.

COMPONENT	CONFIGURATION
Optimizer	Adam [24]
Loss function	Binary Cross-Entropy (BCE)
Batch size	16
Epochs	100
Learning rate	0.001

5. Results and Discussion

5.1. Segmentation Results

The experimental results demonstrate the effectiveness of the proposed segmentation framework across multiple dermoscopic datasets. Overall, the model achieves consistently high performance, with accuracy exceeding 95.00% on all evaluated datasets and reaching 99.31% on DermQuest. Similarly, specificity remains very high, peaking at 99.76% on DermQuest, which indicates a strong ability to correctly identify background pixels and minimize false positives.

Sensitivity values also remain stable across datasets, reaching 93.65% on the Global Dataset and 91.11% on HAM10000. These results confirm the model’s reliability in correctly identifying lesion pixels under varying acquisition conditions.

Structural segmentation metrics further support these findings. The Dice Similarity Coefficient reaches 92.50% and the Intersection over Union (IoU) reaches 87.40% on the most representative datasets, indicating accurate lesion boundary delineation. Additionally, perceptual quality metrics show strong structural consistency between predicted masks and ground truth annotations, with SSIM values up to 97.17% and PSNR reaching 23.35 dB on DermQuest.

Although larger datasets such as HAM10000 (≈ 1579 s) and the Global Dataset (≈ 1890 s) require longer training times, they provide improved statistical stability and stronger generalization capability. In contrast, smaller datasets (DermIS, DermQuest, and PH2) require significantly shorter training times (less than 145 s) while still producing competitive segmentation results.

The complete quantitative results are summarized in Table 6.

Table 6: Detailed segmentation performance of the proposed model on validation and test sets.

METRIC	SPLIT	DERMIS	DERMQUEST	PH2	GLOBAL	HAM10000
Accuracy (%)	Test	95.97	99.14	96.44	96.54	96.97
	Validation	95.70	99.31	93.71	95.10	95.50
Specificity (%)	Test	98.50	99.75	99.10	98.78	98.25
	Validation	97.93	99.76	97.92	97.50	98.00
Recall (%)	Test	88.89	90.75	87.85	93.65	91.11
	Validation	84.32	88.47	89.42	92.20	90.80
PSNR (dB)	Test	17.51	22.57	16.07	17.54	17.70
	Validation	16.37	23.35	14.36	17.54	17.70
Loss	Test	0.03	0.01	0.03	0.04	0.03
	Validation	0.03	0.01	0.03	0.04	0.03
Time (s)	Test	105.32	122.23	145.21	1890.40	1579.43
	Validation	105.32	122.23	145.21	1890.40	1579.43
IoU (%)	Test	81.32	85.42	85.27	87.40	87.34
	Validation	77.80	83.16	86.34	86.90	86.90
DSC (%)	Test	89.27	91.88	91.93	92.50	92.54
	Validation	86.80	90.35	92.44	92.00	92.00
SSIM (%)	Test	89.15	96.46	88.98	90.00	89.90
	Validation	89.14	97.17	84.91	89.60	89.50

Legend: PSNR = Peak Signal-to-Noise Ratio, IoU = Intersection over Union, DSC = Dice Similarity Coefficient, SSIM = Structural Similarity Index.

5.2. Training Dynamics Analysis

The analysis of accuracy and loss curves provides insight into the influence of dataset size and diversity on the learning process.

For the HAM10000 dataset, which contains the largest number of samples, the model exhibits rapid convergence toward stable high accuracy (Figure 5). The corresponding loss curve shows a smooth and consistent decrease, indicating efficient optimization and strong generalization capability.

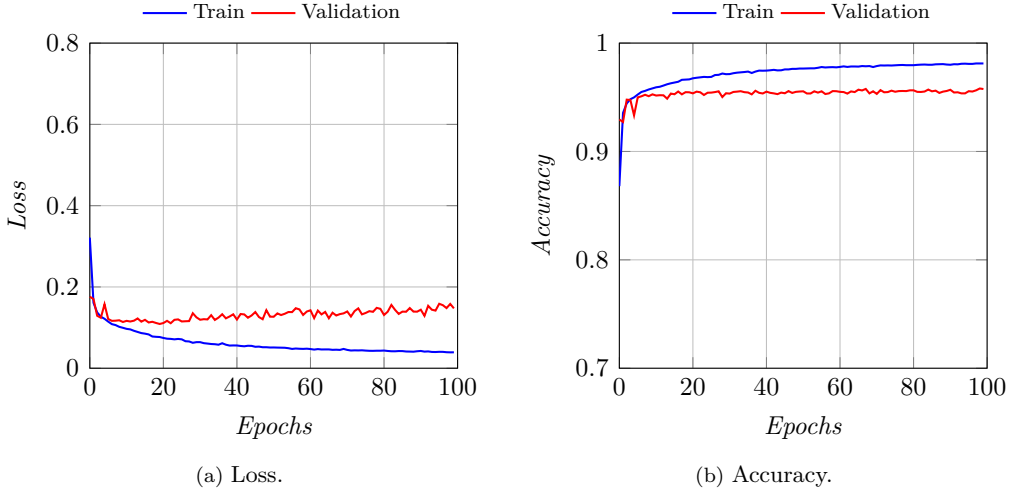


Figure 5: Training and validation loss and accuracy on the **HAM10000** dataset.

In contrast, smaller datasets such as PH2 and DermIS present more pronounced fluctuations in validation curves (Figure 6 and Figure 7). These variations are expected due to the limited number of samples and reduced variability in training data. Nevertheless, the model maintains stable accuracy levels, demonstrating its ability to extract discriminative features even under limited data conditions.

The DermQuest dataset shows extremely rapid convergence (Figure 8), likely due to the relative homogeneity of the dataset and the clearer lesion boundaries present in many images.

Finally, the Global Dataset (Figure 9), constructed by aggregating all individual datasets, benefits from the large statistical contribution of HAM10000 while incorporating additional diversity from smaller datasets. This combination improves the robustness of the training process and enhances cross-dataset generalization.

Figures 10 – 13 show qualitative segmentation results, including input images, ground truth masks, and predicted masks, for representative skin lesion samples from the **HAM10000**, **PH2**, **DermIS**, and **DermQuest**

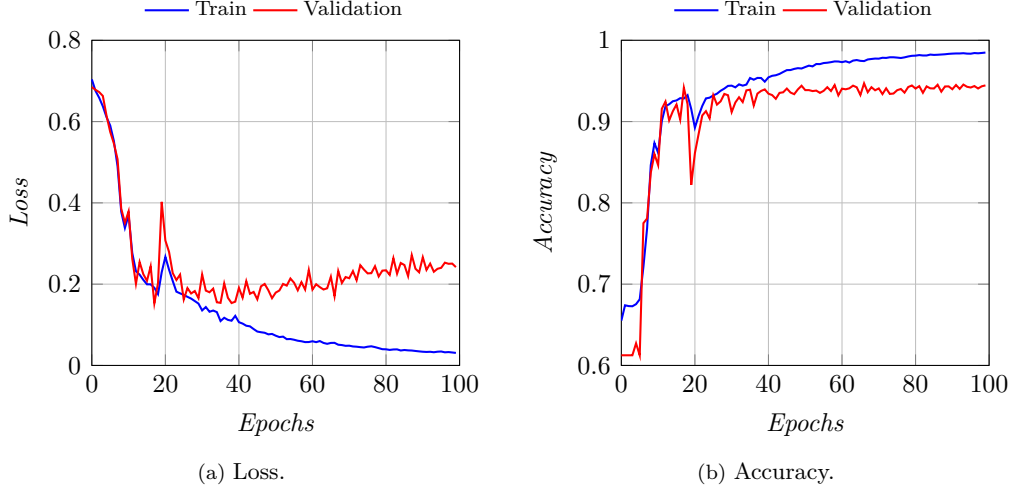


Figure 6: Training and validation loss and accuracy on the **PH2** dataset.

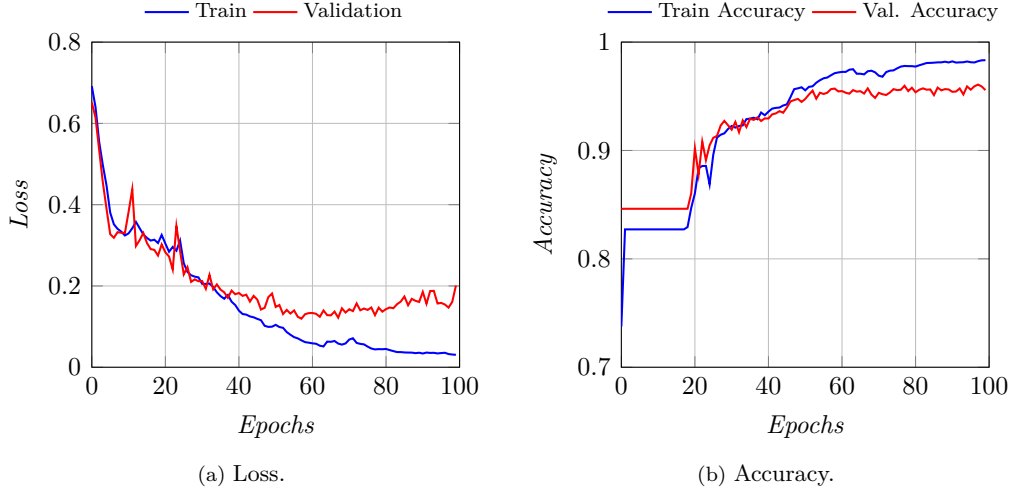


Figure 7: Training and validation loss and accuracy on the **DermIS** dataset.

datasets.

5.3. Comparison with State-of-the-Art Methods

Tables 7, 8, 9 and 10 present a comparative evaluation between the proposed method and several state-of-the-art approaches across different dermoscopic datasets.

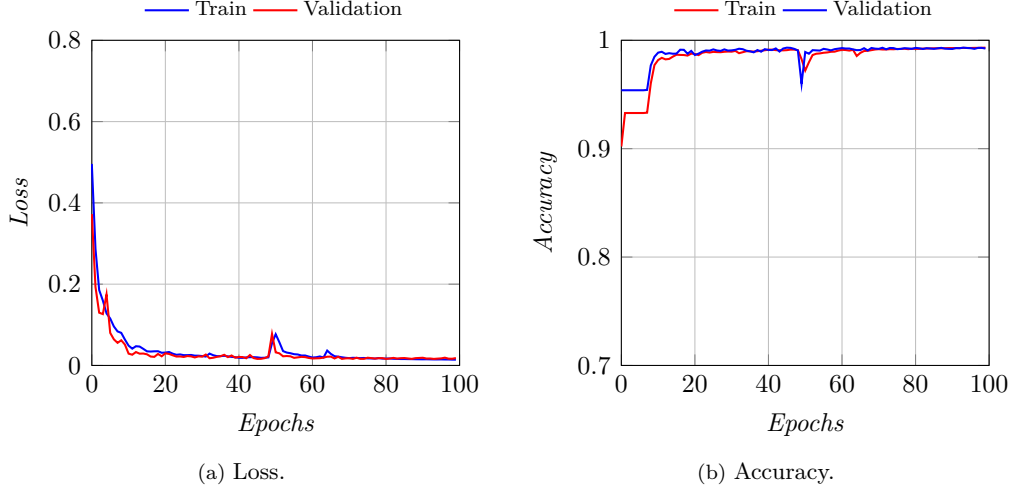


Figure 8: Training and validation loss and accuracy on the **DermQuest** dataset.

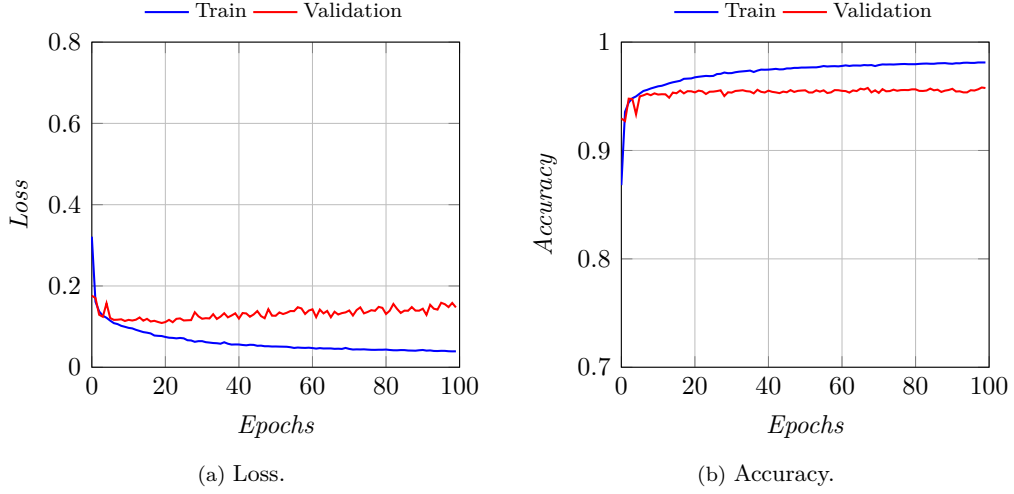


Figure 9: Training and validation loss and accuracy on the **Global** dataset.

For HAM10000 (see table 7), the proposed model achieves a Dice score of 92.54% and an IoU of 87.34%, exceeding the results reported by Monica et al. [25], Araújo et al. [22], and Tschandl et al. [10]. The obtained accuracy of 96.97% remains competitive with the approach of Karri et al. [14], which reports 97.80%.

On the PH2 dataset (see table 8), the proposed method achieves an ac-

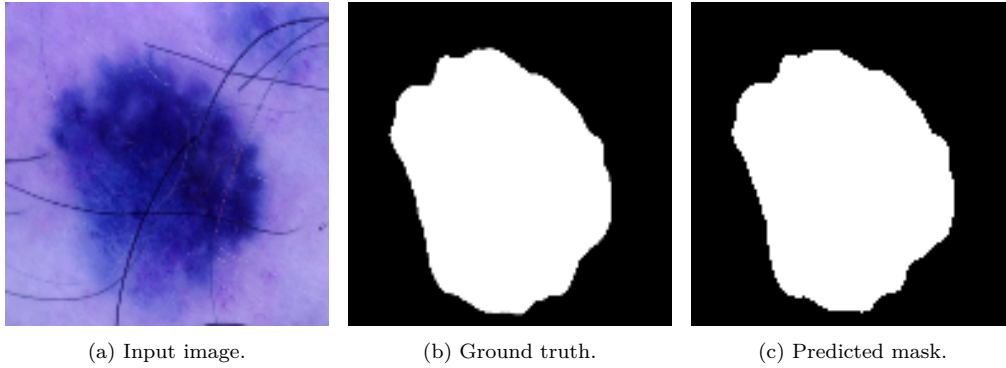


Figure 10: Qualitative segmentation results on the **HAM10000** dataset, showing the input image, ground truth mask, and predicted segmentation mask.

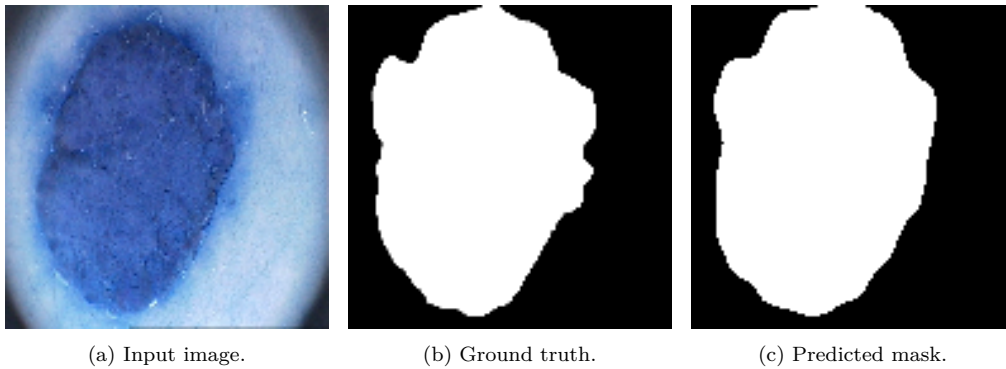


Figure 11: Qualitative segmentation results on the **PH2** dataset, showing the input image, ground truth mask, and predicted segmentation mask.

Table 7: Comparison of the proposed method with *state-of-the-art* segmentation methods on the **HAM10000** dataset.

REFERENCE	YEAR	METHOD	ACC	DSC	IoU
Karri et al. [14]	2023	nSknRSUNet	97.80%	94.00%	–
Hu et al. [12]	2023	SkinSam	94.50%	88.70%	78.40%
Monica et al. [25]	2024	Otsu	–	92.00%	80.00%
Tschandl et al. [10]	2019	U-Net-16	–	87.60%	80.00%
Araújo et al. [22]	2021	U-Net	–	91.30%	–
Proposed method	2026	<i>regularized</i> U-Net	96.97%	92.54%	87.34%

curacy of 96.44%, outperforming Saba et al. [2] (95.41%), Ahmed et al. [29] (94.30%), and Yin et al. [11] (94.34%). While DI-U-Net [11] reports a

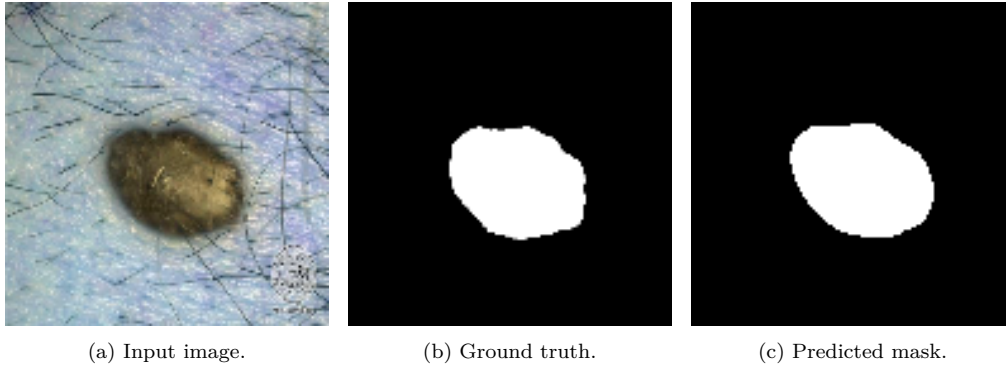


Figure 12: Qualitative segmentation results on the **DermIS** dataset, showing the input image, ground truth mask, and predicted segmentation mask.

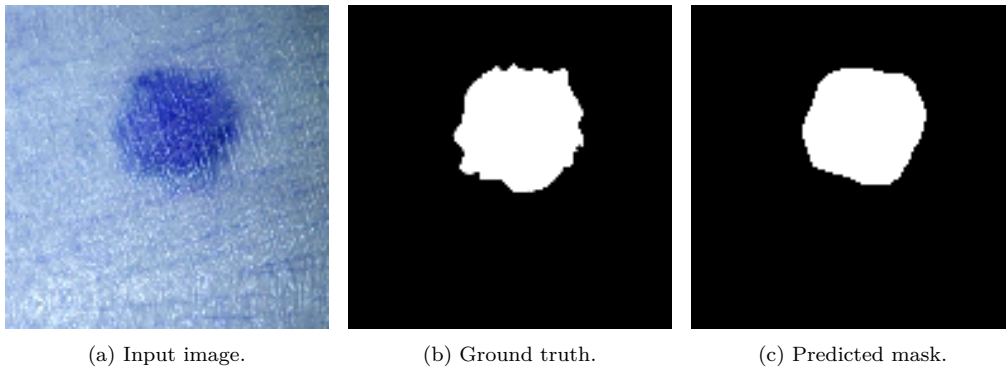


Figure 13: Qualitative segmentation results on the **DermQuest** dataset, showing the input image, ground truth mask, and predicted segmentation mask.

Table 8: Comparison of the proposed method with *state-of-the-art* segmentation methods on the **PH2** dataset.

REFERENCE	YEAR	METHOD	ACC	DSC	IoU
Saba et al. [2]	2019	Inception-V3	95.41%	–	–
Bi et al. [26]	2019	CNN	–	91.20%	85.00%
Duggani et al. [27]	2021	TL	83.63%	79.94%	79.94%
Bi et al. [28]	2017	FCN	90.66%	–	–
Ahmed et al. [29]	2023	Mask-RCNN	94.30%	–	–
Wen Yin et al. [11]	2023	DI-U-Net	94.34%	93.48%	–
Proposed method	2026	<i>regularized</i> U-Net	96.44%	91.93%	85.27%

slightly higher Dice score, the proposed framework provides more balanced performance across multiple metrics, particularly in terms of IoU.

Table 9: Comparison of the proposed method with *state-of-the-art* segmentation methods on the **DermIS** dataset.

REFERENCE	YEAR	METHOD	ACC	DSC	IoU
Araújo et al. [22]	2021	U-Net	–	87.20%	–
Chakkara et al. [30]	2020	WaterSheet+Sobel	89.60%	–	–
Santhosh et al. [31]	2021	ACS	–	86.26%	–
Proposed method	2026	<i>regularized</i> U-Net	95.97%	89.27%	81.32%

On DermIS (see table 9), the proposed method reaches 95.97% accuracy and a Dice score of 89.27%, outperforming previous approaches such as Chakkaravarthy et al. [30] and Santhosh et al. [31]. These results highlight the ability of the proposed framework to adapt to datasets with different visual characteristics.

Table 10: Comparison of the proposed method with *state-of-the-art* segmentation methods on the **DermQuest** dataset.

REFERENCE	YEAR	METHOD	ACC	DSC	IoU
Chakkara et al. [30]	2020	WaterSheet+Sobel	95.30%	–	–
Rout et al. [20]	2019	LinkNet-152	98.40%	80.00%	68.70%
Ashraf et al. [32]	2020	TL-CNN	85.10%	–	85.20%
Santhosh et al. [31]	2021	ACS	–	86.83%	–
Proposed method	2026	<i>regularized</i> U-Net	99.14%	91.88%	85.42%

Finally, on DermQuest (see table 10), the proposed approach achieves the best performance among the compared methods, with an accuracy of 99.14% and a Dice score of 91.88%. It surpasses methods proposed by Rout et al. [20], Chakkaravarthy et al. [30], and Ashraf et al. [32].

Overall, these results demonstrate that the proposed segmentation framework provides competitive and stable performance across heterogeneous dermoscopic datasets.

6. Concluding remarks

This study presented a robust framework for skin lesion segmentation designed to address the challenges associated with dermoscopic image variability and the limited availability of annotated medical data. By integrating

a carefully designed preprocessing pipeline with a dropout-regularized U-Net architecture trained from scratch, the proposed system enables accurate and consistent delineation of lesion boundaries while avoiding potential biases introduced by models pretrained on natural image datasets. This design choice promotes domain-specific feature learning and enhances the robustness of the segmentation process in dermatological imaging contexts.

The experimental evaluation conducted on four publicly available dermoscopic datasets—HAM10000, PH2, DermIS, and DermQuest—demonstrated the effectiveness and stability of the proposed approach. The results showed competitive performance in terms of Dice coefficient, Intersection over Union (IoU), and structural similarity metrics, confirming the model’s ability to produce reliable segmentation maps across heterogeneous datasets and acquisition conditions. These findings highlight that a carefully designed preprocessing strategy combined with a regularized convolutional architecture can achieve strong segmentation performance without relying on complex architectures or large-scale pretrained models.

Furthermore, the construction of a global dataset by aggregating multiple benchmarks enabled a comprehensive evaluation of cross-dataset generalization. The obtained results confirmed the robustness of the proposed framework under diverse imaging conditions, supporting its potential applicability in real-world dermatological analysis scenarios. In this context, the proposed method can serve as an effective preprocessing component for automated skin lesion analysis pipelines and computer-aided diagnosis systems.

From a clinical perspective, accurate and reliable lesion segmentation is a critical step in dermatological image analysis, as it directly influences subsequent tasks such as feature extraction, lesion classification, and malignancy assessment. By providing stable and consistent region-of-interest delineation, the proposed framework can assist dermatologists and support the development of reliable decision-support tools for skin cancer analysis.

The results suggest that carefully designed preprocessing and regularization strategies may provide an effective alternative to increasingly complex segmentation architectures.

Future research will focus on extending the diversity of training datasets, investigating advanced optimization and regularization strategies, and exploring lightweight architectures suitable for real-time clinical applications. Additionally, integrating the proposed segmentation framework with downstream diagnostic modules may contribute to the development of comprehensive automated systems for dermatological image analysis and computer-

aided skin cancer diagnosis.

7. Acknowledgements

We gratefully acknowledge the providers of the four publicly available benchmark dermoscopic datasets used in this study, namely **HAM10000** [16], **PH2** [17], **DermIS** [18], and **DermQuest** [19], for making their data accessible and enabling the evaluation of the proposed segmentation framework.

The authors acknowledge the use of generative artificial intelligence tools for language editing and clarity improvement. All scientific content, interpretations, and conclusions remain the sole responsibility of the authors.

References

- [1] International Agency for Research on Cancer, World Health Organization, Global cancer observatory, <https://gco.iarc.fr/today/en>, accessed on 3 September 2025 (2025).
- [2] T. Saba, M. A. Khan, A. Rehman, S. L. Marie-Sainte, Region extraction and classification of skin cancer: A heterogeneous framework of deep cnn features fusion and reduction, *Journal of medical systems* 43 (9) (2019) 289. doi:10.1007/s10916-019-1413-3.
- [3] M. E. Celebi, H. Iyatomi, G. Schaefer, W. V. Stoecker, Lesion border detection in dermoscopy images, *Computerized medical imaging and graphics* 33 (2) (2009) 148–153. doi:10.1016/j.compmedimag.2008.11.002.
- [4] Z. Wang, C. Wang, L. Peng, K. Lin, Y. Xue, X. Chen, L. Bao, C. Liu, J. Zhang, Y. Xie, Radiomic and deep learning analysis of dermoscopic images for skin lesion pattern decoding, *Scientific reports* 14 (1) (2024) 19781. doi:10.1038/s41598-024-70231-x.
- [5] X. Yang, H. Li, X. Zhou, Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy, *IEEE Transactions on Circuits and Systems I: Regular Papers* 53 (11) (2006) 2405–2414. doi:10.1109/TCSI.2006.884469.

- [6] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proceedings of the IEEE* 86 (11) (1998) 2278–2324. doi:10.1109/5.726791.
- [7] R. Baig, M. Bibi, A. Hamid, S. Kausar, S. Khalid, Deep learning approaches towards skin lesion segmentation and classification from dermoscopic images-a review, *Current Medical Imaging Reviews* 16 (5) (2020) 513–533. doi:10.2174/1573405615666191021124513.
- [8] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241. doi:10.1007/978-3-319-24574-4_28.
- [9] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, D. Feng, a multistage fully convolutional networks, *IEEE transactions on biomedical engineering* 64 (9) (2017) 2065–2074. doi:10.1109/TBME.2017.2659239.
- [10] P. Tschandl, C. Sinz, H. Kittler, Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation, *Computers in biology and medicine* 104 (2019) 111–116. doi:10.1016/j.compbiomed.2018.11.01.
- [11] W. Yin, D. Zhou, R. Nie, Di-unet: dual-branch interactive u-net for skin cancer image segmentation, *Journal of Cancer Research and Clinical Oncology* 149 (17) (2023) 15511–15524. doi:10.1007/s00432-023-05319-4.
- [12] M. Hu, Y. Li, X. Yang, Skinsam: Empowering skin cancer segmentation with segment anything model, *arXiv preprint* (2023). arXiv:2304.13973, doi:10.48550/arXiv.2304.13973.
- [13] K. A. Kumar, C. Vanmathi, Optimization driven model and segmentation network for skin cancer detection, *Computers and Electrical Engineering* 103 (2022) 108359. doi:10.1016/j.compeleceng.2022.108359.
- [14] M. Karri, C. S. R. Annavarapu, U. R. Acharya, Skin lesion segmentation using two-phase cross-domain transfer learning framework, *Computer Methods and Programs in Biomedicine* 231 (2023) 107408. doi:10.1016/j.cmpb.2023.107408.

- [15] P. Parida, R. Rout, Transition region based approach for skin lesion segmentation, *ELCVIA. Electronic letters on computer vision and image analysis* 19 (1) (2020) 0028–37. doi:10.1016/j.procs.2020.04.039.
- [16] P. Tschandl, C. Rosendahl, H. Kittler, The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, *Scientific data* 5 (1) (2018) 180161. doi:10.1038/sdata.2018.161.
- [17] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, J. Rozeira, Ph 2-a dermoscopic image database for research and benchmarking, in: 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), IEEE, 2013, pp. 5437–5440. doi: 10.1109/embc.2013.6610779.
- [18] Dermatology information system (2012).
URL <https://www.dermis.net/dermisroot/en/home/index.htm>
- [19] Dermquest archive (2012).
URL <https://www.galdermahcp.com/dermquest-archive>
- [20] R. Rout, P. Parida, Transition region based approach for skin lesion segmentation, *Procedia Computer Science* 171 (2020) 379–388. doi: 10.1016/j.procs.2020.04.039.
- [21] G. M. S. Himel, M. M. Islam, K. A. Al-Aff, S. I. Karim, M. K. U. Sikder, Skin cancer segmentation and classification using vision transformer for automatic analysis in dermatoscopy-based noninvasive digital system, *International journal of biomedical imaging* 2024 (1) (2024) 3022192. doi:10.1155/2024/3022192.
- [22] R. L. Araújo, F. H. d. Araújo, R. R. e. Silva, Automatic segmentation of melanoma skin cancer using transfer learning and fine-tuning, *Multimedia Systems* 28 (4) (2022) 1239–1250. doi:10.1109/HEALTHCOM49281.2021.9398926.
- [23] H. Ide, T. Kurita, Improvement of learning for cnn with relu activation by sparse regularization, in: 2017 international joint conference on neural networks (IJCNN), IEEE, 2017, pp. 2684–2691. doi: 10.1109/IJCNN.2017.7966185.

- [24] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014). doi:10.48550/arXiv.1412.6980.
- [25] K. Monica, J. Shreeharsha, P. Falkowski-Gilski, B. Falkowska-Gilska, M. Awasthy, R. Phadke, Melanoma skin cancer detection using mask-rcnn with modified gru model, *Frontiers in Physiology* 14 (2024) 1324042. doi:10.3389/fphys.2023.1324042.
- [26] L. Bi, J. Kim, E. Ahn, A. Kumar, D. Feng, M. Fulham, Step-wise integration of deep class-specific learning for dermoscopic image segmentation, *Pattern recognition* 85 (2019) 78–89. doi:10.1016/j.patcog.2018.08.001.
- [27] K. Duggani, M. K. Nath, A technical review report on deep learning approach for skin cancer detection and segmentation, *Data analytics and management: Proceedings of ICDAM (2021)* 87–99doi:10.1007/978-981-15-8335-3_9.
- [28] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, D. Feng, Dermoscopic image segmentation via multistage fully convolutional networks, *IEEE transactions on biomedical engineering* 64 (9) (2017) 2065–2074. doi:10.1109/TBME.2017.2659239.
- [29] N. Ahmed, X. Tan, L. Ma, A new method proposed to melanoma-skin cancer lesion detection and segmentation based on hybrid convolutional neural network, *Multimedia Tools and Applications* 82 (8) (2023) 11873–11896. doi:10.1007/s11042-022-13618-0.
- [30] A. P. Chakkaravarthy, A. Chandrasekar, Anatomical region segmentation method from dermoscopic images of pigmented skin lesions, *International Journal of Imaging Systems and Technology* 30 (3) (2020) 636–652. doi:10.1002/ima.22403.
- [31] M. Santhosh, R. Rubin Silas Raj, V. Rajinikanth, S. C. Satapathy, Image assisted assessment of cancer segment from dermoscopy images, in: *Intelligent Data Engineering and Analytics: Frontiers in Intelligent Computing: Theory and Applications (FICTA 2020)*, Volume 2, Springer, 2020, pp. 701–710. doi:10.1007/978-981-15-5679-1_68.

- [32] R. Ashraf, S. Afzal, A. U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O.-Y. Song, M. Maqsood, Region-of-interest based transfer learning assisted framework for skin cancer detection, *IEEE Access* 8 (2020) 147858–147871. doi:10.1109/ACCESS.2020.3014701.

Chapitre 4

Titre :

Classification Hiérarchique en Deux Phases des Lésions Cutanées Utilisant des Espaces de Caractéristiques CNN Réutilisés et une Fusion de Classifieurs Inter-Étapes

Résumé – Le cancer de la peau, en particulier le mélanome, représente un problème majeur de santé publique en raison de son incidence en constante augmentation. Les méthodes de diagnostic clinique conventionnelles souffrent de limitations notables, notamment la variabilité inter-observateur et un risque d'erreur de diagnostic, ce qui motive l'adoption de techniques d'intelligence artificielle (IA) pour la détection précoce et la classification fiable des lésions cutanées. Dans cette étude, nous proposons un modèle hybride en deux phases basé sur la fusion de classifieurs pour la classification de multiples types de lésions cutanées. Le cadre proposé combine des caractéristiques artisanales avec des représentations profondes extraites de réseaux de neurones convolutionnels (CNN) pré-entraînés. La chaîne de traitement comprend l'augmentation et l'équilibrage des données, l'amélioration de la qualité de l'image, l'extraction de la région d'intérêt, l'extraction de caractéristiques et la réduction de dimensionnalité. Contrairement aux approches multiclasse planes conventionnelles, le cadre proposé réutilise hiérarchiquement les espaces de caractéristiques des CNN et fusionne les classifieurs à travers les étapes, réduisant progressivement l'espace de décision. Cette stratégie de décision structurée améliore la séparabilité des classes et la robustesse, en particulier pour les lésions visuellement similaires. Quatre classifieurs—Machine à Vecteurs de Support Nu (NuSVM), Analyse Discriminante Linéaire (LDA), Naïve Bayes de Bernoulli (BNB) et Régression Logistique (LR)—ont été évalués en utilisant quatre ensembles de données publics : PH2, HAM10000, DermIS et DermQuest. Trois scénarios expérimentaux ont été étudiés : la classification binaire, la classification multiclasse directe et la classification multiclasse en deux phases proposée. Les résultats expérimentaux démontrent la supériorité du modèle proposé, en particulier sur l'ensemble de données PH2, où il a atteint des précisions de 98,43% pour la classification binaire, 87,45% pour la classification multiclasse directe et 96,00% pour la classification multiclasse en deux phases. Ces résultats soulignent l'efficacité de l'approche hybride pour améliorer à la fois la précision et la robustesse du diagnostic dermatologique assisté par l'IA. Bien que validé sur des ensembles de données dermoscopiques, l'architecture hiérarchique proposée est générique et peut être étendue à d'autres tâches de classification d'images médicales.

Mots clés :

Cancer de la Peau, Intelligence Artificielle, Classification Hiérarchique, Réutilisation

de Caractéristiques CNN, Fusion de Classifieurs

Contributions associées :

Z. Mounmie, Y. Yaddaden, A. Bouzouane, "Hierarchical Two-Phase Skin Lesion Classification Using Reused CNN Feature Spaces and Cross-Stage Classifier Fusion" in *Engineering Applications of Artificial Intelligence* (EAAI-S-26-07398). (Statut : Soumis)

Engineering Applications of Artificial Intelligence

Hierarchical Two-Phase Skin Lesion Classification Using Reused Convolutional Neural Network Feature Spaces and Cross-Stage Classifier Fusion

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research paper
Keywords:	Skin Cancer; Artificial Intelligence; Hierarchical Classification; Convolutional Neural Network Feature Reuse; Classifier Fusion
Corresponding Author:	Yacine Yaddaden, Ph. D. University of Quebec at Rimouski CANADA
First Author:	Zakariaou Moumnie, M.Sc. in computer science
Order of Authors:	Zakariaou Moumnie, M.Sc. in computer science
	Yacine Yaddaden, Ph. D.
	Abdenour Bouzouane, Ph. D.
Abstract:	Skin cancer, particularly melanoma, represents a major public health concern due to its steadily increasing incidence. Conventional clinical diagnostic methods suffer from notable limitations, including inter-observer variability and a risk of misdiagnosis, which motivates the adoption of artificial intelligence techniques for early detection and reliable classification of skin lesions. In this study, we propose a two-phase hybrid model based on classifier fusion for the classification of multiple skin lesion types. The framework combines handcrafted features with deep representations extracted from pre-trained convolutional neural networks. The processing pipeline consists of data augmentation and balancing, image quality enhancement, region-of-interest extraction, feature extraction, and dimensionality reduction. Unlike conventional flat multiclass approaches, the proposed framework hierarchically reuses convolutional neural network feature spaces and fuses classifiers across stages, progressively reducing the decision space. This structured decision strategy improves class separability and robustness, particularly for visually similar lesions. Four classifiers—Nu-Support Vector Machine, Linear Discriminant Analysis, Bernoulli Naive Bayes, and Logistic Regression—were evaluated using four benchmark datasets. Three experimental scenarios were investigated: binary classification, direct multi-class classification, and the proposed two-phase multi-class classification. Experimental results demonstrate the superiority of the proposed model, particularly on the PH2 dataset, where it achieved accuracies of 98.43% for binary classification, 87.45% for direct multi-class classification, and 96.00% for two-phase multi-class classification. These findings highlight the effectiveness of the hybrid approach in enhancing both accuracy and robustness for artificial-intelligence-assisted dermatological diagnosis. Although validated on dermoscopic datasets, the proposed hierarchical architecture is generic and can be extended to other medical image classification tasks.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Graphical Abstract

Hierarchical Two-Phase Skin Lesion Classification Using Reused Convolutional Neural Network Feature Spaces and Cross-Stage Classifier Fusion

Zakariaou Moumnine, Yacine Yaddaden, Abdenour Bouzouane

Highlights

Hierarchical Two-Phase Skin Lesion Classification Using Reused Convolutional Neural Network Feature Spaces and Cross-Stage Classifier Fusion

Zakariaou Moumnine, Yacine Yaddaden, Abdenour Bouzouane

- Hierarchical two-phase framework with reused CNN feature spaces.
- Cross-stage classifier fusion for progressive decision refinement.
- Improved discrimination of visually similar skin lesion classes.
- Consistent performance gains across four public datasets.

Hierarchical Two-Phase Skin Lesion Classification Using Reused Convolutional Neural Network Feature Spaces and Cross-Stage Classifier Fusion

Zakariaou Moumnie^a, Yacine Yaddaden^{a,*}, Abdenour Bouzouane^b

^a*Université du Québec à Rimouski, 1595 Bd Alphonse-Desjardins, Lévis, G6V 0A6, QC, Canada*

^b*Université du Québec à Chicoutimi, 555 Bd de l'Université, Chicoutimi, G7H 2B1, QC, Canada*

Abstract

Skin cancer, particularly melanoma, represents a major public health concern due to its steadily increasing incidence. Conventional clinical diagnostic methods suffer from notable limitations, including inter-observer variability and a risk of misdiagnosis, which motivates the adoption of artificial intelligence techniques for early detection and reliable classification of skin lesions.

In this study, we propose a two-phase hybrid model based on classifier fusion for the classification of multiple skin lesion types. The framework combines handcrafted features with deep representations extracted from pre-trained convolutional neural networks. The processing pipeline consists of data augmentation and balancing, image quality enhancement, region-of-interest extraction, feature extraction, and dimensionality reduction.

Unlike conventional flat multiclass approaches, the proposed framework hierarchically reuses convolutional neural network feature spaces and fuses classifiers across stages, progressively reducing the decision space. This structured decision strategy improves class separability and robustness, particularly for visually similar lesions.

Four classifiers—Nu-Support Vector Machine, Linear Discriminant Analysis, Bernoulli Naive Bayes, and Logistic Regression—were evaluated using

*Corresponding author

Email addresses: zakariaou.moumnie@uqar.ca (Zakariaou Moumnie), yacine_yaddaden@uqar.ca (Yacine Yaddaden), abouzoua@uqac.ca (Abdenour Bouzouane)

four public benchmark datasets. Three experimental scenarios were investigated: binary classification, direct multi-class classification, and the proposed two-phase multi-class classification.

Experimental results demonstrate the superiority of the proposed model, particularly on the PH2 dermoscopic image dataset, where it achieved accuracies of 98.43% for binary classification, 87.45% for direct multi-class classification, and 96.00% for two-phase multi-class classification. These findings highlight the effectiveness of the hybrid approach in enhancing both accuracy and robustness for artificial-intelligence-assisted dermatological diagnosis. Although validated on dermoscopic datasets, the proposed hierarchical architecture is generic and can be extended to other medical image classification tasks.

Keywords:

Skin Cancer, Artificial Intelligence, Hierarchical Classification, Convolutional Neural Network Feature Reuse, Classifier Fusion

1. Introduction

Skin cancer is among the most prevalent diseases worldwide and constitutes a major public health concern. Its incidence has steadily increased over recent decades, driven by factors such as prolonged exposure to ultraviolet (UV) radiation, population aging, and evolving lifestyle habits.

In 2016, Canada reported approximately 6,800 new cases of melanoma—the most aggressive, though less frequent, form of skin cancer—resulting in nearly 1,200 deaths [1, 2]. A joint report published in 2019 by the World Health Organization (WHO) and the International Labour Organization (ILO) further highlighted the impact of occupational exposure to solar radiation, reporting nearly 19,000 deaths from non-melanoma skin cancer across 183 countries [3].

In this context, early diagnosis plays a critical role in improving patient outcomes. However, conventional clinical diagnosis largely depends on practitioner expertise, which can lead to inter-observer variability and potential misdiagnosis. Artificial intelligence (AI) offers a promising alternative through automated and standardized skin lesion assessment. Recent studies have demonstrated the effectiveness of AI-based approaches in dermatology, achieving performance comparable to expert dermatologists [4, 5]. These advances motivate the development of robust and clinically applicable computer-aided diagnosis systems.

Current state-of-the-art methods predominantly rely on convolutional neural networks (CNNs), sometimes combined with traditional image processing techniques [6]. While such hybrid frameworks have demonstrated improved diagnostic performance, most existing approaches adopt flat multiclass prediction strategies, which often struggle with visually similar lesion categories and class imbalance. To address these limitations, we introduce a hierarchical two-phase classification framework that progressively constrains the decision space.

Our experimental protocol comprises three interconnected analytical stages. The first phase performs binary classification to distinguish benign from malignant lesions. This is followed by direct multiclass classification. Finally, a two-phase hierarchical strategy is applied, where the outcome of the initial binary stage guides subsequent subclassification. By reusing CNN feature representations across stages and fusing classifiers hierarchically, the proposed framework reduces class ambiguity and improves inter-class separability.

This study presents both methodological and clinical contributions by advancing AI-driven dermoscopic analysis while supporting early diagnosis in practical settings.

The main contributions of this work are:

- A hierarchical two-phase classification framework that reuses CNN feature representations across stages.
- A cross-stage classifier fusion strategy that progressively constrains the decision space.
- Extensive validation on four public datasets demonstrating improved robustness over direct multiclass classification.

The remainder of this paper is organized as follows. Section 2 reviews related work and foundational concepts. Feature extraction and classification methodologies are detailed in Sections 3 and 4, respectively. The proposed skin cancer detection pipeline is presented in Section 5. Experimental settings and public datasets are described in Section 6. Results and discussions are provided in Section 7. Finally, Section 8 concludes the paper and outlines future research directions.

2. Background

2.1. Fundamentals

Given the rising prevalence of skin cancer and the critical importance of early detection, automated dermoscopic diagnosis has emerged as a promising tool to assist dermatologists [5]. However, substantial variability in lesion morphology, appearance, and inter-patient characteristics makes automated classification a highly challenging task [1].

This work builds upon the integration of pre-trained convolutional neural networks (CNNs) with conventional machine learning (ML) techniques to improve skin lesion classification performance. The underlying motivation is to exploit the complementarity between deep representations learned automatically by neural networks and handcrafted structural or textural features explicitly extracted from dermoscopic images.

A skin cancer detection system typically follows a processing pipeline similar to that of a classical pattern recognition framework: (1) dermoscopic images serve as input data; (2) preprocessing is applied to enhance image quality and facilitate subsequent stages; (3) discriminative features are extracted to capture relevant visual information; (4) dimensionality reduction techniques are employed to decrease computational complexity while preserving maximal class separability (5) classification is performed by evaluating multiple learning algorithms and selecting the most effective one; and (5) the final output provides either a binary decision (benign vs. malignant) or a detailed multi-class lesion categorization.

In automated skin lesion analysis, two primary classification paradigms are commonly adopted: binary classification and multi-class classification. These approaches address distinct yet complementary diagnostic objectives within AI-assisted clinical decision support systems. Figure 1 illustrates a generic framework for skin lesion classification.

Binary classification aims to determine whether a lesion is benign or malignant. It represents a critical first screening stage, enabling rapid identification of suspicious cases that require further clinical evaluation by a dermatologist [1]. This approach is commonly employed in large-scale screening programs, where rapid assessment and interpretability are essential [5].

In contrast, **multi-class classification** enables finer-grained lesion discrimination by identifying specific lesion categories present in curated medical image databases [7]. These categories typically include benign keratoses and lentigines (BKL), dermatofibroma (DF), melanocytic nevi (NV)—subdivided

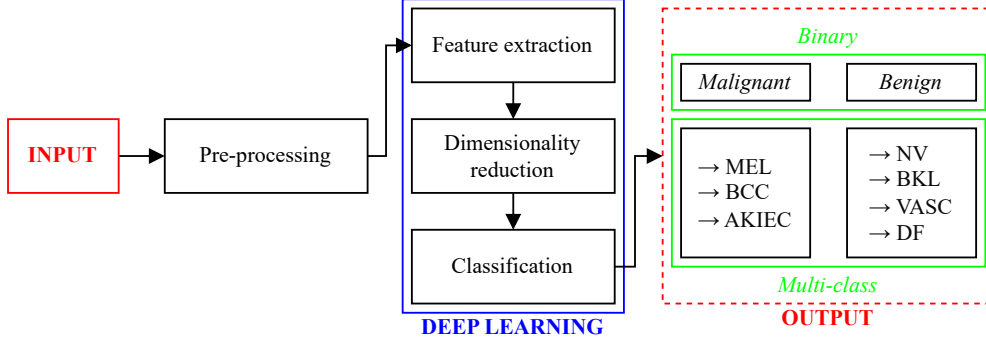


Figure 1: Overview of a common skin lesion classification system.

into common nevi (CNV) and atypical nevi (ANV)—vascular lesions (VASC), actinic keratoses and intraepithelial carcinoma (AKIEC), basal cell carcinoma (BCC), and malignant melanoma (MEL). Among these, melanoma is the most aggressive and life-threatening form of skin cancer, making early detection particularly critical [8, 7].

The proposed two-phase classification framework enables the development of flexible diagnostic systems tailored to diverse clinical scenarios, from rapid initial screening to comprehensive pathological analysis. The hybrid model presented in this work aims to deliver a robust, efficient, and generalizable AI-driven solution for automated dermatological image diagnosis.

2.2. Related Works

Numerous studies have focused on improving classification accuracy for various types of skin cancer. To contextualize the proposed framework, this section reviews representative works in automated skin cancer detection, covering traditional machine learning (ML), convolutional neural networks (CNNs), and hybrid approaches for feature learning and classification.

2.2.1. Binary Classification

Murugan et al. [9] proposed an ML-based skin cancer detection system employing median filtering to remove artifacts such as noise, hair, and bubbles. Lesion segmentation was performed using the watershed algorithm, while ABCD and GLCM features were extracted for classification. Several classifiers, including Random Forest (RF), Support Vector Machine (SVM), and k -Nearest Neighbors (k -NN), were evaluated.

Ashraf et al. [10] introduced an ROI-based intelligent system to differentiate melanoma from nevus using transfer learning. An improved K-means algorithm was applied for lesion segmentation, focusing on melanoma-relevant regions. A CNN-based TL model incorporating data augmentation was employed for classification using DermIS and DermQuest datasets.

Dahou et al. [11] developed an optimized skin cancer detection model using MobileNetV3 for feature extraction. Feature selection was refined through Hunger Games Search (HGS), Particle Swarm Optimization (PSO), and Dynamic Opposite Learning HGS (DOLHGS), leading to improved classification performance.

Nazi et al. [12] proposed a hybrid deep learning framework combining U-Net for lesion segmentation with a DCNN-SVM classifier. Transfer learning and data augmentation were employed to address limited data availability, with pre-trained CNNs serving as feature extractors to enhance diagnostic accuracy.

Xin et al. [13] introduced SkinTrans, a Vision Transformer (ViT)-based model that utilizes multi-scale overlapping sliding windows and attention mechanisms to highlight informative regions. Contrastive learning was further incorporated to improve class separability.

Hosny et al. [14] explored a Residual Deep Convolutional Neural Network (RDCNN) architecture evaluated on raw, preprocessed, and segmented images. A TL-based variant trained on external datasets demonstrated superior robustness and flexibility compared to conventional DL models.

2.2.2. Beyond Binary Classification

Hoang et al. [15] proposed a segmentation-based classification framework combining EW-FCM with Wide-ShuffleNet. Lesions were initially segmented using Entropy Weight and First-order Cumulative Moment, followed by deep classification of lesion types.

Abd Elaziz et al. [16] introduced an optimization-based framework employing MobileNetV3 for feature extraction and a novel Improved Artificial Rabbits Optimizer (IARO) for feature selection. Their model was evaluated on HAM10000, PH2, and ISIC 2016 datasets.

Fraiwan et al. [15] applied transfer learning to classify seven lesion categories from HAM10000 using thirteen pre-trained CNN models. The study highlighted challenges related to class imbalance and limited sample availability in certain categories.

Bassi et al. [17] proposed an ensemble-based framework integrating nine CNN architectures, including AlexNet, GoogLeNet, MobileNetV2, Xception, ResNet variants, InceptionResNet-V2, and DenseNet201. Outputs were fused via a Collective Decision Block to improve multi-class classification performance.

Abayomi-Alli et al. [18] addressed data scarcity using a covariant SMOTE-based augmentation strategy, generating synthetic samples in a nonlinear embedding space to train a SqueezeNet model.

Afza et al. [19] proposed a hierarchical framework combining image enhancement, superpixel-based segmentation, and deep learning. Features extracted via ResNet-50 were optimized using Grasshopper Optimization Algorithm (GOA) and classified with Bernoulli Naive Bayes (BNB).

Collectively, these studies demonstrate that ML-based methods, CNNs, transfer learning, ensemble strategies, and RDCNN architectures are widely investigated for improving diagnostic performance in skin cancer detection.

2.3. Validity with Regard to Previous Work

Existing skin lesion classification approaches span traditional computer vision methods [9], deep neural networks [12, 14], and more recent Transformer-based architectures [13]. Despite significant advances, challenges persist in generalization and fine-grained discrimination among visually similar lesion categories.

To address these limitations, the present study introduces several methodological enhancements.

A two-stage hierarchical decision framework is proposed, consisting of an initial binary classifier followed by an assisted multi-class classifier. This strategy constrains the decision space and reduces inter-class confusion, an approach that remains relatively underexplored in current literature [10, 18].

By integrating pre-trained CNNs with traditional ML methods, the proposed hierarchical architecture offers advantages over standalone deep models. In particular, the multi-level fusion mechanism is expected to provide:

- richer and more discriminative lesion feature representations,
- increased robustness to class imbalance,
- improved generalization capability.

Most existing approaches rely on flat multiclass prediction or standalone deep architectures. In contrast, our method explicitly introduces a hierarchical reuse of CNN features combined with cross-stage classifier fusion. This strategy remains relatively underexplored and provides improved robustness in visually ambiguous lesion categories.

Consequently, the proposed two-phase hierarchical framework represents a promising strategy for automated skin lesion detection and diagnosis, especially in complex clinical scenarios.

3. Feature Extraction Techniques

To extract discriminative representations from dermoscopic images, we employ both classical machine learning (ML)-based feature descriptors and deep features derived from pre-trained convolutional neural network (CNN) architectures.

3.1. Models Based on Pre-trained CNN Architectures

3.1.1. VGG-16 and VGG-19

VGG16 and VGG19 are deep CNN architectures developed by the Visual Geometry Group at the University of Oxford, consisting of 16 and 19 learnable layers, respectively. Both models rely on 3×3 convolutional filters, ReLU activation functions, and max-pooling layers [20, 21]. With an input resolution of 224×224 pixels and approximately 138 million parameters for VGG16, these networks are capable of extracting hierarchical and highly discriminative visual features. VGG19 is also widely adopted in transfer learning scenarios due to its deeper architecture, which may provide richer feature representations compared to VGG16 [21].

3.1.2. ResNet-50

ResNet-50 is a deep residual network designed to mitigate the vanishing gradient problem through residual connections. The architecture comprises 50 layers organized into residual blocks, followed by batch normalization and ReLU activation functions. With approximately 25 million parameters, ResNet-50 is able to learn highly complex feature representations, making it well suited for challenging image classification tasks such as skin lesion analysis [22].

3.1.3. MobileNet

MobileNet is a lightweight CNN architecture specifically designed to reduce computational complexity for resource-constrained environments. It is based on depthwise separable convolutions, resulting in a compact model with approximately 4 million parameters while maintaining competitive performance [23]. Owing to its efficiency and fast inference speed, MobileNet is particularly suitable for real-time and embedded applications.

3.2. Feature Extraction Based on Traditional Methods

3.2.1. Local Binary Patterns

Local Binary Patterns (LBP) is an illumination-invariant texture descriptor that encodes local image structures by comparing each pixel with its neighboring pixels [24]. Several variants, including uniform and rotation-invariant LBP, have been proposed to improve robustness and discriminative power. Despite its effectiveness, LBP remains sensitive to rotation and fine-grained texture variations [25].

3.2.2. Histogram of Oriented Gradients

Histogram of Oriented Gradients (HOG) computes local histograms of intensity gradients, which are subsequently normalized to enhance robustness. This descriptor primarily captures shape-related information and exhibits invariance to illumination changes. However, it is computationally demanding and sensitive to noise [26]. HOG remains a widely used feature extraction technique in object detection and action recognition tasks.

4. Classification Techniques

In this study, a set of supervised learning classifiers is employed to discriminate skin lesion categories. The performance of these models is evaluated based on their ability to exploit discriminative information extracted from both traditional machine learning (ML) features and deep learning (DL)-based representations.

4.1. LDA (Linear Discriminant Analysis)

Linear Discriminant Analysis (LDA) is a statistical classification method that projects high-dimensional data onto a lower-dimensional subspace while maximizing class separability. This is achieved by optimizing Fisher's criterion, defined as:

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (1)$$

where S_B and S_W denote the between-class and within-class scatter matrices, respectively, and w represents the optimal projection vector.

4.2. NuSVM (Support Vector Machine with Linear Kernel)

Nu-Support Vector Machine (NuSVM) is an extension of the standard SVM that introduces a parameter $\nu \in (0, 1]$ to control the trade-off between the number of support vectors and training errors. The optimization problem is formulated as:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^n \xi_i \quad (2)$$

subject to:

$$y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (3)$$

where C is the regularization parameter, x_i denotes the input samples, y_i their corresponding labels, and ξ_i are slack variables allowing soft-margin classification.

4.3. LR (Logistic Regression)

Logistic Regression (LR) is a probabilistic linear classifier that estimates the likelihood of a sample belonging to a given class using the logistic (sigmoid) function:

$$P(y = 1|x) = \frac{1}{1 + e^{-w^T x - b}} \quad (4)$$

where w and b represent the weight vector and bias term, respectively. Model parameters are optimized by minimizing the cross-entropy loss:

$$\mathcal{L} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (5)$$

4.4. BNB (Bernoulli Naive Bayes)

Bernoulli Naive Bayes (BNB) is a probabilistic classifier based on Bayes' theorem, assuming conditional independence among features. It is particularly suitable for binary-valued feature vectors. The posterior probability is expressed as:

$$P(C_k|x) \propto P(C_k) \prod_{i=1}^n P(x_i|C_k)^{x_i} \cdot (1 - P(x_i|C_k))^{1-x_i} \quad (6)$$

where $P(C_k)$ denotes the prior probability of class C_k , and $P(x_i|C_k)$ represents the likelihood of feature x_i given class C_k .

5. Proposed system

The primary objective of this study is to evaluate the performance of various feature descriptors in combination with different classifiers on standard benchmark datasets, as illustrated in Figure 2. Each stage of the proposed pipeline is described in detail below.

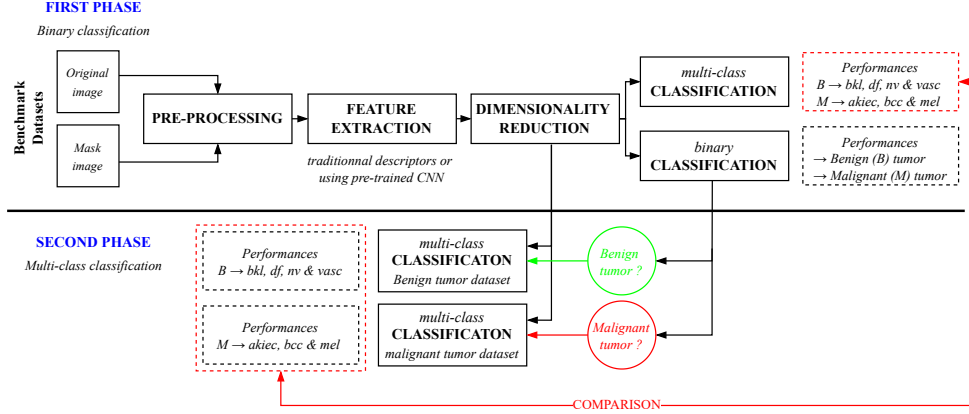


Figure 2: Overview of the proposed system.

The first phase performs coarse benign–malignant discrimination, while the second phase refines predictions within each group. Classifiers selected at each stage are subsequently fused to produce the final decision.

5.1. Pre-processing

Pre-processing constitutes a crucial initial step prior to model training. It comprises three main stages: data augmentation, region-of-interest (ROI) extraction, and histogram equalization for contrast enhancement.

5.1.1. Data Augmentation

Two synchronized generators are employed to independently augment both the dermoscopic images and their corresponding segmentation masks. RGB images undergo random geometric transformations, including rotations up to 30° , translations up to 10%, shear up to 0.2, zoom between 90% and 110%, and horizontal and vertical flips. In addition, photometric variations are applied, such as brightness adjustment between 0.8 and 1.2 and channel shifts up to 20 [27]. Border padding is handled via interpolation from neighboring pixels to avoid edge artifacts.

The associated masks are subjected to identical geometric transformations to ensure pixel-wise alignment, with zero-value padding applied to preserve lesion boundaries.

5.1.2. ROI Extraction

This step aims to automatically isolate lesion regions from binary masks, where lesion pixels are represented by 1 (white) and background pixels by 0 (black). Connected components are identified using 8-connectivity analysis. Images containing multiple disconnected regions are discarded to ensure a single lesion per sample.

For each detected lesion, a minimum bounding rectangle is computed. Let $\mathbf{R}$ denote the set of lesion pixels. The bounding box is defined as:

$$\text{BBox} = \left(\min_{(x,y) \in \mathbf{R}} x, \min_{(x,y) \in \mathbf{R}} y, \max_{(x,y) \in \mathbf{R}} x, \max_{(x,y) \in \mathbf{R}} y \right) \quad (7)$$

This region is then cropped from the original dermoscopic image, focusing exclusively on the lesion area. The resulting image is resized to 224×224 pixels using bilinear interpolation to ensure consistent input dimensions for CNN-based feature extraction.

5.1.3. Histogram Equalization and Resizing

To enhance contrast while preserving chromatic information, histogram equalization is applied exclusively to the Value (V) channel after converting images from the RGB to the HSV color space. This strategy avoids color

distortions commonly introduced by direct histogram equalization in RGB space by maintaining Hue (H) and Saturation (S) components.

Furthermore, all images are resized to a fixed resolution of 128×128 pixels prior to traditional feature extraction to guarantee uniform processing across the dataset.

5.2. Feature Extraction

To conduct a comprehensive comparative evaluation, two complementary feature extraction strategies are adopted. Traditional handcrafted descriptors, including Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG), are employed alongside deep features extracted from pre-trained CNN architectures such as VGG-16, VGG-19, ResNet-50, and MobileNet.

This dual strategy enables direct comparison between conventional representations and deep learned features, providing insights into their respective effectiveness for skin lesion classification.

5.3. Dimensionality Reduction

To further improve computational efficiency and mitigate redundancy in high-dimensional feature spaces, Principal Component Analysis (PCA) is applied to the extracted features [28]. PCA is a classical statistical technique that projects data onto a lower-dimensional subspace while preserving maximal variance.

The method generates uncorrelated variables, known as principal components, with the leading components capturing the majority of data variability. Consequently, less informative dimensions are discarded.

This process involves centering the data, computing the covariance matrix, extracting eigenvectors and eigenvalues, and selecting the components associated with the highest eigenvalues. These principal directions provide a compact and interpretable representation of the data.

By projecting samples onto this reduced subspace, dimensionality is decreased, noise is attenuated, generalization is enhanced, and the performance of subsequent ML classifiers is accelerated.

5.4. Classification

Multiple supervised learning algorithms are employed to categorize skin lesions into predefined classes. These models are trained on labeled samples and evaluated on unseen data to assess their generalization capability.

Specifically, NuSVM (linear kernel Support Vector Machine), Logistic Regression (LR), Bernoulli Naive Bayes (BNB), and Linear Discriminant Analysis (LDA) are considered.

Classifier performance during testing is used to quantify predictive accuracy on novel samples. The overarching goal is to identify the most robust classification strategy capable of addressing key challenges in skin lesion analysis, including inter-class similarity, class imbalance, and substantial visual variability.

6. Evaluation Protocol

6.1. Dataset Description

1. **HAM10000:** The HAM10000 dataset comprises 10,015 dermoscopic images of pigmented skin lesions, of which more than 50% are histopathologically confirmed. It encompasses seven diagnostic categories, grouped into benign and malignant classes, and is widely used as a benchmark for skin lesion analysis [7].
2. **PH2:** Collected at Pedro Hispano Hospital (Portugal), the PH2 dataset contains 200 dermoscopic images of melanocytic lesions acquired at $20\times$ magnification with a resolution of 768×560 pixels. It is commonly employed for evaluating segmentation and classification algorithms [8].
3. **DermIS:** DermIS is a publicly accessible educational dermatology database providing clinical images accompanied by expert annotations and diagnostic descriptions [29].
4. **DermQuest:** Originally developed by Galderma, DermQuest remains a widely referenced resource in skin lesion classification studies, despite no longer being actively maintained. Its image repository is still accessible online [30].

Tables 1, 2, 3 and 4 summarize the characteristics of the original datasets, including the number of classes and sample sizes before and after data augmentation. Figure 3 illustrates representative examples from each dataset along with their corresponding lesion categories.

6.2. Evaluation Metrics

To assess the effectiveness of the proposed framework, several standard classification metrics are employed, including **accuracy**, **precision**, **recall**,

Table 1: Details about **HAM10000** dataset before and after data augmentation.

CLASS	SUBCLASS	BEFORE AUG.	AFTER AUG.
Benign lesions	Melanocytic nevi (NV)	6,705	1,000
	Benign keratosis-like (BKL)	1,099	1,000
	Vascular lesions (VASC)	142	1,000
	Dermatofibroma (DF)	115	1,000
Malignant lesions	Melanoma (MEL)	1,113	1,000
	Basal cell carcinoma (BCC)	514	1,000
	Actinic keratoses (AKIEC)	327	1,000
Total		10,015	7,000

Table 2: Details about **PH2** dataset before and after data augmentation.

CLASS	SUBCLASS	BEFORE AUG.	AFTER AUG.
Benign lesions	Common nevus (CNV)	80	400
	Atypical nevus (ANV)	80	400
Malignant lesions	Melanoma (MEL)	40	400
Total		200	1,200

Table 3: Details about **DermIS** dataset before and after data augmentation.

CLASS	BEFORE AUG.	AFTER AUG.
Melanoma (MEL)	43	220
Not melanoma (NMEL)	26	203
Total	69	423

Table 4: Details about **DermQuest** dataset before and after data augmentation.

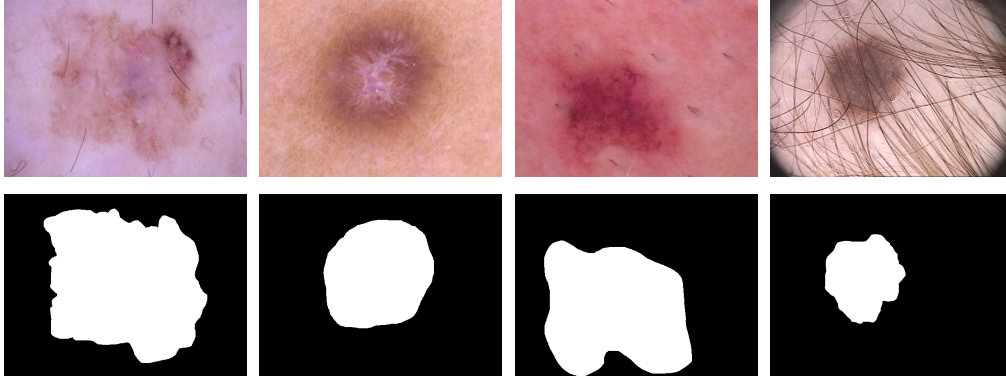
CLASS	BEFORE AUG.	AFTER AUG.
Melanoma (MEL)	76	305
Not melanoma (NMEL)	61	287
Total	137	592

F1-score, and the **confusion matrix**. These measures provide complementary insights into the model’s ability to discriminate among lesion categories.

Accuracy (Eq. 8) quantifies the overall proportion of correctly classified samples:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

Benign Lesions



Malignant Lesions

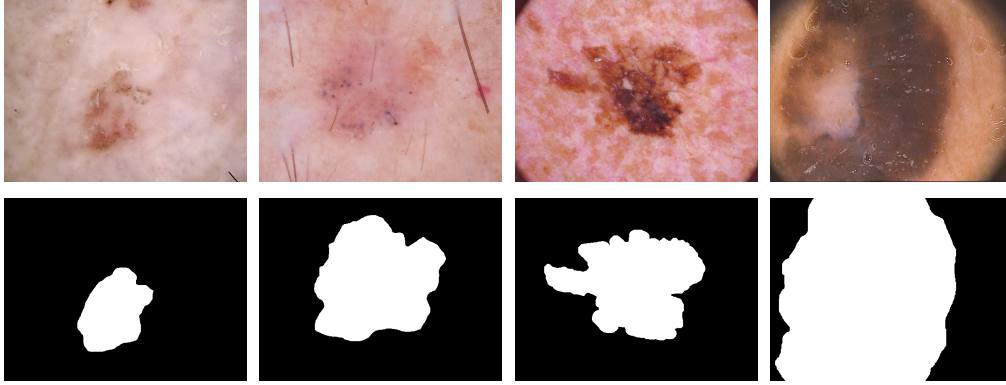


Figure 3: Representative dermoscopic images and corresponding segmentation masks for benign and malignant lesions.

Precision (Eq. 9) indicates the proportion of positive predictions that are truly positive:

$$\text{Precision} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FP}} \quad (9)$$

Recall (Eq. 10), also referred to as *sensitivity*, measures the ability of the model to correctly identify all positive instances:

$$\text{Recall} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}} \quad (10)$$

The F1-score (Eq. 11), defined as the harmonic mean of precision and recall, is particularly suitable for evaluating performance on imbalanced datasets:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

where:

- **TP**: True Positives — correctly classified positive samples
- **TN**: True Negatives — correctly classified negative samples
- **FP**: False Positives — negative samples incorrectly classified as positive
- **FN**: False Negatives — positive samples incorrectly classified as negative

Confusion Matrix: The confusion matrix summarizes prediction outcomes by reporting the number of correctly and incorrectly classified samples for each class. It provides detailed insight into model behavior and highlights systematic misclassification patterns.

6.3. Validation Strategy

To mitigate overfitting and ensure reliable generalization, the dataset is partitioned into two subsets: **80%** for training and **20%** for testing. This train–test split strategy enables an unbiased evaluation of model performance on unseen data.

Although simple, this approach provides a stable and computationally efficient evaluation protocol. Performance discrepancies between training and testing sets further help identify potential overfitting, thereby supporting the development of robust and generalizable models.

This protocol was selected to maintain consistency with prior studies on the same datasets and to ensure sufficient samples for training after augmentation. Future work will consider cross-validation to further assess statistical stability.

7. Results & Discussion

7.1. Experimental Results

In this study, two classification paradigms were investigated. First, binary classification was performed to discriminate between benign and malignant lesions. Second, multiclass classification was conducted, initially by identifying lesion subtypes within each category and subsequently by jointly classifying all lesion types across each dataset.

7.1.1. Binary Classification

For the binary classification task, the highest performance was achieved on the PH2 dataset, where the combination of the pre-trained ResNet-50 model and Logistic Regression (LR) yielded an accuracy of 98.43% (see Table 6). On the HAM10000 dataset, both the VGG-16 + BNB and ResNet-50 + NuSVM configurations achieved an accuracy of 91.48% (see Table 5).

For the DermIS and DermQuest datasets, maximum accuracies of 96.42% and 90.64%, respectively, were obtained primarily using VGG-19, VGG-16, and ResNet-50 feature extractors (see Tables 7 and 8). These results highlight the robustness and generalization capability of deep CNN-based representations for binary skin lesion detection.

7.1.2. Multi-class Classification

To implement the proposed two-phase multiclass classification framework, the feature descriptors and classifiers that achieved the best performance at each preceding stage were retained. After obtaining the overall binary classification results for each dataset, the pipeline proceeded to lesion sub-classification.

For HAM10000, the first phase focused on benign lesions and their corresponding subclasses (BKL, DF, NV, VASC), achieving an accuracy of 95.00% using the MobileNet descriptor combined with NuSVM (see Table 10). Subsequently, malignant lesion subclasses (AKIEC, BCC, MEL) were classified using ResNet-50 with NuSVM, reaching an accuracy of 70.44% (see Table 11). The comparatively lower performance in malignant classification is mainly attributed to the higher visual heterogeneity among malignant lesion types.

For the final multiclass decision, features extracted from the same descriptor were retained, and the best-performing classifiers from the binary and sub-classification stages were fused to obtain predictions across all seven

Table 5: Comparison of binary classification methods on **HAM10000** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	88.72	89.70	88.65	88.72
	NuSVM	76.96	83.84	75.72	76.96
	LR	80.88	81.56	80.77	80.88
	BNB	88.23	90.10	88.09	88.23
ResNet50	BNB	91.42	92.10	91.38	91.42
	NuSVM	89.63	91.13	89.70	89.63
	LR	90.68	91.72	90.62	90.68
	BNB	78.18	78.56	78.18	78.35
VGG16	LDA	88.97	90.44	88.88	88.97
	NuSVM	91.42	92.22	91.37	91.42
	LR	89.21	90.32	89.14	89.21
	BNB	81.12	82.18	81.12	80.98
VGG19	LDA	85.78	86.93	85.67	85.78
	NuSVM	90.44	91.39	90.38	90.44
	LR	88.72	89.58	88.66	88.72
	BNB	82.84	83.63	82.67	82.84
HOG	LDA	84.06	84.77	83.86	84.06
	NuSVM	90.85	90.98	90.84	90.85
	LR	88.23	89.92	88.10	88.23
	BNB	69.60	69.70	69.62	69.60
LBP	LDA	90.11	90.73	90.11	90.09
	NuSVM	76.90	82.54	75.94	76.90
	LR	84.32	87.80	84.30	83.94
	BNB	82.84	86.90	82.35	82.84

HAM10000 classes. The same hierarchical protocol was applied to the PH2 dataset.

To assess the effectiveness of the hierarchical strategy, a direct multi-class classification was first conducted by extracting features from all images and jointly classifying all lesion categories. For PH2, this approach achieved an accuracy of 87.45% using ResNet-50 with LR (see Table 13), while for HAM10000, an accuracy of 73.40% was obtained using ResNet-50 with NuSVM (see Table 12).

In contrast, the proposed two-phase hierarchical framework, combining

Table 6: Comparison of binary classification methods on **PH2** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	91.33	92.57	90.89	91.33
	NuSVM	91.22	92.80	90.18	91.22
	LR	96.87	97.05	96.18	96.87
	BNB	86.53	87.45	86.06	86.53
VGG16	LDA	94.11	94.73	94.11	94.20
	NuSVM	95.31	95.89	94.10	95.31
	LR	89.23	90.20	89.10	89.23
	BNB	77.23	78.44	84.00	77.23
VGG19	LDA	96.87	97.53	96.10	96.80
	NuSVM	89.23	90.20	89.10	89.23
	LR	85.43	86.40	85.12	85.43
	BNB	90.62	91.53	90.12	90.62
HOG	LDA	66.66	67.77	67.10	66.66
	NuSVM	93.75	93.92	93.79	93.75
	LR	87.50	88.14	87.66	87.50
	BNB	65.67	66.53	65.89	65.67
LBP	LDA	77.08	76.49	75.79	77.08
	NuSVM	85.41	85.34	84.98	85.41
	LR	66.66	69.83	68.53	66.66
	BNB	68.75	67.74	67.83	68.75

multi-level feature extraction with classifier fusion, significantly improved performance. Specifically, PH2 achieved an accuracy of 96.00% using MobileNet with NuSVM+LR, as identified during the binary (benign vs. malignant) and sub-class (ANV vs. CNV) stages. For HAM10000, an accuracy of 85.43% was obtained using ResNet-50 with LDA+NuSVM/NuSVM classifiers derived from previous phases, demonstrating a substantial improvement over direct multiclass classification (see Tables 14 and 15).

To facilitate interpretation, confusion matrices are provided for each classification task, illustrating model behavior across lesion categories (see Tables 16, 17, 18, 19, 20, 21, 22, 25, 26, 23, and 24).

Although pre-trained CNN models introduce higher computational cost, their superior accuracy and stability make them particularly suitable for clinical skin lesion analysis.

Table 7: Comparison of binary classification methods on **DermIS** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	75.00	76.20	74.70	75.00
	NuSVM	85.71	88.88	85.41	85.71
	LR	85.71	86.45	85.64	85.71
	BNB	85.71	88.88	85.41	85.71
ResNet50	LDA	75.00	78.12	73.33	75.00
	NuSVM	93.75	94.53	93.77	93.75
	LR	87.50	87.50	87.50	87.50
	BNB	84.37	84.63	84.42	84.37
VGG16	LDA	71.42	73.33	70.80	71.42
	LR	96.42	96.42	96.42	96.42
	NuSVM	92.85	92.85	92.85	92.85
	BNB	85.71	86.41	85.64	85.71
VGG19	LDA	82.14	82.30	82.12	82.14
	NuSVM	96.42	96.42	96.42	96.42
	LR	96.42	96.42	96.42	96.42
	BNB	89.28	89.48	89.27	89.28
HOG	LDA	51.23	52.34	51.73	51.23
	NuSVM	60.71	61.22	60.25	60.71
	LR	58.62	59.23	58.69	58.62
	BNB	55.52	55.98	55.76	55.52
LBP	LDA	67.85	67.94	67.81	67.85
	NuSVM	57.14	57.77	56.25	57.14
	LR	60.71	61.22	60.25	60.71
	BNB	60.20	61.20	60.16	60.20

Experimental results indicate that deep CNN architectures—including MobileNet, ResNet-50, VGG19, and VGG16—consistently outperform traditional handcrafted descriptors (HOG and LBP) in both binary and multi-class scenarios, highlighting their superior capacity to capture complex visual patterns inherent in dermoscopic imagery.

Overall, pre-trained CNNs provide the most reliable feature representations for medical image classification, and the incorporation of a hierarchical two-phase strategy substantially enhances system performance. Comparative visualizations of the direct and hierarchical approaches are presented in

Table 8: Comparison of binary classification methods on **DermQuest** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	75.00	75.39	74.90	75.00
	NuSVM	84.37	84.50	84.35	84.37
	LR	87.50	88.09	87.45	87.50
	BNB	81.25	83.33	80.95	81.25
ResNet50	LDA	90.62	90.78	90.61	90.62
	NuSVM	90.62	90.78	90.61	90.62
	LR	90.62	90.78	90.61	90.62
	BNB	81.25	81.74	81.17	81.25
VGG16	LDA	67.89	68.82	67.95	67.89
	LR	90.62	90.78	90.61	90.62
	NuSVM	87.50	88.09	87.45	87.50
	BNB	71.87	72.67	71.62	71.87
VGG19	LDA	78.12	78.23	78.10	78.12
	NuSVM	90.62	90.78	90.61	90.62
	LR	84.37	84.50	84.35	84.37
	BNB	75.00	75.39	74.90	75.00
HOG	LDA	56.67	57.83	56.90	56.67
	NuSVM	52.73	53.78	52.56	52.73
	LR	56.25	56.63	56.52	56.25
	BNB	55.12	55.67	55.56	55.12
LBP	LDA	68.75	71.81	68.61	68.75
	NuSVM	68.75	70.00	68.25	68.75
	LR	68.75	69.04	68.62	68.75
	BNB	56.67	57.83	56.90	56.67

Figures 4, 5, and 6.

7.2. Comparison with Existing Methods

As reported in Table 27, the proposed framework achieves 91.42% accuracy for binary classification on HAM10000, matching or surpassing state-of-the-art methods such as Waheed et al. [31] (91.00%) and Dheeb et al. [32] (91.05%). These results indicate that conventional deep architectures, including ResNet-50 and VGG-16, are sufficient for binary lesion discrimination without requiring highly complex models.

Table 9: Benign lesion classification results (ANV, CNV) on **PH2** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	68.75	69.04	68.62	68.75
	NuSVM	90.65	90.65	90.65	90.65
	LR	90.65	90.65	90.65	90.65
	BNB	75.00	75.79	74.77	75.00
ResNet50	LDA	70.31	70.44	70.24	70.31
	NuSVM	82.81	83.10	82.77	82.81
	LR	79.68	81.17	79.44	79.68
	BNB	60.37	60.41	60.33	59.37
VGG16	LDA	62.82	63.30	62.76	62.82
	NuSVM	79.68	79.95	79.64	79.68
	LR	85.93	85.98	85.93	85.93
	BNB	67.18	67.33	67.11	67.18
VGG19	LDA	61.11	62.21	61.02	61.11
	NuSVM	81.25	83.38	81.08	81.25
	LR	75.00	75.09	74.97	75.00
	BNB	62.50	62.54	62.46	62.50
HOG	LDA	51.82	52.20	51.78	51.82
	NuSVM	61.71	62.52	61.67	61.71
	LR	60.72	61.09	60.45	60.72
	BNB	53.92	54.21	53.86	53.92
LBP	LDA	64.06	64.07	64.04	64.06
	NuSVM	58.25	58.34	56.07	56.25
	LR	62.25	62.54	62.46	62.25
	BNB	67.18	67.20	67.17	67.18

In contrast, multi-class classification (see Table 28) remains more challenging, with reported accuracies typically ranging from 73.00% to 85.00% [15, 17, 33, 34, 35]. Our two-phase hierarchical strategy attains 85.43%, demonstrating improved class consistency. By progressively narrowing the decision space and reusing CNN-based descriptors within a structured pipeline, the proposed architecture enhances discrimination among visually similar lesion categories.

On the PH2 dataset (see Table 29), our model achieves 98.85% accuracy in binary classification, slightly exceeding the best reported result [40] (98.80%), further validating its robustness. For multi-class (see Table 30) classification,

Table 10: Benign lesion classification results (BKL, DF, NV, VASC) on **HAM10000** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	86.25	86.43	86.30	86.25
	NuSVM	95.00	95.17	95.02	95.00
	LR	91.25	91.45	91.30	91.25
	BNB	61.87	64.10	62.20	61.25
ResNet50	LDA	64.37	64.53	64.80	64.37
	NuSVM	91.87	91.97	91.78	91.87
	LR	90.62	90.52	90.44	90.62
	BNB	78.18	78.56	78.18	78.35
VGG16	LDA	64.35	65.60	64.23	64.35
	NuSVM	94.37	94.44	94.35	94.37
	LR	88.75	88.86	88.64	88.75
	BNB	66.87	69.71	66.12	66.87
VGG19	LDA	85.78	86.93	85.67	85.78
	NuSVM	91.87	92.45	91.87	91.87
	LR	82.50	83.59	82.60	82.50
	BNB	70.68	70.71	70.30	70.68
HOG	LDA	51.40	55.30	52.80	51.40
	NuSVM	53.20	55.80	52.10	53.20
	LR	53.90	57.92	52.10	53.80
	BNB	52.60	54.70	51.22	52.60
LBP	LDA	52.50	56.79	51.31	52.50
	NuSVM	53.80	57.78	52.80	53.80
	LR	52.10	55.80	52.30	52.10
	BNB	52.84	56.90	51.35	52.84

an accuracy of 96.00% is obtained, comparable to Dahou et al. [11] (96.43%) and Afza et al. [19] (95.40%). The hierarchical framework again demonstrates strong generalization, particularly in distinguishing between nevi subtypes and melanoma, while leveraging multiple classifiers within a unified CNN-based feature space.

For DermIS and DermQuest (Tables 31 and 32), the proposed method achieves accuracies of 96.42% and 90.62%, respectively, outperforming existing approaches such as Ashraf et al. [45] and Narayan et al. [46]. Although binary classification does not inherently require a hierarchical design, these

Table 11: Malignant lesion classification results (AKIEC, BCC, MEL) on **HAM10000** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	54.89	56.89	54.70	54.89
	NuSVM	67.29	69.05	65.72	67.29
	LR	64.15	64.89	61.79	64.15
	BNB	64.77	65.51	63.49	64.77
ResNet50	NuSVM	70.44	72.52	69.79	70.44
	LDA	64.56	66.70	61.63	64.56
	LR	66.08	67.61	63.40	66.08
	BNB	65.44	66.30	64.67	65.44
VGG16	LDA	58.57	60.32	57.89	58.57
	NuSVM	61.63	62.08	59.06	61.63
	LR	61.63	63.39	59.90	61.63
	BNB	52.20	51.90	51.67	52.20
VGG19	LDA	56.78	58.53	56.20	56.78
	NuSVM	60.37	62.39	58.49	60.37
	LR	65.40	66.87	64.26	65.40
	BNB	52.20	51.90	51.67	52.20
HOG	LDA	52.20	51.80	50.90	52.20
	NuSVM	51.40	52.10	51.10	52.10
	LR	53.32	53.90	52.80	53.32
	BNB	50.70	51.90	50.80	50.70
LBP	LDA	51.89	52.80	50.70	51.89
	NuSVM	50.70	51.90	50.78	50.70
	LR	51.09	52.40	50.10	51.09
	BNB	51.97	52.41	50.83	51.87

results highlight the robustness of the VGG-19 backbone, even when applied to smaller and more heterogeneous datasets.

Overall, performance variations across datasets underline the impact of data quality and class imbalance, reinforcing the importance of flexible and modular architectures. These comparative results confirm the relevance and effectiveness of the proposed system across diverse experimental settings.

The performance improvement achieved by the hierarchical strategy can be attributed to progressive reduction of class ambiguity. The initial binary

Table 12: Multi-class classification results (seven classes) on **HAM10000** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	61.07	61.53	60.48	61.07
	NuSVM	66.07	66.52	64.82	66.07
	LR	62.85	62.39	62.02	62.85
	BNB	55.92	56.40	53.75	54.92
ResNet50	LDA	63.37	63.53	62.80	63.37
	LR	70.87	70.97	70.78	70.87
	NuSVM	73.4	73.6	74.6	73.5
	BNB	65.38	65.56	64.18	65.35
VGG16	LDA	65.00	64.64	64.33	64.35
	NuSVM	66.42	66.55	65.60	66.42
	LR	64.64	64.02	63.80	64.64
	BNB	54.92	55.92	53.75	54.92
VGG19	LDA	62.85	64.31	62.79	62.85
	NuSVM	65.00	65.19	64.38	65.00
	LR	63.57	63.48	62.82	63.57
	BNB	60.54	60.65	60.20	60.54
HOG	LDA	50.63	50.78	49.90	50.63
	NuSVM	52.72	53.17	52.65	52.72
	LR	50.94	51.89	50.45	50.94
	BNB	51.21	52.32	50.29	51.21
LBP	LDA	49.82	50.82	50.01	49.82
	NuSVM	52.06	52.92	51.96	52.06
	LR	54.92	55.46	54.86	54.92
	BNB	48.91	49.25	48.82	48.91

stage filters visually dissimilar samples, allowing subsequent classifiers to operate on more homogeneous feature distributions, thereby improving inter-class separability.

Unlike several state-of-the-art approaches relying on heavy end-to-end deep models, our framework combines frozen CNN feature extractors with lightweight classifiers, offering a favorable trade-off between performance and computational complexity.

Table 13: Multi-class classification results (three classes) on **PH2** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	LDA	71.87	73.96	71.35	71.87
	NuSVM	79.76	80.77	78.55	78.76
	LR	79.76	80.77	78.55	79.44
	BNB	60.71	62.51	60.49	60.71
ResNet50	LR	87.45	87.35	87.57	87.45
	LDA	85.41	85.71	85.29	85.41
	NuSVM	82.30	82.47	82.34	82.30
	BNB	67.77	68.21	67.39	67.77
VGG16	LDA	58.57	60.32	57.89	58.57
	NuSVM	84.37	84.51	83.89	84.37
	LR	81.25	82.00	80.65	81.25
	BNB	79.16	78.71	78.78	79.16
VGG19	LDA	82.30	83.17	82.47	82.30
	NuSVM	85.30	85.65	84.72	85.30
	LR	81.25	81.05	80.89	81.25
	BNB	62.45	63.65	62.07	62.45
HOG	LDA	49.76	51.80	49.09	49.76
	NuSVM	57.29	58.83	56.83	57.29
	LR	52.32	52.90	51.80	52.32
	BNB	50.43	51.87	50.30	50.43
LBP	LDA	48.56	50.63	49.42	48.56
	NuSVM	50.20	51.10	50.07	50.20
	LR	47.78	49.23	48.90	47.78
	BNB	48.89	49.98	48.87	48.89

Table 14: Two-phase *multi-class* classification results on **HAM10000** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	BNB+NuSVM/NuSVM	84.30	84.10	84.00	84.40
ResNet50	LDA+NuSVM/NuSVM	85.43	85.40	86.40	83.50
VGG16	NuSVM+NuSVM/NuSVM	83.30	83.36	83.75	84.10
VGG19	NuSVM+NuSVM/LR	82.80	82.90	82.87	83.50
HOG	NuSVM+LR/LR	66.80	68.30	65.50	47.03
LBP	LDA+NuSVM/LDA	41.11	42.33	41.41	41.11

Table 15: Two-phase *multi-class* classification results on **PH2** dataset.

METHOD	CLASSIFIER	ACCURACY (%)	PRECISION (%)	F1-SCORE (%)	RECALL (%)
MobileNet	NuSVM+LR	96.00	95.90	95.89	95.93
ResNet50	LR+NuSVM	95.30	95.50	95.20	95.70
VGG16	NuSVM+LR	90.22	90.34	90.13	90.34
VGG19	LDA+NuSVM	93.76	93.10	93.91	93.10
HOG	NuSVM+NuSVM	88.98	89.12	88.56	88.98
LBP	NuSVM+BNB	68.75	67.73	67.50	76.66

Table 16: **HAM10000** (*binary*).

	BENIGN	MALIGNANT
BENIGN	84.30	15.70
MALIGNANT	1.47	98.53

Table 17: **PH2** (*binary*).

	BENIGN	MALIGNANT
BENIGN	100.00	0.00
MALIGNANT	3.13	96.87

Table 18: **DermIS** (*binary*).

	BENIGN	MALIGNANT
BENIGN	92.86	7.14
MALIGNANT	0.00	100.00

Table 19: **DermQuest** (*binary*).

	BENIGN	MALIGNANT
BENIGN	95.00	5.00
MALIGNANT	10.00	90.00

Table 20: **HAM10000** (*malignant*).

	AKIEC	BCC	MEL
AKIEC	79.25	13.21	7.55
BCC	66.04	32.08	1.89
MEL	5.66	9.43	84.91

Table 21: **HAM10000** (*benign*).

	BKL	DF	NV	VASC
BKL	95.00	2.50	0.00	2.50
DF	2.50	97.50	0.00	0.00
NV	2.50	0.00	97.50	0.00
VASC	0.00	10.00	0.00	90.00

Table 22: **PH2** (*benign*).

	ANV	CNV
ANV	90.63	9.38
CNV	9.38	90.63

7.3. Limitations

Limitations include reliance on segmentation masks, reduced performance on highly heterogeneous malignant subclasses, and the use of augmented samples. Future work will address these aspects through weakly supervised learning and uncertainty modeling.

Table 23: **PH2** (*multi-class*).

	ANV	CNV	MEL
ANV	78.13	15.63	6.25
CNV	15.63	84.38	0.00
MEL	3.13	0.00	96.87

Table 24: **PH2** (*multi-class, two-phase*).

	ANV	CNV	MEL
ANV	100.00	0.00	0.00
CNV	6.25	93.75	0.00
MEL	0.00	3.13	96.88

Table 25: **HAM10000** (*multi-class*).

	BKL	DF	NV	VASC	AKIEC	BCC	MEL
BKL	62.50	0.00	0.00	2.50	0.00	5.00	30.00
DF	5.00	62.50	0.00	7.50	12.50	5.00	7.50
NV	2.50	0.00	97.50	0.00	0.00	0.00	0.00
VASC	2.50	0.00	0.00	97.50	0.00	0.00	0.00
AKIEC	2.50	5.00	5.00	0.00	80.00	7.50	0.00
BCC	15.00	10.00	2.50	7.50	17.50	45.00	2.50
MEL	7.50	5.00	0.00	0.00	5.00	10.00	72.50

8. Conclusion and Remarks

This study demonstrates that combining pre-trained CNN architectures with a two-phase hierarchical decision strategy provides an effective framework for automated skin lesion classification. The integration of deep feature extractors (ResNet-50, VGG-16, VGG-19, and MobileNet) with handcrafted descriptors (HOG and LBP), together with classical classifiers (LR, LDA, BNB, and NuSVM), enables the exploitation of both rich deep representations and the flexibility of traditional machine learning approaches.

The proposed hierarchical framework, which first performs binary classification followed by guided multiclass discrimination, significantly reduces the decision space and enhances system robustness against the intrinsic variability and complexity of dermoscopic images.

These methodological contributions reinforce the relevance of hybrid DL–ML paradigms for computer-aided diagnosis (CAD), paving the way toward efficient and clinically adaptable skin lesion assessment systems. Future work will focus on extending the hierarchical architecture through multi-task learning, incorporating uncertainty modeling, and developing strategies to better handle noisy and weakly labeled data.

The hierarchical reuse of CNN features combined with cross-stage clas-

Table 26: **HAM10000** (*multi-class, two-phase*).

	BKL	DF	NV	VASC	AKIEC	BCC	MEL
BKL	100.00	0.00	0.00	0.00	0.00	0.00	0.00
DF	0.00	100.00	0.00	0.00	0.00	0.00	0.00
NV	0.00	0.00	50.00	0.00	30.00	7.50	12.50
VASC	2.50	0.00	2.50	87.50	0.00	0.00	7.50
AKIEC	0.00	0.00	0.00	0.00	100.00	0.00	0.00
BCC	0.00	0.00	0.00	0.00	5.00	85.00	10.00
MEL	2.50	0.00	0.00	15.00	0.00	0.00	82.50

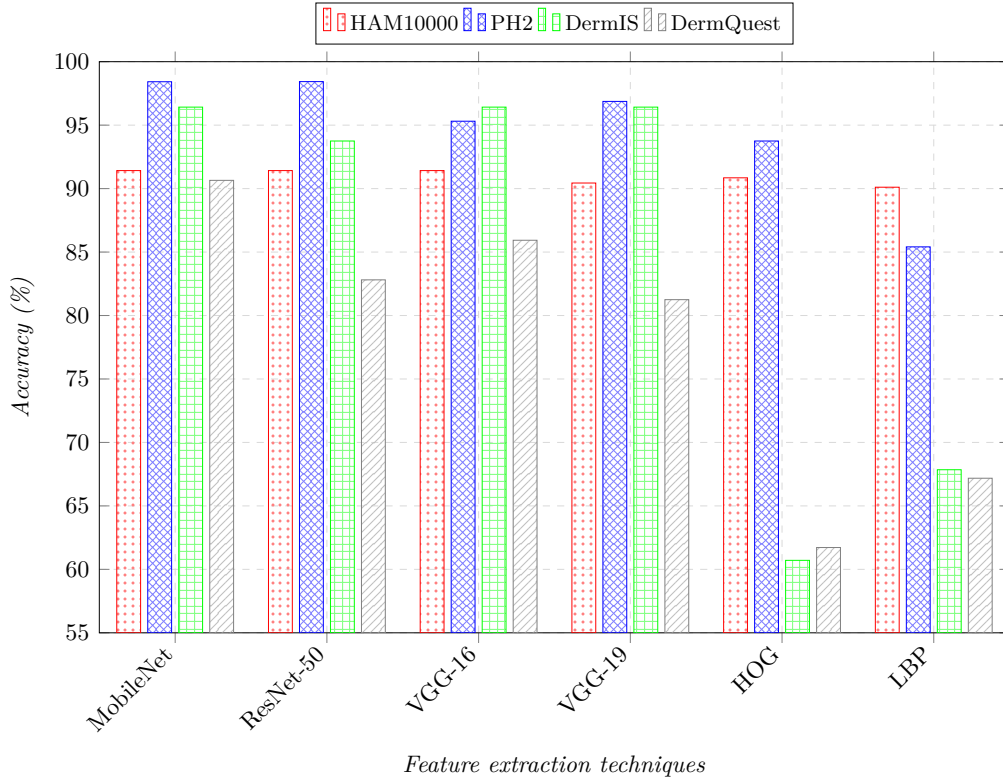


Figure 4: Binary classification accuracy (%) across feature extraction techniques.

sifier fusion constitutes the core contribution of this work, demonstrating that structured decision pipelines can outperform flat multiclass strategies in complex medical imaging scenarios.

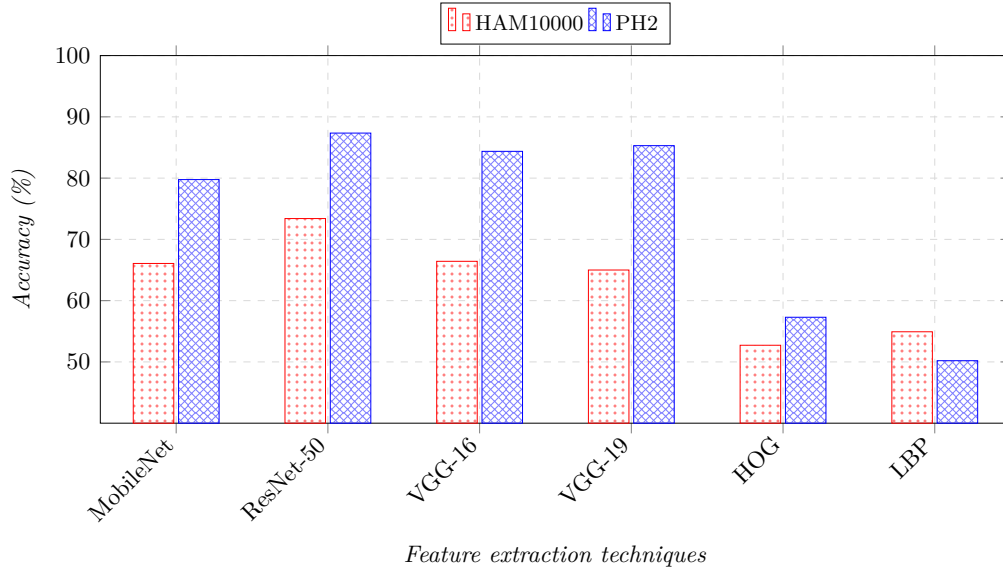


Figure 5: Two-class classification accuracy (%) across feature extraction techniques.

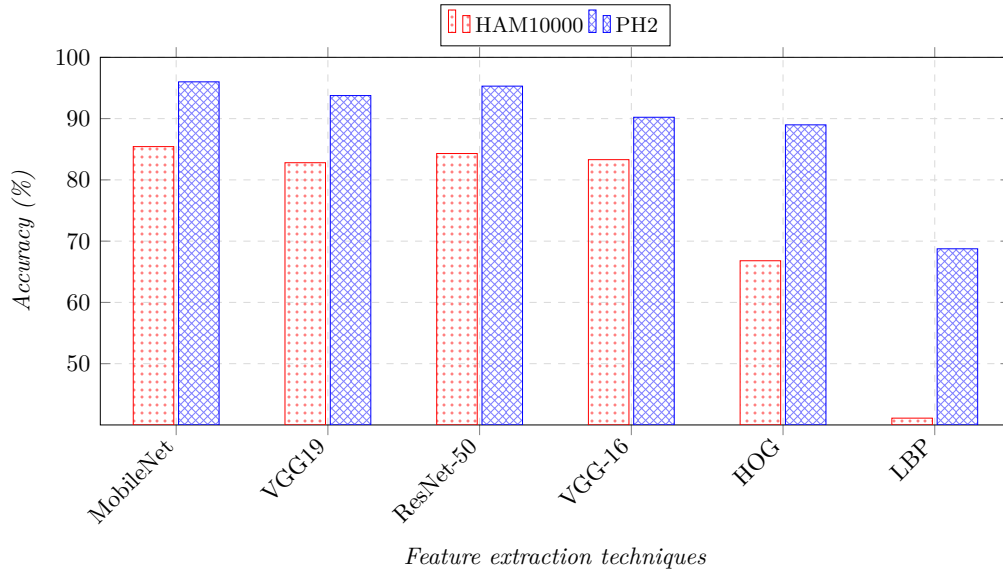


Figure 6: Two-phase *multi-class* classification accuracy (%) across feature extraction techniques.

Table 27: Comparative evaluation of the performance of the proposed system against *state-of-the-art* methods on the **HAM10000** dataset for *binary* classification.

REFERENCE	YEAR	APPROACH	ACCURACY
Ameri et al. [36]	2020	AlexNet	84.00%
Sae-Lim et al. [37]	2019	MobileNet	83.90%
Afza et al. [19]	2022	ResNet-50	85.50%
Dheeb et al. [32]	2025	VGG-16	91.05%
D. P. Sahoo et al. [38]	2022	ResNet50	89.30%
Waheed et al. [31]	2023	CNN	91.00%
Our method	2026	ResNet50, VGG-16	91.42%

Table 28: Comparative evaluation of the performance of the proposed system against *state-of-the-art* methods on the **HAM10000** dataset for *multi-class* classification.

REFERENCE	YEAR	APPROACH	ACCURACY
Fraiwan et al. [15]	2022	DenseNet201	73.50%
Bassi et al. [17]	2019	Ensemble Learning	82.80%
Chaturvedi et al. [33]	2021	MobileNet	83.10%
Mobiny et al. [34]	2019	DenseNet-169	83.59%
Kanani et al. [35]	2019	CNN	77.98%
Gouda et al. [39]	2022	InceptionV3	85.00%
Our method	2026	Two-phase	85.43%

Table 29: Comparative evaluation of the performance of the proposed system against *state-of-the-art* methods on the **PH2** dataset for *binary* classification.

REFERENCE	YEAR	APPROACH	ACCURACY
Nazi et al.[12]	2020	DCNN+SVM	92.00%
Nawaz et al.[41]	2022	Faster-RCNN	96.10%
Akram et al[40]	2020	CNN	98.80%
Adegun et al.[42]	2020	CNN	95.00%
Abayomi et al.[18]	2021	SqueezeNet	92.18%
Our method	2026	ResNet50	98.85%

9. Acknowledgements

The authors would like to thank the providers of the four benchmark skin lesion datasets—HAM10000 [7], PH2 [8], DermIS [29], and DermQuest

Table 30: Comparative evaluation of the performance of the proposed system against *state-of-the-art* methods on the **PH2** dataset for *multi-class* classification.

REFERENCE	YEAR	APPROACH	ACCURACY
Abayomi et al.[18]	2021	SqueezeNet	66.40%
Arora[43]	2022	SURFs+SVM	85.70%
Rodrigues et al.[44]	2020	DenseNet201+KNN	93.16%
Afza et al.[19]	2022	ResNet-50	95.40%
Dahou et al.[11]	2023	DOLHGS	96.43%
Our method	2026	Two-phase	96.00%

Table 31: Comparative evaluation of the performance of the proposed system against *state-of-the-art* methods on the **DermIS** dataset for *binary* classification.

REFERENCE	YEAR	APPROACH	ACCURACY
hosny et al. [14]	2022	RDCNN	90.30%
Ashraf et al. [47]	2020	ROI+CNN	97.40%
Razmjoooy et al. [46]	2023	CNN+MAFBUZO	94.00%
Dhar et al. [48]	2021	CNN	85.00%
Nayaran et al. [45]	2024	ROI+CNN	70.40%
Our method	2026	VGG-19	96.42%

Table 32: Comparative evaluation of the performance of the proposed system against *state-of-the-art* methods on the **DermQuest** dataset for *binary* classification.

REFERENCE	YEAR	APPROACH	ACCURACY
Hosny et al. [14]	2022	RDCNN	90.30%
Selvarasa et al. [49]	2020	Business tool	88.00%
Nayaran et al. [45]	2024	ROI+CNN	91.40%
Dhar et al. [48]	2021	CNN	85.00%
Zhang et al. [6]	2020	CNN-IWOA	85.70%
Our method	2026	VGG-19	90.62%

[30]—which were used to evaluate the performance of the proposed framework.

The authors also acknowledge the use of generative AI tools to assist with grammatical correction, language refinement, and clarification of the manuscript content. The scientific content, experimental design, and conclusions remain entirely the responsibility of the authors.

References

- [1] Société canadienne du cancer, Skin cancer information (2023).
URL <https://cancer.ca/en/cancer-information/cancer-types/skin-non-melanoma/risks>
- [2] Gouvernement du Canada, Sun safety (8 2023).
URL <https://www.canada.ca/en/public-health/services/sun-safety/skin-cancer.html>
- [3] Organisation mondiale de la Santé, Working under the sun causes 1 in 3 deaths from non-melanoma skin cancer, say who and ilo, accessed on February 22, 2025 (11 2023).
URL <https://www.who.int/news/item/08-11-2023-working-under-the-sun-causes-1-in-3-deaths-from-non-melanoma-skin-cancer--say-who-and-ilo>
- [4] R. Dorey, M. Crawford, Artificial intelligence useful in skin cancer screening (2024).
URL <https://ici.radio-canada.ca/nouvelle/2039816/cancer-peau-intelligence-artificielle-medecine>
- [5] S. Thrun, When artificial intelligence detects skin cancer (2017).
URL <https://ici.radio-canada.ca/nouvelle/1013185/intelligence-artificielle-detecte-cancer-peau-universite-standford>
- [6] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network, *Artificial Intelligence in Medicine* 102 (2020) 101756. doi: 10.1016/j.artmed.2019.101756.
- [7] P. Tschandl, The ham10000 dataset: A large collection of multi-source dermoscopic images of common pigmented skin lesions (2018). doi: 10.7910/DVN/DBW86T.
- [8] T. Mendonça, P. M. Ferreira, J. Marques, A. R. S. Marcal, J. Rozeira, PH² - a dermoscopic image database for research and benchmarking, in: *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan,

2013. doi:10.1109/EMBC.2013.6610779.

URL <https://www.fc.up.pt/addi/ph2%20database.html>

- [9] A. Murugan, S. A. H. Nair, K. P. S. Kumar, Detection of skin cancer using svm, random forest, and knn classifiers, *Journal of Medical Systems* 43 (7) (2019) 269. doi:10.1007/s10916-019-1400-8.
- [10] R. Ashraf, S. Afzal, A. U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O.-Y. Song, M. Maqsood, Region-of-interest based transfer learning assisted framework for skin cancer detection, *IEEE Access* 8 (2020) 147858–147871. doi:10.1109/ACCESS.2020.3014701.
- [11] A. Dahou, A. O. Aseeri, A. Mabrouk, R. A. Ibrahim, M. A. Al-Betar, M. A. Elaziz, Optimal skin cancer detection model using transfer learning and hunger games opposed dynamic search, *Diagnostics* 13 (9) (2023) 1579. doi:10.3390/diagnostics13091579.
- [12] Z. Nazi, T. Abir, Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with u-net and dcnn-svm, in: M. Uddin, J. Bansal (Eds.), *Proceedings of the International Joint Conference on Computational Intelligence, Algorithms for Intelligent Systems*, Springer, Singapore, 2020, pp. 369–378. doi:10.1007/978-981-13-7564-4_32.
- [13] C. Xin, Z. Liu, K. Zhao, L. Miao, Y. Ma, X. Zhu, Q. Zhou, S. Wang, L. Li, F. Yang, S. Xu, H. Chen, An improved transformer network for skin cancer classification, *Computers in Biology and Medicine* 149 (2022) 105939. doi:10.1016/j.combiomed.2022.105939.
- [14] K. M. Hosny, M. A. Kassem, Refined deep residual convolutional network for skin lesion classification, *Journal of Digital Imaging* 35 (2) (2022) 258–280. doi:10.1007/s10278-021-00552-0.
- [15] M. Fraiwan, E. Faouri, On automatic skin cancer detection and classification using deep transfer learning, *Sensors* 22 (13) (2022) 4963. doi:10.3390/s22134963.
- [16] M. A. Elaziz, A. Dahou, A. Mabrouk, S. El-Sappagh, A. O. Aseeri, An efficient artificial rabbits optimization based on mutation strategy for skin cancer prediction, *Computers in Biology and Medicine* 163 (2023) 107154. doi:10.1016/j.combiomed.2023.107154.

- [17] S. Bassi, A. Gomekar, Deep learning diagnosis of pigmented skin lesions, in: 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019, pp. 1–6. doi:10.1109/ICCCNT45670.2019.8944601.
- [18] O. O. Abayomi-Alli, R. Damasevicius, S. Misra, R. Maskeliunas, A. Abayomi-Alli, Detection of malignant skin melanoma using image augmentation by oversampling in a variety of lower-dimensional nonlinear embedding, Turkish Journal of Electrical Engineering and Computer Sciences 29 (8) (2021) 4020–4036. doi:10.3906/elk-2101-133.
- [19] F. Afza, M. Sharif, M. Mittal, M. A. Khan, D. J. Hemanth, A hierarchical three-step superpixels and deep learning framework for skin lesion classification, Methods 202 (2022) 88–102, machine Learning Methods for Bio-Medical Image and Signal Processing: Recent Advances. doi:10.1016/j.ymeth.2021.02.013.
- [20] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint (2014). arXiv:1409.1556, doi:10.48550/arXiv.1409.1556.
- [21] G. Boesch, Essential guide to very deep convolutional networks (vgg) (2021).
URL <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/>
- [22] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, arXiv preprint (2015). arXiv:1512.03385.
- [23] X. Gu, A. Salehi, L. Wang, P. J. Coates, N. Sgaramella, K. Nylander, Early detection of squamous cell carcinoma of the oral tongue using multidimensional plasma protein analysis and interpretable machine learning, Journal of Oral Pathology & Medicine 52 (7) (2023) 637–643. doi:10.1111/jop.13461.
- [24] L. Ichim, D. Popescu, Melanoma detection using an objective system based on multiple connected neural networks, IEEE Access 8 (2020) 179189–179202. doi:10.1109/ACCESS.2020.3028248.
- [25] F. Grignaffini, F. Barbuto, L. Piazzo, et al., Machine learning approaches for skin cancer classification from dermoscopic images: A

- systematic review, *Algorithms* 15 (11) (2022) 438. doi:10.3390/a15110438.
- [26] L. Qin, W. Zeng, W. Gao, W. Wang, Local invariant descriptor for image matching, in: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 2, 2005, pp. 1025–1028. doi:10.1109/ICASSP.2005.1415582.
- [27] Pandas Development Team, *Pandas: Utilisation des données manquantes, documentation pandas 0.22.0 Edition, documentation officielle de la bibliothèque Pandas pour la gestion des valeurs manquantes* (2019).
URL https://pandas.pydata.org/pandas-docs/stable/user_guide/missing_data.html
- [28] K. Pearson, Liii. on lines and planes of closest fit to systems of points in space, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2 (11) (1901) 559–572. doi:10.1080/14786440109462720.
- [29] Département de médecine sociale clinique (Université de Heidelberg) and Département de Dermatologie (Université de Erlangen), *Dermatology information system* (2012).
URL <http://www.dermIS.com/>
- [30] Département de médecine sociale clinique (Université de Heidelberg) and Département de Dermatologie (Université de Erlangen), *Dermquest* (2012).
URL <http://www.dermquest.com/>
- [31] S. R. Waheed, S. M. Saadi, M. S. M. Rahim, N. M. Suaib, F. H. Najjar, M. M. Adnan, A. A. Salim, Melanoma skin cancer classification based on cnn deep learning algorithms, *Malaysian Journal of Fundamental and Applied Sciences* 19 (3) (2023) 299–305. doi:10.11113/mjfas.v19n3.2900.
- [32] D. Albashish, N. Almansour, A. Abdullah, H. M. J. Mustafa, M. R. AlSayed, O. Alrashdan, Design an ensemble pretrained deep learning model for classification of melanoma skin cancer images, in:

- 2025 1st International Conference on Computational Intelligence Approaches and Applications (ICCIAA), 2025, pp. 1–7. doi:10.1109/ICCIAA65327.2025.11013331.
- [33] S. Chaturvedi, K. Gupta, P. Prasad, Skin lesion analyzer: A seven-way multi-class skin cancer classification using mobilenet, in: A. Hassanien, R. Bhatnagar, A. Darwish (Eds.), *Advanced Machine Learning Technologies and Applications*, Vol. 1141 of *Advances in Intelligent Systems and Computing*, Springer, Singapore, 2021, pp. 179–189. doi:10.1007/978-981-15-3383-9_15.
 - [34] A. Mobiny, A. Singh, H. Van Nguyen, Risk-aware machine learning classifier for skin lesion diagnosis, *Journal of Clinical Medicine* 8 (8) (2019) 1241. doi:10.3390/jcm8081241.
 - [35] P. Kanani, M. Padole, Deep learning to detect skin cancer using google colab, *International Journal of Engineering and Advanced Technology* 8 (6) (2019) 2176–2183. doi:10.35940/ijeat.F8587.088619.
 - [36] A. Amery, A deep learning approach for skin cancer detection in dermoscopy images, *Journal of Biomedical Physics and Engineering* 10 (6) (2020) 801–806. doi:10.31661/jbpe.v0i0.2004-1107.
 - [37] W. Sae-Lim, W. Wettayaprasit, P. Aiyarak, Convolutional neural networks using mobilenet for skin lesion classification, in: *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2019. doi:10.1109/JCSSE.2019.8864155.
 - [38] D. Sahoo, M. Rout, P. K. Mallick, S. Samanta, Comparative analysis of medical images using transfer learning based deep learning models, in: *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022, pp. 1–8. doi:10.1109/ASSIC55218.2022.10088373.
 - [39] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, N. Z. Jhanjhi, Skin cancer detection based on skin lesion images using deep learning, *Healthcare* 10 (7) (2022) 1183. doi:10.3390/healthcare10071183.
 - [40] T. Akram, H. M. J. Lodhi, S. R. Naqvi, M. Rizwan, M. Ramzan, M. Imran, A multilevel features selection framework for skin lesion classifica-

- tion, *Human-centric Computing and Information Sciences* 10 (2020) 12. doi:10.1186/s13673-020-00216-y.
- [41] M. Nawaz, Z. Mehmood, T. Nazir, R. A. Naqvi, A. U. Rehman, M. Iqbal, T. Saba, Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering, *Microscopy Research and Technique* 85 (1) (2022) 339–351. doi:10.1002/jemt.23908.
 - [42] A. A. Adegun, S. Viriri, Deep learning-based system for automatic melanoma detection, *IEEE Access* 8 (2020) 7160–7172. doi:10.1109/ACCESS.2019.2962812.
 - [43] G. Arora, A. K. Dubey, Z. A. Jaffery, et al., Bag of features and support vector-based early diagnosis of skin cancer, *Neural Computing and Applications* 34 (10) (2022) 8385–8392. doi:10.1007/s00521-020-05212-y.
 - [44] D. d. A. Rodrigues, R. F. Ivo, S. C. Satapathy, S. Wang, J. Hemanth, P. P. Reboucas Filho, A new approach for classification skin lesion based on transfer learning, deep learning, and iot system, *Pattern Recognition Letters* 136 (2020) 8–15. doi:10.1016/j.patrec.2020.05.019.
 - [45] G. S. Narayan, M. R. Bendre, Skin cancer identification using cnn: A lightweight model with fine-grained classification, in: 2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSOciCon), 2024, pp. 1–6. doi:10.1109/MITADTSOciCon60330.2024.10575526.
 - [46] N. Razmjoooy, A. Arshaghi, Application of multilevel thresholding and cnn for the diagnosis of skin cancer utilizing a multi-agent fuzzy buzzard algorithm, *Biomedical Signal Processing and Control* 84 (2023) 104984. doi:10.1016/j.bspc.2023.104984.
 - [47] R. Ashraf, S. Afzal, A. U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O.-Y. Song, M. Maqsood, Region-of-interest based transfer learning assisted framework for skin cancer detection, *IEEE Access* 8 (2020) 147858–147871. doi:10.1109/ACCESS.2020.3014701.
 - [48] P. Dhar, S. Guha, Skin lesion detection using fuzzy approach and classification with cnn, *International Journal of Engineering and Manufacturing* 11 (1) (2021) 11–21. doi:10.5815/ijem.2021.01.02.

- [49] M. Selvarasa, A. Aponso, A critical analysis of computer aided approaches for skin cancer screening, in: 2020 International Conference on Image Processing and Robotics (ICIP), 2020, pp. 1–4. doi: 10.1109/ICIP48927.2020.9367370.

Dear Editor,

We are pleased to submit our manuscript entitled “*Hierarchical Two-Phase Skin Lesion Classification Using Reused Convolutional Neural Network Feature Spaces and Cross-Stage Classifier Fusion*” for consideration for publication in *Engineering Applications of Artificial Intelligence*.

This work addresses a practical challenge in AI-driven medical image analysis: improving multiclass skin lesion classification under visually similar categories and class imbalance, while maintaining computational efficiency suitable for real-world deployment. Although many existing approaches rely on flat multiclass prediction or fully end-to-end deep learning models, such strategies often suffer from class ambiguity, limited robustness, and high computational cost.

To overcome these limitations, we propose a hierarchical two-phase classification framework that progressively constrains the decision space. The approach reuses CNN feature representations across stages and introduces cross-stage classifier fusion to enhance class separability and system robustness. By structuring the decision process hierarchically—first performing benign–malignant discrimination, followed by refined subclass prediction—the framework improves reliability in complex multiclass scenarios while preserving interpretability.

The proposed method was evaluated on four publicly available datasets, demonstrating consistent improvements over direct multiclass classification. Importantly, our framework combines frozen CNN feature extractors with lightweight supervised classifiers, offering a favorable trade-off between predictive performance and computational efficiency. This design makes the approach particularly attractive for engineering-oriented applications and resource-constrained clinical decision support systems.

We believe this manuscript aligns well with the scope of *Engineering Applications of Artificial Intelligence*, as it presents a practical and structured AI framework that bridges deep feature learning with classical machine learning, emphasizing robustness, modularity, and deployability in real-world healthcare environments.

This manuscript is original, has not been published previously, and is not under consideration elsewhere. All authors have approved its submission to *Engineering Applications of Artificial Intelligence*. A preliminary version of this work has been accepted for presentation at ICATEEE’2025. The journal version substantially extends the conference paper by introducing a hierarchical two-phase framework, additional datasets, expanded multiclass evaluation, and deeper engineering analysis regarding robustness and computational efficiency.

Thank you for considering our manuscript. We look forward to your feedback.

Sincerely,

Prof. Yacine Yaddaden, Ph. D
Université du Québec à Rimouski
Campus de Lévis

*Hierarchical Two-Phase Skin Lesion Classification Using Reused Convolutional Neural Network
Feature Spaces and Cross-Stage Classifier Fusion*

Zakariaou Mounnie

M.Sc. student in Computer Science (*expected to graduate soon*)

University of Quebec at Rimouski

zakariaou.mounnie@uqar.ca

Yacine Yaddaden

Professor in computer science

University of Quebec at Rimouski

yacine_yaddaden@uqar.ca

Abdenour Bouzouane

Professor in computer science

University of Quebec at Chicoutimi

abouzoua@uqac.ca

Chapitre 5

Titre :

Un pipeline d'apprentissage profond hybride pour la détection du mélanome utilisant des caractéristiques U-Net modifiées et VGG-19.

Résumé – Le mélanome, forme la plus mortelle du cancer de la peau, demeure difficile à détecter précocement en raison de la variabilité des lésions et de la subjectivité du diagnostic. L'IA, en particulier l'apprentissage profond, offre des perspectives prometteuses pour améliorer la précision diagnostique et l'automatisation. Dans cette étude, nous proposons un cadre hybride robuste basé sur l'IA pour la détection automatisée du mélanome, intégrant des techniques de segmentation profonde et de classification supervisée. Une architecture U-Net personnalisée est employée pour une segmentation précise des lésions, soutenue par des opérations de prétraitement incluant l'augmentation de données, l'équilibrage des classes, l'amélioration du contraste, la normalisation et le redimensionnement. Des caractéristiques discriminantes de haut niveau sont extraites à l'aide des réseaux de neurones convolutifs CNNs pré-entraînés, et une réduction de dimensionnalité est appliquée via l'Analyse en Composantes Principales (PCA) afin d'améliorer l'efficacité des classifieurs. Nous évaluons le modèle sur plusieurs classifieurs supervisés, le Nu-Support Vector Machine (NuSVM) obtenant les meilleures performances. Les résultats expérimentaux sur les bases de données PH2 et HAM10000 démontrent l'efficacité du cadre proposé, avec des scores de Dice de 96,6 % et 95,3 % pour la segmentation, et des précisions de classification de 90,44 % et 96,87 %, respectivement. Ces résultats soulignent le potentiel de la combinaison de modèles de segmentation profonde avec des classifieurs optimisés basés sur des caractéristiques, afin d'améliorer la fiabilité et l'évolutivité des systèmes de diagnostic du mélanome.

Mots clés :

Mélanome, lésions cutanées, apprentissage profond, segmentation, U-Net, apprentissage par transfert, VGG-19, classification supervisée, Nu-SVM, images de dermoscopie, PCA.

Contributions associées :

Z. Mounmie, Y. Yaddaden, A. Bouzouane, "A Hybrid Deep Learning Pipeline for Melanoma Detection Using Modified U-Net and VGG-19 Features," in *Les 2Nd Conférence internationale sur les technologies avancées en génie électrique et électrique (ICATEEE2025)*, IEEE, 2025, pp. 1-7, (Statut : Publié)

A Hybrid Deep Learning Pipeline for Melanoma Detection Using Modified U-Net and VGG-19 Features

Zakariaou Mounmie
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
zakariaou.mounmie@uqar.ca

Yacine Yaddaden
Département de mathématiques,
informatique et génie
Université du Québec à Rimouski
Rimouski, Québec, Canada.
yacine_yaddaden@uqar.ca

Abdenour Bouzouane
Département d'informatique
et de mathématique
Université du Québec à Chicoutimi
Chicoutimi, Québec, Canada.
abouzoua@uqac.ca

Abstract—Melanoma, the deadliest form of skin cancer, remains challenging to detect early due to lesion variability and diagnostic subjectivity. Artificial Intelligence (AI), particularly deep learning, offers promising avenues to enhance diagnostic accuracy and automation. In this study, we propose a robust hybrid AI framework for automated melanoma detection, integrating deep segmentation and supervised classification techniques. A *custom* U-Net architecture is employed for precise lesion segmentation, supported by preprocessing operations including data augmentation, class balancing, contrast enhancement, normalization, and resizing. High-level discriminative features are extracted using transfer learning from *pre-trained* Convolutional Neural Networks (CNNs), and dimensionality reduction is applied via Principal Component Analysis (PCA) to improve classifier efficiency. We evaluate multiple supervised classifiers, with Nu-Support Vector Machine (NuSVM) achieving the best performance. Experimental results on the PH2 and HAM10000 datasets demonstrate the effectiveness of the proposed framework, with Dice segmentation scores of 96.6% and 95.3%, and classification accuracies of 90.44% and 96.87%, respectively. These results highlight the potential of combining deep segmentation models with optimized feature-based classifiers to enhance the reliability and scalability of melanoma diagnosis systems.

Index Terms—Melanoma detection, skin lesion classification, deep learning, U-Net segmentation, transfer learning, VGG-19, supervised classification, NuSVM, dermoscopy images, dimensionality reduction, PCA, medical image analysis.

I. INTRODUCTION

Skin cancer remains a major global health concern due to rising incidence, particularly malignant melanoma with its high metastatic potential. The World Health Organization reports over 132,000 new melanoma cases annually worldwide, with continuously increasing trends [1]. While non-melanocytic skin cancers are more prevalent, melanoma accounts for most skin cancer-related deaths due to its aggressive nature.

Since 2004, treatment advances have dramatically improved 5-year survival rates to nearly 99% for localized melanoma but only about 30% for metastatic cases [2]. Incidence rates are higher in women under 50, but greater in older men, reflecting differences in UV exposure patterns and healthcare access.

In 2022, the United States reported 99,780 new invasive melanoma cases and 7,650 deaths [2].

Current diagnosis relies on clinical examination, dermoscopy, and histopathology, yet faces challenges of subjective interpretation, complex lesion morphology, and invasive, costly analysis [3]. Artificial Intelligence—particularly Machine Learning and Deep Learning—has emerged as a promising solution to improve diagnostic accuracy and accessibility. Convolutional Neural Networks have demonstrated performance comparable to experts, though challenges like overfitting, generalization issues, and interpretability limitations remain [4].

We propose a multi-step AI framework for melanoma detection: lesion segmentation using *custom* U-Net [5]; preprocessing with data augmentation, class balancing, contrast enhancement, and normalization; feature extraction via transfer learning from pre-trained VGG-19 [6] with PCA dimensionality reduction; and classification using NuSVM, Logistic Regression, Bernoulli Naive Bayes, and Linear Discriminant Analysis.

Evaluation on publicly available dermoscopic datasets HAM10000 [7] and PH2 [8] demonstrates the framework's effectiveness in automated melanoma diagnosis.

The paper is organized as follows: Section II presents fundamental concepts; Section III reviews related work; Section IV describes the proposed method; Section V outlines datasets and evaluation protocol; Section VI reports and discusses results; and Section VII provides conclusions and future directions.

II. FUNDAMENTALS

With the rising incidence of skin cancer, early and accurate detection has become paramount. Automated analysis of dermoscopic images offers valuable support to dermatologists but continues to face significant challenges due to the high variability and morphological complexity of skin lesions.

In this study, we propose a hybrid framework that combines semantic segmentation with deep feature extraction from

pre-trained CNNs to enhance lesion classification. Semantic segmentation precisely isolates the Region of Interest (ROI), effectively minimizing background interference and focusing subsequent analysis on relevant areas. This isolation improves the accuracy of the classification step, which assigns diagnostic labels to lesions, thereby facilitating timely and reliable clinical decisions.

A. Semantic segmentation

Semantic segmentation enables pixel-level delineation of suspicious regions within dermoscopic images. The U-Net architecture, characterized by its encoder-decoder design and skip connections, excels at preserving spatial detail even with limited annotated data [5]. Its performance is further improved through preprocessing techniques, transfer learning, careful hyperparameter tuning, and regularization strategies such as dropout and early stopping [9].

B. Lesion classification

Following segmentation, skin lesion classification typically involves several stages: image acquisition, preprocessing, feature extraction, and supervised classification.

Lesions are categorized as benign—such as nevi (NV), benign keratosis (BKL), dermatofibromas (DF), and vascular lesions (VASC)—or malignant, including actinic keratosis (AKIEC), basal cell carcinomas (BCC), and melanomas (MEL), the latter being the most aggressive and deadly form [7], [8].

By integrating precise lesion segmentation with powerful feature-based classification, our approach offers a robust and scalable AI-driven solution for automated skin lesion diagnosis. The following section places this work in context by reviewing recent advances and related contributions in the field.

III. RELATED WORKS

A. Segmentation approaches

Ünver et al. [9] proposed an automated skin lesion segmentation method combining YOLO for lesion detection and GrabCut for segmentation. Their pipeline includes hair removal, YOLO-based localization, GrabCut segmentation, and morphological contour refinement, enabling accurate dermoscopic lesion analysis.

Mingzhe Hu et al. [10] proposed SkinSAM, a fine-tuned Segment Anything Model using Vision Transformers (ViT) for dermoscopic image segmentation. They showed that larger transformers and fine-tuning boost performance, underscoring transformers' growing role in medical imaging.

Tschandl et al. [11] trained a fully convolutional network using ResNet-34 as the encoder in a U-Net architecture. They compared three initialization strategies: random He initialization, ImageNet pre-training, and fine-tuned ResNet-34 weights for skin lesion classification, before training for binary segmentation.

Bi et al. [12] presented a class-specific deep learning framework that learns discriminative features separately for

melanoma and non-melanoma lesions. They employed a step-wise probability-based integration to effectively merge the outputs of specialized models, improving segmentation accuracy.

B. Classification approaches

A range of studies have focused on improving skin lesion classification through traditional machine learning and deep learning techniques.

Arora et al. [13] propose a method combining a bag-of-features approach with SURF descriptors for feature extraction, followed by a quadratic SVM for classification. Their pipeline begins with image preprocessing—smoothing, edge enhancement, and removal of noise and hair—before segmenting the lesion region. Discriminative features are then extracted from the segmented area, and the SVM classifier assigns the lesion to either a normal or cancerous class.

Adegun et al. [14] proposed a DL-based melanoma detection system combining an enhanced encoder-decoder network with skip connections to improve feature extraction. Segmentation was followed by a dedicated binary classification module, termed Lesion-classifier.

Nazi et al. [15] introduced a hybrid method integrating U-Net segmentation with a deep CNN-SVM classifier. To address limited data, they employed transfer learning and data augmentation alongside pre-trained CNNs to extract robust features.

Afza et al. [16] developed a hierarchical pipeline integrating contrast enhancement, superpixel-based segmentation, ResNet50 feature extraction, and Grasshopper Optimization for feature selection, collectively enhancing image quality, lesion isolation, and classification accuracy.

These works illustrate the evolving diversity of skin lesion analysis techniques, often integrating segmentation and classification to enhance diagnostic accuracy. Building on these advances, the next section details our framework's feature extraction and classification strategies.

IV. PROPOSED SYSTEM

The proposed system for skin cancer diagnosis comprises two main components: lesion segmentation for ROI extraction, and lesion classification. Fig. 1 provides an overview of the complete processing pipeline.

A. Pre-processing

Effective preprocessing enhances data quality, mitigates class imbalance, and improves model generalization. Our pipeline includes three key stages:

- 1) **Normalization and Resizing:** Images and masks are normalized to $[0, 1]$ via min-max scaling and resized to 128×128 pixels using bilinear interpolation. This ensures architectural compatibility while preserving spatial details.
- 2) **Data Augmentation:** A dual-generator applies synchronized geometric (rotation, translation, shear, zoom, flipping) and photometric adjustments (brightness, color

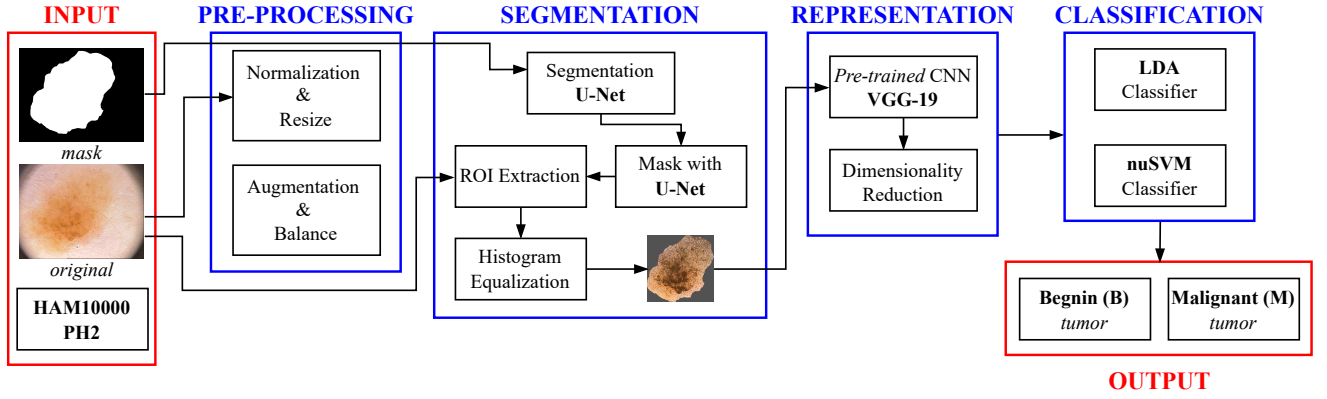


Fig. 1. Overview of the proposed system.

shifts) to images and masks simultaneously. This maintains pixel-wise alignment and enriches dataset diversity, reducing overfitting.

- 3) **Class Balancing:** Targeted augmentation on underrepresented classes synthetically increases their samples, mitigating imbalance and enhancing classification robustness.

B. Segmentation using a custom U-Net

The segmentation module employs a *custom* U-Net architecture tailored for dermoscopic image analysis. It maintains the symmetric encoder-decoder structure with skip connections, combining spatial details from shallow layers with semantic context from deeper ones for precise pixel-level segmentation [5].

However, modifications distinguish our implementation: four encoding/decoding levels instead of five, fewer convolutional filters (16 to 256) to reduce computational cost, and 128×128 input size instead of 256×256 for faster training and hardware compatibility. Convolutional blocks use He initialization, same padding, and ReLU activation, with a final sigmoid layer producing binary masks.

Training uses 50 epochs, batch size 16, Adam optimizer, binary cross-entropy loss, dropout rates of 0.1 – 0.3, and learning rate 0.001. Resulting masks extract ROIs from original images, with histogram equalization applied to enhance local contrast for improved feature extraction and classification.

C. Feature Extraction with Pre-trained VGG-19

We employ VGG-19 [6] as a deep feature extractor, leveraging its 16 convolutional layers trained on ImageNet. To capture high-level dermoscopic patterns, features are extracted up to the last convolutional layer of Block 5, which provides rich hierarchical representations of lesion texture, color, and structure. All ROI-based inputs are resized to 128×128 pixels to ensure architectural compatibility while preserving discriminative content for downstream classification.

D. Dimensionality Reduction

Principal Component Analysis (PCA) [17] reduces dimensionality by computing orthogonal linear combinations of input variables. The first principal component captures maximum data variance, with subsequent components capturing decreasing variance. This reduction decreases model complexity while preserving essential information, improving performance by mitigating overfitting and accelerating training.

E. Classification

Multiple supervised classifiers discriminate benign from malignant lesions:

- 1) **Linear Discriminant Analysis (LDA):** Maximizes class separability through linear feature combinations in lower-dimensional space.
- 2) **Nu-Support Vector Machine (NuSVM):** Uses parameter ν to control support vectors and training errors, enhancing flexibility for imbalanced data.
- 3) **Logistic Regression (LR):** Probabilistic linear model estimating class membership likelihood via logistic function.
- 4) **Bernoulli Naïve Bayes (BNB):** Naïve Bayes variant optimized for binary/boolean features.

This modular design—combining ROI segmentation, deep feature extraction, and classical classifiers—balances interpretability with computational efficiency, leveraging deep learning strengths despite limited annotated data.

V. EVALUATION PROTOCOL

A. Data Description

This study uses two public dermoscopic image datasets: HAM10000 and PH2, both established benchmarks for skin lesion analysis. Table I summarizes sample counts before and after augmentation. Fig. 2 and Fig. 3 show representative examples from both datasets.

- 1) **HAM10000** [7]: Contains 10,015 dermoscopic images of pigmented lesions from diverse populations and

TABLE I
DETAILS ABOUT **HAM10000** AND **PH2** BEFORE AND AFTER DATA
AUGMENTATION.

DATASET	TYPE	CLASS	BEFORE AUG.	AFTER AUG.
HAM10000	Benign	NV	6,705	1,200
		BKL	1,099	1,200
		VASC	142	1,196
		DF	115	1,097
		Total	8,061	4,693
	Malignant	MEL	1,113	1,500
		BCC	514	1,500
		AKIEC	327	1,500
		Total	1,954	4,500
PH2	Benign	CNV	80	400
		ANV	80	400
		Total	160	800
	Malignant	MEL	40	367
		Total	40	367



Fig. 2. Benign lesions: **HAM10000** and **PH2**.

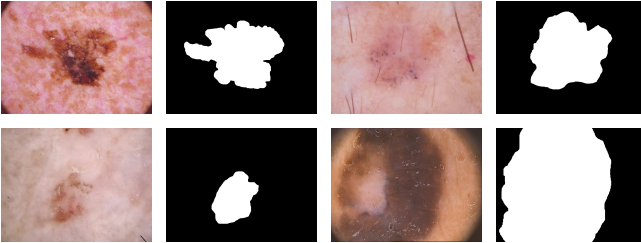


Fig. 3. Malignant lesions: **HAM10000** and **PH2**.

devices, with resolutions varying from 450×600 to 674×1024 pixels. Over 50% are histopathologically confirmed. The seven diagnostic categories include melanocytic nevi (NV), benign keratosis (BKL), dermatofibroma (DF), vascular lesions (VASC), intraepithelial carcinoma (AKIEC), basal cell carcinoma (BCC), and melanoma (MEL).

- 2) **PH2** [8]: Comprises 200 high-resolution dermoscopic images (768×560 pixels, $20\times$ magnification) from Pedro Hispano Hospital, with expert annotations and segmentation masks. Categories include common nevi, atypical nevi, and melanoma.

B. Evaluation Strategy

We used a simple validation scheme for both segmentation and classification. For segmentation, datasets were split into 70% training, 15% validation, and 15% testing. For classification, the split was 80% training and 20% testing. Model performance was evaluated using standard classification metrics, as detailed in the following section.

C. Evaluation Metrics

The segmentation and classification performance was quantitatively assessed using a set of well-established metrics, each capturing a specific aspect of model accuracy and robustness. The **Dice Similarity Coefficient (DSC)** and **Intersection over Union (IoU)** quantify the overlap between the predicted and ground truth masks, with IoU providing a stricter evaluation criterion. **Sensitivity** measures the proportion of true lesion pixels correctly identified, whereas **specificity** reflects the model's ability to accurately classify non-lesion pixels, thereby reducing false positives. **Accuracy** provides an overall estimate of prediction correctness across all pixels. **Precision** indicates the proportion of predicted lesion pixels that are truly positive, while the **F1-score**, defined as the harmonic mean of precision and recall, offers a balanced assessment under class imbalance. Finally, the **Structural Similarity Index Measure (SSIM)** captures perceptual similarity by comparing luminance, contrast, and structural information, thus evaluating spatial fidelity beyond simple pixel-wise overlap. Together, these metrics assess both pixel-level agreement and the structural consistency of the segmented regions.

VI. RESULTS & DISCUSSION

A. Segmentation Results

Tables II and III provide a comparative analysis of segmentation performance on the HAM10000 and PH2 datasets. Our method, based on a *custom* U-Net architecture, demonstrates competitive performance across all evaluation metrics.

On the HAM10000 dataset, our custom U-Net achieves competitive performance with balanced metrics. While Karri et al. [18] lead in accuracy (97.0%) and Dice (93.6%), our method excels in specificity (98.0%), IoU (81.0%), and SSIM (86.4%). Compared to Mingzhe Hu et al. [10] and Tschandl et al. [11], we maintain higher IoU with comprehensive metric reporting, demonstrating robust false positive control and structural preservation for clinical dermoscopic segmentation.

On the PH2 dataset, our custom U-Net method achieves the best performance in accuracy (96.6%), specificity (97.8%), recall (82.4%), and SSIM (85.6%). Although Bi et al. [12] achieve slightly higher Dice (92.1%) and IoU (85.6%) scores, our approach provides comprehensive evaluation across all metrics and demonstrates better control of false positives, which is essential for clinical dermoscopy applications.

B. Classification results

Table IV and Table V present the classification performance of VGG-19-based feature representations combined with four

TABLE II
RESULTS AND COMPARATIVE ANALYSIS WITH STATE-OF-THE-ART METHODS USING THE **HAM10000** DATASET.

REFERENCE	METHOD	YEAR	ACCURACY	SPECIFICITY	DSC	IoU	RECALL	SSIM
Karri et al. [18]	nSKnRSUNet	2023	97.0%	–	93.6%	–	–	–
Mingzhe Hu et al. [10]	SkinSAM	2023	94.5%	–	88.7%	78.4%	–	–
Tschandl et al. [11]	U-Net-16	2019	–	–	87.6%	80.0%	–	–
Proposed	<i>custom</i> U-Net	2025	95.3%	98.0%	88.0%	81.0%	80.0%	86.4%

TABLE III
RESULTS AND COMPARATIVE ANALYSIS WITH STATE-OF-THE-ART METHODS USING THE **PH2** DATASET.

REFERENCE	METHOD	YEAR	ACCURACY	SPECIFICITY	DSC	IoU	RECALL	SSIM
Saba et al. [19]	DCNN	2019	94.4%	–	–	–	–	–
Bi et al. [12]	DCL-PSI	2019	95.3%	94.5%	92.1%	85.6%	–	–
Ünver et al. [9]	YOLO	2021	93.3%	92.6%	84.3%	74.8%	–	–
Proposed	<i>custom</i> U-Net	2025	96.6%	97.8%	90.4%	83.6%	82.4%	85.6%

conventional classifiers: LDA, NuSVM, LR, and BNB, evaluated on the HAM10000 and PH2 datasets, respectively.

TABLE IV
PERFORMANCE OF THE **VGG-19**-BASED REPRESENTATION ON THE **HAM10000** DATASET.

CLASSIFIER	ACCURACY	PRECISION	F1-SCORE	RECALL
LDA	85.78%	86.93%	85.67%	85.78%
NuSVM	90.44%	91.39%	90.38%	90.44%
LR	88.72%	89.58%	88.66%	88.72%
BNB	82.84%	83.63%	82.67%	82.84%

On HAM10000, the VGG-19 + NuSVM combination achieves the best performance with an accuracy of 90.44%, precision of 91.39%, F1-score of 90.38%, and recall of 90.44%.

For PH2, VGG-19 + LDA attains the highest accuracy of 96.87%, precision of 97.53%, and F1-score of 96.10%.

TABLE V
PERFORMANCE OF THE **VGG-19**-BASED REPRESENTATION ON THE **PH2** DATASET.

CLASSIFIER	ACCURACY	PRECISION	F1-SCORE	RECALL
LDA	96.87%	97.53%	96.10%	96.80%
NuSVM	89.23%	90.20%	89.10%	89.23%
LR	85.43%	86.40%	85.12%	85.43%
BNB	90.62%	91.53%	90.12%	90.62%

These results highlight the discriminative power of VGG-19 features and the strong performance of NuSVM and LDA in handling inter-class variability.

Table VI and Table VII show the confusion matrices for HAM10000 and PH2. Performance on PH2 is slightly better, likely due to its more homogeneous nature and smaller class variability. Nonetheless, classification accuracy per class exceeds 90.0% on both datasets, demonstrating the reliability and effectiveness of our method.

TABLE VI
CONFUSION MATRIX FOR **HAM10000**.

	BENIGN	MALIGNANT
BENIGN	90.00%	10.00%
MALIGNANT	9.12%	90.88%

TABLE VII
CONFUSION MATRIX FOR **PH2**.

	BENIGN	MALIGNANT
BENIGN	100.00%	0.00%
MALIGNANT	6.25%	93.75%

Table VIII and IX compares the proposed method with recent state-of-the-art approaches on both datasets.

TABLE VIII
COMPARATIVE ANALYSIS OF OUR METHOD WITH EXISTING APPROACHES OVER RECENT YEARS ON THE **HAM10000** DATASET.

REFERENCE	APPROACH	YEAR	ACCURACY
Tanna et al. [20]	CNN+SVM	2021	84.39%
Afza et al. [16]	ResNet-50	2022	85.80%
Sae-Lim et al. [21]	MobileNet	2019	83.90%
Proposed	VGG-19 + NuSVM	2025	90.44%

TABLE IX
COMPARATIVE ANALYSIS OF OUR METHOD WITH EXISTING APPROACHES OVER RECENT YEARS ON THE **PH2** DATASET.

REFERENCE	APPROACH	YEAR	ACCURACY
Nazi et al. [15]	DCNN+SVM	2019	92.00%
Arora et al. [13]	SURFs+SVM	2022	85.70%
Adegun et al. [14]	CNN	2020	95.00%
Proposed	VGG-19 + LDA	2025	96.87%

Our VGG-19 + NuSVM achieves top accuracy with 90.44% on HAM10000 (see Table VIII), surpassing Tanna et al. [20], Afza et al. [16], and Sae-Lim et al. [21]. On PH2 (see Table IX), VGG-19 + LDA reaches 96.87% accuracy, outperforming Nazi et al. [15], Arora et al. [13], and Adegun et al. [14]. These results validate combining deep features with classical classifiers for skin lesion classification.

VII. CONCLUDING REMARKS

This study proposes an effective methodological framework for skin lesion segmentation and classification, based on a lightweight, *custom* U-Net architecture combined with rigorous preprocessing and training strategies. The extraction of regions of interest (ROI) from segmented lesions enabled

optimized classification performance. Deep features extracted using VGG-19, when paired with conventional classifiers such as NuSVM or LDA, achieved the highest performance, with the choice of the optimal classifier depending on the specific characteristics of each dataset.

The proposed method demonstrates strong performance despite challenges such as class imbalance and lesion heterogeneity. Nevertheless, improving generalization across diverse datasets remains an open issue. Future work will explore the integration of Transformer-based architectures, self-supervised learning approaches, and the design of lightweight models suitable for real-time deployment in clinical environments.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributors of the benchmark datasets HAM10000 [7] and PH2 [8], whose work made this study possible. We also thank the developers of the AI frameworks and tools utilized throughout this research.

REFERENCES

- [1] W. H. O. (WHO), "Skin cancer: Facts and figures," 2023, last accessed on February 22 2025. [Online]. Available: <https://www.who.int/news/item/08-11-2023-working-under-the-sun-causes-1-in-3-deaths-from-non-melanoma-skin-cancer--say-who-and-ilo>
- [2] American Cancer Society, "Cancer facts & figures 2022," 2022, last accessed on June 8 2025. [Online]. Available: <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html>
- [3] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *nature*, vol. 542, no. 7639, pp. 115–118, 2017.
- [4] T. J. Brinker, A. Hekler, A. H. Enk, J. Klode, A. Hauschild, C. Berking, B. Schilling, S. Haferkamp, D. Schadendorf, T. Holland-Letz *et al.*, "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task," *European Journal of Cancer*, vol. 113, pp. 47–54, 2019.
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [7] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions," *Scientific data*, vol. 5, no. 1, pp. 1–9, 2018.
- [8] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira, "Ph 2-a dermoscopic image database for research and benchmarking," in *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2013, pp. 5437–5440.
- [9] H. M. Ünver and E. Ayan, "Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm," *Diagnostics*, vol. 9, no. 3, p. 72, 2019.
- [10] M. Hu, Y. Li, and X. Yang, "Skinsam: Empowering skin cancer segmentation with segment anything model," *arXiv preprint arXiv:2304.13973*, 2023.
- [11] P. Tschandl, C. Sinz, and H. Kittler, "Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation," *Computers in biology and medicine*, vol. 104, pp. 111–116, 2019.
- [12] L. Bi, J. Kim, E. Ahn, A. Kumar, D. Feng, and M. Fulham, "Step-wise integration of deep class-specific learning for dermoscopic image segmentation," *Pattern recognition*, vol. 85, pp. 78–89, 2019.
- [13] G. Arora, A. K. Dubey, Z. A. Jaffery, and A. Rocha, "Bag of feature and support vector machine based early diagnosis of skin cancer," vol. 34, no. 11. Springer, 2022, pp. 8385–8392.
- [14] A. A. Adegun and S. Viriri, "Deep learning-based system for automatic melanoma detection," *IEEE Access*, vol. 8, pp. 7160–7172, 2019.
- [15] Z. A. Nazi and T. A. Abir, "Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with u-net and dnn-svm," in *Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018*. Springer, 2019, pp. 371–381.
- [16] F. Afza, M. Sharif, M. Mittal, M. A. Khan, and D. J. Hemanth, "A hierarchical three-step superpixels and deep learning framework for skin lesion classification," *Methods*, vol. 202, pp. 88–102, 2022.
- [17] K. Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, vol. 2, no. 11, pp. 559–572, 1901.
- [18] M. Karri, C. S. R. Annavarapu, and U. R. Acharya, "Skin lesion segmentation using two-phase cross-domain transfer learning framework," *Computer Methods and Programs in Biomedicine*, vol. 231, p. 107408, 2023.
- [19] T. Saba, M. A. Khan, A. Rehman, and S. L. Marie-Sainte, "Region extraction and classification of skin cancer: A heterogeneous framework of deep cnn features fusion and reduction," *Journal of medical systems*, vol. 43, no. 9, p. 289, 2019.
- [20] R. Tanna and T. Sharma, "Binary classification of melanoma skin cancer using svm and cnn," in *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*. IEEE, 2021, pp. 1–4.
- [21] W. Sae-Lim, W. Wettayaprasit, and P. Aiyarak, "Convolutional neural networks using mobilenet for skin lesion classification," in *2019 16th international joint conference on computer science and software engineering (JCSSE)*. IEEE, 2019, pp. 242–247.

Chapitre 6

Conclusion Générale

La détection du cancer de la peau reste un enjeu majeur à l'échelle internationale, en particulier pour le mélanome dont le diagnostic précoce est essentiel afin de réduire le taux de mortalité lié au cancer de la peau. Cela met en évidence la nécessité de développer des systèmes performants capables de favoriser une identification rapide et fiable des lésions.

Les avancées en intelligence artificielle et en apprentissage automatique ont profondément transformé ce domaine, en permettant la conception de systèmes de détection plus puissants et adaptatifs. Dans ce projet, différentes techniques ont été explorées selon les tâches à accomplir, notamment la segmentation et la classification. La méthode proposée pour la segmentation s'est révélée globalement efficace, en offrant de bons résultats tout en restant peu coûteuse et rapide à entraîner.

En ce qui concerne la classification, les modèles mis en œuvre varient selon qu'il s'agisse d'une tâche binaire ou multiclasse. Bien que les performances obtenues soient également satisfaisantes, leur mise en place s'avère plus coûteuse en ressources et requiert des temps d'entraînement plus importants par rapport à l'étape précédente.

6.1 Objectifs accomplis

Ce mémoire s'est intéressé à la détection du cancer de la peau à travers l'analyse de trois travaux complémentaires, chacun abordant une dimension spécifique du problème :

- ✓ Le premier article, intitulé « *Segmentation robuste multi-bases de données des lésions cutanées à l'aide d'un U-Net régularisé entraîné à partir de zéro avec une stratégie de prétraitement améliorée.* », avait pour objectif principal de proposer un modèle de segmentation des images dermoscopiques, capable d'isoler avec précision la ROI. Cette étape représente une phase préparatoire indispensable à la classification, élément central de ce projet. De plus, le modèle proposé peut être réutilisé dans d'autres contextes que la détection du cancer de la peau.
- ✓ Le deuxième article, intitulé « *Classification Hiérarchique en Deux Phases des Lésions Cutanées Utilisant des Espaces de Caractéristiques CNN Réutilisés et une Fusion de Classifieurs Inter-Étapes.* », explore un large éventail de techniques d'extraction de caractéristiques, allant des méthodes classiques aux CNN pré-entraînées. Il propose une approche en deux phases, fondée sur la fusion de plusieurs classificateurs, afin d'optimiser les performances de classification. Ces caractéristiques ont été mises à profit pour identifier et catégoriser différents types de cancers cutanés.
- ✓ Enfin, le troisième article, intitulé « *Un pipeline d'apprentissage profond hybride pour la détection du mélanome utilisant des caractéristiques U-Net modifiées et VGG-19.* », propose une synthèse entre la segmentation et la classification. Ce travail met en évidence la robustesse et l'efficacité du pipeline développé, confirmant la pertinence de l'approche intégrée.

6.2 Comparaison des approches développées

Comme présenté dans la Revue de littérature, les méthodes de détection du cancer de la peau reposent généralement sur deux étapes principales : la segmentation et la classification. Dans le cadre de ce travail, chaque système a été évalué à l'aide de quatre bases de données publiques.

6.2.1 Segmentation des lésions cutanées

La figure 6.1 illustre la comparaison de l'approche proposée pour la segmentation des lésions cutanées en termes de taux d'exactitude. Les résultats obtenus mettent en évidence une précision élevée pour l'ensemble des jeux de données considérés.

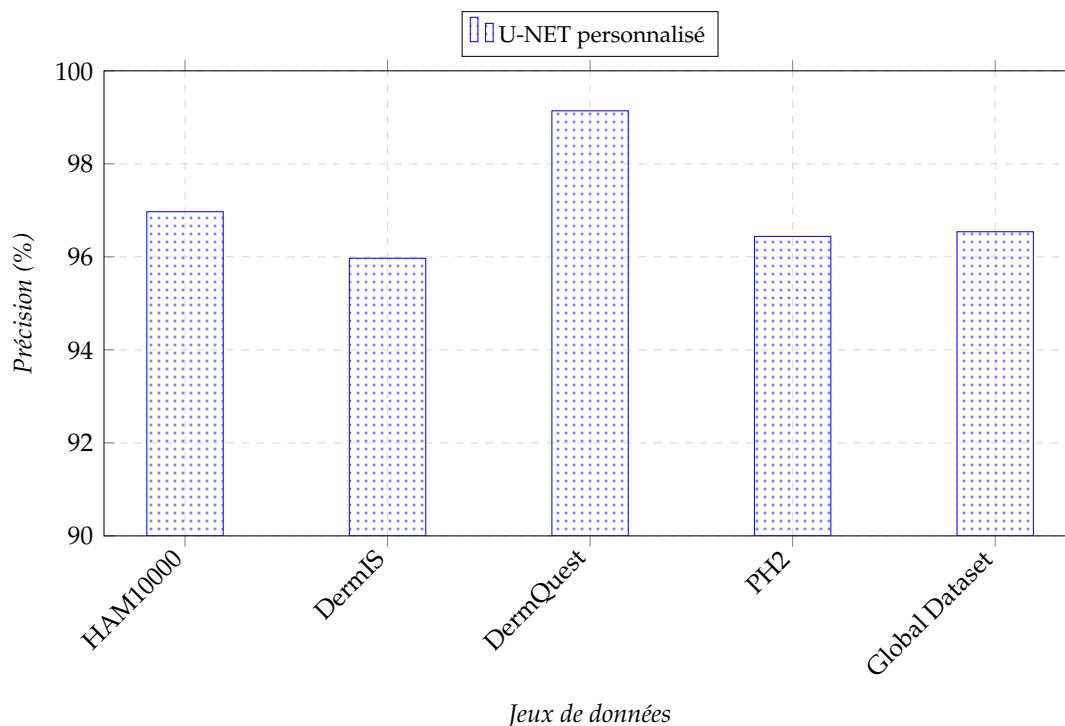


FIGURE 6.1 – Comparaison des performances de segmentation sur différents jeux de données.

6.2.2 Classification des lésions cutanées

Les figures 6.2 et 6.3 présentent la comparaison de l'approche proposée pour la classification binaire et multiclasse des lésions cutanées en termes de taux d'exactitude. Il ressort de ces résultats que la base de données PH2 enregistre les performances les plus élevées.

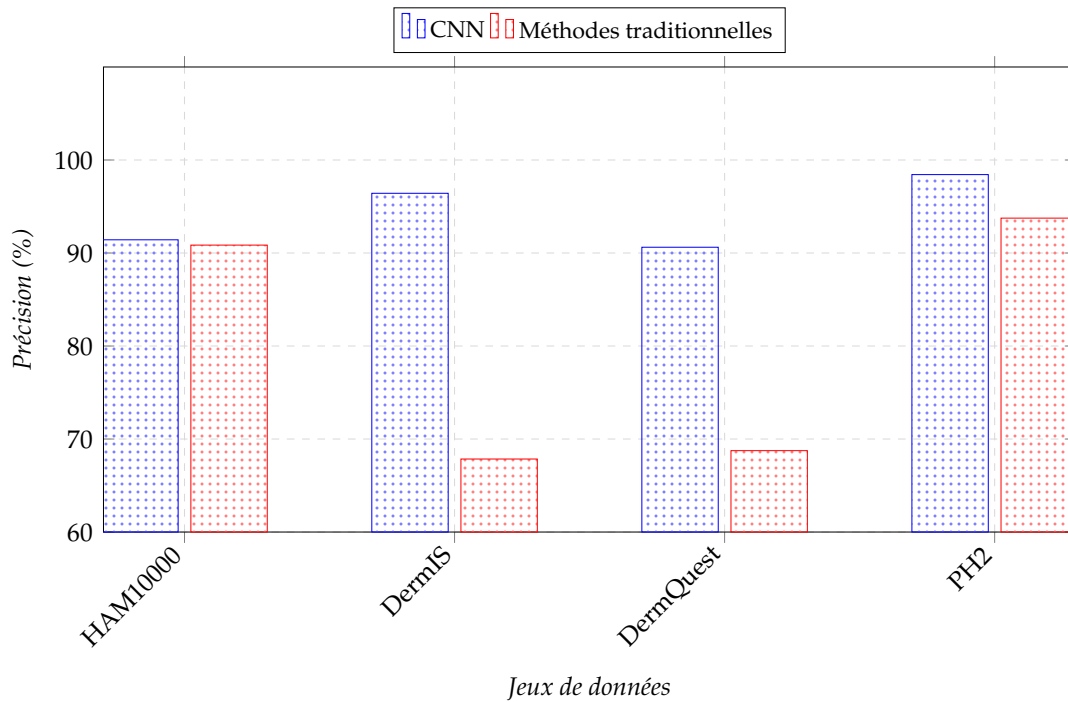


FIGURE 6.2 – Comparaison des performances en classification binaire sur différents jeux de données

6.3 Limitations connues

- × Notre manque d'expertise en dermatologie limite notre compréhension approfondie et notre interprétation précise des différents types de lésions rencontrés dans les bases de données utilisées.
- × La mise en place de deux systèmes distincts — l'un dédié à la segmentation et l'autre à la classification — complexifie le processus global et le rend plus coûteux en termes de ressources, notamment en temps de calcul et en entraînement.

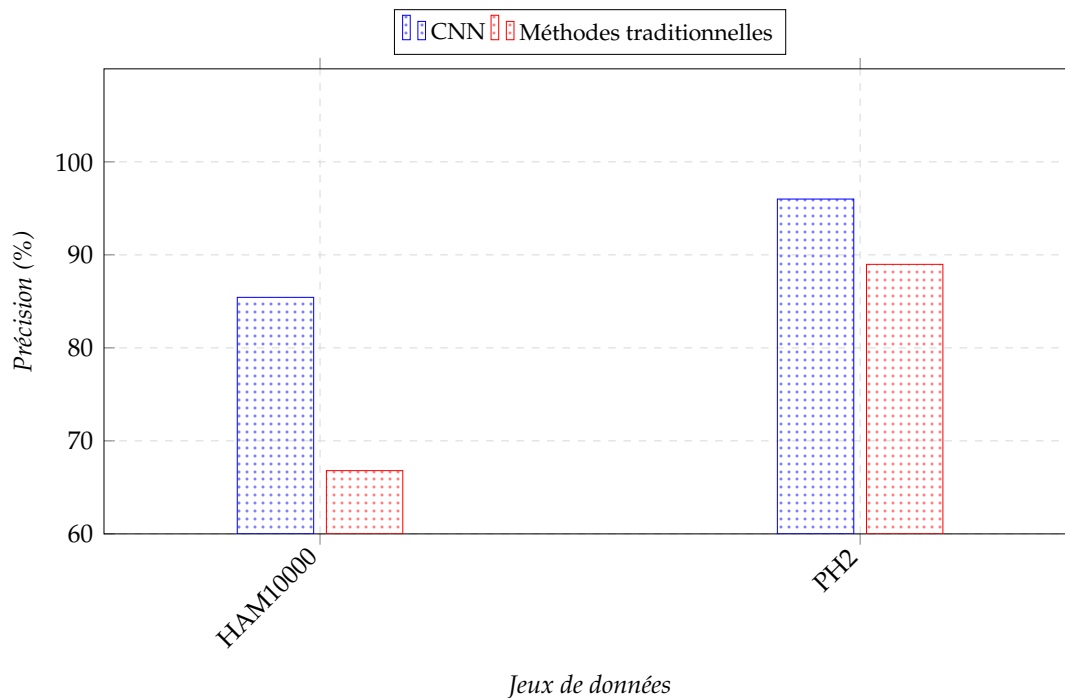


FIGURE 6.3 – Comparaison des performances en classification multi-classes sur différents jeux de données

- × L'utilisation de bases de données publiques, bien qu'elle permette une comparaison objective avec d'autres travaux, présente certaines limites liées à leur taille et à leur diversité, ce qui peut restreindre la robustesse et la généralisabilité de nos conclusions.
- × La variabilité des lésions et leurs fortes similarités visuelles constituent un défi supplémentaire, pouvant réduire le taux de reconnaissance et compliquer la précision des résultats en segmentation et classification.

6.4 Perspectives et travaux futurs

Bien que des progrès considérables aient été réalisés dans la détection et la classification des lésions cutanées, il reste des défis importants à relever.

- ◇ Les systèmes doivent être capables de s'adapter à la diversité des types

de lésions, aux variations de couleur de peau et aux conditions d'acquisition des images (qualité de la caméra, luminosité, bruit).

- ◇ Les futures recherches se concentreront sur le développement de modèles plus robustes et généralisables, capables de maintenir des performances élevées sur des bases de données variées incluant des personnes à la peau foncée et dans des contextes cliniques réels.
- ◇ L'évolution de ce travail inclut également l'intégration de systèmes d'aide à la décision clinique, qui non seulement détectent les lésions suspectes mais proposent aussi un degré de confiance et une interprétation visuelle des résultats.
- ◇ Pour une meilleure acceptation en pratique médicale, il sera crucial de développer des systèmes conviviaux, sécurisés et capables de s'intégrer dans le flux de travail des dermatologues, en minimisant le risque de fausses alertes tout en maximisant la valeur ajoutée pour le diagnostic précoce.

6.5 Retour d'expérience

Au cours de ma maîtrise en informatique, j'ai eu l'opportunité d'approfondir mes connaissances dans le domaine de l'IA. Ce projet n'est pas le premier que je mène dans ce domaine, puisque j'ai déjà réalisé des travaux similaires lors de mon parcours académique antérieur. Toutefois, il se distingue par son caractère singulier, lié à son impact humain tangible et à ses retombées bénéfiques dans le domaine de la santé, notamment en dermatologie. C'est pour ces raisons que j'ai choisi de consacrer mon projet de maîtrise à cette thématique.

L'encadrement dont j'ai bénéficié s'est avéré déterminant. L'engagement, les conseils personnalisés et les échanges constructifs avec mon directeur de

recherche ont constitué un soutien solide, indispensable à l'avancement du projet. Cette expérience a renforcé mon ambition de m'investir davantage dans la recherche en IA, avec pour objectif d'explorer des problématiques innovantes et de contribuer à l'avancée des connaissances dans ce domaine.

Les réalisations obtenues dans ce cadre ont été couronnées par l'attribution de la bourse d'excellence et de la bourse d'approbation employé et supérieure de l'Université du Québec à Chicoutimi (UQAC). Par ailleurs, ma participation à deux éditions successives du *Forum Innovation, Ingénierie, Informatique et Entrepreneuriat (FI3E)* a été récompensée par un troisième prix dans la catégorie informatique la première année, puis un deuxième prix et le prix coup de cœur du public l'année suivante.

En complément, mes expériences professionnelles — notamment en tant qu'*auxiliaire d'enseignement, correcteur et représentant des étudiants des cycles supérieurs en informatique* — ont enrichi mon parcours. Elles ont également éveillé et consolidé mon intérêt pour l'enseignement, que je considère comme une voie complémentaire et stimulante à la recherche scientifique.

Annexe

Titre :

Poster de participation au Forum Innovation, Ingénierie, Informatique et Entrepreneuriat (FI3E) en 2024 & 2025[65]

Introduction — À chaque étape du projet, dès l'obtention de résultats préliminaires, je réalise des présentations afin de recueillir des retours et des commentaires constructifs permettant d'apporter des améliorations. En 2024, mon poster présenté au FI3E portait sur le développement d'une application console intégrant un modèle de détection du cancer de la peau, tandis qu'en 2025, un poster élaboré en collaboration avec un étudiant du même programme était consacré à la reconnaissance des types de lésions, avec la proposition d'un modèle de détection en deux phases. Lors de ces deux participations, nous avons obtenu deux prix dans la catégorie informatique, ainsi que le prix coup de cœur du public.

DÉVELOPPEMENT D'UNE APPLICATION CONSOLE BASÉE SUR L'INTELLIGENCE ARTIFICIELLE POUR LA DÉTECTION DU CANCER DE LA PEAU.

Zakariaou Mounmie and Yacine Yaddaden

Département de mathématiques, informatique et génie, Université du Québec à Rimouski

Problématique

La détection précoce du cancer de la peau est d'une importance cruciale pour améliorer le taux de survie et la prise en charge des patients.

- Le cancer de la peau peut être évitable et traitable s'il est détecté tôt.
- Le diagnostic précis du cancer de la peau peut être complexe et demande l'expertise d'un dermatologue expérimenté.

Objectifs

L'objectif est de développer un système efficace et simple basé sur l'apprentissage automatique et les réseaux de neurones profonds (CNN) pré-entraînés capable de prédire le cancer de la peau. Pour atteindre cet objectif, nous nous sommes fixés des sous-objectifs suivants :

- Préparer un ensemble de données suffisamment diversifié et représentatif pour entraîner le modèle d'apprentissage (Dans notre cas, nous avons utilisé le PH2 et HAM10000).
- Effectuer la segmentation sur l'ensemble des images des deux bases de données
- Équilibrer les jeux de données HAM10000 et PH2
- Redimensionner les images en 128*128 ou 224*224 en fonction des techniques d'extraction utilisées
- Extraire les caractéristiques importantes dans les images avec les réseaux de neurones pré-entraînés comme le VGG-16 et ResNet50 mais aussi avec les méthodes traditionnelles LBP et HOG.
- Entraîner et valider le modèle d'apprentissage avec les métriques d'évaluation connues : Précision, sensibilité, F1-Score, Rappel et la matrice de confusion

Méthologie suivie

1. **la Règle ABCD** : Les acronymes de cette méthode sont basés principalement sur les paramètres liés à : l'Asymétrie de la lésion, l'irrégularité de la Bordure, la variabilité de la Couleur (brun, noir, rouge et bleu) et le Diamètre de la tumeur. [3]
 2. **GLCM (Gray Level Co-occurrence Matrix)** : La matrice de co-occurrence des niveaux de gris (GLCM) examine par une méthode arithmétique en tenant compte de l'association dimensionnelle des pixels [3].
- **Génération du modèle** : RF (Random Forest) pour la base de données ISIC 2016[3] et SDCA Maximum Entropy pour le PH2 [1]

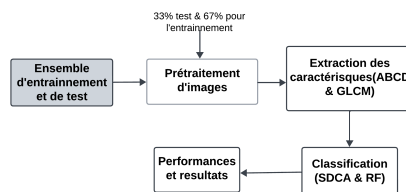
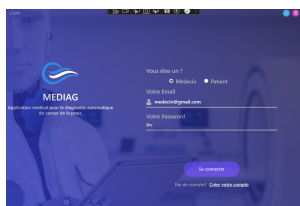


FIGURE 1 – Méthologie proposée.

Interface graphique

L'interface graphique de notre application est présentée dans l'image suivante.



Évaluation

- La base de données utilisée ici sont l'ISIC 2016 et le



FIGURE 2 – Images PH2 et ISIC 2016

- Pour mesurer la performance du modèle, nous avons utilisé l'accuracy, la precision, le rappel et la matrice de confusion.

Résultats

La Table 1 présente les résultats préliminaires obtenus avec le système proposée.

	Accuracy	Précision	Rappel		POSITIF	NEGATIF
PH2	0.58	0.63	0.78		3	12
ISIC2016	0.804	0.807	0.98			
	Résultats préliminaires				Matrice de confusion ISIC2016	
					POSITIF	NEGATIF
					POSITIF	3
					NEGATIF	2

FIGURE 3 – résultats Matrix de confusion

Conclusion

- Le projet est techniquement réalisable et les résultats préliminaires obtenus sont prometteurs et peuvent être améliorés
- En perspective, on prévoit de travailler sur d'autres techniques de prétraitement et d'extraction afin d'améliorer les performances du modèle.

Références

- [1] BARBUTO Francesco, PIAZZO Lorenzo et al. GRIGNAFFINI, Flavia. Machine learning approaches for skin cancer classification from dermoscopic images : A systematic review. *Algorithms*, 2022.
- [2] Teresa Mendonça, Pedro M Ferreira, Jorge S Marques, André RS Marcal, and Jorge Rozeira. Ph 2-a dermoscopic image database for research and benchmarking. In *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 5437–5440. IEEE, 2013.
- [3] NAIR S. Anu H. et KUMAR KP Sanal. MURUGAN, A. Detection of skin cancer using svm, random forest and knn classifiers. *Journal of medical systems*, 2019.
- [4] CHADHA HAMMADI ET YACINE YADDADEN. Développement d'une méthode pour la détection et la classification du type de cancer de la peau avec l'apprentissage automatique.. 2023.

Problématique

- Le cancer de la peau est évitable et traitable s'il est détecté tôt[2].
- Le diagnostic précis du cancer de la peau peut être complexe et demande l'expertise d'un dermatologue expérimenté.
- [1]

Objectifs

- Développer un système hybride efficace pour la détection et la classification des images cancéreuses et non cancéreuses.
- Affiner le modèle afin de l'optimiser pour la reconnaissance des types des cancers
- Le modèle générer sera exploité et intégré dans une application web pour la détection du cancer de la peau.

Méthodologie suivie

1. Pré-traitements

- Augmentation et équilibre des données
- Redimensionnement des images en 128×128
- Segmentation d'images (U-Net)
- Histogramme d'égalisation

2. Extraction des caractéristiques

- **Méthodes classiques** : HOG (Histogram of Oriented Gradients) et LBP (Local Binary Patterns)
- **Réseaux de neurones CNN** : ResNet50, VGG16, MobileNet et VGG19

3. Réduction des dimensions

- Sélection des caractéristiques avec le PCA

4. Classification

- Huit algorithmes utilisés : Extra Trees (ET), Bernoulli Naive Bayes (BNB), Support Vector Classifier (SVC), LightGBM (LGBM), Label Propagation (LP), Linear Discriminant Analysis (LDA), Stochastic Gradient Descent (SGD), Decision Tree Classifier (DT).

5. Évaluation du modèle

- 80% pour l'entraînement et 20% pour le test
- **métriques de performance** : Précision, Accuracy, F1-Score, Rappel et Matrice de confusion

6. Développement de l'application web

- Conception en architecture 3 couches (React + Django + PostgreSQL)
- Authentification sécurisée via JWT
- Intégration du modèle de prédiction via API Django

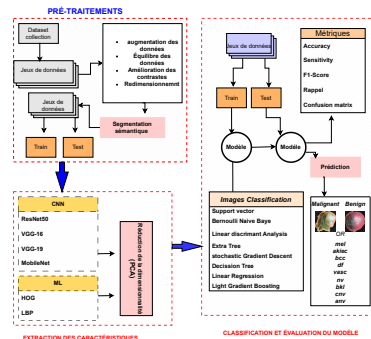


FIGURE 1 – Système proposé

Les bases de données

Nom	Echantillons	Nombre de classes
HAM10000	10015	7
PH2	200	3
DermIS	89	2
DermQuest	137	2

FIGURE 2 – Résumé des bases de données utilisées

FIGURE 3 – DermIS

FIGURE 4 – DermQuest

FIGURE 5 – PH2

FIGURE 6 – HAM10000

Résultats

BD	Méthode	Classificateur	Acc	Pré	Rappel	F1-S
HAM	MobileNet	SVC	92%	91.4%	92.6%	92.1%
	HOG	SVC	76.5%	76.45%	76.7%	76.3%
PH2	MobileNet	SVC	98.85%	98.9%	98.7%	98.5%
	HOG	LR	93.75%	93.39%	93.75%	93.74%
IS	MobileNet	SVC	96.6%	96.3%	96.5%	96.6%
	LBP	SVC	78.37%	78.57%	84.99%	77.54%
Quest	ResNet50	SVC	90.8%	90.7%	90.6%	90.5%
	LBP	ET	87.5%	87.5%	87.4%	87.8%

FIGURE 7 – Résultats de classification binaire

BD	Méthode	Classificateur	Acc	Pré	Rappel	F1-S
HAM	ResNet50	SVC+SVC/SVC	85.43%	85.4%	86.4%	83.5%
	HOG	SVC+SVC/SVC	76.5%	76.45%	76.7%	76.3%
PH2	MobileNet	ET+SVC/SVC	96%	95.9%	95.89%	95.93%
	HOG	SVC+LGBM/DT	66.8%	68.3%	65.5%	67.6%

FIGURE 8 – Résultats de classification multi-classes

- En classification binaire et multi-classes, le CNN MobileNet présente des meilleurs résultats soient 98.85% de précision en classification binaire et 96% en classification multi-classes de précision avec le jeu de données PH2

- Avec les Méthodes ML, HOG donne des meilleurs résultats avec le jeu de données PH2 soient 93.75% de classification binaire et 76.45% en classification multi-classes

	Benign	Malignant
Benign	135	25
Malignant	15	145

(a) Matrice de confusion - HAM10000 (binaire)

	Benign	Malign
Benign	32	0
Malignant	1	31

(b) Matrice de confusion - PH2 (binaire)

	Benign	Malign
Benign	32	0
Malignant	1	31

(a) Matrice de confusion - HAM10000 (binaire) (b) Matrice de confusion - PH2 (binaire)

	bkl	df	nv	vasc	akiec	bcc	mel
bkl	40	0	0	0	0	0	0
df	0	40	0	0	0	0	0
nv	0	0	20	0	12	3	5
vasc	1	0	1	35	0	0	3
akiec	0	0	0	0	40	0	0
bcc	0	0	0	0	2	34	4
mel	1	0	0	6	0	0	33

(d) Matrice de confusion - PH2 (multi-classes)

(c) Matrice de confusion - HAM10000 (7 classes) (d) Matrice de confusion - PH2 (multi-classes)

Application Web

- L'application développée permet une exploitation directe du modèle de classification à travers une interface web accessible.

- **Frontend (React, Redux, TypeScript)** : Interface intuitive pour l'authentification, l'upload d'image, l'affichage et la gestion des diagnostics.
- **Backend (Django REST Framework)** : API sécurisée (JWT), traitement des requêtes, intégration du modèle IA, gestion des utilisateurs et diagnostics.
- **Stockage (PostgreSQL + Cloudinary)** : Base de données relationnelle pour les utilisateurs et diagnostics. Les images sont hébergées sur Cloudinary.
- **Déploiement (Docker, Railway, GitHub)** : Application conteneurisée avec Docker, versionnée sur GitHub, déployée sur Railway avec intégration continue.

Conclusion

- Les résultats confirment l'efficacité des réseaux de neurones convolutifs (CNN) pour extraire les caractéristiques visuelles pertinentes et classer les lésions cutanées avec précision.
- L'application web développée est fonctionnelle, accessible et prête à être utilisée comme outil de diagnostic.

Références

- [1] tanuja Ansari, Uzma Bamo et Sarode. Skin cancer detection using image processing. *Int. Res. J. Eng. Technol.*, 2017.
- [2] GILANI Syed Qasim SYED Tehreem et al. NAQVI, Maryam. Skin cancer detection using deep learning—a review. *Diagnosis*, 2020.

Bibliographie

- [1] Lébini Chabi ADJOBO. « Amélioration du diagnostic du cancer de la peau par classification des images de dermoscopie à l'aide de techniques d'intelligence artificielle ». Thèse de doct. Université Bourgogne Franche-Comté ; Université d'Abomey-Calavi (Bénin), 2023.
- [2] Noor AHMED, Xin TAN et Lizhuang MA. « A new method for melanoma skin cancer lesion detection and segmentation based on hybrid convolutional neural network ». In : *Multimedia Tools and Applications* 82.8 (2023), p. 11873-11896. DOI : 10.1007/s11042-022-13618-0. URL : <https://doi.org/10.1007/s11042-022-13618-0>.
- [3] J. ALVES et al. « Automatic focus assessment on dermoscopic images acquired with smartphones ». In : *Sensors* 19.22 (2019), p. 4957. DOI : 10.3390/s19224957.
- [4] Ghadah ALWAKID et al. « Melanoma Detection Using Deep Learning-Based Classifications ». In : *Healthcare* 10.12 (2022). ISSN : 2227-9032. DOI : 10.3390/healthcare10122481. URL : <https://www.mdpi.com/2227-9032/10/12/2481>.
- [5] A. AMERI. « A deep learning approach to skin cancer detection in dermoscopy images ». In : *Journal of Biomedical Physics & Engineering* 10.6 (2020), p. 801-806.
- [6] Rafael Luz ARAUJO et al. « Automatic segmentation of melanoma skin cancer using deep learning ». In : *2020 IEEE International Conference*

- on E-Health Networking, Application and Services (HEALTHCOM)*. IEEE, mars 2021, p. 1-6. DOI : 10.1109/HEALTHCOM49281.2021.9398926.
- [7] Rehan ASHRAF et al. « Transfer learning–based region of interest–assisted framework for skin cancer detection ». In : *IEEE Access* 8 (2020), p. 147858-147871. DOI : 10.1109/ACCESS.2020.3014701.
- [8] Mohamed BADAWI et al. « Skin Cancer Classification and Segmentation Using Deep Learning ». In : *International Journal of Telecommunications* 4.1 (2024), p. 1-23. DOI : 10.21608/ijt.2024.280957.1045.
- [9] Philippe BAHADORAN, Jean-Paul ORTONNE et Robert BALLOTTI. « Que “trafiquent” les mélanosomes ? » In : *Médecine/Sciences* 18.2 (2002), p. 205-209.
- [10] J. BAUER et al. « Surveillance of patients at high risk for cutaneous malignant melanoma using digital dermoscopy ». In : *British Journal of Dermatology* 152.1 (2005), p. 87-92.
- [11] Julia BENEDETTI. *Structure et fonctions de la peau*. Consulté le 19 mars 2025. 2024. URL : <https://www.merckmanuals.com/fr-ca/accueil/troubles-cutan%C3%A9s/biologie-de-la-peau/structure-et-fonction-de-la-peau>.
- [12] Lei BI et al. « Dermoscopic image segmentation via multistage fully convolutional networks ». In : *IEEE Transactions on Biomedical Engineering* 64.9 (2017), p. 2065-2074. DOI : 10.1109/TBME.2017.2712771.
- [13] I. BIÈCHE. « Biologie moléculaire des cancers ». In : *Immuno-analyse & Biologie Spécialisée* 19.1 (2004), p. 13-22. ISSN : 0923-2532. DOI : 10.1016/j.immbio.2003.11.002. URL : <https://www.sciencedirect.com/science/article/pii/S0923253203001455>.

-
- [14] CANCER RESEARCH UK. *Ensemble, nous vainquons le cancer*. Consulté le 20 avril 2025. 2023. URL : <https://www.cancerresearchuk.org/about-cancer/what-is-cancer/how-cancers-grow>.
- [15] P. CARCAGNÌ et al. « Classification des lésions cutanées en combinant des apprentissages à plusieurs niveaux dans une architecture Dense-Net ». In : *Analysis and Processing of Images - ICIAP 2019*. Sous la dir. d'E. RICCI et al. T. 11751. Lecture Notes in Computer Science. Cham : Springer, 2019, p. 330-341. DOI : 10.1007/978-3-030-30642-7_30.
- [16] M. E. CELEBI et al. « Lesion border detection in dermoscopy images ». In : *Computerized Medical Imaging and Graphics* 33.2 (2009), p. 148-153. DOI : 10.1016/j.compmedimag.2008.11.002.
- [17] A. Prabhu CHAKKARAVARTHY et A. CHANDRASEKAR. « Anatomical region segmentation method from dermoscopic images of pigmented skin lesions ». In : *International Journal of Imaging Systems and Technology* 30.3 (2020), p. 636-652. DOI : 10.1002/ima.22404.
- [18] Noel CODELLA et al. *Skin Lesion Analysis Toward Melanoma Detection 2018 : A Challenge Hosted by the International Skin Imaging Collaboration (ISIC)*. 2019. arXiv : 1902.03368 [cs.CV].
- [19] Noel C. F. CODELLA et al. « Analyse des lésions cutanées en vue de la détection du mélanome : un défi au Symposium international sur l'imagerie biomédicale (ISBI) 2017, organisé par la collaboration internationale en imagerie cutanée (ISIC) ». In : *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, p. 168-172. DOI : 10.1109/ISBI.2018.8363547. URL : <https://challenge.isic-archive.com/data/>.

- [20] CSCIENCE. *Simplifier le diagnostic du cancer de la peau grâce à l'IA*. Consulté le 23 février 2025. 2023. URL : <https://www.cscience.ca/simplifier-le-diagnostic-du-cancer-de-la-peau-grace-a-lia/>.
- [21] DERMATOLOGIE MÉDICALE ESTHÉTIQUE ET LASER. *La dermoscopie*. Consulté le 26 juillet 2025. 2016. URL : <http://docteur-casacci.com/dermoscopie/>.
- [22] *Dermatology Information System (DERMIS)*. 2012. URL : <http://www.dermIS.com/> (visité le 30/12/2024).
- [23] *DermQuest : Skin Cancer Detection*. 2012. URL : <http://vip.uwaterloo.ca/demos/skin-cancer-detection> (visité le 30/12/2024).
- [24] ELSAN. *Cancer de la peau : définition*. Consulté le 22 février 2025. 2021. URL : <https://www.elsan.care/fr/pathologie-et-traitement/cancers/cancer-de-la-peau-definition>.
- [25] EUCERIN. *Comprendre la peau : structure et fonctions cutanées*. Consulté le 19 mars 2025. 2024. URL : <https://www.fr.eucerin.ca/la-peau/connaissances-de-base-sur-la-peau/structure-et-fonction-cutanees>.
- [26] Seham GAMIL et al. « An Efficient AdaBoost Algorithm for Enhancing Skin Cancer Detection and Classification ». In : *Algorithms* 17.8 (2024), p. 353. DOI : 10.3390/a17080353.
- [27] Claus GARBE et al. « European consensus-based interdisciplinary guideline for melanoma. Part 1 : Diagnostics : Update 2022 ». In : *European Journal of Cancer* 170 (2022), p. 236-255. DOI : 10.1016/j.ejca.2022.03.008. URL : <https://www.sciencedirect.com/science/article/pii/S0959804922001526>.

- [28] GOUVERNEMENT DU CANADA. *Sécurité-soleil*. Consulté le 22 février 2025. 2023. URL : <https://www.canada.ca/fr/sante-publique/services/securite-soleil/cancer-peau.html>.
- [29] Manu GOYAL et al. « Segmentation des lésions cutanées dans les images dermoscopiques avec des méthodes d'apprentissage profond d'ensemble ». In : *IEEE Access* 8 (2020), p. 4171-4181. DOI : 10.1109/ACCESS.2019.2960504.
- [30] Julie GOYON. « Matériaux amorphes : des solides qui coulent de façon collective ». Thèse de doctorat. Thèse de doct. Université Sciences et Technologies - Bordeaux I, oct. 2008. URL : <https://theses.hal.science/tel-00360121>.
- [31] David GUTMAN et al. *Skin Lesion Analysis toward Melanoma Detection : A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)*. 2016. arXiv : 1605.01397 [cs.CV].
- [32] Alzbeta HARTINGER. « Détection du cancer de la peau par tomographie d'impédance électrique ». Thèse de doct. École Polytechnique de Montréal, 2012.
- [33] Dilli HEMALATHA, Kaile Nishi LATHA et Pandala. Madhavi LATHA. « Détection du cancer de la peau à l'aide de la technique d'apprentissage profond ». In : *2023 2nd International Conference for Innovation in Technology (INOCON)*. IEEE, 2023, p. 1-5. DOI : 10.1109/INOCON57975.2023.10101344.
- [34] Galib Muhammad Shahriar HIMEL et al. « Skin Cancer Segmentation and Classification Using Vision Transformer for Automatic Analysis in Dermatoscopy-Based Noninvasive Digital System ». In : *International*

- Journal of Biomedical Imaging* 2024 (2024), p. 3022192. DOI : 10 . 1155 / 2024/3022192.
- [35] Khalid M. HOSNY, Mohamed A. KASSEM et Mohamed M. FOAUD. « Classification of skin lesions using transfer learning and augmentation with Alex-net ». In : *PLOS ONE* 14.5 (2019), e0217293. DOI : 10 . 1371/journal.pone.0217293.
- [36] Mingzhe HU, Yuheng LI et Xiaofeng YANG. *SkinSAM : Empowering Skin Cancer Segmentation with Segment Anything Model*. 2023. arXiv : 2304 . 13973 [cs.CV]. URL : <https://arxiv.org/abs/2304.13973>.
- [37] Muhammad Qasim KHAN et al. « Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer ». In : *IEEE Access* 7 (2019), p. 90132-90144. DOI : 10 . 1109/ACCESS.2019.2926837.
- [38] Harald KITTLER et Michael BINDER. « Follow-up of melanocytic skin lesions with digital dermoscopy : risks and benefits ». In : *Archives of Dermatology* 138.10 (2002), p. 1379-1379.
- [39] K. Anup KUMAR et C. VANMATHI. « Optimization driven model and segmentation network for skin cancer detection ». In : *Computers and Electrical Engineering* 103 (2022), p. 108359. DOI : 10 . 1016/j.compeleceng.2022.108359.
- [40] Nabeel F. LATTOOFI et al. « Melanoma Skin Cancer Detection Based on ABCD Rule ». In : *2019 First International Conference of Computer and Applied Sciences (CAS)*. IEEE, 2019, p. 154-157. DOI : 10 . 1109/CAS47993.2019.9075465.
- [41] Santé MAGAZINE. *D'où viennent nos grains de beauté ? À quoi servent-ils ?* Consulté le 23 octobre 2025. Avr. 2024. URL : <https://www.santemagazine.fr/sante/maladies/maladie-de-la-peau/dou-viennent-nos-grains-de-beaute-a-quoi-servent-ils-1055967>.

- [42] Alexandre MÉLISSOPOULOS et Christine LEVACHER. « La peau : structure et physiologie ». In : *Éditions Médicales Internationales* 11 (1998).
- [43] Teresa MENDONÇA et al. *PH² : A dermoscopic image database for research and benchmarking*. 35th International Conference of the IEEE Engineering in Medicine et Biology Society, 2013. DOI : 10.1109/EMBC.2013.6610779. URL : <https://www.fc.up.pt/addi/ph2%20database.html>.
- [44] K. M. MONICA et al. « Melanoma skin cancer detection using mask-RCNN with modified GRU model ». In : *Frontiers in Physiology* 14 (2024). DOI : 10.3389/fphys.2023.1324042.
- [45] Gagare Shrikant NARAYAN et M. R. BENDRE. « Identification du cancer de la peau à l'aide de CNN : un modèle léger avec une classification à grain fin ». In : *2024 MIT International Conference on School of Computing, Arts, Design and Technology (MITADTSoCiCon)*. IEEE, 2024, p. 1-6. DOI : 10.1109/MITADTSoCiCon60330.2024.10575526.
- [46] ORGANISATION MONDIALE DE LA SANTÉ (OMS). *Cancer de la peau : faits et chiffres*. Consulté le 22 février 2025. 2023. URL : <https://www.who.int/fr/news/item/08-11-2023-working-under-the-sun-causes-1-in-3-deaths-from-non-melanoma-skin-cancer--say-who-and-ilo>.
- [47] Kemal POLAT et Kaan Onur KOC. « Détection des maladies de la peau à partir d'une image de dermoscopie à l'aide de la combinaison d'un réseau neuronal convolutif et d'un contre tous ». In : *Journal de l'intelligence artificielle et des systèmes* 2 (2020), p. 80-97. DOI : 10.33969/AIS.2020.21006.
- [48] T. P. POTJER et al. « Genetic testing for melanoma : current insights and clinical implications ». In : *International Journal of Cancer* 144.10 (2019),

- p. 2453-2464. DOI : 10.1002/ijc.31984. URL : <https://doi.org/10.1002/ijc.31984>.
- [49] RADIO CANADA. *Simplifier le diagnostic du cancer de la peau grâce à l'IA*. Consulté le 23 février 2025. 2024. URL : <https://ici.radio-canada.ca/nouvelle/2039816/cancer-peau-intelligence-artificielle-medecine>.
- [50] Aminur Rab RATUL et al. « Skin Lesions Classification Using Deep Learning Based on Dilated Convolution ». In : *bioRxiv* (2019). DOI : 10.1101/860700.
- [51] Veronica ROTEMBERG et al. « A patient-centric dataset of images and metadata for identifying melanomas using clinical context ». In : *Scientific Data* 8.1 (2021), p. 34. DOI : 10.1038/s41597-021-00815-z.
- [52] Ranjita ROUT et Priyadarsan PARIDA. « Transition region based approach for skin lesion segmentation ». In : *Procedia Computer Science* 171 (2020), p. 379-388. DOI : 10.1016/j.procs.2020.04.039.
- [53] M. SANTHOSH et al. « Évaluation assistée par image d'un segment de cancer à partir d'images de dermoscopie ». In : *Intelligent Engineering and Data Analysis*. Sous la dir. de S. SATAPATHY et al. T. 1177. Advances in Intelligent Systems and Computing. Singapore : Springer, 2021, p. 755-763. DOI : 10.1007/978-981-15-5679-1_68.
- [54] Elineide SANTOS et al. « Une méthode de segmentation semi-supervisée des lésions cutanées ». In : *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, juill. 2020, p. 223-228. DOI : 10.1109/IWSSIP48289.2020.9145240.

- [55] E. M. SENAN et M. E. JADHAV. « Techniques de détection des lésions cutanées dans les images de dermoscopie PH2 à l'aide d'un motif binaire local (LBP) ». In : *Recent Trends in Image Processing and Pattern Recognition*. Sous la dir. de K. C. SANTOSH et B. GAWALI. T. 1381. Communications in Computer and Information Science. Singapore : Springer, 2021, p. 15-26. DOI : 10.1007/978-981-16-0493-5_2.
- [56] Kirill SIROTKIN et al. « Improved skin lesion recognition by a Self-Supervised Curricular Deep Learning approach ». In : *arXiv preprint* (2021). arXiv : 2112.12086 [cs.CV].
- [57] SLIDESERVE. *Anatomie et physiologie de la peau normale*. Consulté le 26 juillet 2025. 2025. URL : <https://www.slideserve.com/graham/anatomie-et-physiologie-de-la-peau-normale>.
- [58] SOCIÉTÉ CANADIENNE DU CANCER. *Informations sur le cancer*. Consulté le 22 février 2025. 2023. URL : <https://cancer.ca/fr/cancer-information/cancer-types/skin-non-melanoma/risks>.
- [59] SOCIÉTÉ CANADIENNE DU CANCER. *Les types de tumeurs*. Consulté le 22 mars 2025. 2023. URL : <https://cancer.ca/fr/cancer-information/cancer-types/skin-non-melanoma/risks>.
- [60] Authors not SPECIFIED. « Confocal microscopy – Working principle and applications in dermatology ». In : *Journal of Skin and Sexually Transmitted Diseases* 5 (2022). ISSN : 2582-3175. DOI : 10.25259/JSSTD_23_2022. URL : https://doi.org/10.25259/JSSTD_23_2022.
- [61] Philipp TSCHANDL. *The HAM10000 dataset : A large collection of multi-source dermatoscopic images of common pigmented skin lesions*. Version V4. Harvard Dataverse, 2018. DOI : 10.7910/DVN/DBW86T. URL : <https://doi.org/10.7910/DVN/DBW86T>.

- [62] Halil Murat ÜNVER et Enes AYAN. « Skin Lesion Segmentation in Dermoscopic Images with Combination of YOLO and GrabCut Algorithm ». In : *Diagnostics* (2019). DOI : 10.3390/diagnostics9030072.
- [63] Yongwei WANG et al. « SSD-KD : A self-supervised diverse knowledge distillation method for lightweight skin lesion classification using dermoscopic images ». In : *Medical Image Analysis* 84 (2023), p. 102693. DOI : 10.1016/j.media.2022.102693.
- [64] Wen YIN, Dongming ZHOU et Rencan NIE. « DI-UNet : dual-branch interactive U-Net for skin cancer image segmentation ». In : *Journal of Cancer Research and Clinical Oncology* 149.17 (2023), p. 15511-15524. ISSN : 1432-1335. DOI : 10.1007/s00432-023-05319-4. URL : <https://doi.org/10.1007/s00432-023-05319-4>.
- [65] ZAKARIAOU et YADDADEN. « développement d’une application console basée sur l’intelligence artificielle pour la détection du cancer de la peau. » In : (2025).
- [66] Nadia S. ZGHAL et Nabil DERBEL. « Melanoma Skin Cancer Detection based on Image Processing ». In : *Current Medical Imaging* 16.1 (2020), p. 50-58. DOI : 10.2174/1573405614666180911120546.